SUPPLEMENTARY MATERIALS

Time-varying risk premium in large cross-sectional equity datasets

Patrick Gagliardini, Elisa Ossola and Olivier Scaillet

These supplementary materials provide the proofs of the technical lemmas used in the paper (Appendix 5) and the results of Monte-Carlo experiments that investigate the finite-sample properties of the estimators and test statistics (Appendix 6). We also derive inference for the cost of equity and include some empirical results for Ford Motor, Disney Walt, Motorola and Sony (Appendix 7). Finally, we provide some robustness checks for the empirical analysis (Appendix 8).

Appendix 5: Proofs of the technical lemmas

A.5.1 Proof of Lemma 1 (iii)

We have $\hat{w}_i - w_i = \mathbf{1}_i^{\chi}(\hat{v}_i^{-1} - v_i^{-1}) + (\mathbf{1}_i^{\chi} - 1)v_i^{-1}$ and $\hat{v}_i^{-1} - v_i^{-1} = -\hat{v}_i^{-1}v_i^{-1}(\hat{v}_i - v_i)$. Since v_i is uniformly lower bounded from part (ii), we have $\frac{1}{n}\sum_i |\hat{w}_i - w_i| \leq C\frac{1}{n}\sum_i \mathbf{1}_i^{\chi}\frac{|\hat{v}_i - v_i|}{C - |\hat{v}_i - v_i|} + C\frac{1}{n}\sum_i (1 - \mathbf{1}_i^{\chi})$. The second term in the RHS is $o_p(1)$ from Lemma 4. To prove that the first term is $o_p(1)$ it is sufficient to show:

$$\sup_i \mathbf{1}_i^{\chi}|\hat{v}_i - v_i| = o_p(1). \tag{33}$$

We use Equation (24). Since $\hat{\nu}_1 - \nu = O_p(T^{-c})$, for some $c > 0$ (by repeating the proof of Proposition 2 with known weights equal to 1), $\mathbf{1}_i^{\chi}\|\hat{Q}_{x,i}^{-1}\| \leq C\chi_{1,T}$, $\mathbf{1}_i^{\chi}\tau_{i,T} \leq \chi_{2,T}$, $\|S_{ii}\| \leq M$, and by using Assumption C.5, the uniform bound in (33) follows if we prove:

$$\sup_i \mathbf{1}_i^{\chi}\|\hat{S}_{ii} - S_{ii}\| = O_p(T^{-c}), \tag{34}$$

$$\sup_i \mathbf{1}_i^{\chi}\|\hat{Q}_{x,i}^{-1} - Q_x^{-1}\| = O_p(T^{-c}), \tag{35}$$

$$\sup_i \mathbf{1}_i^{\chi}|\tau_{i,T} - \tau_i| = O_p(T^{-c}), \tag{36}$$

for some $c > 0$. To prove the uniform bound (34), we use Equation (26). As in the proof of Lemma 1 (i), we have $\sup_i T^{-1/2}\|Y_{i,T}\| = O_{p,log}(T^{-\eta/2})$ from Assumption C.1 c), and similarly $\sup_i T^{-1/2}\|W_{1,i,T} + W_{2,i,T}\| = O_{p,log}(T^{-\eta/2})$ and $\sup_i T^{-1/2}\|W_{3,i,T}\| = O_p(T^{-\eta/2})$, from Assumptions C.1 e) and f), respectively. Moreover, $\|\hat{Q}_{x,i}^{(4)}\| \leq M$ and $\mathbf{1}_i^{\chi}\tau_{i,T} \leq \chi_{2,T}$. Thus, from Assumption C.5, bound (34) follows. To prove (35) we use $\hat{Q}_{x,i}^{-1} - Q_x^{-1} = -\tau_{i,T}\hat{Q}_{x,i}^{-1}W_{i,T}Q_x^{-1}$, where $W_{i,T}$ is defined as in Equation (27) and is such that $\sup_i \|W_{i,T}\| = O_{p,log}(T^{-\eta/2})$ from Assumption C.1 b). Finally, (36) follows from $|\tau_{i,T} - \tau_i| \leq \tau_{i,T}\tau_i \left|\frac{1}{T}\sum_t (I_{i,t} - E[I_{i,t}|\gamma_i])\right|$, $\mathbf{1}_i^{\chi}\tau_{i,T} \leq \chi_{2,T}$, $\tau_i \leq M$ and by using $\sup_i \left|\frac{1}{T}\sum_t (I_{i,t} - E[I_{i,t}|\gamma_i])\right| = O_{p,log}(T^{-\eta/2})$ from Assumption C.1 d).

A.5.2 Proof of Lemma 3

A.5.2.1 Part i)

Let us write I_{21} as:

$$\begin{aligned} I_{21} &= \frac{1}{\sqrt{n}}\sum_i \hat{w}_i\tau_{i,T}^2\hat{Q}_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\hat{Q}_{x,i}^{-1} \\ &= \frac{1}{\sqrt{n}}\sum_i \hat{w}_i\tau_{i,T}^2\hat{Q}_x^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\hat{Q}_x^{-1} + \frac{1}{\sqrt{n}}\sum_i \hat{w}_i\tau_{i,T}^2\left(\hat{Q}_{x,i}^{-1} - \hat{Q}_x^{-1}\right)\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\hat{Q}_x^{-1} \\ &\quad + \frac{1}{\sqrt{n}}\sum_i \hat{w}_i\tau_{i,T}^2\hat{Q}_x^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\left(\hat{Q}_{x,i}^{-1} - \hat{Q}_x^{-1}\right) \\ &\quad + \frac{1}{\sqrt{n}}\sum_i \hat{w}_i\tau_{i,T}^2\left(\hat{Q}_{x,i}^{-1} - \hat{Q}_x^{-1}\right)\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\left(\hat{Q}_{x,i}^{-1} - \hat{Q}_x^{-1}\right) \\ &=: \hat{Q}_x^{-1}I_{211}\hat{Q}_x^{-1} + I_{212}\hat{Q}_x^{-1} + \hat{Q}_x^{-1}I_{212}' + I_{213}. \end{aligned}$$

We control the terms separately.

Proof that $I_{211} = \frac{1}{\sqrt{n}}\sum_i w_i\tau_i^2\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right) + O_{p,log}(\sqrt{n}/T) = O_p(1) + O_{p,log}(\sqrt{n}/T)$. We use a decomposition similar to term I_{11} in the proof of Lemma 2:

$$\begin{aligned} I_{211} &= \frac{1}{\sqrt{n}}\sum_i w_i\tau_i^2\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)+\frac{1}{\sqrt{n}}\sum_i\left(\mathbf{1}_i^\chi-1\right)w_i\tau_i^2\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right) \\ &+\frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^\chi w_i\left(\tau_{i,T}^2-\tau_i^2\right)\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right) \\ &+\frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^\chi\left(\hat{v}_i^{-1}-v_i^{-1}\right)\tau_{i,T}^2\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right) \qquad =: I_{2111}+I_{2112}+I_{2113}+I_{2114}. \end{aligned}$$

To prove $I_{2111}=O_p(1)$, take $k,l=1,...,K$, and consider $\zeta_{nT}:=\frac{1}{\sqrt{n}}\sum_i w_i\tau_i^2\left(Y_{i,k,T}Y_{i,l,T}-S_{ii,kl,T}\right)$. Then:

$$\begin{aligned} &E[\zeta_{nT}^2|x_{\underline{T}},I_{\underline{T}},\{\gamma_i\}]=\frac{1}{n}\sum_{i,j}w_iw_j\tau_i^2\tau_j^2 cov\left(Y_{i,k,T}Y_{i,l,T},Y_{j,k,T}Y_{j,l,T}|x_{\underline{T}},I_{\underline{T}},\gamma_i,\gamma_j\right) \\ &= \frac{1}{nT^2}\sum_{i,j}\sum_{t_1,t_2,t_3,t_4} w_iw_j\tau_i^2\tau_j^2 cov\left(\varepsilon_{i,t_1}\varepsilon_{i,t_2},\varepsilon_{j,t_3}\varepsilon_{j,t_4}|x_{\underline{T}},\gamma_i,\gamma_j\right)I_{i,t_1}I_{i,t_2}I_{j,t_3}I_{j,t_4}x_{t_1,k}x_{t_2,l}x_{t_3,k}x_{t_4,l}. \end{aligned}$$

From Assumptions A.1 c), C.3 b) and C.4, it follows $E[\zeta_{nT}^2]=O(1)$. Hence, $\zeta_{nT}=O_p(1)$ and $I_{2111}=O_p(1)$. We can prove that $I_{2112}=o_p(1)$ and $I_{2113}=o_p(1)$ by using arguments similar to terms I_{112} and I_{113} in the proof of Lemma 2. Finally, let us prove that $I_{2114}=O_{p,log}(\sqrt{n}/T)$. Similarly to I_{114} in the proof of Lemma 2, we use

$$\hat{v}_i^{-1}-v_i^{-1}=-v_i^{-2}\left(\hat{v}_i-v_i\right)+\hat{v}_i^{-1}v_i^{-2}\left(\hat{v}_i-v_i\right)^2, \tag{37}$$

and Equation (24). We focus on term:

$$I_{21141}=-\frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^\chi v_i^{-2}\tau_{i,T}^3 c'_{\hat{\nu}_1}\hat{Q}_{x,i}^{-1}\left(\hat{S}_{ii}-S_{ii}\right)\hat{Q}_{x,i}^{-1}c_{\hat{\nu}_1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right),$$

the other contributions to I_{2114} can be controlled similarly. Now, we use Equation (26) and treat x_t as a scalar to ease notation. We have:

$$
\begin{aligned}
I_{21141} &= -\frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} v_i^{-2}\tau_{i,T}^4 c'_{\hat{\nu}_1}\hat{Q}_{x,i}^{-1}W_{1,i,T}\hat{Q}_{x,i}^{-1}c_{\hat{\nu}_1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)\\
&\quad -\frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} v_i^{-2}\tau_{i,T}^4 c'_{\hat{\nu}_1}\hat{Q}_{x,i}^{-1}W_{2,i,T}\hat{Q}_{x,i}^{-1}c_{\hat{\nu}_1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)\\
&\quad +2\frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} v_i^{-2}\tau_{i,T}^5 c'_{\hat{\nu}_1}\hat{Q}_{x,i}^{-1}W_{3,i,T}\hat{Q}_{x,i}^{-1}Y_{i,T}\hat{Q}_{x,i}^{-1}c_{\hat{\nu}_1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)\\
&\quad -\frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} v_i^{-2}\tau_{i,T}^6 c'_{\hat{\nu}_1}\hat{Q}_{x,i}^{-1}\hat{Q}_{x,i}^{(4)}\hat{Q}_{x,i}^{-1}Y_{i,T}Y'_{i,T}\hat{Q}_{x,i}^{-2}c_{\hat{\nu}_1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)\\
&=: -c'_{\hat{\nu}_1}(I_{211411}+I_{211412}+I_{211413}+I_{211414})c_{\hat{\nu}_1}.
\end{aligned}
$$

Let us focus on term I_{211411} and prove that it is $O_{p,log}(\sqrt{n}/T)$. We have:

$$
I_{211411} = \frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} v_i^{-2}\tau_{i,T}^4\hat{Q}_{x,i}^{-2}W_{1,i,T}Y_{i,T}^2 - \frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^{\chi} v_i^{-2}\tau_{i,T}^4\hat{Q}_{x,i}^{-2}W_{1,i,T}S_{ii,T} =: I_{2114111}+I_{2114112}.
$$

Term $I_{2114111}$ is such that:

$$
|E[I_{2114111}|x_{\underline{T}}, I_{\underline{T}}, \{\gamma_i\}]| \leq \frac{C\chi_{1,T}^2\chi_{2,T}^4}{\sqrt{n}T^2}\sum_i\sum_{t_1,t_2,t_3}|E[\eta_{i,t_1}\varepsilon_{i,t_2}\varepsilon_{i,t_3}|x_{\underline{T}},\gamma_i]|,
$$

and

$$
V[I_{2114111}|x_{\underline{T}}, I_{\underline{T}}, \{\gamma_i\}] \leq \frac{C\chi_{1,T}^4\chi_{2,T}^8}{nT^4}\sum_{i,j}\sum_{t_1,\dots,t_6}|cov(\eta_{i,t_1}\varepsilon_{i,t_2}\varepsilon_{i,t_3},\eta_{j,t_4}\varepsilon_{j,t_5}\varepsilon_{j,t_6}|x_{\underline{T}},\gamma_i,\gamma_j)|.
$$

From Assumptions C.2, C.3 f) and C.5, we get $E[I_{2114111}] = O_{log}(\sqrt{n}/T)$ and $V[I_{2114111}] = o(1)$, which implies $I_{2114111} = O_{p,log}(\sqrt{n}/T)$. The other terms making I_{2114} can be controlled similarly, and we get $I_{2114} = O_{p,log}(\sqrt{n}/T)$.

