TI 2025-045/III
Tinbergen Institute Discussion Paper

A General Randomized Test for Alpha

Daniele Massacci[1]
Lucio Sarno[2]
Lorenzo Trapani[3]
Pierluigi Vallarino[4]

1 King's College London

2 University of Cambridge, Centre for Economic Policy Research

3 Università di Pavia, University of Leicester

4 Erasmus University Rotterdam, Tinbergen Institute

A GENERAL RANDOMIZED TEST FOR ALPHA

DANIELE MASSACCI[1], LUCIO SARNO[2], LORENZO TRAPANI[3], PIERLUIGI VALLARINO[4]

Abstract. We propose a methodology to construct tests for the null hypothesis that the pricing errors of a panel of asset returns are jointly equal to zero in a linear factor asset pricing model — that is, the null of "zero alpha". We consider, as a leading example, a model with observable, tradable factors, but we also develop extensions to accommodate for non-tradable and latent factors. The test is based on equation-by-equation estimation, using a *randomized* version of the estimated alphas, which only requires rates of convergence. The distinct features of the proposed methodology are that it does not require the estimation of any covariance matrix, and that it allows for both N and T to pass to infinity, with the former possibly faster than the latter. Further, unlike extant approaches, the procedure can accommodate conditional heteroskedasticity, non-Gaussianity, and even strong cross-sectional dependence in the error terms. We also propose a de-randomized decision rule to choose in favor or against the correct specification of a linear factor pricing model. Monte Carlo simulations show that the test has satisfactory properties and it compares favorably to several existing tests. The usefulness of the testing procedure is illustrated through an application of linear factor pricing models to price the constituents of the S&P 500.

1. Introduction

Testing for the existence of "alpha" — i.e. the component of expected returns that cannot be explained as compensation for systematic risk — is central in asset pricing, and yet poses a number of issues. Available testing procedures are often marred by low power or poor size control, require strong assumptions on the data generating process (DGP), and generally

[1]King's Business School, King's College London, UK
[2]Cambridge Judge Business School and Girton College, University of Cambridge, UK; and Centre for Economic Policy Research (CEPR)
[3]Universita' di Pavia, Italy; and University of Leicester, UK
[4]Econometrics Institute, Erasmus School of Economics, The Netherlands
Key words and phrases. Testing for alpha, randomized tests, linear pricing factor models.
Acknowledgements. We are grateful for comments and suggestions to Marine Carrasco, Marcelo Fernandes, Yingying Li, Oliver Linton, Roberto Renò, Dick van Dijk, Dacheng Xiu, Chen Zhou and to participants in presentations held at: the 2nd FinML conference; the 2025 Annual SoFiE conference; Erasmus University Rotterdam; and ESSEC Business School. The usual disclaimer applies.

involve the estimation or inversion of (often large) covariance matrices (Giglio et al., 2022). In this paper, we propose a general methodology to test for "zero alpha" — that is, for the null that the pricing errors in a factor-based asset pricing model, are jointly equal to zero. In order to make the presentation easier to follow, we mainly focus on the following linear factor pricing model, with *tradable* and *observable* factors

$$y_{i,t} = \alpha_i + \beta_i' f_t + u_{i,t}, \quad 1 \le i \le N, \quad 1 \le t \le T, \tag{1.1}$$

where: $y_{i,t}$ is the excess return on the i-th security at time t; f_t is a K-dimensional vector of (tradable and observable) factors; β_i is a K-dimensional vector of factor loadings; and $u_{i,t}$ is a zero mean error term. Equation (1.1) is the workhorse model employed in asset pricing, and it encompasses several popular specifications such as e.g. the Capital Asset Pricing Model (CAPM) of Sharpe (1964), and the three- and five-factors models of Fama and French (1993) and Fama and French (2015). In the context of (1.1), we propose a novel approach to test for the null hypothesis that all the α_is are jointly equal to zero versus the alternative that at least one α_i is nonzero, viz.

$$\mathbb{H}_0 \quad : \quad \alpha_i = 0 \qquad \text{for all} \quad 1 \le i \le N, \tag{1.2}$$

$$\mathbb{H}_A \quad : \quad \alpha_i \neq 0 \qquad \text{for at least one } 1 \le i \le N. \tag{1.3}$$

Indeed, we build tests for the equivalent version of (1.2)

$$\mathbb{H}_0 : \max_{1 \le i \le N} \ |\alpha_i| = 0. \tag{1.4}$$

Testing for (1.4) is an interesting exercise *per se*. In the context of asset pricing, by (1.1) it follows that $\mathbb{E}(y_{i,t}) = \alpha_i + \beta_i' \mathbb{E}(f_t)$; hence, α_i represents the excess return on the i-th cross-sectional unit not explained by the K factors. In fact, under our factor tradability assumption, α_i is the *pricing error*. Hence, (1.4) corresponds to the null hypothesis that no pricing error is made by (1.1) for any of the N assets under consideration - in essence, being a test for the correct specification of (1.1).[1]

Arguably, the first contribution to propose a test for "zero alpha" in the context of linear pricing models is the paper by Gibbons et al. (1989, GRS henceforth), where an F-test is proposed for the joint null hypothesis that $\alpha_1 = \alpha_2 = ... = \alpha_N = 0$. Such an approach is entirely natural; however, it is predicated upon having several assumptions on the error terms $u_{i,t}$ such as Gaussianity, homogeneity, and independence across the time and the cross-sectional dimensions, and also the restriction that N is fixed with $N < T$, which may be unrealistic in many applications. In turn, this is needed because the test requires a (consistent) estimate of the $N \times N$ covariance matrix of the regression residuals, which subsequently needs to be inverted. Hence, it is apparent that the case $N > T$, often encountered in applied work, cannot be addressed with the GRS test. Fan et al. (2015), Gagliardini et al. (2016) and Pesaran and Yamagata (2024) propose several solutions towards this technical difficulty by developing average-type alpha tests — i.e. tests where individual statistics on each of the α_is are averaged — for the "joint asymptotics" case where $\min\{N, T\} \to \infty$. All these tests can deal with cross-sectional dependence among the errors terms (in essence, they can allow for cross-sectional dependence as long as a Central Limit theory holds for cross-sectional averages). However, their implementation still hinges on estimating a large dimensional, $N \times N$ covariance matrix, which requires several restrictions on the covariance structure of the errors, and which can become computationally intensive when certain estimators are considered (this is e.g. the case for the threshold estimator of Bickel and Levina, 2008, used by Gagliardini et al., 2016). Furthermore, all the tests mentioned above require some restrictions on the errors $u_{i,t}$ in (1.1); namely, all tests

[1]Note that the null in (1.4) differs from that of e.g. Giglio et al. (2021), who test N times whether the i-th cross-sectional unit has *non-positive* α_i. Moreover, from a methodological point of view, the paper by Giglio et al. (2021) is concerned with handling a *multiple* testing problem, while our procedure considers a one-shot test. See also Sullivan et al. (1999) and White (2000) for examples of related, bootstrap-based tests for alpha.

require the assumption of serially independent errors, which is bound to cause problems in the presence of conditional heteroskedasticity. The approaches by Fan et al. (2015) and Gagliardini et al. (2016) further require *Gaussian* errors. In a recent contribution with essentially the same set-up as in our model (1.1), Feng et al. (2022) propose a max-type test for the null that $\max_{1\leq i\leq N}|\alpha_i| = 0$; their test is based on calculating the maximum across the estimated alphas, and therefore its asymptotics hinges on an Extreme-Value-type argument, rather than a "central" one like the tests reviewed above. Their inference is still valid even in the presence of cross-sectional dependence in the error terms, but this can only be limited (i.e., weak); moreover, the assumption of serially independent errors is imposed, although this can be, arguably, relaxed. Ardia and Sessinou (2024) test whether all alphas are zero by combining p-values from asset-specific tests using the Cauchy combination approach proposed by Liu and Xie (2020). While their test can be applied under quite general assumptions, its implementation requires several tuning parameters whose values significantly impact the outcome of the procedure. Finally, Chernov et al. (2025) directly generalize the GRS test to the high-dimensional case by considering a ridge-regularized estimator of the $N \times N$ covariance matrix of the residuals; however, such an approach is still predicated upon assumptions as stringent as those of the aforementioned average-type tests.[2]

Testing methodology and the contribution of this paper

In this paper, we fill the gaps mentioned above, by proposing a test for (1.4) which allows for the errors to have: *(i)* arbitrary levels of cross sectional dependence, e.g. due to strong within- and between-industries (or countries) relations or productivity networks; *(ii)* (weak) serial dependence, including conditional heteroskedasticity phenomena such as, e.g., volatility clustering, or the "leverage effect" (Black, 1976); *(iii)* non-Gaussianity, such

[2]Other relevant references include Gungor and Luger (2016), Ma et al. (2020), and Raponi et al. (2020). Further, the correct specification of linear factor pricing models can also be assessed by testing whether they imply a pricing kernel with zero Hansen-Jagannathan distance (see e.g., Hansen and Jagannathan, 1997, Hodrick and Zhang, 2001, and Carrasco and Nokho, 2024).

as skewness and excess kurtosis, indeed relaxing some of the moment assumptions in the papers mentioned above; and also for *(iv)* the joint asymptotic case where $\min\{N,T\} \to \infty$, without requiring that $N < T$, and in fact allowing for N to even diverge faster than T as they pass to infinity. Importantly, our test statistic only requires estimation of the α_is unit by unit, thus being able to ignore the cross-sectional structure of the data. Therefore, the test does not require the estimation of any $N \times N$ covariance matrices, and requires virtually no tuning.

Whilst the technical details are described in the next sections, here we offer a preview of the main arguments, which are based on the construction of a *randomized* test statistic. In order to construct the test statistic, we estimate the α_is from N separate time series regressions, unit by unit: at no stage do we require joint estimation. Although any consistent estimator can be employed, here we use OLS, obtaining, say, $\widehat{\alpha}_i$; under standard assumptions, it can be expected that $\widehat{\alpha}_i$ will converge to α_i at a rate $T^{-1/2}$. We then pre-multiply each $|\widehat{\alpha}_i|$ by a function of T which diverges as $T \to \infty$, but at a slower rate than $O\left(T^{1/2}\right)$: hence, we obtain N statistics which – by construction – drift to zero when $\alpha_i = 0$, and diverge to positive infinity whenever $\alpha_i \neq 0$. We then perturb the resulting N statistics by adding to each of them a $\mathcal{N}(0,1)$ shock, with the N shocks forming an *i.i.d.* sequence. As a consequence, we obtain an N-dimensional sequence which, under $\mathbb{H}_0$, is (roughly) $i.i.d.\mathcal{N}(0,1)$ conditionally on the sample. Finally, we take the largest of these perturbed statistics as a test statistic for $\mathbb{H}_0$ in (1.4): conditionally on the sample, under $\mathbb{H}_0$ this is distributed as a Gumbel, whereas it diverges under the alternative that at least one asset is mis-priced. This methodology has, at least conceptually, some similarities with the one proposed in Fan et al. (2015), where a sequence is added to a test statistic, constructed so as to drift to zero under the null (thus introducing no distortion in the asymptotics under the null), and diverging under the alternative (so as to boost the power). However, unlike

Fan et al. (2015), we do not need to estimate at any stage the asymptotic covariance matrix of the estimated $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)'$, or define a high-dimensional weight matrix: only the individual, unit-by-unit, estimates of the α_is are required. Moreover, we only require a rate of convergence for the estimate $\widehat{\alpha}_i$. Hence, the assumptions on serial dependence and moment conditions can be relatively mild. Specifically, we assume that pricing factors and errors in (1.1) belong in a class of weakly dependent, stationary processes which encompasses virtually all the commonly used DGPs in econometrics and statistics, such as ARMA models, ARCH/GARCH models and even more complex nonlinear models.

To the best of our knowledge, our test is the first application of randomization to asset pricing, and in particular to the case of alpha testing. While entirely novel to this setting, randomized tests have been used in econometrics and statistics, particularly where a limiting distribution is unavailable or non-pivotal, or where its derivation requires excessively restrictive assumptions; although a comprehensive literature review goes beyond the scope of this paper, we refer to Corradi and Swanson (2006) for a first example of application of randomized tests in econometrics, and to the paper by He et al. (2023) for references.

We make at least four contributions to the current literature. First, whilst we focus on the specific case of "testing for alpha", we propose a novel methodology to construct tests involving a growing number of parameters with no need for joint estimation and, in essence, no need to take the dimensionality of the problem into account. In a similar spirit, although in (1.1) we assume that the common factors f_t are the same across all units, our approach can be readily extended e.g. to the case where pricing factors are heterogeneous across assets. Indeed, the factor structure could be so heterogeneous that no factor influences all the assets, which makes our testing procedure robust to the presence of weak and semi-strong pricing factors, and we do not need any factor to be "strong" in our model.[3] Second,

[3]We say that the k-th pricing factor is *strong* if $\beta_{i,k} \neq 0$ for all assets, i.e. for $i = 1, \ldots, N$. Conversely, we say that this factor is *weak* (resp. *semi-strong*) when $\beta_{i,k} \neq 0$ for $\lfloor N^\gamma \rfloor$ of the assets with $0 < \gamma < 1/2$

in the specific context of testing for alpha, we relax several technical conditions considered in the extant literature, allowing for serial and (even strong) cross-sectional dependence, conditional heteroskedasticity, thicker tails and a larger N/T ratio than allowed for in other contributions – hence, we expand the scope of the application of testing for zero alpha. Third, we strengthen the randomized test by proposing a decision rule which shows excellent control for Type I and Type II errors in simulations. Fourth and last, as mentioned above, we focus on the case of observable and tradable factors, but our methodology can also be extended to other frameworks; in Section 4, we study the extension to the cases of non-tradable and of latent factors; to the best of our knowledge, all extant contributions on testing for alpha do not consider these cases.

The remainder of the paper is organized as follows. We discuss our set-up and assumptions in Section 2. The hypothesis of interest and the randomized testing approach are discussed in Section 3; we report the de-randomized decision rule in Section 3.1, and in Section 3.2, we offer guidelines for the practical implementation of our procedure. In Section 4, we consider extensions to the cases of non-tradable and latent factors (Sections 4.1 and 4.2 respectively). Simulations are in Section 5, while Section 6 contains an empirical illustration. Conclusions and further lines of research are in Section 7. Further Monte Carlo evidence, technical lemmas and proofs are in the Supplement.

NOTATION. We define the probability space $(\Omega, \boldsymbol{B}, \mathbb{P})$, where Ω is the sample space with elements $\omega \in \Omega$, $\boldsymbol{B}$ is the space of events, and $\mathbb{P}$ denotes the probability function. Given a (possibly vector-valued) random variable X, $\mathbb{E}(X)$ is the mean and $\mathcal{V}(X)$ is the covariance viz. $\mathcal{V}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))(X - \mathbb{E}(X))'\right]$; further, we denote the $\mathcal{L}_\nu$-norm of X as $|X|_\nu = (\mathbb{E}|X|^\nu)^{1/\nu}$. The indicator function of a set $\mathcal{A}$ is denoted as $\mathbb{I}(\mathcal{A})$. We use: "a.s" for "almost sure(ly)"; "$\rightarrow$" to denote the ordinary limit; and "$\stackrel{a.s.}{\rightarrow}$" to denote almost

(resp. $1/2 \leq \gamma < 1$). This definition is also the one used by Chudik et al. (2011), and several others in this literature.

sure convergence. Orders of magnitude for an almost surely convergent sequence with multiparameter index Π, say s_{Π} are denoted as $O_{a.s.}(b_{\Pi})$ and $o_{a.s.}(b_{\Pi})$ when, respectively, $\mathbb{P}\left(\limsup_{\Pi\to\infty}\left|b_{\Pi}^{-1}s_{\Pi}\right|<\infty\right)=1$, and $b_{\Pi}^{-1}s_{\Pi}\overset{a.s.}{\to}0$, as $\Pi\to\infty$. Positive, finite constants are denoted as c_0, c_1, ... and their value may change from line to line. Other, relevant notation is introduced later on in the paper.

2. Model and assumptions

Recall our workhorse model (1.1)

$$y_{i,t}=\alpha_i+\beta_i' f_{i,t}+u_{i,t}.$$

We begin with a definition of weak dependence which we use throughout the paper.

Definition 2.1. *The sequence $\{m_t,-\infty<t<\infty\}$ forms an $\mathcal{L}_{\nu}$-decomposable Bernoulli shift if and only if it holds that $m_t=g\left(\eta_t,\eta_{t-1},...\right)$, where: (i) $g:S^{\infty}\to\mathbb{R}^k$ is a non random measurable function; (ii) $\{\eta_t,-\infty<t<\infty\}$ is an i.i.d. sequence with values in a measurable space S; (iii) $|m_t|_{\nu}<\infty$; and (iv) $\left|m_t-m_{t,\ell}^*\right|_{\nu}\leq c_0\ell^{-a}$, for some $c_0>0$ and $a>0$, where $m_{t,\ell}^*=g\left(\eta_t,...,\eta_{t-\ell+1},\eta_{t-\ell,t,\ell}^*,\eta_{t-\ell-1,t,\ell}^*...\right)$, with $\left\{\eta_{s,t,\ell}^*,-\infty<s,\ell,t<\infty\right\}$ i.i.d. copies of η_0 independent of $\{\eta_t,-\infty<t<\infty\}$.*

Decomposable Bernoulli shifts were firstly proposed in Ibragimov (1962), and they have proven a convenient way to model dependent time series, mainly due to their generality and to the fact that it is much easier to verify whether a sequence forms a decomposable Bernoulli shift than, for example, verifying mixing conditions. Virtually all the most common DGPs in econometrics and statistics can be shown to satisfy Definition 2.1. Liu and Lin (2009), *inter alia*, provide various theoretical results, and numerous examples including ARMA models, ARCH/GARCH sequences, and other nonlinear time series models (such as e.g. Random Coefficient AutoRegressive models and threshold models).

We are now ready to present our assumptions.

Assumption 2.1. *For all* $1 \leq i \leq N$ *and some* $\nu > 4$, $\{u_{i,t}, -\infty < t < \infty\}$ *is an* $\mathcal{L}_{\nu}$-*decomposable Bernoulli shift, with* $a > (\nu - 1) / (\nu - 2)$, $\mathbb{E}u_{i,t} = 0$, *and* $\min_{1 \leq i \leq N} \mathbb{E}u_{i,t}^2 > 0$.

Assumption 2.2. *For some* $\nu > 4$, $\{f_t, -\infty < t < \infty\}$ *is an* $\mathcal{L}_{\nu}$-*decomposable,* K-*dimensional Bernoulli shift, with* $a > (\nu - 1) / (\nu - 2)$ *and positive definite covariance matrix* $\mathcal{V}(f_t)$.

Assumption 2.3. *It holds that* $\mathbb{E}(f_t u_{i,t}) = 0$, *for all* $1 \leq i \leq N$.

Assumption 2.1 states that the errors $u_{i,t}$ can be (weakly) dependent across time, and may exhibit conditional heteroskedasticity. In contrast, as mentioned in the introduction, the tests by Fan et al. (2015), Gagliardini et al. (2016) and Pesaran and Yamagata (2024) all assume independence over time of $u_{i,t}$, thus being unable to accommodate conditional heteroskedasticity in asset returns. Moreover, Assumption 2.1 does not pose any restriction on the presence or extent of cross-sectional dependence in the errors: thus, even a multi-factor structure in the errors $u_{i,t}$ — with strong or semi-strong factors — can be allowed for. This can be contrasted with the tests mentioned above (and, indeed, all other tests in this literature), where the amount of cross-sectional dependence in the error terms is always restricted. In turn, this ensures that our approach can be employed even when strong and semi-strong pricing factors have been omitted from (1.1), as long as these are uncorrelated with the pricing factors used in the model. Finally, we require error terms to have four finite moments, as opposed to the Gaussianity assumption in Fan et al. (2015); as is typical in the time series literature (see e.g. the book by Davidson, 1994), the assumption poses a restriction between the amount of dependence (quantified via a) and the existence of higher order moments. By the same token, Assumption 2.2 only requires the regressors to have finite 4-th moment, which is substantially milder than the sub-exponential tails required in Fan et al. (2015) and Feng et al. (2022). Finally, Assumption 2.3 is a standard weak exogeneity requirement that is less restrictive than the independence assumption in e.g. Fan et al. (2015), Feng et al. (2022), and Pesaran and Yamagata (2024).

3. The test

Recall (1.4)

$$\mathbb{H}_0 : \max_{1 \leq i \leq N} |\alpha_i| = 0,$$

and consider the OLS estimator of α_i in (1.1)

$$\widehat{\alpha}_{i,T} = \overline{y}_i - \widehat{\beta}'_{i,T} \overline{f}, \tag{3.1}$$

where $\overline{y}_i = T^{-1} \sum_{t=1}^{T} y_{i,t}$, $\overline{f} = T^{-1} \sum_{t=1}^{T} f_t$, and $\widehat{\beta}_{i,T}$ is the OLS estimator of β_i in (1.1), defined as

$$\widehat{\beta}_{i,T} = \left[\sum_{t=1}^{T} \left(f_t - \overline{f} \right) \left(f_t - \overline{f} \right)' \right]^{-1} \left[\sum_{t=1}^{T} \left(f_t - \overline{f} \right) y_{i,t} \right].$$

Recall that the data admit at least $\nu > 4$ moments. We use the transformation

$$\psi_{i,NT} = \left| \frac{T^{1/\nu} \widehat{\alpha}_{i,T}}{\widehat{s}_{NT}} \right|^{\nu/2}, \tag{3.2}$$

where the rescaling sequence is based on the squared OLS residuals $\widehat{u}^2_{i,t}$:

$$\widehat{s}_{NT} = \sqrt{\frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} \widehat{u}^2_{i,t}}. \tag{3.3}$$

Recall that we only work with rates of convergence, requiring $\psi_{i,NT}$ to drift to zero under the null and diverge under the alternative. Thus, in (3.2), any rescaling sequence which removes the measurement unit from $\widehat{\alpha}_{i,T}$ can be used in principle. For example, one may use an estimate of the variance of the individual $\widehat{\alpha}_{i,T}$, and indeed such estimate does not even need to be consistent. We propose $\widehat{s}_{NT}$ because – by way of a guideline – it turns out to deliver the best performance in our simulations.

We are now ready to discuss the construction of the test statistic. Heuristically, under standard assumptions it should hold that $\widehat{\alpha}_{i,T} - \alpha_i = O_{a.s.}\left(T^{-1/2}\right)$. Hence, under the null of (1.4), it should hold by construction that $\psi_{i,NT} \stackrel{a.s.}{\to} 0$; conversely, under the alternative it should hold that $\psi_{i,NT} \stackrel{a.s.}{\to} \infty$. In order to have a statistic to test for $\mathbb{H}_0$, we now perturb the $\psi_{i,NT}$ by adding a sequence of *i.i.d.* Gaussian variables as follows

$$z_{i,NT} = \psi_{i,NT} + \omega_i,$$

where $\omega_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0,1)$ are generated independently of the sample $\left\{(u_{i,t}, f_t')', 1 \le i \le N, 1 \le t \le T\right\}$. Heuristically, under the null $z_{i,NT}$ should be an *i.i.d.* sequence of standard normals; under the alternative, there should be (at least) one spike due to the fact that, for some i, $\alpha_i \neq 0$. Hence, we base our test on the maximally selected $z_{i,NT}$:

$$Z_{N,T} = \max_{1 \le i \le N} z_{i,NT}.$$

In order to study the asymptotics of $Z_{N,T}$, let

$$b_N = \sqrt{2\log N} - \frac{\log\log N + \ln(4\pi)}{2\sqrt{2\log N}}, \quad \text{and} \quad a_N = \frac{b_N}{1 + b_N^2},$$

and consider the following restriction:

Assumption 3.1. *It holds that* $N = O\left(T^{\frac{1}{2}\left(\frac{\nu}{2}-1\right)-\varepsilon}\right)$ *for some* $\varepsilon > 0$.

Assumption 3.1 poses a constraint on the relative rate of divergence of N and T as they pass to infinity; as can be seen, the more moments the data admit, the larger N can be relative to T. A comparison with a similar requirement in Assumption A1(iii) in Feng et al. (2022) may shed light on this assumption. To begin with, as mentioned above, Assumption 2.3 entails that, in our proofs, we need to expand the moment conditions (in essence, due to the fact that we use Hölder's inequality); if, similarly to Feng et al. (2022), we assumed independence between f_t and $u_{i,t}$, then Assumption 3.1 would become $N = O\left(T^{\nu/2-1-\epsilon}\right)$

for some (arbitrarily small) $\varepsilon > 0$, which coincides with Assumption A1(iii) in Feng et al. (2022). Similarly, the asymptotics in Pesaran and Yamagata (2024) requires $N = o\left(T^2\right)$, under the assumptions of deterministic regressors and at least eight finite moments for the errors. In our case, as long as $\nu \geq 6$, the condition that $N = o\left(T^2\right)$ is satisfied, and therefore we have either the same asymptotic regime with a milder moment condition, or, with the same moment condition, a larger N relative to T. In Section B.1 in the Supplement, we discuss the possibility of relaxing – given ν – Assumption 3.1 to allow for a broader set of combinations of N and T.

Let $\mathbb{P}^*$ denote the probability conditional on the sample, viz. $\left\{\left(u_{i,t}, f_t'\right)', 1 \leq i \leq N, 1 \leq t \leq T\right\}$.

Theorem 3.1. *We assume that Assumptions 2.1-3.1 are satisfied. Then, under $\mathbb{H}_0$ of (1.4), it holds that*

$$\lim_{\min\{N,T\}\to\infty} \mathbb{P}^*\left(\frac{Z_{N,T} - b_N}{a_N} \leq x\right) = \exp\left(-\exp\left(-x\right)\right), \tag{3.4}$$

for almost all realizations of $\left\{\left(u_{i,t}, f_t'\right)', 1 \leq i \leq N, 1 \leq t \leq T\right\}$, and all $-\infty < x < \infty$. Under $\mathbb{H}_A$ of (1.3), it holds that

$$\lim_{\min\{N,T\}\to\infty} \mathbb{P}^*\left(\frac{Z_{N,T} - b_N}{a_N} \leq x\right) = 0, \tag{3.5}$$

for almost all realizations of $\left\{\left(u_{i,t}, f_t'\right)', 1 \leq i \leq N, 1 \leq t \leq T\right\}$, and all $-\infty < x < \infty$.

Theorem 3.1 describes the limiting behavior of the test statistic $Z_{N,T}$ both under the null and under the alternative hypotheses. By (3.4), (the suitably normed version of) $Z_{N,T}$ converges (in distribution, a.s. conditionally on the sample) to a Gumbel distribution. Equation (3.4) implies that asymptotic critical values at nominal level τ are given by

$$c_\tau = b_N - a_N \log\left(-\log\left(1 - \tau\right)\right). \tag{3.6}$$

Similarly, (3.5) roughly states that under the alternative (the suitably normed version of) $Z_{N,T}$ diverges to positive infinity in probability, a.s. conditional on the sample.

3.1. **De-randomized inference.** The results in Theorem 3.1 are different from "standard" inferential theory. In particular, (3.5) entails that

$$\lim_{\min\{N,T\}\to\infty} \mathbb{P}^*\left(Z_{N,T} \geq c_\tau | \mathbb{H}_A\right) = 1,$$

which corresponds to the usual notion of power. On the other hand, the result under the null is more delicate. Whilst it is true that, when using the test, it holds that

$$\lim_{\min\{N,T\}\to\infty} \mathbb{P}^*\left(Z_{N,T} \geq c_\tau | \mathbb{H}_0\right) = \tau,$$

this result is different from the standard notion of size. Indeed, $Z_{N,T}$ is constructed by using the added randomness, $\{\omega_i, 1 \leq i \leq N\}$; as can be seen by inspecting our proofs, the effect of said randomness does not vanish asymptotically. In turn, this entails that, under the null, different researchers using the same data will obtain different values of $Z_{N,T}$, and, consequently, different p-values; indeed, if an infinite number of researchers were to carry out the test, the p-values would follow a uniform distribution on $[0,1]$. This is a well-known feature of randomized tests in general (see e.g. Corradi and Swanson, 2006).
Hence, we propose a decision rule to discern between $\mathbb{H}_0$ and $\mathbb{H}_A$ which is not driven by the added randomness, and is therefore the same across all researchers using the same dataset. Following Horváth and Trapani (2019), each researcher will compute $Z_{N,T}$ over B replications, at each replication $1 \leq b \leq B$ constructing a statistic $Z_{N,T}^{(b)}$ using a random sequence $\omega_i^{(b)} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ for $1 \leq i \leq N$, independent across $1 \leq b \leq B$ and of the sample. Finally, for a nominal level τ, we define the function

$$Q_{N,T,B}(\tau) = B^{-1} \sum_{b=1}^{B} \mathbb{I}\left(Z_{N,T}^{(b)} \leq c_\tau\right), \tag{3.7}$$

i.e. the percentage of times that the researcher does not reject the null at significance level τ. An immediate consequence of Theorem 3.1 is that

$$\lim_{\min\{N,T,B\}\to\infty} \mathbb{P}^* \left(Q_{N,T,B}(\tau) = 1-\tau\right) = 1 \quad \text{under } \mathbb{H}_0, \tag{3.8}$$

$$\lim_{\min\{N,T,B\}\to\infty} \mathbb{P}^* \left(Q_{N,T,B}(\tau) = 0\right) = 1 \quad \text{under } \mathbb{H}_A, \tag{3.9}$$

for almost all realizations of $\left\{(u_{i,t}, f_t')', 1 \leq i \leq N, 1 \leq t \leq T\right\}$. This result holds for all different researchers, and therefore averaging across the replications $1 \leq b \leq B$ washes out the added randomness in $Q_{N,T,B}(\tau)$: as a consequence, all researchers will obtain the same value of $Q_{N,T,B}(\tau)$. As noted in He et al. (2023), $Q_{N,T,B}(\tau)$ corresponds to (the complement to one of) the "fuzzy decision", or "abstract randomized decision rule" reported in equation (1.1a) in Geyer and Meeden (2005). This notion can be illustrated by considering a random variable, say $\mathcal{D}$, which takes two values: "do not reject $\mathbb{H}_0$" with probability $Q_{N,T,B}(\tau)$, and "reject $\mathbb{H}_0$". According to (3.8), asymptotically it holds that, a.s. conditionally on the sample $\mathbb{P}^*(\omega : \mathcal{D} = \text{"reject } \mathbb{H}_0\text{"}|\mathbb{H}_0) = \tau$, across all researchers, which reconciles the procedure with the notion of *size* of a test. Similarly, (3.9) states that, asymptotically, $\mathbb{P}^*(\omega : \mathcal{D} = \text{"reject } \mathbb{H}_0\text{"}|\mathbb{H}_A) = 1$ a.s. conditionally on the sample, which corresponds to the notion of *power* of a test.

