Factor Models with Local Factors - Determining the Number of Relevant Factors

Simon Freyaldenhoven, *Federal Reserve Bank of Philadelphia*[*]

April 9, 2021

Online Appendix

List of Tables

List of Figures

[*]Email: simon.freyaldenhoven@phil.frb.org

A Monte Carlo Simulation

A.1 DGP for Table 1

$$\underset{(500\times n)}{X} = \underset{(500\times 2)}{F}\underset{(2\times n)}{\Lambda^T} + \underset{(500\times n)}{e}.$$

We observe a panel with $T = 500$, where the cross-sectional dimension varies across simulations (see Table 1 in the main paper). The variables exhibit the following factor structure: With $F_{tk} \overset{i.i.d.}{\sim} N(0,1), k = 1, 2$ for all t, Λ is a matrix of ones and zeros such that:

$$\begin{aligned} X_j &= F_1 + e_j, && \text{for } j \in \mathcal{A}_2^c \\ X_j &= F_1 + F_2 + e_j, && \text{for } j \in \mathcal{A}_2. \end{aligned}$$

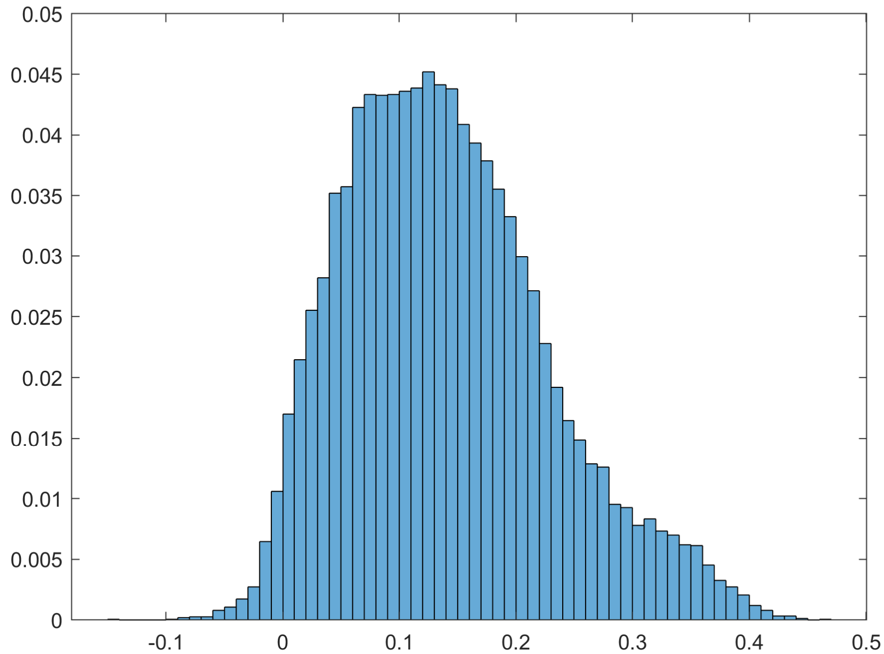
The cardinality of $\mathcal{A}_2$ is varied from $n^{\frac{1}{4}}$ to $n^{\frac{3}{4}}$.

Finally, we allow the idiosyncratic errors to exhibit both cross-sectional as well as intertemporal correlation. We follow Onatski [2010] and model the errors as follows:

$$\begin{aligned} e_{ti} &= \rho e_{t-1,i)} + (1-\rho^2)^{1/2} v_{ti} \\ v_{ti} &= \beta v_{t,i-1)} + (1-\beta^2)^{1/2} u_{ti}, \quad u_{ti} \overset{i.i.d.}{\sim} N(0,1), \end{aligned}$$

with $(\rho, \beta) = (0.3, 0.3)$ to allow for modest correlations in the error terms.

A.2 Empirical Violations of Orthogonality Assumptions

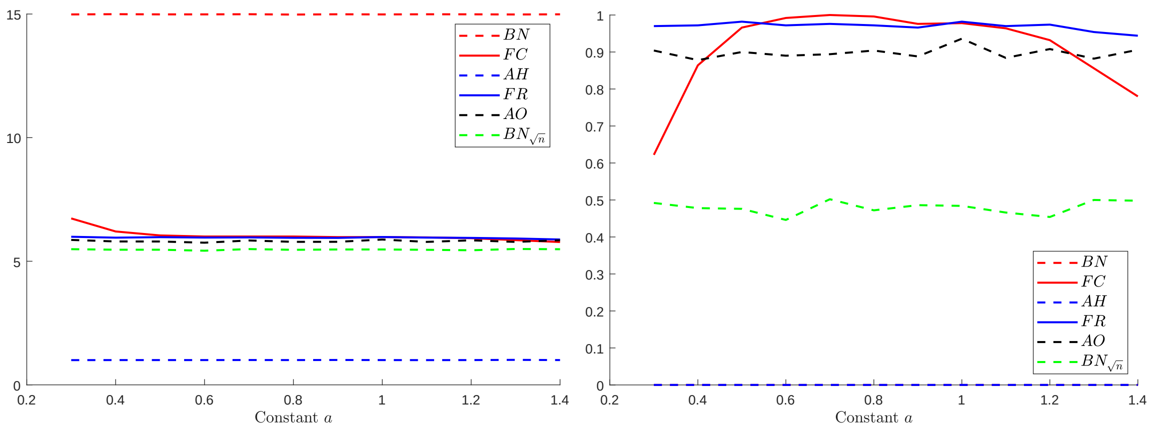


Online Appendix Figure 1: Empirical frequency for $\frac{\lambda_{\cdot k}\lambda_{\cdot l}}{\|\lambda_{\cdot k}\|\|\lambda_{\cdot l}\|}$ for $l \neq k$, $l, k \in \{1, \ldots, 6\}$, in baseline DGP. Figure based on 1000 realizations. Under Assumption 1(b), all off-diagonal entries in $\Lambda'\Lambda$ should be zero.

A.3 Robustness to tuning parameters

Since all our results involve asymptotic rate arguments, multiplicative constants have no effect on our results. Online Appendix Figure 2 shows the finite sample implications for the choice of a in $g(n) = a\sqrt{loglog(n)}$ Throughout the paper, we set $a = 0.7$, in which case $g(n) = 0.7\sqrt{loglog(n)} \approx 1$ for most relevant sample sizes.

Online Appendix Figure 2 illustrates that our results are robust to the choice of a. Note that only estimators based on $\hat{\Upsilon}_k^2$ (TR and TC) are affected by the choice of a. With the possible exception of the ED estimator of Onatski [2010], our proposed estimators dominate existing estimators for all values of a considered.

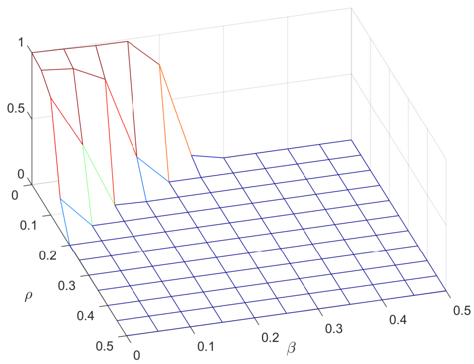


(a) Average estimate for number of factors according to EC, TC, ER, TR, ED and $EC_{\sqrt{n}}$

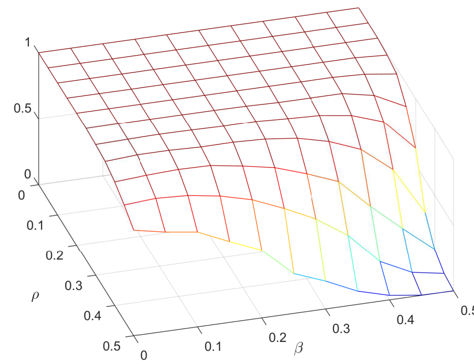
(b) % of simulations correctly estimating number of factors r_1 according to EC, TC, ER, TR, ED and $EC_{\sqrt{n}}$

Online Appendix Figure 2: Empirical behavior of estimators as the tuning parameter a is varied in $g(n) = a\sqrt{loglog(n)}$ to determine the number of elements z in the partial sum used to construct $\hat{\Upsilon}_k^u$. Note that EC, ER, ED and $EC_{\sqrt{n}}$ are unaffected by this choice. Data generated by baseline DGP, with $(n, T) = (300, 500)$, $(\rho, \beta) = (0.3, 0.1)$, $\theta = 1$, and $r_1 = 6$. Figure based on 500 replications.

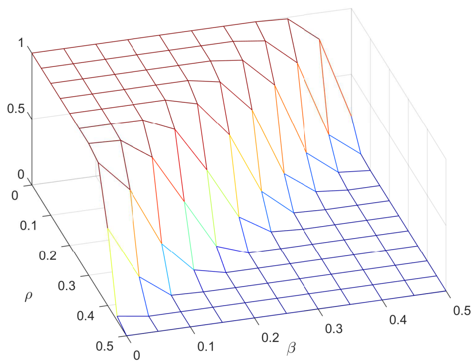
A.4 Results Under Unfavorable DGP



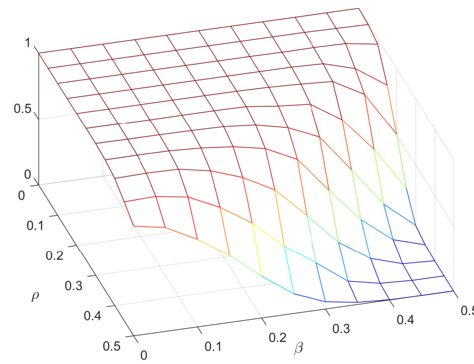
(a) Thresholding based on $\hat{\Upsilon}^0_k$ (EC)



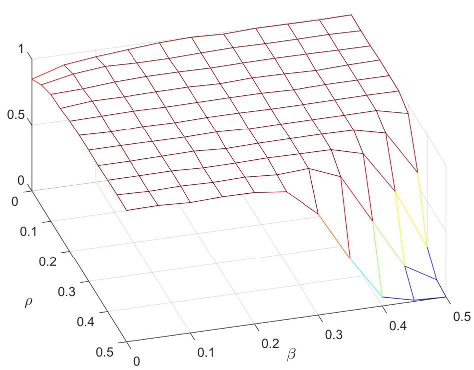
(b) Thresholding based on $\hat{\Upsilon}^2_k$ (TC)



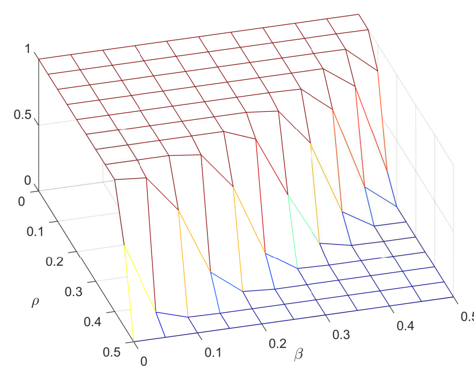
(c) Maximum ratio of two subsequent values of $\hat{\Upsilon}^0_k$ as in Ahn and Horenstein [2013] (ER)



(d) Maximum ratio of two subsequent values of $\hat{\Upsilon}^2_k$ (TR)

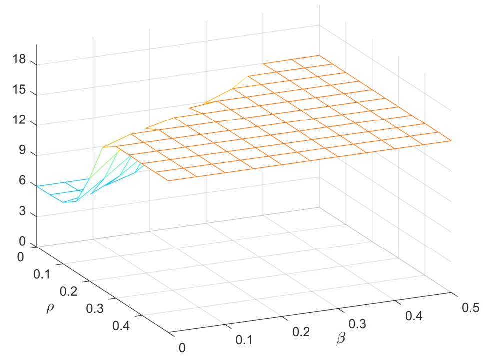


(e) Difference of two subsequent eigenvalues as in Onatski [2010] (ED)

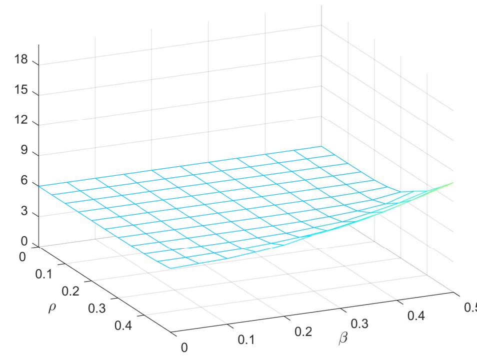


(f) Thresholding based on $\hat{\Upsilon}^0_k$ $(EC_{\sqrt{n}})$

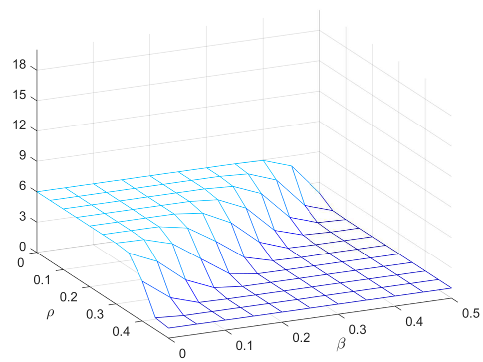
Online Appendix Figure 3: Percentage of simulations correctly estimating the true number of factors as both cross-sectional and intertemporal correlation is varied. All factors affect all outcomes. $r = 6$, $(n, T) = (300, 500)$, $\theta = 1.5$. For each entry in Λ, $\lambda_{ik} = 1 + \nu_{ik}$, where $\eta_{ik} \sim N(0, 1)$. Figure based on 500 replications.