Proof that $I_{212} = o_p(1)$. We have:

$$
\begin{aligned}
I_{212} &= \frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^{\chi} v_i^{-1}\tau_{i,T}^2\left(\hat{Q}_{x,i}^{-1}-\hat{Q}_x^{-1}\right)\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)\\
&\quad +\frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^{\chi}(\hat{v}_i^{-1}-v_i^{-1})\tau_{i,T}^2\left(\hat{Q}_{x,i}^{-1}-\hat{Q}_x^{-1}\right)\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right) =: I_{2121}+I_{2122}.
\end{aligned}
$$

We focus on term I_{2121}, use Equation (27) and treat x_t as a scalar to ease notation. We have:

$$
\begin{aligned}
I_{2121} &= -\frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^{\chi} v_i^{-1}\tau_{i,T}^3\hat{Q}_{x,i}^{-1}W_{i,T}\hat{Q}_x^{-1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right)\\
&\quad +\frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^{\chi} v_i^{-1}\tau_{i,T}^2\hat{Q}_{x,i}^{-1}W_T\hat{Q}_x^{-1}\left(Y_{i,T}Y'_{i,T}-S_{ii,T}\right) =: (I_{21211}+I_{21212})\hat{Q}_x^{-1}.
\end{aligned}
$$

Let us focus on I_{21211}. We have:

$$E[\|I_{21211}\|^2|x_{\underline{T}}, I_{\underline{T}}, \{\gamma_i\}] \leq \frac{C\chi_{1,T}^2\chi_{2,T}^6}{nT^2}\sum_{i,j}\sum_{t_1,...,t_4}\|W_{i,T}\|\|W_{j,T}\||cov(\varepsilon_{i,t_1}\varepsilon_{i,t_2}, \varepsilon_{j,t_3}\varepsilon_{j,t_4}|x_{\underline{T}}, \gamma_i, \gamma_j)|.$$

By the Cauchy-Schwarz inequality, we get:

$$\begin{aligned} E[\|I_{21211}\|^2|\{\gamma_i\}] &\leq C\chi_{1,T}^2\chi_{2,T}^6 \sup_i E[\|W_{i,T}\|^4|\gamma_i]^{1/2} \\ &\quad \frac{1}{nT^2}\sum_{i,j}\sum_{t_1,t_2,t_3,t_4} E\left[|cov(\varepsilon_{i,t_1}\varepsilon_{i,t_2}, \varepsilon_{j,t_3}\varepsilon_{j,t_4}|x_{\underline{T}}, \gamma_i, \gamma_j)|^2|\gamma_i, \gamma_j\right]^{1/2}. \end{aligned}$$

From Assumptions C.1 b), C.3 b), C.4 a), and C.5, we deduce $E[\|I_{21211}\|^2|] = o(1)$, which implies $I_{21211} = o_p(1)$. Similar argument can be used to prove that the other terms making I_{212} are $o_p(1)$.

Proof that $I_{213} = o_p(1)$. This step uses arguments similar as for I_{212}.

A.5.2.2 Part (ii)

We have $I_{22} = \frac{1}{\sqrt{nT}}\sum_i \hat{w}_i\tau_{i,T}^2\hat{Q}_{x,i}^{-1}W_{1,i,T}\hat{Q}_{x,i}^{-1}$, where $W_{1,i,T}$ is as in Equation (26). Write:

$$I_{22} = \frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^\chi v_i^{-1}\tau_{i,T}^2\hat{Q}_{x,i}^{-1}W_{1,i,T}\hat{Q}_{x,i}^{-1} + \frac{1}{\sqrt{nT}}\sum_i \mathbf{1}_i^\chi(\hat{v}_i^{-1} - v_i^{-1})\tau_{i,T}^2\hat{Q}_{x,i}^{-1}W_{1,i,T}\hat{Q}_{x,i}^{-1} =: I_{221} + I_{222}.$$

Let us first consider I_{221}. We have:

$$E[\|I_{221}\|^2|x_{\underline{T}}, I_{\underline{T}}, \{\gamma_i\}] \leq C\chi_{1,T}^4\chi_{2,T}^4\frac{1}{nT^2}\sum_{i,j}\sum_{t_1,t_2}|cov(\eta_{i,t_1}, \eta_{j,t_2}|x_{\underline{T}}, \gamma_i, \gamma_j)|.$$

From Assumptions C.3 a) and C.5, it follows $E[\|I_{222}\|^2] = O_{log}(1/T)$, and thus $I_{222} = O_{p,log}(1/\sqrt{T})$.

Let us now consider term I_{222}. We use Equation (37), and plug in the decompositions (24) and (26). We focus on term $c_{\nu_1}^2 I_{2221}$ of the resulting expansion, where:

$$I_{2221} = -\frac{1}{\sqrt{n}T}\sum_i \mathbf{1}_i^\chi v_i^{-2}\tau_{i,T}^4\hat{Q}_{x,i}^{-4}W_{1,i,T}^2.$$

The other terms can be treated similarly. We have:

$$E[I_{2221}|x_{\underline{T}}, I_{\underline{T}}, \{\gamma_i\}] \leq C\chi_{1,T}^4\chi_{2,T}^4\frac{1}{\sqrt{n}T^2}\sum_i\sum_{t_1,t_2}|cov(\varepsilon_{i,t_1}^2, \varepsilon_{i,t_2}^2|x_{\underline{T}}, \gamma_i)|,$$

and

$$V[I_{2221}|x_{\underline{T}}, I_{\underline{T}}, \{\gamma_i\}] \leq C\chi_{1,T}^8\chi_{2,T}^8 \frac{1}{nT^4}\sum_{i,j}\sum_{t_1,t_2,t_3,t_4} |cov(\eta_{i,t_1}\eta_{i,t_2}, \eta_{j,t_3}\eta_{j,t_4}|x_{\underline{T}}, \gamma_i, \gamma_j)|.$$

From Assumptions C.3 a) and C.5, it follows $E[I_{2221}] = O_{log}(\sqrt{n}/T)$. By Assumptions C.3 d) and C.5 we can prove that $V[I_{2221}] = o(1)$, and it follows $I_{2221} = O_p(\sqrt{n}/T)$.

A.5.2.3 Part (iii)

We have $I_{23} = -\frac{2}{\sqrt{n}T}\sum_i \hat{w}_i\tau_{i,T}^3\hat{Q}_{x,i}^{-3}W_{3,i,T}Y_{i,T} + \frac{1}{\sqrt{n}T}\sum_i \hat{w}_i\tau_{i,T}^4\hat{Q}_{x,i}^{-4}\hat{Q}_{x,i}^{(4)}Y_{i,T}^2$, where $W_{3,i,T}$ and $\hat{Q}_{x,i}^{(4)}$ are as in Equation (26) and we treat x_t as a scalar to ease notation. By similar arguments as in part (ii) we can prove that $I_{23} = O_{p,log}(\sqrt{n}/T)$.

A.5.2.4 Part (iv)

The statement follows from Lemma 1 (ii)-(iii), $\mathbf{1}_i^{\chi}\tau_{i,T} \leq \chi_{2,T}$, $\mathbf{1}_i^{\chi}\|\hat{Q}_{x,i}^{-1}\| \leq C\chi_{1,T}$, bound (34), $\|S_{ii}\| \leq M$ and Assumption C.5.

A.5.2.5 Part (v)

The statement follows from Equation (21), Lemma 1 (iv), $I_1 = O_p(1)$ and $\frac{1}{n}\sum_i \hat{w}_i\tau_{i,T}^2E_2'\hat{Q}_{x,i}^{-1}Y_{i,T}Y_{i,T}'\hat{Q}_{x,i}^{-1} = O_{p,log}(1)$.

A.5.3 Proof of Lemma 4

We have $\mathbb{P}[\mathbf{1}_i^{\chi} = 0] \leq \mathbb{P}[\tau_{i,T} \geq \chi_{2,T}] + \mathbb{P}\left[CN\left(\hat{Q}_{x,i}\right) \geq \chi_{1,T}\right] =: P_{1,nT} + P_{2,nT}$. Let us first control $P_{1,nT}$. We have $P_{1,nT} \leq \mathbb{P}\left[\frac{1}{T}\sum_t I_{i,t} \leq \chi_{2,T}^{-1}\right] \leq \mathbb{P}\left[\frac{1}{T}\sum_t\left(I_{i,t} - \tau_i^{-1}\right) \leq \chi_{2,T}^{-1} - M^{-1}\right]$, where we use $\tau_i \leq M$ for all i (Assumption C.4 c)). Then, for $0 < \delta < M^{-1}/2$ and T large such that $M^{-1} - \chi_{2,T}^{-1} > \delta$, we get the upper bound $P_{1,nT} \leq \mathbb{P}\left[\left|\frac{1}{T}\sum_t\left(I_{i,t} - \tau_i^{-1}\right)\right| \geq \delta\right]$. By using that $\tau_i^{-1} = E[I_{i,t}|\gamma_i]$ and $\mathbb{P}\left[\left|\frac{1}{T}\sum_t\left(I_{i,t} - \tau_i^{-1}\right)\right| \geq \delta\right] = E\left[\mathbb{P}\left[\left|\frac{1}{T}\sum_t\left(I_{i,t} - E[I_{i,t}|\gamma_i]\right)\right| \geq \delta|\gamma_i\right]\right] \leq$

$\sup_{\gamma\in[0,1]} \mathbb{P}\left[\left|\frac{1}{T}\sum_t \left(I_t(\gamma) - E[I_t(\gamma)]\right)\right| \geq \delta\right]$, from Assumption C.1 d) it follows $P_{1,nT} = O(T^{-\bar{b}})$, for any $\bar{b} > 0$.