Theorem 3.2. *We assume that Assumptions 2.1-3.1 are satisfied, and that $B = O\left(\log^2 N\right)$. Then it holds that*

$$\left|\frac{Q_{N,T,B}(\tau) - (1-\tau)}{\sqrt{\alpha(1-\tau)}}\right| = O_{a.s.}\left(\sqrt{\frac{2\log\log B}{B}}\right), \tag{3.10}$$

under H_0, for almost all realizations of $\left\{(u_{i,t}, f_t')', 1 \leq i \leq N, 1 \leq t \leq T\right\}$. Under H_A, it holds that $Q_{N,T,B}(\tau) = o_{a.s.}(1)$, for almost all realizations of $\left\{(u_{i,t}, f_t')', 1 \leq i \leq N, 1 \leq t \leq T\right\}$.

Building on Theorem 3.2, a "de-randomized" decision rule can be proposed. By (3.10), under $\mathbb{H}_0$ there exists a triplet of random variables (N_0, T_0, B_0) such that

$$Q_{N,T,B}(\tau) \geq 1 - \tau - \sqrt{\tau(1-\tau)}\sqrt{\frac{2\log\log B}{B}}, \tag{3.11}$$

for all (N, T, B) with $N \geq N_0$, $T \geq T_0$, and $B \geq B_0$. Similarly, under $\mathbb{H}_A$ there exists a triplet of random variables (N_0, T_0, B_0) such that $Q_{N,T,B}(\tau) \leq \epsilon$, for all $\epsilon > 0$ and (N, T, B) with $N \geq N_0$, $T \geq T_0$, and $B \geq B_0$. This dichotomous behavior can be exploited to construct a decision rule based on $Q_{N,T,B}(\tau)$: $\mathbb{H}_0$ is not rejected when $Q_{N,T,B}(\tau)$ exceeds a threshold, whereas it is rejected otherwise. In theory, one could use the threshold defined in (3.10), but this, albeit valid asymptotically, is likely to be overly conservative in finite samples. A less conservative decision rule in favor of the null could be

$$Q_{N,T,B}(\tau) \geq (1-\tau) - f(B), \tag{3.12}$$

with $f(B)$ a user-specified, non-increasing function of B such that

$$\lim_{B\to\infty} f(B) = 0 \text{ and } \limsup_{B\to\infty} (f(B))^{-1}\sqrt{\frac{2\log\log B}{B}} = 0. \tag{3.13}$$

3.2. **From theory to practice: guidelines and recommendations.** The procedure proposed in Section 3 depends on some nuisance parameters (chiefly, on how many moments ν the data admit), and on some tuning quantities (chiefly, the number of replications B and the function $f(B)$ in the de-randomized approach). In this section, we offer a set of guidelines/suggestions which could inform the practical application of these procedures.

We begin by discussing the choice of ν. There are at least two ways in which ν can be estimated:

(1) A *direct* approach, based on using a tail index estimator. This approach, based on the plug-in principle offers an estimate of ν which, under standard assumptions, is

consistent; however, the properties of tail index estimators may be rather poor in finite samples (Embrechts et al., 2013a).

(2) An *indirect* approach, based upon noting that – in order to construct $\psi_{i,NT}$ – a lower bound (as opposed to an exact value) for ν would suffice. In order to find such a bound, e.g. the tests by Trapani (2016) and Degiannakis et al. (2023) could be employed to test for the null hypothesis $\mathbb{H}_0 : \mathbb{E}\left|y_{i,t}\right|^{\nu_0} = \infty$. Upon rejecting the null, it follows that $\nu \geq \nu_0$, and therefore ν_0 can be used in (3.2).[4]

In addition — as we do in our empirical illustration — one can set $\nu = 4$, i.e. the lowest possible value allowed by our asymptotic theory, which we would recommend when the sample size T does not afford reliable inference.

Turning to the specifics of the de-randomization, we note that:

(1) The choice of B is constrained by the condition $B = O\left(\log^2 N\right)$. In practice, the function $Q_{N,T,B}\left(\tau\right)$ is computed by averaging an *i.i.d.* sequence of uniformly distributed random variables, and therefore it can be expected that asymptotics approximations will be quite accurate even for (relatively) small B. In our simulations, we employ $B = \log^2 N$, which we recommend as a guideline.

(2) The choice of $f\left(B\right)$ is based on (3.13), and the simulations in He et al. (2023) show that the de-randomized decision rule is relatively robust to the specification of $f\left(B\right)$. In our simulations, we suggest $f\left(B\right) = B^{1/4}$, which is also found to deliver the best results in He et al. (2023).

[4] In principle, even a test for a specific distribution (such as a test for Gaussianity) could be employed in this approach.

4. Extensions: non-tradable and latent factors

As mentioned in the introduction, the main contribution of this paper is a methodology to test for no pricing errors; we have focussed on (1.1) and assumed factors are observable and tradable only for simplicity. In this section, we show that our methods can be readily extended to more complex settings, provided that a consistent estimator of the α_is is available. As illustrative examples, we consider the case of non-tradable factors based on Fama-MacBeth estimation, and the case of latent factors, based on principal component analysis (PCA). These extensions are not considered, to the best of our knowledge, in any contribution in the current literature; conversely, our methodology can readily accommodate for them, in essence obtaining the same results as in the case of observable and tradable factors. We only report the main results on the "one shot" tests; assumptions and technicalities are relegated to Section B.2 in the Supplement, and the extension to de-randomization can be done by following *verbatim* Section 3.1. Henceforth, we define $\mathbb{M}_{1_N} = \mathbb{I}_N - N^{-1}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'$, where $\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones.

4.1. Non-tradable factors and Fama-MacBeth estimation.

Consider the case of a linear factor pricing model based on K observable, *non-tradable* factors

$$y_{i,t} = \alpha_i + \beta_i'\lambda + \beta_i'v_t + u_{i,t}, \tag{4.1}$$

where $v_t = f_t - \mathbb{E}(f_t)$ and $\lambda \in \mathbb{R}^K$ is the vector of risk premia for the K factors f_t. Estimation of α_i is based on Algorithm 3 in Giglio et al. (2021).

Step 1: Estimate β_i via OLS from the time-series regressions $y_{i,t} = \alpha_i + \beta_i f_t + u_{i,t}$,

$$\widehat{\beta}_i = \left[\sum_{t=1}^{T}\left(f_t - \overline{f}\right)\left(f_t - \overline{f}\right)'\right]^{-1}\left[\sum_{t=1}^{T}\left(f_t - \overline{f}\right)\left(y_{i,t} - \overline{y}_i\right)\right], \tag{4.2}$$

with $\overline{y}_i = T^{-1}\sum_{t=1}^{T} y_{i,t}$, and define $\widehat{\boldsymbol{\beta}} = \left(\widehat{\beta}_1, ..., \widehat{\beta}_N\right)'$.

Step 2: Define $\widehat{\lambda}$ as the OLS estimates of a regression of $\overline{\mathbf{y}} = (\overline{y}_1, ..., \overline{y}_N)'$ onto a vector of ones and $\widehat{\boldsymbol{\beta}}$

$$\widehat{\lambda} = \left(\widehat{\boldsymbol{\beta}}' \mathbb{M}_{1_N} \widehat{\boldsymbol{\beta}}\right)^{-1} \left(\widehat{\boldsymbol{\beta}}' \mathbb{M}_{1_N} \bar{\mathbf{y}}\right). \tag{4.3}$$

Step 3: The estimator of α_i is given by

$$\widehat{\alpha}_i^{FM} = \overline{y}_i - \widehat{\beta}_i' \widehat{\lambda}. \tag{4.4}$$

Based on $\widehat{\alpha}_i^{FM}$ defined in (4.4), we can construct the same test statistic as before, based on

$$\psi_{i,NT}^{FM} = \left| \frac{T^{1/\nu} \widehat{\alpha}_i^{FM}}{\widehat{s}_{NT}^{FM}} \right|^{\nu/2},$$

where the rescaling sequence $\widehat{s}_{NT}^{FM}$ is constructed as in (3.3), using the residuals $\widehat{u}_{i,t}^{FM} = y_{i,t} - \left(\widehat{\alpha}_i^{FM} + \widehat{\beta}_i' f_t\right)$. Defining $z_{i,NT}^{FM} = \psi_{i,NT}^{FM} + \omega_i$,[5] our test can be based on

$$Z_{N,T}^{FM} = \max_{1 \le i \le N} z_{i,NT}^{FM}. \tag{4.5}$$

Theorem 4.1. *We assume that the assumptions of Theorem 3.1 are satisfied, and that Assumptions B.1 and B.2 in Section B.2 of the Supplement also hold. Then, the same result as in Theorem 3.1 holds.*

4.2. **Latent factors.** Consider a linear factor pricing model based on K latent factors

$$y_{i,t} = \alpha_i + \beta_i' \lambda + \beta_i' v_t + u_{i,t}, \tag{4.6}$$

where $v_t = f_t - \mathbb{E}(f_t)$ is *not observable*, β_i is a $K \times 1$ vector of loadings and, as above, λ is the vector of risk premia for the K latent factors f_t.

[5] As before, $\omega_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ generated independently of the sample $\left\{(u_{i,t}, f_t')', 1 \le i \le N, 1 \le t \le T\right\}$.

Write $\widetilde{\mathbf{y}}_t = \boldsymbol{\beta}\widetilde{v}_t + \widetilde{\mathbf{u}}_t$, where $\widetilde{\mathbf{y}}_t = \mathbf{y}_t - \overline{\mathbf{y}}$ with $\mathbf{y}_t = (y_{1,t}, ..., y_{N,t})'$, $\widetilde{v}_t = v_t - \left(T^{-1}\sum_{t=1}^{T} v_t\right)$, $\widetilde{\mathbf{u}}_t$ is defined analogously, and $\boldsymbol{\beta} = (\beta_1, ..., \beta_N)'$. Define also the $N \times N$ sample second moment matrix

$$\widehat{\boldsymbol{\Sigma}}_y = \frac{1}{NT}\sum_{t=1}^{T} \widetilde{\mathbf{y}}_t \widetilde{\mathbf{y}}_t'. \tag{4.7}$$

The estimation of α_i follows Algorithm 4 in Giglio et al. (2021).

Step 1: Estimate $\boldsymbol{\beta}$ using PCA, with estimator $\widehat{\boldsymbol{\beta}}^{PC}$ given by the eigenvectors corresponding to the first K eigenvalues of $\widehat{\boldsymbol{\Sigma}}_y$ under the constraint $\left(\widehat{\boldsymbol{\beta}}^{PC}\right)' \widehat{\boldsymbol{\beta}}^{PC} = N\mathbb{I}_K$.[6]

Steps 2 and 3: These steps are the same as in the previous section, with $\widehat{\lambda}^{PC} = \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\bar{\mathbf{y}}\right)$, and

$$\widehat{\alpha}_i^{PC} = \overline{y}_i - \left(\widehat{\beta}_i^{PC}\right)' \widehat{\lambda}^{PC}. \tag{4.8}$$

Let $C_{N,T} = \min\{N, T\}$. Based on $\widehat{\alpha}_i^{PC}$ defined in (4.8), we define

$$\psi_{i,NT}^{PC} = \left|\frac{C_{N,T}^{1/\nu}\widehat{\alpha}_i^{PC}}{\widehat{s}_{NT}^{PC}}\right|^{\nu/2},$$

where $\widehat{s}_{NT}^{PC}$ is constructed as in (3.3), using $\widehat{u}_{i,t}^{PC} = y_{i,t} - \left(\widehat{\alpha}_i^{PC} + \widehat{\beta}_i^{PC\prime}\widehat{f}_t^{PC}\right)$ and $\widehat{f}_t^{PC} = N^{-1}\widehat{\boldsymbol{\beta}}^{PC\prime}\widetilde{\mathbf{y}}_t$. Letting $z_{i,NT}^{PC} = \psi_{i,NT}^{PC} + \omega_i$, with ω_i defined as above, the test is based on

$$Z_{N,T}^{PC} = \max_{1\le i\le N} z_{i,NT}^{PC}. \tag{4.9}$$

Theorem 4.2. *We assume that the assumptions of Theorem 3.1 are satisfied, and that Assumptions B.3-B.6 in Section B.2 of the Supplement also hold. Then, the same result as in Theorem 3.1 holds.*

[6]Our discussion implicitly assumes that K is known. Of course, this is not the case in practice, where K has to be determined by the user. This is ordinarily done using consistent estimators such as those of Bai and Ng (2002), Ahn and Horenstein (2013), and Trapani (2018). Note that the result in Theorem 4.2 holds unchanged when K is estimated using these consistent estimators.

As a final remark, we conjecture that Theorem 4.2 holds with minor modifications to Assumptions B.3-B.6 when one estimates factors and loadings with the Risk-Premium PCA (RP-PCA) approach of Lettau and Pelger (2020a,b). In fact, just like our Theorem 4.2, their theory for the strong factors case relies on conditions that are extremely similar to those of Bai (2003). Similar considerations also hold for the Projected PCA approach of Fan et al. (2016), which is increasingly being used in asset pricing studies (see Kim et al., 2021 and Hong et al., 2025, among others).

5. Simulations

We study the finite sample properties of our test based on (1.1). We use a similar DGP to Feng et al. (2022):

$$y_{i,t} = \alpha_i + \sum_{p=1}^{3} \beta_{i,p} f_{p,t} + u_{i,t}, \tag{5.1}$$

$$f_t = \overline{f} + \Phi f_{t-1} + \zeta_t, \tag{5.2}$$

with $f_t = (f_{1,t}, f_{2,t}, f_{3,t})'$, $\overline{f} = (0.53, 0.19, 0.19)'$, $\Phi = \text{diag}\{-0.1, 0.2, -0.2\}$ and $\zeta_t \overset{i.i.d.}{\sim} \mathcal{N}_3(0, I_3)$. Loadings are generated as $\beta_{i,1} \overset{i.i.d.}{\sim} \mathcal{U}(0.3, 1.8)$, $\beta_{i,2} \overset{i.i.d.}{\sim} \mathcal{U}(-1, 1)$, and $\beta_{i,3} \overset{i.i.d.}{\sim} \mathcal{U}(-0.6, 0.9)$ for all i. We allow for strong cross-sectional dependence in the innovations $\mathbf{u}_t = (u_{1,t}, \ldots, u_{N,t})'$ via a factor model:

$$\mathbf{u}_t = \boldsymbol{\gamma} g_t + \boldsymbol{\xi}_t, \quad \text{with} \quad g_t = \phi_g g_{t-1} + \chi_t, \tag{5.3}$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)'$ for $\gamma_i \overset{i.i.d.}{\sim} \mathcal{U}(0.7, 0.9)$, $\phi_g = 0.4$, and $\chi_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, with $\{\chi_t, 1 \leq t \leq T\}$ generated independently of $\{\zeta_t, 1 \leq t \leq T\}$. Note that, since $\gamma_i > 0$ for all i, it follows that the omitted factor g_t is a strong factor. In (5.3), the N-dimensional random vectors $\{\boldsymbol{\xi}_t, 1 \leq t \leq T\}$ are generated independently of $\{\zeta_t, 1 \leq t \leq T\}$ and $\{\chi_t, 1 \leq t \leq T\}$

using various distributions with mean zero and covariance matrix Σ_ξ. We consider the following three set-ups:

(1) **The Gaussian case**: $\boldsymbol{\xi}_t \overset{i.i.d.}{\sim} \mathcal{N}_N(0, I_N)$.

(2) **The Student's t case**: where $\xi_{i,t}$ follows a Students's t distribution with $d = 5.5$ degrees of freedom, zero mean and unit scale, independent across i. In this case, $\xi_{i,t}$ and $y_{i,t}$ have regularly varying tails;

(3) **The GARCH case**: we generate $\boldsymbol{\xi}_t = \mathbf{H}_t \boldsymbol{z}_t$, with: $\boldsymbol{z}_t = (z_{1,t}, ..., z_{N,t})'$ and $z_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$; and $\mathbf{H}_t = \text{diag}\{h_{1,t}, \ldots, h_{N,t}\}$ with $h_{i,t}^2 = \omega_i + \alpha_i \xi_{i,t}^2 + \beta_i h_{i,t-1}^2$, with $\omega_i \overset{i.i.d.}{\sim} \mathcal{U}(0.01, 0.05)$, $\alpha_i \overset{i.i.d.}{\sim} \mathcal{U}(0.01, 0.04)$ and $\beta_i \overset{i.i.d.}{\sim} \mathcal{U}(0.85, 0.95)$.[7]

In all scenarios, we report empirical rejection frequencies under the null and under the alternative, for nominal level $\tau = 5\%$, using $N \in \{100, 200, 500\}$ and $T \in \{100, 200, 300, 400, 500\}$. As far as the alternative hypothesis is concerned, we consider a rather sparse alternative where $\alpha_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ for 5% percent of the cross-sectional units $1 \le i \le N$. We construct the test statistic using a notional value of $\nu = 5$ in (3.2), and consider both the "one-shot" test in Section 3, and the de-randomized version in Section 3.1. For the latter, we examine results using two decision rules: the first one is based on the inequality in (3.11) while the second one employs the relation in (3.12) with $f(B) = B^{-1/4}$. We compare our test with the tests by Fan et al. (2015, FLY henceforth), Gagliardini et al. (2016, GOS), Feng et al. (2022, FLLM), Pesaran and Yamagata (2024, PY), and Ardia and Sessinou (2024, AS).[8,9]

[7]These parameter values imply that $\xi_{i,t}$ has finite sixth moment for any i.

[8]The test of FLY is based on estimating the covariance matrix of u_t with the POET method of Fan et al. (2013). As in Fan et al. (2015), we consider the *soft* thresholding function; results based on hard and scad thresholding (Fan and Li, 2001) are numerically identical and available upon request. Finally, we set tuning parameters of the AS test as suggested in their Monte Carlo exercise (in particular, and using their notation, we set $L = 0$ and $\psi = 1/3$).

[9]Also, Monte Carlo results in Feng et al. (2022) and Pesaran and Yamagata (2024) show that their tests consistently outperform that of Gungor and Luger (2016). Hence, we omit comparisons with this last approach.

We start from the Gaussian case; results using the one-shot test are in Table 5.1, whereas in Table 5.2 we report rejection frequencies from the de-randomization approach. Our test is the only one that controls the size for all combinations of N and T, whereas the other tests are consistently oversized. Table 5.2 suggest that the de-randomization procedure based on $f(B) = B^{-1/4}$ also works very well. Indeed, empirical rejection frequencies are extremely close to zero under the null and quickly converge to one under the alternative; the latter is particularly true when N gets large, thus showing that our approach is particularly suitable for high-dimensional settings. As expected, the threshold based on the LIL leads to higher empirical rejection frequencies under both the null and the alternative. As far as power is concerned, the right panels of the table show that our tests performs satisfactorily, whilst at the same time guaranteeing size control. We note that the test by FLLM outperforms ours in terms of power in most cases, but it is also oversized. Turning to the case of data with heavier tails, Tables 5.3 and 5.4 report empirical rejection frequencies for the Student's t case. Results for the randomized test are in line with those of Table 5.1. As in the Gaussian case, the other tests are oversized. The de-randomized procedure works as expected also in the case of heavier tails. Similar considerations hold for power as for the Gaussian case. Finally, results under the GARCH case are in Tables 5.5 and 5.6; size and power of all tests behave as in the other cases, and so does the decision rule based on Theorem 3.1.

In a further set of experiments we consider power under twelve different alternatives, each based on a different percentage of mis-priced assets: 1%, 2%, 3%, ..., 9%, 10%, 15% and 20%. For brevity, we only focus on sample sizes $T = 100$ and $N = 500$, which are the most relevant for the empirical analysis of Section 6. Figure 5.1 reports the empirical rejection frequencies for the twelve levels of sparsity (the horizontal axis reports the percentage of mis-priced assets) across all DGPs. Tests by FLY and FLLM always achieve unit power, while our approach performs almost equally well (and better than the other tests), as its power converges to one almost immediately.

TABLE 5.1. Empirical rejection frequencies for the test in Theorem 3.1, Gaussian case.

		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0.4$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.058	0.042	0.040	0.040	0.039	0.935	0.960	0.969	0.972	0.975
	FLLM	0.124	0.112	0.101	0.088	0.095	0.986	0.995	0.996	0.998	0.998
	FLY	0.353	0.218	0.150	0.129	0.138	0.996	0.997	0.998	1.000	1.000
	GOS	0.212	0.200	0.179	0.185	0.193	0.792	0.945	0.970	0.979	0.988
	PY	0.215	0.195	0.178	0.183	0.191	0.785	0.941	0.970	0.979	0.988
	AS	0.104	0.075	0.066	0.077	0.089	0.222	0.547	0.852	0.959	0.976
200	Thm. 1	0.071	0.039	0.036	0.036	0.035	0.997	0.999	1.000	1.000	1.000
	FLLM	0.151	0.087	0.099	0.102	0.104	1.000	1.000	1.000	1.000	1.000
	FLY	0.386	0.181	0.153	0.125	0.117	1.000	1.000	1.000	1.000	1.000
	GOS	0.214	0.176	0.183	0.186	0.217	0.863	0.987	0.995	0.999	1.000
	PY	0.219	0.174	0.182	0.185	0.217	0.865	0.986	0.995	0.999	1.000
	AS	0.113	0.083	0.085	0.082	0.078	0.225	0.657	0.944	0.995	0.999
500	Thm. 1	0.070	0.046	0.040	0.038	0.038	1.000	1.000	1.000	1.000	1.000
	FLLM	0.121	0.134	0.120	0.109	0.089	1.000	1.000	1.000	1.000	1.000
	FLY	0.458	0.268	0.190	0.141	0.125	1.000	1.000	1.000	1.000	1.000
	GOS	0.181	0.222	0.224	0.213	0.192	0.936	1.000	1.000	1.000	1.000
	PY	0.192	0.220	0.221	0.212	0.191	0.946	1.000	1.000	1.000	1.000
	AS	0.100	0.106	0.112	0.099	0.102	0.248	0.768	0.979	0.999	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$.

TABLE 5.2. Empirical rejection frequencies for the decision rule of Theorem 3.1, Gaussian case.

N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.067	0.011	0.001	0.000	0.000	0.962	0.976	0.982	0.984	0.989
	$f(B) = B^{-1/4}$	0.004	0.000	0.000	0.000	0.000	0.930	0.958	0.971	0.972	0.975
200	LIL	0.059	0.006	0.000	0.000	0.000	0.999	0.999	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.010	0.000	0.000	0.000	0.000	0.996	0.997	0.999	1.000	1.000
500	LIL	0.381	0.224	0.162	0.104	0.077	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.012	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}$.

In Section A in the Supplement we report further simulations. In Section A.1, we report results for the case $\phi_g = 0$; in Section A.2, we investigate the cases where g_t is either a weak or a semi-strong omitted factor; in Section A.3, we investigate the finite sample properties of our test when only one factor is strong, while the others are at most semi-strong; in

Table 5.3. Empirical rejection frequencies for the test in Theorem 3.1, Student's t case.

		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0.4$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.058	0.046	0.039	0.037	0.038	0.888	0.930	0.947	0.956	0.963
	FLLM	0.125	0.094	0.102	0.103	0.100	0.977	0.992	0.999	1.000	1.000
	FLY	0.337	0.175	0.139	0.118	0.113	0.996	0.999	1.000	1.000	1.000
	GOS	0.000	0.000	0.000	0.000	0.000	0.794	0.934	0.973	0.983	0.991
	PY	0.242	0.196	0.202	0.198	0.208	0.826	0.934	0.973	0.983	0.991
	AS	0.093	0.079	0.088	0.079	0.088	0.170	0.474	0.798	0.934	0.965
200	Thm. 1	0.056	0.040	0.037	0.037	0.035	0.982	0.993	0.997	0.997	0.998
	FLLM	0.104	0.114	0.108	0.109	0.110	0.999	1.000	1.000	1.000	1.000
	FLY	0.379	0.228	0.156	0.139	0.119	1.000	1.000	1.000	1.000	1.000
	GOS	0.000	0.000	0.000	0.202	0.219	0.858	0.980	0.993	0.998	1.000
	PY	0.240	0.217	0.197	0.201	0.217	0.891	0.979	0.993	0.998	1.000
	AS	0.102	0.077	0.082	0.101	0.096	0.192	0.553	0.904	0.989	0.993
500	Thm. 1	0.068	0.039	0.038	0.040	0.039	1.000	1.000	1.000	1.000	1.000
	FLLM	0.135	0.103	0.123	0.114	0.111	1.000	1.000	1.000	1.000	1.000
	FLY	0.488	0.232	0.177	0.138	0.132	1.000	1.000	1.000	1.000	1.000
	GOS	0.271	0.211	0.203	0.206	0.211	0.942	0.998	1.000	1.000	1.000
	PY	0.270	0.209	0.201	0.204	0.207	0.969	0.999	1.000	1.000	1.000
	AS	0.097	0.085	0.099	0.095	0.108	0.223	0.637	0.966	0.996	0.999

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$.

Table 5.4. Empirical rejection frequencies for the decision rule of Theorem 3.1, Student's t case.

		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.056	0.007	0.001	0.000	0.000	0.945	0.966	0.974	0.981	0.983
	$f(B) = B^{-1/4}$	0.003	0.001	0.000	0.000	0.000	0.885	0.927	0.946	0.957	0.971
200	LIL	0.040	0.001	0.000	0.000	0.000	0.996	0.997	0.999	1.000	1.000
	$f(B) = B^{-1/4}$	0.004	0.000	0.000	0.000	0.000	0.985	0.994	0.995	0.996	0.998
500	LIL	0.433	0.230	0.140	0.108	0.078	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.013	0.001	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}$.

Sections A.4 and A.5 we report a set of experiments for the cases of non-tradable factors (studied in Section 4.1) and latent factors (studied in Section 4.2) respectively.

TABLE 5.5. Empirical rejection frequencies for the test in Theorem 3.1, GARCH case.

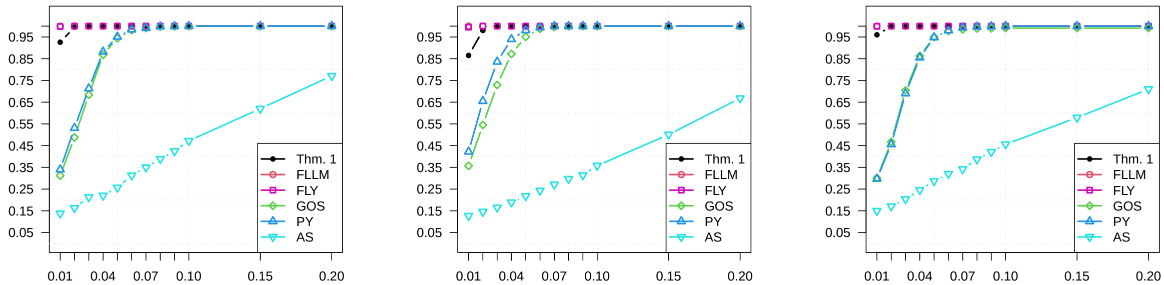
		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0.4$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.079	0.050	0.044	0.042	0.040	0.965	0.975	0.98	0.987	0.987
	FLLM	0.124	0.112	0.101	0.088	0.095	0.986	0.995	0.996	0.998	0.998
	FLY	0.313	0.174	0.154	0.135	0.123	1.000	1.000	1.000	1.000	1.000
	GOS	0.212	0.200	0.179	0.185	0.193	0.792	0.945	0.970	0.979	0.988
	PY	0.215	0.195	0.178	0.183	0.191	0.785	0.941	0.970	0.979	0.988
	AS	0.104	0.075	0.066	0.077	0.089	0.222	0.547	0.852	0.959	0.976
200	Thm. 1	0.070	0.042	0.036	0.033	0.034	1.000	1.000	1.00	1.000	1.000
	FLLM	0.151	0.087	0.099	0.102	0.104	1.000	1.000	1.000	1.000	1.000
	FLY	0.315	0.212	0.144	0.136	0.120	1.000	1.000	1.000	1.000	1.000
	GOS	0.214	0.176	0.183	0.186	0.217	0.863	0.987	0.995	0.999	1.000
	PY	0.219	0.174	0.182	0.185	0.217	0.865	0.986	0.995	0.999	1.000
	AS	0.113	0.083	0.085	0.082	0.078	0.225	0.657	0.944	0.995	0.999
500	Thm. 1	0.109	0.053	0.043	0.038	0.039	1.000	1.000	1.00	1.000	1.000
	FLLM	0.121	0.134	0.120	0.109	0.089	1.000	1.000	1.000	1.000	1.000
	FLY	0.471	0.247	0.165	0.132	0.103	1.000	1.000	1.000	1.000	1.000
	GOS	0.181	0.222	0.224	0.213	0.192	0.936	1.000	1.000	1.000	1.000
	PY	0.192	0.220	0.221	0.212	0.191	0.946	1.000	1.000	1.000	1.000
	AS	0.100	0.106	0.112	0.099	0.102	0.248	0.768	0.979	0.999	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$; $\alpha_i \stackrel{i.i.d.}{\sim} N(0,1)$ for 5% of the cross-sectional units, under the alternative.

TABLE 5.6. Empirical rejection frequencies for the decision rule of Theorem 3.1, GARCH case.

