JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **31**: 929–960 (2016)
Published online 16 July 2015 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.2476

EXPONENT OF CROSS-SECTIONAL DEPENDENCE: ESTIMATION AND INFERENCE

NATALIA BAILEY,[a*] GEORGE KAPETANIOS[a] AND M. HASHEM PESARAN[b,c]

[a] *Queen Mary, University of London, UK*
[b] *University of Southern California, Los Angeles, CA, USA*
[c] *Trinity College, University of Cambridge, UK*

SUMMARY

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

Received 1 April 2015; Revised 2 February 2015

Supporting information may be found in the online version of this article.

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, and regional and national economies. Researchers in many fields have turned to network theory, and 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.

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

* Correspondence to: Natalia Bailey, School of Economics and Finance, Queen Mary, University of London, Mile End Road, London E1 4NS, UK. E-mail: n.bailey@qmul.ac.uk

Let x_{it} denote a double array of random variables indexed by $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$, over space and time, respectively, and, without loss of generality, assume that $E(x_{it}) = 0$. Then the covariance matrix of $\boldsymbol{x}_t = (x_{1t}, x_{2t}, \ldots, x_{Nt})'$ is given by $\boldsymbol{\Sigma}_N = E\left(\boldsymbol{x}_t \boldsymbol{x}_t'\right) = (\sigma_{ij,x})$, with its largest eigenvalue denoted by $\lambda_{\max}(\boldsymbol{\Sigma}_N)$. The variables x_{it} are said to be strongly cross-sectionally correlated if $\lambda_{\max}(\boldsymbol{\Sigma}_N)$ rises with N, and they are said to be weakly cross-sectionally correlated if $\lambda_{\max}(\boldsymbol{\Sigma}_N)$ is bounded in N. This is clearly an important distinction and forms the basis of most factor models considered in the finance and macroeconometric literature (Forni *et al.*, 2000; Forni and Lippi, 2001; Bai and Ng, 2002; Bai, 2003).

In particular, standard factor models assume that $\lambda_{\max}(\boldsymbol{\Sigma}_N) = O(N)$, while spatial models typically assume that $\lambda_{\max}(\boldsymbol{\Sigma}_N) = O(1)$. In practice, one would expect to encounter degrees of cross-sectional dependence that lie between these two extremes. Also, in empirical applications where the degree of cross-sectional dependence is weak, it might not be possible to distinguish different models of cross-sectional dependence in terms of $\lambda_{\max}(\boldsymbol{\Sigma}_N)$. For example, $\lambda_{\max}(\boldsymbol{\Sigma}_N)$ is bounded in N irrespective of whether x_{it} are cross-sectionally independent or spatially dependent. For this reason, and as we shall see below, it is only possible to identify and consistently estimate α for values of $\alpha > 1/2$. Accordingly, we consider models of cross-sectional dependence for which $\lambda_{\max}(\boldsymbol{\Sigma}_N) = O(N^{\alpha})$, and $1/2 < \alpha \leq 1$, and investigate the problem of estimating α. It is important that empirical analysis of cross-sectional dependence is firmly based on observations rather than on an a priori chosen value of α.

We propose estimating α using the variance of the cross-sectional average of the observed data, $\bar{x}_t = T^{-1}\Sigma_{t=1}^{T} x_{it}$, and present bias-corrected estimators of α under a multiple factor setting. We derive their asymptotic properties and consider a number of extensions that allow for the presence of temporal dependence in the factors or the idiosyncratic component, and weak cross-sectional dependence in the latter. It is also worth pointing out that our estimators of α do not use explicitly a factor structure. The factor representation is only needed as a vehicle to derive the theoretical properties of the estimator and to give α a unique interpretation as a measure of cross-sectional dependence. We use this vehicle because working with covariances directly would involve high-level assumptions and would potentially lead to stricter conditions such as the need for T to rise faster than N. A further crucial reason for using the factor model is that, as proven in Theorem 4 of Chamberlain and Rothschild (1983), a covariance matrix that has a finite number of eigenvalues that tend to infinity as N increases, has a unique factor representation. This makes the factor model a canonical model for analysing cross-sectional dependence associated with covariance matrices with a finite number of exploding eigenvalues.

To illustrate the properties of the proposed estimators of α and their asymptotic distributions, we carry out a detailed Monte Carlo study that considers a battery of robustness checks. Finally, we provide a number of empirical applications investigating the degree of inter-linkages of real and financial variables in the global economy. We explore the extent to which macroeconomic variables are interconnected across and within countries, and present recursive estimates of α applied to excess returns on securities included in Standard & Poor's 500 (S&P 500) index.

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

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

2. PRELIMINARIES AND MOTIVATIONS

In this section, we introduce the concept of the exponent of cross-sectional dependence and our proposed estimator of it. We start our discussion by considering a simple measure of cross-sectional dependence based on cross-sectional averages defined by $\bar{x}_t = N^{-1}\sum_{i=1}^{N} x_{it}$. The limiting behaviour of $\bar{x}_t$ is of interest in its own right and provides information on the nature and degree of cross-sectional dependence. In the case of asset returns, this determines the extent to which risk, associated with investing in particular portfolios of assets, is diversifiable. In the case of firm sales, this is of interest in relation to the effect of idiosyncratic, firm level, shocks onto aggregate macroeconomic variables such as gross domestic product (GDP). In the case where x_{it} are cross-sectionally independent, using CLT, one obtains the result that $\mathrm{var}(\bar{x}_t) = O\left(N^{-1}\right)$. However, in the more general and realistic case where x_{it} are cross-sectionally correlated, we have that $\mathrm{var}(\bar{x}_t)$ declines at a rate that is a function of α, where α is defined by

$$0 < c_1 < \lim_{N\to\infty} N^{-\alpha}\lambda_{\max}(\mathbf{\Sigma}_N) < c_2 < \infty \tag{1}$$

We note that $\mathrm{var}(\bar{x}_t)$ cannot decline at a rate faster than N^{-1}. It is also easily seen that $\mathrm{var}(\bar{x}_t)$ cannot decline at a rate slower than $N^{\alpha-1}$, $0 \le \alpha \le 1$. To see this, we explore the link between $\lambda_{\max}(\mathbf{\Sigma}_N)$ and $\mathrm{var}(\bar{x}_t)$. Note that $\bar{x}_t = N^{-1}\boldsymbol{\iota}'\boldsymbol{x}_t$, where $\boldsymbol{\iota}$ is an $N \times 1$ vector of ones. Then, we have

$$\mathrm{var}(\bar{x}_t) = N^{-2}\boldsymbol{\iota}'\mathbf{\Sigma}_N\boldsymbol{\iota} \le N^{-2}\boldsymbol{\iota}'\boldsymbol{\iota}\lambda_{\max}(\mathbf{\Sigma}_N) = N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N)$$

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

It is interesting to note that the above measures of cross-sectional dependence are also related to the degree of pervasiveness of factors in unobserved factor models often used in the literature to model cross-sectional dependence. Factor models have a long pedigree both as a conceptual device for summarizing multivariate datasets as well as an empirical framework with sound theoretical underpinnings both in finance and economics. Conventionally, these make the distinction between the 'common component', which has a pervasive effect on the data so that α, as defined in equation (1), is assumed to equal unity, and the 'idiosyncratic component', whose impact is localized in nature, i.e. $\alpha = 0$. Recent econometric research on factor models include Bai and Ng (2002), Bai (2003), Forni *et al.* (2000, 2009), Forni and Lippi (2001), Pesaran (2006) and Stock and Watson (2002).[1]

As an illustration consider the single-factor model

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

where x_{it} depends on a single unobserved factor f_{1t}, with the associated factor loadings, β_{i1}, and cross-sectionally independent idiosyncratic errors, u_{it}. The extent of cross-sectional dependence in

[1] While Forni *et al.* (2000) and Forni and Lippi (2001) study the eigenvalues of the spectral density matrix, Forni *et al.* (2009) focus on the eigenvalues of the covariance matrix, which reflect closely the assumptions of Chamberlain and Rothschild (1983). In turn, Bai and Ng (2002) and Bai (2003) make assumptions on the sum of the covariances of the errors.

DOI: 10.1002/jae

x_{it} crucially depends on the nature of the factor loadings. It is easily seen that

$$\lambda_{\max}(\boldsymbol{\Sigma}_N) = O\left(\sum_{i=1}^{N} \beta_{i1}^2\right) = O\left[\left(\sup_j |\beta_{j1}|\right)\sum_{i=1}^{N}|\beta_{i1}|\right] = O\left(\sum_{i=1}^{N}|\beta_{i1}|\right)$$

when $\sup_j |\beta_{j1}| < K$. Also $\text{var}(\bar{x}_t) = O\left\{\max\left[\left(N^{-1}\sum_{i=1}^{N}\beta_{i1}\right)^2, N^{-1}\right]\right\}$.[2] The degree of cross-sectional dependence will be strong if the average value of β_{i1} is bounded away from zero. In such a case, $N^{-1}\lambda_{\max}(\boldsymbol{\Sigma}_N)$ and $\text{var}(\bar{x}_t)$ are both $O(1)$, which yields $\alpha = 1$.