Let us now consider $P_{2,nT}$. By using $\|\hat{Q}_{x,i}\| \leq M$ (Assumption C.4 a)), we get $eig_{max}(\hat{Q}_{x,i}) \leq M$, and thus $CN\left(\hat{Q}_{x,i}\right) \leq M^{1/2}\left[eig_{min}(\hat{Q}_{x,i})\right]^{-1/2}$. Hence $P_{2,nT} \leq \mathbb{P}\left[eig_{min}(\hat{Q}_{x,i}) \leq M/\chi^2_{1,T}\right]$. By using that $eig_{min}(\hat{Q}_{x,i}) \geq eig_{min}(Q_x) - \|\hat{Q}_{x,i} - Q_x\|$, we get $P_{2,nT} \leq \mathbb{P}\left[\|\hat{Q}_{x,i} - Q_x\| \geq eig_{min}(Q_x) - M/\chi^2_{1,T}\right]$. Now, let $0 < \delta \leq eig_{min}(Q_x)/2$ and T large such that $eig_{min}(Q_x) - M/\chi^2_{1,T} > \delta$. Then, by using $\mathbb{P}\left[\|\hat{Q}_{x,i} - Q_x\| \geq \delta\right] \leq \mathbb{P}\left[\left|\frac{1}{T}\sum_t I_{i,t}(x_t x_t - Q_x)\right| \geq \sqrt{\delta}\right] + \mathbb{P}\left[\tau_{i,T} \geq \sqrt{\delta}\right]$ we get $P_{2,nT} \leq P\left[\left|\frac{1}{T}\sum_t I_{i,t}(x_t x_t - Q_x)\right| \geq \sqrt{\delta}\right] + O(T^{-\bar{b}})$. The first term in the RHS is $O(T^{-\bar{b}})$ by using $\mathbb{P}\left[\left|\frac{1}{T}\sum_t I_{i,t}(x_t x_t - Q_x)\right| \geq \sqrt{\delta}\right] \leq \sup_{\gamma\in[0,1]} \mathbb{P}\left[\left|\frac{1}{T}\sum_t I_t(\gamma)(x_t x_t - Q_x)\right| \geq \sqrt{\delta}\right]$ and Assumption C.1 b). Then, $P_{2,nT} = O(T^{-\bar{b}})$, for any $\bar{b} > 0$.

A.5.4 Proof of Lemma 5

Let $W_T(\gamma) := \frac{1}{T}\sum_t (I_t(\gamma) - E[I_t(\gamma)])$ and $r_T := T^{-a}$ for $0 < a < \eta/2$. Since $|W_T(\gamma)| \leq 1$ for all $\gamma \in [0,1]$, we have:

$$\begin{aligned}
\sup_{\gamma\in[0,1]} E[|W_T(\gamma)|^4] &\leq \sup_{\gamma\in[0,1]} E[|W_T(\gamma)|] = \sup_{\gamma\in[0,1]} \int_0^1 \mathbb{P}[|W_T(\gamma)| \geq \delta] d\delta \leq r_T + \sup_{\gamma\in[0,1]} \int_{r_T}^1 \mathbb{P}[|W_T(\gamma)| \geq \delta] d\delta \\
&\leq r_T + C_1 T \int_{r_T}^1 \exp\left\{-C_2 \delta^2 T^{\eta}\right\} d\delta + C_3 \exp\left\{-C_4 T^{\bar{\eta}}\right\} \int_{r_T}^1 \frac{1}{\delta} d\delta \\
&\leq r_T + C_1 T \exp\left\{-C_2 r_T^2 T^{\eta}\right\} + C_3 \exp\left\{-C_4 T^{\bar{\eta}}\right\} \log(1/r_T) = o(1),
\end{aligned}$$

from Assumption C.1 d).

A.5.5 Proof of Lemma 6

By definition of $\tilde{S}_{ij}$, we have

$$\begin{aligned}
\frac{1}{n}\sum_{i,j}\left\|\tilde{S}_{ij}-S_{ij}\right\| &= \frac{1}{n}\sum_{i,j}\left\|\hat{S}_{ij}\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa\right\}}-S_{ij}\right\| \\
&\leq \frac{1}{n}\sum_{i,j}\left\|S_{ij}\mathbf{1}_{\{\|S_{ij}\|\geq\kappa\}}-S_{ij}\right\|+\frac{1}{n}\sum_{i,j}\left\|\hat{S}_{ij}\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa\right\}}-S_{ij}\mathbf{1}_{\{\|S_{ij}\|\geq\kappa\}}\right\| \\
&=: I_{31}+I_{32}.
\end{aligned}$$

By Assumption A.4,

$$I_{31}=\frac{1}{n}\sum_{i,j}\|S_{ij}\|\,\mathbf{1}_{\{\|S_{ij}\|<\kappa\}}\leq\max_{i}\sum_{j}\|S_{ij}\|^{q}\,\kappa^{1-q}\leq\kappa^{1-q}c_{0}\left(n\right)=O_{p}\left(\kappa^{1-q}n^{\delta}\right), \tag{38}$$

where $c_0(n) := \max_i \sum_j \|S_{ij}\|^q = O_p(n^\delta)$.

Let us now consider I_{32}:

$$\begin{aligned}
I_{32} &= \frac{1}{n}\sum_{i,j}\left\|\hat{S}_{ij}\right\|\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa,\|S_{ij}\|<\kappa\right\}}+\frac{1}{n}\sum_{i,j}\|S_{ij}\|\,\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|<\kappa,\|S_{ij}\|\geq\kappa\right\}} \\
&\quad+\frac{1}{n}\sum_{i,j}\left\|\hat{S}_{ij}-S_{ij}\right\|\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa,\|S_{ij}\|\geq\kappa\right\}} \\
&\leq \max_{i}\sum_{j}\left\|\hat{S}_{ij}\right\|\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa,\|S_{ij}\|<\kappa\right\}}+\max_{i}\sum_{j}\|S_{ij}\|\,\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|<\kappa,\|S_{ij}\|\geq\kappa\right\}} \\
&\quad+\max_{i}\sum_{j}\left\|\hat{S}_{ij}-S_{ij}\right\|\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa,\|S_{ij}\|\geq\kappa\right\}} =: I_{33}+I_{34}+I_{35}.
\end{aligned}$$

From Assumption A.4, we have:

$$I_{35}\leq\max_{i,j}\left\|\hat{S}_{ij}-S_{ij}\right\|\max_{i}\sum_{j}\|S_{ij}\|^{q}\,\kappa^{-q}=O_{p}\left(\psi_{nT}c_{0}\left(n\right)\kappa^{-q}\right). \tag{39}$$

Let us study I_{33}:

$$I_{33} \leq \max_{i}\sum_{j}\left\|\hat{S}_{ij}-S_{ij}\right\|\mathbf{1}_{\left\{\left\|\hat{S}_{ij}\right\|\geq\kappa,\|S_{ij}\|<\kappa\right\}}+\max_{i}\sum_{j}\|S_{ij}\|\,\mathbf{1}_{\{\|S_{ij}\|<\kappa\}} =: I_{36}+I_{37}.$$

By Assumption A.4,

$$I_{37}\leq\kappa^{1-q}c_{0}\left(n\right). \tag{40}$$

Now take $v \in (0,1)$. Let $N_i(\epsilon) := \sum_j \mathbf{1}_{\{\|\hat{S}_{ij}-S_{ij}\|>\epsilon\}}$, for $\epsilon > 0$, then

$$\begin{aligned} I_{36} &= \max_i \sum_j \left\|\hat{S}_{ij} - S_{ij}\right\| \mathbf{1}_{\{\|\hat{S}_{ij}\|\geq\kappa, \|S_{ij}\|\leq v\kappa\}} + \max_i \sum_j \left\|\hat{S}_{ij} - S_{ij}\right\| \mathbf{1}_{\{\|\hat{S}_{ij}\|\geq\kappa, v\kappa<\|S_{ij}\|<\kappa\}} \\ &\leq \max_{i,j} \left\|\hat{S}_{ij} - S_{ij}\right\| \max_i N_i\left((1-v)\kappa\right) + \max_{i,j}\left\|\hat{S}_{ij} - S_{ij}\right\| c_0(n)(v\kappa)^{-q}. \end{aligned}$$

Moreover, by the Chebyschev inequality, for any positive sequence R_{nT} we have:

$$\mathbb{P}\left[\max_i N_i(\epsilon) \geq R_{nT}\right] \leq n\mathbb{P}\left[N_i(\epsilon) \geq R_{nT}\right] \leq \frac{n}{R_{nT}} E[N_i(\epsilon)] \leq \frac{n^2}{R_{nT}} \max_{i,j} \mathbb{P}\left[\left\|\hat{S}_{ij} - S_{ij}\right\| \geq \epsilon\right],$$

which implies $\max_i N_i(\epsilon) = O_p\left(n^2 \max_{i,j} \mathbb{P}\left[\left\|\hat{S}_{ij} - S_{ij}\right\| \geq \epsilon\right]\right)$. Thus,

$$I_{36} = O_p\left(\psi_{nT} n^2 \Psi_{nT}\left((1-v)\kappa\right) + \psi_{nT} c_0(n)(v\kappa)^{-q}\right). \tag{41}$$

Finally, we consider I_{34}. We have

$$\begin{aligned} I_{34} &\leq \max_i \sum_j \left(\left\|\hat{S}_{ij} - S_{ij}\right\| + \left\|\hat{S}_{ij}\right\|\right) \mathbf{1}_{\{\|\hat{S}_{ij}\|<\kappa, \|S_{ij}\|\geq\kappa\}} \\ &\leq \max_{i,j}\left\|\hat{S}_{ij} - S_{ij}\right\| \max_i \sum_j \mathbf{1}_{\{\|S_{ij}\|\geq\kappa\}} + \kappa \max_i \sum_j \mathbf{1}_{\{\|S_{ij}\|\geq\kappa\}} \\ &= O_p\left(\psi_{nT} c_0(n) \kappa^{-q} + c_0(n)\kappa^{1-q}\right). \end{aligned} \tag{42}$$

Combining (38)-(42) the result follows.

