MAXIMUM CONDITIONAL ALPHA TEST FOR CONDITIONAL MULTI-FACTOR MODELS

Jun Zhang, Wei Lan, Xinyan Fan and Wen Chen

Southwestern University of Finance and Economics,

Renmin University of China and Nanjing Agricultural University

Supplementary Material

The Supplementary Material consists of ten parts (Sections S1–S10). Section S1 provides eight useful lemmas, Section S2 provides the proof of Theorem 1, Section S3 provides the proof of Theorem 2, Section S4 provides the proof of Proposition 1, Section S5 provides the proof of Theorem 3, Section S6 describes the test portfolios, Section S7 presents the empirical evidence for the time-varying coefficients and sparse alternatives based on real data, Section S8 reports the simulation results of conditional multi-factor models with latent factors, Section S9 reports the simulation results of the MAX test proposed by Feng et al. (2022), and Section S10 provides the simulation results for a student-t distribution error.

S1 Eight Useful Lemmas

We only present the proof of Lemmas 5 and 8. Lemmas 1, 3, and 4 are borrowed from Ma et al. (2020), Lemma 2 is borrowed from Cai et al.

(2014), Lemma 6 can be obtained directly through Bonferroni inequality (Wang, 2012), and the Lemma 7 is directly borrowed from Fan and Han (2017).

Lemma 1. *Define* $\rho_{i0t} = \alpha_i(t/T) - \boldsymbol{\gamma}_{i0}^{0\top}\widetilde{\boldsymbol{B}}(t/T)$ *and* $\rho_{ijt} = \beta_{ij}(t/T) - \boldsymbol{\gamma}_{ij}^{0\top}\boldsymbol{B}(t/T)$ *for* $1 \leq j \leq d$ *and* $1 \leq i \leq N$*. Then, under Assumption (A.1), there exist* $\boldsymbol{\gamma}_{i0}^{0} \in \mathbb{R}^L$ *and* $\boldsymbol{\gamma}_{ij}^{0} \in \mathbb{R}^L$ *such that* $\sup_{1\leq t\leq T}|\rho_{i0t}| = O(L^{-r})$ *and* $\sup_{1\leq t\leq T}|\rho_{ijt}| = O(L^{-r})$ *as* $T \to \infty$.

Lemma 2. *(Bonferroni Inequality). Let* $\mathcal{A} = \bigcup_{k=1}^{N}\mathcal{A}_k$*. For any* $K < [N/2]$*, we have*

$$\sum_{k=1}^{2K}(-1)^{k-1}\mathcal{V}_k \leq P(\mathcal{A}) \leq \sum_{k=1}^{2K-1}(-1)^{k-1}\mathcal{V}_k,$$

where $\mathcal{V}_k = \sum_{1\leq i_1<\cdots<i_k\leq N} P(\mathcal{A}_{i_1}\bigcup\cdots\bigcup\mathcal{A}_{i_k})$.

Lemma 3. *Under Assumption (A.3), there exist constants* $0 < c_z \leq C_z < \infty$ *with probability one,*

$$c_zL^{-1} \leq \lambda_{\min}(\mathbb{Z}^\top\mathbb{Z}/T) \leq \lambda_{\max}(\mathbb{Z}^\top\mathbb{Z}/T) \leq C_zL^{-1},$$

$$c_zL \leq \lambda_{\min}\{(\mathbb{Z}^\top\mathbb{Z}/T)^{-1}\} \leq \lambda_{\max}\{(\mathbb{Z}^\top\mathbb{Z}/T)^{-1}\} \leq C_zL,$$

as $T \to \infty$.

Lemma 4. *Under Assumption (A.3) (i), there exist constants* $0 < c_m \leq C_m \leq 1$ *such that* $c_mT \leq \boldsymbol{1}_T^\top\boldsymbol{M}_{\mathbb{Z}}\boldsymbol{1}_T \leq C_mT$.

Lemma 5. *Under Assumptions (A.1)-(A.4), we have*

$$P\left\{\max_{1\le i,j\le N}|\widehat{\sigma}_{ij}-\sigma_{ij}|>C_\sigma\big(\sqrt{L\frac{\log(N)}{T}}+L^{-r}\big)\right\}\to 0,$$

for some constant $C_\sigma > 0$.

Proof: Under Assumptions (A.2)-(A.4), by the Bernstein inequality, there exist two finite positive constants C_1 and C_2 such that the event

$$\mathbb{A}=\left\{\max_i\left|\frac{1}{T}\sum_{i=1}^{T}\varepsilon_{it}\varepsilon_{jt}-\sigma_{ij}\right|<C_1\sqrt{\frac{\log N}{T}};\frac{1}{T}\sum_{t=1}^{T}\|\mathbf{f}_t\|^2<C_2\right\}$$

occurs with the probability approaching one. In addition, there further exist two finite positive constants C_3 and C_4 such that the event $\mathbb{B}=\left\{\max_{i,j}\frac{1}{T}|\sum_{t=1}^{T}f_{jt}\varepsilon_{it}|<C_3\sqrt{\frac{\log N}{T}};\max_i\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2<C_4\right\}$ occurs with the probability approaching one. Then, on the event $\mathbb{A}\cap\mathbb{B}$, by the triangular and Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned}\max_{i,j}|\widehat{\sigma}_{ij}-\sigma_{ij}|\le C_1\sqrt{\frac{\log(N)}{T}}&+2\max_i\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\widehat{\varepsilon}_{it}-\varepsilon_{it})^2C_4}\\&+\max_i\frac{1}{T}\sum_{t=1}^{T}(\varepsilon_{it}-\widehat{\varepsilon}_{it})^2,\end{aligned}$$

where $\widehat{\varepsilon}_{it}$ is the t-th element of $\mathbf{M}_{\widetilde{\mathbb{Z}}}\mathbf{R}_i$.

Denote $\boldsymbol{\rho}_{it}=\rho_{it0}+\sum_{j=1}^{d}\rho_{ijt}f_{jt}$, $\boldsymbol{\rho}_i=(\boldsymbol{\rho}_{i1},\ldots,\boldsymbol{\rho}_{iT})^\top$, $\mathbf{P}_{\widetilde{\mathbb{Z}}}=\widetilde{\mathbb{Z}}\big(\widetilde{\mathbb{Z}}^\top\widetilde{\mathbb{Z}}\big)^{-1}\widetilde{\mathbb{Z}}^\top$, and $\mathbf{P}_{\mathbb{Z}} = \mathbb{Z}\big(\mathbb{Z}^\top\mathbb{Z}\big)^{-1}\mathbb{Z}^\top$. Then, we have $\mathbf{R}_i = \alpha_{i,ACA}\mathbf{1}_T + \mathbb{Z}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i + \boldsymbol{\rho}_i$, and

$$\begin{aligned}
&\max_i T^{-1}\sum_{t=1}^{T}(\widehat{\varepsilon}_{it} - \varepsilon_{it})^2 \\
&= \max_i T^{-1}(-\mathbf{P}_{\widetilde{\mathbb{Z}}}\boldsymbol{\varepsilon}_i + \mathbf{M}_{\widetilde{\mathbb{Z}}}\boldsymbol{\rho}_i)^\top(-\mathbf{P}_{\widetilde{\mathbb{Z}}}\boldsymbol{\varepsilon}_i + \mathbf{M}_{\widetilde{\mathbb{Z}}}\boldsymbol{\rho}_i) \\
&\le \max_i 2T^{-1}\boldsymbol{\varepsilon}_i^\top\mathbf{P}_{\widetilde{\mathbb{Z}}}\boldsymbol{\varepsilon}_i + 2T^{-1}\boldsymbol{\rho}_i^\top\mathbf{M}_{\widetilde{\mathbb{Z}}}\boldsymbol{\rho}_i \\
&\le \max_i 2T^{-1}\boldsymbol{\varepsilon}_i^\top\mathbf{P}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i + 2T^{-1}\boldsymbol{\varepsilon}_i^\top(\mathbf{P}_{\mathbb{Z}} - \mathbf{P}_{\widetilde{\mathbb{Z}}})\boldsymbol{\varepsilon}_i + 2T^{-1}\boldsymbol{\rho}_i^\top\boldsymbol{\rho}_i. \qquad \text{(S1.1)}
\end{aligned}$$

We next prove the three parts involved in (S1.1) separately. For the first part of (S1.1), according to Lemma 3, there exist finite positive constants C_5–C_7 such that

$$\begin{aligned}
&\max_i T^{-1}\boldsymbol{\varepsilon}_i^\top\mathbf{P}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i \\
&\le \max_i T^{-2}\boldsymbol{\varepsilon}_i^\top\mathbb{Z}\mathbb{Z}^\top\boldsymbol{\varepsilon}_i\lambda_{\max}\Big[(\frac{1}{T}\mathbb{Z}^\top\mathbb{Z})^{-1}\Big] \\
&\le \max_i C_5LT^{-2}\boldsymbol{\varepsilon}_i^\top\mathbb{Z}\mathbb{Z}^\top\boldsymbol{\varepsilon}_i \\
&= \max_i C_5LT^{-2}\Big[\sum_{l=1}^{L}(\sum_{t\in\{|l(t)-l|\le q-1\}}\widetilde{B}_l(t/T)\varepsilon_{it})^2 + \sum_{j=1}^{d}\sum_{l=1}^{L}(\sum_{t\in\{|l(t)-l|\le q-1\}}f_{jt}B_l(t/T)\varepsilon_{it})^2\Big] \\
&\le C_6L^{-1}\sum_{l=1}^{L}\Big\{\max_i(\frac{L}{T}\sum_{t\in\{|l(t)-l|\le q-1\}}\varepsilon_{it})^2 + \max_{i,j}(\frac{L}{T}\sum_{t\in\{|l(t)-l|\le q-1\}}f_{jt}\varepsilon_{it})^2\Big\} \\
&\le C_7L\frac{\log(N)}{T}. \qquad \text{(S1.2)}
\end{aligned}$$

For the second part of (S1.1), we have

$$\max_i T^{-1}\boldsymbol{\varepsilon}_i^\top(\mathbf{P}_{\mathbb{Z}} - \mathbf{P}_{\widetilde{\mathbb{Z}}})\boldsymbol{\varepsilon}_i = \max_i T^{-1}\boldsymbol{\varepsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T(\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1}\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i.$$

By Hoeffding inequality, we have

$$
\begin{aligned}
&P\big[\max_i T^{-1}\boldsymbol{\varepsilon}_i^\top(\mathbf{P}_{\mathbb{Z}}-\mathbf{P}_{\widetilde{\mathbb{Z}}})\boldsymbol{\varepsilon}_i \geq C_8\log(N)/T\big]\\
\leq&\sum_{i=1}^N P\big[T^{-1/2}|(\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1/2}\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i| \geq C_8\sqrt{\log(N)/T}\big]\\
\leq&\exp(-C_9\log(N)).
\end{aligned}
$$

For the third part of (S1.1), there exists finite positive constant C_8 such that

$$
\max_i T^{-1}\boldsymbol{\rho}_i^\top\boldsymbol{\rho}_i \leq \max_i T^{-1}(d+1)\sum_{t=1}^T\{\rho_{it0}^2+\sum_{j=1}^d\rho_{ijt}^2 f_{jt}^2\} < C_8 L^{-2r}, \quad \text{(S1.3)}
$$

where the last inequality is a result of Lemma 1. Combining the results in (S1.2) and (S1.3), we have $\max_{i,j}|\widehat{\sigma}_{ij}-\sigma_{ij}| \leq C_\sigma\big(\sqrt{\frac{L\log(N)}{T}}+L^{-r}\big)$ for positive constant C_σ with the probability approaching one, which completes the proof of this lemma.