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

Although mathematically convenient, the assumption that all factor loadings vary with N (almost uniformly) is rather restrictive in many economic applications. Therefore, we will not consider it in detail but only briefly as an alternative formulation. In this paper, we consider a baseline formulation where we assume that only $[N^\alpha]$ of the N factor loadings are individually important ($[N^\alpha]$ is the integer part of $N^\alpha, 0 < \alpha \leq 1$), in the sense that they are bounded away from zero. In effect, the factor loadings $\boldsymbol{\beta} = (\beta_{11}, \beta_{21}, \dots, \beta_{N1})'$ are grouped into two categories: a strong category $\left((\beta_{11}, \beta_{21}, \dots, \beta_{[N^\alpha]1})'\right)$ with non-zero means, and a weak category $\left((\beta_{([N^\alpha]+1)1}, \beta_{([N^\alpha]+2)1}, \dots, \beta_{N1})'\right)$ with negligible effects and a mean that tends to zero with N. Under this setup, $N^{-1}\lambda_{\max}(\boldsymbol{\Sigma}_N) = O\left(N^{\alpha-1}\right)$, as long as $\max_i \beta_{i1} = o_p\left(N^d\right)$, for all $d > 0, \text{var}(\bar{x}_t) = O\left(N^{2\alpha-2}\right)$ and the standard deviation of $\bar{x}_t$, denoted by $\text{SD}(\bar{x}_t)$ is $O\left[\max\left(N^{\alpha-1}, N^{-1/2}\right)\right]$. Note that at least $N^{1/2}$ of the loadings must have non-zero means for the covariances in Σ_N to dominate the diagonal of Σ_N and result in a rate of decline for $\text{SD}(\bar{x}_t)$ that is $O\left(N^{\alpha-1}\right)$. If fewer than $N^{1/2}$ of the loadings have non-zero means, then $\text{SD}(\bar{x}_t) = O\left(N^{-1/2}\right)$. The presence of at least $N^{1/2}$ loadings with non-zero means implies that $\alpha > 1/2$. In that case, and as long as the mean of the loadings from the strong category is non-zero, then $N^{-1}\lambda_{\max}(\boldsymbol{\Sigma}_N)$ and $\text{SD}(\bar{x}_t)$ decline at the same rate. As a result, in the context of the factor model in (2), α has a unique role as a measure of cross-sectional dependence. It is important to note that if the sum of the means of the loadings from the strong category over m factors, say μ_v, is equal to zero, then $\text{SD}(\bar{x}_t) = O\left(N^{-1/2}\right)$ for all α including the case of $\alpha = 1$. The implication is that even a strong factor model allows full portfolio diversification at the same rate as if no factors had been present. Seen from this perspective, the case where $\mu_v = 0$ does not seem very plausible, at least in the case of macro and financial datasets where full diversification of risk does not seem to be a possibility.

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

Consider now the following multiple-factor generalization of our basic setup:

$$x_{it} = \sum_{j=1}^{m} \beta_{ij} f_{jt} + u_{it} = \boldsymbol{\beta}_i' \boldsymbol{f}_t + u_{it}, \text{for } i = 1, 2, \dots, N; t = 1, 2, \dots, T$$

[2] A similar analysis can be made using the column sum norm of $\boldsymbol{\Sigma}_N$, defined by $\|\boldsymbol{\Sigma}_N\|_1 = \sup_j \sum_{i=1}^{N} |\sigma_{ij,x}|$.

where $\boldsymbol{f}_t = (f_{1t}, f_{2t}, \ldots, f_{mt})'$ is an $m \times 1$ vector of unobserved factors, and $\boldsymbol{\beta}_i$ is the associated vector of factor loadings (m is fixed). Stacking over cross-section units we get

$$\boldsymbol{x}_t = \boldsymbol{\beta}' \boldsymbol{f}_t + \mathbf{u}_t, \text{ for } t = 1, 2, \ldots, T \tag{3}$$

where $\boldsymbol{x}_t = (x_{1t}, x_{2t}, \ldots, x_{Nt})'$, $\boldsymbol{f}_t$ is specified above, $\mathbf{u}_t = (u_{1t}, u_{2t}, \ldots, u_{Nt})'$ and $\boldsymbol{\beta}$ is the associated matrix of factor loadings: $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_N)'$, $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{im})'$. We specify the loadings as follows:

$$\begin{aligned} \beta_{ij} &= v_{ij} \text{ for } i = 1, 2, \ldots, [N^{\alpha_j}], \\ \beta_{ij} &= \tilde{v}_{ij}, \text{ for } i = [N^{\alpha_j}] + 1, [N^{\alpha_j}] + 2, \ldots, N \end{aligned} \tag{4}$$

where $1/2 < \alpha_1 \leq 1$, and $\{v_{ij}\}_{i=1}^{[N^\alpha]}$ is an independent and identically distributed (i.i.d.) sequence of random variables with mean $\mu_{v_j} \neq 0$ and variance $0 < \sigma^2_{v_j} < \infty$. Without loss of generality, $\alpha := \alpha_1 \geq \alpha_j$, $j = 2, \ldots, m$. Also, $\sum_{i=[N^\alpha]+1}^{N} \tilde{v}_{ij} = O_p(1)$. Further conditions are discussed in the next section. At this point the above conditions are sufficient to motivate our estimator. As discussed above, the factor loadings in equation (4) are classified into two groupings: a category with pervasive effects that have a non-zero mean μ_v and a category whose impact is non-pervasive and fades as N increases. This loading setup infers that $N^{-1} \sum_{i=1}^{N} \beta_{ij}^2 = O_p\left(N^{\alpha_j - 1}\right)$, which is more general than the standard assumption in the factor literature that requires $N^{-1} \sum_{i=1}^{N} \beta_{ij}^2$ to have a strictly positive limit (see, for example, Assumption B of Bai and Ng, 2002). The standard assumption is satisfied only if $\alpha_j = 1$. Also, this implies a rate of decline for $\mathrm{SD}(\bar{x}_t)$ of $O\left(N^{\alpha-1}\right)$ so long as $\mu_v \neq 0$ and at least $N^{1/2}$ of the loadings have non-zero mean. As mentioned earlier, if $\mu_{v_j} = 0$, $j = 1, \ldots, m$, then $\mathrm{SD}(\bar{x}_t) = O\left(N^{-1/2}\right)$ for all α including the case $\alpha = 1$, but we do not see this case as very plausible, at least for macro and financial datasets.

Given the above setting, $\boldsymbol{\Sigma}_\beta = E(\boldsymbol{\beta}\boldsymbol{\beta}') - E(\boldsymbol{\beta})E(\boldsymbol{\beta}')$, with $\lambda_{\max}(\boldsymbol{\Sigma}_\beta) < K < \infty$. Further, $E(\mathbf{u}_t) = \mathbf{0}$, $\boldsymbol{\Sigma}_u = E\left(\mathbf{u}_t \mathbf{u}_t'\right)$, with $\lambda_{\max}(\boldsymbol{\Sigma}_u) < K < \infty$, $\mu_{f_j} = E(f_{jt}) = 0$, $\sigma^2_{f_j} = E(f_{jt} - \mu_{f_j})^2 = 1$, $j = 1, \ldots, m$. Finally, f_{jt} are distributed independently of $\boldsymbol{\beta}$ and of the idiosyncratic errors, $u_{it'}$, for all i, t and t'. Hence

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

Consider now the cross-sectional averages of the observables $\bar{x}_t = \boldsymbol{\iota}' \boldsymbol{x}_t / N$. Then

$$\sigma^2_{\bar{x}} = \mathrm{var}(\bar{x}_t) = N^{-2} \boldsymbol{\iota}' \mathrm{cov}(\boldsymbol{x}_t) \boldsymbol{\iota} = N^{-2} \boldsymbol{\iota}' \left[\boldsymbol{\Sigma}_\beta + \boldsymbol{\Sigma}_u\right] \boldsymbol{\iota} + \left[\frac{\boldsymbol{\iota}' E(\boldsymbol{\beta})}{N}\right] \left[\frac{\boldsymbol{\iota}' E(\boldsymbol{\beta})}{N}\right]' \tag{5}$$

But under (4)

$$N^{-1} \boldsymbol{\iota}' E(\boldsymbol{\beta}) = O(N^{\alpha-1}) + O(N^{-1})$$

Also

$$N^{-2} \boldsymbol{\iota}' \boldsymbol{\Sigma}_\beta \boldsymbol{\iota} \leq \left[N^{\alpha-2}\right] \lambda_{\max}\left(\boldsymbol{\Sigma}_\beta\right)$$