A.5.6 Proof of Lemma 7

By using $\hat{\varepsilon}_{i,t} = \varepsilon_{i,t} - x_t'\left(\hat{\beta}_i - \beta_i\right)$ and $\hat{S}_{ij}^0 = \frac{1}{T_{ij}}\sum_t I_{ij,t}\varepsilon_{i,t}\varepsilon_{j,t}x_t x_t'$, we have:

$$\begin{aligned} \hat{S}_{ij} &= \hat{S}_{ij}^0 - \frac{1}{T_{ij}}\sum_t I_{ij,t}\varepsilon_{i,t}x_t'\left(\hat{\beta}_j - \beta_j\right)x_t x_t' - \frac{1}{T_{ij}}\sum_t I_{ij,t}\varepsilon_{j,t}x_t'\left(\hat{\beta}_i - \beta_i\right)x_t x_t' \\ &\quad + \frac{1}{T_{ij}}\sum_t I_{ij,t}\left(\hat{\beta}_i - \beta_i\right)' x_t x_t'\left(\hat{\beta}_j - \beta_j\right)x_t x_t' \\ &=: \hat{S}_{ij}^0 - A_{ij} - B_{ij} + C_{ij}, \end{aligned}$$

where $A_{ij} = B_{ji}$. Then, for any i, j, we have $\left\| \hat{S}_{ij} - S_{ij} \right\| \leq \left\| \hat{S}^0_{ij} - S_{ij} \right\| + \|A_{ij}\| + \|B_{ij}\| + \|C_{ij}\|$. We get for any $\xi \geq 0$:

$$\begin{aligned} \Psi_{nT}(\xi) &\leq \max_{i,j} \mathbb{P}\left[\left\|\hat{S}^0_{ij} - S_{ij}\right\| \geq \frac{\xi}{4}\right] + \max_{i,j} \mathbb{P}\left[\|A_{ij}\| \geq \frac{\xi}{4}\right] + \max_{i,j} \mathbb{P}\left[\|B_{ij}\| \geq \frac{\xi}{4}\right] \\ &\quad + \max_{i,j} \mathbb{P}\left[\|C_{ij}\| \geq \frac{\xi}{4}\right] = \Psi^0_{nT}(\xi/4) + 2P_{1,nT}(\xi/4) + P_{2,nT}(\xi/4), \end{aligned} \tag{43}$$

where $\Psi^0_{nT}(\xi/4) := \max_{i,j} \mathbb{P}\left[\left\|\hat{S}^0_{ij} - S_{ij}\right\| \geq \frac{\xi}{4}\right]$, $P_{1,nT}(\xi/4) := \max_{i,j} \mathbb{P}\left[\|A_{ij}\| \geq \frac{\xi}{4}\right]$, and $P_{2,nT}(\xi/4) := \max_{i,j} \mathbb{P}\left[\|C_{ij}\| \geq \frac{\xi}{4}\right]$. Let us bound the three terms in the RHS of Inequality (43).

a) Bound of $\Psi^0_{nT}(\xi/4)$. We use that $\hat{S}^0_{ij} - S_{ij} = \frac{1}{T_{ij}} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - S_{ij}\right)$ $= \tau_{ij,T} \frac{1}{T} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - E\left[\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' | \gamma_i \gamma_j\right]\right)$ and $\tau_{ij} \leq M$. Then:

$$\begin{aligned} \|\hat{S}^0_{ij} - S_{ij}\| &\leq M \left\| \frac{1}{T} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - E\left[\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' | \gamma_i \gamma_j\right]\right) \right\| \\ &\quad + |\tau_{ij,T} - \tau_{ij}| \left\| \frac{1}{T} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - E\left[\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' | \gamma_i \gamma_j\right]\right) \right\|. \end{aligned}$$

We deduce:

$$\begin{aligned} &\Psi^0_{nT}(\xi/4) \\ \leq\ &\max_{i,j} \mathbb{P}\left[\left\| \frac{1}{T} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - E\left[\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' | \gamma_i \gamma_j\right]\right) \right\| \geq \frac{\xi}{8M}\right] + \max_{i,j} \mathbb{P}\left[|\tau_{ij,T} - \tau_{ij}| \geq \sqrt{\frac{\xi}{8}}\right] \\ &+ \max_{i,j} \mathbb{P}\left[\left\| \frac{1}{T} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - E\left[\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' | \gamma_i \gamma_j\right]\right) \right\| \geq \sqrt{\frac{\xi}{8}}\right] \\ \leq\ &2 \max_{i,j} \mathbb{P}\left[\left\| \frac{1}{T} \sum_t I_{ij,t} \left(\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' - E\left[\varepsilon_{i,t}\varepsilon_{j,t} x_t x_t' | \gamma_i \gamma_j\right]\right) \right\| \geq \frac{\xi}{8M}\right] + \max_{i,j} \mathbb{P}\left[|\tau_{ij,T} - \tau_{ij}| \geq \sqrt{\frac{\xi}{8}}\right] \\ =:\ &2P_{3,nT} + P_{4,nT}, \end{aligned}$$

for small ξ. We use $P_{3,nT} \leq \sup_{\gamma,\gamma' \in [0,1]} \mathbb{P}\left[\left\| \frac{1}{T} \sum_t I_t(\gamma) I_t(\gamma') \left(\varepsilon_t(\gamma)\varepsilon_t(\gamma') x_t x_t' - E\left[\varepsilon_t(\gamma)\varepsilon_t(\gamma') x_t x_t'\right]\right) \right\| \geq \frac{\xi}{8M}\right]$ and Assumption C.1 e) to get $P_{3,nT} \leq C_1 T \exp\left\{-C_2^* \xi^2 T^\eta\right\} + C_3^* \xi^{-1} \exp\left\{-C_4 T^{\bar{\eta}}\right\}$, for some constants $C_1, C_2^*, C_3^*, C_4 > 0$. To bound $P_{4,nT}$, we use $\tau_{ij} \leq M$ and $|\tau_{ij,T} - \tau_{ij}| \leq \tau_{ij} \tau_{ij,T} |\tau_{ij,T}^{-1} - \tau_{ij}^{-1}| \leq$

$\tau_{ij}\dfrac{|\tau_{ij,T}^{-1}-\tau_{ij}^{-1}|}{\tau_{ij}^{-1}-|\tau_{ij,T}^{-1}-\tau_{ij}^{-1}|}\leq 2M^2|\tau_{ij,T}^{-1}-\tau_{ij}^{-1}|$, if $|\tau_{ij,T}^{-1}-\tau_{ij}^{-1}|\leq M^{-1}/2$. Thus, we have $P_{4,nT}\leq 2\max_{i,j}\mathbb{P}\left[|\tau_{ij,T}^{-1}-\tau_{ij}^{-1}|\geq \frac{1}{2M^2}\sqrt{\frac{\xi}{8}}\right]$, for small ξ. By using $\tau_{ij,T}^{-1}=\frac{1}{T}\sum_t I_{ij,t}$ and $\tau_{ij}^{-1}=E[I_{ij,t}|\gamma_i,\gamma_j]$, from Assumption C.1 d) we get:

$$\begin{aligned}\max_{i,j}\mathbb{P}\left[|\tau_{ij,T}^{-1}-\tau_{ij}^{-1}|\geq \frac{1}{2M^2}\sqrt{\frac{\xi}{8}}\right] &\leq \sup_{\gamma,\gamma'\in[0,1]}\mathbb{P}\left[\left|\frac{1}{T}\sum_t(I_t(\gamma)I_t(\gamma')-E[I_t(\gamma)I_t(\gamma')])\right|\geq \frac{1}{2M^2}\sqrt{\frac{\xi}{8}}\right]\\ &\leq C_1T\exp\left\{-C_2^*\xi T^{\eta}\right\}+C_3^*\xi^{-1/2}\exp\left\{-C_4T^{\bar{\eta}}\right\}.\end{aligned}$$

We deduce:

$$\Psi_{nT}^0\left(\xi/4\right)\leq C_1^*T\exp\left\{-C_2^*\xi^2T^{\eta}\right\}+C_3^*\xi^{-1}\exp\left\{-C_4T^{\bar{\eta}}\right\}. \tag{44}$$

b) Bound of $P_{1,nT}\left(\xi/4\right)$. For some constant C, we have

$$\|A_{ij}\|\leq C\tau_{ij,T}\max_{k,l,m}\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{t,k}x_{t,l}x_{t,m}\right|\left\|\hat{\beta}_j-\beta_j\right\|.$$

Let $\chi_{3,T}=(\log T)^a$, for $a>0$. From a similar argument as in the proof of Lemma 4, and Assumption C.1 d), we have $\max_{i,j}\mathbb{P}\left[\tau_{ij,T}\geq\chi_{3,T}\right]=O(T^{-\bar{b}})$, for any $\bar{b}>0$. Thus,

$$\begin{aligned}&P_{1,nT}\left(\xi/4\right)\\ \leq\;&\max_{i,j}\mathbb{P}\left[\tau_{ij,T}\max_{k,l,m}\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{t,k}x_{t,l}x_{t,m}\right|\left\|\hat{\beta}_j-\beta_j\right\|\geq\frac{\xi}{4C}\right]\\ \leq\;&\max_{i,j}\mathbb{P}\left[\tau_{ij,T}\geq\chi_{3,T}\right]+\max_{i,j}\mathbb{P}\left[\max_{k,l,m}\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{t,k}x_{t,l}x_{t,m}\right|\geq\sqrt{\frac{\xi}{4\chi_{3,T}C}}\text{ and }\tau_{ij,T}\leq\chi_{3,T}\right]\\ &+\max_{i,j}\mathbb{P}\left[\left\|\hat{\beta}_j-\beta_j\right\|\geq\sqrt{\frac{\xi}{4\chi_{3,T}C}}\text{ and }\tau_{ij,T}\leq\chi_{3,T}\right]\\ \leq\;&(K+1)^3\max_{i,j}\max_{k,l,m}\mathbb{P}\left[\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{t,k}x_{t,l}x_{t,m}\right|\geq\sqrt{\frac{\xi}{4\chi_{3,T}C}}\right]\\ &+\mathbb{P}\left[\left\|\hat{\beta}_j-\beta_j\right\|\geq\sqrt{\frac{\xi}{4\chi_{3,T}C}}\text{ and }\tau_{j,T}\leq\chi_{3,T}\right]+O(T^{-\bar{b}}).\end{aligned} \tag{45}$$

By Assumption C.1 f),

$$\begin{aligned}\max_{i,j}\max_{k,l,m}\mathbb{P}\left[\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{t,k}x_{t,l}x_{t,m}\right|\geq\sqrt{\frac{\xi}{4\chi_{3,T}C}}\right] &\leq C_1T\exp\left\{-\frac{C_2^*\xi}{\chi_{3,T}}T^{\eta}\right\}\\ &+C_3^*\sqrt{\frac{\chi_{3,T}}{\xi}}\exp\left\{-C_4T^{\bar{\eta}}\right\}.\end{aligned} \tag{46}$$

Let us now focus on $\mathbb{P}\left[\left\|\hat{\beta}_j - \beta_j\right\| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}} \text{ and } \tau_{j,T} \leq \chi_{3,T}\right]$. By using

$$\left\|\hat{\beta}_j - \beta_j\right\| \leq \chi_{3,T}\left\|Q_x^{-1}\right\|\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| + \chi_{3,T}\left\|\hat{Q}_{x,j}^{-1} - Q_x^{-1}\right\|\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\|$$

when $\tau_{j,T} \leq \chi_{3,T}$, we get

$$\begin{aligned}
& \mathbb{P}\left[\left\|\hat{\beta}_j - \beta_j\right\| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}} \text{ and } \tau_{j,T} \leq \chi_{3,T}\right] \\
\leq\ & \mathbb{P}\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| \geq \frac{1}{2}\sqrt{\frac{\xi}{4\chi_{3,T}C}}\chi_{3,T}^{-1}\left\|Q_x^{-1}\right\|^{-1}\right] \\
& + \mathbb{P}\left[\left\|\hat{Q}_{x,j}^{-1} - Q_x^{-1}\right\|\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| \geq \frac{1}{2}\sqrt{\frac{\xi}{4\chi_{3,T}C}}\chi_{3,T}^{-1}\right] \\
\leq\ & \mathbb{P}\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| \geq \sqrt{\frac{\xi}{16\chi_{3,T}^3 C}}\left\|Q_x^{-1}\right\|^{-1}\right] \\
& + \mathbb{P}\left[\left\|\hat{Q}_{x,j}^{-1} - Q_x^{-1}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\right] + \mathbb{P}\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\right] \\
\leq\ & 2\mathbb{P}\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| \geq \sqrt{\frac{\xi}{16\chi_{3,T}^3 C}}\left\|Q_x^{-1}\right\|^{-1}\right] + \mathbb{P}\left[\left\|\hat{Q}_{x,j}^{-1} - Q_x^{-1}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\right], \quad (47)
\end{aligned}$$

for small ξ. From Assumption C.1c), the first probability in the RHS of Inequality (47) is such that:

$$\mathbb{P}\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_t\varepsilon_{j,t}\right\| \geq \sqrt{\frac{\xi}{16\chi_{3,T}^3 C}}\left\|Q_x^{-1}\right\|^{-1}\right] \leq C_1 T \exp\left\{-\frac{C_2^*\xi}{\chi_{3,T}^3}T^{\eta}\right\} + C_3^*\sqrt{\frac{\chi_{3,T}^3}{\xi}}\exp\left\{-C_4 T^{\bar{\eta}}\right\}. \quad (48)$$

To bound the second probability in the RHS of Inequality (47) we use the next Lemma.