Lemma 6. *Let $\chi_{i,T}$ be a random variable with a chi-squared distribution of degree T, then we should have $\max_i|T^{-1}\chi_{i,T}-1| = O_p(\sqrt{\log(N)/T})$ for any $1\leq i\leq N$.*

Lemma 7. *For any arbitrary symmetric matrices $\mathbf{A}$ and $\mathbf{B}$, we have*

$$
|\lambda_{i\mathbf{A}}-\lambda_{i\mathbf{B}}| \leq \|\mathbf{A}-\mathbf{B}\| \quad \text{and} \quad \|\boldsymbol{\zeta}_{i\mathbf{A}}-\boldsymbol{\zeta}_{i\mathbf{B}}\| \leq \frac{\sqrt{2}\|\mathbf{A}-\mathbf{B}\|}{\min\{|\lambda_{i-1,\mathbf{A}}-\lambda_{i\mathbf{B}}|,|\lambda_{i+1,\mathbf{A}}-\lambda_{i\mathbf{B}}|\}},
$$

where $\lambda_{i\mathbf{C}}$'s and $\boldsymbol{\zeta}_{i\mathbf{C}}$'s are the eigenvalues and eigenvectors of any arbitrary matrix $\mathbf{C}$ for $i\geq 1$, and $\boldsymbol{\zeta}_0=\infty$.

Lemma 8. *Under Assumptions (A.1), (A.3), (A.5), and (A.6), we further assume* $\widetilde{\mu}_{e-1} - \widetilde{\mu}_e \geq d_N$ *for some positive constant* d_N *and for any* $e = 2, \cdots, v$. *If* $\log(N) = o(L)$, *we can obtain that*

$$\max_{1\leq i\leq N} |\widehat{\sigma}_{\boldsymbol{\epsilon},ii} - \sigma_{\boldsymbol{\epsilon},ii}| = O_p(L^{-r} + L^{1/2}T^{-1/2}).$$

PROOF. Recall that $\mathbf{R}_i = \alpha_{i,ACA}\mathbf{1}_T + \mathbb{Z}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i + \boldsymbol{\rho}_i$, and $\widehat{\boldsymbol{\varepsilon}}_i = \mathbf{M}_{\mathbb{Z}}\mathbf{R}_i = \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T\alpha_{i,ACA} + \mathbf{M}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i + \mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}_i$. Note that $\widehat{\sigma}_{\boldsymbol{\epsilon},ii} = T^{-1}\widehat{\boldsymbol{\epsilon}}_i^\top\widehat{\boldsymbol{\epsilon}}_i$ with $\widehat{\boldsymbol{\epsilon}}_i^\top = \mathbf{M}_{\widehat{\mathbf{X}}}\widehat{\boldsymbol{\varepsilon}}_i^\top$. As a result, we have $\widehat{\sigma}_{\boldsymbol{\epsilon},ii} = T^{-1}\widehat{\boldsymbol{\varepsilon}}_i^\top\mathbf{M}_{\widehat{\mathbf{X}}}\widehat{\boldsymbol{\varepsilon}}_i$, which leads to

$$\begin{aligned}
&\max_i |\widehat{\sigma}_{\boldsymbol{\epsilon},ii} - \sigma_{\boldsymbol{\epsilon},ii}| \\
&= \left|T^{-1}\widehat{\boldsymbol{\varepsilon}}_i^\top\mathbf{M}_{\widehat{\mathbf{X}}}\widehat{\boldsymbol{\varepsilon}}_i - T^{-1}\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i + T^{-1}\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i - \sigma_{\boldsymbol{\epsilon},ii}\right| \\
&\leq \max_i \left|T^{-1}\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i - \sigma_{\boldsymbol{\epsilon},ii}\right| + \max_i \left|T^{-1}\widehat{\boldsymbol{\varepsilon}}_i^\top\mathbf{M}_{\widehat{\mathbf{X}}}\widehat{\boldsymbol{\varepsilon}}_i - T^{-1}\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i\right| \\
&= \widetilde{I}_1 + \widetilde{I}_2.
\end{aligned} \tag{S1.4}$$

We first consider $\widetilde{I}_1$. Note that $\boldsymbol{\epsilon}_i$ follows a multivariate normal distribution with mean zero and covariance matrix $\sigma_{\boldsymbol{\epsilon},ii}I_T$. As a result, $\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i/\sigma_{\boldsymbol{\epsilon},ii}$ follows a chi-square distribution of degree $T-L$. Thus, according to Lemma 6, we can obtain that

$$\begin{aligned}
\widetilde{I}_1 =& \max_i \left|T^{-1}\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i - \sigma_{\boldsymbol{\epsilon},ii}\right| \\
&\leq \max_i \left|T^{-1}\boldsymbol{\epsilon}_i^\top\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i - (T-L)\sigma_{\boldsymbol{\epsilon},ii}/T\right| + L\max_i |\sigma_{\boldsymbol{\epsilon},ii}|/T \\
&= \max_i \sigma_{\boldsymbol{\epsilon},ii}\left|T^{-1}\chi_{T-L} - (T-L)/T\right| + L\max_i \sigma_{\boldsymbol{\epsilon},ii}/T \\
&= O_p(\sqrt{\log(N)/T} + LT^{-1}).
\end{aligned} \tag{S1.5}$$

For $\widetilde{I}_2$, under the null hypothesis, we have we have $\widehat{\boldsymbol{\varepsilon}}_i = \mathbf{M}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i + \mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}_i$ and $\boldsymbol{\varepsilon}_i = \mathbf{X}\boldsymbol{\lambda}_i + \boldsymbol{\epsilon}_i$. Then,

$$\begin{aligned}
\widetilde{I}_2 =& \max_i \left\{|T^{-1}\widehat{\boldsymbol{\varepsilon}}_i^\top \mathbf{M}_{\widehat{\mathbf{X}}}\widehat{\boldsymbol{\varepsilon}}_i - T^{-1}\boldsymbol{\epsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i|\right\} \\
=& \max_i\{|T^{-1}\boldsymbol{\epsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i - T^{-1}\boldsymbol{\epsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\boldsymbol{\epsilon}_i|\} + \max_i |T^{-1}\boldsymbol{\lambda}_i^\top \mathbf{X}^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i| \\
&+ \max_i |T^{-1}\boldsymbol{\rho}_i^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}_i| + \max_i |2T^{-1}\boldsymbol{\epsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i| + \max_i 2|T^{-1}\boldsymbol{\epsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}_i| \\
&+ \max_i |2T^{-1}\boldsymbol{\rho}_i^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i| \\
=& J_1 + J_2 + J_3 + J_4 + J_5 + J_6.
\end{aligned}$$

Then, we give the order of J_i for $i = 1, \ldots, 6$ in the following four steps. We consider J_1 to J_3 in Steps I to III, respectively, and J_4 to J_6 in Step IV.

STEP I. Before we give the order of J_1, we first bound the difference $\mathbf{P}_{\widehat{\mathbf{X}}}$ and $\mathbf{P}_{\mathbf{X}}$. We have that

$$\|\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}}\| \le \|\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\widetilde{\mathbf{X}}}\| + \|\mathbf{P}_{\widetilde{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}}\|.$$

According to Theorem 2 of Wang (2012), we have

$$\|\mathbf{P}_{\widetilde{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}}\| \le \mathrm{tr}^{1/2}\{(\mathbf{P}_{\widetilde{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}})^2\} = O_p(T^{-1/2}).$$

Next, we provide the order of $\|\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\widetilde{\mathbf{X}}}\|$ according to Lemma 7. Accordingly, by Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\left\|\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\widetilde{\mathbf{X}}}\right\|_2^2 \\
&= T^{-2}\left\|\widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top - \widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^\top\right\|^2 = T^{-2}\left\|\widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top - \widehat{\mathbf{X}}\widetilde{\mathbf{X}}^\top + \widehat{\mathbf{X}}\widetilde{\mathbf{X}}^\top - \widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^\top\right\|^2 \\
&\le 2T^{-2}\left\|\widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top - \widehat{\mathbf{X}}\widetilde{\mathbf{X}}^\top\right\|^2 + 2T^{-2}\left\|\widehat{\mathbf{X}}\widetilde{\mathbf{X}}^\top - \widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^\top\right\|^2 = 4T^{-1}\|\widehat{\mathbf{X}} - \widetilde{\mathbf{X}}\|^2.
\end{aligned}$$

Accordingly, by Lemma 7, we have

$$\|T^{-1/2}\widehat{\mathbf{X}} - T^{-1/2}\widetilde{\mathbf{X}}\| \leq \sqrt{2v}\|(TN)^{-1}\widehat{\mathcal{E}}\widehat{\mathcal{E}}^\top - (TN)^{-1}\mathcal{E}\mathcal{E}^\top\|/K_{\min},$$

where $K_{\min} = \min_{e\in\{1\ldots,v\}}\{|\widehat{\mu}_{e-1} - \widetilde{\mu}_e|, |\widehat{\mu}_{e+1} - \widetilde{\mu}_e|\}$. Under H_0, $\boldsymbol{\alpha}_{ACA} = 0$, we obtain

$$\begin{aligned}\left\|(NT)^{-1}\widehat{\mathcal{E}}\widehat{\mathcal{E}}^\top - (NT)^{-1}\mathcal{E}\mathcal{E}^\top\right\| \leq & \left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}\boldsymbol{\rho}^\top\mathbf{M}_{\mathbb{Z}}\right\| + 2\left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\mathcal{E}\boldsymbol{\rho}^\top\mathbf{M}_{\mathbb{Z}}\right\| \\ & + \left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\mathcal{E}\mathcal{E}^\top\mathbf{M}_{\mathbb{Z}} - (NT)^{-1}\mathcal{E}\mathcal{E}^\top\right\|.\end{aligned} \tag{S1.6}$$

Moreover, one can easily verify that $\lambda_{\max}\left(N^{-1}T^{-1}\boldsymbol{\rho}\boldsymbol{\rho}^\top\right) = O_p(L^{-2r})$ using (S1.3) , where $\boldsymbol{\rho} = (\boldsymbol{\rho}_1, \cdots, \boldsymbol{\rho}_T)^\top \in \mathbb{R}^{T\times N}$. Then,

$$\left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}\boldsymbol{\rho}^\top\mathbf{M}_{\mathbb{Z}}\right\| \leq \lambda_{\max}(\mathbf{M}_{\mathbb{Z}})\left\|(NT)^{-1}\boldsymbol{\rho}\boldsymbol{\rho}^\top\right\| = O_p\left(L^{-2r}\right).$$

Moreover, we also can easily verify that $\lambda_{\max}\left(N^{-1}T^{-1}\mathcal{E}\mathcal{E}^\top\right) = O_p(1)$ by Assumption (A.6). Then, we can also obtain

$$\begin{aligned}\left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\mathcal{E}\boldsymbol{\rho}^\top\mathbf{M}_{\mathbb{Z}}\right\| &\leq \lambda_{\max}(\mathbf{M}_{\mathbb{Z}})\left\|(NT)^{-1}\mathcal{E}\boldsymbol{\rho}^\top\right\| \\ &\leq (NT)^{-1}\lambda_{\max}^{1/2}\left(\mathcal{E}\mathcal{E}^\top\right)\lambda_{\max}^{1/2}\left(\boldsymbol{\rho}\boldsymbol{\rho}^\top\right) = O_p\left(L^{-r}\right).\end{aligned}$$

For the third term of (S1.6), we have that

$$\begin{aligned}\left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\mathcal{E}\mathcal{E}^\top\mathbf{M}_{\mathbb{Z}} - (NT)^{-1}\mathcal{E}\mathcal{E}^\top\right\| \leq & 2(NT)^{-1}\|\mathbf{P}_{\mathbb{Z}}\mathcal{E}\mathcal{E}^\top\| + (NT)^{-1}\|\mathbf{P}_{\mathbb{Z}}\mathcal{E}\mathcal{E}^\top\mathbf{P}_{\mathbb{Z}}\|^2 \\ \leq & 2(NT)^{-1}\|\mathbf{P}_{\mathbb{Z}}\mathcal{E}\|\|\mathcal{E}\| + (NT)^{-1}\|\mathbf{P}_{\mathbb{Z}}\mathcal{E}\|^2\end{aligned}$$