Using the above results in equation (5) we now have

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

DOI: 10.1002/jae

where

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

and μ_v^2 is defined in terms of μ_{v_j} in a way that will be discussed in detail in the next section. By assumption $\lambda_{\max}\left(\boldsymbol{\Sigma}_\beta\right) < K < \infty$, and hence under $1 \geq \alpha > 1/2$, we have

$$\sigma_{\bar{x}}^2 = \text{var}(\bar{x}_t) = \mu_v^2\left[N^{2\alpha-2}\right] + N^{-1}c_N + O(N^{\alpha-2}) \tag{8}$$

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

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

or

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

Initially, α can be identified from equation (9) using a consistent estimator of $\sigma_{\bar{x}}^2$, given by

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

where $\bar{x} = T^{-1}\sum_{t=1}^{T}\bar{x}_t$. This gives rise to the following estimator of α:

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

which is consistent and has a rate of convergence of $\ln(N)^{-1}$. Note here that the fourth term on the right-hand side of equation (9) is of a smaller order of magnitude than the previous three terms and can be ignored. However, it is important that the estimator of α also allows for the third term in equation (9). This can be achieved by replacing μ_v^2 with a suitable estimator. There are many alternatives for this estimation, which are discussed in detail in the next section. Our chosen estimator of μ_v^2 is obtained through identifying the significant slope coefficients of the cross-sectional averages, $\bar{x}_t$, from the ordinary least squares (OLS) regression of each unit x_{it} on $\bar{x}_t$, and we denote it by $\hat{\mu}_v^2$ (see Section 3.1 for details of the procedure).

Next, we discuss correcting the bias arising from the final term in equation (9). This is easily achieved in the case of exact factor models where the idiosyncratic errors are cross-sectionally independent, and $\boldsymbol{\Sigma}_u$ is a diagonal matrix. In this case, $c_N \equiv \bar{\sigma}_N^2 = N^{-1}\sum_{i=1}^{N}\sigma_i^2$, where σ_i^2 is the ith diagonal term of $\boldsymbol{\Sigma}_u$, and a consistent estimator of it is given by

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

DOI: 10.1002/jae

where $\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2$, $\hat{u}_{it} = x_{it} - \hat{\delta}_i \bar{x}_t$, and $\hat{\delta}_i$ denotes the OLS estimator of the regression coefficient of x_{it} on $\bar{x}_t$. Note that while $\hat{c}_N$, as an estimator for c_N, is motivated by appealing to an exact factor model, mild deviations from this model can be dealt with by using an alternative estimator for c_N, as discussed in Section 3.2. Using consistent estimators of $\sigma_{\bar{x}}^2$, μ_v^2, and c_N, we propose the following bias-adjusted estimator:

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

3. THEORETICAL DERIVATIONS

3.1. Main Results

In this section we present our formal theoretical results. Our first set of results characterizes the asymptotic behaviour of $\hat{\alpha}$. We make the following assumptions, where we state the full set of conditions, which were partly discussed in Section 2, for convenience.

Assumption 1. The factor loadings are given by

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

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

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

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

where $\boldsymbol{v}_{ft}$ is a sequence of i.i.d. random variables with mean zero and a finite variance matrix, $\boldsymbol{\Sigma}_{v_f}$, and uniformly finite φth moments for some $\varphi > 4$. The matrix coefficients, $\boldsymbol{\psi}_{fj}$, satisfy the absolute summability condition

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

such that $\{\zeta(\varphi-2)\}/\{2(\varphi-1)\} \geq 1/2$. $\boldsymbol{f}_t$ is distributed independently of the idiosyncratic errors, $u_{it'}$, for all i, t and t', and $f_{jt} \perp f_{st}$, $j \neq s$, $j, s = 1, \ldots, m$.

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

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

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

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

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

Assumptions 2 and 3 are mostly straightforward specifications of the factor and error processes assuming a linear structure with sufficient restrictions to enable the use of central limit theorems. Note that Assumption 3 rules out the existence of cross-sectional dependence in the error terms. This may be considered restrictive, but relaxing it is not straightforward. While this condition will be relaxed in Section 3.2, we choose not to discuss the fully general case at this point, as it will detract from the main exposition with complicated but, ultimately, not very significant methodological amendments. Further, the proposed solution that is presented in the next section, while effective in small samples, cannot be fully justified theoretically for small values of α. This issue is discussed in detail in the next section.

First, note that

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

and

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

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

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

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

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

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

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

where $\boldsymbol{S}_{ff} = (s_{jo,f}) = \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}_T\right)\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}_T\right)', j,o = 1,2,\ldots,m, \bar{\boldsymbol{f}}_T = T^{-1}\sum_{t=1}^{T}\boldsymbol{f}_t, \boldsymbol{\Sigma}_{ff} = diag\left(\sigma_{f_j}^2\right) = I, \boldsymbol{\mu}_f = E(\boldsymbol{f}_t) = (\mu_{f_1},\ldots,\mu_{f_m})'$, and $\boldsymbol{\mu}_v = [E(\boldsymbol{v}_j)] = (\mu_{v_1},\ldots,\mu_{v_m})', \boldsymbol{v}_j = \left(v_{1j},\ldots,v_{[N^{\alpha_j}]j}\right)'$. Further, define $\mu_v^2 = \sum_{j=1}^{m}\mu_{v_j}^2$.

Our exposition in Section 2 suggests that $\hat{\alpha}$, as an estimator of α, is subject to two sources of bias, $\frac{\ln(\mu_v^2)}{2\ln(N)}$ and $\frac{c_N}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}$, where the latter bias corresponds to the last part of equation (13) in the multiple-factor case. This can be corrected using a first-order accurate estimator given by $\frac{\hat{c}_N}{N\hat{\sigma}_{\bar{x}}^2}$ or a second-order bias correction given by $\frac{\hat{c}_N}{\ln(N)N\hat{\sigma}_{\bar{x}}^2}\left(1+\frac{\hat{c}_N}{N\hat{\sigma}_{\bar{x}}^2}\right)$, where $\hat{c}_N$ is defined in equation (12). We denote the estimators that make use of these corrections by

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

and

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

We now introduce the main theorem of the paper.

Theorem 1.

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

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

where

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

d_T is defined by equation (20)

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

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

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

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

(c) Continue to assume that $\alpha = \alpha_1 = \alpha_2 = \ldots = \alpha_m > 1/2$, and $\alpha > 1/2$, then

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

(d) Further, if either

DOI: 10.1002/jae

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

or if

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

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

$$\begin{aligned} \omega &= \lim_{N,T\to\infty}\left[\frac{\min(N^\alpha, T)}{T}V_{\overline{f_1^2}} + \frac{\min(N^\alpha, T)}{N^\alpha}\frac{4\sigma_{v_1}^2}{\mu_{v_1}^2}\right] \\ V_{\overline{f_1^2}} &= \mathrm{var}\left(\tilde{f}_{1t}^2\right) + 2\sum_{i=1}^{\infty}\mathrm{cov}\left(\tilde{f}_{1t}^2, \tilde{f}_{1t-i}^2\right) \end{aligned} \tag{26}$$

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

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

(e) Finally, if $\alpha = \alpha_1 > \alpha_2 \geq \alpha_3 \ldots \geq \alpha_m \geq 0$ but neither equation (24) nor (25) holds, then equations (21), (22) and (23) hold with ω replacing ω_m, and

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

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

1. Run the OLS regression of x_{it} on a constant and $\bar{x}_t$ and denote the estimated coefficient of $\bar{x}_t$ by $\hat{\delta}_i$, for $i = 1, 2, \ldots, N$.
2. Compute the t-ratio associated with the ith coefficient, $\hat{\delta}_i, i = 1, 2, \ldots, N$, as $z_{\hat{\delta}_i} = \hat{\delta}_i/\mathrm{s.e.}\left(\hat{\delta}_i\right)$.
3. Construct

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

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

4. Estimate μ_v by

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

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

The critical values, $c_{p_i,N}$, can be set using the multiple testing approaches of Bonferroni (1935, 1936) or Holm (1979). Both approaches deal with the multiple testing problem without making any assumptions about the cross-dependence of the underlying N individual t tests.[3] But Holm's approach is less conservative and uses different critical values across the units. To be more specific, let $t_i = \left|z_{\hat{\delta}_i}\right|$, for $i = 1, 2, \ldots, N$, and sort these t-ratios in descending order, such that $t_{(1)} > t_{(2)} > \ldots > t_{(N)}$, with associated critical values, $c_{p_{(i)},N}$. Suppose also that under the null hypothesis $\beta_{i1} = 0$, $z_{\hat{\delta}_i}$ is asymptotically distributed as $N(0,1)$, with the cumulative distribution function $\Phi(.)$. Then, under Bonferroni's approach, $c_{p_{(i)},N} = \Phi^{-1}\left(1 - \frac{p}{2N}\right)$, which is the same for all units, whereas under Holm's approach $c_{p_{(i)},N} = \Phi^{-1}\left(1 - \frac{p}{2(N-i)}\right)$ corresponding to $t_{(i)}$.