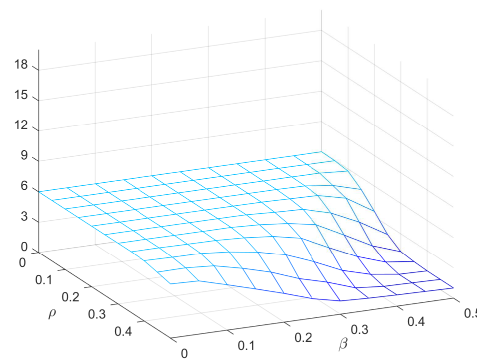
(a) Thresholding based on $\hat{\Upsilon}^0_k$ (EC)



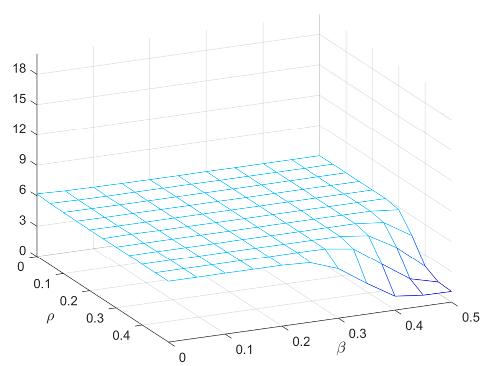
(b) Thresholding based on $\hat{\Upsilon}^2_k$ (TC)



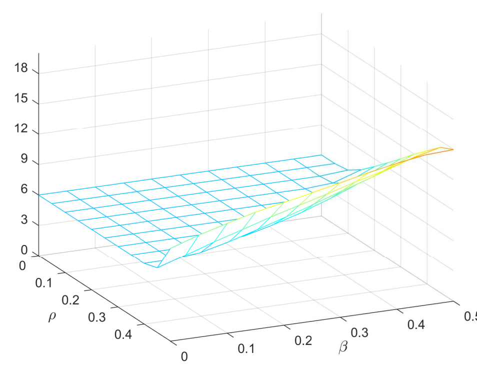
(c) Maximum ratio of two subsequent values of $\hat{\Upsilon}^0_k$ as in Ahn and Horenstein [2013] (ER)



(d) Maximum ratio of two subsequent values of $\hat{\Upsilon}^2_k$ (TR)

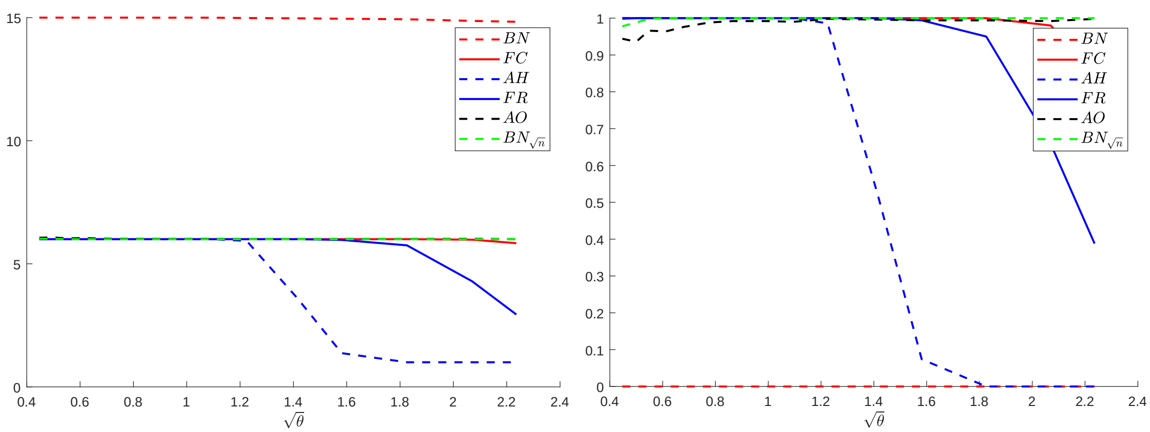


(e) Difference of two subsequent eigenvalues as in Onatski [2010] (ED)



(f) Thresholding based on $\hat{\Upsilon}^0_k$ $(EC_{\sqrt{n}})$

Online Appendix Figure 4: Average number of factors estimated as both cross-sectional and intertemporal correlation is varied. All factors affect all outcomes. $r = 6$, $(n, T) = (300, 500)$, $\theta = 1.5$. For each entry in Λ, $\lambda_{ik} = 1 + \nu_{ik}$, where $\eta_{ik} \sim N(0, 1)$. Figure based on 500 replications.



(a) Average estimated number of factors according to TC, TR, EC, ER and ED

(b) % of correctly estimated number of factors according to TC, TR, EC, ER and ED

Online Appendix Figure 5: Empirical behavior of estimators as the relative variance of idiosyncratic noise increases. All factors affect all outcomes. $r = 6$, $(n, T) = (300, 500)$, $(\rho, \beta) = (0.3, 0.1)$. For each entry in Λ, $\lambda_{ik} = 1 + \nu_{ik}$, where $\eta_{ik} \sim N(0, 1)$. Figure based on 500 replications.

n	T	ER	TR	PC	$PC_{\sqrt{n}}$	TC	ED
100	100	1.06 / 0.00	1.82 / 0.00	15 / 0.00	9.78 / 0.00	8.94 / 0.00	1.42 / 0.00
100	150	1.04 / 0.00	1.64 / 0.02	15 / 0.00	7.32 / 0.10	6.96 / **0.30**	1.72 / 0.00
150	100	1 / 0.00	1.92 / 0.00	15 / 0.00	8.66 / 0.00	10.2 / 0.00	1.62 / 0.00
150	250	1.02 / 0.00	2.4 / 0.04	15 / 0.00	4.04 / 0.02	5.66 / **0.54**	2.06 / 0.02
150	500	1.02 / 0.00	3.8 / 0.24	14.6 / 0.00	3.54 / 0.00	5.36 / **0.42**	3.1 / 0.12
300	250	1 / 0.00	4.78 / 0.24	15 / 0.00	3.08 / 0.00	5.68 / **0.46**	2.36 / 0.00
300	500	1 / 0.00	5.56 / 0.78	15 / 0.00	2.94 / 0.00	5.86 / **0.86**	4.66 / 0.50
300	750	1 / 0.00	5.96 / **0.96**	13 / 0.00	3.04 / 0.00	5.96 / **0.96**	5.74 / 0.86
500	250	1 / 0.00	4.5 / 0.16	15 / 0.00	2.88 / 0.00	5.9 / **0.54**	2.86 / 0.02
500	500	1 / 0.00	5.98 / **0.98**	15 / 0.00	2.86 / 0.00	5.98 / **0.98**	5.22 / 0.66
500	750	1 / 0.00	5.98 / **0.98**	14.1 / 0.00	2.9 / 0.00	5.98 / **0.98**	5.94 / 0.94
1000	1e+03	1 / 0.00	6 / **1.00**	13.3 / 0.00	2.88 / 0.00	6 / **1.00**	6 / **1.00**

Online Appendix Table 1: Performance of different estimators as the sample size is varied along a grid of (n, T). All factors affect all outcomes, with $(\rho, \beta) = (0.3, 0.1)$, $\theta = 1.5$, and $r = 6$. Each entry depicts a combination $\hat{r}/\%$, where $\hat{r}$ is the average number of estimated factors, and $\%$ is the percentage correctly classifying $r = 6$. In each row, the highest percentage is highlighted. For each entry in Λ, $\lambda_{ik} = 1 + \nu_{ik}$, where $\eta_{ik} \sim N(0, 1)$. In each row, the highest percentage is highlighted. Table based on 500 replications.

B Mathematical Appendix

B.1 Auxiliary Lemmata

Lemma 4. *Under Assumptions 1-3, for all n and T,*

$$\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\mathbb{E}\left(\frac{e_s'e_t}{n}\right)^2 \leq C.$$

Proof. See Lemma 1(i) in Bai and Ng [2002], using Assumption 3 (b). □

Lemma 5. *Under Assumptions 1-4, for any fixed K, let A be a $T \times K$ matrix A such that $A'A = TI_K$. Define $\alpha_1 = \max_k \alpha_k,\ k = 1, 2, \ldots, K$.*

Then:

$$\frac{1}{T^2}\left|trace(A'F\Lambda'e'A)\right| = O_p(n^{\frac{1}{2}\alpha_1}).$$

Proof.

$$\begin{aligned}
\left|trace(A'F\Lambda'e'A)\right| &= \left|trace(A'[\sum_{k=1}^{r} F_k\lambda_{\cdot k}'e']A)\right| \\
&= \left|\sum_{k=1}^{r} trace(A'[F_k\lambda_{\cdot k}'e']A)\right| \\
&\leq \left|\sum_{k=1}^{r} \|A'\|\|F_k\|\|\lambda_{\cdot k}'e'\|\|A\|\right| \\
&= \left|\sum_{k=1}^{r} \|A\|^2\|F_k\|\sqrt{\sum_t(\sum_i \lambda_{ik}e_{ti})^2}\right|
\end{aligned}$$

By Assumption 4(a), the innermost sum is $O_p(n^{\frac{1}{2}\alpha_k})$. We conclude:

$$\begin{aligned}
\frac{1}{T^2}\left|trace(A'F\Lambda'e'A)\right| &\leq \left|\sum_{k=1}^{r} \|\frac{1}{\sqrt{T}}A\|^2\|\frac{1}{\sqrt{T}}F_k\|\sqrt{\frac{1}{T}\sum_t(\sum_i \lambda_{ik}e_{ti})^2}\right| \\
&= \left|\sum_{k=1}^{r} O_p(1) \times O_p(n^{\frac{1}{2}\alpha_k})\right| \\
&= O_p(n^{\frac{1}{2}\alpha_1}),
\end{aligned}$$

which completes the proof. We note that in most cases at least one factor will be strong, corresponding to $\alpha_1 = 1$. In that case, the above rate becomes $O_p(\sqrt{n})$.

$\square$

Lemma 6. *Under Assumptions 1- 4, for any fixed K, let A be a $T \times K$ matrix A such that $A'A = TI_K$. Define $\alpha_1 = \max_k \alpha_k$, k=1, 2, ..., K.*

$$\sup_A (A' \frac{XX'}{T} A - A' \frac{F\Lambda'\Lambda F'}{T} A) = O_p(n^{\frac{1}{2}\alpha_1})$$

Proof.

$$\begin{aligned}
\sup_A \big(A' \frac{XX'}{T^2} A - A' \frac{F\Lambda'\Lambda F'}{T^2} A\big) &= \sup_A A' \big(\frac{e\Lambda F'}{T^2} + \frac{F\Lambda' e'}{T^2} + \frac{ee'}{T^2}\big) A \\
&\leq \sup_A A' \big(\frac{e\Lambda F'}{T^2} + \frac{F\Lambda' e'}{T^2}\big) A + \sup_A A' \frac{ee'}{T^2} A \\
&= O_p(n^{\frac{1}{2}\alpha_1}) + O_p(1),
\end{aligned}$$

where the last equality follows from Lemma 5 and Assumption 3(e). $\square$

Lemma 7. *Denoting the singular value decomposition of $\frac{1}{\sqrt{T}} X$ by $U\Sigma V'$, let $\hat{F}_1 = \sqrt{T} U_1$. Then, under Assumptions 1- 4:*

$$\frac{\hat{F}_1' F_1}{T} = 1 + O_p(n^{-\frac{1}{2}\alpha_1}),$$

and for $l = 1, \ldots, r$

$$\frac{\hat{F}_1' F_l}{T} = O_p(n^{-\frac{1}{4}\alpha_1}).$$

Proof. Decompose $\hat{F}_1$ as follows:

$$\hat{F}_1 = F(\frac{F'F}{T})^{-\frac{1}{2}} \xi_1 + \hat{V} \text{ such that } \hat{V}' F = 0. \tag{1}$$

Since $\frac{\hat{F}_1'\hat{F}_1}{T} = \xi_1'\xi_1 + \frac{\hat{V}'\hat{V}}{T}$, this implies $\xi_1'\xi_1 \leq 1$. Further,

$$\begin{aligned}
\hat{F}_1' \frac{F\Lambda'\Lambda F'}{T^2} \hat{F}_1 &= [F(\frac{F'F}{T})^{-\frac{1}{2}}\xi_1 + \hat{V}]' \Big(\frac{F\Lambda'\Lambda F'}{T^2}\Big) [F(\frac{F'F}{T})^{-\frac{1}{2}}\xi_1 + \hat{V}] \\
&= \xi_1' (\frac{F'F}{T})^{\frac{1}{2}} \Lambda'\Lambda (\frac{F'F}{T})^{\frac{1}{2}} \xi_1 \\
&= \xi_1' I_r \Lambda'\Lambda I_r \xi_1 \\
&= \xi_1' D_r^{(n)} \xi_1,
\end{aligned}$$

by Assumptions 1(b) and (c), and

$$\begin{aligned}
\frac{1}{T^2}\big(\hat{F}_1'F\Lambda'\Lambda F'\hat{F}_1 - \hat{F}_1'XX'\hat{F}_1\big) &= \frac{1}{T^2}\big(\hat{F}_1'F\Lambda'\Lambda F'\hat{F}_1 - F_1'F\Lambda'\Lambda F'F_1\big) \\
&\quad + \frac{1}{T^2}\big(F_1'F\Lambda'\Lambda F'F_1 - \hat{F}_1'XX'\hat{F}_1\big) \\
&= O_p(n^{\frac{1}{2}\alpha_1}) \qquad \text{(by Lemma 6).}
\end{aligned}$$

The second term on the RHS is simply the difference between the largest eigenvalue of XX'/T and $F\Lambda'\Lambda F'/T$. Following the reasoning in the proof of Theorem 1 that difference is $O_p(n^{\frac{1}{2}\alpha_1})$. It follows that the first term on the RHS is also $O_p(n^{\frac{1}{2}\alpha_1})$. We therefore obtain

$$\begin{aligned}
\frac{1}{T^2}\big(\hat{F}_1'F\Lambda'\Lambda F'\hat{F}_1 - F_1'F\Lambda'\Lambda F'F_1\big) &= \xi_1' D_r^{(n)}\xi_1 - d_1 && (2) \\
&= (\xi_{11}^2 - 1)d_1 + \sum_{l=2}^{r} \xi_{1l}^2 d_l && (3) \\
&= O_p(n^{\frac{1}{2}\alpha_1}). && (4)
\end{aligned}$$

We distinguish two cases. Case 1: $(\xi_{11}^2 - 1) = O_p(n^{-\frac{1}{2}\alpha_1})$. Since $\xi_{1l}^2 > 0\ \forall l$, and $\xi_1'\xi_1 \leq 1$, it follows that also $\xi_{1l}^2 = O_p(n^{-\frac{1}{2}\alpha_1})$ for $l = 2, \ldots, r$.