Next, we bound the order of $\|\mathbf{P}_{\mathbb{Z}}\mathcal{E}\|$. According to the definition of $\mathbf{P}_{\mathbb{Z}}$, the eigen-decomposition of $\mathbf{P}_{\mathbb{Z}}$ can be written as $U_P U_P^\top$, where U_P is the $N \times L$

matrix consisting of the first L eigenvectors of $\mathbf{P}_{\mathbb{Z}}$. Then,

$$\|\mathbf{P}_{\mathbb{Z}}\mathcal{E}\| \leq \|U_P^\top \mathcal{E}\| \leq \|U_P^\top \mathbf{X}\boldsymbol{\Lambda}\| + \|U_P^\top \boldsymbol{\epsilon}\| \leq \|U_P^\top \mathbf{X}\|\|\boldsymbol{\Lambda}\| + \|U_P^\top \boldsymbol{\epsilon}\|.$$

Since elements in $\mathbf{X}_t$ and $\boldsymbol{\epsilon}_i$ are independent normal random variables, according to Tomioka and Suzuki (2014), we have $\|U_P^\top \mathbf{X}\| = O_p(\sqrt{L})$ and $\|U_P^\top \boldsymbol{\epsilon}\| = O_p(\sqrt{NL})$. According to Assumption (A.6), $\|\boldsymbol{\Lambda}\| = O(\sqrt{N})$. Thus

$$\left\|(NT)^{-1}\mathbf{M}_{\mathbb{Z}}\mathcal{E}\mathcal{E}^\top \mathbf{M}_{\mathbb{Z}} - (NT)^{-1}\mathcal{E}\mathcal{E}^\top\right\| \leq O_p(\sqrt{L/T}).$$

Accordingly, we can obtain that

$$\left\|(NT)^{-1}\widehat{\mathcal{E}}\widehat{\mathcal{E}}^\top - (NT)^{-1}\mathcal{E}\mathcal{E}^\top\right\| = O_p(\sqrt{L/T} + L^{-r}).$$

Next, we consider $K_{\min}$. Note that

$$|\widehat{\mu}_{e-1} - \widetilde{\mu}_e| = |\widehat{\mu}_{e-1} - \widetilde{\mu}_{e-1} + \widetilde{\mu}_{e-1} - \widetilde{\mu}_e| \geq |\widetilde{\mu}_{e-1} - \widetilde{\mu}_e| - |\widehat{\mu}_{e-1} - \widetilde{\mu}_{e-1}|.$$

By the assumption that $|\widetilde{\mu}_{e-1} - \widetilde{\mu}_e| \geq d_N$ for any $e = 2, \ldots, v$. Moreover, by Lemma 7 again, we have

$$|\widehat{\mu}_{e-1} - \widetilde{\mu}_{e-1}| \leq \left\|(TN)^{-1}\widehat{\mathcal{E}}\widehat{\mathcal{E}}^\top - (TN)^{-1}\mathcal{E}\mathcal{E}^\top\right\| = o_p(1).$$

Consequently, with the probability approaching one,

$$2K_{\min} \geq |\widehat{\mu}_{e-1} - \widetilde{\mu}_e| \geq d_N.$$

Combining the above results, we can obtain that

$$\left\|\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}}\right\| = O_p(L^{-r} + \sqrt{L/T}). \tag{S1.7}$$

Now we consider J_1. We have

$$J_1 = \max_i \left|T^{-1}\boldsymbol{\varepsilon}_i^\top \mathbf{M}_{\mathbb{Z}}(\mathbf{I}_T - \mathbf{M}_{\widehat{\mathbf{X}}})\mathbf{M}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i\right| = \max_i \left|T^{-2}\boldsymbol{\varepsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top \mathbf{M}_{\mathbb{Z}}\boldsymbol{\varepsilon}_i\right|.$$

By the definition of $\widehat{\mathbf{X}}$, we have $\mathbf{M}_{\mathbb{Z}}\widehat{\mathbf{X}} = \widehat{\mathbf{X}}$. Thus, we have

$$\begin{aligned} J_1 =& \max_i \left|T^{-2}\boldsymbol{\varepsilon}_i^\top \widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top \boldsymbol{\varepsilon}_i\right| \leq \max_i T^{-1}\left|\boldsymbol{\varepsilon}_i^\top(\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}})\boldsymbol{\varepsilon}_i\right| + \max_i T^{-1}\left|\boldsymbol{\varepsilon}_i^\top \mathbf{P}_{\mathbf{X}}\boldsymbol{\varepsilon}_i\right| \\ \leq& O_p(\|\mathbf{P}_{\widehat{\mathbf{X}}} - \mathbf{P}_{\mathbf{X}}\|) + O_p(\sqrt{\log(N)/T}) = O_p(L^{-r} + \sqrt{L/T}). \end{aligned}$$

Step II. Now we consider J_2. Recall that $\mathrm{tr}\{\mathbf{X}^\top \mathbf{M}_{\widetilde{\mathbf{X}}}\mathbf{X}\} = O_p(1)$ (Wang, 2012). We obtain

$$\begin{aligned} J_2 &= \max_i T^{-1}\|\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i\|^2 \leq 2\max_i T^{-1}\|\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{P}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i\|^2 + 2\max_i T^{-1}\|\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{X}\boldsymbol{\lambda}_i\|^2 \\ &\leq 2\max_i T^{-1}\|\mathbf{P}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i\|^2 + 2\max_i T^{-1}\|(\mathbf{M}_{\widehat{\mathbf{X}}} - \mathbf{M}_{\mathbf{X}})\mathbf{X}\boldsymbol{\lambda}_i\|^2 \\ &= O_p(L/T(1 + \sqrt{\log(N)/L}) + L^{-2r} + L/T) = O_p(L^{-2r} + L/T). \end{aligned}$$

Step III. For J_3, we have

$$\max_i T^{-1}\boldsymbol{\rho}_i^\top \mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}_i \leq \max_i T^{-1}\boldsymbol{\rho}_i^\top \boldsymbol{\rho}_i = O_p(L^{-2r}).$$

Step IV. For J_4, by Cauchy-Schwarz inequality, we have

$$J_4 \leq \max_i \|T^{-1/2}\boldsymbol{\epsilon}_i^\top \mathbf{M}_{\mathbb{Z}}\|\|T^{-1/2}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i\| = O_p(\sqrt{L/T} + L^{-r}),$$

where the last equality is due to the order of J_2 and Lemma 6. Similarly, we have

$$J_5 \leq \max_i \|T^{-1/2}\boldsymbol{\epsilon}_i^\top\|\|T^{-1/2}\mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\boldsymbol{\rho}_i\| = O_p(L^{-r}).$$

For J_6,

$$J_6 \le \max_i \|T^{-1/2}\boldsymbol{\rho}_i^\top \mathbf{M}_{\mathbb{Z}}\| \|T^{-1/2}\mathbf{M}_{\widehat{\mathbf{X}}}\mathbf{M}_{\mathbb{Z}}\mathbf{X}\boldsymbol{\lambda}_i\| \le O_p(L^{-r}\sqrt{L/T} + L^{-2r}).$$

Combining all of these results above, we have $\widetilde{I}_2 = O_p(L^{-r} + \sqrt{L/T})$, which completes the proof of Lemma 8.

S2 Proof of Theorem 1

Denote $\boldsymbol{\theta} = (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1}\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T$ and let θ_t be the t-th element of $\boldsymbol{\theta}$. After simple calculation, we have

$$\begin{aligned}
\widehat{\alpha}_{i,ACA} &= (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1}\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{R}_i \\
&= \alpha_{i,ACA} + \boldsymbol{\varepsilon}_i^\top \boldsymbol{\theta} + \boldsymbol{\rho}_i^\top \boldsymbol{\theta} \\
&= \alpha_{i,ACA} + \sum_{t=1}^T \varepsilon_{it}\theta_t + \sum_{t=1}^T \boldsymbol{\rho}_{it}\theta_t.
\end{aligned}$$

Define $\mathbf{h} = (h_1, \cdots, h_T)^\top = \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T$, $V_{it} = \varepsilon_{it}h_t/\sigma_{ii}^{1/2}$, $U_{it} = V_{it} + \boldsymbol{\rho}_{it}h_t/\sigma_{ii}^{1/2}$. Let $\widehat{V}_{it} = V_{it}I(|V_{it}| \le \tau_T)$ for $t = 1, \cdots, T$ and $i = 1, \cdots, N$, where $\tau_T = 2\xi_t^{-1}\eta^{-1/2}\sqrt{\log(N+T)}$. Here, η is the constant defined in Assumption (A.2), and $\xi_t \to 0$ will be specified later. Further, define $W_i = \sum_{t=1}^T U_{it}/(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{1/2}$, $\widetilde{W}_i = \sum_{t=1}^T V_{it}/(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{1/2}$, and $\widehat{W}_i = \sum_{t=1}^T \widehat{V}_{it}/(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{1/2}$. We next prove the theorem in two steps. In the first step, we prove that for any $x \in \mathbb{R}$, as $N \to \infty$,

$$P\left[\max_{1\leq i\leq N} \widehat{W}_i^2 - 2\log(N) + \log\{\log(N)\} \leq x\right] \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\}.$$

In the second step, we show that the difference between $\widehat{W}_i$ and W_i is negligible such that $\max_i |W_i - \widehat{W}_i| = o_p(1/\log(N))$.

STEP I. We first prove that for any $x \in \mathbb{R}$, as $N \to \infty$,

$$P\left[\max_{1\leq i\leq N} \widehat{W}_i^2 - 2\log(N) + \log\{\log(N)\} \leq x\right] \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\}. \tag{S2.1}$$

Let $x_N = (2\log(N) - \log\{\log(N)\} + x)^{1/2}$ and $F = (\mathbf{f}_t, t = 1, \cdots, T)$. It follows from Lemma 2 that for any fixed $K \leq [N/2]$,

$$\sum_{k=1}^{2K}(-1)^{k-1}\sum_{1\leq i_1<\cdots<i_k\leq N} P(|\widehat{W}_{i_1}| \geq x_N, \cdots, |\widehat{W}_{i_k}| \geq x_N \big| F) \leq P(\max_{1\leq i\leq N} |\widehat{W}_i| \geq x_N \big| F)$$
$$\leq \sum_{k=1}^{2K-1}(-1)^{k-1}\sum_{1\leq i_1<\cdots<i_k\leq N} P(|\widehat{W}_{i_1}| \geq x_N, \cdots, |\widehat{W}_{i_k}| \geq x_N \big| F). \tag{S2.2}$$

Define $|\widehat{\mathbf{W}}|_{\min} = \min_{1\leq b\leq k} |\widehat{W}_{i_b}|$. Then, under Assumptions (A.2) and (A.4), by Theorem 1 in Zaitsev (1987), we have

$$\begin{aligned} P(|\widehat{\mathbf{W}}|_{\min} \geq x_N \big| F) \leq & P\left\{|\mathcal{Z}|_{\min} \geq x_N - \iota_T \log^{-1/2}(N)\right\} \\ & + c_1 k^{5/2} \exp\left\{-\frac{T^{1/2}\iota_T}{c_2 k^3 \tau_T \log^{1/2}(N)}\right\}, \end{aligned} \tag{S2.3}$$

where c_1 and c_2 are finite positive constants, ι_T is to be specified later, and $\mathcal{Z} = (\mathcal{Z}_{i_1}, \cdots, \mathcal{Z}_{i_k})^\top$ is a k-dimensional normal vector with the covariance matrix $\Pi_{\mathcal{Z}}$ satisfying $cov(\mathcal{Z}_{i_k}, \mathcal{Z}_{i_j}) = \Pi_{i_k i_j}$, where Π is the correlation matrix of $\mathbf{E}_t$. Because $\log(N) = o(T^{1/9})$, we can let $\iota_T \to 0$ sufficiently slowly, such that

$$c_1 k^{5/2} \exp\left\{-\frac{T^{1/2}\iota_T}{c_2 k^3 \tau_T \log(N)^{1/2}}\right\} = O(N^{-\xi}), \tag{S2.4}$$

for any large $\xi > 0$. It then follows from (S2.2), (S2.3), and (S2.4) that

$$P(\max_{1\le i\le N} |\widehat{W}_i| \ge x_N \big| F) \le \sum_{k=1}^{2K-1}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le N} P\left\{|\mathcal{Z}|_{min} \ge x_N - \iota_T \log(N)^{-1/2}\right\} + o(1). \tag{S2.5}$$

Similarly, using Theorem 1 in Zaitsev (1987) again, we can obtain

$$P(\max_{1\le i\le N} |\widehat{W}_i| \ge x_N \big| F) \ge \sum_{k=1}^{2K}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le N} P\left\{|\mathcal{Z}|_{min} \ge x_N - \iota_T \log(N)^{-1/2}\right\} - o(1). \tag{S2.6}$$

By (S2.5), (S2.6), and the proof of Theorem 1 in Cai et al. (2014), we have

$$P\left[\max_{1\le i\le N} \widehat{W}_i^2 - 2\log(N) + \log\{\log(N)\} \le x \big| F\right] \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\},$$

where the right term is not affected by F. Thus, (S2.1) is proved.