		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.085	0.013	0.003	0.000	0.001	0.983	0.986	0.994	0.996	0.995
	$f(B) = B^{-1/4}$	0.010	0.002	0.000	0.000	0.000	0.956	0.975	0.979	0.984	0.988
200	LIL	0.093	0.016	0.005	0.002	0.001	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.019	0.000	0.000	0.000	0.000	0.999	1.000	1.000	1.000	1.000
500	LIL	0.474	0.276	0.178	0.137	0.094	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.059	0.005	0.001	0.000	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}$; $\alpha_i \stackrel{i.i.d.}{\sim} N(0,1)$ for 5% of the cross-sectional units, under the alternative.



(a) Gaussian DGP, $\phi_g = 0.40$. (b) Student's t DGP, $\phi_g = 0.40$. (c) GARCH(1,1) DGP, $\phi_g = 0.40$.

Figure 5.1. Power curves. The horizontal axis reports the percentages of mis-priced assets.

6. Empirical illustration

We illustrate our procedure by testing whether several linear factor pricing models correctly price the constituents of the S&P 500 index. Pricing individual stocks is a challenging task as their returns are known to have non-normal distributions, display substantial heteroskedasticity and correlation structures that make standard asset pricing tests inadequate. Our testing procedure, however, is well suited for this task, given the generality of the assumptions required by our test. Moreover, individual stocks have poorly estimated betas, which can affect statistical power. Blume (1975) shows that estimation error in the betas generates an errors-in-variables problem that biases the betas downward; however, Ang et al. (2020) argue that, although the measurement errors in the betas are larger in individual stocks relative to e.g. portfolio returns, the larger cross-sectional spread in the betas of individual stocks more than offsets this error, leading to a more accurate estimate of the market risk premium.

6.1. **Data, models of interest, and estimation procedure.** We collect monthly data on all stocks that were part of the S&P 500 for at least five years between January 1985 and December 2024, thus ending up with $N = 1019$ stocks. For the i-th security, we compute the return over the t-th month as

$$r_{i,t} = 100\frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} + \frac{DY_{i,t}}{12}, \tag{6.1}$$

for $P_{i,t}$ the end of the month stock price and $DY_{i,t}$ the percent per annum dividend yield.[10] Given this panel of stock returns, we test six linear factor pricing models that are all encompassed by the following regression

$$y_{i,t} = \alpha_i + \beta_{i,1}\text{MKT}_t + \beta_{i,2}\text{SMB}_t + \beta_{i,3}\text{HML}_t + \beta_{i,4}\text{RMW}_t + \beta_{i,5}\text{CMA}_t + \beta_{i,6}\text{MOM}_t + u_{i,t}, \tag{6.2}$$

where MKT_t is the excess return on the market, SMB_t the *size* factor, HML_t the *book-to-market* factor, RMW_t the *profitability* factor, CMA_t the *investment strategy* factor, and MOM_t the *momentum* factor.[11] Both the excess return $y_{i,t}$ and the market factor use the one-month Treasury bill rate as risk-free rate. The first model that we test is the CAPM, which is obtained when only the market factor is considered. We then consider a two-factor (FF2 henceforth) model based on the market and momentum factors, as well as the usual three- and five-factors models of Fama and French (1993, FF3) and Fama and French (2015, FF5). Finally, we augment these last two models with the momentum factor to obtain a four- and a six-factors model that we refer to as FF4 and FF6, respectively. In line with Fan et al. (2015) and Pesaran and Yamagata (2024), we carry out estimation using 5-years rolling windows to assess the specification of the pricing models; the use of a rolling window mitigates the impact of possible time variation in the factor loadings. Hence, we use $T = 60$ (i.e. 5 years) in estimation, while N ranges between 437 and 568.[12]

[10] To ensure that the index accurately represents the US stock market, S&P regularly updates the composition of the index. We account for these changes by revising the set of included assets every month. Security data are sourced from *Datastream*, while we obtain those on the pricing factors (and on the risk-free rate) from the website of Professor Kenneth French.

[11] All these factors are measured as (functions of) returns on stock portfolios, thus being observable and tradable. We refer to the website of Professor French for an exhaustive explanation of their construction.

[12] Within each window, we discard securities with at least one missing observation.

6.2. **Results.** Table 6.1 reports the percentage of windows for which we can reject the null that a pricing model is correctly specified. These empirical rejection frequencies are computed using our de-randomized procedure based on $B = \log^2 N$ repetitions of the test. Each trial is run at the 5% significance level and uses $\nu = 4$ in the calculation of $\psi_{i,NT}$. The choice of ν is dictated by the fact that testing and estimation methods for the number of moments tend to perform poorly when T is as low as 60. Hence, we fix ν to the lowest possible value such that our inference is valid.[13]

TABLE 6.1. Results of the de-randomized procedure for constituents of the S&P 500 index.

	CAPM	FF2	FF3	FF4	FF5	FF6
Full Sample	0.607	0.624	0.710	0.724	0.840	0.840
The Asian financial crisis (Jul-1997 to Dec-1998)	1.000	1.000	1.000	1.000	1.000	1.000
The Dot-com Bubble Burst (Mar-2000 to Oct-2002)	0.875	0.844	1.000	1.000	1.000	1.000
The Global Financial Crisis (Dec-2007 to Jun-2009)	1.000	1.000	1.000	1.000	1.000	1.000
The COVID-19 pandemic (Jan-2020 to May-2021)	0.765	0.588	0.529	0.647	0.647	0.647

Percentage of rolling windows for which the null is rejected. We consider both overall percentages, i.e. for the whole sample between January 1985 and December 2024, and those specific to relevant subperiods. The significance level is always $\tau = 5\%$ and all the windows contain $T = 60$ observations. Calculation of $\psi_{i,NT}$ is based on $\nu = 4$, and the de-randomized procedure uses $B = \log(N)^2$ repetitions of the test and thresholds based on $f(B) = B^{-1/4}$.

The first row of Table 6.1 shows rejection frequencies for the whole sample. According to these results, the CAPM is the least rejected model (i.e. the one that produces the lowest percentage of rejections of the null of zero pricing errors). This seems to indicate that the factor that matters most for pricing individual stocks — among the ones considered here — is simply the market factor, whereas additional ones reduce the pricing power of the model, possibly due to their scant explanatory power, which — adding noise to the estimation procedure — renders the OLS estimates less accurate. In essence, these results

[13]While our inference requires $\nu > 4$ to be valid, it is worth recalling that other tests typically impose stricter moment requirements The only exception is that of Feng et al. (2022), which also requires four finite moments. Hence, no alternative test is available when our inference is invalid due to the tail behavior of the stocks.

indicate that there is little value added in terms of pricing power in the momentum factor and the Fama-French factors beyond the market factor.

The remaining rows of Table 6.1 show rejection frequencies during periods of market turmoil: the Asian financial crisis, the Burst of the Dot-com Bubble, the Global Financial Crisis, and the COVID-19 pandemic.[14] We see that none of these models can price large cap US stocks during the Asian financial crisis and the Global Financial Crisis. The CAPM and the two-factor model are the only ones able to price the assets of interest during part of early 2000s crisis. Finally, we see that the size and book-to-market factors play an important role during the COVID-19 pandemic, as the three-factor models is the best over this period. Notably, this is the only instance where the CAPM is outperformed by larger pricing models.[15] Finally, Figure 6.1 shows values of $Q_{N,T,B}(0.05)$ from equation (3.7) for the six pricing models and across rolling windows. This complements results in Table 6.1, as it allows us to fully understand under which market circumstances models do a better job at pricing the panel of individual stock returns. Grey shaded areas highlight the previous four periods of market distress, as well as the brief recession of the early 1990s. The black dashed line is the threshold based on $f(B) = B^{-1/4}$, and we decide in favor of the null whenever $Q_{N,T,B}(0.05)$ exceeds it. Several results can be noted. First, the performance of all the

[14]For the COVID-19 pandemic, we consider rolling windows between January 2020 and May 2021, where the end date corresponds to the termination of lockdown policies in most of the world. The dates for the other periods of market turmoil are set as in Pesaran and Yamagata (2024).

[15]To interpret this result note that, while the COVID-19 period shares with other crises the fact that the stock market yielded low returns and was characterized by high volatility and low liquidity, it also has features that make it very distinct. Indeed, during the COVID-19 pandemic households were required to stay home to slow the spread of the virus and a variety of firms were severely restricted in producing their goods and services, essentially constraining output production and limiting consumption decisions. This distinctive feature of the COVID-19 period implies that, while uncertainty about the end of the pandemic was very high, in the short term recession fears and low growth expectations strongly characterize that period (Gormsen and Koijen, 2023). It is not surprising, therefore, that value stocks — generally considered long-horizon investments — came under huge pressure as economic uncertainty prompted investors to shorten their time horizons, and indeed we find that the value factor is the key driver of the slightly better performance of larger models during this period (further results are available upon request). This is consistent with the evidence in Campbell et al. (2025), who document that value experienced an historically striking drop in performance during the pandemic.

models improves after the Global Financial Crisis. Indeed, only the CAPM (black line) and the two-factor model (red line) correctly price the stocks of interest before January 2010, albeit only for limited periods. Indeed, for most of the sample prior to the Global Financial Crisis all models are rejected. While the three- (green line) and the four-factor (light-blue line) models start to correctly price our panel right after the Global Financial Crisis, we have to wait some more years to appreciate the pricing relevance of the investment and profitability factors (FF5 in cyan, FF6 in purple). In light of these considerations, it is not surprising that we observe a strong comovement across results for all the models only after the late 2010s. On the other hand, such a similar behavior is always present for the CAPM and the two-factors model. Overall, there seems to be a general tendency of all asset pricing models to perform after the Global Financial Crisis, suggesting that the excess returns provided by individual stocks have increasingly reflected compensation for the systematic risk captured by the factors considered, while unpriced or idiosyncratic risk has reduced. Yet, as a note of caution, it is clear that even during the last 15 years of the sample, there are limited periods when none of the models prices our panel of stock returns fully.

7. Conclusions and discussion

In this paper, we propose a novel, general methodology to test for "zero alpha" in linear factor pricing models, with observable or latent, tradable or non-tradable pricing factors. The test relies on a randomization scheme and can be used even in the presence of conditional heteroskedasticity, strong cross-sectional dependence and non-Gaussianity in the error terms, allowing for the number of assets, N, to pass to infinity even at a faster rate than the sample size T. Extensive Monte Carlo analysis shows that the test has very good power properties, and that it is — compared to other extant approaches — the only one that controls the Type I Error probability in all scenarios.

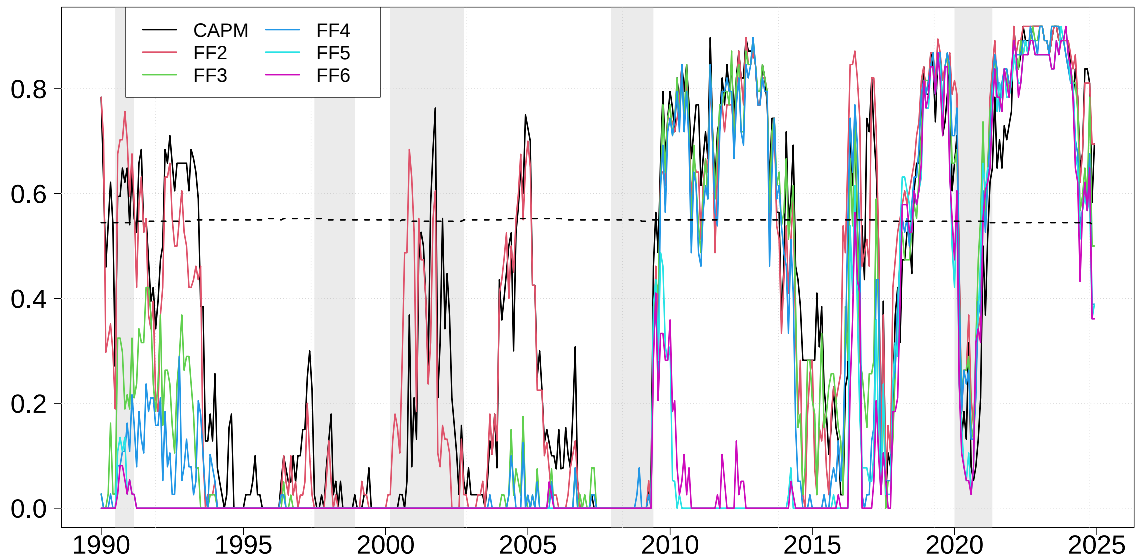


Figure 6.1. Values of $Q_{N,T,B}(0.05)$ with $\nu = 4$ for the six pricing models across 5-Yrs rolling windows. Results are based on constituents of the S&P 500 index. The horizontal dashed line represents the threshold based on $f(B) = B^{-1/4}$. Grey shaded areas correspond to: the US recession from Jul-1990 to Mar-1991, the Asian Financial Crisis (Jul-1997 to Dec-1998), the bust of the Dot-com Bubble and the months after September 11, 2001 (Mar-2000 to Oct-2002), the Global Financial Crisis (Dec-2007 to Jun-2009) and the COVID-19 pandemic (Jan-2020 to May-2021).

The proposed methodology requires, in essence, only a consistent estimator of the α_is, and its rate of convergence. Hence, several extensions of this methodology can be envisaged, in addition to the ones already presented in Section 4, e.g. by considering alternative estimation techniques. For example, when factors are selected from the "factors zoo", a linear factor pricing model can be specified after an initial factor-screening procedure based on variable selection algorithms such as the LASSO or the OCMT procedure of Chudik et al. (2018). As long as a consistent estimator for the α_is exists in this context, a test for the null of (1.4) can be developed. Finally, building on the theory developed herein, a randomized test could be built to test for the null hypotheses $\mathbb{H}_{0,i}$: $\alpha_i = 0$ for $i = 1, \ldots, N$ — or replacing the equality with an inequality, such as $\mathbb{H}_{0,i}$: $\alpha_i \leq 0$, as in Giglio et al. (2021) — while controlling for the well known multiple testing problem. Indeed, the individual test statistics would be perturbed by adding randomness independently across i, thus making the randomized statistics (conditionally) independent across i, which would render the application of several procedures for multiple testing more natural. Such an extension would be particularly valuable for testing mutual/hedge-funds performances, as well as for the assessment of trading strategies.

References

Ahn, S. C. and A. R. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*, 1203–1227.

Ang, A., J. Liu, and K. Schwarz (2020). Using stocks or portfolios in tests of factor models. *Journal of Financial and Quantitative Analysis 55*(3), 709–750.

Ardia, D. and R. Sessinou (2024). Robust inference in large panels and Markowitz portfolios. *Available at SSRN*.

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

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

Bailey, N., G. Kapetanios, and M. H. Pesaran (2021). Measurement of factor strength: Theory and practice. *Journal of Applied Econometrics 36*(5), 587–613.

Barigozzi, M. and L. Trapani (2022). Testing for common trends in nonstationary large datasets. *Journal of Business & Economic Statistics 40*(3), 1107–1122.

Berkes, I., S. Hörmann, and J. Schauer (2011). Split invariance principles for stationary processes. *Annals of Probability 39*(6), 2441–2473.

Bickel, P. J. and E. Levina (2008). Regularized estimation of large covariance matrices. *Annals of Statistics 36*(1), 199–227.

Black, F. (1976). Studies of stock price volatility changes. In *Proceedings from the American statistical association, business and economic statistics section*, pp. 177–181.

Blume, M. E. (1975). Betas and their regression tendencies. *The Journal of Finance 30*(3), 785–795.

Borjesson, P. and C.-E. Sundberg (1979). Simple approximations of the error function q (x) for communications applications. *IEEE Transactions on Communications 27*(3), 639–643.

Campbell, J. Y., S. Giglio, and C. Polk (2025). What drives booms and busts in value? Technical report, National Bureau of Economic Research.

Carrasco, M. and C. Nokho (2024). Hansen-Jagannathan distance with many assets. *Available at SSRN 5078000*.

Chernov, M., B. T. Kelly, S. Malamud, and J. Schwab (2025). A test of the efficiency of a given portfolio in high dimensions. Technical report, National Bureau of Economic Research.

Chudik, A., G. Kapetanios, and M. H. Pesaran (2018). A one covariate at a time, multiple testing approach to variable selection in high-dimensional linear regression models. *Econometrica 86*(4), 1479–1512.

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

Corradi, V. and N. R. Swanson (2006). The effects of data transformation on common cycle, cointegration, and unit root tests: Monte Carlo and a simple test. *Journal of Econometrics 132*, 195–229.

Davidson, J. (1994). *Stochastic limit theory: An introduction for econometricians*. OUP Oxford.

Degiannakis, S., G. Filis, G. Siourounis, and L. Trapani (2023). Superkurtosis. *Journal of Money, Credit and Banking 55*(8), 2061–2091.

Embrechts, P., C. Klüppelberg, and T. Mikosch (2013a). *Modelling extremal events: For insurance and finance*, Volume 33. Springer Science & Business Media.

Embrechts, P., C. Klüppelberg, and T. Mikosch (2013b). *Modelling extremal events: for insurance and finance*, Volume 33. Springer Science & Business Media.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Fan, J. and R. Li (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association 96*(456), 1348–1360.

Fan, J., Y. Liao, and M. Mincheva (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B 75*(4), 603–680.

Fan, J., Y. Liao, and W. Wang (2016). Projected principal component analysis in factor models. *Annals of statistics 44*(1), 219.

Fan, J., Y. Liao, and J. Yao (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica 83*(4), 1497–1541.

Feng, L., W. Lan, B. Liu, and Y. Ma (2022). High-dimensional test for alpha in linear factor pricing models with sparse alternatives. *Journal of Econometrics 229*(1), 152–175.

Gagliardini, P., E. Ossola, and O. Scaillet (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica 84*(3), 985–1046.

Geyer, C. J. and G. D. Meeden (2005). Fuzzy and randomized confidence intervals and p-values. *Statistical Science 20*(4), 358–366.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica 57*(5), 1121–1152.

Giglio, S., B. Kelly, and D. Xiu (2022). Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics 14*(1), 337–368.

Giglio, S., Y. Liao, and D. Xiu (2021). Thousands of alpha tests. *Review of Financial Studies 34*(7), 3456–3496.

Gormsen, N. J. and R. S. Koijen (2023). Financial markets and the covid-19 pandemic. *Annual Review of Financial Economics 15*(1), 69–89.

Gungor, S. and R. Luger (2016). Multivariate tests of mean-variance efficiency and spanning with a large number of assets and time-varying covariances. *Journal of Business & Economic Statistics 34*(2), 161–175.

Hall, P. (1979). On the rate of convergence of normal extremes. *Journal of Applied Probability 16*(2), 433–439.

Hansen, L. P. and R. Jagannathan (1997). Assessing specification errors in stochastic discount factor models. *Journal of Finance 52*(2), 557–590.

He, Y., X. Kong, L. Trapani, and L. Yu (2023). One-way or two-way factor model for matrix sequences? *Journal of Econometrics 235*(2), 1981–2004.

He, Y., X. Kong, L. Trapani, and L. Yu (2024). Online change-point detection for matrix-valued time series with latent two-way factor structure. *Annals of Statistics 52*(4), 1646–1670.

Hodrick, R. J. and X. Zhang (2001). Evaluating the specification errors of asset pricing models. *Journal of Financial Economics 62*(2), 327–376.

Hong, J., J. Li, and L. Sarno (2025). Dynamic currency mispricing and arbitrage profits. *Available at SSRN 5285379*.

Horváth, L. and L. Trapani (2019). Testing for randomness in a random coefficient autoregression model. *Journal of Econometrics 209*(2), 338–352.

Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory of Probability & Its Applications 7*(4), 349–382.

Kim, S., R. A. Korajczyk, and A. Neuhierl (2021). Arbitrage portfolios. *The Review of Financial Studies 34*(6), 2813–2856.

Lettau, M. and M. Pelger (2020a). Estimating latent asset-pricing factors. *Journal of Econometrics 218*(1), 1–31.

Lettau, M. and M. Pelger (2020b). Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies 33*(5), 2274–2325.

Liu, W. and Z. Lin (2009). Strong approximation for a class of stationary processes. *Stochastic Processes and Their Applications 119*(1), 249–280.

Liu, Y. and J. Xie (2020). Cauchy combination test: A powerful test with analytic p-value calculation under arbitrary dependency structures. *Journal of the American Statistical Association 115*(529), 393–402.

Ma, S., W. Lan, L. Su, and C.-L. Tsai (2020). Testing alphas in conditional time-varying factor models with high-dimensional assets. *Journal of Business & Economic Statistics 38*(1), 214–227.

Massacci, D. and L. Trapani (2022). High dimensional threshold regression with common stochastic trends. *Available at SSRN 4133488*.

Merikoski, J. K. and R. Kumar (2004). Inequalities for spreads of matrix sums and products. *Applied Mathematics E-Notes 4*(150-159), 9.

Moricz, F. (1983). A general moment inequality for the maximum of the rectangular partial sums of multiple series. *Acta Mathematica Hungarica 41*, 337–346.

Pesaran, M. H. and T. Yamagata (2024). Testing for alpha in linear factor pricing models with a large number of securities. *Journal of Financial Econometrics 22*(2), 407–460.

Petrov, V. V. (1995). Limit theorems of probability theory: sequences of independent random variables.

Potra, F. A. (1985). On superadditive rates of convergence. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique 19*(4), 671–685.

Raponi, V., C. Robotti, and P. Zaffaroni (2020). Testing beta-pricing models using large cross-sections. *Review of Financial Studies 33*(6), 2796–2842.

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

Sullivan, R., A. Timmermann, and H. White (1999). Data-snooping, technical trading rule performance, and the bootstrap. *Journal of Finance 54*(5), 1647–1691.

Trapani, L. (2016). Testing for (in)finite moments. *Journal of Econometrics 191*, 57–68.

Trapani, L. (2018). A randomized sequential procedure to determine the number of factors. *Journal of the American Statistical Association 113*(523), 1341–1349.

White, H. (2000). A reality check for data snooping. *Econometrica 68*(5), 1097–1126.

A. Further simulations

This section extends the Monte Carlo analysis of the main body by studying the finite sample performances of the test based on Theorem 3.1 under different data generating processes. Specifically, Sections A.1 and A.2 modify the persistence and factor structure for the error terms of the pricing models $u_{i,t}$, while Section A.3 discusses the implications of changes in the strength of the pricing factors $f_t = (f_{1,t}, f_{2,t}, f_{3,t})'$.

A.1. **No persistence in the omitted factor.** We complement the results of Section 5 under $\phi_g = 0$ in (5.3) - that is, the omitted (strong) common factor has no persistence. Tables A.1, A.3 and A.5 contain the empirical rejection frequencies under the null (left panels) and the alternative (right panels) for our one-shot test. For average type tests (FLY, GOS and PY) results are in line with those of the main body, though the degree of over-rejection under $\mathbb{H}_0$ is smaller. The p-value combination of AS is still mildly oversized, while the max-type approach of FLLM performs on par with ours.

Results for the "strong" decision rule of Theorem 3.1 are in Tables A.2, A.4 and A.6. As desired, empirical rejection frequencies quickly converge to zero under the null and for all cases considered. As in the main body, convergence is faster when using $f(B) = B^{-1/4}$. Similarly good results hold under the alternative, where the rejection frequencies always converge to one.

Finally, Figure A.1 reports empirical powers when $N = 500$ and $T = 100$ for twelve levels of sparsity under the alternative (these levels are the same as for Figure 5.1 in Section 5). We see that our tests almost always exhibits unit power, while some of the others incur in a loss of power as the alternative becomes sparser.

A.2. **Different strengths of the omitted factor.** We consider the same three-factor pricing model of Section 5. In the main body, $\gamma_i \overset{i.i.d.}{\sim} \mathcal{U}(0.7, 0.9)$ which implies that $\gamma_i > 0$ for any cross-sectional unit. We now depart from that assumption in two ways: first, we

Table A.1. Empirical rejection frequencies for the test in Theorem 3.1, Gaussian case.

		$\phi_\nu = 0$; $\alpha_i = 0$ for all i					$\phi_\nu = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.039	0.038	0.037	0.039	0.039	0.935	0.963	0.973	0.975	0.979
	FLLM	0.045	0.045	0.035	0.023	0.024	0.988	0.995	0.997	0.998	0.998
	FLY	0.249	0.148	0.096	0.087	0.100	0.995	0.996	0.998	1.000	1.000
	GOS	0.081	0.071	0.063	0.057	0.065	0.805	0.946	0.976	0.982	0.992
	PY	0.087	0.068	0.060	0.055	0.065	0.810	0.946	0.976	0.982	0.992
	AS	0.068	0.069	0.064	0.061	0.072	0.219	0.685	0.941	0.988	0.992
200	Thm. 1	0.049	0.034	0.033	0.034	0.033	0.998	0.999	1.000	1.000	1.000
	FLLM	0.055	0.027	0.027	0.028	0.028	1.000	1.000	1.000	1.000	1.000
	FLY	0.296	0.131	0.108	0.079	0.091	1.000	1.000	1.000	1.000	1.000
	GOS	0.088	0.062	0.064	0.065	0.061	0.888	0.986	0.997	0.999	1.000
	PY	0.096	0.059	0.062	0.064	0.061	0.902	0.986	0.997	0.999	1.000
	AS	0.079	0.059	0.080	0.078	0.082	0.253	0.833	0.984	1.000	1.000
500	Thm. 1	0.044	0.036	0.037	0.037	0.036	1.000	1.000	1.000	1.000	1.000
	FLLM	0.035	0.044	0.036	0.027	0.029	1.000	1.000	1.000	1.000	1.000
	FLY	0.385	0.196	0.141	0.106	0.089	1.000	1.000	1.000	1.000	1.000
	GOS	0.071	0.084	0.087	0.078	0.068	0.951	1.000	1.000	1.000	1.000
	PY	0.082	0.083	0.085	0.077	0.067	0.973	1.000	1.000	1.000	1.000
	AS	0.069	0.089	0.099	0.084	0.090	0.291	0.913	0.999	1.000	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$.

Table A.2. Empirical rejection frequencies for the decision rule of Theorem 3.1, Gaussian case.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.016	0.000	0.000	0.000	0.000	0.969	0.980	0.984	0.985	0.989
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000	0.969	0.980	0.984	0.985	0.989
200	LIL	0.009	0.000	0.000	0.000	0.000	0.999	0.999	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.002	0.000	0.000	0.000	0.000	0.999	0.999	1.000	1.000	1.000
500	LIL	0.276	0.137	0.080	0.050	0.028	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.001	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}$.

consider the case where only $\lfloor N^{0.4} \rfloor$ assets have a non-zero loading on g_t; secondly, we look at a situation where $\lfloor N^{0.8} \rfloor$ cross-sectional units have a non-zero loading on g_t. These two experiments correspond to the case of a weak and of a semi-strong omitted pricing factors,

Table A.3. Empirical rejection frequencies for the test in Theorem 3.1, Student's t case.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.051	0.042	0.038	0.038	0.038	0.893	0.931	0.953	0.960	0.966
	FLLM	0.052	0.033	0.036	0.032	0.037	0.979	0.992	0.999	1.000	1.000
	FLY	0.248	0.116	0.094	0.082	0.073	0.995	0.999	1.000	1.000	1.000
	GOS	0.099	0.067	0.071	0.073	0.082	0.786	0.940	0.976	0.983	0.990
	PY	0.116	0.077	0.073	0.071	0.082	0.842	0.944	0.976	0.982	0.988
	AS	0.054	0.056	0.058	0.055	0.060	0.175	0.555	0.869	0.963	0.980
200	Thm. 1	0.045	0.039	0.036	0.034	0.035	0.984	0.993	0.997	0.997	0.999
	FLLM	0.043	0.048	0.032	0.038	0.040	0.999	1.000	1.000	1.000	1.000
	FLY	0.282	0.162	0.119	0.103	0.080	1.000	1.000	1.000	1.000	1.000
	GOS	0.098	0.077	0.068	0.076	0.083	0.881	0.984	0.996	0.999	1.000
	PY	0.118	0.083	0.069	0.075	0.082	0.923	0.985	0.996	0.999	1.000
	AS	0.075	0.060	0.076	0.067	0.061	0.184	0.647	0.954	0.994	0.996
500	Thm. 1	0.050	0.038	0.039	0.039	0.037	1.000	1.000	1.000	1.000	1.000
	FLLM	0.053	0.038	0.041	0.037	0.042	1.000	1.000	1.000	1.000	1.000
	FLY	0.420	0.179	0.141	0.108	0.081	1.000	1.000	1.000	1.000	1.000
	GOS	0.162	0.095	0.074	0.068	0.066	0.958	1.000	1.000	1.000	1.000
	PY	0.125	0.095	0.077	0.068	0.065	0.988	1.000	1.000	1.000	1.000
	AS	0.077	0.070	0.084	0.084	0.083	0.225	0.791	0.989	0.997	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$.

Table A.4. Empirical rejection frequencies for the decision rule of Theorem 3.1, Student's t case.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.016	0.002	0.001	0.000	0.000	0.948	0.968	0.977	0.982	0.982
	$f(B) = B^{-1/4}$	0.001	0.001	0.000	0.000	0.000	0.895	0.935	0.952	0.964	0.972
200	LIL	0.016	0.000	0.000	0.000	0.000	0.996	0.997	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000	0.988	0.994	0.996	0.997	0.998
500	LIL	0.356	0.137	0.093	0.065	0.039	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.003	0.001	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}$.

respectively. Notably, the largest eigenvalue of the covariance matrix of $\mathbf{u}_t$ is bounded in the weak factor case, while it diverges to infinity in the semi-strong one. We consider the same sample sizes and alternative hypothesis as in Section 5.