Case 2: If $(\xi_{11}^2 - 1)$ is larger, the terms in 3 must cancel. Because $d_1 > d_2 > \ldots > 0$, this implies that $\xi_{11}^2 - 1 = O_p(n^{-\frac{1}{2}\alpha_1})$. Since $\xi_1'\xi_1 \leq 1$, it follows that $\xi_{1l}^2 = O_p(n^{-\frac{1}{2}\alpha_1})$ for $l = 2, \ldots, r$ and $\frac{\hat{V}'\hat{V}}{T} = O_p(n^{-\frac{1}{2}\alpha_1})$. Since also, by (1),

$$\begin{aligned}
\frac{\hat{F}_1'F}{T} &= \frac{1}{T}\big[F\big(\frac{F'F}{T}\big)^{-1/2}\xi_1 + \hat{V}\big]'F \\
&= \xi_1'\big(\frac{F'F}{T}\big)^{-1/2}\frac{F'F}{T} \\
&= \xi_1',
\end{aligned}$$

it follows that $\left(\frac{\hat{F}_1'F_1}{T}\right)^2 - 1 = O_p(n^{-\frac{1}{2}\alpha_1})$, and therefore $\frac{\hat{F}_1'F_1}{T} = 1 + O_p(n^{-\frac{1}{2}\alpha_1})$. It then also follows that, for $l = 2, \ldots, r$: $\frac{\hat{F}_1'F_l}{T} = O_p(n^{-\frac{1}{4}\alpha_1})$. □

Lemma 8. *Denoting the singular value decomposition of* $\frac{1}{\sqrt{T}}X$ *by* $U\Sigma V'$, *let* $\hat{F} = \sqrt{T}U_{[1:K]}$. *Then, under Assumptions 1-4, for each* $k = 1, \ldots, K$ *and* $l = 1, \ldots, r$*:*

- *For* $k < l$*:* $\frac{\hat{F}_k'F_l}{T} = \bar{O}_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_k})$
- *For* $k = l$*:* $\frac{\hat{F}_k'F_l}{T} = 1 + \bar{O}_p(n^{\frac{1}{2}\alpha_1 - \alpha_l})$
- *For* $k > l$*:* $\frac{\hat{F}_k'F_l}{T} = \bar{O}_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_l})$.

Proof. The result for the first row of $\frac{\hat{F}'F}{T}$ is given in Lemma 7. For the remaining columns we repeat the steps above in orthonormal subspaces. Our strategy is therefore similar to the one followed in Stock and Watson [2002]. However, allowing for varying factor strengths requires a more nuanced consideration of the subsequent principal components. Additionally, unlike Stock and Watson [2002], we explicitly derive the rates of convergence for all quantities of interest.

Using the same reasoning as in the previous lemma, we decompose $\hat{F}_k$, the kth column of $\hat{F}$, as follows:

$$\hat{F}_k = F(\frac{F'F}{T})^{-\frac{1}{2}}\xi_k + \hat{V}_k \text{ such that } \hat{V}_k'F = \mathbf{0}.$$

This implies $\xi_k'\xi_k \leq 1$,

$$\hat{F}_k' \frac{F\Lambda'\Lambda F'}{T^2} \hat{F}_k = \xi_k' I_r D_r I_r \xi_k,$$

and

$$\begin{aligned}
&\frac{1}{T^2}\big(\hat{F}_k'F\Lambda'\Lambda F'\hat{F}_k - \hat{F}_k'XX'\hat{F}_k\big) \\
&= \frac{1}{T^2}\big(\hat{F}_k'F\Lambda'\Lambda F'\hat{F}_k - F_k'F\Lambda'\Lambda F'F_k\big) + \frac{1}{T^2}\big(F_k'F\Lambda'\Lambda F'F_k - \hat{F}_k'XX'\hat{F}_k\big) \\
&= O_p(n^{\frac{1}{2}\alpha_1}),
\end{aligned}$$

again using Lemma 6. Following the reasoning in the proof of Theorem 1, the second term on the

RHS is $O_p(n^{\frac{1}{2}\alpha_k})$. This implies for the first term that

$$\frac{1}{T^2}\left(\hat{F}_k' F\Lambda'\Lambda F'\hat{F}_k - F_k' F\Lambda'\Lambda F' F_k\right) = \xi_k' D_r \xi_k - d_k$$
$$= (\xi_{kk}^2 - 1)d_k + \sum_{l\neq k}\xi_{kl}^2 d_l$$
$$= O_p(n^{\frac{1}{2}\alpha_1}).$$

Because $d_1 > d_2 > \ldots > 0$, this implies that

- For $k = l$: $\xi_{kk}^2 = 1 + O_p(n^{\frac{1}{2}\alpha_1 - \alpha_l})$
- For $k < l$: $\xi_{kl}^2 = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k})$
- For $k > l$: $\xi_{kl}^2 = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_l})$ (Since $\xi_k'\xi_k \leq 1$).

We further note that $\frac{\hat{F}_k' F}{T} \leq 1$ and hence $\xi_{kl}^2 = O_p(1)\ \forall l$. This also implies a lower bound on the factor strength, indicated by α_k, for which ξ_k^2 is guaranteed to converge: $\alpha_k > \frac{1}{2}\alpha_1$. □

Lemma 9. *With $\hat{F}$ defined as before, define a $(r \times K)$ matrix $H = \Lambda'\Lambda\frac{F'\hat{F}}{T}\hat{D}_K^{-1}$, where $\hat{D}_K$ is a diagonal matrix with the K largest eigenvalues of $\frac{X'X}{T}$ on the main diagonal. Then, under Assumptions 1-4:*

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{F}_{tk} - H_{k\cdot}' F_t)^2 = O_p(n^{1-2\alpha_k}),$$

where $H_{k\cdot}'$ denotes the kth row of H'.

Proof. Note that by the properties of eigenvectors and eigenvalues $\hat{F} = \frac{XX'}{T}\hat{F}\hat{D}_K^{-1}$. Then:

$$\hat{F} - FH = \frac{XX'}{T}\hat{F}\hat{D}_K^{-1} - F\Lambda'\Lambda\frac{F'\hat{F}}{T}\hat{D}_K^{-1}$$
$$= \frac{1}{T}(XX' - F\Lambda'\Lambda F')\hat{F}\hat{D}_K^{-1}$$
$$= \frac{1}{T}(ee' + e\Lambda F' + F\Lambda' e')\hat{F}\hat{D}_K^{-1}.$$

This is related to the decomposition first derived in Bai and Ng [2002] and used extensively in the literature since its introduction (e.g. Bai [2003], Choi [2012]). The following derivations therefore follow those in Bai and Ng [2002] and Bai [2003], who consider only strong factors. For a particular t we may write:

$$\hat{F}_t - H'F_t = \frac{1}{T}\hat{D}_K^{-1}\hat{F}'(ee_t + e\Lambda F_t + F\Lambda' e_t) \tag{5}$$
$$= \hat{D}_K^{-1}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{F}_s e_s' e_t + \frac{1}{T}\sum_{s=1}^{T}\hat{F}_s F_s'\Lambda' e_t + \frac{1}{T}\sum_{s=1}^{T}\hat{F}_s e_s'\Lambda F_t\right).$$

Because $(I + II + III)^2 \leq 3(I^2 + II^2 + III^2)$, by Cauchy-Schwarz and submultiplicity of the norm: $\|\hat{F}_t - H'F_t\|^2 \leq \|\hat{D}_r^{-1}\|^2 3(I_t + II_t + III_t)$, where:

$$I_t = \frac{1}{T^2}\|\sum_{s=1}^{T}\hat{F}_s e_s' e_t\|^2$$
$$II_t = \frac{1}{T^2}\|\sum_{s=1}^{T}\hat{F}_s F_s'\Lambda' e_t\|^2$$
$$III_t = \frac{1}{T^2}\|\sum_{s=1}^{T}\hat{F}_s e_s'\Lambda F_t\|^2.$$

Thus $\frac{1}{T}\sum_{t=1}^{T}\|\hat{F}_t - H'F_t\|^2 \leq \|\hat{D}_K^{-1}\|^2\frac{1}{T}\sum_{t=1}^{T}3(I_t + II_t + III_t)$, while for each individual factor estimate $\hat{F}_k$, $k = 1, 2, \ldots, r$, $\frac{1}{T}\sum_{t=1}^{T}(\hat{F}_{tk} - H_k' F_{tk})^2 \leq \|\hat{d}_k^{-1}\|^2\frac{1}{T}\sum_{t=1}^{T}3(I_{tk} + II_{tk} + III_{tk})$, with the r-by-1 vector $\hat{F}_s$ replaced by the scalar $\hat{F}_{sk}$ in each of I_t, II_t and III_t above.

Consider each of the above three terms separately:

$$\frac{1}{T}\sum_{t=1}^{T}I_{tk} = \frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}e_s' e_t\|^2$$
$$\leq \frac{1}{T}\sum_{t=1}^{T}\left(\|\frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}[e_s' e_t - \mathbb{E}(e_s' e_t)]\|^2 + \|\frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}\,\mathbb{E}(e_s' e_t)\|^2\right)$$
$$= O_p(n).$$

Since this part does not involve any non-standard assumptions (it does not involve the factor loadings), the last equality follows using the same arguments as in the proof of Theorem 1 in Bai and Ng [2002] using Lemma 4 and Assumption 3(c). Details are not worth repeating. For the next part:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} II_{tk} &= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{T^2}\|\sum_{s=1}^{T}\hat{F}_{sk}F_s'\Lambda' e_t\|^2 \\
&\leq \frac{1}{T}\sum_{t=1}^{T}\left[\|\Lambda' e_t\|^2\left(\frac{1}{T}\sum_{s=1}^{T}\|F_s\|^2\right)\left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk}\|^2\right)\right] \\
&= \frac{1}{T}\sum_{t=1}^{T}\left[\|\Lambda' e_t\|^2 O_p(1)\right] \\
&= O_p(n^{\alpha_1}),
\end{aligned}$$

by Assumption 4(a). Finally, for III_{tk} one can show in a similar manner that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} III_{tk} &= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{T^2}\|\sum_{s=1}^{T}\hat{F}_{sk}e_s'\Lambda F_t\|^2 \\
&= O_p(n^{\alpha_1}).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\|\hat{F}_{tk} - H_{k\cdot}'F_t\|^2 &\leq \hat{d}_k^{-2}3(I_{tk} + II_{tk} + III_{tk}) \\
&\leq O_p(n^{-2\alpha_k})\Big(O_p(n) + O_p(n^{\alpha_1}) + O_p(n^{\alpha_1})\Big) \\
&= O_p(n^{1-2\alpha_k}).
\end{aligned}$$

□

***Proof of Lemma** 3.* First note that

$$H = \Lambda'\Lambda\frac{F'\hat{F}}{T}\hat{D}_K^{-1} = \begin{bmatrix} \frac{d_1}{\hat{d}_1}\frac{F_1'\hat{F}_1}{T} & \frac{d_1}{\hat{d}_2}\frac{F_1'\hat{F}_2}{T} & \cdots & \frac{d_1}{\hat{d}_K}\frac{F_1'\hat{F}_K}{T} \\ \frac{d_2}{\hat{d}_1}\frac{F_2'\hat{F}_1}{T} & \frac{d_2}{\hat{d}_2}\frac{F_2'\hat{F}_2}{T} & & \vdots \\ \vdots & & \ddots & \\ \frac{d_r}{\hat{d}_1}\frac{F_r'\hat{F}_1}{T} & \cdots & & \frac{d_r}{\hat{d}_K}\frac{F_r'\hat{F}_K}{T} \end{bmatrix},$$

where d_k and $\hat{d}_k$ denote the kth entry on the diagonal of $\Lambda'\Lambda$ and $\hat{D}_K$ respectively. Consider entry H_{lk} at position (l,k). First note that $H_{kk} = \frac{d_k}{\hat{d}_k}\frac{F_k'\hat{F}_k}{T} = (1 + O_p(n^{-\frac{1}{2}\alpha_k}))(1 + \bar{O}_p(n^{\frac{1}{2}\alpha_1-\alpha_k})) =$

$1 + O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k})$ by Lemma 8. Next, consider the case $\alpha_k \geq \alpha_l$. By Lemma 8

$$H_{lk} = \frac{d_l}{\hat{d}_k}\frac{F_l'\hat{F}_k}{T} = O_p(n^{\alpha_l - \alpha_k})\bar{O}_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_k}) = \bar{O}_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_k}).$$