Step II. We next prove that $\max_i |W_i - \widehat{W}_i| = o_p\{1/\log(N)\}$. Note that $\max_i |W_i - \widehat{W}_i| \le \max_i(|\widetilde{W}_i - \widehat{W}_i| + |W_i - \widetilde{W}_i|)$. We then consider the two terms $\max_i |\widetilde{W}_i - \widehat{W}_i|$ and $\max_i |W_i - \widetilde{W}_i|$ separately. We first prove that $\max_{1\le i\le N}\left|\widetilde{W}_i-\widehat{W}_i\right| = o_p\{1/\log(N)\}$. We have that

$$P\left\{\max_{1\le i\le N}\left|\widetilde{W}_i-\widehat{W}_i\right| \ge \frac{1}{\log(N)}\right\}$$

$$\le P(\max_{1\le i\le N}\max_{1\le t\le T}|V_{it}|\ge \tau_T)$$

$$\le \sum_{1\le i\le N}\sum_{1\le t\le T} P(|V_{it}|\ge\tau_T).$$

For any $\xi_t\to 0$, we obtain

$$\begin{aligned}P(|V_{it}|\ge\tau_T) =& P(|\varepsilon_{it}/\sigma_{ii}^{1/2}|\ge|h_t^{-1}|\tau_T)\\ =& P(|\varepsilon_{it}/\sigma_{ii}^{1/2}|\ge|h_t^{-1}|\tau_T, |h_t^{-1}|\ge\xi_t)\\ &+P(|\varepsilon_{it}/\sigma_{ii}^{1/2}|\ge|h_t^{-1}|\tau_T, |h_t^{-1}|<\xi_t)\\ \le& P(|\varepsilon_{it}/\sigma_{ii}^{1/2}|\ge\xi_t\tau_T)+P(|h_t|\ge\xi_t^{-1}).\end{aligned}$$

We consider the above two parts separately. We first calculate $P(|\varepsilon_{it}/\sigma_{ii}^{1/2}|\ge \xi_t\tau_T)$. Under Assumption (A.2), by the Markov inequality, we then obtain

$$P(|\varepsilon_{it}/\sigma_{ii}^{1/2}|\ge\xi_t\tau_T)\le\mathcal{K}\exp(-\eta\xi_t^2\tau_T^2)=\mathcal{K}(N+T)^{-4}. \qquad \text{(S2.7)}$$

We next calculate $P(|h_t|\ge\xi_t^{-1})$. Define $\boldsymbol{\kappa}=(\mathbb{Z}^\top\mathbb{Z})^{-1}(\mathbb{Z}^\top\mathbf{1}_T)=(\mathbb{Z}^\top\mathbb{Z}/T)^{-1}(\mathbb{Z}^\top\mathbf{1}_T/T)=(\kappa_1,\cdots,\kappa_{(1+d)L})^\top$ and $\widetilde{\boldsymbol{\kappa}}=\left\{\mathbb{E}(\mathbb{Z}^\top\mathbb{Z})/T\right\}^{-1}\mathbb{E}(\mathbb{Z}^\top\mathbf{1}_T/T)$. By the definition of h_t, we have $P(|h_t|\ge\xi_t^{-1})=P(|1-\sum_{k=1}^{(1+d)L}\kappa_k Z_{tk}|\ge\xi_t^{-1})\le P(|\sum_{k=1}^{(1+d)L}\kappa_k Z_{tk}|\ge\xi_t^{-1}-1)$. Under Assumption (A.3) and $L=o(T^{1/3})$, by the result in Ma et al. (2020), we have $\|\mathbb{E}(\mathbb{Z}^\top\mathbf{1}_T/T)\|_\infty=O(L^{-1})$, $\|\left\{\mathbb{E}(\mathbb{Z}^\top\mathbb{Z})/T\right\}^{-1}\|_\infty=O(L)$, $\|\mathbb{Z}^\top\mathbf{1}_T/T-\mathbb{E}(\mathbb{Z}^\top\mathbf{1}_T/T)\|_\infty=O_p(\log T/\sqrt{TL})$

and $\|(\mathbb{Z}^\top\mathbb{Z}/T)^{-1} - \left\{\mathbb{E}(\mathbb{Z}^\top\mathbb{Z})/T\right\}^{-1}\|_\infty = O_p(L^2\log T/\sqrt{TL})$. Then, we obtain $\|\widetilde{\boldsymbol{\kappa}}\|_\infty \leq \|\left\{\mathbb{E}(\mathbb{Z}^\top\mathbb{Z})/T\right\}^{-1}\|_\infty \|\mathbb{E}(\mathbb{Z}^\top\mathbf{1}_T/T)\|_\infty = O(1)$, and

$$\begin{aligned}
\|\boldsymbol{\kappa} - \widetilde{\boldsymbol{\kappa}}\|_\infty \leq & \|\mathbb{Z}^\top\mathbf{1}_T/T - \mathbb{E}(\mathbb{Z}^\top\mathbf{1}_T/T)\|_\infty\|\left\{\mathbb{E}(\mathbb{Z}^\top\mathbb{Z})/T\right\}^{-1}\|_\infty \\
& + \|(\mathbb{Z}^\top\mathbb{Z}/T)^{-1} - \left\{\mathbb{E}(\mathbb{Z}^\top\mathbb{Z})/T\right\}^{-1}\|_\infty\|\mathbb{Z}^\top\mathbf{1}_T/T\|_\infty \\
= & O_p(L^{1/2}\log T/T^{1/2}).
\end{aligned}$$

Thus, $\boldsymbol{\kappa} \xrightarrow{p} \widetilde{\boldsymbol{\kappa}}$ as $T \to \infty$, and κ_i is bounded by a positive constant C_κ with the probability tending to one. By the fact that $|\sum_{l=1}^L B_l(t/T)|$ is bounded by a positive constant $C^{'}$ and Assumption (A.3) (ii), we have

$$\begin{aligned}
P(|h_t| \geq \xi_t^{-1}) &\leq P(|\sum_{k=1}^{(1+d)L}\kappa_k Z_{tk}| \geq \xi_t^{-1} - 1) \\
&\leq P(C_\kappa \sum_{k=1}^{(1+d)L}|Z_{tk}| \geq \xi_t^{-1} - 1) \\
&\leq P\left\{C_\kappa C^{'}(1 + \sum_{j=1}^{d}|f_{jt}|) \geq \xi_t^{-1} - 1\right\} \qquad \text{(S2.8)} \\
&\leq P\left\{C_\kappa C^{'}(1 + d\max_{1\leq j\leq d}|f_{jt}|) \geq \xi_t^{-1} - 1\right\} \\
&\leq \exp\left[-\left\{b_1^{-1}d^{-1}C_\kappa^{-1}C^{'-1}(\xi_t^{-1} - 1 - C_\kappa C^{'})\right\}^{a_1}\right].
\end{aligned}$$

Combining the results (S2.7) and (S2.8), we have

$$\begin{aligned}
&P\left\{\max_{1\leq i\leq N}\left|\widetilde{W}_i - \widehat{W}_i\right| \geq \frac{1}{\log(N)}\right\} \\
&\leq NT\left\{P(|\varepsilon_{it}/\sigma_{ii}^{1/2}| \geq \xi_t\tau_T) + P(|h_t| \geq \xi_t^{-1})\right\} \\
&\leq \mathcal{K}NT(T+N)^{-4} + NT\exp\left[-\left\{b_1^{-1}d^{-1}C_\kappa^{-1}C^{'-1}(\xi_t^{-1} - 1 - C_\kappa C^{'})\right\}^{a_1}\right] \to 0
\end{aligned}$$

(S2.9)

by setting $\xi_t = o\left[\{\log(N) + \log(T)\}^{-1/a_1}\right]$.

We next prove that $\max_{1\le i\le N}\left|\widetilde{W}_i - W_i\right| = o_p(1/\log(N))$. Under Assumption (A.3), there exist two finite positive constants $\widetilde{C}_z$ and $\breve{C}_z$, such that

$$\begin{aligned}
\max_i |W_i - \widetilde{W}_i| =& (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1/2} \max_i \sum_{t=1}^{T} |\rho_{it} h_t| / \sigma_{ii}^{1/2} \\
\le& \widetilde{C}_z T^{-1/2} \max_i \sum_{t=1}^{T} |(\rho_{it0} + \sum_{j=1}^{d} \rho_{ijt} f_{jt})(1 - \sum_{k=1}^{(1+d)L} \kappa_k Z_{tk})| \\
\le& \breve{C}_z T^{-1/2} L^{-r} \sum_{t=1}^{T} (1 + \sum_{j=1}^{d} |f_{jt}|)^2 \\
=& O_p(L^{-r} T^{1/2}) = o_p(1/\log(N)).
\end{aligned} \tag{S2.10}$$

Thus, by (S2.9) and (S2.10), we obtain $\max_i |W_i - \widehat{W}_i| \le \max_i(|\widetilde{W}_i - \widehat{W}_i| + |W_i - \widetilde{W}_i|) = o_p(1/\log(N))$.

Accordingly, we obtain

$$\left|\max_{1\le i\le N} W_i^2 - \max_{1\le i\le N} \widehat{W}_i^2\right| \le 2 \max_{1\le i\le N} |W_i| \max_{1\le i\le N} \left|W_i - \widehat{W}_i\right| + \max_{1\le i\le N} \left|W_i - \widehat{W}_i\right|^2 = o_p(1),$$

which immediately leads to

$$\begin{aligned}
&P\left[\max_{1\le i\le N} W_i^2 - 2\log(N) + \log\{\log(N)\} \le x\right] \\
&= P\left[\max_{1\le i\le N} \widehat{W}_i^2 - 2\log(N) + \log\{\log(N)\} \le x + \max_{1\le i\le N} \widehat{W}_i^2 - \max_{1\le i\le N} W_i^2\right] \\
&= P\left[\max_{1\le i\le N} \widehat{W}_i^2 - 2\log(N) + \log\{\log(N)\} \le x + o_p(1)\right].
\end{aligned} \tag{S2.11}$$

Recall that $W_i^2 = (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA})^2/\sigma_{ii}$. Then, by (S2.1) and (S2.11), we have

$$\begin{aligned}
&P\left[\max_{1\le i\le N} \frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA})^2}{\sigma_{ii}} - 2\log(N) + \log\{\log(N)\} \le x\right] \\
&\qquad\qquad \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\}.
\end{aligned} \tag{S2.12}$$

Accordingly, to prove the theorem, it suffices to show that

$$\left|\max_{1\le i\le N} \frac{\left(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA}\right)^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1}\widehat{\sigma}_{ii}} - \max_{1\le i\le N} \frac{\left(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA}\right)^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1}\sigma_{ii}}\right| = o_p(1). \tag{S2.13}$$