Lemma 12 *For any two non-singular matrices A and B such that $\|A - B\| < \frac{1}{2}\|A^{-1}\|^{-1}$ we have:*

$$\|B^{-1} - A^{-1}\| \leq 2\|A^{-1}\|^2\|A - B\|.$$

From Lemma 12, we get:

$$\begin{aligned}
\mathbb{P}\left[\left\|\hat{Q}_{x,j}^{-1}-Q_x^{-1}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\right] &\leq \mathbb{P}\left[\left\|\hat{Q}_{x,j}-Q_x\right\| \geq \frac{1}{2}\left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\|Q_x^{-1}\|^{-2}\right] \\
&\quad +\mathbb{P}\left[\left\|\hat{Q}_{x,j}-Q_x\right\| \geq \frac{1}{2}\|Q_x^{-1}\|^{-1}\right] \\
&\leq 2\mathbb{P}\left[\left\|\hat{Q}_{x,j}-Q_x\right\| \geq \frac{1}{2}\left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\|Q_x^{-1}\|^{-2}\right],
\end{aligned}$$

for small $\xi > 0$. From Assumption C.1b),

$$\begin{aligned}
\mathbb{P}\left[\left\|\hat{Q}_{x,j}-Q_x\right\| \geq \frac{1}{2}\left(\frac{\xi}{16\chi_{3,T}^3 C}\right)^{1/4}\|Q_x^{-1}\|^{-2}\right] &\leq C_1 T \exp\left\{-C_2^*\sqrt{\frac{\xi}{\chi_{3,T}^3}}T^{\eta}\right\} \\
&\quad +2C_3^*\left(\frac{\chi_{3,T}^3}{\xi}\right)^{1/4}\exp\left\{-C_4 T^{\bar{\eta}}\right\}. \qquad (49)
\end{aligned}$$

Then, from (45)-(49) we get:

$$P_{1,nT}(\xi/4) \leq C_1^* T \exp\left\{-C_2^*\xi T^{\eta}/\chi_{3,T}^3\right\} + \frac{C_3^*\chi_{3,T}^{3/2}}{\sqrt{\xi}}\exp\left\{-C_4 T^{\bar{\eta}}\right\} + O(T^{-\bar{b}}), \qquad (50)$$

for small $\xi > 0$ and some constants $C_1^*, C_2^*, C_3^*, C_4 > 0$.

c) Bound of $P_{2,nT}(\xi/4)$. We have from Assumption C.4

$$\begin{aligned}
\|C_{ij}\| &\leq \left\|\hat{\beta}_i-\beta_i\right\|\left\|\hat{\beta}_j-\beta_j\right\|\sup_{k,l,m,p}\left|\frac{1}{T_{ij}}\sum_t I_{ij,t}x_{t,k}x_{t,l}x_{t,m}x_{t,p}\right| \\
&\leq C\left\|\hat{\beta}_i-\beta_i\right\|\left\|\hat{\beta}_j-\beta_j\right\|.
\end{aligned}$$

Thus, we have:

$$P_{2,nT}(\xi/4) \leq \max_{i,j}\mathbb{P}\left[C\left\|\hat{\beta}_i-\beta_i\right\|\left\|\hat{\beta}_j-\beta_j\right\| \geq \frac{\xi}{4}\right] \leq 2\mathbb{P}\left[\left\|\hat{\beta}_i-\beta_i\right\| \geq \left(\frac{\xi}{4C}\right)^{1/2}\right].$$

By the same arguments as above, we get:

$$P_{2,nT}(\xi/4) \leq C_1^* T \exp\left\{-C_2^*\xi T^{\eta}/\chi_{3,T}^3\right\} + \frac{C_3^*\chi_{3,T}^{3/2}}{\sqrt{\xi}}\exp\left\{-C_4 T^{\bar{\eta}}\right\}, \qquad (51)$$

for small $\xi > 0$ and some constants $C_1^*, C_2^*, C_3^*, C_4 > 0$.

d) Conclusion. From inequalities (43), (44), (50) and (51) we deduce:

$$\Psi_{nT}(\xi) \leq C_1^* T \exp\left\{-C_2^* \xi_T^2 T^\eta\right\} + \frac{C_3^*}{\xi_T} \exp\left\{-C_4 T^{\bar{\eta}}\right\} + O(T^{-\bar{b}}),$$

where $\xi_T := \min\{\xi, \sqrt{\xi/\chi_{3,T}^3}\}$, for small $\xi > 0$ and constants $C_1^*, C_2^*, C_3^*, C_4 > 0$. For $\xi = (1-v)\kappa$ and $\kappa = M\sqrt{\frac{\log n}{T^\eta}}$, we get $\xi_T = (1-\nu)\kappa$ for large T and

$$\begin{aligned} n^2 \Psi_{nT}((1-v)\kappa) &\leq C_1^* n^2 T \exp\left\{-C_2^* M^2 (1-v)^2 \log n\right\} + \frac{n^2 C_3^*}{(1-v)M}\sqrt{\frac{T^\eta}{\log n}} \exp\left\{-C_4^* T^{\bar{\eta}}\right\} \\ &\quad + O(n^2 T^{-\bar{b}}) = O(1), \end{aligned}$$

for $\bar{b}$ and M sufficiently large, when $n, T \to \infty$ such that $n = O(T^{\bar{\gamma}})$ for $\bar{\gamma} > 0$.

Finally, let us prove that $\psi_{nT} = O_p\left(\sqrt{\frac{\log n}{T^\eta}}\right)$. Let $\epsilon > 0$. Then,

$$\begin{aligned} \mathbb{P}\left[\psi_{nT} \geq \sqrt{\frac{\log n}{T^\eta}}\epsilon\right] &\leq n^2 \max_{i,j} \mathbb{P}\left[\left\|\hat{S}_{ij} - S_{ij}\right\| \geq \sqrt{\frac{\log n}{T^\eta}}\epsilon\right] \\ &= n^2 \Psi_{nT}\left(\sqrt{\frac{\log n}{T^\eta}}\epsilon\right) \leq n^2 \Psi_{nT}((1-v)\kappa) = O(1), \end{aligned}$$

for large ϵ. The conclusion follows.

A.5.7 Proof of Lemma 8

Under the null hypothesis $\mathcal{H}_0$, and by definition of the fitted residual $\hat{e}_i$, we have

$$\begin{aligned} \hat{e}_i &= a_i - b_i'\hat{\nu} + \hat{c}_\nu'\left(\hat{\beta}_i - \beta_i\right) \\ &= a_i - b_i'\nu + \hat{c}_\nu'\left(\hat{\beta}_i - \beta_i\right) - b_i'(\hat{\nu} - \nu) \\ &= \hat{c}_\nu'\left(\hat{\beta}_i - \beta_i\right) - b_i'(\hat{\nu} - \nu). \end{aligned} \tag{52}$$

By definition of $\hat{Q}_e$, it follows

$$\begin{aligned} \hat{Q}_e &= \frac{1}{n}\sum_i \hat{w}_i\left[\hat{c}_\nu'\left(\hat{\beta}_i - \beta_i\right)\right]^2 - 2(\hat{\nu} - \nu)'\frac{1}{n}\sum_i \hat{w}_i b_i\left(\hat{\beta}_i - \beta_i\right)'\hat{c}_\nu + (\hat{\nu} - \nu)'\frac{1}{n}\sum_i \hat{w}_i b_i b_i'(\hat{\nu} - \nu) \\ &=: \frac{1}{n}\sum_i \hat{w}_i\left[\hat{c}_\nu'\left(\hat{\beta}_i - \beta_i\right)\right]^2 - 2I_{71} + I_{72}. \end{aligned}$$

Let us study the second term in the RHS:

$$I_{71} = \frac{1}{\sqrt{nT}} (\hat{\nu} - \nu)' \frac{1}{\sqrt{n}} \sum_i \hat{w}_i \tau_{i,T} b_i Y_{i,T}' \hat{Q}_{x,i}^{-1} \hat{c}_\nu =: \frac{1}{\sqrt{nT}} (\hat{\nu} - \nu)' I_{711} \hat{c}_\nu,$$

where $I_{711} = O_p(1)$ by the same arguments used to control term I_1 in the proof of Proposition 3. We have $\hat{\nu} - \nu = O_{p,log}\left(\frac{1}{\sqrt{nT}} + \frac{1}{T}\right)$ and $\hat{c}_\nu = O_p(1)$ by Lemma 3 (v). Thus, $I_{71} = O_{p,log}\left(\frac{1}{nT} + \frac{1}{T\sqrt{nT}}\right)$.

Let us now consider I_{72}. From Lemma 1 (ii)-(iii) and Lemma 3 (v), we have $I_{72} = O_{p,log}\left(\frac{1}{nT} + \frac{1}{T^2}\right)$. The conclusion follows.

A.5.8 Proof of Lemma 9

Under $\mathcal{H}_1$, and using Equation (52), we have $\hat{e}_i = e_i + \hat{c}_\nu' \left(\hat{\beta}_i - \beta_i\right) - b_i' (\hat{\nu} - \nu)$. By definition of $\hat{Q}_e$, it follows:

$$\begin{aligned} \hat{Q}_e &= \frac{1}{n} \sum_i \hat{w}_i e_i^2 + 2 \frac{1}{n} \sum_i \hat{w}_i \hat{c}_\nu' \left(\hat{\beta}_i - \beta_i\right) e_i - 2 (\hat{\nu} - \nu)' \frac{1}{n} \sum_i \hat{w}_i b_i e_i \\ &\quad + \frac{1}{n} \sum_i \hat{w}_i \left[\hat{c}_\nu' \left(\hat{\beta}_i - \beta_i\right)\right]^2 - 2 (\hat{\nu} - \nu)' \frac{1}{n} \sum_i \hat{w}_i b_i \left(\hat{\beta}_i - \beta_i\right)' \hat{c}_\nu + (\hat{\nu} - \nu)' \frac{1}{n} \sum_i \hat{w}_i b_i b_i' (\hat{\nu} - \nu) \\ &=: I_{81} + I_{82} + I_{83} + I_{84} + I_{85} + I_{86}. \end{aligned} \tag{53}$$

From Equations (24) and (26) and similar arguments as in Section A.2.3 c), we have $I_{81} = \frac{1}{n} \sum_i w_i e_i^2 + O_{p,log}\left(\frac{1}{\sqrt{T}}\right)$. By similar arguments as for term I_1 in the proof of Proposition 3, we have $I_{82} = \frac{1}{\sqrt{nT}} \left(\frac{1}{\sqrt{n}} \sum_i \hat{w}_i \tau_{i,T} e_i Y_{i,T}' \hat{Q}_{x,i}^{-1}\right) \hat{c}_\nu = O_p\left(\frac{1}{\sqrt{nT}}\right)$. By using $\frac{1}{n} \sum_i \hat{w}_i b_i e_i = \frac{1}{n} \sum_i w_i b_i e_i + O_{p,log}\left(\frac{1}{\sqrt{T}}\right) = O_p\left(\frac{1}{\sqrt{n}}\right) + O_{p,log}\left(\frac{1}{\sqrt{T}}\right)$ and $\hat{\nu} - \nu_\infty = O_{p,log}\left(\frac{1}{\sqrt{n}} + \frac{1}{T}\right)$, we get $I_{83} = O_{p,log}\left(\frac{1}{n} + \frac{1}{\sqrt{nT}} + \frac{1}{\sqrt{T^3}}\right)$. Similar as for I_{82} we have $I_{85} = O_{p,log}\left(\frac{1}{n\sqrt{T}} + \frac{1}{\sqrt{nT^3}}\right)$. From $\hat{\nu} - \nu_\infty = O_{p,log}\left(\frac{1}{\sqrt{n}} + \frac{1}{T}\right)$, we have $I_{86} = O_{p,log}\left(\frac{1}{n} + \frac{1}{T^2}\right)$. The conclusion follows.