Finally, from Lemma 9:

$$\frac{(\hat{F}_k' - FH_{\cdot k})'(\hat{F}_k' - FH_{\cdot k})}{T} = \frac{\hat{F}_k'\hat{F}_k}{T} + H_{\cdot k}'\frac{F'F}{T}H_{\cdot k} - 2\frac{\hat{F}_k'F}{T}H_{\cdot k} = O_p(n^{1-2\alpha_k}). \tag{6}$$

Further,

$$\frac{\hat{F}_k'\hat{F}_k}{T} + H_{\cdot k}'\frac{F'F}{T}H_{\cdot k} - 2\frac{\hat{F}_k'F}{T}H_{\cdot k} = 1 + \sum_{l=1}^{r} H_{lk}^2 - 2\sum_{l=1}^{r}\frac{\hat{F}_k'F_l}{T}H_{lk} \tag{7}$$

$$= 1 + H_{kk}^2 - 2\frac{\hat{F}_k'F_k}{T}H_{kk} + \sum_{l\neq k}^{r} H_{lk}^2 - 2\sum_{l\neq k}^{r}\frac{\hat{F}_k'F_l}{T}H_{lk}. \tag{8}$$

Since $H_{kk} = 1 + O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k})$, it follows that $1 + H_{kk}^2 - 2\frac{\hat{F}_k'F_k}{T}H_{kk} = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k})$. Combining this with (6)-(8), we obtain

$$\sum_{l\neq k}^{r} H_{lk}^2 - 2\sum_{l\neq k}^{r}\frac{\hat{F}_k'F_l}{T}H_{lk} + O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k}) = O_p(n^{1-2\alpha_k})$$

$$\sum_{l\neq k}^{r}\left(H_{lk}^2 - 2\frac{\hat{F}_k'F_l}{T}H_{lk}\right) = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k}) + O_p(n^{1-2\alpha_k})$$

$$\sum_{l\neq k}^{r}\left(\left(\frac{d_l}{\hat{d}_k}\right)^2\frac{F_l'\hat{F}_k}{T}^2 - 2\left(\frac{\hat{F}_k'F_l}{T}\right)^2\frac{d_l}{\hat{d}_k}\right) = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k}) + O_p(n^{1-2\alpha_k})$$

$$\sum_{l\neq k}^{r}\frac{d_l}{\hat{d}_k}\left(\frac{F_l'\hat{F}_k}{T}\right)^2\left[\frac{d_l}{\hat{d}_k} - 2\right] = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k}) + O_p(n^{1-2\alpha_k}).$$

Split the sum above into three parts according to the relationship between α_k and α_l and start with the elements for which $\alpha_k > \alpha_l$. Then,

$$\frac{d_l}{\hat{d}_k}\left(\frac{F_l'\hat{F}_k}{T}\right)^2\left[\frac{d_l}{\hat{d}_k} - 2\right] = O_p(n^{\alpha_l - \alpha_k})O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k})O_p(1) = o_p(n^{\frac{1}{2}\alpha_1 - \alpha_k}).$$

Next, consider elements in the sum for which $\alpha_k = \alpha_l$. Then,

$$\frac{d_l}{\hat{d}_k}\left(\frac{F_l'\hat{F}_k}{T}\right)^2\left[\frac{d_l}{\hat{d}_k} - 2\right] = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k})O_p(1) = O_p(n^{\frac{1}{2}\alpha_1 - \alpha_k}).$$

Finally consider the remaining terms. First note that for the remaining sum the upper limit for the entire sum still holds, as the terms in the first two cases are small enough. Further note that all terms in this remaining sum are positive with probability 1. Thus, each term is bounded by its overall sum and for all k such that $\alpha_k < \alpha_l$:

$$\frac{d_l}{\hat{d}_k}\left(\frac{F_l'\hat{F}_k}{T}\right)^2\left[\frac{d_l}{\hat{d}_k}-2\right] = O_p(n^{\frac{1}{2}\alpha_1-\alpha_k}) + O_p(n^{1-2\alpha_k}). \tag{9}$$

Since the LHS in (9) is equal to H_{lk}^2 up to a negligible term, this establishes that $H_{\cdot k} = \iota_k + O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k})$ in this last case, thus finishing the proof. □

Proof of Lemma *2.* Revisit the decomposition from Lemma 9. It follows that

$$\begin{aligned}\hat{F}_{tk} - H_{k\cdot}'F_t &= \hat{d}_k^{-1}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}e_s'e_t + \frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}F_s'\Lambda'e_t + \frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}e_s'\Lambda F_t\right)\\ &= \hat{d}_k^{-1}\left(I_{tk} + II_{tk} + III_{tk},\right).\end{aligned}$$

Start with I_{tk} and decompose as follows:

$$\begin{aligned}I_{tk} &= \frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}e_s'e_t\\ &\leq \frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)e_s'e_t + \frac{1}{T}H_{k\cdot}'\sum_{s=1}^{T}F_se_s'e_t\\ &\leq \frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)[e_s'e_t - \mathbb{E}(e_s'e_t)] + \frac{1}{T}H_{k\cdot}'\sum_{s=1}^{T}F_s[e_s'e_t - \mathbb{E}(e_s'e_t)]\\ &\qquad + \frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)\,\mathbb{E}(e_s'e_t) + \frac{1}{T}H_{k\cdot}'\sum_{s=1}^{T}F_s\,\mathbb{E}(e_s'e_t).\end{aligned}$$

For the first part:

$$\begin{aligned}&\|\frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)[e_s'e_t - \mathbb{E}(e_s'e_t)]\|\\ &\qquad\leq \left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk} - H_{k\cdot}'F_s\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^{T}[e_s'e_t - \mathbb{E}(e_s'e_t)]^2\right)^{1/2}.\end{aligned} \tag{10}$$

By Lemma 9 the first term is $O_p(n^{\frac{1}{2}-\alpha_k})$. For the second term inside the brackets of (10):

$$\frac{1}{T}\sum_{s=1}^{T}[e_s'e_t - \mathbb{E}(e_s'e_t)]^2 = \frac{n}{T}\sum_{s=1}^{T}[\frac{1}{\sqrt{n}}e_s'e_t - \mathbb{E}(e_s'e_t)]^2.$$

This is $O_p(n)$ by Assumption 3(c), and thus the first part of the decomposition of I_t is $O_p(n^{\frac{1}{2}-\alpha_k})O_p(\sqrt{n}) = O_p(n^{1-\alpha_k})$. For the second part in the decomposition of I_{tk}:

$$\begin{aligned} H_{k\cdot}'\frac{1}{T}\sum_{s=1}^{T}F_s[e_s'e_t - \mathbb{E}(e_s'e_t)] &= [\iota_k + O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k})]O_p(1) \\ &= O_p(1) + O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k}), \end{aligned}$$

by Assumption 3(d). Next consider the third part of I_{tk}:

$$\begin{aligned} \left|\frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)\,\mathbb{E}(e_s'e_t)\right| &\leq \left(\frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)^2\right)^{\frac{1}{2}}\frac{n}{\sqrt{T}}\left(\sum_{s=1}^{T}\mathbb{E}(\frac{e_s'e_t}{n})^2\right)^{\frac{1}{2}} \\ &= \frac{n}{\sqrt{T}}O_p(n^{\frac{1}{2}-\alpha_k})O_p(1) = O_p(n^{1-\alpha_k}), \end{aligned}$$

by Lemma 9 and Assumption 3(b). Finally, for the last part of I_{tk}, $\frac{1}{T}H_{k\cdot}'\sum_{s=1}^{T}F_s\,\mathbb{E}(e_s'e_t) = O_p(1) + O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k})$, since

$$\mathbb{E}\left|\sum_{s=1}^{T}F_sE(e_s'e_t)\right| \leq \max_s\|F_s\|\sum_{s=1}^{T}|\,\mathbb{E}(e_s'e_t)| \leq C$$

by Assumption 3(b) and using the fact that $\max_s\|F_s\| < C$. It follows that

$$I_{tk} = O_p(n^{1-\alpha_k}).$$

Next, consider II_{tk}:

$$\begin{aligned} II_{tk} &= \frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}F_s'\Lambda'e_t \\ &= \frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk} - H_{k\cdot}'F_s)F_s'\Lambda'e_t + H_{k\cdot}'\frac{1}{T}\sum_{s=1}^{T}F_sF_s'\Lambda'e_t. \end{aligned}$$

For the second part:

$$\begin{aligned}
H'_{k\cdot}\frac{1}{T}\sum_{s=1}^{T}F_sF'_s\Lambda' e_t &= H'_{k\cdot}(\frac{1}{T}\sum_{s=1}^{T}F_sF'_s)(\Lambda' e_t) = H'_{k\cdot}\Lambda' e_t \\
&= [\iota_k + O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k})]O_p(n^{\frac{1}{2}\alpha_1}) \\
&= O_p(n^{\frac{1}{2}\alpha_1}) + O_p(n^{\frac{3}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}+\frac{1}{2}\alpha_1-\alpha_k}).
\end{aligned}$$

For the first part:

$$\|\frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H'_{k\cdot}F_s)F'_s\Lambda' e_t\| \leq \left(\frac{1}{T}\sum_{s=1}^{T}\|(\hat{F}_{sk}-H'_{k\cdot}F_s)\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^{T}\|F'_s\Lambda' e_t\|^2\right)^{1/2}.$$

Further:

$$\frac{1}{T}\sum_{s=1}^{T}\|F'_s\Lambda' e_t\|^2 \leq \|\Lambda' e_t\|^2\frac{1}{T}\sum_{s=1}^{T}\|F_s\|^2 = O_p(n^{\alpha_1}),$$

and by Lemma 9 $\left(\frac{1}{T}\sum_{s=1}^{T}\|(\hat{F}_{sk}-H'_{k\cdot}F_s)\|^2\right)^{1/2} = O_p(n^{\frac{1}{2}-\alpha_k})$. Therefore:

$$\begin{aligned}
II_{tk} &= \frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}F'_s\Lambda' e_t \\
&= \frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H'_{k\cdot}F_s)F'_s\Lambda' e_t + H'_{k\cdot}\frac{1}{T}\sum_{s=1}^{T}F_sF'_s\Lambda' e_t \\
&\leq O_p(n^{\frac{1}{2}\alpha_1}) + O_p(n^{\frac{3}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}+\frac{1}{2}\alpha_1-\alpha_k}).
\end{aligned}$$

Finally, consider III_{tk}:

$$\begin{aligned}
III_{tk} &= \frac{1}{T}\sum_{s=1}^{T}\hat{F}_{sk}e'_s\Lambda F_t \\
&= \frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H'_{k\cdot}F_s)e'_s\Lambda F_t + H'_{k\cdot}\frac{1}{T}\sum_{s=1}^{T}F_se'_s\Lambda F_t.
\end{aligned}$$

Start with the first term:

$$\begin{aligned}
\|\frac{1}{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H'_{k\cdot}F_s)e'_s\Lambda F_t\| &\leq \left(\frac{1}{T}\sum_{s=1}^{T}\|(\hat{F}_{sk}-H'_{k\cdot}F_s)\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{s=1}^{T}\|e'_s\Lambda\|^2\right)^{1/2}\|F_t\| \\
&= O_p(n^{\frac{1}{2}-\alpha_k})O_p(n^{\frac{1}{2}\alpha_1})O_p(1) = O_p(n^{\frac{1}{2}+\frac{1}{2}\alpha_1-\alpha_k}).
\end{aligned}$$

For the second term:

$$H'_{k\cdot}\frac{1}{T}\sum_{s=1}^{T}F_s e'_s\Lambda F_t = H'_{k\cdot}\frac{n^{\frac{1}{2}\alpha_1}}{\sqrt{T}}\big(\frac{1}{n^{\frac{1}{2}\alpha_1}\sqrt{T}}\sum_{s=1}^{T}F_s e'_s\Lambda\big)F_t$$
$$= [\iota_k + O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k})]O_p(n^{\frac{1}{2}\alpha_1-\frac{1}{2}})O_p(1)$$
$$= O_p(n^{\frac{1}{2}\alpha_1-\frac{1}{2}}) + O_p(n^{\frac{1}{2}\alpha_1-\alpha_k}) + O_p(n^{\frac{3}{4}\alpha_1-\frac{1}{2}-\frac{1}{2}\alpha_k}),$$

using Assumption 4(b). It follows that

$$III_{tk} = O_p(n^{\frac{1}{2}+\frac{1}{2}\alpha_1-\alpha_k}) + O_p(n^{\frac{1}{2}\alpha_1-\frac{1}{2}}) + O_p(n^{\frac{1}{2}\alpha_1-\alpha_k}) + O_p(n^{\frac{3}{4}\alpha_1-\frac{1}{2}-\frac{1}{2}\alpha_k})$$
$$= O_p(n^{\frac{1}{2}+\frac{1}{2}\alpha_1-\alpha_k}).$$

Combining these partial results we obtain that

$$\hat{F}_{tk} - H'_{k\cdot}F_t = \hat{d}_k^{\,-1}(I_t + II_t + III_t)$$
$$= O_p(n^{-\alpha_k})\bigg(O_p(n^{1-\alpha_k}) + O_p(n^{\frac{1}{2}\alpha_1}) + O_p(n^{\frac{3}{4}\alpha_1-\frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}+\frac{1}{2}\alpha_1-\alpha_k})\bigg)$$
$$= O_p(n^{-\alpha_k})\bigg(O_p(n^{1-\alpha_k}) + O_p(n^{\frac{1}{2}\alpha_1})\bigg)$$
$$= O_p(n^{1-2\alpha_k}) + O_p(n^{\frac{1}{2}\alpha_1-\alpha_k}).$$