By the triangle inequality, (S2.13) follows as

$$\begin{aligned}
&\left|\max_{1\le i\le N} \frac{\left(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA}\right)^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1}\widehat{\sigma}_{ii}} - \max_{1\le i\le N} \frac{\left(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA}\right)^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1}\sigma_{ii}}\right| \\
&\le \max_{1\le i\le N} \left|\frac{\left(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA}\right)^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1}\sigma_{ii}}\right| \left|\frac{\sigma_{ii}}{\widehat{\sigma}_{ii}} - 1\right| \\
&\le \max_{1\le i\le N} \left|\frac{\left(\widehat{\alpha}_{i,ACA} - \alpha_{i,ACA}\right)^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1}\sigma_{ii}}\right| \max_{1\le i\le N} \left|\frac{\sigma_{ii}}{\widehat{\sigma}_{ii}} - 1\right|.
\end{aligned}$$

First, by (S2.12), we have

$$\max_{1\le i\le N}\left|\frac{(\widehat{\alpha}_{i,ACA}-\alpha_{i,ACA})^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1}\sigma_{ii}}\right| = O_p\{\log(N)\}. \tag{S2.14}$$

In addition, by Lemma 5, we have

$$\begin{aligned}\max_{1\le i\le N}\left|\frac{\sigma_{ii}}{\widehat{\sigma}_{ii}}-1\right| &\le \max_{1\le i\le N}\frac{1}{|\widehat{\sigma}_{ii}|}\max_{1\le i\le N}|\sigma_{ii}-\widehat{\sigma}_{ii}| \le C\max_{1\le i\le N}|\sigma_{ii}-\widehat{\sigma}_{ii}| \\ &= O_p\left\{\sqrt{L\log(N)/T}+L^{-r}\right\}.\end{aligned} \tag{S2.15}$$

Combining the results in (S2.14) and (S2.15), we have

$$\begin{aligned}&\left|\max_{1\le i\le N}\frac{(\widehat{\alpha}_{i,ACA}-\alpha_{i,ACA})^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1}\widehat{\sigma}_{ii}} - \max_{1\le i\le N}\frac{(\widehat{\alpha}_{i,ACA}-\alpha_{i,ACA})^2}{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)^{-1}\sigma_{ii}}\right| \\ &= O_p\left[\frac{L^{1/2}\{\log(N)\}^{3/2}}{T^{1/2}}+L^{-r}\log(N)\right] = o_p(1).\end{aligned}$$

Thus, (S2.13) is proved. Consequently, we have

$$\begin{aligned}P\left[\max_{1\le i\le N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)(\widehat{\alpha}_{i,ACA}-\alpha_{i,ACA})^2}{\widehat{\sigma}_{ii}} - 2\log(N)+\log\{\log(N)\}\le x\right] \\ \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\},\end{aligned} \tag{S2.16}$$

which completes the proof of Theorem 1.

S3 Proof of Theorem 2

Using the results proved in Theorem 1, we define

$$\mathrm{MCA} = \max_{1\le i\le N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)\widehat{\alpha}_{i,ACA}^2}{\widehat{\sigma}_{ii}} \text{ and } \mathrm{MCA}_1 = \max_{1\le i\le N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)(\widehat{\alpha}_{i,ACA}-\alpha_{i,ACA})^2}{\widehat{\sigma}_{ii}}.$$

Then, by the proof of Theorem 1, we have

$$P\left[\mathrm{MCA}_1 - 2\log(N) + \log\{\log(N)\} \leq x\right] \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\}.$$

This result implies that

$$P\left[\mathrm{MCA}_1 \leq 2\log(N) - \frac{1}{2}\log\{\log(N)\}\right] \to 1,$$

by setting $x = \frac{1}{2}\log\{\log(N)\}$. Note that $\max_{1\leq i\leq N}|\alpha_{i,ACA}/\sigma_{ii}^{1/2}| \geq \sqrt{8\log(N)/(c_m T)}$ by the definition of $\mathcal{U}(2\sqrt{2}/\sqrt{c_m})$. Consequently, by the triangle inequality, we have

$$\begin{aligned}
\mathrm{MCA} &\geq \max_{1\leq i\leq N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)\alpha_{i,ACA}^2}{2\widehat{\sigma}_{ii}} - \mathrm{MCA}_1 \\
&\geq \max_{1\leq i\leq N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)\alpha_{i,ACA}^2}{2\sigma_{ii}} - \mathrm{MCA}_1 - \max_{1\leq i\leq N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)\alpha_{i,ACA}^2}{2\sigma_{ii}}\frac{|\widehat{\sigma}_{ii}-\sigma_{ii}|}{\sigma_{ii}}.
\end{aligned}$$

According to Lemma 5, we have $\max_i|\widehat{\sigma}_{ii}-\sigma_{ii}| = O_p(L^{-r} + \sqrt{L/T}) = O_p(\sqrt{L/T})$ since $L^{-r}\sqrt{T}\log(N) = o_p(1)$. If $\max_i (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)\alpha_{i,ACA}^2/2\sigma_{ii} = o_p(\sqrt{T/L})$, then we have

$$\begin{aligned}
\mathrm{MCA} &\geq \max_{1\leq i\leq N}\frac{(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}}\mathbf{1}_T)\alpha_{i,ACA}^2}{2\sigma_{ii}} - \mathrm{MCA}_1 - o_p(1) \\
&\geq 4\log(N) - 2\log(N) + \frac{1}{2}\log\{\log(N)\} \\
&\geq 2\log(N) - \log\{\log(N)\} + q_\lambda,
\end{aligned}$$

which implies

$$P(\Psi_\lambda = 1) = P\Big\{\mathrm{MCA} \geq 2\log(N) - \log\big(\log(N)\big) + q_\lambda\Big\} \to 1.$$

as $N, T \to \infty$. If $(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)\alpha_{i,ACA}^2/2\sigma_{ii} \geq C_\alpha(\sqrt{T/L})$, for some constants C_α, then

$$\begin{aligned}\text{MCA} \geq & C_\alpha\sqrt{T/L} - 2\log(N) + \frac{1}{2}\log\log(N) - C_\alpha\sqrt{T/L}o_p(1) \\ \geq & C_\alpha/2\sqrt{T/L} \geq 2\log(N) - \log\{\log(N)\} + q_\lambda\end{aligned}$$

which completes the proof of Theorem 2.

S4 Proof of Proposition 1

One can easily verify that $\mathcal{S}(k_N, \varpi) \subseteq \mathcal{U}(2\sqrt{2}/\sqrt{c_m})$. Then, by the results of Theorem 2, the power of MCA converges to one for this case. Accordingly, it suffices to show that $P\big(\mathrm{T}_H > z_{1-\lambda}\big) \to \lambda$.

Under the proposition assumptions, one can also verify that assumptions (A1)–(A3) and conditions (C.1)–(C.2) used in Ma et al. (2020) are valid. Then, by the results of Theorem 2 in Ma et al. (2020), we have $\widehat{\sigma}_{NT}^{-1} = O_p(N^{1/2})$, and the HDA test statistic T_H follows the asymptotic normal distribution with mean γ^0 and variance one under the alternative hypotheses, where $\gamma^0 = \lim_{(N,T)\to\infty} \widehat{\sigma}_{NT}^{-1} N^{-1} T^{-1} \sum_{i=1}^N \alpha_{i,ACA}^2 (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^2$. Accordingly, as $\min(N, T) \to \infty$,

$$P\big(\mathrm{T}_H > z_{1-\lambda}\big) \to \Phi\big(\gamma^0 - z_{1-\lambda}\big),$$

where $\Phi(\cdot)$ denotes the cumulative distribution of a standard normal distribution. Therefore, to prove the theorem, it suffices to show $\gamma^0 \to 0$ when $\boldsymbol{\alpha}_{ACA} \in \mathcal{S}\big(k_N, \varpi\big)$ with the probability approaching one.

According to Lemma 4, we have

$$\begin{aligned}
\gamma^0 &\leq \lim_{(N,T)\to\infty} \widehat{\sigma}_{NT}^{-1} N^{-1} T^{-1} (C_m T)^2 \sum_{i=1}^{N} \alpha_{i,ACA}^2 \\
&\leq \lim_{(N,T)\to\infty} \widehat{\sigma}_{NT}^{-1} N^{-1} C_m^2 T N^p \left\{8N^{\varpi}/(c_m T)\right\} \max_{1\leq i\leq N} \sigma_{ii} \\
&\leq \lim_{(N,T)\to\infty} 8 C_m^2 c_m^{-1} (N^{p+\varpi-1/2}) \max_{1\leq i\leq N} \sigma_{ii} = 0.
\end{aligned}$$

Combining the results above, we have $P\big(\mathrm{T}_H > z_{1-\lambda}\big) \to \lambda$, which completes the entire proof.

S5 Proof of Theorem 3

Define $W_i^* = \sigma_{\boldsymbol{\epsilon},ii}^{-1/2} (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{1/2} (\widehat{\alpha}_{i,ACA} - (\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1} \mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \widehat{\mathbf{X}} \widehat{\boldsymbol{\lambda}}_i)$. We prove Theorem 3 in two steps. In Step I, we provide the distribution of $\max_i W_i^{*2}$. In Step II, we provide the distribution of $\widetilde{\mathrm{MCA}}$.

Step I. After a simple calculation, we have

$$\begin{aligned}
W_i^* &= \widetilde{\boldsymbol{\theta}}^\top \mathbf{R}_i - \widetilde{\boldsymbol{\theta}}^\top \widehat{\mathbf{X}} \widehat{\boldsymbol{\lambda}}_i - \widetilde{\boldsymbol{\theta}}^\top \mathbf{X}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \boldsymbol{\epsilon}_t + \widetilde{\boldsymbol{\theta}}^\top \mathbf{X}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \boldsymbol{\epsilon}_i \\
&= \widetilde{\boldsymbol{\theta}}^\top (\mathbf{I}_T - \mathbf{P}_{\mathbf{X}}) \boldsymbol{\epsilon}_i + \widetilde{\boldsymbol{\theta}}^\top \boldsymbol{\rho}_i + \widetilde{\boldsymbol{\theta}}^\top (\mathbf{X} \boldsymbol{\lambda}_i - \widehat{\mathbf{X}} \widehat{\boldsymbol{\lambda}}_i) + \widetilde{\boldsymbol{\theta}}^\top \mathbf{X}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \boldsymbol{\epsilon}_i,
\end{aligned}$$

where $\widetilde{\boldsymbol{\theta}} = (\sigma_{\boldsymbol{\epsilon},ii} \mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{-1/2} \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T$.

Similarly to in the proof of Theorem 1, we define $h^* = \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T$, $V_{it}^* = \epsilon_{it} h_t^* / \sigma_{\boldsymbol{\epsilon},ii}^{1/2}$. Let $\widehat{V}_{it}^* = V_{it}^* I(|V_{it}^*| \leq \tau_T)$ for $t = 1, \cdots, T$ and $i =$

$1, \cdots, N$, where $\tau_T = 2\xi_t^{-1}\sqrt{\log(N+T)}$, where $\xi_t \to 0$ is the same as that in Theorem 1. Further, define $\widetilde{W}_i^* = \sum_{t=1}^T V_{it}^*/(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{1/2}$, $\bar{W}_i^* = \sum_{t=1}^T V_{it}^*/(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{1/2}$, and $\widehat{W}_i^* = \sum_{t=1}^T \widehat{V}_{it}^*/(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)^{1/2}$. To prove that $\max_i W_i^{*2}$ converges to the type I extreme value distribution, we first prove that $\widehat{W}_i^{*2}$ converges to the type I extreme value distribution. Then we prove $\max_{1 \le i \le N} |W_i^* - \widehat{W}_i^*| = o_p(1/\log(N))$.

Using the same techniques as those in the proof of Theorem 1, we can prove that

$$P\left[\max_{1 \le i \le N} \widehat{W}_i^{*2} - 2\log(N) + \log\{\log(N)\} \le x\right] \to \exp\left\{-\frac{1}{\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right\}.$$

For conciseness, we omit the details here.