TABLE A.5. Empirical rejection frequencies for the test in Theorem 3.1, GARCH case.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.056	0.043	0.040	0.039	0.039	0.967	0.983	0.989	0.99	0.992
	FLLM	0.045	0.045	0.035	0.023	0.024	0.988	0.995	0.997	0.998	0.998
	FLY	0.209	0.109	0.096	0.098	0.088	1.000	1.000	1.000	1.000	1.000
	GOS	0.081	0.071	0.063	0.057	0.065	0.805	0.946	0.976	0.982	0.992
	PY	0.087	0.068	0.060	0.055	0.065	0.810	0.946	0.976	0.982	0.992
	AS	0.068	0.069	0.064	0.061	0.072	0.219	0.685	0.941	0.988	0.992
200	Thm. 1	0.048	0.035	0.033	0.032	0.032	1.000	1.000	1.000	1.00	1.000
	FLLM	0.055	0.027	0.027	0.028	0.028	1.000	1.000	1.000	1.000	1.000
	FLY	0.234	0.139	0.104	0.091	0.085	1.000	1.000	1.000	1.000	1.000
	GOS	0.088	0.062	0.064	0.065	0.061	0.888	0.986	0.997	0.999	1.000
	PY	0.096	0.059	0.062	0.064	0.061	0.902	0.986	0.997	0.999	1.000
	AS	0.079	0.059	0.080	0.078	0.082	0.253	0.833	0.984	1.000	1.000
500	Thm. 1	0.065	0.041	0.035	0.036	0.035	1.000	1.000	1.000	1.00	1.000
	FLLM	0.035	0.044	0.036	0.027	0.029	1.000	1.000	1.000	1.000	1.000
	FLY	0.388	0.189	0.129	0.100	0.083	1.000	1.000	1.000	1.000	1.000
	GOS	0.071	0.084	0.087	0.078	0.068	0.951	1.000	1.000	1.000	1.000
	PY	0.082	0.083	0.085	0.077	0.067	0.973	1.000	1.000	1.000	1.000
	AS	0.069	0.089	0.099	0.084	0.090	0.291	0.913	0.999	1.000	1.000

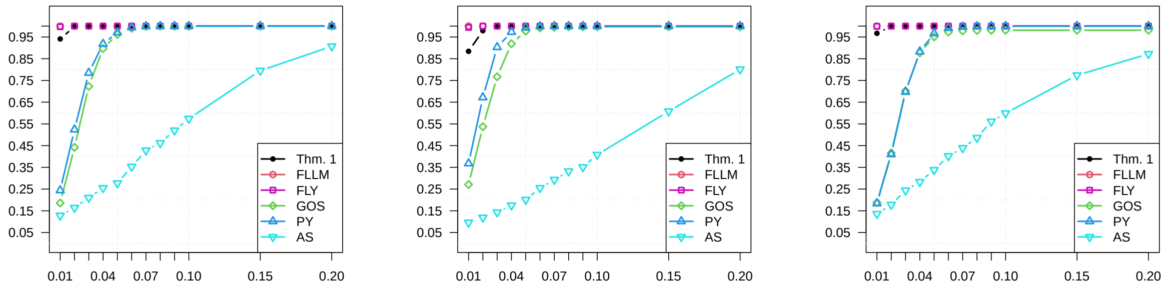
Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$; $\alpha_i \overset{i.i.d.}{\sim} N(0,1)$ for 5% of the cross-sectional units, under the alternative.

TABLE A.6. Empirical rejection frequencies for the decision rule of Theorem 3.1, GARCH case.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.026	0.003	0.000	0.000	0.000	0.985	0.989	0.995	0.996	0.995
	$f(B) = B^{-1/4}$	0.002	0.000	0.000	0.000	0.000	0.965	0.978	0.988	0.988	0.993
200	LIL	0.036	0.002	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.004	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
500	LIL	0.370	0.142	0.088	0.060	0.039	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.018	0.002	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}$; $\alpha_i \overset{i.i.d.}{\sim} N(0,1)$ for 5% of the cross-sectional units, under the alternative.

Empirical rejection frequencies for the weak factor case are in Table A.7. Results for our test and for those of Feng et al. (2022) and Fan et al. (2015) are very similar to those in



(a) Gaussian DGP, $\phi_g = 0$.

(b) Student's t DGP, $\phi_g = 0$.

(c) GARCH(1,1) DGP, $\phi_g = 0$.

Figure A.1. Power curves. The horizontal axis reports the percentages of mis-priced assets.

the main body (Table 5.3) both under both the null and the alternative hypothesis. Actual sizes of the GOS and PY tests are much closer to the nominal one, suggesting that strong-cross sectional dependence in the residuals was the driver of their overrejections. Their powers are substantially unaltered. The test of Ardia and Sessinou (2024) performs well in terms of size, but still exhibits a lack of power when T is small. All previous considerations hold unchanged with respect to the value of ϕ_g.

Table A.8 shows results for the semi-strong case. We would like to point out the importance of this data generating process, as Bailey et al. (2021) described tens of semi-strong pricing factor for the cross-section of US excess returns. Empirical rejection frequencies are very close to those of the main body, as the FLY, GOS, PY, and FLLM tests all become oversized when $\phi_g = 0.4$. This is most likely an effect of strong cross-sectional dependence in the residuals, as implied by a diverging eigenvalue in their covariance matrix. Results on the test by of Ardia and Sessinou (2024) are equivalent to those of Tables 5.3 and A.7.

A.3. **Strong and semi-strong pricing factors.** We now consider the same data generating process of Section A.2 but assuming that the omitted factor g_t is strong. This time, however, we assume that $\beta_{i,2} = \beta_{i,3} = 0$ for $N - \lfloor N^{0.8} \rfloor$ randomly chosen assets (cf. Bailey et al., 2021, who found that only the market factor is strong, while other 140 factors are at most semi-strong). Empirical rejection frequencies for this data generating process are

TABLE A.7. Empirical rejection frequencies for the test in Theorem 3.1, Student's t innovations with weak omitted common factor.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.040	0.027	0.035	0.035	0.033	0.941	0.959	0.968	0.976	0.980
	FLLM	0.048	0.057	0.044	0.045	0.034	0.984	0.998	0.999	1.000	1.000
	FLY	0.279	0.168	0.114	0.087	0.074	0.990	0.998	0.999	1.000	1.000
	GOS	0.084	0.069	0.063	0.056	0.057	0.983	0.994	0.998	0.999	1.000
	PY	0.063	0.053	0.058	0.052	0.053	0.981	0.994	0.997	0.999	1.000
	AS	0.050	0.053	0.064	0.050	0.059	0.153	0.513	0.881	0.967	0.985
200	Thm. 1	0.051	0.035	0.040	0.033	0.030	0.996	0.999	1.000	1.000	1.000
	FLLM	0.061	0.046	0.032	0.040	0.054	1.000	1.000	1.000	1.000	1.000
	FLY	0.325	0.146	0.107	0.082	0.085	1.000	1.000	1.000	1.000	1.000
	GOS	0.093	0.074	0.060	0.049	0.066	0.996	1.000	1.000	1.000	1.000
	PY	0.061	0.057	0.052	0.043	0.059	1.000	1.000	1.000	1.000	1.000
	AS	0.062	0.058	0.055	0.054	0.059	0.163	0.643	0.965	0.996	0.999
500	Thm. 1	0.030	0.054	0.031	0.032	0.031	1.000	1.000	1.000	1.000	1.000
	FLLM	0.071	0.051	0.049	0.061	0.043	1.000	1.000	1.000	1.000	1.000
	FLY	0.431	0.173	0.122	0.098	0.084	1.000	1.000	1.000	1.000	1.000
	GOS	0.099	0.064	0.062	0.061	0.053	0.993	1.000	1.000	1.000	1.000
	PY	0.044	0.046	0.050	0.053	0.048	1.000	1.000	1.000	1.000	1.000
	AS	0.046	0.043	0.050	0.054	0.054	0.177	0.807	0.994	0.999	1.000
		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0.4$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.048	0.036	0.028	0.031	0.035	0.944	0.963	0.971	0.976	0.978
	FLLM	0.055	0.068	0.056	0.050	0.046	0.984	0.999	1.000	1.000	1.000
	FLY	0.313	0.190	0.133	0.100	0.085	0.990	0.999	1.000	1.000	1.000
	GOS	0.120	0.110	0.089	0.079	0.080	0.984	0.995	0.997	0.999	1.000
	PY	0.088	0.089	0.082	0.073	0.074	0.983	0.994	0.997	0.999	1.000
	AS	0.045	0.056	0.065	0.055	0.059	0.153	0.512	0.874	0.967	0.984
200	Thm. 1	0.038	0.040	0.035	0.031	0.040	0.997	0.999	1.000	1.000	1.000
	FLLM	0.070	0.053	0.044	0.049	0.063	1.000	1.000	1.000	1.000	1.000
	FLY	0.346	0.170	0.123	0.101	0.100	1.000	1.000	1.000	1.000	1.000
	GOS	0.120	0.097	0.088	0.077	0.096	0.996	1.000	1.000	1.000	1.000
	PY	0.082	0.080	0.079	0.072	0.091	1.000	1.000	1.000	1.000	1.000
	AS	0.064	0.058	0.055	0.052	0.058	0.170	0.648	0.963	0.997	0.999
500	Thm. 1	0.051	0.050	0.036	0.036	0.030	1.000	1.000	1.000	1.000	1.000
	FLLM	0.073	0.059	0.050	0.067	0.045	1.000	1.000	1.000	1.000	1.000
	FLY	0.469	0.192	0.141	0.109	0.096	1.000	1.000	1.000	1.000	1.000
	GOS	0.122	0.088	0.079	0.083	0.070	0.993	1.000	1.000	1.000	1.000
	PY	0.065	0.059	0.061	0.075	0.065	1.000	1.000	1.000	1.000	1.000
	AS	0.049	0.042	0.049	0.052	0.052	0.178	0.809	0.994	0.999	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$. Only $\lfloor N^{0.4} \rfloor$ cross-sectional entities have a non-zero loading on the omitted common factor.

Table A.8. Empirical rejection frequencies for the test in Theorem 3.1; Student's t innovations with semi-strong omitted common factor.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.038	0.029	0.037	0.035	0.033	0.940	0.957	0.968	0.978	0.983
	FLLM	0.046	0.044	0.041	0.047	0.044	0.986	0.996	0.999	1.000	1.000
	FLY	0.267	0.147	0.114	0.090	0.081	0.996	0.998	0.999	1.000	1.000
	GOS	0.087	0.063	0.060	0.061	0.064	0.979	0.990	0.996	0.996	0.999
	PY	0.073	0.058	0.058	0.058	0.061	0.973	0.990	0.996	0.996	0.999
	AS	0.052	0.064	0.066	0.051	0.049	0.139	0.551	0.879	0.969	0.984
200	Thm. 1	0.051	0.037	0.041	0.036	0.030	0.995	0.998	0.998	0.998	0.999
	FLLM	0.053	0.048	0.042	0.048	0.046	1.000	1.000	1.000	1.000	1.000
	FLY	0.326	0.161	0.123	0.093	0.075	1.000	1.000	1.000	1.000	1.000
	GOS	0.119	0.078	0.056	0.060	0.054	0.995	1.000	1.000	1.000	1.000
	PY	0.091	0.069	0.048	0.059	0.053	0.998	1.000	1.000	1.000	1.000
	AS	0.042	0.046	0.045	0.051	0.058	0.154	0.653	0.965	0.996	0.999
500	Thm. 1	0.037	0.052	0.037	0.034	0.032	1.000	1.000	1.000	1.000	1.000
	FLLM	0.043	0.056	0.037	0.052	0.037	1.000	1.000	1.000	1.000	1.000
	FLY	0.413	0.170	0.140	0.105	0.099	1.000	1.000	1.000	1.000	1.000
	GOS	0.136	0.086	0.067	0.054	0.071	0.991	1.000	1.000	1.000	1.000
	PY	0.094	0.065	0.058	0.053	0.067	1.000	1.000	1.000	1.000	1.000
	AS	0.046	0.055	0.050	0.057	0.065	0.168	0.794	0.992	0.999	1.000
		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0.4$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.050	0.038	0.032	0.031	0.034	0.931	0.958	0.970	0.978	0.982
	FLLM	0.081	0.069	0.069	0.071	0.070	0.985	0.998	0.999	1.000	1.000
	FLY	0.332	0.183	0.140	0.122	0.109	0.995	0.999	1.000	1.000	1.000
	GOS	0.183	0.179	0.166	0.166	0.162	0.978	0.990	0.996	0.997	0.999
	PY	0.168	0.168	0.159	0.161	0.158	0.975	0.990	0.996	0.997	0.999
	AS	0.051	0.066	0.069	0.053	0.046	0.133	0.521	0.870	0.963	0.982
200	Thm. 1	0.046	0.042	0.033	0.033	0.040	0.995	0.997	0.999	1.000	1.000
	FLLM	0.083	0.086	0.086	0.080	0.070	1.000	1.000	1.000	1.000	1.000
	FLY	0.390	0.194	0.152	0.116	0.094	1.000	1.000	1.000	1.000	1.000
	GOS	0.214	0.188	0.169	0.158	0.162	0.992	1.000	1.000	1.000	1.000
	PY	0.197	0.182	0.163	0.152	0.158	0.997	1.000	1.000	1.000	1.000
	AS	0.041	0.060	0.056	0.049	0.060	0.142	0.631	0.958	0.993	0.998
500	Thm. 1	0.062	0.053	0.037	0.036	0.031	1.000	1.000	1.000	1.000	1.000
	FLLM	0.083	0.087	0.078	0.078	0.072	1.000	1.000	1.000	1.000	1.000
	FLY	0.480	0.213	0.163	0.124	0.116	1.000	1.000	1.000	1.000	1.000
	GOS	0.263	0.213	0.190	0.195	0.205	0.988	1.000	1.000	1.000	1.000
	PY	0.220	0.192	0.178	0.190	0.202	1.000	1.000	1.000	1.000	1.000
	AS	0.064	0.055	0.055	0.062	0.069	0.168	0.760	0.991	1.000	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$. Only $\lfloor N^{0.8} \rfloor$ cross-sectional entities have a non-zero loading on the omitted common factor.

reported in Table A.9. Results on all tests are substantially equivalent to those of Table 5.3, thus validating results of the main body also in the case where only one pricing factor is strong. This holds true irrespectively of the value of ϕ_g.

A.4. **Non-tradable factors.** We now study the finite sample properties of the testing procedure described in Theorem 4.1. To do it, we consider a three-factor pricing model similar to that of Section A.2 but with no factor structure in $u_{i,t}$, viz.

$$y_{i,t} = \alpha_i + \beta'\lambda + \sum_{p=1}^{3} \beta_{i,p} v_{p,t} + u_{i,t},$$

$$v_t = \Phi v_{t-1} + \zeta_t,$$

where we have constructed the DGP in terms of v_t for coherence with Section 4.1 (see equation (4.1), in particular). Similarly to the main body, $\mathbf{u}_t = (u_{1,t}, \ldots, u_{N,t})'$ follows one of the following three specifications:

(1) **The Gaussian case**: $\mathbf{u}_t \overset{i.i.d.}{\sim} \mathcal{N}_N(0, I_N)$.

(2) **The Student's t case**: $u_{i,t}$ follows a Students's t distribution with $d = 5.5$ degrees of freedom, zero mean and unit scale, independent across i. In this case, $u_{i,t}$ and $y_{i,t}$ have regularly varying tails;

(3) **The GARCH case**: we generate $\mathbf{u}_t = \mathbf{H}_t \mathbf{z}_t$, with: $\mathbf{z}_t = (z_{1,t}, ..., z_{N,t})'$ and $z_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$; and $\mathbf{H}_t = \text{diag}\{h_{1,t}, \ldots, h_{N,t}\}$ with $h_{i,t}^2 = \omega_i + \alpha_i \xi_{i,t}^2 + \beta_i h_{i,t-1}^2$.

Values of Φ, $\beta = (\beta_1, \beta_2, \beta_3)'$, and of GARCH parameters are set as in Section 5, while $\lambda_p \overset{i.i.d.}{\sim} \mathcal{U}(0, 1/2)$ for $p = 1, 2, 3$.

Table A.10 reports empirical rejection frequencies under the null (left panels) and under the alternative (right panel) for the test in Theorem 4.1. As in Section 5, we set the nominal size to $\tau = 5\%$ and study powers for the same significance level. Because our test is the only one with a fully-fledged asymptotic theory for non-tradabale factors, we do not report rejection frequencies for other approaches. Our procedure satisfactorily controls the size for any sample size and specification of the residuals $u_{i,t}$. Empirical powers are consistently one. Table A.11 presents rejection frequencies for a de-randomized procedure similar to that of Theorem 3.2 but based on Theorem 4.1 rather than 3.1. No matter the residuals' properties, empirical rejection frequencies always converge to zero as the sample size increases. As for the analysis in Section 5, this convergence is much faster using critical values based on $f(B) = B^{-1/4}$.

A.5. **Latent factors.** We now investigate the finite sample properties of the test in Theorem 4.2 when considering the DGP of Section A.4 with unobserved factors. Again, size and power are studied for nominal level $\tau = 5\%$. Results for empirical rejections frequencies under the null (left panels) and under the alternative (right panels) are reported in Table A.12. Again, the testing procedure exhibits satisfactory size control and excellent power properties. The test is a bit oversized in the GARCH case, particularly when $T = 100$. Given the results on the other estimators (standard OLS and Fama-MacBeth), this over-rejection seem to be due to problems in PCA-based estimation of factors and loadings under this particular GARCH DGP. Similar conclusions hold when we look at results on the de-randomized procedure in Table A.13, where we see that the approach works fine up to some over-rejection of the null in the GARCH case.

TABLE A.9. Empirical rejection frequencies for the test in Theorem 3.1; Student's t innovations with one strong and two semi-strong pricing factors.

		$\phi_g = 0$; $\alpha_i = 0$ for all i					$\phi_g = 0$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.036	0.029	0.038	0.035	0.034	0.893	0.932	0.951	0.961	0.969
	FLLM	0.041	0.033	0.028	0.032	0.041	0.980	0.995	1.000	1.000	1.000
	FLY	0.232	0.154	0.115	0.097	0.101	0.994	0.999	1.000	1.000	1.000
	GOS	0.069	0.071	0.074	0.068	0.074	0.794	0.947	0.973	0.981	0.988
	PY	0.087	0.072	0.072	0.068	0.072	0.850	0.949	0.973	0.981	0.987
	AS	0.068	0.057	0.066	0.050	0.065	0.148	0.552	0.872	0.966	0.982
200	Thm. 1	0.046	0.039	0.039	0.036	0.031	0.984	0.996	0.998	0.997	0.998
	FLLM	0.047	0.042	0.032	0.033	0.035	1.000	1.000	1.000	1.000	1.000
	FLY	0.284	0.144	0.114	0.098	0.085	1.000	1.000	1.000	1.000	1.000
	GOS	0.065	0.071	0.059	0.070	0.070	0.892	0.986	0.998	0.999	1.000
	PY	0.100	0.081	0.057	0.070	0.070	0.934	0.988	0.998	0.999	1.000
	AS	0.069	0.073	0.064	0.075	0.074	0.186	0.644	0.966	0.996	0.999
500	Thm. 1	0.034	0.053	0.033	0.035	0.032	1.000	1.000	1.000	1.000	1.000
	FLLM	0.049	0.037	0.028	0.028	0.032	1.000	1.000	1.000	1.000	1.000
	FLY	0.385	0.193	0.138	0.106	0.086	1.000	1.000	1.000	1.000	1.000
	GOS	0.097	0.060	0.061	0.063	0.067	0.972	0.999	1.000	1.000	1.000
	PY	0.127	0.070	0.062	0.063	0.065	0.992	1.000	1.000	1.000	1.000
	AS	0.075	0.068	0.083	0.077	0.084	0.215	0.782	0.997	0.999	1.000
		$\phi_g = 0.4$; $\alpha_i = 0$ for all i					$\phi_g = 0.4$; $\alpha_i \sim N(0,1)$ for 5% of units				
N	Test \T	100	200	300	400	500	100	200	300	400	500
100	Thm. 1	0.053	0.039	0.032	0.032	0.036	0.889	0.934	0.945	0.961	0.973
	FLLM	0.116	0.090	0.101	0.096	0.095	0.978	0.994	1.000	1.000	1.000
	FLY	0.332	0.183	0.140	0.122	0.109	0.994	1.000	1.000	1.000	1.000
	GOS	0.205	0.174	0.190	0.202	0.198	0.804	0.938	0.977	0.981	0.988
	PY	0.233	0.174	0.186	0.201	0.195	0.845	0.934	0.974	0.981	0.987
	AS	0.073	0.077	0.080	0.061	0.085	0.159	0.467	0.799	0.937	0.969
200	Thm. 1	0.050	0.045	0.039	0.034	0.043	0.984	0.994	0.998	0.999	1.000
	FLLM	0.121	0.105	0.095	0.099	0.103	1.000	1.000	1.000	1.000	1.000
	FLY	0.390	0.194	0.152	0.116	0.094	1.000	1.000	1.000	1.000	1.000
	GOS	0.193	0.202	0.179	0.186	0.202	0.877	0.984	0.996	0.999	1.000
	PY	0.240	0.201	0.178	0.184	0.200	0.909	0.984	0.996	0.999	1.000
	AS	0.087	0.094	0.090	0.090	0.109	0.185	0.565	0.914	0.993	0.994
500	Thm. 1	0.060	0.054	0.040	0.039	0.034	1.000	1.000	1.000	1.000	1.000
	FLLM	0.137	0.102	0.105	0.118	0.094	1.000	1.000	1.000	1.000	1.000
	FLY	0.480	0.213	0.163	0.124	0.116	1.000	1.000	1.000	1.000	1.000
	GOS	0.216	0.176	0.182	0.195	0.184	0.946	0.999	1.000	1.000	1.000
	PY	0.246	0.181	0.181	0.193	0.182	0.977	0.999	1.000	1.000	1.000
	AS	0.094	0.078	0.084	0.092	0.091	0.211	0.653	0.961	0.996	0.999

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}$. All the assets are exposed to the first pricing factor, while only $\lfloor N^{0.8} \rfloor$ cross-sectional entities have a non-zero loading on the remaining ones.

Table A.10. Empirical rejection frequencies for the test in Theorem 4.1.

	Gaussian case									
	$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 5% of units				
$N \backslash T$	100	200	300	400	500	100	200	300	400	500
100	0.041	0.033	0.025	0.028	0.024	1.000	1.000	1.000	1.000	1.000
200	0.048	0.037	0.038	0.027	0.045	1.000	1.000	1.000	1.000	1.000
500	0.043	0.039	0.033	0.045	0.032	1.000	1.000	1.000	1.000	1.000
	Student's t case									
	$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 5% of units				
$N \backslash T$	100	200	300	400	500	100	200	300	400	500
100	0.026	0.033	0.034	0.037	0.022	1.000	1.000	1.000	1.000	1.000
200	0.035	0.032	0.040	0.033	0.038	1.000	1.000	1.000	1.000	1.000
500	0.058	0.040	0.030	0.038	0.040	1.000	1.000	1.000	1.000	1.000
	GARCH case									
	$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 5% of units				
$N \backslash T$	100	200	300	400	500	100	200	300	400	500
100	0.055	0.045	0.032	0.035	0.043	1.000	1.000	1.000	1.000	1.000
200	0.081	0.037	0.042	0.025	0.019	1.000	1.000	1.000	1.000	1.000
500	0.091	0.038	0.040	0.036	0.045	1.000	1.000	1.000	1.000	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi_{i,NT}^{FM}$.

Table A.11. Empirical rejection frequencies for a de-randomized procedure based on Theorem 4.1.

		Gaussian case									
		$\phi_\nu = 0$; $\alpha_i = 0$ for all i					$\phi_\nu = 0$; $\alpha_i \sim N(0,1)$ for 1% of units				
N	C.V. \T	100	200	300	400	500	100	200	300	400	500
100	LIL	0.001	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
200	LIL	0.002	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.001	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
500	LIL	0.277	0.086	0.053	0.037	0.017	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
		Student's t case									
		$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 1% of units				
100	LIL	0.009	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
200	LIL	0.005	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.001	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
500	LIL	0.308	0.104	0.049	0.038	0.019	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
		GARCH noise									
		$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 1% of units				
100	LIL	0.039	0.005	0.001	0.001	0.000	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.003	0.002	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
200	LIL	0.067	0.009	0.000	0.001	0.001	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.010	0.004	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000
500	LIL	0.453	0.172	0.083	0.061	0.034	1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.026	0.003	0.000	0.001	0.000	1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi^{FM}_{i,NT}$.

Table A.12. Empirical rejection frequencies for the test in Theorem 4.1.

	Gaussian case									
	$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 5% of units				
$N \backslash T$	100	200	300	400	500	100	200	300	400	500
100	0.047	0.033	0.025	0.028	0.024	1.000	1.000	1.000	1.000	1.000
200	0.052	0.037	0.038	0.027	0.045	1.000	1.000	1.000	1.000	1.000
500	0.050	0.040	0.033	0.045	0.032	1.000	1.000	1.000	1.000	1.000
	Student's t case									
	$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 5% of units				
$N \backslash T$	100	200	300	400	500	100	200	300	400	500
100	0.033	0.032	0.033	0.033	0.022	1.000	1.000	1.000	1.000	1.000
200	0.042	0.033	0.041	0.032	0.038	1.000	1.000	1.000	1.000	1.000
500	0.065	0.040	0.030	0.038	0.041	1.000	1.000	1.000	1.000	1.000
	GARCH case									
	$\alpha_i = 0$ for all i					$\alpha_i \sim N(0,1)$ for 5% of units				
$N \backslash T$	100	200	300	400	500	100	200	300	400	500
100	0.080	0.047	0.030	0.034	0.043	1.000	1.000	1.000	1.000	1.000
200	0.131	0.051	0.043	0.027	0.019	1.000	1.000	1.000	1.000	1.000
500	0.193	0.064	0.049	0.039	0.046	1.000	1.000	1.000	1.000	1.000

Note: The nominal size is 5% and powers are assessed at 5% level of significance; frequencies are computed across $M = 1000$ Monte Carlo samples and we set $\nu = 5$ when computing $\psi^{PC}_{i,NT}$.

Table A.13. Empirical rejection frequencies for a de-randomized procedure based on Theorem 4.2.

		Gaussian case										
		$\phi_\nu = 0$; $\alpha_i = 0$ for all i						$\phi_\nu = 0$; $\alpha_i \sim N(0,1)$ for 1% of units				
N	C.V. \T	100	200	300	400	500		100	200	300	400	500
100	LIL	0.005	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
200	LIL	0.007	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
500	LIL	0.399	0.134	0.087	0.047	0.032		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000	
		Student's t case										
		$\alpha_i = 0$ for all i						$\alpha_i \sim N(0,1)$ for 1% of units				
100	LIL	0.007	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
200	LIL	0.015	0.000	0.001	0.000	0.000		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.002	0.000	0.001	0.000	0.000		1.000	1.000	1.000	1.000	1.000
500	LIL	0.378	0.138	0.088	0.054	0.036		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.003	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
		GARCH case										
		$\alpha_i = 0$ for all i						$\alpha_i \sim N(0,1)$ for 1% of units				
100	LIL	0.096	0.005	0.002	0.000	0.000		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.025	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
200	LIL	0.174	0.039	0.004	0.002	0.000		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.045	0.008	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
500	LIL	0.727	0.342	0.174	0.112	0.072		1.000	1.000	1.000	1.000	1.000
	$f(B) = B^{-1/4}$	0.120	0.017	0.002	0.004	0.001		1.000	1.000	1.000	1.000	1.000

Note: Results using either LIL-based critical values (LIL) or critical values based on $f(B) = B^{-1/4}$. The de-randomized statistic is based on nominal level $\tau = 5\%$. We set $B = \log(N)^2$ for the calculation of $Q_{N,T,B}(\tau)$ and $\nu = 5$ for that of $\psi_{i,NT}^{PC}$.

B. Complements to Sections 3.2 and 4

Here and henceforth, the Euclidean norm of a vector is denoted as $\|\cdot\|$. Given an $m \times n$ matrix $\mathbf{A}$ with element a_{ij} we use the following notation for its norms: $\|\mathbf{A}\|$ is the Euclidean/spectral norm, defined as $\|\mathbf{A}\| \leq \sqrt{\lambda_{\max}(\mathbf{A}'\mathbf{A})}$; $\|\mathbf{A}\|_F$ is the Frobenious norm; $\|\mathbf{A}\|_1$ is the $\mathcal{L}_1$-norm defined as $\|\mathbf{A}\|_1 = \max_{1\leq j\leq n}\sum_{i=1}^m |a_{ij}|$; the $\mathcal{L}_\infty$-norm $\|\mathbf{A}\|_\infty$ is defined as $\|\mathbf{A}\|_\infty = \max_{1\leq i\leq m}\sum_{j=1}^n |a_{ij}|$.

B.1. **Extension of the asymptotic regime in Assumption 3.1.** We discuss how to extend the asymptotic regime required in Assumption 3.1. In particular, we show hoe we can further relax the assumption to $N = O\left(T^{\nu/4-\varepsilon}\right)$, which allows for larger values of N compared to Assumption 3.1, and to the corresponding Assumption A1(iii) in Feng et al. (2022). In such a case, we would need to redefine $\psi_{i,NT}$ as

$$\psi_{i,NT} = \left|\frac{T^{\delta}\widehat{\alpha}_{i,T}}{\widehat{s}_{NT}}\right|^{\nu/2}, \tag{B.1}$$

where δ is a user-chosen quantity such that

$$0 < \delta < \frac{1}{2} - \frac{2}{\nu}\frac{\log N}{\log T}. \tag{B.2}$$

Equation (B.2) does not suggest a decision rule *per se*, but only an upper bound for δ, which is a tuning parameter. The rationale underpinning (B.1) is based on the fact that - upon inspecting the proofs of Theorem 3.1 and Lemma C.8 - we require that, under the null, $\sum_{i=1}^N \psi_{i,NT} = o_{a.s.}(1)$. In turn, this follows as long as $N\left|T^{\delta}\widehat{\alpha}_{i,T}\right|^{\nu/2}$ drifts to zero; intuitively, under the null $\widehat{\alpha}_{i,T}$ drifts to zero at a rate $O_{a.s.}\left(T^{-1/2}\right)$, and therefore $N\left|T^{\delta}\widehat{\alpha}_{i,T}\right|^{\nu/2} = O_{a.s.}\left(NT^{(\delta-1/2)\nu/2}\right) = o_{a.s.}(1)$ by the definition of δ in (B.2).