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

3.2. Extensions

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

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

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

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

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

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

and

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

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

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

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

where

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

and $\hat{e}_{it} = x_{it} - \hat{\varrho}_i \widehat{pc}_t$, $\widehat{pc}_t$ is the first principal component of $x_{it}, i = 1, 2, \ldots, N$, and $\hat{\varrho}_i$ denotes the OLS estimator of the regression coefficient of x_{it} on $\widehat{pc}_t$. The use of cross-sectional averages, $\bar{x}_t$, in place of $\widehat{pc}_t$ to compute $\hat{e}_{it}$ does not help in estimation of c_N since $\sum_{i=1}^{N}\left(x_{it} - \hat{\delta}_i \bar{x}_t\right) = 0$, where $\hat{\delta}_i$ is the OLS slope coefficient in the regression of x_{it} on $\bar{x}_t$ and suggests setting $\tilde{c}_N$ to zero. In a multiple-factor setting, additional principal components are needed to filter out any remaining cross-sectional error dependencies. Proving the consistency of $\tilde{c}_N$ is challenging. For the values of α where use of this estimator is needed ($a < 2/3$) it is not even clear whether factors can be estimated consistently. Kapetanios and Marcellino (2010) are not able to show this result and to the best of our knowledge it has not been proven elsewhere. Even if such a result held, it would not automatically ensure the consistency of $\tilde{c}_N$. Perhaps more relevantly, in that region of α its estimation is challenging even under strict assumptions, since its identification, while theoretically possible, is difficult. We note that we present Monte Carlo results based on $\tilde{c}_N$ when we carry out Monte Carlo experiments with cross-sectionally dependent idiosyncratic components. However, we have also carried out these experiments using $\hat{c}_N$ and these results bear out the theoretical outcome that, for $\alpha \geq 2/3, \tilde{c}_N$ and $\hat{c}_N$ provide asymptotically equivalent estimators.[4] Therefore, we believe that, both on theoretical and practical grounds, proving the consistency of $\tilde{c}_N$ is beyond the scope of the paper.

Up to now we have analysed estimators of the exponent of cross-sectional dependence assuming that factor loadings take the form given in Assumption 1. We briefly examine an alternative formulation (discussed in Section 2) which is mathematically convenient, although it is more difficult to justify from an economic perspective as it assumes that all factor loadings fall at the same rate. More specifically, consider the following alternative formulation for a one-factor setting.

[4] These results are available upon request.

DOI: 10.1002/jae

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

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

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

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

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

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

and its first bias-corrected version is given by

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

Under the alternative formulation equation (32), there is no need for further bias corrections. Then, the next Corollary follows (a proof is provided in supplementary Appendix II).

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

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

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

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

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

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

where $v_{Ni} = \breve{v}_i + \zeta_{Ni}$ and $\{\breve{v}_i\}_{i=1}^{N}$ is an i.i.d. sequence of random variables with mean $\mu_{\breve{v}} \neq 0$ and variance $\sigma_{\breve{v}}^2 < \infty$. Lemma 14 provides general conditions for this assumption, under which our theoretical results hold. In this remark, we explore a leading case of departure from Assumption 5 that is covered by Lemma 14. Without loss of generality, we order the cross-section units such that $\zeta_{Ni} = N^{(1-\alpha)/2}\eta_i$ for $i = 1, 2, \ldots, M$, 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 $M = o\left(N^{\alpha}\right)$.

4. MONTE CARLO STUDY

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

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

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

$$f_{jt} = \rho_j f_{j,t-1} + \sqrt{1-\rho_j^2}\zeta_{jt}, \ \ j = 1, 2, \ \text{for } t = -49, -48, \ldots, 0, 1, \ldots, T \tag{37}$$

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

The shocks follow an AR(1) process:

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

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

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

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

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

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

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

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

DOI: 10.1002/jae

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

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

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

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

and

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

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

Experiment 1. Here we use a basic design of equation (36) where the factors f_{jt}, for $j = 1, 2$, are serially uncorrelated; namely, we set $\rho_j = 0.0$ for $j = 1, 2$, in equation (37).

Experiment 2. Under this experiment we use the same design as in Experiment 1, but allow for temporal dependence in the factors; namely we set $\rho_j = 0.5$ for $j = 1, 2$, in equation (37).

Experiment 3. Under this experiment we use the same design as in Experiment 1, but we allow for departure of the idiosyncratic errors from normality and generate the idiosyncratic errors as $\varepsilon_{it} \sim$ i.i.d.$[(\chi^2(2) - 2)/4], i = 1, 2, \ldots, N$.

Experiment 4. The design for this experiment is as in Experiment 1, but allows the errors, u_{it}, to be cross-sectionally dependent according to a first-order spatial autoregressive model. Let $\mathbf{u}_t = (u_{1t}, u_{2t}, \ldots, u_{Nt})'$, and set $\mathbf{u}_t$ as

$$\mathbf{u}_t = \mathbf{Q}\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_t = \sigma_\varepsilon \boldsymbol{\eta}_t; \quad \boldsymbol{\eta}_t \sim \text{i.i.d. } N(\mathbf{0}, \mathbf{I}_N)$$

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

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

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

For all experiments, we consider the values of $\alpha = 0.70, 0.75, \ldots, 0.90, 0.95, 1.00$, $N = 50, 100, 200, 500, 1000$ and $T = 100, 200, 500$, and base them on 2000 replications. For each replication, the values of $\alpha, \alpha_2, d_i, \rho_j, \phi_i, \varsigma$ and $\mathbf{S}$ are given as set out above. These parameters are fixed across all replications. The values of $v_{ij}, j = 1, 2$ are drawn randomly (N of them) for each replication.

DOI: 10.1002/jae

In all experiments, we present bias and RMSE results for the bias-adjusted estimator $\mathring{\alpha}$ given by equation (13), where μ_{v_1} is estimated using Holm's approach to address the associated multiple testing problem. For experiments 1–3, we use $\hat{c}_N$ given by equation (12) to estimate c_N, while for experiment 4 we use $\tilde{c}_N$, given by equation (30). All results are scaled up by 100.

4.1. Summary of the Results

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

The results for Experiment 2, where the factors are allowed to be serially correlated, are summarized in the right-hand panel of Table A.1. As compared to the baseline case, we see a marginal deterioration in the results, particularly for relatively small values of N, T and α, but these differences tend to vanish as N and T are increased.

The results of Experiment 3, where the idiosyncratic errors, u_{it}, are allowed to be non-normal, are summarized in the left-hand panel of Table A.2. As can be seen, the results are slightly affected by the non-normality of the error terms when α is relatively small. Consistent with the baseline case of Experiment 1, both the bias and RMSE of $\mathring{\alpha}$ fall gradually as N, T and α are increased.

Finally, the effects of allowing for weak cross-sectional dependence in the idiosyncratic errors, u_{it}, on estimation of α are summarized in the right-hand panel of Table A.2 for Experiment 4. 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 those reported in Table A.1, for the baseline experiments.[7] However, one would expect greater distortions as θ is increased, although the effects of introducing weak dependence in the idiosyncratic errors are likely to be less pronounced if higher values of α are considered. For values of α near the borderline value of $1/2$, it will become particularly difficult to distinguish between factor- and spatial-dependent structures. In order to illustrate this point we also consider the case when $\theta = 0.4$. Results in Table A6 of supplementary Appendix VI show some deterioration in the bias, RMSE, size and power of the α estimator, especially for smaller α and N.

In line with Experiment 1, we show the full set of bias, RMSE, size and power results based on $\mathring{\alpha}$ for the remaining Experiments 2–4. All additional results and their discussion can be found in supplementary Appendix VI (see Tables A2–A5).

The Monte Carlo results clearly illustrate the potential utility of the estimation and inferential procedure proposed in the paper for the analysis of cross-sectional dependence. The results are broadly in agreement with the theory and are reasonably robust to departures from the basic model assumptions.

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

DOI: 10.1002/jae

Although the results tend to deteriorate slightly when we consider serially correlated factors or weak error cross-sectional dependence, the estimated values of α tend to retain a high degree of accuracy even for moderate sample sizes. It is also worth bearing in mind that in most empirical applications the interest will be on estimates of α that are close to unity, as it is for these values that a factor structure makes sense, as compared to spatial or other network models of cross-sectional dependence. It is therefore helpful that the small-sample performance of the proposed estimator improves when values of α close to unity are considered.