Next, we prove that $\max_i |W_i^* - \widehat{W}_i^*| = o_p(1/\log(N))$. Note that $\max_i |W_i^* - \widehat{W}_i^*| \le \max_i(|W_i^* - \widetilde{W}_i^*| + |\widetilde{W}_i^* - \bar{W}_i^*| + |\bar{W}_i^* - \widehat{W}_i^*|)$. Then, we consider the three terms separately.

For $\max_i |W_i^* - \widetilde{W}_i^*|$, we have

$$\begin{aligned}
\max_i |W_i^* - \widetilde{W}_i^*| =& \max_i |\widetilde{\boldsymbol{\theta}}^\top \boldsymbol{\rho}_i + \widetilde{\boldsymbol{\theta}}^\top (\mathbf{X}\boldsymbol{\lambda}_i - \widehat{\mathbf{X}}\widehat{\boldsymbol{\lambda}}_i) + \widetilde{\boldsymbol{\theta}}^\top \mathbf{X}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \boldsymbol{\epsilon}_i| \\
=& \max_i |\widetilde{\boldsymbol{\theta}}^\top \left(\mathbf{X}\boldsymbol{\lambda}_i - \mathbf{P}_{\widehat{\mathbf{X}}} \mathbf{X}\boldsymbol{\lambda}_i + (\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}})\boldsymbol{\epsilon}_i + \mathbf{M}_{\widehat{\mathbf{X}}} \boldsymbol{\rho}_i\right)| \\
=& \max_i |\widetilde{\boldsymbol{\theta}}^\top (\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}}) \mathbf{X}\boldsymbol{\lambda}_i| + \max_i |\widetilde{\boldsymbol{\theta}}^\top (\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}}) \boldsymbol{\epsilon}_i| \\
&+ \max_i |\widetilde{\boldsymbol{\theta}}^\top (\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}}) \boldsymbol{\rho}_i| \\
=& I^{(1)} + I^{(2)} + I^{(3)}.
\end{aligned}$$

For $I^{(1)}$, according to Assumption (A.7), we have that

$$I^{(1)} = \max_i \|\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}}\| \|\widetilde{\boldsymbol{\theta}}^\top \mathbf{X}\boldsymbol{\lambda}_i\|$$
$$\leq \max_i \|\boldsymbol{\lambda}_i\| \|\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}}\| \|\widetilde{\boldsymbol{\theta}}^\top \mathbf{X}\|.$$

According to the normal assumption, we have that $\|\widetilde{\boldsymbol{\theta}}^\top \mathbf{X}\| = O_p(1)$. Recall that $\|\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{X}}}\| = O_p(\sqrt{L/T} + L^{-r}) = O_p(\sqrt{L/T})$ since $L^{-r}\sqrt{T}\log(N) = o(1)$. Under Assumptions (A.6) and (A.7), we have $I^{(1)} = o_p(1/\log(N))$. Similarly, we have $I^{(2)} = o_p(1/\log(N))$. And $I^{(3)} = o_p(1/\log(N))$ can be obtained as the same as that in the proof of Theorem 1.

For $\max_i |\bar{W}_i^* - \widehat{W}_i^*|$, we have that

$$P\left\{\max_{1\leq i\leq N} \left|\bar{W}_i^* - \widehat{W}_i^*\right| \geq \frac{1}{\log(N)}\right\}$$
$$\leq P(\max_{1\leq i\leq N} \max_{1\leq t\leq T} |V_{it}^*| \geq \tau_T)$$
$$\leq \sum_{1\leq i\leq N} \sum_{1\leq t\leq T} P(|V_{it}^*| \geq \tau_T).$$

For any $\xi_t \to 0$, we obtain

$$\begin{aligned} P(|V_{it}^*| \geq \tau_T) =& P(|\epsilon_{it}/\sigma_{\boldsymbol{\epsilon},ii}^{1/2}| \geq |h_t^{*-1}|\tau_T) \\ =& P(|\epsilon_{it}/\sigma_{\boldsymbol{\epsilon},ii}^{1/2}| \geq |h_t^{*-1}|\tau_T, |h_t^{*-1}| \geq \xi_t) \\ &+ P(|\epsilon_{it}/\sigma_{\boldsymbol{\epsilon},ii}^{1/2}| \geq |h_t^{*-1}|\tau_T, |h_t^{*-1}| < \xi_t) \\ \leq& P(|\epsilon_{it}/\sigma_{\boldsymbol{\epsilon},ii}^{1/2}| \geq \xi_t\tau_T) + P(|h_t^*| \geq \xi_t^{-1}). \end{aligned}$$

We consider the above two parts separately. First, we calculate $P(|\epsilon_{it}/\sigma_{\boldsymbol{\epsilon},ii}^{1/2}| \geq$

$\xi_t \tau_T)$. According to the normal assumption, we have

$$P(|\epsilon_{it}/\sigma_{ii}^{1/2}| \geq \xi_t \tau_T) \leq \mathcal{K}(N+T)^{-4}. \tag{S5.1}$$

Next, we calculate $P(|h_t^*| \geq \xi_t^{-1})$. Recall that $\boldsymbol{\kappa} = (\mathbb{Z}^\top \mathbb{Z})^{-1}(\mathbb{Z}^\top \mathbf{1}_T) = (\mathbb{Z}^\top \mathbb{Z}/T)^{-1}(\mathbb{Z}^\top \mathbf{1}_T/T) = (\kappa_1, \cdots, \kappa_{(1+d)L})^\top$ and $\widetilde{\boldsymbol{\kappa}} = \left\{\mathbb{E}(\mathbb{Z}^\top \mathbb{Z})/T\right\}^{-1} \mathbb{E}(\mathbb{Z}^\top \mathbf{1}_T/T)$. According to the proof of Theorem 1, we have that $\boldsymbol{\kappa} \to \widetilde{\boldsymbol{\kappa}}$ with probability tending to 1, and κ_i is bounded. Following the definition of h_t^*, we have

$$\begin{aligned} &P(|h_t^*| \geq \xi_t^{-1}) = P(|1 - \sum_{k=1}^{(1+d)L} \kappa_k Z_{tk} - \mathbf{X}_t^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T| \geq \xi_t^{-1}) \\ \leq &P(|\sum_{k=1}^{(1+d)L} \kappa_k Z_{tk}| \geq (\xi_t^{-1} - 1)/2) + P(|\mathbf{X}_t^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T| \geq (\xi_t^{-1} - 1)/2). \end{aligned} \tag{S5.2}$$

The first term of (S5.2) is less than $\exp(-C_1^*(\xi_t^{-1} - 1 - C_2^*)^{a_1})$, for some constants C_1^* and C_2^*, which is proved in the proof of Theorem 1. For the second term of (S5.2), according to Assumption (A.5), we have that $\mathbb{E}(T^{-1}\mathbf{X}^\top \mathbf{X}) = \mathbf{I}_v$, $\mathrm{Var}(T^{-1/2}\mathbf{X}^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T) = T^{-1}\mathbb{E}(\mathrm{Var}(\mathbf{X}^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T | \mathbb{Z})) \leq T^{-1}\mathbb{E}(\mathbf{1}_T^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T)\mathbf{I}_v$. Denote $\varpi_1 = T^{-1/2}(T^{-1}\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{X}^\top \mathbf{M}_{\mathbb{Z}} \mathbf{1}_T$. Thus, the elements in ϖ_1 are bounded in probability. Under the normal assumption of $\mathbf{X}_t$, via 'Hoeffding inequality', we have

$$\begin{aligned}
&P(|\mathbf{X}_t^\top(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T| \geq (\xi_t^{-1}-1)/2)\\
\leq &P(|\sum_{k=1}^{v} T^{-1/2}X_{tv}\varpi_t| \geq (\xi_t^{-1}-1)/2)\\
\leq &P(|\sum_{k=1}^{v} T^{-1/2}X_{tv}C_3^*| \geq (\xi_t^{-1}-1)/2)\\
\leq &\exp(-C_4^*T(\xi_t^{-1}-1)^2),
\end{aligned}$$

for some constants C_3^* and C_4^*.

Then, we have

$$\begin{aligned}
&P\left\{\max_{1\leq i\leq N}\left|\bar{W}_i^* - \widehat{W}_i^*\right| \geq \frac{1}{\log(N)}\right\}\\
&\leq NT\left\{P(|\epsilon_{it}/\sigma_{\boldsymbol{\epsilon},ii}^{1/2}| \geq \xi_t\tau_T) + P(|h_t^*| \geq \xi_t^{-1})\right\}\\
&\leq \mathcal{K}NT(T+N)^{-4} + NT\exp(-C_1^*(\xi_t^{-1}-1-C_2^*)^{a_1}) + NT\exp(-C_4^*T(\xi_t^{-1}-1)^2) \to 0.
\end{aligned}$$

For $\max_i|\widetilde{W}_i^* - \bar{W}_i^*|$, we have

$$\max_i|\widetilde{W}_i^* - \bar{W}_i^*| = \max_i|\bar{W}_i^*||\frac{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{M}_{\mathbf{X}}\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T}{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T} - 1| \leq 3\log(N)|\frac{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{P}_{\mathbf{X}}\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T}{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T}|.$$

As $\frac{1}{T}\mathbf{X}^\top\mathbf{X} \xrightarrow{p} I_v$, and $\mathrm{Var}(\mathbf{X}^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T/\sqrt{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T})$ is bounded, we have that $\frac{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{P}_{\mathbf{X}}\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T}{\mathbf{1}_T^\top\mathbf{M}_{\mathbb{Z}}\mathbf{1}_T} = O_p(1/T)$. Then $\max_i|\widetilde{W}_i^* - \bar{W}_i^*| = o_p(1/\log(N))$.

Accordingly, we obtain

$$\left|\max_{1\leq i\leq N} W_i^{*2} - \max_{1\leq i\leq N}\widehat{W}_i^{*2}\right| \leq 2\max_{1\leq i\leq N}|W_i^*|\max_{1\leq i\leq N}\left|W_i^* - \widehat{W}_i^*\right| + \max_{1\leq i\leq N}\left|W_i^* - \widehat{W}_i^*\right|^2 = o_p(1),$$

which immediately leads to

$$
\begin{aligned}
&P\left[\max_{1\le i\le N} W_i^{*2} - 2\log(N) + \log\{\log(N)\} \le x\right] \\
&= P\left[\max_{1\le i\le N} \widehat{W}_i^{*2} - 2\log(N) + \log\{\log(N)\} \le x + \max_{1\le i\le N} \widehat{W}_i^{*2} - \max_{1\le i\le N} W_i^{*2}\right] \\
&= P\left[\max_{1\le i\le N} \widehat{W}_i^{*2} - 2\log(N) + \log\{\log(N)\} \le x + o_p(1)\right].
\end{aligned}
\tag{S5.3}
$$

STEP II. Recall that

$$\widetilde{\text{MCA}} = W_i^{*2}\sigma_{\boldsymbol{\epsilon},ii}/\hat{\sigma}_{\boldsymbol{\epsilon},ii}.$$

Thus to prove Theorem 3, we only need $\max_i |\sigma_{\boldsymbol{\epsilon},ii}/\hat{\sigma}_{\boldsymbol{\epsilon},ii} - 1| = o_p(1/\log(N))$, which can be proven using Lemma (8).

S6 Test portfolios

We use a total of 334 bivariate-sorted portfolios from model factors and prominent return anomalies as test assets in the empirical analyses. We start from a set of 100 portfolios: 25 5×5 portfolios sorted by size and book-to-market ratio, 25 5×5 portfolios sorted by size and operating profitability, 25 5×5 portfolios sorted by size and investment, and 25 5×5 portfolios sorted by size 5×5 and momentum. We then add to these 100 portfolios 234 additional portfolios obtained from anomalies. In particular, we try to include several prominent return anomalies, and they are as follows: (i) accruals (see Sloan, 1996), hereafter AC; (ii) market β (see Black et al., 1972; Fama and MacBeth, 1973), hereafter β; (iii) net share issues (see Ikenberry et al., 1995; Loughran and Ritter, 1995), hereafter NSI; (iv) daily variance (see Ang et al., 2006), hereafter Var; (v) daily residual variance (see Ang et al., 2006, Fu, 2009), hereafter RVar; (vi) short-term reversal (see Jegadeesh, 1990; Lehmann, 1990), hereafter STR; (vii) long-term reversal (see De Bondt and Thaler, 1985), hereafter LTR. Specifically, the sets of 234 additional portfolios include: 25 5×5 size-AC portfolios, 25 5×5 size-β portfolios, 35 5×7 size-NSI portfolios, 25 5×5 size-Var portfolios, 25 5×5 size-RVar portfolios, 25 5×5 size-STR portfolios, 25 5×5 size-LTR portfolios, and 49 industry portfolios employed by Ahmed et al. (2019). The data of the 334 test portfolios and the descriptions of the portfolio construction are available from Kenneth French's website.