The same arguments hold in the case of nontradable and latent factors, upon replacing $T^{1/\nu}$ with T^{δ} in the definition of $\psi^{FM}_{i,NT}$, and $C^{1/\nu}_{NT}$ with C^{δ}_{NT} in the definition of $\psi^{PC}_{i,NT}$, respectively.

B.2. **Further assumptions - Section 4.** We consider the following assumptions for the case of *non-tradable factors.* Let

$$\mathbf{S}_\beta = \frac{1}{N}\sum_{i=1}^{N}\left(\beta_i - \overline{\beta}\right)\left(\beta_i - \overline{\beta}\right)', \tag{B.3}$$

where we define

$$\overline{\beta} = \frac{1}{N}\sum_{i=1}^{N}\beta_i. \tag{B.4}$$

Assumption B.1. *It holds that: (i)* λ *and* β_i *are fixed with* $\|\lambda\| < \infty$ *and* $\max_{1\le i\le N}\|\beta_i\| < \infty$*; and (ii)* $\mathbf{S}_\beta$ *is positive definite for all values of* N*.*

Assumption B.2. *It holds that: (i)* $\{u_{i,t}, 1 \le t \le T\}$ *and* $\{f_t, 1 \le t \le T\}$ *are two mutually independent groups; and (ii)* $\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{T}\sum_{t=1}^{T}\left|\mathbb{E}\left(u_{i,t}u_{j,s}\right)\right| \le c_0 NT$*.*

We further consider the following assumptions - which complement and extend the assumptions in the main paper - for the case of *latent factors.*

Assumption B.3. *It holds that: (i)* λ *and* $\boldsymbol{\beta}$ *are nonrandom with* $\|\lambda\| < \infty$ *and* $\|\boldsymbol{\beta}\| < \infty$*; (ii)* $\boldsymbol{\beta}'\boldsymbol{\beta} = N\mathbb{I}_K$*; and (iii)* $\mathbf{S}_\beta$ *is positive definite for all values of* N*.*

Assumption B.4. *It holds that* $\mathbb{E}\left(\widetilde{v}_t\widetilde{v}_t'\right)$ *is a positive definite matrix.*

Assumption B.5. *Let* $\gamma_{s,t} = \sum_{i=1}^{N}\mathbb{E}\left(u_{i,t}u_{i,s}\right)/N$*. It holds that: (i)* $\sum_{s=1}^{T}\left|\gamma_{s,t}\right| < c_0$ *for* $1 \le t \le T$*; (ii)* $\mathbb{E}\left|\sum_{i=1}^{N}\left(u_{i,s}u_{i,t} - \gamma_{s,t}\right)\right|^2 < c_0 N$ *for* $1 \le s, t \le T$*; (iii)* $\mathbb{E}\left\|\sum_{i=1}^{N}\beta_i u_{i,t}\right\|^4 < c_0 N^2$ *for* $1 \le t \le T$*; (iv)* $\sum_{i=1}^{N}\left|\mathbb{E}\left(u_{i,t}u_{j,s}\right)\right| \le c_0$ *for all* $1 \le t, s \le T$ *and* $1 \le j \le N$*; (v)* $\sum_{i=1}^{N}\sum_{s=1}^{T}\left|\mathbb{E}\left(u_{i,t}u_{j,s}\right)\right| \le c_0$ *for all* $1 \le t \le T$ *and* $1 \le j \le N$*; (vi)*

$$\mathbb{E}\left\|\sum_{t=1}^{T} w'\left(\mathbf{u}_t u_{i,t} - \mathbb{E}\left(\mathbf{u}_0 u_{i,0}\right)\right)\right\|^{\nu/2} \le c_0\left(NT\right)^{\nu/4},$$

for any w *such that* $\|w\| = O\left(N^{1/2}\right)$*.*

Assumption B.6. *It holds that* $\{v_t, 1 \leq t \leq T\}$ *and* $\{u_{i,t}, 1 \leq t \leq T\}$ *are two mutually independent groups, for* $1 \leq i \leq N$.

C. Technical lemmas

Henceforth, we denote the distribution function of the standard normal calculated at $-\infty < x < \infty$ as $\Phi(x)$.

C.1. **Preliminary lemmas.** We begin with a Baum-Katz-type theorem which is also reported in Massacci and Trapani (2022).

Lemma C.1. *Consider a multi-index partial sum process* $U_{S_1,..,S_h} = \sum_{i_2=1}^{S_2} \cdots \sum_{i_h=1}^{S_h} \xi_{i_1,...,i_h}$, *and assume that, for some* $q \geq 1$

$$\mathbb{E}\sum_{i_1=1}^{S_1} |U_{S_1,..,S_h}|^q \leq c_0 S_1 \prod_{j=2}^{h} S_j^{d_j},$$

where $d_j \geq 1$ *for all* $1 \leq j \leq h$. *Then it holds that*

$$\limsup_{\min\{S_1,...,S_h\}\to\infty} \frac{\sum_{i_1=1}^{S_1} |U_{S_1,..,S_h}|^q}{S_1 \prod_{j=2}^{h} S_j^{d_j} \left(\prod_{j=1}^{h} \log S_j\right)^{2+\epsilon}} = 0 \ a.s.,$$

for all $\epsilon > 0$.

Proof. We begin by noting that the function

$$g(x_1, ..., x_h) = x_1 \prod_{j=2}^{h} x_j^{d_j},$$

is superadditive. Consider the vector $(y_1, ..., y_h)$ such that $y_i \geq x_i$ for all $1 \leq i \leq h$. Then, for any two s and t such that $x_1 + s \leq y_1 + t$

$$\frac{1}{s}\left[g(x_1 + s, ..., x_h) - g(x_1, ..., x_h)\right] = \prod_{j=2}^{h} x_j^{d_j},$$

$$\frac{1}{t}\left[g(y_1 + t, ..., y_h) - g(y_1, ..., y_h)\right] = \prod_{j=2}^{h} x_j^{d_j},$$

whence it trivially follows that

$$\frac{1}{s}\left[g\left(x_1+s,....,x_h\right)-g\left(x_1,....,x_h\right)\right]=\frac{1}{t}\left[g\left(y_1+t,....,y_h\right)-g\left(y_1,....,y_h\right)\right].$$

He et al. (2023) also showed that, for any two nonzero s and t such that $x_i+s\leq y_i+t$, $2\leq i\leq h$

$$\frac{1}{s}\left[g\left(x_1+s,....,x_h\right)-g\left(x_1,....,x_h\right)\right]\leq\frac{1}{t}\left[g\left(y_1+t,....,y_h\right)-g\left(y_1,....,y_h\right)\right].$$

Thus, $g\left(x_1,....,x_h\right)$ is an S-convex function (see Definition 2.1 and Proposition 2.3 in Potra, 1985), and therefore it is superadditive (by Proposition 2.9 in Potra, 1985). Hence we can apply the maximal inequality in Corollary 4 in Moricz (1983) with - (in his notation) $f(R)=S_1\prod_{j=2}^{h}S_j^{d_j}$ and $\phi(\cdot)=c_0$. Letting

$$V_{i_1,\dots,i_h}=\sum_{j_1=1}^{i_1}\left|\sum_{j_2=1}^{i_2}\cdots\sum_{j_h=1}^{i_h}\xi_{j_1,\dots,j_h}\right|^q,$$

it follows that

$$\mathbb{E}\max_{1\leq i_1\leq S_1,....,1\leq i_h\leq S_h}V_{i_1,\dots,i_h}\leq c_0S_1\prod_{j=2}^{h}S_j^{d_j}\left(\prod_{j=1}^{h}\log S_j\right).$$

Hence we have

$$\begin{aligned}&\sum_{S_1=1}^{\infty}\cdots\sum_{S_h=1}^{\infty}\frac{1}{\prod_{j=1}^{h}S_j}\mathbb{P}\left(\max_{1\leq i_1\leq S_1,....,1\leq i_h\leq S_h}V_{i_1,\dots,i_h}\geq\varepsilon S_1\prod_{j=2}^{h}S_j^{d_j}\left(\prod_{j=1}^{h}\log S_j\right)^{2+\epsilon}\right)\\ \leq\ &\varepsilon^{-1}\sum_{S_1=1}^{\infty}\cdots\sum_{S_h=1}^{\infty}\frac{1}{S_1^2\prod_{j=2}^{h}S_j^{d_j+1}\left(\prod_{j=1}^{h}\log S_j\right)^{2+\epsilon}}\mathbb{E}\max_{1\leq i_1\leq S_1,....,1\leq i_h\leq S_h}V_{i_1,\dots,i_h}\end{aligned}$$

$$\leq \quad c_0\varepsilon^{-1}\sum_{S_1=1}^{\infty}\cdots\sum_{S_h=1}^{\infty}\frac{1}{\prod_{j=1}^{h}S_j\left(\prod_{j=1}^{h}\log S_j\right)^{1+\epsilon}}\leq c_1\varepsilon^{-1}.$$

The desired result now follows by repeating the proof of Lemma A.1 in Barigozzi and Trapani (2022). □

The following estimate on the growth rate of moments of partial sums will be used throughout the paper, and it can be contrasted with Proposition 4.1 in Berkes et al. (2011).

Lemma C.2. *Let w_t be a centered, $\mathcal{L}_p$-decomposable Bernoulli shift with $p > 2$ and $a > (p-1)/(p-2)$. Then it holds that*

$$E\left(\sum_{t=1}^{m}w_t\right)^p\leq c_0m^{p/2}. \tag{C.1}$$

Proof. We begin by showing that

$$E\left(\sum_{t=1}^{m}w_t\right)^2\leq c_0m. \tag{C.2}$$

By stationarity, we can write

$$\begin{aligned}E\left(\sum_{t=1}^{m}w_t\right)^2 &= E\left(\sum_{t=1}^{m}\sum_{s=1}^{m}w_tw_s\right)=mE\left(w_0^2\right)+2\sum_{t=1}^{m}(m-t)\,E\left(w_tw_0\right)\\ &\leq mE\left(w_0^2\right)+2\sum_{t=1}^{m}\left|E\left(w_tw_0\right)\right|.\end{aligned}$$

Consider now the coupling $\widetilde{w}_{t,t}$, and note that

$$E\left(w_tw_0\right)=E\left(\left(w_t-\widetilde{w}_{t,t}\right)w_0\right)+E\left(\widetilde{w}_{t,t}w_0\right)=E\left(\left(w_t-\widetilde{w}_{t,t}\right)w_0\right),$$

on account of the independence between $\widetilde{w}_{t,t}$ and w_0. Further

$$\left|E\left(\left(w_t-\widetilde{w}_{t,t}\right)w_0\right)\right|\leq\left|w_0\right|_2\left|w_t-\widetilde{w}_{t,t}\right|_2\leq c_0t^{-a},$$

and therefore

$$\sum_{t=1}^{m} \left| E\left(w_t w_0\right)\right| = O\left(m\right).$$

The desired result now follows by putting everything together. We now show the main result. Define $\widetilde{w}_{t,\ell}$ with $\ell = \lfloor m^{\varsigma} \rfloor$, where

$$\text{(C.3)} \qquad \frac{1}{2a} < \varsigma < \frac{p-2}{2\left(p-1\right)}.$$

It holds that

$$\begin{aligned} E\left(\sum_{t=1}^{m} w_t\right)^p &\leq 2^{p-1}\left(E\left(\sum_{t=1}^{m} \widetilde{w}_{t,\ell}\right)^p + E\left(\sum_{t=1}^{m} \left(w_t - \widetilde{w}_{t,\ell}\right)\right)^p\right) \\ &\leq 2^{p-1}\left(E\left(\sum_{t=1}^{m} \widetilde{w}_{t,\ell}\right)^p + E\left(\sum_{t=1}^{m} \left|w_t - \widetilde{w}_{t,\ell}\right|\right)^p\right). \end{aligned}$$

We have

$$E\left(\sum_{t=1}^{m} \left|w_t - \widetilde{w}_{t,\varsigma}\right|\right)^p \leq m^{p-1} \sum_{t=1}^{m} E\left|w_t - \widetilde{w}_{t,\varsigma}\right|^p \leq c_0 m^{p-1} m \ell^{-pa} \leq c_1 m^{p/2},$$

on account of (C.3). We now estimate $E\left(\sum_{t=1}^{m} \widetilde{w}_{t,\ell}\right)^p$; consider the $\lfloor m/\ell \rfloor + 1$ blocks

$$\mathcal{B}_i = \sum_{t=\ell(i-1)+1}^{\ell i} \widetilde{w}_{t,\ell},\ 1 \leq i \leq \lfloor m/\ell \rfloor \quad \text{and} \quad \mathcal{B}_{\lfloor m/\ell \rfloor + 1} = \sum_{t=\lfloor m/\ell \rfloor + 1}^{m} \widetilde{w}_{t,\ell}.$$

Note that, by construction, the sequence of blocks $\mathcal{B}_i$ with i even is an independent sequence, and so is the sequence of the $\mathcal{B}_i$s with odd i. Hence we can write

$$\sum_{t=1}^{m} \widetilde{w}_{t,\ell} = \sum_{i=1}^{\lfloor m/\ell \rfloor / 2} \mathcal{B}_{2i} + \sum_{i=1}^{\lfloor m/\ell \rfloor / 2} \mathcal{B}_{2(i-1)+1} + \mathcal{B}_{\lfloor m/\ell \rfloor + 1}.$$

Thus

$$E\left(\sum_{t=1}^{m} w_t\right)^p \leq 3^{p-1}\left(E\left(\sum_{i=1}^{\lfloor m/\ell \rfloor / 2} \mathcal{B}_{2i}\right)^p + E\left(\sum_{i=1}^{\lfloor m/\ell \rfloor / 2} \mathcal{B}_{2(i-1)+1}\right)^p + E\left(\mathcal{B}_{\lfloor m/\ell \rfloor + 1}\right)^p\right)$$

$$\leq \quad 3^{p-1}\left(E\left|\sum_{i=1}^{\lfloor m/\ell\rfloor/2}\mathcal{B}_{2i}\right|^p + E\left|\sum_{i=1}^{\lfloor m/\ell\rfloor/2}\mathcal{B}_{2(i-1)+1}\right|^p + E\left|\mathcal{B}_{\lfloor m/\ell\rfloor+1}\right|^p\right)$$

On account of the independence of the $\mathcal{B}_{2i}$s across i, we can use Rosenthal's inequality (see e.g. Theorem 2.9 in Petrov, 1995), whence

(C.4) $$E\left|\sum_{i=1}^{\lfloor m/\ell\rfloor/2}\mathcal{B}_{2i}\right|^p \leq c(p)\left(E\sum_{i=1}^{\lfloor m/\ell\rfloor/2}|\mathcal{B}_{2i}|^p + \left|\sum_{i=1}^{\lfloor m/\ell\rfloor/2}E\left(\mathcal{B}_{2i}^2\right)\right|^{p/2}\right),$$

where $c(p)$ is a positive, finite constant that depends only on p. We already know from (C.2) that

$$E\left(\mathcal{B}_{2i}^2\right) \leq c_0\ell,$$

and therefore

$$\left|\sum_{i=1}^{\lfloor m/\ell\rfloor/2}E\left(\mathcal{B}_{2i}^2\right)\right|^{p/2} \leq c_0 m^{p/2},$$

for some c_0. Further

$$\begin{aligned}
& E\sum_{i=1}^{\lfloor m/\ell\rfloor/2}|\mathcal{B}_{2i}|^p \\
= \quad & E\sum_{i=1}^{\lfloor m/\ell\rfloor/2}\left|\sum_{t=\ell(2i-1)+1}^{2\ell i}\widetilde{w}_{t,\ell}\right|^p \leq c_0\left\lfloor\frac{m}{\ell}\right\rfloor\ell^{p-1}\sum_{t=\ell(2i-1)+1}^{2\ell i}E\left|\widetilde{w}_{t,\ell}\right|^p \\
\leq \quad & c_1\frac{m}{\ell}\ell^p \leq c_2 m^{\varsigma(p-1)+1} \leq c_3 m^{p/2},
\end{aligned}$$

by the definition of ς in (C.3). Putting all together, (C.4) now yields

$$E\left|\sum_{i=1}^{\lfloor m/\ell\rfloor/2}\mathcal{B}_{2i}\right|^p \leq c_0 m^{p/2},$$

and the same holds for the odd blocks $\mathcal{B}_{2(i-1)+1}$, and, similarly, for $E\left|\mathcal{B}_{\lfloor m/\ell\rfloor+1}\right|^p$. Hence the final result follows. □

C.2. **Lemmas for Section 3.**

Lemma C.3. *We assume that Assumption 2.2 is satisfied. Then it holds that*

$$\overline{f} = \frac{1}{T}\sum_{t=1}^{T} f_t = \mathbb{E} f_t + o_{a.s.}(1).$$

Proof. We report the proof for the case $K = 1$, for simplicity and without loss of generality. The proof follows from standard arguments; indeed

$$\frac{1}{T}\sum_{t=1}^{T} f_t = \mathbb{E} f_t + \frac{1}{T}\sum_{t=1}^{T}(f_t - \mathbb{E} f_t).$$

Recall that, by Assumption 2.2, $f_t - \mathbb{E} f_t$ is a centered, $\mathcal{L}_\nu$-decomposable Bernoulli shift; thus, by Lemma C.2

$$\mathbb{E}\left|\sum_{t=1}^{T}(f_t - \mathbb{E} f_t)\right|^{\nu} \leq c_0 T^{\nu/2},$$

whence Lemma C.1 readily entails that

$$\frac{1}{T}\sum_{t=1}^{T}(f_t - \mathbb{E} f_t) = o_{a.s.}(1).$$

□

Lemma C.4. *We assume that Assumption 2.1 is satisfied. Then it holds that*

$$\sum_{i=1}^{N}\left|\sum_{t=1}^{T} u_{i,t}\right|^{\gamma} = o_{a.s.}\left(NT^{\gamma/2}(\log N \log T)^{2+\epsilon}\right),$$

for all $\epsilon > 0$ and all $\gamma \leq \nu$.

Proof. We estimate convergence rate of

$$\sum_{i=1}^{N}\mathbb{E}\left|\sum_{t=1}^{T} u_{i,t}\right|^{\gamma}.$$

By Assumption 2.1, we can use Lemma C.2, which entails that, for all $1 \leq i \leq N$

$$\mathbb{E}\left|\sum_{t=1}^{T} u_{i,t}\right|^{\gamma} \leq c_{\nu} T^{\gamma/2},$$

where c_{ν} is a positive, finite constant which depends only on ν, whence

$$\sum_{i=1}^{N} \mathbb{E}\left|\sum_{t=1}^{T} u_{i,t}\right|^{\gamma} \leq c_{0} N T^{\gamma/2}.$$

The desired result now readily obtains from Lemma C.1. □

Lemma C.5. *We assume that Assumption 2.2 is satisfied. Then it holds that*

$$\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)\left(f_t-\overline{f}\right)' = \mathcal{V}\left(f\right)+o_{a.s.}\left(1\right).$$

Proof. As above, we report the proof for the case $K=1$, for simplicity and without loss of generality. It holds that

$$\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2 = \frac{1}{T}\sum_{t=1}^{T} f_t^2 - \overline{f}^2.$$

Consider f_t^2; Assumption 2.2*(i)* immediately entails that $\{f_t^2, -\infty < t < \infty\}$ is an $\mathcal{L}_{\nu/2}$-decomposable Bernoulli shift with rate $a > 1$. Indeed, letting

$$f_t = g^{(f)}\left(\eta_t^{(f)}, \eta_{t-1}^{(f)}, \ldots\right),$$

where $g^{(f)}: S^{\infty} \to \mathbb{R}^K$ is a non random measurable function and $\left\{\eta_t^{(f)}, -\infty < t < \infty\right\}$ is an *i.i.d.* sequence with values in a measurable space S, and consider the coupling construction

$$f_t' = g^{(f)}\left(\eta_t^{(f)}, \ldots, \eta_{t-\ell+1}^{(f)}, \eta_{t-\ell,t,\ell}^{*(f)}, \eta_{t-\ell-1,t,\ell}^{*(f)} \cdots\right),$$

with $\left\{\eta_{s,t,\ell}^{*(f)}, -\infty < s, \ell, t < \infty\right\}$ *i.i.d.* copies of $\eta_0^{(f)}$ independent of $\left\{\eta_t^{(f)}, -\infty < t < \infty\right\}$. Then we have

$$\begin{aligned} & \left|f_t^2 - \left(f_t'\right)^2\right|_{\nu/2} \\ = & \left|\left(f_t + f_t'\right)\left(f_t - f_t'\right)\right|_{\nu/2} \leq \left|f_t + f_t'\right|_\nu \left|f_t - f_t'\right|_\nu \\ \leq & \ 2\left|f_t\right|_\nu \left|f_t - f_t'\right|_\nu \leq c_0 \ell^{-a}, \end{aligned}$$

having used the Cauchy-Schwartz inequality, Minkowski's inequality, and the facts that - by Assumption 2.2 - $|f_t|_\nu = |f_t'|_\nu < \infty$ and $|f_t - f_t'|_\nu \leq c_0 \ell^{-a}$, with $a > 1$. Hence

$$T^{-\nu/2} \mathbb{E}\left|\sum_{t=1}^T \left(f_t^2 - \mathbb{E} f_t^2\right)\right|^{\nu/2} \leq c_{\nu/2} T^{-\nu/4},$$

by Lemma C.2, from which it follows from standard arguments that

$$\frac{1}{T}\sum_{t=1}^T f_t^2 = \mathbb{E} f_t^2 + o_{a.s.}(1).$$

By the same token, it is not hard to see that

$$\overline{f} = \mathbb{E}\left(f_t\right) + o_{a.s.}(1).$$

Thus we have

$$\frac{1}{T}\sum_{t=1}^T \left(f_t - \overline{f}\right)^2 = \mathbb{E} f_t^2 - \left(\mathbb{E} f_t\right)^2 + o_{a.s.}(1),$$

and the desired result obtains from Assumption 2.2*(ii)*. □

Lemma C.6. *We assume that Assumptions 2.1-2.3 are satisfied. Then it holds that*

$$\sum_{i=1}^N \left\|\sum_{t=1}^T f_t u_{i,t}\right\|^\gamma = o_{a.s.}\left(N T^{\gamma/2} \left(\log N \log T\right)^{2+\epsilon}\right),$$

for every $\epsilon > 0$ *and* $\gamma \leq \nu/2$.

Proof. Let $K = 1$ with no loss of generality. We begin by showing that $\{f_t u_{i,t}, -\infty < t < \infty\}$ is an $\mathcal{L}_{\nu/2}$-decomposable Bernoulli shift with rate $a > 1$. Recall that, by Assumption 2.3, $E(f_t u_{i,t}) = 0$, and

$$\begin{aligned} f_t &= g^{(f)}\left(\eta_t^{(f)}, \eta_{t-1}^{(f)}, ...\right), \\ u_{i,t} &= g^{(u_i)}\left(\eta_t^{(i)}, \eta_{t-1}^{(i)}, ...\right), \end{aligned}$$

where $g^{(f)} : S^\infty \to \mathbb{R}^K$ and $g^{(u_i)} : S^\infty \to \mathbb{R}$, $1 \le i \le N$, are non random measurable function and $\left\{\eta_t^{(f)}, -\infty < t < \infty\right\}$ and $\left\{\eta_t^{(i)}, -\infty < t < \infty\right\}$ are *i.i.d.* sequences with values in a measurable space S, and consider the coupling constructions

$$\begin{aligned} f_t' &= g^{(f)}\left(\eta_t^{(f)}, ..., \eta_{t-\ell+1}^{(f)}, \eta_{t-\ell,t,\ell}^{*(f)}, \eta_{t-\ell-1,t,\ell}^{*(f)} \cdots\right), \\ u_{i,t}' &= g^{(u_i)}\left(\eta_t^{(i)}, ..., \eta_{t-\ell+1}^{(i)}, \eta_{t-\ell,t,\ell}^{*(i)}, \eta_{t-\ell-1,t,\ell}^{*(i)} \cdots\right), \end{aligned}$$

where $\left\{\eta_{s,t,\ell}^{*(f)}, -\infty < s, \ell, t < \infty\right\}$ and $\left\{\eta_{s,t,\ell}^{*(i)}, -\infty < s, \ell, t < \infty\right\}$ are *i.i.d.* copies of $\eta_0^{(f)}$ and $\eta_0^{(i)}$ respectively, independent of $\left\{\eta_t^{(f)}, -\infty < t < \infty\right\}$ and $\left\{\eta_t^{(i)}, -\infty < t < \infty\right\}$. Then we have, by elementary arguments

$$\begin{aligned} & \left|f_t u_{i,t} - f_t' u_{i,t}'\right|_\gamma \\ \le\ & \left|(f_t - f_t') u_{i,t}'\right|_\gamma + \left|f_t' (u_{i,t} - u_{i,t}')\right|_\gamma + \left|(f_t - f_t')(u_{i,t} - u_{i,t}')\right|_\gamma \\ \le\ & |u_{i,t}|_{2\gamma} |f_t - f_t'|_{2\gamma} + |f_t|_{2\gamma} \left|u_{i,t} - u_{i,t}'\right|_{2\gamma} + |f_t - f_t'|_{2\gamma} \left|u_{i,t} - u_{i,t}'\right|_{2\gamma} \\ \le\ & |u_{i,t}|_\nu |f_t - f_t'|_\nu + |f_t|_\nu \left|u_{i,t} - u_{i,t}'\right|_\nu + |f_t - f_t'|_\nu \left|u_{i,t} - u_{i,t}'\right|_\nu \\ \le\ & c_0 \ell^{-a} + c_1 \ell^{-a} + c_2 \ell^{-2a} \le c_3 \ell^{-a}. \end{aligned}$$

Then it holds that

$$\sum_{i=1}^N \mathbb{E}\left|\sum_{t=1}^T f_t u_{i,t}\right|^\gamma \le c_0 N T^{\gamma/2},$$

having used Lemma C.2. The desired result now follows from Lemma C.1. □

Lemma C.7. *We assume that Assumptions 2.1-2.3 are satisfied. Then it holds that*

$$\liminf_{\min\{N,T\}\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{u}_{i,t}^2 \quad > \quad 0,$$

$$\limsup_{\min\{N,T\}\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{u}_{i,t}^2 \quad < \quad \infty.$$

Proof. The proof uses several arguments used also elsewhere, so we omit passages when possible to avoid repetitions. It holds that

$$\begin{aligned}
& \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{u}_{i,t}^2 \\
= \quad & \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{i,t}^2 + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\alpha}_i-\alpha_i\right)^2 + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\beta}_i-\beta_i\right)^2 f_t^2 \\
& +\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\alpha}_i-\alpha_i\right)u_{i,t} + \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\beta}_i-\beta_i\right)f_t u_{i,t} \\
& +\frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\alpha}_i-\alpha_i\right)\left(\widehat{\beta}_i-\beta_i\right)f_t \\
= \quad & I+II+III+IV+V+VI.
\end{aligned}$$

It holds that

$$I = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}u_{i,t}^2 + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(u_{i,t}^2-\mathbb{E}u_{i,t}^2\right) = I_a + I_b.$$

By Assumption 2.1, it follows immediately that $0 < I_a < \infty$; also, it is easy to see that $u_{i,t}^2-\mathbb{E}u_{i,t}^2$ is a centered, $\mathcal{L}_{\nu/2}$-decomposable Bernoulli shift (see the arguments in the proof of Lemma C.5), and therefore, by Lemma C.2

$$\mathbb{E}\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(u_{i,t}^2-\mathbb{E}u_{i,t}^2\right)\right|^2$$

$$\leq \frac{1}{NT^2}\sum_{i=1}^{N}\mathbb{E}\left|\sum_{t=1}^{T}\left(u_{i,t}^2-\mathbb{E}u_{i,t}^2\right)\right|^2 \leq c_0 T^{-1},$$

whence Lemma C.1 yields $I_b = o_{a.s.}(1)$. Note also that

$$\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}f_t^2\right),$$

with $T^{-1}\sum_{t=1}^{T}f_t^2 = O_{a.s.}(1)$ by Lemma C.3 and

$$\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)^2 = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)u_{i,t}\right)^2}{\left(\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2\right)^2}.$$

We know from Lemma C.5 that

$$\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2 = c_0 + o_{a.s.}(1),$$

with $c_0 > 0$. Further, using Lemma C.6, it follows that

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)u_{i,t}\right\|^2 = o_{a.s.}(1),$$

whence $III = o_{a.s.}(1)$. The same arguments as in the proof of Lemma C.8 entail that $II = o_{a.s.}(1)$. Finally, a routine application of Hölder's inequality yields that $IV - VI = o_{a.s.}(1)$. □

Lemma C.8. *We assume that Assumptions 2.1-3.1 are satisfied. Then, under the null in (1.4) it holds that*

$$\sum_{i=1}^{N}\psi_{i,NT} = o_{a.s.}(1).$$

Proof. Let - for simplicity and with no loss of generality - $K = 1$. Recall (3.1), whence also

$$\widehat{\alpha}_i = \alpha_i - \left(\widehat{\beta}_i-\beta_i\right)\overline{f} + \overline{u}_i = -\left(\widehat{\beta}_i-\beta_i\right)\overline{f} + \overline{u}_i,$$

under $\mathbb{H}_0$. Hence we have

$$\begin{aligned}\sum_{i=1}^{N}\psi_{i,NT} &= \frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}|\widehat{\alpha}_i|^{\nu/2}\\ &\leq \frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}|\overline{u}_i|^{\nu/2}+\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}\left|\left(\widehat{\beta}_i-\beta_i\right)\overline{f}\right|^{\nu/2}\\ &= \frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}|\overline{u}_i|^{\nu/2}+\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}\left|\frac{\sum_{t=1}^{T}\left(f_t\overline{f}-\overline{f}^2\right)u_{i,t}}{\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2}\right|^{\nu/2},\end{aligned}$$

We know from Lemma C.7 that there exists a positive, finite constant c_0 and a couple of random variables (N_0, T_0) such that, for all $N \geq N_0$ and $T \geq T_0$

$$\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}|\overline{u}_i|^{\nu/2}\leq c_0T^{1/2}\sum_{i=1}^{N}|\overline{u}_i|^{\nu/2};$$

using Lemma C.4, it follows that

$$\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}|\overline{u}_i|^{\nu/2}=o_{a.s.}\left(NT^{1/2}T^{-\nu/4}\left(\log N\log T\right)^{2+\epsilon}\right)=o_{a.s.}\left(1\right),$$

by Assumption 3.1. Also

$$\begin{aligned}&\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\sum_{i=1}^{N}\left|\frac{\sum_{t=1}^{T}\left(f_t\overline{f}-\overline{f}^2\right)u_{i,t}}{\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2}\right|^{\nu/2}\\ \leq\;&\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\frac{\sum_{i=1}^{N}\left|\overline{f}\frac{1}{T}\sum_{t=1}^{T}f_tu_{i,t}\right|^{\nu/2}}{\left|\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2\right|^{\nu/2}}+\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\frac{\sum_{i=1}^{N}\left|\overline{f}^2\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right|^{\nu/2}}{\left|\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2\right|^{\nu/2}}.\end{aligned}$$

Lemmas C.3, C.5 and C.7 entail that there exists a positive, finite constant c_0 and a random variable T_0 such that, for all $T \geq T_0$

$$\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}}\frac{\sum_{i=1}^{N}\left|\overline{f}\frac{1}{T}\sum_{t=1}^{T}f_tu_{i,t}\right|^{\nu/2}}{\left|\frac{1}{T}\sum_{t=1}^{T}\left(f_t-\overline{f}\right)^2\right|^{\nu/2}}\;\leq\;c_0T^{1/2}\sum_{i=1}^{N}\left|\frac{1}{T}\sum_{t=1}^{T}f_tu_{i,t}\right|^{\nu/2},$$

$$\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}} \frac{\sum_{i=1}^{N} \left| \overline{f}^2 \frac{1}{T} \sum_{t=1}^{T} u_{i,t} \right|^{\nu/2}}{\left| \frac{1}{T} \sum_{t=1}^{T} \left(f_t - \overline{f} \right)^2 \right|^{\nu/2}} \leq c_0 T^{1/2} \sum_{i=1}^{N} \left| \frac{1}{T} \sum_{t=1}^{T} u_{i,t} \right|^{\nu/2} .$$

We already know from the above that the second term is $o_{a.s.}(1)$. Using Lemma C.6, it finally follows that

$$\frac{T^{1/2}}{|\hat{s}_{NT}|^{\nu/2}} \frac{\sum_{i=1}^{N} \left| \overline{f} \frac{1}{T} \sum_{t=1}^{T} f_t u_{i,t} \right|^{\nu/2}}{\left| \frac{1}{T} \sum_{t=1}^{T} \left(f_t - \overline{f} \right)^2 \right|^{\nu/2}} = o_{a.s.} \left(N T^{1/2} T^{-\nu/4} \left(\log N \log T \right)^{2+\epsilon} \right),$$

which converges to zero almost surely under Assumption 3.1. The desired result now obtains by putting all together. □

C.3. **Lemmas for Section 4.1.** We now report a series of lemmas for the case, discussed in Section 4.1, of nontradable factors. In order for the notation not to be overly burdensome, we will assume - unless otherwise stated - $K = 1$ whenever possible and with no loss of generality.