□

Lemma 10. *Under Assumptions 1-4, with $\hat{F}$ and H defined as in the previous lemmata:*

$$\frac{(\hat{F}_k - FH_{\cdot k})'e_i}{T} = O_p(n^{1-2\alpha_k}).$$

Proof.

$$\frac{(\hat{F}_k - FH_{\cdot k})'e_i}{T} = \frac{1}{T}\sum_{t=1}^{T}(\hat{F}_{tk} - H'_{k\cdot}F_t)e_{ti}$$
$$= \hat{d}_k^{-1}\bigg(\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_{sk}e'_s e_t e_{ti} + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_{sk}F'_s\Lambda' e_t e_{ti} + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_{sk}e'_s\Lambda F_t e_{ti}\bigg)$$
$$= \hat{d}_k^{-1}\bigg(I_k + II_k + III_k\bigg)$$

$$
\begin{aligned}
I_k &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_{sk}e_s'e_te_{ti} \\
&= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)e_s'e_te_{ti} + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}H_{k\cdot}'F_se_s'e_te_{ti} \\
&= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)[e_s'e_t-\mathbb{E}(e_s'e_t)]e_{ti} + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}H_{k\cdot}'F_s[e_s'e_t-\mathbb{E}(e_s'e_t)]e_{ti} \\
&\qquad + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)\,\mathbb{E}(e_s'e_t)e_{ti} + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}H_{k\cdot}'F_s\,\mathbb{E}(e_s'e_t)e_{ti}.
\end{aligned}
$$

Consider these four terms in turn:

$$
\begin{aligned}
&\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)[e_s'e_t-\mathbb{E}(e_s'e_t)]e_{ti} \\
&\qquad \le \left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk}-H_{k\cdot}'F_s\|^2\right)^{\frac{1}{2}}\left(\frac{1}{T}\sum_{s=1}^{T}\big(\frac{1}{T}\sum_{t=1}^{T}[e_s'e_t-\mathbb{E}(e_s'e_t)]e_{ti}\big)^2\right)^{\frac{1}{2}} \\
&\qquad \le \sqrt{n}\left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk}-H_{k\cdot}'F_s\|^2\right)^{\frac{1}{2}}\left(\frac{1}{T}\sum_{s=1}^{T}\big(\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\sqrt{n}}[e_s'e_t-\mathbb{E}(e_s'e_t)]e_{ti}\big)^2\right)^{\frac{1}{2}} \\
&\qquad \le \sqrt{n}O_p(n^{\frac{1}{2}-\alpha_k})O_p(1) = O_p(n^{1-\alpha_k}),
\end{aligned}
$$

where the boundedness of the last term follows from Assumption 3(c). For the next term, ignoring H, take expectations:

$$
\begin{aligned}
\mathbb{E}\left[\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}F_s[e_s'e_t-\mathbb{E}(e_s'e_t)]e_{ti}\right] &= \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{nT}}\sum_{s=1}^{T}F_s[e_s'e_t-\mathbb{E}(e_s'e_t)]\right)e_{ti}\right] \\
&\le \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(\|\frac{1}{\sqrt{nT}}\sum_{s=1}^{T}F_s[e_s'e_t-\mathbb{E}(e_s'e_t)]\|^2\right)^{\frac{1}{2}}\big(\mathbb{E}(e_{ti})^2\big)^{\frac{1}{2}} \\
&= O(1).
\end{aligned}
$$

For the third term:

$$
\begin{aligned}
&\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)\,\mathbb{E}(e_s'e_t)e_{ti} \\
&\qquad \le \left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk}-H_{k\cdot}'F_s\|^2\right)^{\frac{1}{2}}\left(\frac{n}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}|\,\mathbb{E}(\frac{e_s'e_t}{n})|^2\frac{1}{T}\sum_{t=1}^{T}e_{ti}^2\right)^{\frac{1}{2}} \\
&\qquad = O_p(n^{\frac{1}{2}-\alpha_k})O_p(\sqrt{n}) = O_p(n^{1-\alpha_k}),
\end{aligned}
$$

using Lemma 4. Finally, ignoring H, take expectations of the last term:

$$\mathbb{E}\left[\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}F_s\,\mathbb{E}(e_s'e_t)e_{ti}\right]\leq\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\|F_s\|^2\right)^{\frac{1}{2}}\mathbb{E}(\frac{e_s'e_t}{n})\left(\mathbb{E}\,e_{ti}^2\right)^{\frac{1}{2}}=O(1),$$

since both the first and third term in the final sum is bounded and using Assumption 3(b). Therefore $I_k=O_p(n^{1-\alpha_k})+O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k})+O_p(n^{\frac{1}{2}-\alpha_k})=O_p(n^{1-\alpha_k})$. Next consider II_k:

$$\begin{aligned}II_k&=\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{F}_{sk}F_s'\Lambda'e_te_{ti}\\&=\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)F_s'\Lambda'e_te_{ti}+\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}H_{k\cdot}'F_sF_s'\Lambda'e_te_{ti}.\end{aligned}$$

Again consider both terms separately and start with the second:

$$\begin{aligned}&\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}H_{k\cdot}'F_sF_s'\Lambda'e_te_{ti}=H_{k\cdot}'\left(\frac{1}{T}\sum_{s=1}^{T}F_sF_s'\right)\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{n}\lambda_je_{tj}e_{ti}\\&=H_{k\cdot}'\left(\frac{1}{T}\sum_{s=1}^{T}F_sF_s'\right)\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{n}\lambda_j[e_{tj}e_{ti}-\mathbb{E}(e_{tj}e_{ti})]+\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{n}\lambda_j\,\mathbb{E}(e_{tj}e_{ti})\right)\\&\leq H_{k\cdot}'\left(\frac{1}{T}\sum_{s=1}^{T}F_sF_s'\right)\left(\frac{C}{\sqrt{T}\sqrt{n}}\sum_{t=1}^{T}\sum_{j=1}^{n}[e_{tj}e_{ti}-\mathbb{E}(e_{tj}e_{ti})]+\sum_{j=1}^{n}\lambda_j\,\mathbb{E}(\frac{e_j'e_i}{T})\right)\\&=[\iota_k+O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k})+O_p(n^{\frac{1}{2}-\alpha_k})][O_p(1)+O(1)],\end{aligned}$$

where the boundedness of the last term follows from Assumption 3(b). Similarly for the first term:

$$\begin{aligned}&\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}(\hat{F}_{sk}-H_{k\cdot}'F_s)F_s'\Lambda'e_te_{ti}\\&\qquad\leq\left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk}-H_{k\cdot}'F_s\|^2\right)^{\frac{1}{2}}\left(\frac{1}{T}\sum_{s=1}^{T}\left(\frac{1}{T}\sum_{t=1}^{T}F_s'\Lambda'e_te_{ti}\right)^2\right)^{\frac{1}{2}}\\&\qquad\leq\left(\frac{1}{T}\sum_{s=1}^{T}\|\hat{F}_{sk}-H_{k\cdot}'F_s\|^2\right)^{\frac{1}{2}}\left(\frac{1}{T}\sum_{s=1}^{T}\left(F_s'\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{n}\lambda_je_{tj}e_{ti}\right)^2\right)^{\frac{1}{2}}\\&\qquad=O_p(n^{\frac{1}{2}-\alpha_k})O_p(1),\end{aligned}$$

using the same arguments as above. We conclude that $II_k=O_p(1)+O_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k})+O_p(n^{\frac{1}{2}-\alpha_k})$. Finally, using similar arguments as in the proof of II_k, one can show that the same bounds apply to III_k, and it follows that

$$\frac{(\hat{F}_k - FH_{\cdot k})'e_i}{T} = \hat{d}_k^{-1}\Big(I_k + II_k + III_k\Big)$$
$$= O_p(n^{-\alpha_k})\Big(O_p(n^{1-\alpha_k}) + O_p(1) + O_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_k}) + O_p(n^{\frac{1}{2}-\alpha_k})\Big)$$
$$= O_p(n^{1-2\alpha_k}).$$

□

Lemma 11. *Under Assumptions 1-4, let* $z = n^\tau g(n), \tau \in [0.5, 1]$, *such that (i)* $g(n) \to \infty$ *and (ii)* $g(n)/n^\epsilon \to 0$ *for any* $\epsilon > 0$ *as* $n \to \infty$. *With slight abuse of notation, the estimated loadings* $\hat{\lambda}_{ik}$ *are ordered such that, for each* k, $\left|\hat{\lambda}_{1k}\right| \geq \left|\hat{\lambda}_{2k}\right| \geq \ldots \geq \left|\hat{\lambda}_{nk}\right|$. *Then*

(a) If $\alpha_k > \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$: $\frac{1}{z}\sum_{i=1}^z \hat{\lambda}_{ik}^2 - \frac{1}{z}\sum_{i=1}^z \lambda_{ik}^2 = \bar{O}_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_k}) + \bar{O}_p(n^{1-2\alpha_k})$

(b) If $\alpha_k \leq \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$: $\frac{1}{z}\sum_{i=1}^z \hat{\lambda}_{ik}^2 - \frac{1}{z}\sum_{i=1}^z \lambda_{ik}^2 = O_p\left(\frac{n^{\alpha_k}}{n^\tau g(n)}\right)$.

Proof. By Theorem 3:

$$\hat{\lambda}_{ik} - \lambda_{ik} = \bar{O}_p(n^{\frac{1}{4}\alpha_1 - \frac{1}{2}\alpha_k}) + \bar{O}_p(n^{1-2\alpha_k}).$$

Since

$$\frac{1}{z}\sum_{i=1}^z \hat{\lambda}_{ik}^2 - \frac{1}{z}\sum_{i=1}^z \lambda_{ik}^2 = \frac{1}{z}\sum_{i=1}^z (\hat{\lambda}_{ik}^2 - \lambda_{ik}^2),$$

this is just an average (squared) deviation and the result in part (a) immediately follows.

Next consider the case $\alpha_k \leq \tau$:

$$\frac{1}{z}\sum_{i=1}^z \lambda_{ik}^2 \leq \frac{1}{z}\sum_{i=1}^n \lambda_{ik}^2 = \frac{n^{-\tau}}{g(n)}\psi_k(\Lambda'\Lambda) = O_p\left(\frac{n^{\alpha_k - \tau}}{g(n)}\right)$$

and, similarly

$$\frac{1}{z}\sum_{i=1}^z \hat{\lambda}_{ik}^2 \leq \frac{1}{z}\sum_{i=1}^n \hat{\lambda}_{ik}^2 = \frac{n^{-\tau}}{g(n)}\psi_k\left(\frac{X'X}{T}\right) = O_p\left(\frac{n^{\alpha_k - \tau}}{g(n)}\right).$$

Combined they imply the stated bound on the difference.

Finally, consider the case $\max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\} < \alpha_k \leq \tau$. In those situations both bounds above apply and imply convergence to zero. For $\alpha_k > \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$, the first bound is the tighter one and thus applies. □

Lemma 12. *There exists a constant $c > 0$ such that $lim_{n\to\infty} P(\hat{\Upsilon}^u_{zk}/n^{(1+\frac{1}{2}u)\alpha_k - \frac{1}{2}u} < c) = 0$ for $k = 1, \ldots, r_{max}$.*

Proof. First note that

$$\frac{1}{z}\sum^{z}\hat{\lambda}^2_{ik} \geq \frac{1}{n}\sum^{n}\hat{\lambda}^2_{ik} = \frac{1}{n}\psi_k\left(\frac{X'X}{T}\right).$$

It follows that

$$\hat{\Upsilon}^u_{zk} = \psi_k(\frac{X'X}{T})\left(\frac{1}{z}\sum_i^z \frac{\hat{\lambda}^2_{ik}}{\sqrt{\frac{1}{n}\sum_{i=1}^n \hat{\lambda}^2_{ik}}}\right)^u \geq \psi_k(\frac{X'X}{T})\left(\frac{1}{z}\sum_i^z \hat{\lambda}^2_{ik}\right)^{\frac{1}{2}u}$$
$$\geq \psi_k(\frac{X'X}{T})\left[\frac{1}{n}\psi_k(\frac{X'X}{T})\right]^{\frac{1}{2}u} = n^{-\frac{1}{2}u}\psi_k(\frac{X'X}{T})^{1+\frac{1}{2}u}.$$

For $\alpha_k > 0$, there exists a $c_1 > 0$ such that

$$lim_{n\to\infty} P\left(\psi_k(\frac{X'X}{T})/n^{\alpha_k} < c_1\right) = 0$$

and thus

$$lim_{n\to\infty} P\left(\hat{\Upsilon}^u_{zk}/n^{(1+\frac{1}{2}u)\alpha_k - \frac{1}{2}u} \geq c\right) = 1.$$

Finally, if $\alpha_k = 0$, since $\psi_k(\frac{X'X}{T}) > c_{eig} > 0$ for $k = r+1, \ldots, [dn]$ this implies that there exists a positive constant c_2 such that

$$\hat{\Upsilon}^u_{zk} \geq c_2 n^{-\frac{1}{2}u}.$$

□

B.2 Proofs of Corollary 4 and Theorem 4

Proof of Corollary 4. First consider $k = r_1 + 1, \ldots, r_{max}$. Then, by Theorem 1, $\psi_k\left(\frac{XX'}{T}\right) = O_p(1)$ and thus there exists a finite $c_1 > 0$, $lim_{n\to\infty} P\left(\psi_k\left(\frac{XX'}{T}\right) \geq c_1\right) = 0$. Further, by Assumption 3 (e), there exists a constant $c_2 > 0$, such that $P\left(\psi_{k^*}\left(\frac{XX'}{T}\right) \geq c_2\right) = 1$ for $k^* =$