S7 Real Data Application

We now apply the MCA test to assess the efficiency hypothesis in U.S.'s stock markets, and compare it with the HDA test from Ma et al. (2020) to illustrate the superiority of the MCA test in dealing with sparse alternatives.

S7.1 Data description

Consider the three-factor model proposed by Fama and French (1993),

$$R_{it} - R_{ft} = \alpha_{it} + \beta_{i1t}(\mathrm{MKT}_t - R_{ft}) + \beta_{i2t}\mathrm{SMB}_t + \beta_{i3t}\mathrm{HML}_t + \varepsilon_{it},$$

for $i = 1, \cdots, N$ and $t = 1, \cdots, T$, where R_{it} is the return for test asset i, R_{ft} is the risk-free rate, MKT_t, SMB_t, and HML_t are constituted market, size, and value factors at month t, respectively. A total of 334 bivariate-sorted portfolios are investigated, which have been commonly used in the financial literature (e.g., Fama and French (2015, 2016, 2018, 2020); Kozak et al. (2018); Ahmed et al. (2019); Feng et al. (2020)). A more detailed description of the portfolios is provided in Section S6 in the supplementary material. The time series of factors and portfolios are from Kenneth French's website for the period January 1981 to December 2020 (480 months).

S7.2 Are alphas and betas time-varying?

Before testing the Fama-French three-factor model, it is necessary to check whether the values of alphas and betas are indeed time-varying. Here we first divide the 40 years into four 10 year sub-periods. Then we examine the constancy of the alphas and betas over the full period and four sub-periods based on the constant coefficient (CC) test proposed by Ma et al. (2020). The results are provided in Table S.1, which contains the rejection

Table S.1: The rejection rates of the CC test based on the Fama-French three-factor model over the full period and four sub-periods at the three significance levels $\lambda = 0.1, 0.05, 0.01$.

Time period	$\lambda = 0.1$	$\lambda = 0.05$	$\lambda = 0.01$
1/1981–12/2020	0.940	0.934	0.919
1/1981–12/1990	0.488	0.407	0.338
1/1991–12/2000	0.862	0.805	0.757
1/2001–12/2010	0.823	0.707	0.650
1/2011–12/2020	0.497	0.392	0.329

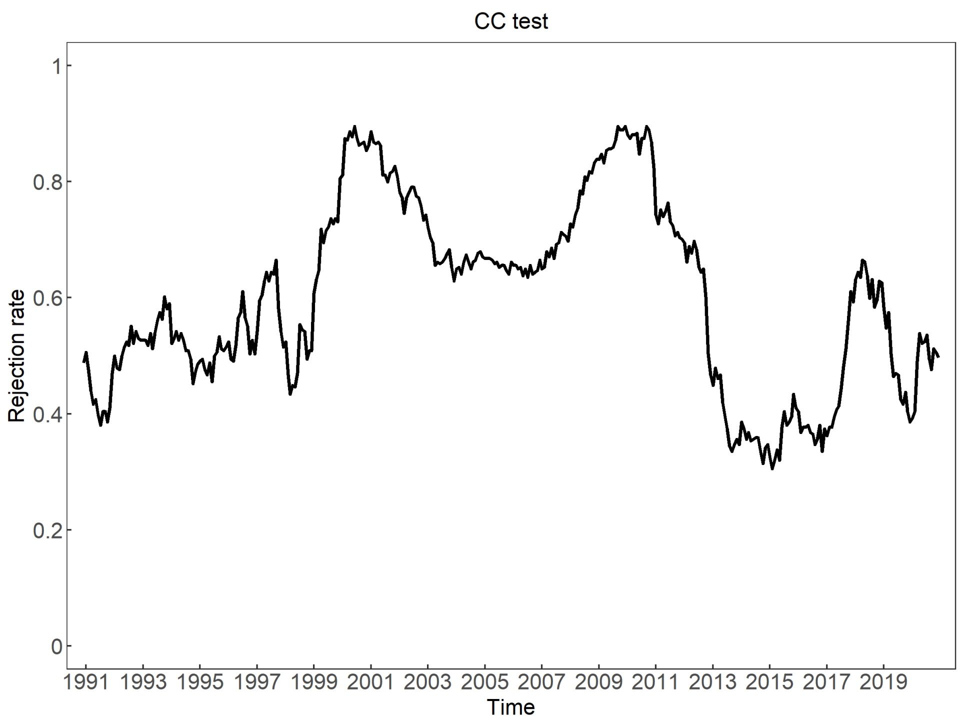

Figure 1: The rejection rates of the CC test based on the Fama-French three-factor model using 120-month rolling windows with a significance level of 5%.

ratios of the CC test at the three significant levels $\lambda = 0.1, 0.05, 0.01$. For the full period, the CC test indicates a decisive rejection of alphas and betas homogeneity in time series, and the rejection rates are close to 1 at all significance levels. Similar results have been found in the four subperiods, especially the subperiods 1/1991-12/2000 and 1/2001-12/2010. To further corroborate the conclusion, CC test is also conducted over the 120 month rolling windows. The results are provided in Figure 1. As shown in Figure 1, the rejection rates are always large, which supports that alphas and betas are time-varying. These evidences indicate that the conditional time-varying multi-factor model is more suitable than the traditional time-invariant multi-factor model.

S7.3 Empirical Results

We employ the MCA and HDA tests to assess the market efficiency of the U.S.'s stock market. We also report the results for our proposed $\widetilde{\text{MCA}}$ test from Section 3 (MCA1 hereafter). We first assess the market efficiency over the full 40-year period and each 10-year sub-period. Then, we study the changes in the market efficiency over time through the 120 month rolling windows. Here, the order of B-splines is set at 3 for all estimation windows, and the number of interior knots n is determined via BIC, as discussed in Section 2.2. The results are provided in Table S.2 and Figure 2. The number of significant nonzero elements in $\widehat{\boldsymbol{\alpha}}_{ACA}$, $|S|$, is defined as $|S| = \# \left|\widehat{\alpha}_{i,ACA}/\widehat{\sigma}_{ii}^{1/2}\right| \geq \sqrt{4\log(N)/T}$, $i = 1, \cdots, 334$.

For the full 40-year period and the 1/1981-12/1990 sub-period, the number of significant nonzero components in $\widehat{\boldsymbol{\alpha}}_{ACA}$ is sufficiently large, 17 and 38, respectively, which indicates that $\widehat{\boldsymbol{\alpha}}_{ACA}$ is medium dense. Both the MCA and HDA tests reject the market efficiency assumption at a significant level of 0.01. By contrast, in the 1/2001-12/2010 sub-period, only one portfolio with a significant nonzero $\widehat{\alpha}_{i,ACA}$ value is detected, which indicates the strong sparsity of $\widehat{\boldsymbol{\alpha}}_{ACA}$. In this sub-period, the p-values of the MCA and HDA tests are 0.000 and 0.189, respectively. The MCA test can reject the null hypothesis at a significant level of 0.01, while the HDA test cannot. The MCA test is more powerful. Furthermore, we find that MCA1 is highly powerful and rejects the null hypothesis in any cases.

The p-values of the MCA, MCA1, and HDA tests and the $|S|$ sequences over time are shown in Figure 2. The time interval can be divided into two parts: 1991-1999 and 1999-2020. In the initial time interval 1991-1999, the number of significant nonzero components is sufficiently large, between 10 and 38, which indicates that $\widehat{\boldsymbol{\alpha}}_{ACA}$ is medium dense. In this case, the proportion of non-zero alphas is between 3.0% and 11.4%, which is similar

Table S.2: The mean-variance efficiency tests are based on the Fama-French three-factor model. The values of test statistics for the MCA, MCA1, and HDA are denoted as MCA, MCA1, and HDA, respectively. The p values of the MCA, MCA1, and HDA tests are denoted as p(MCA), p(MCA1), and p(HDA), respectively.

Time period	MCA	p(MCA)	MCA1	p(MCA1)	HDA	p(HDA)	$\|S\|$
1/1981–12/2020	47.358	0.000	171.973	0.000	11.475	0.000	17
1/1981–12/1990	114.830	0.000	124.732	0.000	8.546	0.000	38
1/1991–12/2000	34.485	0.000	35.755	0.000	3.142	0.002	6
1/2001–12/2010	21.980	0.001	24.465	0.000	1.314	0.189	1
1/2011–12/2020	29.490	0.000	53.380	0.000	3.911	0.000	4

to the 8.3% obtained in the simulation study (for $N = 500$, and $p = 0.6$). All three tests tend to reject the null hypothesis. A large number of anomaly portfolio excess returns decline or even disappear starting in the mid 1990s, especially after 2000 (e.g., Jones and Pomorski (2017); Fama and French (2021)). In the next time interval 1999-2020, the number of significant nonzero components becomes small, which indicates that $\boldsymbol{\alpha}_{ACA}$ tends to be sparse. In this case, the proportion of non-zero alphas is 3.0%, close to the 2.4% obtained in the simulation study (for $N = 500$, and $p = 0.4$). The MCA and MCA1 tests tend to reject the null hypothesis, while the HDA tends to accept it. Based on these results, we conclude that the proposed MCA test is more inclined to reject the null hypothesis than the HDA test.

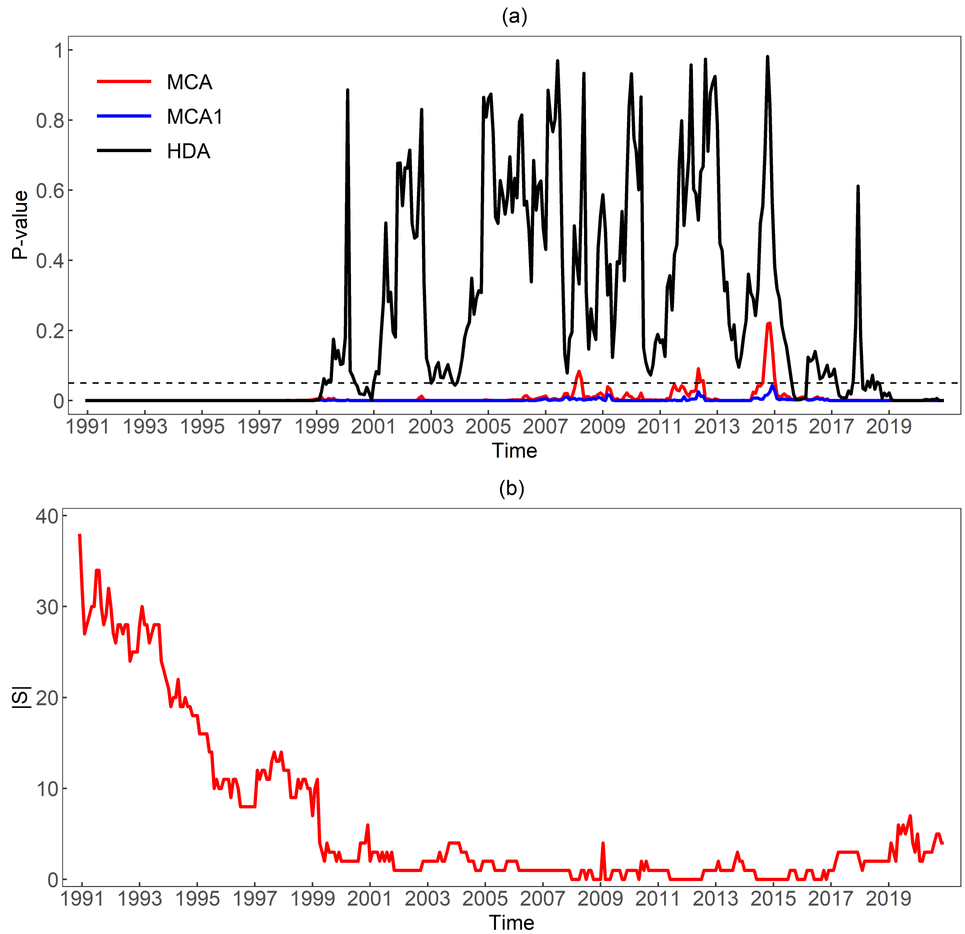

Figure 2: Dynamic movement of the mean-variance efficiency by testing the Fama-French three-factor model using a 120-month rolling window. The blue horizontal line corresponds to the significance level $\lambda = 0.05$. (a) Time variation in p-values of the MCA, MCA1, and HDA tests; (b) Number of significant nonzero components in $\widehat{\boldsymbol{\alpha}}_{ACA}$.