Recall the short-hand notation $\mathbf{S}_\beta$ defined in (B.3), let $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)'$ and

$$\widehat{\mathbf{S}}_\beta = \frac{1}{N} \sum_{i=1}^{N} \left(\widehat{\beta}_i - \frac{1}{N} \sum_{i=1}^{N} \widehat{\beta}_i \right) \left(\widehat{\beta}_i - \frac{1}{N} \sum_{i=1}^{N} \widehat{\beta}_i \right)', \tag{C.5}$$

and

$$\mathbf{S}_f = \frac{1}{T} \sum_{t=1}^{T} \left(f_t - \overline{f} \right) \left(f_t - \overline{f} \right)', \tag{C.6}$$

and note that, after standard passages

$$\begin{aligned} \widehat{\lambda} &= \lambda + \frac{1}{N} \mathbf{S}_\beta^{-1} \boldsymbol{\beta}' \mathbb{M}_{1_N} \boldsymbol{\alpha} + \overline{v} + \frac{1}{N} \left(\widehat{\mathbf{S}}_\beta^{-1} - \mathbf{S}_\beta^{-1} \right) \boldsymbol{\beta}' \mathbb{M}_{1_N} \boldsymbol{\alpha} \\ &\quad + \frac{1}{N} \widehat{\mathbf{S}}_\beta^{-1} \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right)' \mathbb{M}_{1_N} \boldsymbol{\alpha} + \frac{1}{N} \widehat{\mathbf{S}}_\beta^{-1} \widehat{\boldsymbol{\beta}}' \mathbb{M}_{1_N} \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \lambda \\ &\quad + \frac{1}{N} \widehat{\mathbf{S}}_\beta^{-1} \widehat{\boldsymbol{\beta}}' \mathbb{M}_{1_N} \left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}} \right) \overline{v} + \frac{1}{N} \widehat{\mathbf{S}}_\beta^{-1} \boldsymbol{\beta}' \mathbb{M}_{1_N} \overline{\mathbf{u}} \end{aligned} \tag{C.7}$$

$$+\frac{1}{N}\widehat{\mathbf{S}}_{\beta}^{-1}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\overline{\mathbf{u}}.$$

Finally, under both $\mathbb{H}_0$ and $\mathbb{H}_A$ we have

$$\widehat{\alpha}_i^{FM}=\alpha_i+\beta_i'\overline{v}+\overline{u}_i-\left(\widehat{\beta}_i-\beta_i\right)'\lambda-\beta_i'\left(\widehat{\lambda}-\lambda\right)-\left(\widehat{\beta}_i-\beta_i\right)'\left(\widehat{\lambda}-\lambda\right). \tag{C.8}$$

Lemma C.9. *We assume that Assumptions 2.1-2.3 and B.1 and B.2 are satisfied. Then it holds that*

$$N^{-1}\left\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2 = o_{a.s.}\left(\frac{(\log N\log T)^{2+\epsilon}}{T}\right), \tag{C.9}$$

$$\left\|\widehat{\mathbf{S}}_{\beta}-\mathbf{S}_{\beta}\right\| = o_{a.s.}\left(\frac{(\log N\log T)^{2+\epsilon}}{T}\right)+o_{a.s.}\left(\frac{(\log N\log T)^{1+\epsilon}}{\sqrt{NT}}\right), \tag{C.10}$$

for all $\epsilon>0$.

Proof. Recall that we use $K=1$, and note that

$$\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}=\mathbf{S}_f^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t\left(v_t-\overline{v}\right)\right),$$

having defined $\mathbf{u}_t=(u_{1,t},...,u_{N,t})'$ and $\overline{v}=T^{-1}\sum_{t=1}^{T}v_t$, whence also

$$\widehat{\beta}_i-\beta_i=\mathbf{S}_f^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\left(v_t-\overline{v}\right)u_{i,t}\right).$$

Note that

$$\begin{aligned}&N^{-1}\left\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2\\ \leq\ &N^{-1}\sum_{i=1}^{N}\left|\widehat{\beta}_i-\beta_i\right|^2=\mathbf{S}_f^{-2}\times N^{-1}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left(v_t-\overline{v}\right)u_{i,t}\right)^2,\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E}\left[N^{-1}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left(v_t-\overline{v}\right)u_{i,t}\right)^2\right] \\
= & \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}\left[\left(v_t-\overline{v}\right)\left(v_s-\overline{v}\right)\right]E\left(u_{i,t}u_{i,s}\right) \\
\leq & \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\mathbb{E}\left|v_t-\overline{v}\right|^2\right)\left|E\left(u_{i,t}u_{i,s}\right)\right| \leq c_0 T,
\end{aligned}$$

now the desired result follows by using Lemma C.5. Turning to (C.10), note that

$$\widehat{\mathbf{S}}_\beta = \frac{1}{N}\sum_{i=1}^{N}\widehat{\beta}_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\beta}_i\right)^2, \qquad \mathbf{S}_\beta = \frac{1}{N}\sum_{i=1}^{N}\beta_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\right)^2$$

with

$$\widehat{\beta}_i^2 = \left(\beta_i + \mathbf{S}_f^{-1}\frac{1}{T}\sum_{t=1}^{T}u_{i,t}(v_t-\overline{v})\right)^2. \tag{C.11}$$

Hence

$$\begin{aligned}
\left\|\widehat{\mathbf{S}}_\beta - \mathbf{S}_\beta\right\| \leq & \left\|\frac{1}{N}\sum_{i=1}^{N}\widehat{\beta}_i^2 - \frac{1}{N}\sum_{i=1}^{N}\beta_i^2\right\| + \left\|\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i+\beta_i\right)\right\|\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\right\| \\
= & I + II.
\end{aligned}$$

Using (C.11)

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}\widehat{\beta}_i^2 = & \frac{1}{N}\sum_{i=1}^{N}\beta_i^2 + \frac{1}{NT^2}\sum_{i=1}^{N}\mathbf{S}_f^{-2}\left(\sum_{t=1}^{T}u_{i,t}\left(v_t-\overline{v}\right)\right)^2 + \mathbf{S}_f^{-1}\frac{2}{NT}\sum_{i=1}^{N}\beta_i\sum_{t=1}^{T}u_{i,t}\left(v_t-\overline{v}\right) \\
= & I_a + I_b + I_c;
\end{aligned}$$

the same passages as above readily yield

$$I_b = o_{a.s.}\left(\frac{\left(\log N \log T\right)^{2+\epsilon}}{T}\right);$$

also, using Assumptions B.1 and B.2

$$
\begin{aligned}
& \mathbb{E}\left(\frac{1}{NT}\sum_{i=1}^{N}\beta_i\sum_{t=1}^{T}u_{i,t}\left(v_t-\overline{v}\right)\right)^2 \\
= \; & \frac{1}{N^2T^2}\mathbb{E}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\beta_i\beta_j\sum_{t=1}^{T}\sum_{s=1}^{T}\left(v_t-\overline{v}\right)\left(v_s-\overline{v}\right)u_{i,t}u_{j,s}\right) \\
\leq \; & \frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\mathbb{E}\left(v_t-\overline{v}\right)^2\right)\left|\mathbb{E}\left(u_{i,t}u_{j,s}\right)\right| \leq c_0\left(NT\right)^{-1},
\end{aligned}
$$

so that

$$
I_c = o_{a.s.}\left(\frac{\left(\log N\log T\right)^{1+\epsilon}}{\left(NT\right)^{1/2}}\right).
$$

By the same token, turning to II we have

$$
\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)=\mathbf{S}_f^{-1}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left(v_t-\overline{v}\right)u_{i,t}\right),
$$

where, by the same logic as above, it follows that

$$
\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\right\| = o_{a.s.}\left(\frac{\left(\log N\log T\right)^{1+\epsilon}}{\left(NT\right)^{1/2}}\right).
$$

The desired result follows from noting

$$
\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i+\beta_i\right)\right\| \leq \left\|\frac{2}{N}\sum_{i=1}^{N}\beta_i\right\| + \left\|\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\right\|.
$$

□

Lemma C.10. *We assume that Assumptions 2.1-2.3 and B.1 and B.2 are satisfied. Then it holds that*

$$
\sum_{i=1}^{N}\left\|\widehat{\beta}_i-\beta_i\right\|^{\nu/2} = o_{a.s.}\left(NT^{-\nu/4}\left(\log N\log T\right)^{2+\epsilon}\right),
$$

for all $\epsilon > 0$.

Proof. Recall that

$$\widehat{\beta}_i - \beta_i = \mathbf{S}_f^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}(v_t - \overline{v})\, u_{i,t}\right).$$

Then

$$\sum_{i=1}^{N}\left\|\widehat{\beta}_i - \beta_i\right\|^{\nu/2} \leq c_0\left(\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}v_t u_{i,t}\right\|^{\nu/2} + \sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}\overline{v}u_{i,t}\right\|^{\nu/2}\right),$$

and we can readily show - by following the arguments above - that

$$\mathbb{E}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}v_t u_{i,t}\right\|^{\nu/2} \leq c_0 N T^{-\nu/4},$$

so that

$$\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}v_t u_{i,t}\right\|^{\nu/2} = o_{a.s.}\left(NT^{-\nu/4}\left(\log N \log T\right)^{2+\epsilon}\right).$$

Recall that $\overline{v} = o_{a.s.}T^{-1/2}\left(\log T\right)^{1+\epsilon}$, and note

$$\mathbb{E}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right\|^{\nu/2} \leq c_0 N T^{-\nu/4},$$

so that ultimately we receive the desired result by putting all together. □

We now report two lemmas on the rates of $\widehat{\lambda} - \lambda$ under the null and under the alternative.

Lemma C.11. *We assume that $\mathbb{H}_0$ of (1.4) holds, and that Assumptions 2.1-2.3 and B.1 and B.2 are satisfied. Then it holds that*

$$\widehat{\lambda} - \lambda = o_{a.s.}\left(T^{-1/2}\left(\log T\right)^{1+\epsilon}\right),$$

for all $\epsilon > 0$.

Proof. Recall (C.7). Under $\mathbb{H}_0$, it holds that $\boldsymbol{\alpha} = 0$ and therefore

$$\widehat{\lambda} \;=\; \lambda + \overline{v} + \frac{1}{N}\widehat{\mathbf{S}}_{\beta}^{-1}\widehat{\boldsymbol{\beta}}'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\lambda$$

$$
\begin{aligned}
&+\frac{1}{N}\widehat{\mathbf{S}}_{\beta}^{-1}\widehat{\boldsymbol{\beta}}'\mathbb{M}_{1_N}\left(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}\right)\overline{v}+\frac{1}{N}\widehat{\mathbf{S}}_{\beta}^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\overline{\mathbf{u}} \\
&+\frac{1}{N}\widehat{\mathbf{S}}_{\beta}^{-1}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\overline{\mathbf{u}} \\
= \;& \lambda+I+II+III+IV+V.
\end{aligned}
$$

We begin by noting that, from standard passages, $I = o_{a.s.}\left(T^{-1/2}\left(\log T\right)^{1+\epsilon}\right)$. Note that, combining (C.10) and Assumption B.1*(ii)*

$$
\begin{aligned}
\left\|\widehat{\mathbf{S}}_{\beta}^{-1}-\mathbf{S}_{\beta}^{-1}\right\| &\leq \left\|\left(\widehat{\mathbf{S}}_{\beta}\pm\mathbf{S}_{\beta}\right)^{-1}\right\|\left\|\mathbf{S}_{\beta}^{-1}\right\|\left\|\widehat{\mathbf{S}}_{\beta}-\mathbf{S}_{\beta}\right\| \\
&= o_{a.s.}\left(\frac{\left(\log N\log T\right)^{2+\epsilon}}{T}\right)+o_{a.s.}\left(\frac{\left(\log N\log T\right)^{1+\epsilon}}{\sqrt{NT}}\right),
\end{aligned}
$$

and therefore we have

$$
\left\|\widehat{\mathbf{S}}_{\beta}^{-1}\right\| = O_{a.s.}\left(1\right).
$$

Consider now

$$
\begin{aligned}
&\left\|\frac{1}{N}\widehat{\mathbf{S}}_{\beta}^{-1}\widehat{\boldsymbol{\beta}}'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\lambda\right\| \\
\leq \;& \frac{1}{N}\left\|\widehat{\mathbf{S}}_{\beta}^{-1}\pm\mathbf{S}_{\beta}^{-1}\right\|\left\|\left(\widehat{\boldsymbol{\beta}}\pm\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right\|\left\|\lambda\right\|.
\end{aligned}
$$

We have

$$
\begin{aligned}
&\left\|\left(\widehat{\boldsymbol{\beta}}\pm\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right\| \\
\leq \;& \left\|\boldsymbol{\beta}'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right\|+\left\|\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right\| \\
= \;& a+b.
\end{aligned}
$$

Consider a, and let $\boldsymbol{\beta}'\mathbb{M}_{1_N} = \mathbf{w}'$ for short; we have

$$\frac{1}{N}\mathbf{w}'\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) = \frac{1}{N}\sum_{i=1}^{N} w_i\left(\widehat{\beta}_i - \beta_i\right) = \frac{1}{N}\sum_{i=1}^{N} w_i \frac{1}{T\mathbf{S}_f}\sum_{t=1}^{T} u_{i,t}\left(v_t - \overline{v}\right),$$

and therefore

$$\begin{aligned}
& \mathbb{E}\left|\frac{1}{N}\mathbf{w}'\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right|^2 \\
= \ & \frac{1}{N^2T^2\mathbf{S}_f^2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \mathbb{E}\left[\sum_{t=1}^{T}\sum_{s=1}^{T} u_{i,t}u_{i,s}\left(v_t-\overline{v}\right)\left(v_s-\overline{v}\right)\right] \\
\leq \ & c_0\frac{1}{N^2T^2\mathbf{S}_f^2}\mathbb{E}\left[\left(v_t-\overline{v}\right)^2\right]\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\mathbb{E}\left(u_{i,t}u_{i,s}\right)\right| \leq c_1\frac{1}{NT},
\end{aligned}$$

so that

$$a = o_{a.s.}\left(\frac{\left(\log N \log T\right)^{1+\epsilon}}{\sqrt{NT}}\right).$$

Also

$$\begin{aligned}
\frac{1}{N}\left\|\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right\| \ &\leq \ \frac{1}{N}\left\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2 + \frac{1}{N^2}\left[\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\right]^2 \\
&= \ o_{a.s.}\left(\frac{\left(\log N \log T\right)^{2+\epsilon}}{T}\right),
\end{aligned}$$

following the proof of Lemma C.9. Hence, we obtain that

$$II = o_{a.s.}\left(\frac{\left(\log N \log T\right)^{1+\epsilon}}{\sqrt{NT}}\right) + o_{a.s.}\left(\frac{\left(\log N \log T\right)^{2+\epsilon}}{T}\right).$$

The same logic yields that III is dominated by II. Turning to IV, it holds that

$$\frac{1}{N}\boldsymbol{\beta}'\mathbb{M}_{1_N}\overline{\mathbf{u}} = \frac{1}{N}\sum_{i=1}^{N}\beta_i\frac{1}{T}\sum_{t=1}^{T}u_{i,t} - \frac{1}{N}\sum_{i=1}^{N}\beta_i\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{i,t};$$

we have

$$
\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)^2\right]\\
= &\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{i=1}^{N}\beta_i\beta_j\sum_{t=1}^{T}\sum_{t=1}^{T}\mathbb{E}\left(u_{i,t}u_{j,s}\right)\\
\leq &\frac{1}{N^2T^2}\sum_{i=1}^{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{t=1}^{T}\left|\mathbb{E}\left(u_{i,t}u_{j,s}\right)\right| \leq c_0\frac{1}{NT},
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{i,t}\right)^2\right]\\
\leq &\frac{1}{N^2T^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\mathbb{E}\left(u_{i,t}u_{j,s}\right) \leq c_0\frac{1}{NT},
\end{aligned}
$$

by Assumption B.2*(ii)*, so that

(C.12) $$IV = o_{a.s.}\left(\frac{(\log N \log T)^{1+\epsilon}}{\sqrt{NT}}\right).$$

We conclude by only sketching the arguments for V; seeing as

$$
\begin{aligned}
&\frac{1}{N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\overline{\mathbf{u}}\\
= &\frac{1}{N}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\overline{\mathbf{u}} = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\\
\leq &\left(\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)^2\right)^{1/2},
\end{aligned}
$$

and noting

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)^2\right] \leq c_0T^{-1},$$

using Lemma C.9 it follows that

$$V = o_{a.s.}\left(\frac{(\log N \log T)^{2+\epsilon}}{T}\right).$$

The desired result now follows. □

Lemma C.12. *We assume that $\mathbb{H}_A$ of (1.3) holds, and that Assumptions 2.1-2.3 and B.1 and B.2 are satisfied. Then it holds that*

$$\widehat{\lambda} - \lambda = \frac{1}{N}\mathbf{S}_\beta^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha} + o_{a.s.}\left(T^{-1/2}(\log T)^{1+\epsilon}\right),$$

for all $\epsilon > 0$.

Proof. Considering

$$\begin{aligned}
\widehat{\lambda} &= \lambda + \frac{1}{N}\mathbf{S}_\beta^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha} + \overline{v} + \frac{1}{N}\left(\widehat{\mathbf{S}}_\beta^{-1} - \mathbf{S}_\beta^{-1}\right)\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha} \\
&\quad + \frac{1}{N}\widehat{\mathbf{S}}_\beta^{-1}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\boldsymbol{\alpha} + \frac{1}{N}\widehat{\mathbf{S}}_\beta^{-1}\widehat{\boldsymbol{\beta}}'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\lambda \\
&\quad + \frac{1}{N}\widehat{\mathbf{S}}_\beta^{-1}\widehat{\boldsymbol{\beta}}'\mathbb{M}_{1_N}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)\overline{v} + \frac{1}{N}\widehat{\mathbf{S}}_\beta^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\overline{\mathbf{u}} \\
&\quad + \frac{1}{N}\widehat{\mathbf{S}}_\beta^{-1}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\overline{\mathbf{u}} \\
&= \lambda + I + II + III + IV + V + VI + VII + VIII,
\end{aligned}$$

the only terms that require some analysis are III and IV. However, we already know that

$$\left\|\widehat{\mathbf{S}}_\beta^{-1} - \mathbf{S}_\beta^{-1}\right\| = o_{a.s.}\left(\frac{(\log N \log T)^{2+\epsilon}}{T}\right) + o_{a.s.}\left(\frac{(\log N \log T)^{1+\epsilon}}{\sqrt{NT}}\right);$$

further

$$\begin{aligned}
&\frac{1}{N}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\mathbb{M}_{1_N}\boldsymbol{\alpha} \\
= &\frac{1}{N}\sum_{i=1}^N\left(\alpha_i - \frac{1}{N}\sum_{i=1}^N\alpha_i\right)\left(\widehat{\beta}_i - \beta_i\right) = \frac{1}{N}\sum_{i=1}^N\widetilde{w}_i\frac{1}{T}\sum_{t=1}^T(v_t - \overline{v})\,u_{i,t},
\end{aligned}$$

which can be shown to be $o_{a.s.}\left((NT)^{-1/2}(\log N \log T)^{1+\epsilon}\right)$. □

Lemma C.13. *We assume that Assumptions 2.1-2.3, B.1 and B.2 are satisfied. Then it holds that*

$$\begin{aligned}
\liminf_{\min\{N,T\}\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{u}_{i,t}^{FM}\right)^2 &> 0, \\
\limsup_{\min\{N,T\}\to\infty} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{u}_{i,t}^{FM}\right)^2 &< \infty.
\end{aligned}$$

Proof. The proof is very similat to that of Lemma C.7, and we report only the main passages to save space. It holds that

$$\begin{aligned}
& \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{u}_{i,t}^{FM}\right)^2 \\
= & \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{i,t}^2 + \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\alpha}_i^{FM}-\alpha_i\right)^2 + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\beta}_i-\beta_i\right)^2 f_t^2 \\
& + \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\alpha}_i^{FM}-\alpha_i\right)u_{i,t} + \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\beta}_i-\beta_i\right)f_t u_{i,t} \\
& + \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{\alpha}_i^{FM}-\alpha_i\right)\left(\widehat{\beta}_i-\beta_i\right)f_t \\
= & I + II + III + IV + V + VI.
\end{aligned}$$

The rates of terms I, III and V are the same as in the proof of Lemma C.7. Note that $II \geq 0$; we derive an upper bound for it using (C.8). Noting that

$$\begin{aligned}
& \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\alpha}_i^{FM}-\alpha_i\right)^2 \\
\leq & c_0\left[\frac{1}{N}\sum_{i=1}^{N}\beta_i^2\overline{v}^2 + \frac{1}{N}\sum_{i=1}^{N}\overline{u}_i^2 + \left(\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)^2\right)\lambda^2\right.
\end{aligned}$$

$$\left(\frac{1}{N}\sum_{i=1}^{N}\beta_i^2\right)\left(\widehat{\lambda}-\lambda\right)^2+\left(\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)^2\right)\left(\widehat{\lambda}-\lambda\right)^2\Bigg],$$

the results above readily entail that $II = o_{a.s.}(1)$. Similarly

$$\begin{aligned}
IV &= \frac{2}{NT}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i u_{i,t}\right)\overline{v}+\frac{2}{N}\sum_{i=1}^{N}\overline{u}_i\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)-\frac{2}{NT}\left(\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\sum_{t=1}^{T}u_{i,t}\right)\lambda \\
&\quad -\frac{2}{NT}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i u_{i,t}\right)\left(\widehat{\lambda}-\lambda\right)-\frac{2}{NT}\left(\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\left(\sum_{t=1}^{T}u_{i,t}\right)\right)\left(\widehat{\lambda}-\lambda\right) \\
&= IV_a+IV_b+IV_c+IV_d+IV_e.
\end{aligned}$$

Since it is immediate to see that

$$\mathbb{E}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_i u_{i,t}\right)^2\leq c_0 NT,$$

we have

$$IV_a = o_{a.s.}\left(\frac{\left(\log N\log^2 T\right)^{1+\epsilon}}{N^{1/2}T}\right);$$

by the same arguments

$$\mathbb{E}\left(\frac{2}{N}\sum_{i=1}^{N}\overline{u}_i\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)\right)=\frac{2}{N}\sum_{i=1}^{N}\mathbb{E}\left(\overline{u}_i^2\right)\leq c_0T^{-1},$$

and therefore

$$IV_b = o_{a.s.}\left(\frac{(\log T)^{2+\epsilon}}{T}\right).$$

The other terms can be shown to be dominated by using the arguments above. Noting that

$$\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)\leq\left(\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\beta}_i-\beta_i\right)^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)^2\right)^{1/2},$$

it is easy to see that $VI = o_{a.s.}(1)$. The desired result now follows. □

Lemma C.14. *We assume that Assumptions 2.1-2.3, B.1 and B.2 are satisfied. Then it holds that, under* $\mathbb{H}_0$

$$\sum_{i=1}^{N} \psi_{i,NT}^{FM} = o_{a.s.}(1).$$

Proof. Consider the case $K = 1$, and recall (C.8), which under $\mathbb{H}_0$ becomes

$$\widehat{\alpha}_i^{FM} = \beta_i \overline{v} + \overline{u}_i - \left(\widehat{\beta}_i - \beta_i\right)\lambda - \beta_i\left(\widehat{\lambda} - \lambda\right) - \left(\widehat{\beta}_i - \beta_i\right)\left(\widehat{\lambda} - \lambda\right).$$

Hence we have

$$\begin{aligned}
& \sum_{i=1}^{N} \psi_{i,NT}^{FM} \\
= \ & \frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \sum_{i=1}^{N} \left|\widehat{\alpha}_i^{FM}\right|^{\nu/2} \\
\leq \ & c_0 \left[\frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \left(\sum_{i=1}^{N} |\beta_i|^{\nu/2}\right) |\overline{v}|^{\nu/2} + \frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \sum_{i=1}^{N} |\overline{u}_i|^{\nu/2} + \frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \left(\sum_{i=1}^{N} \left|\widehat{\beta}_i - \beta_i\right|^{\nu/2}\right) |\lambda|^{\nu/2} \right. \\
& \left. + \frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \left(\sum_{i=1}^{N} |\beta_i|^{\nu/2}\right) \left|\widehat{\lambda} - \lambda\right|^{\nu/2} + \frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \left(\sum_{i=1}^{N} \left|\widehat{\beta}_i - \beta_i\right|^{\nu/2}\right) \left|\widehat{\lambda} - \lambda\right|^{\nu/2} \right] \\
= \ & I + II + III + IV + V.
\end{aligned}$$

By Assumption B.1*(i)*, $\sum_{i=1}^{N} |\beta_i|^{\nu/2} = O(N)$, so that

$$I = o_{a.s.}\left(NT^{1/2-\nu/4} (\log T)^{(1+\epsilon)\nu/2}\right) = o_{a.s.}(1),$$

seeing as $\nu > 4$. Also, we have already shown in the proof of Lemma C.8 that $II = o_{a.s.}(1)$. Moreover, Lemma C.10 yields

$$\frac{T^{1/2}}{|\widehat{s}_{NT}^{FM}|^{\nu/2}} \left(\sum_{i=1}^{N} \left|\widehat{\beta}_i - \beta_i\right|^{\nu/2}\right) |\lambda|^{\nu/2} = o_{a.s.}\left(NT^{1/2}T^{-\nu/4} (\log N \log T)^{2+\epsilon}\right) = o_{a.s.}(1).$$

Finally, by Assumption B.1*(i)* and Lemma C.11 entail

$$IV = o_{a.s.}\left(NT^{1/2-\nu/4}\left(\log T\right)^{(1+\epsilon)\nu/2}\right) = o_{a.s.}\left(1\right).$$

Finally, it is easy to see that V is dominated by III. The desired result now follows by putting all together. □

C.4. **Lemmas for Section 4.2.** We now report a series of lemmas for the case, discussed in Section 4.2, of latent factors. As in the previous subsection, in the proofs we will assume $K = 1$ whenever possible.