5. EMPIRICAL APPLICATIONS

In this section, we provide estimates of the exponent of cross-sectional dependence, α, for a number of panel datasets used extensively in economics and finance.[8] Specifically, we consider three types of datasets: quarterly cross-country data used in global modelling, large quarterly datasets used in empirical factor model literature, and monthly stock returns on the constituents of the S&P 500 index. We denote the typical elements of these datasets by y_{it}. The observations were standardized as $x_{it} = (y_{it} - \bar{y}_i)/s_i$, where $\bar{y}_i$ and s_i are the sample means and standard deviations of y_{it} for $t = 1, 2, \ldots, T$.

Before providing estimates of the exponent of cross-sectional dependence for these datasets, however, we first need to verify that the degree of cross-dependence in these datasets is sufficiently large. Recall that α is identifiable only if $\alpha > 1/2$. To this end we first apply the recent test of weak cross-sectional dependence (CD) developed by Pesaran (2015) to these datasets. The CD test statistic is defined by

$$CD_{NT} = \left[\frac{TN(N-1)}{2}\right]^{1/2} \widehat{\bar{\rho}}_N \tag{38}$$

where

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

and $\hat{\rho}_{ij}$ is the pair-wise correlation coefficient of x_{it} and x_{jt}. Pesaran (2015) shows that when $T = O(N^d)$ for some $0 < d \leq 1$, then the implicit null of the CD test is given by $0 \leq \alpha < (2-d)/4$, and it is asymptotically distributed as $N(0, 1)$. In our applications, N and T are of the same order of magnitude and $d \approx 1$.[9]

5.1. Cross-Country Dependence of Macro Variables

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

[8] In all empirical applications, we use Holm's approach when implementing the procedure described on pages 10–11. Results using Bonferroni's method are available upon request.

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

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

DOI: 10.1002/jae

Table I. Exponent of cross-country dependence of macro variables

	N	T	$\mathring{\alpha}^*_{0.05}$	$\mathring{\alpha}$	$\mathring{\alpha}^*_{0.95}$
Real GDP growth, q/q	33	122	0.923	0.977	1.031
Inflation, q/q	33	123	0.915	0.978	1.041
Real equity prices, q/q	26	122	0.928	1.001	1.074

*90% level confidence bands. q/q stands for quarter on quarter change.

The CD statistics turned out to be 44.32, 88.34 and 90.69 for output growth, inflation and real equity prices, respectively; these are hugely statistically significant and reject the null hypothesis of weak cross-sectional dependence for all the three datasets, justifying the use of our procedure for estimation of α. Table I presents the bias-corrected estimates, $\mathring{\alpha}$, computed using available cross-country time series, x_{it}, over the period 1979:Q2 to 2009:Q4. Table I also reports the 90% confidence bands constructed following the procedure set out in supplementary Appendix VI. Although there are 33 countries in the GVAR dataset, not all variables are available for all the countries. For example, real equity prices are available only for 26 of the 33 countries.

Looking at the results of Table I for $\mathring{\alpha}$, we observe that the point estimates for all variables considered fall in a small range and indicate that approximately one-seventh of the variables are cross-sectionally weakly correlated, while the remaining ones belong to the strongly correlated group.[11] The exponent of cross-sectional dependence for real equity prices at 1.001 points to financial variables being strongly correlated. Similar estimates are also obtained for the macro variables. For real GDP growth and inflation, we obtain the estimates 0.977 and 0.978 respectively. The confidence bands all lie above 0.5 and do include unity (though marginally), suggesting that in these examples a factor structure might be a good approximation for modelling global dependencies. However, in some instances the value of $\alpha = 1$, typically assumed in the empirical factor literature, might be exaggerating the importance of the common factors for modelling cross-sectional dependence at the expense of other forms of dependencies that originate from trade or financial inter-linkages that are more local or regional rather than global in nature.

5.2. Within-Country Dependence of Macroeconomic Variables

An important strand in the empirical factor literature, influenced by the theoretical and empirical work of Stock and Watson (2002), uses factor models to estimate and forecast a few key macro variables such as output growth, inflation or unemployment rate with a large number of macro variables, which 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 datasets used in Eklund *et al.* (2010). For the USA, the dataset comprises 95 variables and cover the period 1960:Q2 to 2008:Q3. For the UK, the dataset covers 94 variables spanning the period 1977:Q1 to 2008:Q2.

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

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

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

DOI: 10.1002/jae

Table II. Exponent of within-country dependence of macro variables

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

*90% level confidence bands.

5.3. 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 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 S&P 500 index of the large-cap US equities market, and estimate α recursively using rolling samples of size 60 months (5 years). Owing 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 476 securities at the end of each month for the rolling samples of size 5 years. The 1-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 ith security in the sample inclusive of dividend payments (if any).[12] Recursive estimates of α were then computed using the standardized observations $x_{it} = (\tilde{r}_{it} - \bar{\tilde{r}}_i)/s_i$, where $\bar{\tilde{r}}_i$ 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 5 years rolling windows are given in Figure 1.[13] We also computed rolling standard errors for the estimates, $\mathring{\alpha}_t$, which, as discussed in Section VI of the supplementary Appendix, are conservative bands. Based on these standard errors, the 95% confidence bands of the recursive estimates were, on average, ± 0.03 around the point estimates for the rolling sample size considered. These bands are not shown in Figure 1, since we aim to highlight the time variations in the estimates of α.[14]

[12] For further details of data sources and definitions see Pesaran and Yamagata (2012).

[13] As in the previous two applications, we computed the CD statistic for all rolling 5-year windows. It is highly significant and justifies the use of our estimation procedure in all samples. Results of the rolling CD statistics are available upon request.

[14] The rolling estimates of α including their 95% confidence bands are shown in Figure 5 of supplementary Appendix VII.

DOI: 10.1002/jae

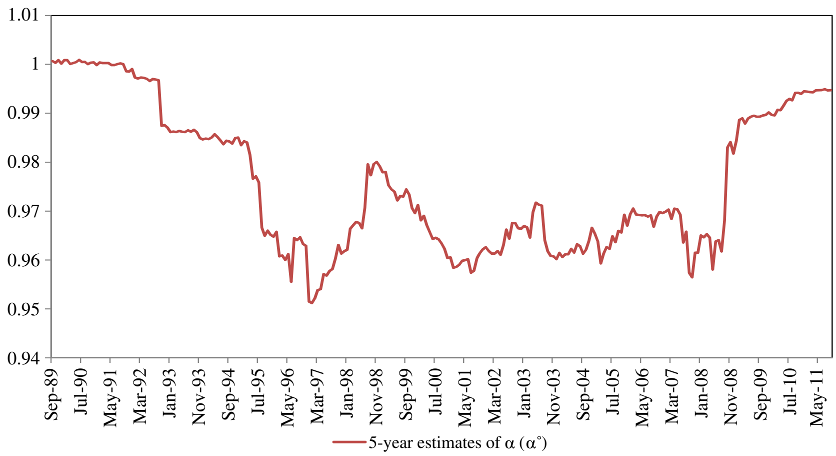

Figure 1. $\mathring{\alpha}_t$ associated with S&P 500 securities' excess returns: 5-year rolling samples

The figure covers 23 years of monthly recursive estimates of α, and yet these fall in a relatively narrow range of 0.951–1.001. These estimates clearly show a high degree of inter-linkages across individual securities, but at times the null hypothesis that $\alpha = 1$ is clearly rejected. The ratio of the number of times that the α estimate falls short of unity to the number of periods considered $(Pr(\mathring{\alpha} < 1))$ amounts to 0.917. Computing the same probability using the α estimates corresponding to the upper limit of the conservative confidence bands points to a relatively high value of 0.438.

More importantly, there are clear trends in the estimates of α. They fall from a high of 1.00 in 1990 to below 0.96 just before the burst of the dot-com bubble in 1999–2000. The period 1997–2000 saw some relatively pronounced fluctuations in α due to smaller crises caused by the Asian economic turmoil, LTCM and the bursting of the dot-com bubble. Over the period 2000–2008 the estimates of α hovered around the value of 0.965, before falling again slightly towards the end of 2008 at the time of the market crash, and then rising again to a level of 0.99 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 the securities market over the 1989–2011 period.

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 $(\overline{\hat{\rho}}_N)$ on securities included in S&P 500 index, using the 5-year rolling windows.[15] As the plots in these figures show, our estimates of α closely follow the rolling estimates of $\bar{\rho}_N$.

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 the CAPM and assumes that the single factor in CAPM regressions can be approximated by a stock market index. Under these assumptions, a direct estimate of α is given by $\hat{\alpha}_d = \ln(\hat{M})/\ln(N)$, where

[15] Denote 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 $\overline{\hat{\rho}}_N = [1/N(N-1)]\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}$, where N is the number of securities under consideration. Almost identical estimates are also obtained if we use returns instead of excess returns.