$r_1+1,\ldots,r_{max}$. Then, for any finite $c_3>0$,

$$\begin{aligned}
&lim_{n\to\infty}P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)}>c_3log(n)\right)\\
&=lim_{n\to\infty}\Bigg[P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)}>c_3log(n)\middle|\psi_{k+1}\left(\frac{XX'}{T}\right)<c_2\right)P\left(\psi_{k+1}\left(\frac{XX'}{T}\right)<c_2\right)\\
&\qquad+P\left((\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)}>c_3log(n)\middle|\psi_{k+1}\left(\frac{XX'}{T}\right)\geq c_2\right)P\left(\psi_{k+1}\left(\frac{XX'}{T}\right)\geq c_2\right)\Bigg]\\
&=lim_{n\to\infty}P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)}>c_3log(n)\middle|\psi_{k+1}\left(\frac{XX'}{T}\right)\geq c_2\right)+0\\
&\leq lim_{n\to\infty}P\left(\psi_k\left(\frac{XX'}{T}\right)>c_2c_3log(n)\right)=0.
\end{aligned}$$

Next, consider $k=1,\ldots,r_1-1$. We already established that, for any finite $q_1>0$,

$$lim_{n\to\infty}P\left(\psi_k\left(\frac{XX'}{T}\right)>q_1\sqrt{n}\right)=1.$$

It then immediately follows that there exists an $h > 0$ such that

$$
\begin{aligned}
& lim_{n\to\infty} P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)} > h\sqrt{n}\right) \\
& = lim_{n\to\infty}\Bigg[P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)} > h\sqrt{n}\middle|\psi_{k+1}\left(\frac{XX'}{T}\right) < q_1\sqrt{n}\right) P\left(\psi_{k+1}\left(\frac{XX'}{T}\right) < q_1\sqrt{n}\right) \\
& \quad + P\left(\frac{\psi_r\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)} > h\sqrt{n}\middle|\psi_{k+1}\left(\frac{XX'}{T}\right) \geq q_1\sqrt{n}\right) P\left(\psi_{r+1}\left(\frac{XX'}{T}\right) \geq q_1\sqrt{n}\right)\Bigg] \\
& = lim_{n\to\infty} P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)} \geq h\sqrt{n}\middle|\psi_{k+1}\left(\frac{XX'}{T}\right) \geq q_1\sqrt{n}\right) + 0 \\
& \leq lim_{n\to\infty} P\left(\psi_r\left(\frac{XX'}{T}\right) > q_1 hn\right).
\end{aligned}
$$

But since there exists a finite $q_2 > 0$ with $lim_{n\to\infty} P\left(\psi_k\left(\frac{XX'}{T}\right) > q_2 n\right) = 0$, letting $h = \frac{q_2}{q_1}$ establishes $lim_{n\to\infty} P\left(\frac{\psi_k\left(\frac{XX'}{T}\right)}{\psi_{k+1}\left(\frac{XX'}{T}\right)} > h\sqrt{n}\right) = 0$. Finally, consider $k = r_1$. By Assumption 5, $\alpha_k > .5$ and thus $lim_{n\to\infty} P\left(\psi_k\left(\frac{XX'}{T}\right) > q_1\sqrt{n}\right) = 1$ for any finite $q_1 > 0$. On the other hand, $\psi_{r_1+1}\left(\frac{XX'}{T}\right) = O_p(1)$ and thus there exists a $q_2 > 0$, such that $P\left(\psi_{r_1+1}\left(\frac{XX'}{T}\right) \geq q_2\right) = 0$.

Then, for any finite $q_3 > 0$

$$
\begin{aligned}
&lim_{n\to\infty} P\left(\frac{\psi_{r_1}\left(\frac{XX'}{T}\right)}{\psi_{r_1+1}\left(\frac{XX'}{T}\right)} > q_3\sqrt{n}\right) \\
&= lim_{n\to\infty}\Bigg[P\left(\frac{\psi_{r_1}\left(\frac{XX'}{T}\right)}{\psi_{r_1+1}\left(\frac{XX'}{T}\right)} > q_3\sqrt{n}\middle|\psi_{r_1+1}\left(\frac{XX'}{T}\right) < q_2\right) P\left(\psi_{r_1+1}\left(\frac{XX'}{T}\right) < q_2\right) \\
&\quad + P\left(\frac{\psi_{r_1}\left(\frac{XX'}{T}\right)}{\psi_{r_1+1}\left(\frac{XX'}{T}\right)} > q_3\sqrt{n}\middle|\psi_{r_1+1}\left(\frac{XX'}{T}\right) \geq q_2\right) P\left(\psi_{r_1+1}\left(\frac{XX'}{T}\right) \geq q_2\right)\Bigg] \\
&= lim_{n\to\infty} P\left(\frac{\psi_{r_1}\left(\frac{XX'}{T}\right)}{\psi_{r_1+1}\left(\frac{XX'}{T}\right)} > q_3\sqrt{n}\middle|\psi_{r_1+1}\left(\frac{XX'}{T}\right) < q_2\right) + 0 \\
&\geq lim_{n\to\infty} P\left(\psi_{r_1}\left(\frac{XX'}{T}\right) > q_2 q_3\sqrt{n}\right) = 1.
\end{aligned}
$$

Choosing $q_3 = \frac{h}{q_2}$, this completes the proof. □

A key step in the proof of Corollary 4 makes use of the following class of (infeasible) quantities Υ_{zk}^u. For $u \in [0, 2]$:

$$\Upsilon_{zk}^u = \psi_k\left(\frac{\Lambda F' F \Lambda'}{T}\right) S_{zk}^u = \psi_k\left(\frac{\Lambda F' F \Lambda'}{T}\right)\left(\frac{1}{z}\sum_i^z \frac{\lambda_{ik}^2}{\sqrt{\frac{1}{n}\sum_{i=1}^n \lambda_{ik}^2}}\right)^u,$$

where, with some abuse of notation, the squared loadings λ_{ik}^2 are sorted in decreasing order. We obtain the following lemma.

Lemma 13. *Under Assumptions 1-2, choose a threshold $z = n^\tau g(n), \tau \in [0, 1]$, such that (i) $g(n) \to \infty$ and (ii) $g(n)/n^\epsilon \to 0$ for any $\epsilon > 0$ as $n \to \infty$. Then, for any given factor $k \leq r$, with $u \in [0, 2]$:*

(a) If $\alpha_k > \tau$: $\Upsilon_{zk}^u \asymp n^{(1-\frac{1}{2}u)\alpha_k + \frac{1}{2}u}$

(b) If $\alpha_k \leq \tau$: $\Upsilon_{zk}^u \asymp n^{(1+\frac{1}{2}u)\alpha_k + (\frac{1}{2}-\tau)u} g(n)^{-u}$.

Further, for $k = r+1, \ldots, r_{max}$: $\Upsilon_{zk}^u = 0$.

Proof. Using Assumption 1 we can rewrite Υ^u_{zk} as follows:

$$\begin{aligned}\Upsilon^u_{zk} &= \psi_k\Big(\frac{\Lambda F'F\Lambda'}{T}\Big)S^u_{zk} = \psi_k\Big(\frac{\Lambda F'F\Lambda'}{T}\Big)\Big(\frac{1}{z}\sum_i^z \frac{\lambda^2_{ik}}{\sqrt{\frac{1}{n}\sum_{i=1}^n \lambda^2_{ik}}}\Big)^u \\ &= \psi_k\big(\Lambda'\Lambda\big)\big(\sum_{i=1}^n \lambda^2_{ik}\big)^{-\frac{1}{2}u}\big(\frac{n^{\frac{1}{2}}}{z}\sum_i^z \lambda^2_{ik}\big)^u \\ &= \psi_k\big(\Lambda'\Lambda\big)^{1-\frac{1}{2}u} n^{\frac{1}{2}u}\big(\frac{1}{z}\sum_i^z \lambda^2_{ik}\big)^u. \end{aligned} \tag{11}$$

First consider scenario (a). With $\alpha_k > \tau$, the last part of (11) is simply an average of the square of the z largest loadings. Combining Assumption 2(a) with the fact that $|\lambda_{ik}| < C\ \forall i$, we immediately have $\Upsilon^u_{zk} \asymp n^{(1-\frac{1}{2}u)\alpha_k} n^{\frac{1}{2}u}$.

Next, for part (b), let $\alpha_k \le \tau$: There are only $|\mathcal{A}_k| \asymp n^{\alpha_k}$ "large" loadings in the sum of equation (11), and Assumption 2 implies that

$$\frac{1}{z}\sum_i^z \lambda^2_{ik} = \frac{1}{z}\sum_{i\in\mathcal{A}_k} \lambda^2_{ik} + \frac{1}{z}\sum_{i\notin\mathcal{A}_k} \lambda^2_{ik} \asymp \frac{n^{\alpha_k-\tau}}{g(n)},$$

and it follows that $\Upsilon^u_{zk} \asymp n^{(1-\frac{1}{2}u)\alpha_k} n^{\frac{1}{2}u} \frac{n^{(\alpha_k-\tau)u}}{g(n)^u}$.

For $k > r$, $\lambda_{ik} = 0\ \forall i$, and this completes the proof. □

Proof of Theorem 4. First note that:

$$\hat{\Upsilon}^u_{zk} - \Upsilon^u_{zk} = n^{\frac{1}{2}u}\Big[\psi_k\big(\frac{X'X}{T}\big)^{1-\frac{1}{2}u}\big(\frac{1}{z}\sum_i^z \hat{\lambda}^2_{ik}\big)^u - \psi_k\big(\frac{\Lambda F'F\Lambda'}{T}\big)^{1-\frac{1}{2}u}\big(\frac{1}{z}\sum_i^z \lambda^2_{ik}\big)^u\Big].$$

Because $ab - cd = (a-c)d + (b-d)c + (a-c)(b-d)$ we may write

$$\hat{\Upsilon}^u_{zk} - \Upsilon^u_{zk} = n^{\frac{1}{2}u}\Big[I + II + III\Big],$$

where

$$\begin{aligned} I &= \Big(\psi_k\big(\frac{X'X}{T}\big)^{1-\frac{1}{2}u} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{T}\big)^{1-\frac{1}{2}u}\Big)\big(\frac{1}{z}\sum_i^z \lambda^2_{ik}\big)^u \\ II &= \Big(\big(\frac{1}{z}\sum_i^z \hat{\lambda}^2_{ik}\big)^u - \big(\frac{1}{z}\sum_i^z \lambda^2_{ik}\big)^u\Big)\psi_k\big(\frac{\Lambda F'F\Lambda'}{T}\big)^{1-\frac{1}{2}u} \\ III &= \Big(\psi_k\big(\frac{X'X}{T}\big)^{1-\frac{1}{2}u} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{T}\big)^{1-\frac{1}{2}u}\Big)\Big(\big(\frac{1}{z}\sum_i^z \hat{\lambda}^2_{ik}\big)^u - \big(\frac{1}{z}\sum_i^z \lambda^2_{ik}\big)^u\Big). \end{aligned}$$

First consider the difference in I:

$$
\begin{aligned}
&\psi_k\big(\frac{X'X}{T}\big)^{1-\frac{1}{2}u} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{T}\big)^{1-\frac{1}{2}u} \\
&= n^{(1-\frac{1}{2}u)\alpha_k}\left[\left(\psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big) + \psi_k\big(\frac{X'X}{Tn^{\alpha_k}}\big) - \psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big)\right)^{1-\frac{1}{2}u} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big)^{1-\frac{1}{2}u}\right] \\
&= n^{(1-\frac{1}{2}u)\alpha_k}\left[\left(\psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big) + \varepsilon_{\psi_k}\right)^{1-\frac{1}{2}u} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big)^{1-\frac{1}{2}u}\right],
\end{aligned}
$$

where $\varepsilon_{\psi_k} = O_p(n^{-\frac{1}{2}\alpha_k})$ following the reasoning in the proof of Theorem 1. Using Newton's generalized binomial theorem:

$$
\begin{aligned}
&\left(\psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big) + \varepsilon_{\psi_k}\right)^{1-\frac{1}{2}u} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big)^{1-\frac{1}{2}u} \\
&= \sum_{w=0}^{\infty} \frac{\frac{\Gamma(2-\frac{1}{2}u)}{\Gamma(2-\frac{1}{2}u-w)}}{w!}\psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big)^{1-\frac{1}{2}u-w}\varepsilon_{\psi_k}^{w} - \psi_k\big(\frac{\Lambda F'F\Lambda'}{Tn^{\alpha_k}}\big)^{1-\frac{1}{2}u} \\
&= O_p(n^{-\frac{1}{2}\alpha_k}) + o_p(n^{-\frac{1}{2}\alpha_k}).
\end{aligned}
$$