Recall that - with reference to (4.6) - $\boldsymbol{\beta} = (\beta_1, ..., \beta_N)'$. Let $\widehat{\Phi}$ be the diagonal matrix containing, in descending order, the K largest eigenvalues of $\widehat{\boldsymbol{\Sigma}}_y$. Then, by definition

$$\widehat{\boldsymbol{\Sigma}}_y \widehat{\boldsymbol{\beta}}^{PC} = \widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi},$$

which implies the following expansion

$$\begin{aligned}
\text{(C.13)} \quad \widehat{\boldsymbol{\beta}}^{PC} &= \widehat{\boldsymbol{\Sigma}}_y \widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1} \\
&= \left[\frac{1}{NT}\sum_{t=1}^{T}\left(\boldsymbol{\beta}\widetilde{v}_t + \widetilde{\mathbf{u}}_t\right)\left(\boldsymbol{\beta}\widetilde{v}_t + \widetilde{\mathbf{u}}_t\right)'\right]\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1} \\
&= \boldsymbol{\beta}\mathbf{H} + \frac{1}{N}\boldsymbol{\beta}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1} + \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{v}_t'\right)\boldsymbol{\beta}'\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1} \\
&\quad + \left[\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{\mathbf{u}}_t'\right]\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1},
\end{aligned}$$

with the constraint $\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\widehat{\boldsymbol{\beta}}^{PC} = N \times \mathbb{I}_K$ and having defined

$$\mathbf{H} = \left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{v}_t'\right)\frac{\boldsymbol{\beta}'\widehat{\boldsymbol{\beta}}^{PC}}{N}\widehat{\Phi}^{-1}.$$

Similarly, we have also the unit-by-unit version of (C.13)

$$\widehat{\beta}_i^{PC} = \mathbf{H}'\beta_i + \widehat{\Phi}^{-1}\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{v}_t'\right)\beta_i + \widehat{\Phi}^{-1}\widehat{\boldsymbol{\beta}}^{PC\prime}\boldsymbol{\beta}\left(\frac{1}{NT}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{u}_{i,t}\right) + \widehat{\Phi}^{-1}\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{u}_{i,t}\right). \tag{C.14}$$

Then, considering

$$\widehat{\lambda}^{PC} = \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\bar{\mathbf{y}}\right),$$

we have

$$\begin{aligned}\widehat{\lambda}^{PC} &= \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\bar{\mathbf{y}}\right)\\ &= \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\alpha}\right) + \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}\right)\lambda\\ &\quad + \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}\right)\overline{v} + \left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\left(\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\overline{\mathbf{u}}\right).\end{aligned} \tag{C.15}$$

Lemma C.15. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that* $\left\|\widehat{\Phi}^{-1}\right\| = O_{a.s.}(1)$.

Proof. Let $\Phi = diag\{\Phi_1, ..., \Phi_K\}$ denote the diagonal matrix containing the K largest eigenvalues of $\boldsymbol{\beta}\mathbb{E}(\widetilde{v}_t\widetilde{v}_t')\boldsymbol{\beta}'/N$. Then, it is immediate to verify that the assumptions of Lemma 2.2 in Trapani (2018) hold, and

$$\widehat{\Phi}_j = \Phi_j + o_{a.s.}(1),$$

for all $1 \leq j \leq K$. Seeing as, using the multiplicative Weyl's inequality (theorem 7 in Merikoski and Kumar 2004), and Assumptions B.3 and B.4

$$\Phi_K \geq \rho_K(\boldsymbol{\beta}'\boldsymbol{\beta}/N)\,\rho_{\min}(\mathbb{E}(\widetilde{v}_t\widetilde{v}_t')) \geq c_0 > 0,$$

where $\rho_k(A)$ denotes the k-th largest eigenvalue of matrix A, the desired result follows. □

Lemma C.16. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that*

$$\frac{1}{N}\left\|\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right\|^2=o_{a.s.}\left(\frac{(\log N\log T)^{2+\epsilon}}{T}\right)+O\left(\frac{1}{N^2}\right).$$

Proof. By (C.13), we have

$$\begin{aligned}
&\left\|\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right\|^2\\
\leq\ & c_0\left(\frac{1}{N^2}\left\|\boldsymbol{\beta}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2+\frac{1}{N^2}\left\|\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{v}_t'\right)\boldsymbol{\beta}'\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2\right.\\
&\left.+\left\|\left(\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{\mathbf{u}}_t'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2\right)\\
=\ & I+I'+II.
\end{aligned}$$

It holds that

$$I\leq\frac{1}{N^2}\|\boldsymbol{\beta}\|^2\left\|\widehat{\boldsymbol{\beta}}^{PC}\right\|^2\left\|\widehat{\Phi}^{-1}\right\|^2\left\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right\|^2\leq c_0\left\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right\|^2,$$

on account of Assumption B.3 and Lemma C.15. Also note that

$$\left\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right\|^2\leq\left\|\frac{1}{T}\sum_{t=1}^{T}v_t\mathbf{u}_t'\right\|^2+\|\overline{v}\|^2\|\overline{\mathbf{u}}\|^2.$$

It holds that

$$\left\|\frac{1}{T}\sum_{t=1}^{T}v_t\mathbf{u}_t'\right\|^2=\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}v_tu_{i,t}\right)^2,$$

and

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}v_tu_{i,t}\right)^2\right]\\
=\ &\frac{1}{T^2}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\mathbb{E}\left(v_tv_su_{i,t}u_{i,s}\right)\leq\frac{1}{T^2}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\mathbb{E}\left(v_tv_s\right)\mathbb{E}\left(u_{i,t}u_{i,s}\right)
\end{aligned}$$

$$\leq \frac{1}{T^2}\mathbb{E}\left(v_0^2\right)\sum_{i=1}^{N}\sum_{t,s=1}^{T}\left|\mathbb{E}\left(u_{i,t}u_{i,s}\right)\right| \leq c_0\frac{N}{T}.$$

Also, we know that $\|\overline{v}\| = o_{a.s.}\left(T^{-1/2}\left(\log T\right)^{1+\epsilon}\right)$, and

$$\|\overline{\mathbf{u}}\|^2 = \sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)^2 = \frac{1}{T^2}\sum_{i=1}^{N}\sum_{t,s=1}^{T}u_{i,t}u_{i,s},$$

with

$$\mathbb{E}\left(\frac{1}{T^2}\sum_{i=1}^{N}\sum_{t,s=1}^{T}u_{i,t}u_{i,s}\right) \leq \frac{1}{T^2}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\left|\mathbb{E}\left(u_{i,t}u_{i,s}\right)\right| \leq c_0\frac{N}{T},$$

so that $\|\overline{\mathbf{u}}\| = o_{a.s.}\left(N^{1/2}T^{-1/2}\left(\log T\log N\right)^{1+\epsilon}\right)$. Putting all together, we ultimately receive

$$I = o_{a.s.}\left(\frac{N}{T}\left(\log N\log T\right)^{2+\epsilon}\right).$$

Turning to II, note

$$\begin{aligned}
&\left\|\left(\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{\mathbf{u}}_t'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2 \\
\leq\ & c_0\left\|\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t' - \mathbb{E}\left(\mathbf{u}_0\mathbf{u}_0'\right)\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2 + c_0\left\|\frac{1}{N}\mathbb{E}\left(\mathbf{u}_0\mathbf{u}_0'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2 + c_0\left\|\frac{1}{N}\overline{\mathbf{u}}\overline{\mathbf{u}}'\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\right\|^2 \\
=\ & II_a + II_b + II_c.
\end{aligned}$$

We already know from the above that $\|\overline{\mathbf{u}}\|^2 = o_{a.s.}\left(NT^{-1}\left(\log N\log T\right)^{2+\epsilon}\right)$, which readily yields

$$II_c = o_{a.s.}\left(NT^{-2}\left(\log N\log T\right)^{4+\epsilon}\right).$$

Also

$$\left\|\mathbb{E}\left(\mathbf{u}_0\mathbf{u}_0'\right)\right\|^2 \leq \left\|\mathbb{E}\left(\mathbf{u}_0\mathbf{u}_0'\right)\right\|_1^2 = \left(\max_{1\leq i\leq N}\sum_{j=1}^{N}\left|\mathbb{E}\left(u_{i,0}u_{j,0}\right)\right|\right)^2 \leq c_0,$$

so that

$$II_b = O\left(\frac{1}{N}\right).$$

Finally

$$\left\| \frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t' - \mathbb{E}\left(\mathbf{u}_0\mathbf{u}_0'\right) \right\|_F^2 = \frac{1}{N^2T^2}\sum_{i,j=1}^{N}\left(\sum_{t,s=1}^{T} Cov\left(u_{i,t}u_{j,t}, u_{i,s}u_{j,s}\right)\right) \leq c_0\frac{1}{T},$$

and therefore

$$II_a = o_{a.s.}\left(\frac{N}{T}\left(\log N \log T\right)^{2+\epsilon}\right).$$

The desired result obtains by putting all together. □

Lemma C.17. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that* $\|\mathbf{H}\| = O_{a.s.}(1)$, *and* $\|\mathbf{H}^{-1}\| = O_{a.s.}(1)$.

Proof. Note

$$\|\mathbf{H}\| \leq \left\| \frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{v}_t' \right\| \frac{\|\boldsymbol{\beta}\| \left\|\widehat{\boldsymbol{\beta}}^{PC}\right\|}{N} \left\|\widehat{\Phi}^{-1}\right\|.$$

Standard arguments yield $\left\|T^{-1}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{v}_t'\right\| = O_{a.s.}(1)$; further, $\left\|\widehat{\boldsymbol{\beta}}^{PC}\right\| = N^{1/2}$ by construction, and $\|\boldsymbol{\beta}\| = N^{1/2}$ by Assumption B.3. The desired result now follows by Lemma C.15. As far as the second part of the lemma is concerned, recall the identification restriction $\boldsymbol{\beta}'\boldsymbol{\beta} = N\mathbb{I}_K$, and that, by construction $\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\widehat{\boldsymbol{\beta}}^{PC} = N\mathbb{I}_K$. Then we have

$$\begin{aligned}
\mathbb{I}_K &= \frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\widehat{\boldsymbol{\beta}}^{PC} \\
&= \frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H} + \boldsymbol{\beta}\mathbf{H}\right)'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H} + \boldsymbol{\beta}\mathbf{H}\right) \\
&= \mathbf{H}'\left(\frac{1}{N}\boldsymbol{\beta}'\boldsymbol{\beta}\right)\mathbf{H} + \mathbf{H}'\frac{1}{N}\boldsymbol{\beta}'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right) + \left[\mathbf{H}'\frac{1}{N}\boldsymbol{\beta}'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\right]' \\
&\quad + \frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right) \\
&= \mathbf{H}'\mathbf{H} + \mathbf{H}'\frac{1}{N}\boldsymbol{\beta}'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right) + \left[\mathbf{H}'\frac{1}{N}\boldsymbol{\beta}'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\right]' \\
&\quad + \frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)'\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right).
\end{aligned}$$

Using Lemma C.16 repeatedly, it is easy to see that

$$\mathbf{H}'\frac{1}{N}\boldsymbol{\beta}'\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)+\left[\mathbf{H}'\frac{1}{N}\boldsymbol{\beta}'\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)\right]'+\frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)'\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)=o_{a.s.}(1),$$

and therefore

$$\mathbf{H}'\mathbf{H}=\mathbb{I}_K+o_{a.s.}(1),$$

so that $\mathbf{H}^{-1}=\mathbf{H}'$. This proves that $\mathbf{H}$ is invertible. □

Lemma C.18. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that, under* $\mathbb{H}_0$

$$\widehat{\lambda}^{PC}-\mathbf{H}^{-1}\lambda=o_{a.s.}\left(\frac{(\log T\log N)^{1+\epsilon}}{T^{1/2}}\right)+O\left(\frac{1}{N}\right),$$

for all $\epsilon>0$.

Proof. By (C.15), under the null it holds that

$$\begin{aligned}\widehat{\lambda}^{PC} &= \left(\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}}{N}\right)^{-1}\left(\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}}{N}\right)\lambda \\ &\quad+\left(\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}}{N}\right)^{-1}\left(\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}}{N}\right)\overline{v} \\ &\quad+\left(\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}}{N}\right)^{-1}\left(\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N}\right) \\ &= I+II+III.\end{aligned}\tag{C.16}$$

Note, to begin with, that

$$\begin{aligned}&\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC} \\ =\ &\mathbf{H}'\mathbf{S}_\beta\mathbf{H}+\mathbf{H}'\boldsymbol{\beta}'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)+\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)'\mathbb{M}_{1_N}\boldsymbol{\beta}\mathbf{H} \\ &+\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)'\mathbb{M}_{1_N}\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right);\end{aligned}$$

thus, using Lemma C.16, it is easy to see that

$$\frac{1}{N}\left\|\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}-\mathbf{H}'\mathbf{S}_{\beta}\mathbf{H}\right\|=o_{a.s.}\left(\frac{(\log N\log T)^{1+\epsilon}}{T^{1/2}}\right)+O\left(\frac{1}{N}\right).$$

Assumption B.3*(iii)* and Lemma C.17 guarantee that $\mathbf{H}'\mathbf{S}_{\beta}\mathbf{H}$ is invertible, and therefore we may write

$$\left\|\left(\frac{1}{N}\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\right)^{-1}\right\|=O_{a.s.}(1). \tag{C.17}$$

Note now that, using (C.13) "in reverse", viz.

$$\begin{aligned}\boldsymbol{\beta} \;=\; & \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}-\frac{1}{N}\boldsymbol{\beta}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\mathbf{H}^{-1}\\ & -\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{v}_t'\right)\boldsymbol{\beta}'\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\mathbf{H}^{-1}\\ & -\left[\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{\mathbf{u}}_t'\right]\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\mathbf{H}^{-1},\end{aligned} \tag{C.18}$$

we have

$$\begin{aligned}& \widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}\\ = \;& \widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}-\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\frac{1}{N}\boldsymbol{\beta}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{\mathbf{u}}_t'\right)\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\mathbf{H}^{-1}\\ & -\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{v}_t'\right)\boldsymbol{\beta}'\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\mathbf{H}^{-1}-\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\left[\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{\mathbf{u}}_t'\right]\widehat{\boldsymbol{\beta}}^{PC}\widehat{\Phi}^{-1}\mathbf{H}^{-1}.\end{aligned}$$

Following exactly the same steps as in the proof of Lemma C.16, it can be shown that

$$\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}}{N}=\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\widehat{\boldsymbol{\beta}}^{PC}}{N}\mathbf{H}^{-1}+o_{a.s.}\left(\frac{(\log N\log T)^{1+\epsilon}}{T^{1/2}}\right)+O\left(\frac{1}{N}\right), \tag{C.19}$$

so that, in (C.16)

$$I = \mathbf{H}^{-1}\lambda + o_{a.s.}\left(\frac{(\log N \log T)^{1+\epsilon}}{T^{1/2}}\right) + O\left(\frac{1}{N}\right).$$

Indeed, by the same token it also holds that

$$\text{(C.20)} \qquad \frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\boldsymbol{\beta}}{N} = \mathbf{H}'\mathbf{S}_\beta + o_{a.s.}\left(\frac{(\log N \log T)^{1+\epsilon}}{T^{1/2}}\right) + O\left(\frac{1}{N}\right) = O_{a.s.}(1).$$

Recalling that $\overline{v} = o_{a.s.}\left(T^{-1/2}(\log T)^{1+\epsilon}\right)$, using (C.17) and (C.20) it follows that

$$II = o_{a.s.}\left(T^{-1/2}(\log T)^{1+\epsilon}\right).$$

Finally, we study

$$\frac{\widehat{\boldsymbol{\beta}}^{PC\prime}\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N} = \mathbf{H}'\frac{\boldsymbol{\beta}\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N} + \frac{\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)'\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N}.$$

We already know from the proof of (C.12) that

$$\left\|\frac{\boldsymbol{\beta}'\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N}\right\| = o_{a.s.}\left(\frac{(\log N \log T)^{1+\epsilon}}{\sqrt{NT}}\right).$$

Also, note that

$$\left|\frac{\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)'\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N}\right| \leq \left\|\frac{\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}}{N^{1/2}}\right\|\left\|\frac{\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N^{1/2}}\right\|;$$

we know from Lemma C.16 that

$$\left\|\frac{\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}}{N^{1/2}}\right\| = o_{a.s.}\left(\frac{(\log N \log T)^{1+\epsilon}}{T^{1/2}}\right) + O\left(\frac{1}{N}\right),$$

and, by standard passages

$$\left\|\frac{\mathbb{M}_{1_N}\overline{\mathbf{u}}}{N^{1/2}}\right\| \leq \left\|\frac{\overline{\mathbf{u}}}{N^{1/2}}\right\| + \frac{1}{N}N^{1/2}\left|\frac{\sum_{i=1}^N \overline{u}_i}{N^{1/2}}\right|$$

$$= \frac{1}{N^{1/2}}\left(\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}u_{i,t}\right)^2\right)^{1/2} + \left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{i,t}\right|$$

$$= o_{a.s.}\left(\frac{(\log N\log T)^{1+\epsilon}}{\sqrt{T}}\right) + o_{a.s.}\left(\frac{(\log N\log T)^{1+\epsilon}}{\sqrt{NT}}\right).$$

The final result follows by putting all together. □

Lemma C.19. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that, under* $\mathbb{H}_A$

$$\widehat{\lambda}^{PC} - \mathbf{H}^{-1}\lambda = \frac{1}{N}\left(\mathbf{H}'\mathbf{S}_\beta\mathbf{H}\right)^{-1}\mathbf{H}'\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha} + o_{a.s.}\left(\frac{(\log T\log N)^{1+\epsilon}}{T^{1/2}}\right) + O\left(\frac{1}{N}\right),$$

for all $\epsilon > 0$.

Proof. The proof follows by combining the arguments in Lemmas C.12 and C.18. □

Lemma C.20. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that*

$$\sum_{i=1}^{N}\left\|\widehat{\beta}_i^{PC} - \mathbf{H}'\beta_i\right\|^{\nu/2} = o_{a.s.}\left(NT^{-\nu/4}(\log N\log T)^{(1+\epsilon)\nu/2}\right) + O\left(N^{1-\nu/2}\right),$$

for all $\epsilon > 0$.

Proof. Using (C.14), it holds that

$$\sum_{i=1}^{N}\left\|\widehat{\beta}_i^{PC} - \mathbf{H}'\beta_i\right\|^{\nu/2}$$

$$\leq c_0\left\|\widehat{\Phi}^{-1}\right\|^{\nu/2}\left\|\widehat{\boldsymbol{\beta}}^{PC}\right\|^{\nu/2}\left\|\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\mathbf{u}}_t\widetilde{v}_t'\right\|^{\nu/2}\sum_{i=1}^{N}\|\beta_i\|^{\nu/2}$$

$$+c_0\left\|\widehat{\Phi}^{-1}\right\|^{\nu/2}\left\|\widehat{\boldsymbol{\beta}}^{PC}\right\|^{\nu/2}\|\boldsymbol{\beta}\|^{\nu/2}\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\widetilde{v}_t\widetilde{u}_{i,t}\right\|^{\nu/2}$$

$$+c_0 \left\| \widehat{\Phi}^{-1} \right\|^{\nu/2} \sum_{i=1}^{N} \left\| \frac{1}{NT} \sum_{t=1}^{T} \left(\widehat{\boldsymbol{\beta}}^{PC} \right)' \widetilde{\mathbf{u}}_t \widetilde{u}_{i,t} \right\|^{\nu/2}$$
$$= \quad I + II + III.$$

We begin by noting that, by Lemma C.15, $\left\| \widehat{\Phi}^{-1} \right\|^{\nu/2} = O_{a.s.}(1)$; further, $\left\| \widehat{\boldsymbol{\beta}}^{PC} \right\|^{\nu/2} = c_0 N^{\nu/4}$ by construction, and $\|\boldsymbol{\beta}\|^{\nu/2} = c_0 N^{\nu/4}$ by the identification restriction $\boldsymbol{\beta}'\boldsymbol{\beta} = N\mathbb{I}_K$. Finally, Assumption B.3 entails $\sum_{i=1}^{N} \|\beta_i\|^{\nu/2} = O(N)$. Consider now

$$\left\| \frac{1}{NT} \sum_{t=1}^{T} \widetilde{\mathbf{u}}_t \widetilde{v}_t' \right\|^{\nu/2} \leq \left\| \frac{1}{NT} \sum_{t=1}^{T} \mathbf{u}_t v_t' \right\|^{\nu/2} + \left\| \frac{1}{N} \overline{\mathbf{u}} \overline{v} \right\|^{\nu/2}.$$

We know from the above that

$$\left\| \frac{1}{NT} \sum_{t=1}^{T} \mathbf{u}_t v_t' \right\| = o_{a.s.} \left(N^{-1/2} T^{-1/2} (\log T \log N)^{1+\epsilon} \right),$$

and therefore

$$\left\| \frac{1}{NT} \sum_{t=1}^{T} \mathbf{u}_t v_t' \right\|^{\nu/2} = o_{a.s.} \left(T^{-\nu/4} N^{-\nu/4} (\log T \log N)^{(1+\epsilon)\nu/2} \right).$$

Also, seeing as (as shown above) it holds that $\overline{v} = o_{a.s.} \left(T^{-1/2} (\log T)^{1+\epsilon} \right)$ and $\|\overline{\mathbf{u}}\| = o_{a.s.} \left(N^{1/2} T^{-1/2} (\log T \log N)^{1+\epsilon} \right)$, we have

$$\left\| \frac{1}{N} \overline{\mathbf{u}} \overline{v} \right\|^{\nu/2} = o_{a.s.} \left(T^{-\nu/2} N^{-\nu/4} (\log T \log N)^{(1+\epsilon)\nu/2} \right).$$

Hence, combining all the results above, it follows that

$$I = o_{a.s.} \left(N T^{-\nu/4} (\log T \log N)^{(1+\epsilon)\nu/2} \right).$$

Turning to II, note that

$$\sum_{i=1}^{N} \left\| \frac{1}{NT} \sum_{t=1}^{T} \widetilde{v}_t \widetilde{u}_{i,t} \right\|^{\nu/2} \leq \sum_{i=1}^{N} \left\| \frac{1}{NT} \sum_{t=1}^{T} v_t u_{i,t} \right\|^{\nu/2} + \sum_{i=1}^{N} \left\| \frac{1}{N} \overline{v} \overline{u}_i \right\|^{\nu/2}$$

with

$$\sum_{i=1}^{N}\mathbb{E}\left\|\frac{1}{NT}\sum_{t=1}^{T}v_t u_{i,t}\right\|^{\nu/2}=c_0 N^{1-\nu/2}T^{-\nu/4},$$

using Lemma C.2. Hence

$$\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}v_t u_{i,t}\right\|^{\nu/2}=o_{a.s.}\left(N^{1-\nu/2}T^{-\nu/4}\left(\log N\log T\right)^{2+\epsilon}\right),$$

and

$$II=o_{a.s.}\left(NT^{-\nu/4}\left(\log N\log T\right)^{2+\epsilon}\right).$$

Also, by exactly the same passages

$$\mathbb{E}\sum_{i=1}^{N}\left\|\frac{1}{N}\overline{u}_i\right\|^{\nu/2}=c_0 N^{1-\nu/2}T^{-\nu/4},$$

and therefore

$$II=o_{a.s.}\left(NT^{-\nu/4}\left(\log N\log T\right)^{2+\epsilon}\right),$$

so that I dominated II. Finally, considering III it holds that

$$\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\widetilde{\mathbf{u}}_t\widetilde{u}_{i,t}\right\|^{\nu/2}\leq\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\mathbf{u}_t u_{i,t}\right\|^{\nu/2}+\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\overline{\mathbf{u}u}_i\right\|^{\nu/2}.$$

Consider

$$\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}\right)'\widetilde{\mathbf{u}}_t u_{i,t}\right\|^{\nu/2}\leq\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)'\mathbf{u}_t u_{i,t}\right\|^{\nu/2}+\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\mathbf{H}'\boldsymbol{\beta}'\mathbf{u}_t u_{i,t}\right\|^{\nu/2}.$$

We have

$$\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)'\mathbf{u}_t u_{i,t}\right\|^{\nu/2}\leq\left\|\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right\|^{\nu/2}\sum_{i=1}^{N}\left\|\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t u_{i,t}\right\|^{\nu/2}$$

with

$$\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_{t}u_{i,t}\right\Vert ^{\nu/2}\leq\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\left(\mathbf{u}_{t}u_{i,t}-\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right)\right\Vert ^{\nu/2}+\sum_{i=1}^{N}\left\Vert \frac{1}{N}\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert ^{\nu/2}.$$

It holds that

$$\begin{aligned}
& \left(NT\right)^{-\nu/2}\sum_{i=1}^{N}\mathbb{E}\left\Vert \sum_{t=1}^{T}\left(\mathbf{u}_{t}u_{i,t}-\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right)\right\Vert ^{\nu/2}\\
= & \left(NT\right)^{-\nu/2}\sum_{i=1}^{N}\mathbb{E}\left(\sum_{j=1}^{N}\left(\sum_{t=1}^{T}\left(u_{j,t}u_{i,t}-\mathbb{E}\left(u_{j,0}u_{i,0}\right)\right)\right)^{2}\right)^{\nu/4}\\
\leq & \ c_{0}\left(NT\right)^{-\nu/2}N^{\nu/4-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\sum_{t=1}^{T}\left(u_{j,t}u_{i,t}-\mathbb{E}\left(u_{j,0}u_{i,0}\right)\right)\right)^{\nu/2}\leq c_{0}T^{-\nu/4}N^{1-\nu/4},
\end{aligned}$$

and, since $\left\Vert \mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert \leq\left\Vert \mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert _{1}\leq c_{0}$

$$\sum_{i=1}^{N}\left\Vert \frac{1}{N}\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert ^{\nu/2}=O\left(N^{1-\nu/2}\right).$$

Combining these results with Lemma C.16, it follows that

$$\begin{aligned}
& \sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}-\boldsymbol{\beta}\mathbf{H}\right)'\mathbf{u}_{t}u_{i,t}\right\Vert ^{\nu/2}\\
= & \ o_{a.s.}\left[\left(N^{\nu/4}T^{-\nu/4}\left(\log N\log T\right)^{(1+\epsilon)\nu/2}+N^{-\nu/4}\right)\left(\left(T^{-\nu/4}N^{1-\nu/4}\left(\log N\log T\right)^{1+\epsilon}\right)+N^{1-\nu/2}\right)\right].
\end{aligned}$$

Also note that

$$\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\mathbf{H}'\boldsymbol{\beta}'\mathbf{u}_{t}u_{i,t}\right\Vert ^{\nu/2}=O_{a.s.}\left(1\right)\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\beta}'\mathbf{u}_{t}u_{i,t}\right\Vert ^{\nu/2},$$

and

$$\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\beta}'\mathbf{u}_{t}u_{i,t}\right\Vert ^{\nu/2}\leq\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\beta}'\left(\mathbf{u}_{t}u_{i,t}-\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right)\right\Vert ^{\nu/2}+\sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\beta}'\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert ^{\nu/2}$$

It is immediate to see that, by Assumption B.5*(vi)*

$$\sum_{i=1}^{N}\mathbb{E}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\beta}^{\prime}\left(\mathbf{u}_{t}u_{i,t}-\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right)\right\Vert^{\nu/2}\leq c_{0}N\left(NT\right)^{-\nu/4}.$$

Also

$$\begin{aligned}
& \sum_{i=1}^{N}\left\Vert \frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\beta}^{\prime}\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert^{\nu/2} \\
= & \sum_{i=1}^{N}\left\Vert \frac{1}{N}\boldsymbol{\beta}^{\prime}\mathbb{E}\left(\mathbf{u}_{0}u_{i,0}\right)\right\Vert^{\nu/2}=\sum_{i=1}^{N}\left\Vert \frac{1}{N}\sum_{j=1}^{N}\beta_{j}\mathbb{E}\left(u_{j,0}u_{i,0}\right)\right\Vert^{\nu/2} \\
\leq & \sum_{i=1}^{N}\left\Vert \max_{1\leq j\leq N}\left|\beta_{j}\right|\frac{1}{N}\sum_{j=1}^{N}\left|\mathbb{E}\left(u_{j,0}u_{i,0}\right)\right|\right\Vert^{\nu/2}\leq c_{0}N^{1-\nu/2},
\end{aligned}$$

by virtue of Assumption B.5*(iv)*. By the same logic as above, it can be shown that the term $\sum_{i=1}^{N}\left\Vert (NT)^{-1}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\beta}}^{PC}\right)^{\prime}\overline{\mathbf{u}u}_{i}\right\Vert^{\nu/2}$ is dominated. Putting all together, the final result obtains. □

Lemma C.21. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that*

$$\begin{aligned}
\liminf_{\min\{N,T\}\rightarrow\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{u}_{i,t}^{PC}\right)^{2} & > & 0, \\
\limsup_{\min\{N,T\}\rightarrow\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{u}_{i,t}^{PC}\right)^{2} & < & \infty.
\end{aligned}$$

Proof. We let $K=1$ for simplicity and without loss of generality. Let $\widehat{\mathbf{u}}_{t}^{PC}=\left(\widehat{u}_{1,t}^{PC},...,\widehat{u}_{N,t}^{PC}\right)^{\prime}$. It holds that

$$\begin{aligned}
& \widehat{\mathbf{u}}_{t}^{PC} \\
= & \widetilde{\mathbf{y}}_{t}-\widehat{\boldsymbol{\beta}}^{PC}\widehat{f}_{t}^{PC}=\boldsymbol{\beta}\widetilde{v}_{t}+\widetilde{\mathbf{u}}_{t}-\widehat{\boldsymbol{\beta}}^{PC}\widehat{f}_{t}^{PC}
\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{\beta}\widetilde{v}_t + \widetilde{\mathbf{u}}_t - \frac{1}{N}\widehat{\boldsymbol{\beta}}^{PC}\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\boldsymbol{\beta}\widetilde{v}_t + \widetilde{\mathbf{u}}_t\right) \\
&= \boldsymbol{\beta}\widetilde{v}_t + \widetilde{\mathbf{u}}_t - \frac{1}{N}\left[\boldsymbol{\beta}\mathbf{H} + \left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\right]\widehat{\boldsymbol{\beta}}^{PC\prime}\left[\widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1} + \boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}\right]\widetilde{v}_t \\
&\quad - \frac{1}{N}\widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}'\boldsymbol{\beta}'\widetilde{\mathbf{u}}_t - \frac{1}{N}\widehat{\boldsymbol{\beta}}^{PC}\left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right]'\widetilde{\mathbf{u}}_t \\
&= \widetilde{\mathbf{u}}_t - \left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\mathbf{H}^{-1}\widetilde{v}_t - \frac{1}{N}\boldsymbol{\beta}\mathbf{H}\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}\right)\widetilde{v}_t \\
&\quad - \frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}\right)\widetilde{v}_t - \frac{1}{N}\widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}'\boldsymbol{\beta}'\widetilde{\mathbf{u}}_t \\
&\quad - \frac{1}{N}\widehat{\boldsymbol{\beta}}^{PC}\left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right]'\widetilde{\mathbf{u}}_t.
\end{aligned}
$$