A.5.9 Proof of Lemma 10

By applying MN Theorem 2 p.35, Theorem 10 p. 55 and using $W_{n,1} = I_n$, we have

$$\begin{aligned}
Ab = vec\,(Ab) &= \left(b' \otimes A\right) vec\,(I_n) \\
&= vec\left[\left(b' \otimes A\right) vec\,(I_n)\right] \\
&= \left(vec\,(I_n)' \otimes I_m\right) vec\left(b' \otimes A\right) \\
&= \left(vec\,(I_n)' \otimes I_m\right)\left(I_n \otimes W_{n,1} \otimes I_m\right)\left(vec\left(b'\right) \otimes vec\,(A)\right) \\
&= \left(vec\,(I_n)' \otimes I_m\right)\left(I_n \otimes I_m\right) vec\left(vec\,(A)\,b'\right) \\
&= \left(vec\,(I_n)' \otimes I_m\right) vec\left(vec\,(A)\,b'\right).
\end{aligned}$$

A.5.10 Proof of Lemma 11

A.5.10.1 Assumption APR.4 (i)

We use that $eig_{\max}(A) \leq \max_{i=1,\dots,n} \sum_{j=1} |a_{i,j}|$ for any matrix $A = [a_{ij}]_{i,j=1,\dots,n}$. Then, for any sequence (γ_i) in $[0,1]$ we have:

$$eig_{\max}(\Sigma_{\varepsilon,1,n}) \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |Cov[\varepsilon_t(\gamma_i), \varepsilon_t(\gamma_j)]| \leq C \max_{m=1,\dots,J_n} \sum_{j=1}^{n} 1\{\gamma_j \in I_m\} \tag{54}$$

where $C := \sup_{\gamma \in [0,1]} E[\varepsilon_t(\gamma)^2]$. Define:

$$\mathcal{J} = \left\{ (\gamma_i) : \max_{m=1,\dots,J_n} \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} = o(1) \right\}.$$

Then Assumption APR.4 (i) holds if $\mu_\Gamma(\mathcal{J}) = 1$. From Theorem 2.1.1 in Stout (1974), it is enough to show that $\sum_{n=1}^{\infty} \mu_\Gamma \left(\max_{m=1,\dots,J_n} \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} > \varepsilon \right) < \infty$, for any $\varepsilon > 0$. Now, since $\max_{m=1,\dots,J_n} B_m = o(1)$, we have $\mu_\Gamma \left(\max_{m=1,\dots,J_n} \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} > \varepsilon \right) \leq \mu_\Gamma \left(\max_{m=1,\dots,J_n} \left| \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} - B_m \right| > \varepsilon/2 \right)$, for large n. Thus, we get:

$$\mu_\Gamma \left(\max_{m=1,\dots,J_n} \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} > \varepsilon \right) \leq J_n \max_{m=1,\dots,J_n} \mu_\Gamma \left(\left| \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} - B_m \right| > \varepsilon/2 \right),$$

for large n. To bound the probability in the RHS, we use $|1\{\gamma_i \in I_m\} - B_m| \leq 1$ and the Hoeffding's inequality (see Bosq (1998), Theorem 1.2) to get:

$$\mu_\Gamma \left(\left| \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} - B_m \right| > \varepsilon/2 \right) \leq 2 \exp\left(-n\varepsilon^2/8\right).$$

Then, since $J_n \leq n$, we get:

$$\sum_{n=1}^{\infty} \mu_\Gamma \left(\max_{m=1,\ldots,J_n} \frac{1}{n} \sum_{i=1}^{n} 1\{\gamma_i \in I_m\} > \varepsilon \right) \leq 2 \sum_{n=1}^{\infty} n \exp\left(-n\varepsilon^2/8\right) < \infty,$$

and the conclusion follows.

A.5.10.2 Assumption A.1

Conditions a) and b) are clearly satisfied under BD.1, BD.3 and BD.4. Let us now consider condition c). We have $\sigma_{ij,t} = E[\varepsilon_t(\gamma_i)\varepsilon_t(\gamma_j)|\gamma_i, \gamma_j] =: \sigma_{ij}$ independent of t. Thus, $E[\sigma_{ij,t}^2|\gamma_i, \gamma_j]^{1/2} = \sigma_{ij}$. By BD.1, BD.4 and the Cauchy-Schwarz inequality $\sigma_{ij} = \sum_{m=1}^{J_n} 1\{\gamma_i, \gamma_j \in I_m\} E[\varepsilon_t(\gamma_i)\varepsilon_t(\gamma_j)|\gamma_i, \gamma_j] \leq C \sum_{m=1}^{J_n} 1\{\gamma_i, \gamma_j \in I_m\}$, where $C = \sup_{\gamma \in [0,1]} E[\varepsilon_t(\gamma)^2]$. Hence, we get:

$$\begin{aligned} E\left[\frac{1}{n}\sum_{i,j} E[\sigma_{ij,t}^2|\gamma_i,\gamma_j]^{1/2}\right] &\leq C\frac{1}{n}\sum_{i}\sum_{m=1}^{J_n} E[1\{\gamma_i \in I_m\}] + C\frac{1}{n}\sum_{i\neq j}\sum_{m=1}^{J_n} E[1\{\gamma_i,\gamma_j \in I_m\}] \\ &= C\sum_{m=1}^{J_n} B_m + C(n-1)\sum_{m=1}^{J_n} B_m^2 = O\left(1 + n\sum_{m=1}^{J_n} B_m^2\right). \end{aligned}$$

From BD.2, the RHS is $O(1)$, and condition c) in Assumption A.1 follows.

A.5.10.3 Assumption A.2

Let us consider condition a). Under BD.1 and BD.3, we have $S_{ij} = \sigma_{ij} Q_x$ and $S_b = \lim_{n\to\infty} E\left[\frac{1}{n}\sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} \sigma_{ij}(Q_x \otimes b_i b_j')\right]$. This limit is finite (if it exists), since from BD.4 we have $\left\|\frac{1}{n}\sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} \sigma_{ij}(Q_x \otimes b_i b_j')\right\| \leq C\frac{1}{n}\sum_{i,j}|\sigma_{i,j}|$, and $E\left[\frac{1}{n}\sum_{i,j}|\sigma_{i,j}|\right] = O(1)$ from Assumption A.1.

Moreover:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_i \tau_i Y_{i,T} \otimes b_i = \frac{1}{\sqrt{Tn}}\sum_{t=1}^{T}\sum_{i=1}^{n} w_i \tau_i I_{i,t}\left(x_t \otimes b_i\right)\varepsilon_{i,t} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{n,t},$$

where $\xi_{n,t} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} w_i \tau_i I_{i,t}\left(x_t \otimes b_i\right)\varepsilon_{i,t}$. The triangular array $(\xi_{n,t})$ is a martingale difference sequence w.r.t. the sigma-field $\mathcal{F}_{n,t} = \{f_{\underline{t}}, \varepsilon_{i,\underline{t}}, \gamma_i, i = 1, ..., n\}$. From a multivariate version of Corollary 5.26 in White (2001), the CLT in condition a) follows if we show:

(i) $\frac{1}{T}\sum_{t=1}^{T} E[\xi_{n,t}\xi_{n,t}^{'}] \to S_b$,

(ii) $\frac{1}{T}\sum_{t=1}^{T}\left(\xi_{n,t}\xi_{n,t}^{'} - E[\xi_{n,t}\xi_{n,t}^{'}]\right) = o_p(1)$,

(iii) $\sup_{t=1,...,T} E[\|\xi_{n,t}\|^{2+\delta}] = O(1)$, for some $\delta > 0$.

Moreover, we prove the alternative characterization of the asymptotic variance-covariance matrix:

(iv) $S_b =$ a.s.-$\lim_{n\to\infty}\frac{1}{n}\sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}}\sigma_{ij}(Q_x \otimes b_i b_j^{\prime})$.

Let us check these conditions. (i) Let $\mathcal{G}_n = \{\gamma_i, i = 1, ..., n\}$. We have:

$$\begin{aligned}
\frac{1}{T}\sum_{t} E[\xi_{n,t}\xi_{n,t}^{'}|\mathcal{G}_n] &= \frac{1}{Tn}\sum_{t}\sum_{i,j} w_i w_j \tau_i \tau_j E\left[I_{i,t}I_{j,t}\left(x_t x_t^{'} \otimes b_i b_j^{'}\right)\varepsilon_{i,t}\varepsilon_{j,t}|\gamma_i, \gamma_j\right] \\
&= \frac{1}{Tn}\sum_{t}\sum_{i,j} w_i w_j \tau_i \tau_j E[I_{i,t}I_{j,t}|\gamma_i, \gamma_j]\left(E[x_t x_t^{'}] \otimes b_i b_j^{'}\right) E[\varepsilon_{i,t}\varepsilon_{j,t}|\gamma_i, \gamma_j] \\
&= \frac{1}{n}\sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{i,j}}\sigma_{ij}\left(Q_x \otimes b_i b_j^{'}\right).
\end{aligned}$$

By taking expectation on both sides, condition (i) follows.