DOI: 10.1002/jae

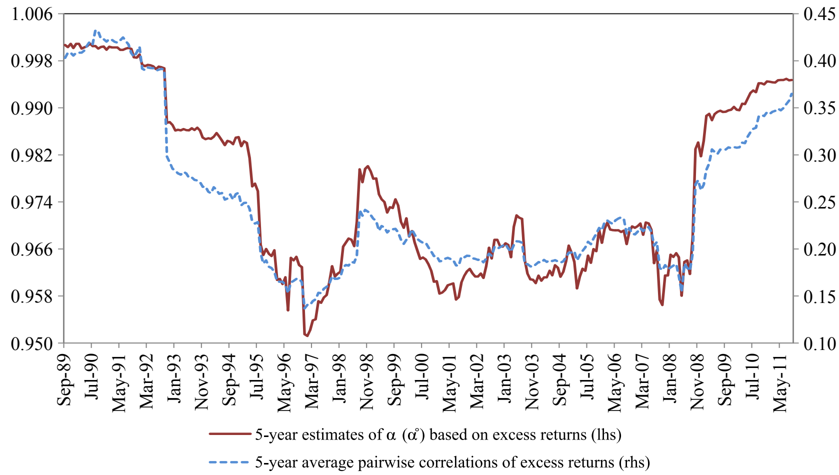

Figure 2. Average pair-wise correlations of excess returns for securities in the S&P 500 index and the associated $\mathring{\alpha}_t$ estimate computed using 5-year rolling samples

$\hat{M}$ denotes the estimated number of non-zero betas, and N is the total number of securities under consideration.[16] $\hat{M}$ can be consistently estimated (as N and $T \to \infty$) by the number of t-tests of $\beta_i = 0$ in the CAPM regressions

$$r_{it} - r_{ft} = a_i + \beta_i \left(r_{mt} - r_{ft}\right) + u_{it}, \text{ for } i = 1, 2, \ldots, N; t = 1, 2, \ldots, T \tag{39}$$

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},[17] and select 1% as the significance level of the tests. Such estimates of α obtained recursively using the 5-year rolling windows are shown in Figure 3. For ease of comparison, this plot also includes our (indirect) estimates of α based on the same datasets (except for the market return, r_{mt}, which is not used). The two sets of estimates co-move over most of the period, especially prior to the dot-com bubble and during the recent financial crisis. The correlation coefficient of the two sets of estimates is 0.815. 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 $\alpha > 1/2$, and in consequence is subject to a higher degree of uncertainty.[18] Nevertheless, it is reassuring that the direct and indirect estimates of α in this application tend to move together closely.

There is also a further consideration when comparing the estimates of α and α_d. Under CAPM the errors, u_{it}, in equation (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

[16] Note that $M = [N^{\alpha}]$, where M is the true number of non-zero betas.
[17] The 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).
[18] The 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.

DOI: 10.1002/jae

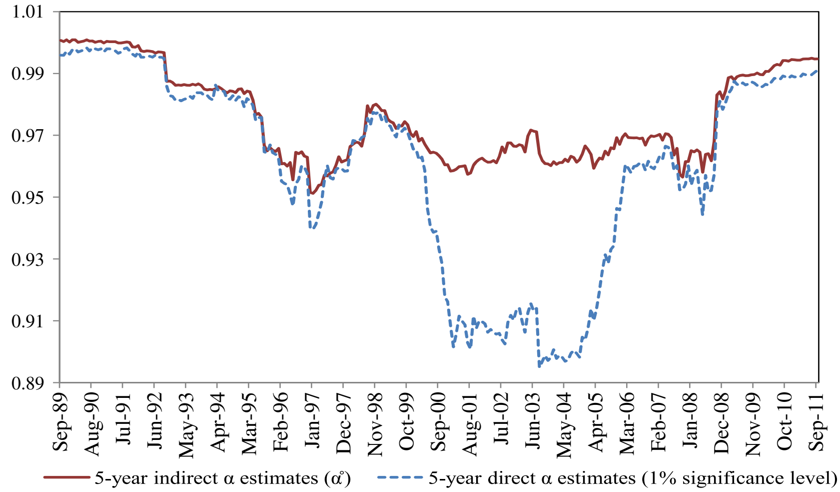

Figure 3. Direct ($\hat{\alpha}_d$) and indirect ($\mathring{\alpha}$) estimates of cross-sectional exponent of the market factor (using excess returns on S&P 500 securities) based on 5-year rolling samples

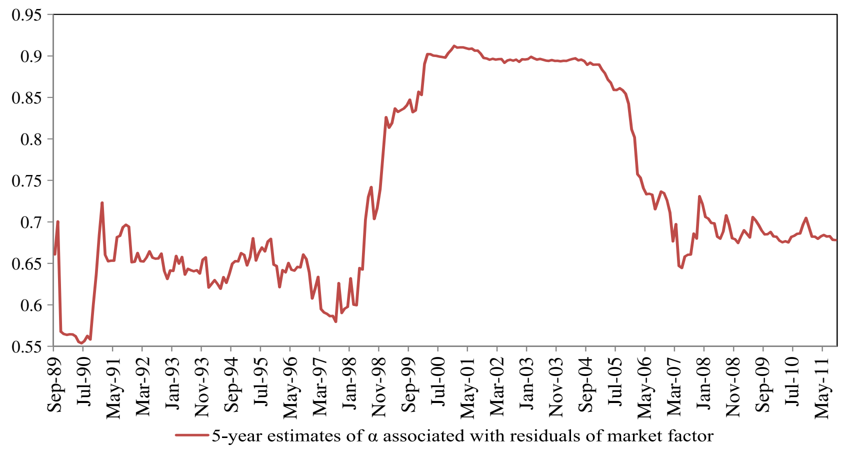

Figure 4. Estimates of cross-sectional exponent of residuals ($\mathring{\alpha}_u$) from CAPM regressions using 5-year rolling samples

compute consistent estimates of α_u.[19] The bias-adjusted estimates of α_u, denoted by $\mathring{\alpha}_u$, and computed using standardized residuals over 5-year rolling samples, are displayed in Figure 4.[20] 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 of 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 a peak of 0.92 in the middle of 2000, staying at around that level for the period up to 2006; it then begins to fall significantly after the start of the recent financial crisis, and currently stands at around 0.67.

[19] A formal proof and analysis when α is estimated from regression residuals is beyond the scope of the present paper.
[20] As before, the rolling estimates of α_u including their 95% confidence bands are shown in Figure 6 of supplementary Appendix VII.

Although special care must be exercised when interpreting these estimates (both because α_u is estimated using residuals and the fact that $\mathring{\alpha}$ 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 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.

6. CONCLUSIONS

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

In this paper, we propose a relatively simple method of measuring the extent of interconnections in large panel datasets in terms of a single parameter that we refer to as the exponent of cross-sectional dependence. We find that this exponent can accommodate a wide spectrum of cross-sectional dependencies in macro and financial datasets. We propose consistent estimators of the cross-sectional exponent and derive their asymptotic distribution. The inference problem is complex, as it involves handling a variety of bias terms and, from an econometric point of view, has noteworthy characteristics such as non-standard 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 datasets. In the first two cases, we find that the results of the empirical analysis accord with prior intuition. For individual securities in the S&P 500 index, the estimates of cross-sectional exponents are systematically high but at times 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-sectional exponent of $\alpha = 1$, which, as our empirical applications suggest, may be too restrictive, and it is important that implications of this assumption's failure are investigated. Chudik *et al.* (2011), Kapetanios and Marcellino (2010) and Onatski (2012) discuss some of these implications, namely that when $1/2 < \alpha < 1$ factor estimates are consistent but their rates of convergence are different (slower) as compared to the case where $\alpha = 1$, and in particular their asymptotic distributions may need to be modified. In some cases, such as when $\alpha < 3/4$ it is not even clear if factor estimates are consistent. Further, when $\alpha < 1$, methods used to determine the number of factors in large datasets, discussed for example by Bai and Ng (2002), Onatski (2009), Kapetanios (2010), Alessi *et al.* (2010) and Ahn and Horenstein (2013), are invalid and can select the wrong number of factors, even asymptotically.[21] 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).

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

ACKNOWLEDGEMENTS

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

REFERENCES

Ahn SC, Horenstein AR. 2013. Eigenvalue ratio test for the number of factors. *Econometrica* **81**(3): 1203–1227.

Alessi L, Barigozzi M, Capasso M. 2010. Improved penalization for determining the number of factors in approximate static factor models. *Statistics and Probability Letters* **80**(23-24): 1806–1813.

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

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

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

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

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

Cesa-Bianchi A, Pesaran M H, Rebucci A, Xu T. 2012. China's emergence in the world economy and business cycles in Latin America. *Economia* **12**(2): 1–75.

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

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

Chudik A, Pesaran MH, Tosetti E. 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: Oxford.

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

Efron B. 2010. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction*. Cambridge University Press: Cambridge, UK.