We bound the following terms:

$$
\begin{aligned}
&\frac{1}{NT}\sum_{t=1}^{T}\left\|\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\mathbf{H}^{-1}\widetilde{v}_t\right\|^2 \\
\leq\ & c_0\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t^2\right)\frac{\left\|\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right\|^2}{N} = o_{a.s.}\left(\frac{(\log N \log T)^{2+\epsilon}}{T}\right) + O\left(\frac{1}{N^2}\right),
\end{aligned}
$$

by Lemma C.16;

$$
\begin{aligned}
&\frac{1}{NT}\sum_{t=1}^{T}\left\|\frac{1}{N}\boldsymbol{\beta}\mathbf{H}\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}\right)\widetilde{v}_t\right\|^2 \\
\leq\ & c_0\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_t^2\right)\frac{\|\boldsymbol{\beta}\|^2}{N}\frac{\left\|\widehat{\boldsymbol{\beta}}^{PC\prime}\right\|^2}{N}\frac{\left\|\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}\right\|^2}{N} \\
=\ & o_{a.s.}\left(\frac{(\log N \log T)^{2+\epsilon}}{T}\right) + O\left(\frac{1}{N^2}\right),
\end{aligned}
$$

again using Lemma C.16;

$$
\frac{1}{NT}\sum_{t=1}^{T}\left\|\frac{1}{N}\left(\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta}\mathbf{H}\right)\widehat{\boldsymbol{\beta}}^{PC\prime}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC}\mathbf{H}^{-1}\right)\widetilde{v}_t\right\|^2
$$

$$\leq \quad c_0 \left(\frac{1}{T} \sum_{t=1}^{T} \widetilde{v}_t^2 \right) \left(\frac{\left\| \boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC} \mathbf{H}^{-1} \right\|^2}{N} \right)^2 \frac{\left\| \widehat{\boldsymbol{\beta}}^{PC\prime} \right\|^2}{N},$$

and therefore it is dominated by the previous terms;

$$\mathbb{E} \frac{1}{NT} \sum_{t=1}^{T} \left\| \frac{1}{N} \widehat{\boldsymbol{\beta}}^{PC} \mathbf{H}' \boldsymbol{\beta}' \widetilde{\mathbf{u}}_t \right\|^2 \leq c_0 \frac{1}{N^2 T} \sum_{t=1}^{T} \mathbb{E} \left(\sum_{j=1}^{N} \beta_j \widetilde{u}_{j,t} \right)^2 \leq c_1 N^{-1},$$

so that this term is bounded by $o_{a.s.} \left(N^{-1} \left(\log T \log^2 N \right)^{2+\epsilon} \right)$;

$$\begin{aligned} & \frac{1}{NT} \sum_{t=1}^{T} \left\| \frac{1}{N} \widehat{\boldsymbol{\beta}}^{PC} \left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta} \mathbf{H} \right]' \widetilde{\mathbf{u}}_t \right\|^2 \\ = \quad & \frac{1}{N^3 T} \sum_{t=1}^{T} \widetilde{\mathbf{u}}_t' \left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta} \mathbf{H} \right] \widehat{\boldsymbol{\beta}}^{PC\prime} \widehat{\boldsymbol{\beta}}^{PC} \left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta} \mathbf{H} \right]' \widetilde{\mathbf{u}}_t \\ = \quad & \frac{1}{N^2 T} \sum_{t=1}^{T} \widetilde{\mathbf{u}}_t' \left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta} \mathbf{H} \right] \left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta} \mathbf{H} \right]' \widetilde{\mathbf{u}}_t \\ \leq \quad & \left(\frac{1}{NT} \sum_{t=1}^{T} \left\| \widetilde{\mathbf{u}}_t \right\|^2 \right) \frac{\left\| \boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}^{PC} \mathbf{H}^{-1} \right\|^2}{N}; \end{aligned}$$

it is not hard to see that

$$\frac{1}{NT} \sum_{t=1}^{T} \left\| \widetilde{\mathbf{u}}_t \right\|^2 = \frac{1}{NT} \sum_{t=1}^{T} \mathbb{E} \left(\sum_{i=1}^{N} \widetilde{u}_{i,t}^2 \right) = O_{a.s.}(1),$$

and therefore

$$\frac{1}{NT} \sum_{t=1}^{T} \left\| \frac{1}{N} \widehat{\boldsymbol{\beta}}^{PC} \left[\widehat{\boldsymbol{\beta}}^{PC} - \boldsymbol{\beta} \mathbf{H} \right]' \widetilde{\mathbf{u}}_t \right\|^2 = o_{a.s.} \left(\frac{\left(\log N \log T \right)^{2+\epsilon}}{T} \right) + o_{a.s.} \left(\frac{\left(\log T \log^2 N \right)^{2+\epsilon}}{N} \right).$$

After some algebra and repeated use of the Cauchy-Schwartz inequality, the above entails that

$$\widehat{s}_{NT}^{PC}=\frac{1}{NT}\sum_{t=1}^{T}\left\|\widehat{\mathbf{u}}_{t}^{PC}\right\|^{2}=\frac{1}{NT}\sum_{t=1}^{T}\left\|\widetilde{\mathbf{u}}_{t}\right\|^{2}+o_{a.s.}\left(\frac{\left(\log N\log T\right)^{2+\epsilon}}{T}\right)+O\left(\frac{1}{N^{2}}\right).$$

From hereon, the proof follows by similar arguments as above. □

Lemma C.22. *We assume that Assumptions 2.1-2.3, and B.3-B.6 are satisfied. Then it holds that*

$$\sum_{i=1}^{N}\psi_{i,NT}^{PC}=o_{a.s.}\left(1\right).$$

Proof. The proof is essentially the same as the proof of Lemmas C.8 and C.14, and we simply discuss the different parts.

Under both $\mathbb{H}_0$ and $\mathbb{H}_A$, it holds that

(C.21)

$$\widehat{\alpha}_{i}^{PC}=\alpha_{i}+\beta_{i}'\overline{v}+\overline{u}_{i}-\left(\widehat{\beta}_{i}-\mathbf{H}'\beta_{i}\right)'\lambda-\beta_{i}'\mathbf{H}\left(\widehat{\lambda}-\mathbf{H}^{-1}\lambda\right)-\left(\widehat{\beta}_{i}-\mathbf{H}'\beta_{i}\right)'\left(\widehat{\lambda}-\mathbf{H}^{-1}\lambda\right).$$

Under $\mathbb{H}_0$ we have

$$\begin{aligned}
&\sum_{i=1}^{N}\psi_{i,NT}^{PC}\\
=\;&\frac{C_{NT}^{1/2}}{\left|\widehat{s}_{NT}^{PC}\right|^{\nu/2}}\sum_{i=1}^{N}\left|\widehat{\alpha}_{i}^{PC}\right|^{\nu/2}\\
\leq\;&c_{0}\left[\frac{C_{NT}^{1/2}}{\left|\widehat{s}_{NT}^{PC}\right|^{\nu/2}}\left(\sum_{i=1}^{N}\left\|\beta_{i}\right\|^{\nu/2}\right)\left\|\overline{v}\right\|^{\nu/2}+\frac{C_{NT}^{1/2}}{\left|\widehat{s}_{NT}^{PC}\right|^{\nu/2}}\sum_{i=1}^{N}\left|\overline{u}_{i}\right|^{\nu/2}+\frac{C_{NT}^{1/2}}{\left|\widehat{s}_{NT}^{PC}\right|^{\nu/2}}\left(\sum_{i=1}^{N}\left\|\widehat{\beta}_{i}-\mathbf{H}'\beta_{i}\right\|^{\nu/2}\right)\left\|\lambda\right\|^{\nu/2}\right.\\
&\left.+\frac{C_{NT}^{1/2}}{\left|\widehat{s}_{NT}^{PC}\right|^{\nu/2}}\left(\sum_{i=1}^{N}\left\|\beta_{i}\right\|^{\nu/2}\right)\left\|\widehat{\lambda}-\mathbf{H}^{-1}\lambda\right\|^{\nu/2}+\frac{C_{NT}^{1/2}}{\left|\widehat{s}_{NT}^{PC}\right|^{\nu/2}}\left(\sum_{i=1}^{N}\left\|\widehat{\beta}_{i}-\mathbf{H}'\beta_{i}\right\|^{\nu/2}\right)\left\|\widehat{\lambda}-\mathbf{H}^{-1}\lambda\right\|^{\nu/2}\right]\\
=\;&I+II+III+IV+V.
\end{aligned}$$

Starting from Lemmas C.18 and C.21, we can show that I, II and IV are $o_{a.s.}(1)$ proceeding as in the proof of Lemma C.14 - indeed, on account of Lemma C.16 IV contains the extra term

$$\frac{C_{NT}^{1/2}}{|\widehat{S}_{NT}^{PC}|^{\nu/2}}\left(\sum_{i=1}^{N}\|\beta_i\|^{\nu/2}\right)O\left(N^{-\nu/2}\right)=O\left(C_{NT}N^{1-\nu/2}\right),$$

but this can be shown to be $o(1)$ by routine calculations. As far as III is concerned, Assumption B.4, and Lemmas C.20-C.21 imply, after some algebra

$$III=C_{NT}^{1/2}\left\{o_{a.s.}\left(NT^{-\nu/4}\left(\log N\log T\right)^{(1+\epsilon)\nu/2}\right)+O\left(N^{1-\nu/2}\right)\right\}=o_{a.s.}(1),$$

where recall $\nu>4$. That $V=o_{a.s.}(1)$ readily follows from Lemma C.18 and the result on III. □

D. Proofs

Proof of Theorem 3.1. We begin by proving (3.4). The proof follows a similar approach to the proof of Theorem 3 in He et al. (2024), which we refine. To begin with, note that, for all $-\infty < x < \infty$

$$\mathbb{P}^*\left(\frac{Z_{N,T}-b_N}{a_N}\leq x\right)=P^*\left(Z_{N,T}\leq a_Nx+b_N\right),$$

where recall that $z_{i,NT}=\psi_{i,NT}+\omega_i$. Seeing as ω_i is, by construction, independent across i and independent of the sample, it follows that

$$\mathbb{P}^*\left(Z_{N,T}\leq a_Nx+b_N\right)=\prod_{i=1}^N\mathbb{P}^*\left(z_{i,NT}\leq a_Nx+b_N\right)=\prod_{i=1}^N\mathbb{P}^*\left(\omega_i\leq a_Nx+b_N-\psi_{i,NT}\right).$$

Let $\Phi(\cdot)$ denote the standard normal distribution; we have

$$\prod_{i=1}^N\mathbb{P}^*\left(\omega_i\leq a_Nx+b_N-\psi_{i,NT}\right)=\exp\left(\sum_{i=1}^N\log\Phi\left(a_Nx+b_N-\psi_{i,NT}\right)\right). \tag{D.1}$$

Note now that

$$\log\Phi\left(a_Nx+b_N-\psi_{i,NT}\right)=\log\Phi\left(a_Nx+b_N\right)+\log\frac{\Phi\left(a_Nx+b_N-\psi_{i,NT}\right)}{\Phi\left(a_Nx+b_N\right)}; \tag{D.2}$$

using Lagrange's theorem, there exists an $a_i^*\in\left(a_Nx+b_N-\psi_{i,NT},a_Nx+b_N\right)$ such that $\Phi\left(a_Nx+b_N-\psi_{i,NT}\right)=\Phi\left(a_Nx+b_N\right)-\varphi\left(a_i^*\right)\psi_{i,NT}$, where $\varphi(\cdot)$ denotes the density function of the standard normal, so that ultimately

$$\log\frac{\Phi\left(a_Nx+b_N-\psi_{i,NT}\right)}{\Phi\left(a_Nx+b_N\right)}=\log\left(1-\frac{\varphi\left(a_i^*\right)}{\Phi\left(a_Nx+b_N\right)}\psi_{i,NT}\right)=\log\left(1-c_i\psi_{i,NT}\right).$$

By elementary arguments, it follows that

$$\exp\left(\sum_{i=1}^N\log\frac{\Phi\left(a_Nx+b_N-\psi_{i,NT}\right)}{\Phi\left(a_Nx+b_N\right)}\right)$$

$$
\begin{aligned}
&= \exp\left(\sum_{i=1}^{N}\log\left(1-c_i\psi_{i,NT}\right)\right) = \exp\left(\frac{N}{N}\log\left(\prod_{i=1}^{N}\left(1-c_i\psi_{i,NT}\right)\right)\right) \\
&\leq \exp\left(N\log\left(\frac{1}{N}\sum_{i=1}^{N}\left(1-c_i\psi_{i,T}\right)\right)\right) = \exp\left(N\log\left(1-\left(\frac{1}{N}\sum_{i=1}^{N}c_i\psi_{i,NT}\right)\right)\right) \\
&= \exp\left(\sum_{h=1}^{\infty}N^{-h+1}\frac{(-1)^h\left(\sum_{i=1}^{N}c_i\psi_{i,NT}\right)^h}{h}\right),
\end{aligned}
$$

having used the arithmetic/geometric mean inequality to move from the second to the third line, and a Taylor expansion of $\log(1+x)$ around $x=0$ in the last line. Since $c_i \leq (2\pi)^{-1/2}\left[\Phi\left(a_N x+b_N\right)\right]^{-1} \leq \bar{c}$, and, by Lemma C.8,

$$
\mathbb{P}\left(\omega : \lim_{\min\{N,T\}\to\infty}\sum_{i=1}^{N}\psi_{i,NT}=0\right)=1,
$$

we can assume that $\lim_{\min\{N,T\}\to\infty}\sum_{i=1}^{N}\psi_{i,NT}=0$, it follows from elementary arguments that

$$
\text{(D.3)} \qquad \lim_{\min\{N,T\}\to\infty}\exp\left(\sum_{i=1}^{N}\log\frac{\Phi\left(a_N x+b_N-\psi_{i,NT}\right)}{\Phi\left(a_N x+b_N\right)}\right)=1.
$$

Thus we have

$$
\begin{aligned}
&\lim_{\min\{N,T\}\to\infty}\prod_{i=1}^{N}P^*\left(\omega_i \leq a_N x+b_N-\psi_{i,NT}\right) \\
&= \left(\lim_{N\to\infty}\Phi^N\left(a_N x+b_N\right)\right)\times\left(\lim_{\min\{N,T\}\to\infty}\exp\left(\sum_{i=1}^{N}\log\frac{\Phi\left(a_N x+b_N-\psi_{i,T}\right)}{\Phi\left(a_N x+b_N\right)}\right)\right) \\
&= \exp\left(-\exp\left(-x\right)\right),
\end{aligned}
$$

using the relations in (D.1) - (D.2) to move from the first to the second line, and the Fisher–Tippett–Gnedenko Theorem (see Theorem 3.2.3 in Embrechts et al., 2013b, among others) along with the limit in (D.3) to obtain the final result.

We now turn to showing (3.5). Under the alternative, there exists a set of $1 \leq m \leq N$ indices $\mathcal{I} = \{i_1, \ldots, i_m\} \subseteq \{1, \ldots, N\}$ such that $|\alpha_i| > 0$ whenever $i \in \mathcal{I}$; in these cases, $\psi_{i,NT}$ diverges almost surely at the rate $T^{1/2}$, i.e.

$$T^{-1/2}\psi_{i,NT} \overset{a.s.}{\to} c > 0 \text{ whenever } i \in \mathcal{I}.$$

Hence, we can assume that

$$\lim_{T\to\infty} T^{-1/2}\psi_{i,NT} = c > 0,$$

whenever $i \in \mathcal{I}$. Note now that, for any $-\infty < x < \infty$ we have

$$\begin{aligned} P^*\left(Z_{N,T} \leq a_N x + b_N\right) &= \prod_{i=1}^{N} P^*\left(z_{i,T} \leq a_N x + b_N\right) \leq \prod_{i\in\mathcal{I}} P^*\left(\omega_i \leq a_N x + b_N - \psi_{i,NT}\right) \\ &= \prod_{i\in\mathcal{I}} \Phi\left(a_N x + b_N - \psi_{i,NT}\right). \end{aligned}$$

Equation (5) in Borjesson and Sundberg (1979) entails that

$$\Phi\left(a_N x + b_N - \psi_{i,NT}\right) \leq \frac{\exp\left(-\frac{1}{2}\left(a_N x + b_N - \psi_{i,NT}\right)^2\right)}{\sqrt{2\pi}\left|a_N x + b_N - \psi_{i,NT}\right|}. \tag{D.4}$$

Seeing as, by construction, $a_N x + b_N = O(\sqrt{2\log N})$ for each $-\infty < x < \infty$, by Assumption 3.1 it follows that $a_N x + b_N - \psi_{i,NT} \overset{a.s.}{\to} -\infty$ whenever $i \in \mathcal{I}$. Hence, as $\min\{N, T\} \to \infty$ it holds that

$$0 \leq \Phi\left(a_N x + b_N - \psi_{i,NT}\right) \leq \frac{\exp\left(-\frac{1}{2}\left(a_N x + b_N - \psi_{i,NT}\right)^2\right)}{\sqrt{2\pi}\left|a_N x + b_N - \psi_{i,NT}\right|} \overset{a.s.}{\to} 0,$$

which, by dominated convergence, entails that $\Phi\left(a_N x + b_N - \psi_{i,T}\right) = o_{a.s.}(1)$. As long as $\mathcal{I}$ is not empty, this immediately entails that

$$\lim_{\min\{N,T\}\to\infty} \mathbb{P}^*\left(Z_{N,T} \leq a_N x + b_N\right) = 0,$$

for almost all realisations of $\left\{(u_{i,t}, f_t')', 1 \leq i \leq N, 1 \leq t \leq T\right\}$.

In conclusion, we note that the theorem still holds, with the proof virtually unchanged for the more general statistic in (B.1). The only difference consists in replacing $T^{-1/2}$ with $T^{-\delta\nu/2}$ when discussing the behavior under the alternative. No further calculations or arguments are required with respect to the current proof. $\square$

Proof of Theorem 3.2. Write, for short, $Q_{N,T,B}(\tau) = Q_\tau$. Recall

$$Q_\tau = \frac{1}{B}\sum_{b=1}^{B} \mathbb{I}\left(Z_{N,T}^{(b)} \leq c_\tau\right),$$

and let $X_{N,T}^{(b)} = \mathbb{I}\left(Z_{N,T}^{(b)} \leq c_\tau\right)$ for short. Note that, by similar passages as in the proof of Theorem 3.1

$$\begin{aligned}
&\mathbb{E}^*\left(X_{N,T}^{(b)}\right) \\
= &\ \mathbb{P}^*\left(Z_{N,T}^{(b)} \leq c_\tau\right) = \prod_{i=1}^{N} \mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau - \psi_{i,NT}\right) \\
= &\ \exp\left(\sum_{i=1}^{N} \log \mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau - \psi_{i,NT}\right)\right) \\
= &\ \exp\left(\sum_{i=1}^{N} \log\left[\mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau\right)\left(1 - \frac{\mathbb{P}^*\left(c_\tau - \psi_{i,NT} \leq \omega_i^{(b)} \leq c_\tau\right)}{\mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau\right)}\right)\right]\right) \\
= &\ \exp\left(\sum_{i=1}^{N} \log \mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau\right) - c_0 \sum_{i=1}^{N} \mathbb{P}^*\left(c_\tau - \psi_{i,NT} \leq \omega_i^{(b)} \leq c_\tau\right)\right) \\
= &\ \exp\left(\sum_{i=1}^{N} \log \mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau\right)\right) \exp\left(-c_1 \sum_{i=1}^{N} \psi_{i,NT}\right)
\end{aligned} \tag{D.5}$$

for some positive, finite constants c_0 and c_1, and having used the fact that $\psi_{i,NT}$ implies $\mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau - \psi_{i,NT}\right) = \mathbb{P}^*\left(\omega_i^{(b)} \leq c_\tau\right) - \mathbb{P}^*\left(c_\tau - \psi_{i,NT} \leq \omega_i^{(b)} \leq c_\tau\right)$ to move from the fourth to the fifth line.

We now start by showing (3.10). It holds that

$$
\text{(D.6)}\quad \frac{Q_{\tau}-(1-\tau)}{\sqrt{\tau(1-\tau)}} = \frac{Q_{\tau}-(1-\tau)}{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}\frac{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}{\sqrt{\tau(1-\tau)}}
$$

$$
= \frac{Q_{\tau}-\mathbb{E}^{*}(Q_{\tau})}{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}\frac{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}{\sqrt{\tau(1-\tau)}}+\frac{\mathbb{E}^{*}(Q_{\tau})-(1-\tau)}{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}\frac{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}{\sqrt{\tau(1-\tau)}}
$$

$$
= I+II,
$$

where $\mathcal{V}^{*}(Q_{\tau})$ denotes the variance of Q_{τ} conditional on the sample, with

$$
\mathcal{V}^{*}(Q_{\tau})=\mathbb{E}^{*}\left(X_{N,T}^{(b)}\right)\left[1-\mathbb{E}^{*}\left(X_{N,T}^{(b)}\right)\right].
$$

On account of (D.5), it follows that $\mathbb{E}^{*}\left(X_{N,T}^{(b)}\right)=(1-\tau)+o_{a.s.}(1)$, and therefore it also follows that $\mathcal{V}^{*}(Q_{\tau})=\tau(1-\tau)+o_{a.s.}(1)$, whence the Law of the Iterated Logarithm ultimately yields that, as far as I in (D.6) is concerned

$$
\left|\frac{Q_{\tau}-\mathbb{E}^{*}(Q_{\tau})}{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}\frac{\left(\mathcal{V}^{*}(Q_{\tau})\right)^{1/2}}{\sqrt{\tau(1-\tau)}}\right|=O_{a.s.}\left(\sqrt{\frac{2\log\log B}{B}}\right).
$$

We now turn to studying II in (D.6). We start by establishing that:

$$
\mathbb{E}^{*}(Q_{\tau})-(1-\tau)=o_{a.s.}\left(\sqrt{\frac{2\log\log B}{B}}\right).
$$

Using (D.5), we receive

$$
\mathbb{E}^{*}(Q_{\tau}) = \mathbb{E}^{*}\left(\frac{1}{B}\sum_{b=1}^{B}X_{N,T}^{(b)}\right)=\mathbb{E}^{*}\left(X_{N,T}^{(b)}\right)
$$

$$
= \exp\left(\sum_{i=1}^{N}\log\mathbb{P}^{*}\left(\omega_{i}^{(b)}\leq c_{\tau}\right)\right)\exp\left(-c_{1}\sum_{i=1}^{N}\psi_{i,NT}\right)
$$

$$
= \left(\mathbb{P}^{*}\left(\omega_{i}^{(b)}\leq c_{\tau}\right)\right)^{N}\exp\left(-c_{1}\sum_{i=1}^{N}\psi_{i,NT}\right).
$$

Note now that

$$
\begin{aligned}
& \sqrt{\frac{B}{\log\log B}} \left| \mathbb{E}^* \left(Q_\tau \right) - (1-\tau) \right| \\
\leq \; & \sqrt{\frac{B}{\log\log B}} \left| \left(\mathbb{P}^* \left(\omega_i^{(b)} \leq c_\tau \right) \right)^N - (1-\tau) \right| \\
& + (1-\tau) \sqrt{\frac{B}{\log\log B}} \left| \exp \left(-c_1 \sum_{i=1}^{N} \psi_{i,NT} \right) - 1 \right| = II_a + II_b.
\end{aligned}
$$

Using equation (10) in Hall (1979) - with, in his notation, $x = -\log\left(-\log\left(1-\tau\right)\right)$ - it holds that

$$
II_a \leq c_0 \sqrt{\frac{B}{\log\log B}} \frac{1}{\log N} = o\left(1\right),
$$

having used the fact that $B = O\left(\left(\log N \right)^2 \right)$. Further, by a standard application of the Mean Value Theorem

$$
II_b \leq c_0 \sqrt{\frac{B}{\log\log B}} \left| \sum_{i=1}^{N} \psi_{i,T} \right| = o_{a.s}\left(1\right),
$$

by Assumption 3.1. Hence we have

$$
\sqrt{\frac{B}{\log\log B}} \left| \mathbb{E}^* \left(Q_\tau \right) - (1-\tau) \right| = o_{a.s}\left(1\right),
$$

which immediately yields the desired result.

Under the alternative, we write

$$
\begin{aligned}
Q_\tau &= \mathbb{E}^* \left(Q_\tau \right) + Q_\tau - \mathbb{E}^* \left(Q_\tau \right) \\
&= \frac{1}{B} \sum_{b=1}^{B} \mathbb{E}^* \left(X_{N,T}^{(b)} \right) + \frac{1}{B} \sum_{b=1}^{B} \left(X_{N,T}^{(b)} - \mathbb{E}^* \left(X_{N,T}^{(b)} \right) \right) \\
&= \mathbb{E}^* \left(X_{N,T}^{(1)} \right) + \frac{1}{B} \sum_{b=1}^{B} \left(X_{N,T}^{(b)} - \mathbb{E}^* \left(X_{N,T}^{(b)} \right) \right) = I + II.
\end{aligned}
$$

We know from the proof of Theorem 3.1 that, under $\mathbb{H}_A$, $I = o_{a.s.}(1)$. Moreover, note that, due to $X_{N,T}^{(b)}$ being *i.i.d.* across $1 \leq b \leq B$

$$\mathcal{V}^* \left(\frac{1}{B} \sum_{b=1}^{B} \left(X_{N,T}^{(b)} - \mathbb{E}^* \left(X_{N,T}^{(b)} \right) \right) \right) = B^{-1} \mathcal{V}^* \left(X_{N,T}^{(1)} \right) \leq c_0 B^{-1},$$

a.s., and therefore, by the Law of the Total Variance, it also holds that

$$\mathcal{V} \left(\frac{1}{B} \sum_{b=1}^{B} \left(X_{N,T}^{(b)} - \mathbb{E}^* \left(X_{N,T}^{(b)} \right) \right) \right) \leq c_0 B^{-1};$$

Lemma C.1 then entails that $II = o_{a.s.}(1)$. The desired result now follows automatically.

□

Proof of Theorem 4.1. The proof is very similar to that of Theorem 3.1, and therefore we only report its main arguments. Under $\mathbb{H}_0$, the result follows immediately from Lemma C.14. Under $\mathbb{H}_A$, using (C.8) it follows that, for $1 \leq i \leq N$

$$\widehat{\alpha}_i^{FM} = \alpha_i + \beta_i' \overline{v} + \overline{u}_i - \left(\widehat{\beta}_i - \beta_i \right)' \lambda - \beta_i' \left(\widehat{\lambda} - \lambda \right) - \left(\widehat{\beta}_i - \beta_i \right)' \left(\widehat{\lambda} - \lambda \right),$$

where recall that, by Lemma C.12

$$\widehat{\lambda} - \lambda = \frac{1}{N} \mathbf{S}_\beta^{-1} \boldsymbol{\beta}' \mathbb{M}_{1_N} \boldsymbol{\alpha} + o_{a.s.} \left(T^{-1/2} \left(\log T \right)^{1+\epsilon} \right).$$

We consider, for simplicity, the case where $\alpha_1 \neq 0$, and $\alpha_i = 0$ for $i \geq 2$. Other, more complicated cases, can be studied by the same token, and we only add some discussion at the end of this proof. In such a case, it holds that

$$\begin{aligned} \widehat{\alpha}_1^{FM} &= \alpha_1 + \frac{1}{N} \mathbf{S}_\beta^{-1} \boldsymbol{\beta}' \mathbb{M}_{1_N} \boldsymbol{\alpha} + r_{N,T}, \\ \widehat{\alpha}_i^{FM} &= \frac{1}{N} \mathbf{S}_\beta^{-1} \boldsymbol{\beta}' \mathbb{M}_{1_N} \boldsymbol{\alpha} + r_{N,T}, \end{aligned}$$

where the remainder term can be studied along the same lines as under the null. Note that, under the case considered

$$\frac{1}{N}\mathbf{S}_{\beta}^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha} =\frac{1}{N}\frac{N-1}{N}\mathbf{S}_{\beta}^{-1}\beta_1\alpha_1 - \frac{1}{N^2}\mathbf{S}_{\beta}^{-1}\sum_{i=1}^{N}\beta_i\alpha_1 = O\left(\frac{1}{N}\right),$$

and therefore $\widehat{\alpha}_1^{FM} = \alpha_1$ plus a negligible remainder. The proof now is the same as the proof of Theorem 3.1. In conclusion, note that, under general alternatives, we need to guarantee that

$$\max_{1\leq i\leq N}\left|\alpha_i + \frac{1}{N}\mathbf{S}_{\beta}^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha}\right| > 0, \tag{D.7}$$

in order to be able to apply the arguments in the proof of Theorem 3.1. However, this condition is always satisfied. Indeed, the only way it cannot hold is if

$$\alpha_i + \frac{1}{N}\mathbf{S}_{\beta}^{-1}\boldsymbol{\beta}'\mathbb{M}_{1_N}\boldsymbol{\alpha} = 0, \tag{D.8}$$

for all $1 \leq i \leq N$, which, in turn, requires under $\alpha_i = \alpha$ for all $1 \leq i \leq N$. However, in such a case, $\mathbb{M}_{1_N}\boldsymbol{\alpha} = \mathbf{0}$, and therefore the only way (D.8) can be satisfied is if $\alpha_i = 0$ for all $1 \leq i \leq N$; but this cannot hold under the alternative hypothesis. □

Proof of Theorem 4.2. The proof is essentially the same as that of Theorem 4.1, *mutatis mutandis.* □