We can thus distinguish between two cases as follows:
For $\alpha_k > \tau$: $I = n^{(1-\frac{1}{2}u)\alpha_k}[O_p(n^{-\frac{1}{2}\alpha_k}) + o_p(n^{-\frac{1}{2}\alpha_k})]O_p(1) = O_p(n^{(\frac{1}{2}-\frac{1}{2}u)\alpha_k})$.
For $\alpha_k \leq \tau$: $I = n^{(1-\frac{1}{2}u)\alpha_k}[O_p(n^{-\frac{1}{2}\alpha_k}) + o_p(n^{-\frac{1}{2}\alpha_k})]O_p\left(\frac{n^{\alpha_k}}{n^{\tau}g(n)^u}\right) = O_p(n^{(\frac{1}{2}+\frac{1}{2}u)\alpha_k-\tau u}g(n)^{-u})$.
Next, consider the difference in II. For (a), with $\alpha_k > \tau$:

$$
\begin{aligned}
&\big(\frac{1}{z}\sum_i^z \hat{\lambda}_{ik}^2\big)^u - \big(\frac{1}{z}\sum_i^z \lambda_{ik}^2\big)^u \\
&= \big(\frac{1}{z}\sum_i^z \lambda_{ik}^2 + \frac{1}{z}\sum_i^z \hat{\lambda}_{ik}^2 - \frac{1}{z}\sum_i^z \lambda_{ik}^2\big)^u - \big(\frac{1}{z}\sum_i^z \lambda_{ik}^2\big)^u \\
&= \big(\frac{1}{z}\sum_i^z \lambda_{ik}^2 + \frac{1}{z}\sum_i^z(\hat{\lambda}_{ik}^2 - \lambda_{ik}^2)\big)^u - \big(\frac{1}{z}\sum_i^z \lambda_{ik}^2\big)^u \\
&= u\big[\frac{1}{z}\sum_i^z \lambda_{ik}^2\big)^{u-1}\big(\frac{1}{z}\sum_i^z(\hat{\lambda}_{ik}^2 - \lambda_{ik}^2)\big] \\
&\qquad + \mathbf{1}_{\{u>1\}}\frac{u(u-1)}{2}\big[\big(\frac{1}{z}\sum_i^z \lambda_{ik}^2\big)^{u-2}\big(\frac{1}{z}\sum_i^z(\hat{\lambda}_{ik}^2 - \lambda_{ik}^2)\big)^2\big] + \ldots,
\end{aligned}
$$

where the third equality follows from the generalized binomial theorem for nonnegative exponents.

Later terms will be dominated. Thus:

$$
\begin{aligned}
II &= \big[\mathbf{1}_{\{u>0\}}[\bar{O}_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + \bar{O}_p(n^{1-2\alpha_k})] \\
&\qquad\qquad\qquad\qquad + \mathbf{1}_{\{u>1\}}[\bar{O}_p(n^{\frac{1}{2}\alpha_1-\alpha_k}) + \bar{O}_p(n^{2-4\alpha_k})]\big]O_p(n^{(1-\frac{1}{2}u)\alpha_k}) \\
&= \mathbf{1}_{\{u>0\}}[\bar{O}_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) + \bar{O}_p(n^{1-2\alpha_k})]O_p(n^{(1-\frac{1}{2}u)\alpha_k}).
\end{aligned}
$$

Similarly, by Lemma 11, the same rate holds if $\max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\} < \alpha_k \leq \tau$. On the other hand, if $\alpha_k \leq \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$, by Lemma 11:

$$
\begin{aligned}
(\frac{1}{z}\sum_i^z \hat{\lambda}_{ik}^2)^u - (\frac{1}{z}\sum_i^z \lambda_{ik}^2)^u &= \frac{1}{z^u}\left[\left(\sum_i^z \hat{\lambda}_{ik}^2\right)^u - \left(\sum_i^z \lambda_{ik}^2\right)^u\right] \\
&= \frac{1}{n^{\tau u} g(n)^u}\mathbf{1}_{\{u>0\}}[O_p(n^{u\alpha_k}) - O_p(n^{u\alpha_k}] \\
&= \mathbf{1}_{\{u>0\}}O_p(n^{(\alpha_k-\tau)u}g(n)^{-u}),
\end{aligned}
$$

which in turn implies that

$$
II = \mathbf{1}_{\{u>0\}}O_p(\frac{n^{(\alpha_k-\tau)u}}{g(n)^u})O_p(n^{(1-\frac{1}{2}u)\alpha_k}) = \mathbf{1}_{\{u>0\}}O_p(\frac{n^{(1+\frac{1}{2}u)\alpha_k-\tau u}}{g(n)^u}).
$$

Using the derivations above, it is straightforward to see that $III = O_p(II)$.

We therefore conclude that, for $\alpha_k > \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$:

$$\begin{aligned}
\Upsilon^u_{zk} - \hat{\Upsilon}^u_{zk} &= n^{\frac{1}{2}u}\Big[I + II + III\Big] \\
&= n^{\frac{1}{2}u}\Big[O_p(n^{(\frac{1}{2}-\frac{1}{2}u)\alpha_k}) + \mathbf{1}_{\{u>0\}}[\bar{O}_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}) \\
&\qquad + \bar{O}_p(n^{1-2\alpha_k})]O_p(n^{(1-\frac{1}{2}u)\alpha_k})\Big] \\
&= n^{\frac{1}{2}u}\Big[O_p(n^{(\frac{1}{2}-\frac{1}{2}u)\alpha_k}) + \mathbf{1}_{\{u>0\}}\big[O_p(min\{n^{(1-\frac{1}{2}u)\alpha_k}, n^{(1-\frac{1}{2}u)\alpha_k+\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}\}) \\
&\qquad + O_p(min\{n^{(1-\frac{1}{2}u)\alpha_k}, n^{(1-\frac{1}{2}u)\alpha_k+1-2\alpha_k}\big]\Big] \\
&= n^{(1-\frac{1}{2}u)\alpha_k+\frac{1}{2}u}\Big[O_p(n^{-\frac{1}{2}\alpha_k}) + \mathbf{1}_{\{u>0\}}\big[O_p(min\{1, n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k}\}) \\
&\qquad + O_p(min\{1, n^{1-2\alpha_k}\})\big],\Big] \\
&= n^{(1-\frac{1}{2}u)\alpha_k+\frac{1}{2}u}\Big[O_p(n^{-\frac{1}{2}\alpha_k}) + \mathbf{1}_{\{u>0\}}\big[\bar{O}_p(n^{\frac{1}{4}\alpha_1-\frac{1}{2}\alpha_k})\bar{O}_p(n^{1-2\alpha_k})\big]\Big].
\end{aligned}$$

For (c), with $\alpha_k \leq \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$:

$$\begin{aligned}
\Upsilon^u_{zk} - \hat{\Upsilon}^u_{zk} &= n^{\frac{1}{2}u}\Big[I + II + III\Big] \\
&= n^{\frac{1}{2}u}\Big[O_p(\frac{n^{(\frac{1}{2}+\frac{1}{2}u)\alpha_k-\tau u}}{g(n)^u}) + \mathbf{1}_{\{u>0\}}O_p(\frac{n^{(1+\frac{1}{2}u)\alpha_k-\tau u}}{g(n)^u})\Big] \\
&= \frac{n^{(1+\frac{1}{2}u)\alpha_k+(\frac{1}{2}-\tau)u}}{g(n)^u}\Big[O_p(n^{-\frac{1}{2}\alpha_k}) + \mathbf{1}_{\{u>0\}}O_p(1)\Big].
\end{aligned}$$

We conclude, combining the above with Lemma 13, that

1. For $\alpha_k > \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$: $\hat{\Upsilon}^u_{zk} = \Upsilon^u_{zk} + o_p(\Upsilon^u_{zk})$,
2. For $0 < \alpha_k \leq \max\{\frac{1+\tau}{3}, \frac{\alpha_1+4\tau}{6}\}$: $\hat{\Upsilon}^u_{zk} = \Upsilon^u_{zk} + \mathbf{1}_{\{u>0\}}O_p(\Upsilon^u_{zk}) + o_p(\Upsilon^u_{zk})$,
3. For $\alpha_k = 0$ and $k > r$: $\hat{\Upsilon}^u_{zk} = O_p(n^{(\frac{1}{2}-\tau)u}g(n)^{-u})$,

which finishes the proof. □

B.3 Arbitrage Pricing Theory

We assume that the n-vector of demeaned asset returns $R_t - \mathbb{E}(R_t)$ for a given t follows a factor structure with potentially local factors as in the previous sections:

$$R_i - \mathbb{E}(R_i) = \underset{(1\times r)(r\times 1)}{\lambda_i' \; F} + \underset{(1\times 1)}{e_i} = \underset{(1\times K)(K\times 1)}{\lambda_i^{K'} \; F^K} + e_i^K, \tag{12}$$

treating the factors as random and the errors as uncorrelated with the factors. Equation (12) emphasizes again that, in the framework of this paper, we can always choose to move some of the weaker factors into the error structure at the expense of more correlation in the error term. Denote the return of a portfolio by $R^p = \sum_i^n w_i R_i$, with $\sum_i^n w_i = 1$. We formalize the term "well-diversified" by imposing a bound on the sup-norm of the weights: $|w_i| \leq W_n \; \forall i$. Following Green and Hollifield [1992], we say that exact APT pricing holds if the mean returns are in the span of the factor loadings and a constant vector:

$$\mathbb{E}(R_j) = (1 - \sum_k^K \lambda_{jk})\, \mathbb{E}(R_0^*) + \sum_k^K \lambda_{jk}\, \mathbb{E}(R_k^*),$$

where the portfolios $R_k^*, k = 0, \ldots, K^*$ are "factor-mimicking" portfolios. Their construction is detailed in the proof of Proposition 1 below, and conditions for their existence are given in Huberman et al. [1987]. Similarly, we define exact APT to hold in the limit, if, as n increases, there exist sequences of feasible factor-mimicking portfolios R_{nk}^*[1], such that for any fixed j

$$lim_{n\to\infty}\, \mathbb{E}(R_j) - [(1 - \sum_k^K \lambda_{jk})\, \mathbb{E}(R_{n0}^*) + \sum_k^K \lambda_{jk}\, \mathbb{E}(R_{nk}^*)] = 0.$$

Finally denote by ν_n the return on the global minimum variance portfolio when there are n assets and assume that the mean-variance frontier does not become vertical in the limit, such that there remains a meaningful trade-off between mean and variance.[2] We then obtain the following proposition:

Proposition 1. *Consider the sequence of efficient (minimum variance) portfolios for some mean return $\mu \neq lim_{n\to\infty}\nu_n$. If*

(i) $W_n = o(\frac{1}{n^\gamma})$, $\gamma > \frac{1}{2}$ *for every such portfolio, and*

(ii) $\lim_{n\to\infty}(\max_j \sum_{i=1}^n \left|Cov(e_i, e_j)\right|) = O(\sqrt{n})$,

[1] R_{n0}^* will be the minimum-variance portfolio with zero loadings

[2] This is the equivalent of the "absence of arbitrage" assumption in the Hilbert space setting of Chamberlain and Rothschild [1983].

then exact APT pricing holds in the limit with respect to the strongest K factors, where K is defined such that $\alpha_k > \gamma$ for $k = 1, 2, \ldots, K$ and $\alpha_k \leq \gamma$ for $k \geq K + 1$.

Proof. The proof largely follows the proof of Theorem 3 in Green and Hollifield [1992]. Define the set of demeaned portfolios

$$\Xi_n = \{R^p - \mathbb{E}(R^p) : R^p = \sum_{i=1}^{n} w_i R_i, \sum_{i=1}^{n} w_i = 1\}$$

and construct the factor-mimicking portfolios by projecting the zero vector and the strongest K factors $k = 1, \ldots, K$ onto Ξ_n, such that:

$$F_k = R^*_{nk} - \mathbb{E}(R^*_{nk}) + \xi_{nk},$$

where $E(\xi_{nk} R_j) = 0$ for $j = 1, \ldots, n$. For asset j, consider the combination of K factor-mimicking portfolios with the same factor risk:

$$R^{K*}_{nj} = (1 - \sum_{k}^{K} \lambda_{jk}) R^*_{n0} + \sum_{k}^{K} \lambda_{jk} R^*_{nk}.$$

Let

$$\begin{aligned}
\Pi^K_{nj} &= R_j - R^{K*}_{nj} \\
&= R_j - \mathbb{E}(R_j) + \mathbb{E}(R_j) - (1 - \sum_{k}^{K} \lambda_{jk}) R^*_{n0} - \sum_{k}^{K} \lambda_{jk} R^*_{nk} \\
&= c^K_j + \sum_{k}^{K} \lambda_{jk} F_k + e^K_j - (1 - \sum_{k}^{K} \lambda_{jk})[R^*_{n0} - \mathbb{E}(R^*_{n0})] - \sum_{k}^{K} \lambda_{jk}[R^*_{nk} - \mathbb{E}(R^*_{nk})] \\
&= c^K_j + (1 - \sum_{k}^{K} \lambda_{jk}) \xi^{n0} + \sum_{k}^{K} \lambda_{jk} \xi_{nk} + e^K_j, \qquad (13)
\end{aligned}$$

with

$$c^K_j = \mathbb{E}(R_j) - \left((1 - \sum_{k}^{K} \lambda_{jk}) \, \mathbb{E}(R^*_{n0}) + \sum_{k}^{K} \lambda_{jk} \, \mathbb{E}(R^*_{nk}) \right).$$

Recalling that W_n denotes the sup-norm on the asset weights w_i, we can invoke the following result by Green and Hollifield [1992].