Eklund J, Kapetanios G, Price S. 2010. Forecasting in the presence of recent structural change. Working paper 406. Bank of England.

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

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

Forni M, Giannone D, Lippi M, Reichlin L. 2009. Opening the black box: structural factor models versus structural VARs. *Econometric Theory* **25**: 1319–1347.

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

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, Marcellino M. 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.

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

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

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

Pesaran MH. 2015. Testing weak cross-sectional dependence in large panels. *Econometric Reviews* **34**: 1089–1117.
Pesaran MH, Yamagata T. 2012. Testing CAPM with a large number of assets. Working paper 1210, University of Cambridge.
Pesaran MH, Schuermann T, Weiner S. 2004. Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business and Economic Statistics* **22**(2): 129–162.
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 SA. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* **13**: 341–360.
Ross SA. 1977. The capital asset pricing model (CAPM), short-sale restrictions and related issues. *Journal of Finance* **32**: 177–183.
Sharpe W. 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* **19**(3): 425–442.
Stock JH, Watson MW. 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* **20**(1): 147–162.

APPENDIX A: PROOFS OF THEOREMS

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

A.1. Proof of Theorem 1

We start by noting that

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

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

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

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

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

where

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

DOI: 10.1002/jae

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

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

Hence

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

Then,

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

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

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

or

$$\begin{aligned}2\ln(N)\left(\hat{\alpha}-\alpha\right) - \ln\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N\right) &= \frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N'\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} \\ &\quad + \frac{\left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\boldsymbol{\beta}}_N'\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} + O_p\left(N^{-\alpha}\right) + o_p(B_{N,T})\end{aligned} \tag{A.3}$$

where

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

DOI: 10.1002/jae

Consider the first term of the right-hand side of equation (A.3). We have

$$\frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t-\bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N'\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N}=\frac{\frac{2}{\sqrt{TN}}N^{\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\left[\boldsymbol{\Sigma}_{ff}^{-1/2}\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t-\bar{\boldsymbol{f}}\right)\sqrt{N}\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N'\boldsymbol{S}_{ff}^{1/2}\boldsymbol{S}_{ff}^{1/2}\boldsymbol{\Sigma}_{ff}^{-1/2}\bar{\boldsymbol{\beta}}_N} \tag{A.4}$$

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

$$\boldsymbol{\Sigma}_{ff}^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^T\left(\boldsymbol{f}_t-\bar{\boldsymbol{f}}\right)\left(\sqrt{N}\bar{u}_t\right)\to_p N\left(0,\bar{\sigma}_N^2 I_m\right) \tag{A.5}$$

where $\overline{\sigma}_N^2$ is defined in lemma 2.

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

$$\begin{aligned}
&\frac{1}{\bar{\boldsymbol{\beta}}_N'\bar{\boldsymbol{\beta}}_N}-\frac{1}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N}\\
&=\frac{1}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N+2N^{\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho+N^{-2}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho}-\frac{1}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N}\\
&=\frac{-N^{\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho-N^{-2}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho}{N^{4\alpha-4}\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N\right)^2+N^{3\alpha-4}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N+N^{2\alpha-4}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N}\\
&=-\left[N^{2-3\alpha}\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N\right)^{-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho+N^{2-4\alpha}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N\right)^{-2}\right]\\
&\quad\times\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N+N^{-\alpha}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho+N^{-2\alpha}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho\right)^{-1}\\
&=O_p\left(N^{2-3\alpha}\right)
\end{aligned}$$

and hence

$$\frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t-\bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N'\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N}=O_p\left(T^{-1/2}N^{1/2-\alpha}\right)+O_p\left(T^{-1/2}N^{1/2-2\alpha}\right) \tag{A.6}$$

Consider now the second term on the right-hand side of equation (A.3). We use equation (A.6) again. Note that since, by Lemma 1 and Theorems 17.5 and 19.11 of Davidson (1994), $\sqrt{NT}\bar{u}=O_p(1)$, and, since $\boldsymbol{S}_{ff}\boldsymbol{\Sigma}_{ff}^{-1}=1+O_p(T^{-1/2})$ where $0<\boldsymbol{\Sigma}_{ff}<\infty$,

$$\frac{\bar{u}^2}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)} \tag{A.7}$$

$$=\frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)} \tag{A.8}$$

$$=O_p\left(T^{-1}N^{1-2\alpha}\right) \tag{A.9}$$

Similarly,

$$\begin{aligned}
&\frac{\frac{1}{T}\sum_{t=1}^{T}\bar{u}_t^2}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\\
&=\frac{\frac{1}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\sqrt{N}\bar{u}_t\right)^2-\bar{\sigma}_N^2\right]+\sqrt{T}\bar{\sigma}_N^2\right\}}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\\
&=\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]+\sqrt{T}\right\}}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\\
&=\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\\
&\quad+\frac{\bar{\sigma}_N^2}{N\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}
\end{aligned} \tag{A.10}$$

Note that

$$\frac{\bar{\sigma}_N^2}{N\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}-\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}=O_p(N^{1-3\alpha}) \tag{A.11}$$

Also, by Lemma 3,

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

and

$$\begin{aligned}
&\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]\right\}}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\\
&\quad=O_p(T^{-1/2}N^{1-2\alpha})+O_p(T^{-1/2}N^{1-3\alpha})
\end{aligned} \tag{A.12}$$

So

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

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

$$\hat{\alpha}-\alpha=O_p\left(\frac{1}{\ln(N)}\right) \tag{A.13}$$

DOI: 10.1002/jae

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

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

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

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

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

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

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

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

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

Similarly,

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

and

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

DOI: 10.1002/jae

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

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

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

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

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

Also, by Lemmas 7 and 9 we have

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

and

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

which prove the remainder of the theorem.

DOI: 10.1002/jae

Table A.1. Bias and RMSE (×100) for the $\mathring{\alpha}$ estimate of the cross-sectional exponent: case of cross-sectionally independent idiosyncratic errors, $N = 50, 100, 200, 500, 1000$ and $T = 100, 200, 500$ ($\alpha_2 = 2\alpha/3$, f_{jt} and $u_{it} \sim$ i.i.d. $N(0,1)$, $v_{ij} \sim$ i.i.d.U$(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1, 2$, $\mu_v = 0.87$, $\mu_{v_2} = 0.71$, $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2-\alpha)}\mu_{v_2}^2}$)