Let us now consider condition (ii). Define $\zeta_{n,T} = \frac{1}{T}\sum_{t}(\xi_{n,t,k}\xi_{n,t,l} - E[\xi_{n,t,k}\xi_{n,t,l}])$, where $\xi_{n,t,k}$ is the k-th element of $\xi_{n,t}$. Since $E[\zeta_{n,T}] = 0$, it is enough to show $V[\zeta_{n,T}] = o(1)$, for any k, l. We show this for $k = l$, the proof for $k \neq l$ is similar. For expository purpose we omit the index k, and we write $x_{t,k}^2 \equiv x_t^2$. We have:

$$V[\zeta_{n,T}] = \frac{1}{T^2}\sum_{t} V[\xi_{n,t}^2] + \frac{1}{T^2}\sum_{t\neq s} Cov\left(\xi_{n,t}^2, \xi_{n,s}^2\right), \tag{55}$$

where:

$$\xi_{n,t}^2 = \frac{1}{n} \sum_{i,j} w_i w_j \tau_i \tau_j I_{i,t} I_{j,t} x_t^2 b_i b_j \varepsilon_{i,t} \varepsilon_{j,t}.$$

- Consider first the terms $Cov(\xi_{n,t}^2, \xi_{n,s}^2)$ for $t \neq s$. By the variance decomposition formula:

$$Cov(\xi_{n,t}^2, \xi_{n,s}^2) = E\left[Cov(\xi_{n,t}^2, \xi_{n,s}^2 | \mathcal{G}_n)\right] + Cov\left[E(\xi_{n,t}^2|\mathcal{G}_n), E(\xi_{n,s}^2|\mathcal{G}_n)\right].$$

We have $Cov(\xi_{n,t}^2, \xi_{n,s}^2|\mathcal{G}_n) = 0$ from the i.i.d. assumption over time. Moreover:

$$E[\xi_{n,t}^2|\mathcal{G}_n] = \frac{1}{n} \sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} Q_x \sigma_{ij} b_i b_j = \frac{1}{n} \sum_{m=1}^{J_n} \sum_{i,j} \alpha_{ij} \sigma_{ij} 1\{\gamma_i, \gamma_j \in I_m\},$$

where $\alpha_{ij} = w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} b_i b_j Q_x$ and $Q_x = E[x_t^2]$. Thus:

$$Cov\left[E(\xi_{n,t}^2|\mathcal{G}_n), E(\xi_{n,s}^2|\mathcal{G}_n)\right] = \frac{1}{n^2} \sum_{m,p=1}^{J_n} \sum_{i,j,k,l} Cov\left(\alpha_{ij}\sigma_{ij} 1\{\gamma_i, \gamma_j \in I_m\}, \alpha_{kl}\sigma_{kl} 1\{\gamma_k, \gamma_l \in I_p\}\right).$$

In the above sum, the terms such that sets $\{i, j\}$ and $\{k, l\}$ do not have a common element, vanish. Consider now the sum of the terms such that $i = k$ (terms such that $i = l$, or $j = k$, or $j = l$ are symmetric). Therefore, let us focus on the sum

$$\begin{aligned} S_n &:= \frac{1}{n^2} \sum_{m,p=1}^{J_n} \sum_{i,j,l} Cov\left(\alpha_{ij}\sigma_{ij} 1\{\gamma_i, \gamma_j \in I_m\}, \alpha_{il}\sigma_{il} 1\{\gamma_i, \gamma_l \in I_p\}\right) \\ &= \frac{1}{n^2} \sum_{m=1}^{J_n} \sum_{i,j,l} Cov\left(\alpha_{ij}\sigma_{ij} 1\{\gamma_i, \gamma_j \in I_m\}, \alpha_{il}\sigma_{il} 1\{\gamma_i, \gamma_l \in I_m\}\right) \\ &\quad -\frac{1}{n^2} \sum_{m,p=1, m\neq p}^{J_n} \sum_{i,j,l} E\left[\alpha_{ij}\sigma_{ij} 1\{\gamma_i, \gamma_j \in I_m\}\right] E\left[\alpha_{il}\sigma_{il} 1\{\gamma_i, \gamma_l \in I_p\}\right]. \end{aligned}$$

From BD.4, we have $\alpha_{ij} \leq C$ and $\sigma_{ij} \leq C$. Thus, we get $S_n = O\left(\frac{1}{n^2} \sum_{m=1}^{J_n} \sum_{i,j,l} E[1\{\gamma_i, \gamma_j, \gamma_l \in I_m\}]\right) + O\left(\frac{1}{n^2} \sum_{m,p=1,m\neq p}^{J_n} \sum_{i,j,l} E\left[1\{\gamma_i, \gamma_j \in I_m\}\right] E\left[1\{\gamma_i, \gamma_l \in I_p\}\right]\right)$. By using that $\sum_{i,j,l} E[1\{\gamma_i, \gamma_j, \gamma_l \in I_m\}] = O\left(nB_m + n^2 B_m^2 + n^3 B_m^3\right)$ and $\sum_{i,j,l} E\left[1\{\gamma_i, \gamma_j \in I_m\}\right] E\left[1\{\gamma_i, \gamma_l \in I_p\}\right] = O\left(nB_m B_p + \right.$

$n^2(B_m^2 B_p + B_m B_p^2) + n^3 B_m^2 B_p^2)\big)$, we get $S_n = O\left(1/n + \sum_{m=1}^{J_n} B_m^2 + n\sum_{m=1}^{J_n} B_m^3 + n\left(\sum_{m=1}^{J_n} B_m^2\right)^2\right)$. The RHS is $o(1)$ from BD.2. Thus, we have shown that:

$$Cov(\xi_{n,t}^2, \xi_{n,s}^2) = o(1), \tag{56}$$

uniformly in $t \neq s$.

- Consider now $V[\xi_{n,t}^2]$. By the variance decomposition formula:

$$V[\xi_{n,t}^2] = E\left[V(\xi_{n,t}^2|\mathcal{G}_n)\right] + V\left[E(\xi_{n,t}^2|\mathcal{G}_n)\right].$$

By similar arguments as above, we have $V\left[E(\xi_{n,t}^2|\mathcal{G}_n)\right] = o(1)$ uniformly in t. Consider now term $E\left[V(\xi_{n,t}^2|\mathcal{G}_n)\right]$. We have:

$$\begin{aligned} V(\xi_{n,t}^2|\mathcal{G}_n) &= \frac{1}{n^2}\sum_{i,j,k,l} w_i w_j w_k w_l \tau_i \tau_j \tau_k \tau_l b_i b_j b_k b_l \\ &\quad \cdot Cov\left(I_{i,t}I_{j,t}x_t^2\varepsilon_{i,t}\varepsilon_{j,t}, I_{k,t}I_{l,t}x_t^2\varepsilon_{k,t}\varepsilon_{l,t}|\gamma_i,\gamma_j,\gamma_k,\gamma_l\right). \end{aligned}$$

Moreover:

$$\begin{aligned} &Cov\left(I_{i,t}I_{j,t}x_t^2\varepsilon_{i,t}\varepsilon_{j,t}, I_{k,t}I_{l,t}x_t^2\varepsilon_{k,t}\varepsilon_{l,t}|\gamma_i,\gamma_j,\gamma_k,\gamma_l\right) \\ &= E\left[I_{i,t}I_{j,t}I_{k,t}I_{l,t}|\gamma_i,\gamma_j,\gamma_k,\gamma_l\right] E\left[\varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{k,t}\varepsilon_{l,t}|\gamma_i,\gamma_j,\gamma_k,\gamma_l\right] E[x_t^4] - \sigma_{ij}\sigma_{kl}\tau_{ij}^{-1}\tau_{kl}^{-1}E[x_t^2]^2. \end{aligned}$$

From the block dependence structure in BD.1, the expectation $E\left[\varepsilon_{i,t}\varepsilon_{j,t}\varepsilon_{k,t}\varepsilon_{l,t}|\gamma_i,\gamma_j,\gamma_k,\gamma_l\right]$ is different from zero only if a pair of indices are in a same block I_m, and the other pair is also in a same block I_p, say, possibly with $m = p$. Similarly, $\sigma_{ij}\sigma_{kl}$ is different from zero only if γ_i and γ_j are in the same block and γ_k and γ_l are in the same block. From BD.4, we deduce that $V(\xi_{n,t}^2|\mathcal{G}_n) \leq C\frac{1}{n^2}\sum_{i,j,k,l}\sum_{m,p=1}^{J_n} 1\{\gamma_i,\gamma_j \in I_m\}1\{\gamma_k,\gamma_l \in I_p\}$, where in the double sum the elements with $m \neq p$ are not zero only if the pairs (γ_i,γ_j) and (γ_k,γ_l) have no element in common. Thus:

$$\begin{aligned} E\left[V(\xi_{n,t}^2|\mathcal{G}_n)\right] &= O\left(\frac{1}{n^2}\sum_{i,j,k,l}\sum_{m=1}^{J_n} E[1\{\gamma_i,\gamma_j,\gamma_k,\gamma_l \in I_m\}]\right) \\ &\quad + O\left(\frac{1}{n^2}\sum_{i,j,k,l: i\neq k,l; j\neq k,l}\ \sum_{m,p=1: m\neq p}^{J_n} E[1\{\gamma_i,\gamma_j \in I_m\}]E[1\{\gamma_k,\gamma_l \in I_p\}]\right). \end{aligned}$$

By using $\sum_{i,j,k,l} \sum_{m=1}^{J_n} E[1\{\gamma_i, \gamma_j, \gamma_k, \gamma_l \in I_m\}] = O\left(\sum_{m=1}^{J_n} (nB_m + n^2 B_m^2 + n^3 B_m^3 + n^4 B_m^4)\right)$ and $\sum_{i,j,k,l} \sum_{m,p=1}^{J_n} E[1\{\gamma_i, \gamma_j \in I_m\}]E[1\{\gamma_k, \gamma_l \in I_p\}] = O\left(\sum_{m,p=1}^{J_n} (n^2 B_m B_p + n^3 B_m^2 B_p + n^4 B_m^2 B_p^2)\right)$, we get:

$$E\left[V(\xi_{n,t}^2|\mathcal{G}_n)\right] = O\left(1 + n\sum_{m=1}^{J_n} B_m^2 + (n\sum_{m=1}^{J_n} B_m^2)^2 + n^2 \sum_{m=1}^{J_n} B_m^4\right).$$

By BD.2, $n \max_{m=1,\dots,n} B_m^2 = O(1)$, and we get $E\left[V(\xi_{n,t}^2|\mathcal{G}_n)\right] = O(1)$.

Thus, we have shown:

$$V(\xi_{n,t}^2) = O(1), \tag{57}$$

uniformly in t.

From (55), (56) and (57), we get $V[\zeta_{nT}] = o(1)$, and condition (ii) follows. From (57) and by using $E[\xi_{n,t}^2] = O(1)$, condition (iii) follows for $\delta = 2$. Finally, condition (iv) follows from $\frac{1}{n}\sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} \sigma_{ij} b_i b_j' = (1 + \lambda' V[f_t]\lambda)^{-2} \frac{1}{n}\sum_{i,j} \frac{1}{\tau_{ij}} \frac{\sigma_{i,j}}{\sigma_{ii}\sigma_{jj}} b_i b_j'$ and the next Lemma 13.

Lemma 13 *Under Assumptions BD.1-BD.4:* $\frac{1}{n}\sum_{i,j} \frac{1}{\tau_{ij}} \frac{\sigma_{i,j}}{\sigma_{ii}\sigma_{jj}} b_i b_j' \to L$, *$P$-a.s., where:*

$$L = \lim_{n\to\infty} E\left[\frac{1}{n}\sum_{i,j} \frac{1}{\tau_{ij}} \frac{\sigma_{i,j}}{\sigma_{ii}\sigma_{jj}} b_i b_j'\right] = \int_0^1 \omega(\gamma) d\gamma + \lim_{n\to\infty} n \sum_{m=1}^{J_n} \int_{I_m} \int_{I_m} \omega(\gamma,\gamma') d\gamma d\gamma',$$

with $\omega(\gamma,\gamma') := E[I_t(\gamma)I_t(\gamma')] \frac{E[\varepsilon_t(\gamma)\varepsilon_t(\gamma')]}{E[\varepsilon_t(\gamma)^2]E[\varepsilon_t(\gamma')^2]} b(\gamma)b(\gamma')'$ *and* $\omega(\gamma) := \omega(\gamma,\gamma)$.

Then, we have proved part a). Part b) follows by a standard CLT.