Theorem (Theorem 1 of Green and Hollifield [1992])**.** *The efficient portfolio with mean $\mu \neq \nu$ is* *well diversified (i.e. $|w_i| \leq W_n \ \forall i$) if and only if the return, R^*, on every portfolio with weights that sum to one, satisfies*

$$\left|\mathbb{E}(R^*) - \mathbb{E}(R_z)\right| \leq \left|\frac{W_n}{\gamma_n}\right| \sum_{i=1}^{n} \left|Cov(R^*, R_i)\right|,$$

and the payoff, Π^*, on every hedge position with weights that sum to zero, satisfies

$$\left|\mathbb{E}(\Pi^*)\right| \leq \left|\frac{W_n}{\gamma_n}\right| \sum_{i=1}^{n} \left|Cov(\Pi^*, R_i)\right|,$$

where γ_n is uniformly bounded away from zero by the assumption of no asymptotic arbitrage.

Therefore, if the efficient frontier contains a well-diversified portfolio, this implies that

$$\left|\mathbb{E}(\Pi_{nj}^K)\right| \leq \left|\frac{W_n}{\gamma_n}\right| \sum_{i=1}^{n} \left|Cov(\Pi_{nj}^K, R_i)\right|,$$

because Π_j^n is the return on a hedge position with weights summing to zero. By (13), $Cov(\Pi_{nj}^K, R_i) = Cov(e_i^K, e_j^K)$ and thus:

$$\begin{aligned}
\left|\mathbb{E}(\Pi_{nj}^K)\right| &\leq \left|\frac{W_n}{\gamma_n}\right| \sum_{i=1}^{n} \left|Cov(e_i^K, e_j^K)\right| \\
&= \left|\frac{W_n}{\gamma_n}\right| \sum_{i=1}^{n} \left|Cov(F^w \lambda_i + e_i, F^w \lambda_j + e_j)\right| \\
&\leq \left|\frac{W_n}{\gamma_n}\right| \left(\sum_{i=1}^{n} \sum_{k=K+1}^{r} \lambda_{ik}\lambda_{jk} + \sum_{i=1}^{n} \left|Cov(e_i, e_j)\right| \right) \\
&= \left|\frac{W_n}{\gamma_n}\right| \left(\sum_{k=K+1}^{r} \lambda_{jk} \sum_{i=1}^{n} \lambda_{ik} + \sum_{i=1}^{n} \left|Cov(e_i, e_j)\right| \right) \\
&= \left|\frac{W_n}{\gamma_n}\right| \left(\sum_{k=K+1}^{r} \lambda_{jk} \left[\sum_{i \in \mathcal{A}_k} \lambda_{ik} + \sum_{i \notin \mathcal{A}_k} \lambda_{ik} \right] + \sum_{i=1}^{n} \left|Cov(e_i, e_j)\right| \right. \\
&= \left|\frac{W_n}{\gamma_n}\right| \left(\sum_{k=K+1}^{r} \lambda_{jk} \left[O(n^{\alpha_k}) + O(\sqrt{n}) \right] + \sum_{i=1}^{n} \left|Cov(e_i, e_j)\right| \right. \\
&\leq \left|\frac{W_n}{\gamma_n}\right| \left(\sum_{k=K+1}^{r} O(n^{\alpha_k}) + O(\sqrt{n}) + O(\sqrt{n}) \right) \\
&\leq \left|\frac{W_n}{\gamma_n}\right| \Big(O(n^{\alpha_{K+1}}) + O(\sqrt{n}) + O(\sqrt{n}) \Big).
\end{aligned}$$

We therefore conclude that

$$lim_{n\to\infty}\left[\mathbb{E}(R_j)-\left((1-\sum_k^K\lambda_{jk})\,\mathbb{E}(R^*_{n0})+\sum_k^K\lambda_{jk}\,\mathbb{E}(R^*_{nk})\right)\right]$$
$$=lim_{n\to\infty}\left|\mathbb{E}(\Pi^K_{nj})\right|=lim_{n\to\infty}W_nO\left(\max(n^{\alpha_{K+1}},\sqrt{n}\right)=0,$$

whenever $W_n=o\left(\min(n^{-\alpha_{K+1}},n^{-\frac{1}{2}})\right)$. This completes the proof. □

Proposition 1 states that exact APT holds in the limit if the efficient portfolios are well diversified. Further, the number of factors that are priced depends directly on the degree of diversification of the portfolios on the efficient frontier. The better diversified these portfolios are (the smaller W_n), the smaller the number of factors that have a non-zero factor premium.

In particular, with $W_n=o(\frac{1}{\sqrt{n}})$, which yields diversification in the sense of Chamberlain and Rothschild [1983] and Chamberlain [1983], Proposition 1 establishes that exact APT pricing holds in the limit with respect to the r_1 factors affecting proportionally more than $\sqrt{n}$ of the assets (factors with $\alpha_k>.5$).

Proposition 1 holds under more general conditions than the approximate factor model of Chamberlain and Rothschild [1983]. We do not require all eigenvalues of the error covariance matrix to be bounded, but explicitly allow for additional, weaker factors. Instead of ruling out the existence of such weaker factors, Proposition 1 establishes that they will not be priced.

B.4 Aggregate Fluctuations in the Economy

Firm i produces a quantity S_{it} of the consumption good. Firm-level growth rates have a factor structure as follows:

$$\frac{\Delta S_{i,t+1}}{S_{it}}=\frac{S_{i,t+1}-S_{it}}{S_{it}}=\lambda_iF_{t+1}+\sigma_i\varepsilon_{i,t+1}, \tag{14}$$

where $\sigma_i<\infty$ is firm i's volatility, and the $\varepsilon_{i,t+1}$ are uncorrelated random variables with mean zero and unit variance. Thus, firms' growth rates may be correlated through the presence of the first component. However, we do not impose the factors to be pervasive and likely $\lambda_{ik}=0$ for most firm-factor combinations. Intuitively, these factors can correspond to economy-wide shocks but also industry shocks, including shocks that affect as few as two firms. Thus (14) is quite general.

In this stylized model, GDP growth is given by:

$$\frac{\Delta Y_{t+1}}{Y_t} = \frac{1}{Y_t}\sum_{i=1}^{n} \Delta S_{i,t+1} = \sum_{i=1}^{n} \frac{S_{it}}{Y_t}[\lambda_i F_{t+1} + \varepsilon_{i,t+1}]$$
$$= \sum_{i=1}^{n} \frac{S_{it}}{Y_t}\lambda_i F_{t+1} + \sum_{i=1}^{n} \frac{S_{it}}{Y_t}\varepsilon_{i,t+1}.$$

It follows that the variance of GDP growth at time $(t+1)$ conditional on time t information is equal to

$$Var_t\left(\sum_{i=1}^{n} \frac{S_{it}}{Y_t}\lambda_i F_{t+1} + \sum_{i=1}^{n} \frac{S_{it}}{Y_t}\varepsilon_{i,t+1}\right) = Var_t\left(\sum_{i=1}^{n} \frac{S_{it}}{Y_t}\lambda_i F_{t+1}\right) + Var_t\left(\sum_{i=1}^{n} \frac{S_{it}}{Y_t}\varepsilon_{i,t+1}\right)$$
$$= Var_t\left(\sum_{i=1}^{n} \frac{S_{it}}{Y_t}\sum_{k=1}^{r} \lambda_{ik} F_{k,t+1}\right) + \sum_{i=1}^{n}\left(\frac{S_{it}}{Y_t}\right)^2 \sigma_i^2.$$

For ease of notation, consider firms of equal size ($S_{it} = \frac{Y_t}{n}$) and normalize the factors such that $Var(F_{kt}) = 1$. Further assume that, for a given k, the factor loadings are 1 on a subset of size $|\mathcal{A}_k| \asymp n^{\alpha_k}$ and zero everywhere else[3]. Then:

$$Var_t\left(\frac{\Delta Y_{t+1}}{Y_t}\right) = \sum_{k=1}^{r}\left(\sum_{i\in\mathcal{A}_k}\frac{1}{n}\right)^2 + \sum_{i=1}^{n}\frac{1}{n^2}\sigma_i^2$$
$$\asymp \sum_{k=1}^{r} n^{2\alpha_k - 2} + O_p(\frac{1}{n}). \tag{15}$$

Absent any factors ($r = 0$), clearly $\sigma_{GDP} = \sqrt{Var_t(\frac{\Delta Y_{t+1}}{Y_t})} = \frac{\sigma}{\sqrt{n}}$, and idiosyncratic fluctuations disappear in the aggregate at rate $\sqrt{n}$. Next, consider an economy with r shocks, where r_1 is the number of factors with $\alpha_k > .5$:

$$Var_t\left(\frac{\Delta Y_{t+1}}{Y_t}\right) \asymp \sum_{k=1}^{r_1} n^{2\alpha_k - 2} + \sum_{k=r_1+1}^{r} n^{2\alpha_k - 2} + \frac{\sigma}{n}$$
$$= \sum_{k=1}^{r_1} n^{2\alpha_k - 2} + O_p(\frac{1}{n}).$$

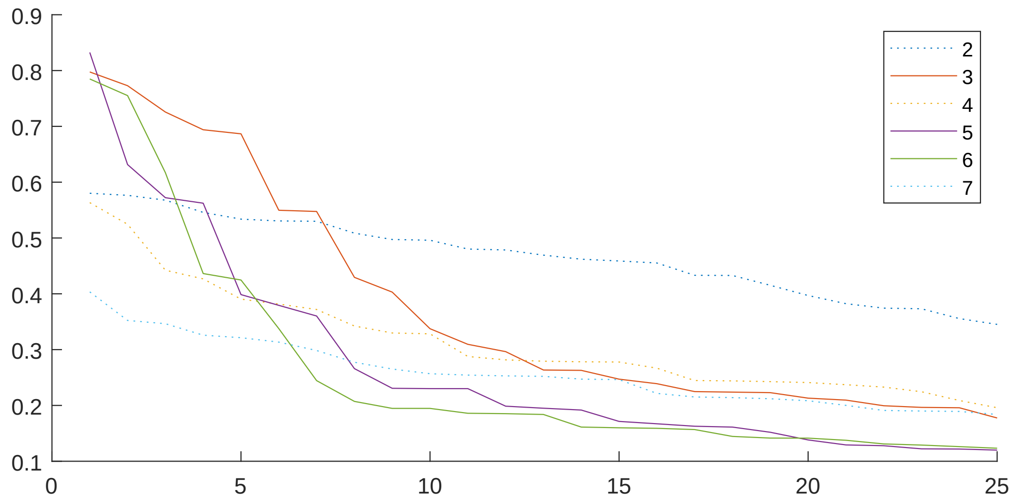
Equation (15) establishes that the important shocks are those with $\alpha_k > \frac{1}{2}$ and that the standard rate of convergence breaks down whenever shocks exist that affect more than $\sqrt{n}$ firms.

[3]Defining the loadings instead in a more general way as in Assumption 2 does not alter any conclusions.

This is in line with the granularity conditions derived in Gabaix [2011], who considers heterogeneous firm sizes that may grow with n. Intuitively, with the growth rate of the economy given by the sum of both the idiosyncratic and factor shocks in our context, we can think of the sector shocks as additional but larger firms. Then the economy consists of $n + r$ components (with $r << n$). Proposition 2 in Gabaix [2011] establishes that $\sigma_{GDP} \asymp \frac{1}{\sqrt{n}}$ only if the largest firm has a relative weight of at most $W_n = O(\frac{1}{\sqrt{n}})$. This corresponds exactly to the limit on sector size stated above.

The key implication for the purposes of this paper is that, in order to understand the origins of fluctuations, the important shocks are precisely those that affect proportionally more than $\sqrt{n}$ firms.

C Additional Results for Empirical Application



Online Appendix Figure 6: Absolute value of 25 largest loadings for factors 2-7. Each line corresponds to the largest 25 loadings (in absolute value) for a specific factor. Solid lines correspond to the "local" factors 3, 5, and 6.

Factor 3	Factor 5	Factor 6
PPI: Int. Material: Supplies & Components	**Nonfarm: Unit Nonlabor Payments**	**tb6m-tb3m**
PPI: Industrial Commodities	**Nonfarm: Unit Labor Cost**	**GS1-Tb3m**
PPI: Finished Consumer Goods	**Nonfarm: Real Compensation Per Hour**	**GS10-Tb3m**
PPI: Crude Petroleum Defl by PCE(LFE)	**BS: Real Compensation Per Hour**	**S&P'S STOCK PRICE INDEX**
Gasoline and other energy goods	PPI: Finished Consumer Foods	**DOW JONES IA**
BS: Real Compensation Per Hour	Food & beverages for off-premises consump	Consumer Loans, All Commercial Banks
Nonfarm: Real Compensation Per Hour	**Nonfarm: Output Per Hour of All Persons**	**BAA-GS10 Spread**
BS: Implicit Price Deflator	PPI: Finished Consumer Goods	
ISM Manufacturing: Prices Paid Index		

Online Appendix Table 2: Variables corresponding to largest loadings for factors 3, 5 and 6, the most local factors. Red coloring indicates a negative loading, while black indicates a positive loading. For factor 3, we note that six of the nine variables, printed in bold, represent price indices as classified in the handbook chapter of Stock and Watson [2016]. Additionally the fourth entry, while classified as an "Oil market variable," also represents a price index. The remaining two variables are both classified as "Productivity and Earnings" and it is worth noting that they have the opposite sign. Next, of the five series classified as "Productivity and Earnings" in the data, all five of these are associated with factor 5, emphasized in bold. The remaining three entries are all price indices. The 6th factor is highly concentrated on spreads and stock market indicators (again emphasized in bold). In fact, this factor is associated with a negative return on the stock market and an increase in the interest rate spread. The 6th factor could thus be interpreted as indicating a flight from stocks into safe assets, such as bonds.