Experiment 1: serially uncorrelated factors ($\rho_j = 0.0, j = 1, 2$)									Experiment 2: serially correlated factors ($\rho_j = 0.5, j = 1, 2$)								
	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00		α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
$N\backslash T$					100				$N\backslash T$					100			
50	Bias	2.20	1.04	0.42	0.09	−0.18	0.03	−0.23	50	Bias	3.10	1.83	1.08	0.46	0.01	0.09	−0.26
	RMSE	3.54	2.47	1.80	1.28	0.90	0.54	0.24		RMSE	4.62	3.35	2.39	1.68	1.12	0.68	0.29
100	Bias	0.97	0.45	0.02	0.27	0.15	−0.02	−0.04	100	Bias	2.52	1.64	0.89	0.82	0.50	0.16	−0.07
	RMSE	2.02	1.44	0.93	0.71	0.48	0.29	0.06		RMSE	3.61	2.60	1.74	1.33	0.88	0.48	0.10
200	Bias	0.52	0.43	0.19	0.15	0.04	0.03	0.03	200	Bias	1.52	1.22	0.76	0.52	0.26	0.13	0.01
	RMSE	1.48	0.99	0.63	0.46	0.29	0.17	0.03		RMSE	2.56	1.87	1.25	0.88	0.54	0.31	0.06
500	Bias	0.13	0.05	0.09	0.03	0.04	0.03	0.05	500	Bias	1.18	0.70	0.49	0.25	0.15	0.06	0.03
	RMSE	0.68	0.46	0.35	0.23	0.16	0.10	0.06		RMSE	1.91	1.22	0.82	0.50	0.32	0.17	0.05
1000	Bias	0.00	0.01	0.05	0.01	0.03	0.00	0.06	1000	Bias	0.84	0.48	0.32	0.18	0.10	0.02	0.03
	RMSE	0.53	0.35	0.25	0.17	0.12	0.07	0.06		RMSE	1.42	0.87	0.57	0.36	0.23	0.12	0.05
					200									200			
50	Bias	4.62	2.95	1.76	0.96	0.29	0.23	−0.29	50	Bias	5.54	3.78	2.35	1.39	0.64	0.44	−0.31
	RMSE	5.41	3.77	2.58	1.75	1.10	0.71	0.30		RMSE	6.39	4.60	3.14	2.15	1.35	0.88	0.31
100	Bias	2.51	1.55	0.62	0.60	0.31	0.05	−0.10		Bias	3.50	2.42	1.33	1.12	0.68	0.33	−0.10
	RMSE	3.12	2.19	1.28	0.99	0.61	0.33	0.10		RMSE	4.25	3.08	1.94	1.51	0.98	0.58	0.10
200	Bias	0.45	0.54	0.28	0.21	0.08	0.04	−0.02	200	Bias	1.88	1.44	0.88	0.62	0.32	0.16	−0.02
	RMSE	1.14	0.96	0.61	0.45	0.28	0.15	0.02		RMSE	2.56	1.89	1.22	0.87	0.53	0.28	0.02
500	Bias	0.24	0.12	0.14	0.06	0.06	0.04	0.02		Bias	1.25	0.73	0.54	0.31	0.21	0.11	0.02
	RMSE	0.61	0.41	0.31	0.19	0.14	0.08	0.02		RMSE	1.67	1.04	0.73	0.46	0.30	0.16	0.02
1000	Bias	0.10	0.08	0.10	0.06	0.07	0.03	0.03	1000	Bias	0.86	0.53	0.39	0.24	0.17	0.08	0.03
	RMSE	0.38	0.26	0.21	0.14	0.10	0.05	0.03		RMSE	1.20	0.74	0.52	0.33	0.23	0.11	0.03
					500									500			
50	Bias	10.04	6.54	4.18	2.61	1.43	0.63	−0.35	50	Bias	9.57	6.63	4.66	3.02	1.57	0.74	−0.34
	RMSE	10.36	6.91	4.54	2.96	1.80	0.97	0.35		RMSE	10.01	7.03	5.06	3.42	1.97	1.08	0.35
100	Bias	4.02	3.06	1.75	1.38	0.79	0.31	−0.13	100	Bias	5.91	4.20	2.60	1.94	1.13	0.46	−0.13
	RMSE	4.28	3.30	2.05	1.63	1.02	0.53	0.13		RMSE	6.28	4.50	2.89	2.19	1.35	0.65	0.13
200	Bias	1.75	1.56	0.88	0.64	0.31	0.11	−0.05	200	Bias	3.25	2.54	1.55	1.06	0.56	0.26	−0.05
	RMSE	2.08	1.83	1.14	0.86	0.51	0.23	0.05		RMSE	3.65	2.81	1.80	1.26	0.73	0.37	0.05
500	Bias	0.90	0.47	0.31	0.13	0.09	0.04	−0.01	500	Bias	1.90	1.14	0.77	0.44	0.29	0.15	−0.01
	RMSE	1.17	0.72	0.48	0.27	0.17	0.09	0.01		RMSE	2.19	1.36	0.93	0.56	0.37	0.19	0.01
1000	Bias	0.66	0.23	0.14	0.07	0.06	0.0 2	0.00	1000	Bias	1.12	0.67	0.47	0.28	0.19	0.08	0.00
	RMSE	0.90	0.42	0.25	0.15	0.10	0.04	0.00		RMSE	1.35	0.83	0.57	0.35	0.23	0.11	0.00

Table A.2. Bias and RMSE (×100) for the $\mathring{\alpha}$ estimate of the cross-sectional exponent: case of two serially independent factors, $N = 50, 100, 200, 500, 1000$ and $T = 100, 200, 500$ ($\alpha_2 = 2\alpha/3$, f_{jt} and $u_{it} \sim$ i.i.d. $N(0,1)$, $v_{ij} \sim$ i.i.d.U$(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1, 2$, $\mu_v = 0.87$, $\mu_{v_2} = 0.71$, $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2-\alpha)}\mu_{v_2}^2}$)

Experiment 3: non-normal idiosyncratic errors ($\varepsilon_{it} \sim$ i.i.d.$\chi^2(2)$)									Experiment 4: spatially dependent idiosyncratic errors ($\theta = 0.2$)								
	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00		α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
$N\backslash T$					100				$N\backslash T$					100			
50	Bias	2.26	1.09	0.43	0.07	−0.20	0.00	−0.24	50	Bias	3.57	2.11	1.20	0.74	0.35	0.45	0.10
	RMSE	3.64	2.59	1.83	1.29	0.91	0.53	0.25		RMSE	4.77	3.33	2.27	1.61	1.06	0.74	0.11
100	Bias	1.24	0.60	0.12	0.30	0.17	−0.01	−0.05	100	Bias	1.62	0.89	0.37	0.53	0.39	0.18	0.09
	RMSE	2.29	1.56	0.98	0.73	0.49	0.29	0.06		RMSE	2.45	1.73	1.10	0.90	0.63	0.35	0.10
200	Bias	0.27	0.36	0.17	0.13	0.02	0.02	0.03	200	Bias	0.31	0.39	0.22	0.19	0.09	0.09	0.09
	RMSE	1.26	0.89	0.57	0.41	0.27	0.16	0.04		RMSE	1.23	0.97	0.66	0.49	0.33	0.20	0.09
500	Bias	0.14	0.04	0.08	0.02	0.03	0.02	0.05	500	Bias	0.13	0.03	0.07	0.02	0.04	0.05	0.07
	RMSE	0.72	0.49	0.35	0.24	0.17	0.10	0.06		RMSE	0.75	0.54	0.40	0.27	0.19	0.12	0.07
1000	Bias	0.02	0.01	0.05	0.01	0.03	0.00	0.05	1000	Bias	−0.01	−0.02	0.03	0.00	0.03	0.01	0.06
	RMSE	0.50	0.34	0.25	0.17	0.12	0.07	0.06		RMSE	0.56	0.39	0.28	0.20	0.14	0.08	0.07
					200									200			
50	Bias	4.73	2.97	1.79	1.00	0.33	0.22	−0.29	50	Bias	6.83	4.52	2.95	1.89	0.98	0.70	0.04
	RMSE	5.61	3.82	2.60	1.77	1.10	0.68	0.30		RMSE	7.52	5.17	3.57	2.46	1.50	0.99	0.04
100	Bias	2.10	1.29	0.49	0.51	0.30	0.04	−0.10	100	Bias	2.78	1.79	0.85	0.81	0.53	0.23	0.05
	RMSE	2.85	1.97	1.18	0.89	0.60	0.32	0.10		RMSE	3.42	2.39	1.43	1.14	0.76	0.40	0.05
200	Bias	0.64	0.64	0.31	0.23	0.08	0.04	−0.02	200	Bias	0.87	0.77	0.42	0.33	0.17	0.12	0.04
	RMSE	1.40	1.09	0.66	0.48	0.29	0.16	0.02		RMSE	1.52	1.19	0.74	0.55	0.35	0.21	0.04
500	Bias	0.28	0.13	0.14	0.06	0.06	0.04	0.02	500	Bias	0.28	0.12	0.14	0.07	0.08	0.07	0.04
	RMSE	0.67	0.42	0.31	0.20	0.14	0.08	0.02		RMSE	0.69	0.45	0.33	0.22	0.16	0.10	0.04
1000	Bias	0.11	0.07	0.10	0.06	0.07	0.03	0.03	1000	Bias	0.08	0.04	0.08	0.05	0.07	0.04	0.04
	RMSE	0.41	0.26	0.20	0.14	0.11	0.05	0.03		RMSE	0.44	0.30	0.22	0.15	0.12	0.06	0.04
					500									500			
50	Bias	8.70	5.81	3.87	2.47	1.31	0.66	−0.34	50	Bias	13.52	9.01	5.90	3.75	2.20	1.23	0.01
	RMSE	9.16	6.23	4.26	2.86	1.68	1.00	0.34		RMSE	13.93	9.39	6.24	4.07	2.51	1.46	0.01
100	Bias	4.64	3.33	1.92	1.45	0.84	0.30	−0.13	100	Bias	6.01	4.16	2.48	1.87	1.13	0.54	0.01
	RMSE	4.97	3.58	2.19	1.68	1.07	0.51	0.13		RMSE	6.36	4.38	2.71	2.06	1.30	0.69	0.01
200	Bias	2.15	1.77	0.95	0.67	0.32	0.12	−0.05	200	Bias	2.47	1.97	1.10	0.79	0.43	0.21	0.02
	RMSE	2.52	2.03	1.21	0.88	0.50	0.25	0.05		RMSE	2.78	2.22	1.35	0.99	0.59	0.31	0.02
500	Bias	0.88	0.45	0.30	0.13	0.09	0.04	−0.01	500	Bias	0.89	0.46	0.31	0.15	0.11	0.07	0.02
	RMSE	1.18	0.69	0.47	0.27	0.17	0.09	0.01		RMSE	1.21	0.74	0.49	0.29	0.19	0.11	0.02
1000	Bias	0.30	0.15	0.13	0.07	0.06	0.02	0.00	1000	Bias	0.27	0.13	0.11	0.06	0.06	0.03	0.01
	RMSE	0.55	0.32	0.22	0.14	0.09	0.04	0.00		RMSE	0.53	0.31	0.21	0.14	0.10	0.05	0.01

DOI: 10.1002/jae