S8 Simulation Results for Maximum Conditional Alpha Test with Latent Factors

We conduct Monte Carlo experiments to illustrate the finite sample performance of the proposed test under conditional multi-factor models with latent factors. Specifically, we simulate the random error ε_{it} from a latent factor model, that is,

$$\varepsilon_{it} = \boldsymbol{\lambda}_i^\top \mathbf{X}_t + \epsilon_{it},$$

for $i = 1, \cdots, N$, $t = 1, \cdots, T$, where $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \cdots, \varepsilon_{Nt})^\top \in \mathbb{R}^N$, $\boldsymbol{\lambda}_i = (\lambda_{i1}, \cdots, \lambda_{iv})^\top \in \mathbb{R}^v$, and $\mathbf{X}_t = (X_{1t}, \cdots, X_{vt})^\top \in \mathbb{R}^v$ is the low dimension of v unknown latent factors. Here, $\mathbf{X}_t$ and $\boldsymbol{\lambda}_i$ are independently drawn from a standard normal distribution. We consider two different numbers of latent factors ($v = 1, 3$), that is, the latent factor models include one factor or three factors. The experimental results are summarized in Table S.3, which show that the empirical size and power performances of the proposed factor-adjusted MCA test are stable.

S9 Simulation Results of the MAX Test

To assess the performance of the MAX test (Feng et al., 2022) in testing alpha coefficients under sparse alternatives, simulation studies were conducted for Examples 1 and 2. The simulations considered three different sample sizes, three different numbers of test assets, and four different error distributions. The summarized results are presented in Table S.4, and they reveal significant size distortions in the MAX test. This outcome is expected since the MAX test is specifically designed for time-invariant factor loadings, while Examples 1 and 2 involve time-varying factor loadings.

S10 Simulation Results for Student-t Distribution Errors

The sizes and powers of the MCA and HDA tests with student-t distribution errors are summarized in Table S.5. We consider the standardized t_5 distribution: $\widetilde{e}_{it} \sim t(5)/\sqrt{5/3}$. We find that the simulation results are quantitatively similar to the results for normal, exponential, and mixture distribution errors.

Bibliography

Ahmed, S., Bu, Z., and Tsvetanov, D. (2019). Best of the best: A comparison of factor models. *Journal of Financial and Quantitative Analysis*, 54(4):1713–1758.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Black, F., Jensen, M. C., and Scholes, M. (1972). The capital asset pricing model: some empirical tests, in michael c. jensen, ed.: Studies in the theory of capital markets. *Praeger, New York*, page 79121.

Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, pages 349–372.

De Bondt, W. F. and Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, 40(3):793–805.

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

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

Fama, E. F. and French, K. R. (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies*, 29(1):69–103.

Fama, E. F. and French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, 128(2):234–252.

Fama, E. F. and French, K. R. (2020). Comparing cross-section and time-series factor models. *The Review of Financial Studies*, 33(5):1891–1926.

Fama, E. F. and French, K. R. (2021). The value premium. *The Review of Asset Pricing Studies*, 11(1):105–121.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Fan, J. and Han, X. (2017). Estimation of the false discovery proportion with unknown dependence. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 79(4):1143.

Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, 75(3):1327–1370.

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

Fu, F. (2009). Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics*, 91(1):24–37.

Ikenberry, D., Lakonishok, J., and Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics*, 39(2-3):181–208.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance*, 45(3):881–898.

Jones, C. S. and Pomorski, L. (2017). Investing in disappearing anomalies. *Review of Finance*, 21(1):237–267.

Kozak, S., Nagel, S., and Santosh, S. (2018). Interpreting factor models. *The Journal of Finance*, 73(3):1183–1223.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, 105(1):1–28.

Loughran, T. and Ritter, J. R. (1995). The new issues puzzle. *The Journal of Finance*, 50(1):23–51.

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

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, pages 289–315.

Tomioka, R. and Suzuki, T. Spectral norm of random tensors. *arXiv preprint arXiv:1407.1870, 2014.*

Wang, H. (2012). Factor profiled sure independence screening. *Biometrika*, 99(1):15–28.

Zaitsev, A. Y. (1987). On the gaussian approximation of convolutions under multidimensional analogues of sn bernstein's inequality conditions. *Probability theory and related fields*, 74(4):535–566.

Table S.3: Size and Power of the proposed MCA test under conditional multi-factor models with latent factors for Example 1 with normal distribution errors.

		size	power				
T	N		p=1	p=0.8	p=0.6	p=0.4	p=0.2
			$v = 1$				
120	50	0.052	0.552	0.746	0.864	0.926	0.954
	100	0.056	0.608	0.796	0.908	0.928	0.966
	200	0.048	0.708	0.872	0.944	0.966	0.97
	500	0.04	0.778	0.944	0.978	0.972	0.978
240	50	0.048	0.606	0.734	0.898	0.958	0.988
	100	0.054	0.782	0.868	0.958	0.976	0.996
	200	0.048	0.866	0.95	0.996	0.99	0.996
	500	0.056	0.962	0.992	0.992	0.998	0.998
360	50	0.048	0.66	0.77	0.918	0.982	0.998
	100	0.06	0.836	0.886	0.984	0.996	0.998
	200	0.05	0.942	0.978	1	0.998	1
	500	0.044	0.984	0.998	1	0.996	1
			$v = 3$				
120	50	0.046	0.466	0.708	0.828	0.906	0.956
	100	0.06	0.574	0.794	0.904	0.93	0.984
	200	0.062	0.662	0.854	0.944	0.954	0.978
	500	0.06	0.774	0.936	0.968	0.97	0.952
240	50	0.058	0.582	0.768	0.878	0.96	0.986
	100	0.056	0.792	0.846	0.958	0.982	0.998
	200	0.044	0.884	0.948	0.974	0.984	0.996
	500	0.056	0.974	0.988	0.998	0.996	1
360	50	0.054	0.654	0.776	0.914	0.976	0.994
	100	0.046	0.834	0.886	0.986	0.99	0.998
	200	0.052	0.928	0.968	0.992	0.992	0.998
	500	0.05	0.98	0.994	0.998	0.996	1

Table S.4: The empirical sizes of the MAX test for testing conditional alphas with a nominal level of 5%, where Normal, Exponential, Mixture and t distributions refer to the distribution from which the error term is generated.

		Example 1				Example 2			
T	N	Normal	Exp	Mix	t	Normal	Exp	Mix	t
120	50	0.19	0.23	0.172	0.198	0.246	0.346	0.28	0.278
	100	0.276	0.278	0.23	0.26	0.312	0.406	0.36	0.358
	200	0.392	0.38	0.336	0.308	0.398	0.528	0.476	0.47
	500	0.44	0.436	0.442	0.41	0.58	0.624	0.582	0.578
240	50	0.174	0.204	0.194	0.196	0.228	0.286	0.242	0.246
	100	0.252	0.242	0.222	0.242	0.316	0.326	0.346	0.306
	200	0.278	0.304	0.246	0.306	0.38	0.432	0.376	0.39
	500	0.402	0.428	0.388	0.334	0.494	0.524	0.494	0.5
360	50	0.194	0.182	0.188	0.18	0.242	0.276	0.254	0.216
	100	0.236	0.26	0.236	0.22	0.326	0.29	0.272	0.3
	200	0.282	0.332	0.24	0.288	0.378	0.42	0.39	0.356
	500	0.384	0.346	0.336	0.35	0.496	0.51	0.552	0.514

Table S.5: Size and power of MCA and HDA tests from Examples 1 and 2 with t distribution errors.

		size		power(dense)		power(medium dense)				power(sparse)			
				p=1		p=0.8		p=0.6		p=0.4		p=0.2	
T	N	MCA	HDA	MCA	HDA	MCA	HDA	MCA	HDA	MCA	HDA	MCA	HDA
Example 1													
120	50	0.046	0.058	0.762	0.894	0.732	0.828	0.712	0.674	0.802	0.616	0.848	0.458
	100	0.052	0.058	0.84	0.928	0.81	0.872	0.808	0.726	0.77	0.42	0.918	0.484
	200	0.05	0.062	0.904	0.954	0.828	0.868	0.852	0.71	0.888	0.39	0.9	0.23
	500	0.049	0.054	0.898	0.946	0.894	0.898	0.88	0.644	0.87	0.24	0.918	0.138
240	50	0.046	0.05	0.758	0.892	0.748	0.852	0.752	0.624	0.826	0.526	0.95	0.446
	100	0.053	0.044	0.81	0.932	0.814	0.874	0.864	0.684	0.848	0.334	0.936	0.386
	200	0.046	0.038	0.898	0.956	0.854	0.894	0.836	0.58	0.906	0.274	0.966	0.16
	500	0.052	0.042	0.934	0.982	0.928	0.91	0.898	0.544	0.922	0.152	0.966	0.08
360	50	0.042	0.05	0.76	0.892	0.742	0.83	0.738	0.596	0.842	0.452	0.94	0.32
	100	0.038	0.036	0.83	0.94	0.848	0.896	0.852	0.68	0.886	0.33	0.964	0.3
	200	0.046	0.048	0.89	0.948	0.864	0.882	0.868	0.542	0.91	0.196	0.982	0.124
	500	0.052	0.04	0.94	0.986	0.918	0.92	0.924	0.484	0.954	0.126	0.982	0.074
Example 2													
120	50	0.04	0.056	0.742	0.886	0.678	0.826	0.644	0.648	0.672	0.542	0.802	0.47
	100	0.038	0.051	0.82	0.924	0.78	0.856	0.76	0.718	0.73	0.452	0.86	0.454
	200	0.052	0.045	0.868	0.948	0.808	0.862	0.77	0.624	0.786	0.342	0.874	0.246
	500	0.056	0.037	0.908	0.954	0.876	0.884	0.806	0.512	0.83	0.238	0.866	0.136
240	50	0.036	0.044	0.728	0.89	0.698	0.822	0.666	0.642	0.724	0.496	0.832	0.382
	100	0.052	0.038	0.774	0.924	0.818	0.888	0.772	0.66	0.762	0.326	0.906	0.344
	200	0.042	0.044	0.888	0.954	0.824	0.89	0.854	0.628	0.826	0.226	0.944	0.14
	500	0.046	0.04	0.914	0.964	0.906	0.904	0.88	0.488	0.892	0.14	0.944	0.07
360	50	0.048	0.052	0.702	0.886	0.716	0.848	0.69	0.608	0.718	0.42	0.852	0.332
	100	0.052	0.044	0.824	0.926	0.808	0.892	0.76	0.646	0.764	0.276	0.942	0.296
	200	0.048	0.048	0.878	0.96	0.86	0.91	0.826	0.564	0.834	0.194	0.948	0.124
	500	0.062	0.038	0.94	0.982	0.904	0.916	0.89	0.444	0.906	0.098	0.95	0.064