A.5.10.4 Assumption A.3

Assumption A.3 is satisfied since the errors are i.i.d. and have zero third moment (Assumption BD.1).

A.5.10.5 Assumption A.4

We have to show that $\max_i \sum_j \|S_{ij}\|^q = O_p(n^\delta)$, for any $q \in (0,1)$ and $\delta > 1/2$. From $S_{ij} = \sigma_{ij} Q_x$, and an argument similar to (54):

$$\max_i \sum_j \|S_{ij}\|^q \leq C \max_{m=1,\ldots,J_n} \sum_{j=1}^n 1\{\gamma_j \in I_m\} \leq Cn \max_{m=1,\ldots,J_n} B_m + C \max_{m=1,\ldots,J_n} \left| \sum_{j=1}^n [1\{\gamma_j \in I_m\} - B_m] \right|,$$

for any $q > 0$. Let us derive (probability) bounds for the two terms in the RHS. From BD.2:

$$n \max_m |B_m| \leq \sqrt{n} \left(n \sum_m |B_m|^2 \right)^{1/2} = O\left(\sqrt{n}\right).$$

Let $\varepsilon_n := n^\delta$, with $\delta > 1/2$. Then:

$$\begin{aligned} P\left[\max_{m=1,\ldots,J_n} \left| \sum_{j=1}^n [1\{\gamma_j \in I_m\} - B_m] \right| \geq \varepsilon_n \right] &\leq J_n \max_{m=1,\ldots,J_n} P\left[\left| \sum_{j=1}^n [1\{\gamma_j \in I_m\} - B_m] \right| \geq \varepsilon_n \right] \\ &\leq 2J_n \exp(-\varepsilon_n^2/(2n)) = o(1), \end{aligned}$$

from the Hoeffding's inequality (see Bosq (1998), Theorem 1.2), and $J_n \leq n$. Thus, we have shown that $\max_{m=1,\ldots,J_n} \left| \sum_{j=1}^n [1\{\gamma_j \in I_m\} - B_m] \right| = o_p(n^\delta)$, and the conclusion follows.

A.5.10.6 Assumption A.5

We have $S_{ii,T} = \sigma_{ii} \hat{Q}_{x,i}$ and $S_{ij} = \sigma_{ij} Q_x$. Let us denote by $\mathcal{H} = \sigma\left((f_t), (I_t(\gamma)), \gamma \in [0,1], \gamma_i, i = 1, 2, \ldots\right)$ the information in the factor path, the indicators paths and the individual random effects. The proof is in two steps.

STEP 1: We first show that conditional on $\mathcal{H}$ we have

$$\Upsilon_{nT} := \frac{1}{\sqrt{n}} \sum_i w_i \tau_i^2 \left[Y_{i,T} \otimes Y_{i,T} - \widetilde{S}_{ii,T} \right] \Rightarrow N(0, \Omega), \quad n, T \to \infty, \tag{58}$$

P-a.s., where $\widetilde{S}_{ii,T} = \sigma_{ii} vec(\hat{Q}_{x,i})$ and $\Omega = \lim_{n \to \infty} E\left[\frac{1}{n} \sum_{i,j} w_i w_j \frac{\tau_i^2 \tau_j^2}{\tau_{ij}^2} \sigma_{ij}^2 \right] \left[Q_x \otimes Q_x + (Q_x \otimes Q_x) W_{K+1} \right].$

For this purpose, we apply the Lyapunov CLT for heterogenous independent arrays (see Davidson (1994), Theorem 23.11). Write

$$\Upsilon_{nT} = \frac{1}{\sqrt{n}} \sum_i \sum_{m=1}^{J_n} 1\{\gamma_i \in I_m\} w_i \tau_i^2 \left[Y_{i,T} \otimes Y_{i,T} - \widetilde{S}_{ii,T} \right] = \frac{1}{\sqrt{J_n}} \sum_{m=1}^{J_n} W_{m,nT},$$

where

$$W_{m,nT} := \sqrt{\frac{J_n}{n}} \sum_i 1\{\gamma_i \in I_m\} w_i \tau_i^2 \left[Y_{i,T} \otimes Y_{i,T} - \widetilde{S}_{ii,T} \right].$$

Conditional on $\mathcal{H}$, the variables $W_{m,nT}$, for $m = 1, ..., J_n$ are independent, with zero mean. The conclusion follows if we prove:

(i) $\lim_{n,T} \frac{1}{J_n} \sum_m V\left[W_{m,nT} | \mathcal{H}\right] = \Omega$, P-a.s, and

(ii) $\lim_{n,T} \frac{1}{J_n^{3/2}} \sum_m E\left[\|W_{m,nT}\|^3 | \mathcal{H} \right] = 0$, P-a.s..

To show (i), we use:

$$\begin{aligned} V\left[W_{m,nT} | \mathcal{H}\right] &= \frac{J_n}{n} \sum_{i,j \in I_m} w_i w_j \tau_i^2 \tau_j^2 Cov\left[Y_{i,T} \otimes Y_{i,T}, Y_{j,T} \otimes Y_{j,T} | \mathcal{H}\right] \\ &= \frac{J_n}{n} \sum_{i,j \in I_m} w_i w_j \tau_i^2 \tau_j^2 \left\{ E\left[\left(Y_{i,T} \otimes Y_{i,T}\right)\left(Y_{j,T} \otimes Y_{j,T}\right)' | \mathcal{H} \right] - \widetilde{S}_{ii,T} \widetilde{S}_{jj,T}' \right\}, \end{aligned}$$

where $\sum_{i,j \in I_m}$ denotes double sum over all $i, j = 1, ..., n$ such that $\gamma_i, \gamma_j \in I_m$. Now, we have by the independence property over time:

$$\begin{aligned} & E\left[\left(Y_{i,T} \otimes Y_{i,T}\right)\left(Y_{j,T} \otimes Y_{j,T}\right)' | \mathcal{H} \right] \\ = \; & \frac{1}{T^2} \sum_t \sum_s \sum_p \sum_q E\left[\varepsilon_{i,t}\varepsilon_{i,p}\varepsilon_{j,s}\varepsilon_{j,q} | (f_t), \gamma_i, \gamma_j\right] I_{i,t} I_{i,p} I_{j,s} I_{j,q} \left(x_t x_s' \otimes x_p x_q' \right) \\ = \; & E\left[\varepsilon_{it}^2 \varepsilon_{jt}^2 | \gamma_i, \gamma_j\right] \frac{1}{T^2} \sum_t I_{i,t} I_{j,t} \left(x_t x_t' \otimes x_t x_t' \right) + \sigma_{ij}^2 \frac{1}{T^2} \sum_t \sum_{p \neq t} I_{ij,t} I_{ij,p} \left(x_t x_t' \otimes x_p x_p' \right) \\ & + \sigma_{ii}^2 \sigma_{jj}^2 \frac{1}{T^2} \sum_t \sum_{s \neq t} I_{i,t} I_{j,s} \left(x_t x_s' \otimes x_t x_s' \right) + \sigma_{ij}^2 \frac{1}{T^2} \sum_t \sum_{s \neq t} I_{ij,t} I_{ij,s} \left(x_t x_s' \otimes x_s x_t' \right) \\ =: \; & E\left[\varepsilon_{it}^2 \varepsilon_{jt}^2 | \gamma_i, \gamma_j\right] A_{1,T} + \sigma_{ij}^2 A_{2,T} + \sigma_{ii}^2 \sigma_{jj}^2 A_{3,T} + \sigma_{ij}^2 A_{4,T}. \end{aligned}$$

Moreover, $A_{1,T} = \frac{T_{ij}}{T^2} \sum_t \frac{I_{ij,t}}{T_{ij}} \left(x_t x_t' \otimes x_t x_t' \right) = O\left(T_{ij}/T^2\right) = O(1/T)$, uniformly in $\mathcal{H}$. Let us define $\hat{Q}_{x,ij} = \frac{1}{T_{ij}} \sum_t I_{ij,t} x_t x_t'$, then

$$
\begin{aligned}
A_{2,T} &= \frac{1}{T^2} \sum_t \sum_p I_{ij,t} I_{ij,p} \left(x_t x_t^{'} \otimes x_p x_p^{'} \right) - A_{1,T} = \frac{1}{\tau_{ij,T}^2} \left(\hat{Q}_{x,ij} \otimes \hat{Q}_{x,ij} \right) + O\left(1/T\right), \\
A_{3,T} &= \frac{1}{T^2} \sum_t \sum_s I_{i,t} I_{j,s} \left(x_t x_s^{'} \otimes x_t x_s^{'} \right) - A_{1,T} = vec\left(\hat{Q}_{x,i}\right) vec\left(\hat{Q}_{x,j}\right)^{'} + O\left(1/T\right),
\end{aligned}
$$

and

$$
\begin{aligned}
A_{4,T} &= \frac{1}{T^2} \sum_t \sum_s I_{ij,t} I_{ij,s} \left(x_t x_s^{'} \otimes x_s x_t^{'} \right) - A_{1,T} \\
&= \frac{1}{T^2} \sum_t \sum_s I_{ij,t} I_{ij,s} \left(x_t \otimes x_s \right) \left(x_s \otimes x_t \right)^{'} - A_{1,T} \\
&= \frac{1}{T^2} \sum_t \sum_s I_{ij,t} I_{ij,s} \left(x_t \otimes x_s \right) \left(x_t \otimes x_s \right)^{'} W_{K+1} - A_{1,T} \\
&= \frac{1}{\tau_{ij,T}^2} \left(\hat{Q}_{x,ij} \otimes \hat{Q}_{x,ij} \right) W_{K+1} + O\left(1/T\right).
\end{aligned}
$$

Then, it follows that:

$$
\begin{aligned}
V\left[W_{m,nT} | \mathcal{H}\right] &= \frac{J_n}{n} \left[\sum_{i,j \in I_m} w_i w_j \frac{\tau_i^2 \tau_j^2}{\tau_{ij,T}^2} \sigma_{ij}^2 \left(\hat{Q}_{x,ij} \otimes \hat{Q}_{x,ij} + \hat{Q}_{x,ij} \otimes \hat{Q}_{x,ij} W_{K+1} \right) \right] \\
&\quad + O\left(\frac{J_n}{n} \frac{1}{T} \sum_{i,j \in I_m} w_i w_j \tau_i^2 \tau_j^2 \right),
\end{aligned}
$$

where the O term is uniform w.r.t. $\mathcal{H}$. Thus, we get:

$$
\begin{aligned}
\frac{1}{J_n} \sum_m V\left[W_{m,nT} | \mathcal{H}\right] &= \left(\frac{1}{n} \sum_{i,j} w_i w_j \frac{\tau_i^2 \tau_j^2}{\tau_{ij}^2} \sigma_{ij}^2 \right) \left(Q_x \otimes Q_x + Q_x \otimes Q_x W_{K+1} \right) \\
&\quad + \frac{1}{n} \sum_m \sum_{i,j \in I_m} w_i w_j \tau_i^2 \tau_j^2 \sigma_{ij}^2 \alpha_{ij} + O\left(\frac{1}{T} \frac{1}{n} \sum_m \sum_{i,j \in I_m} w_i w_j \tau_i^2 \tau_j^2 \right),
\end{aligned}
$$

where the $\alpha_{ij} = \frac{1}{\tau_{ij,T}^2} \left(\hat{Q}_{x,ij} \otimes \hat{Q}_{x,ij} + \hat{Q}_{x,ij} \otimes \hat{Q}_{x,ij} W_{K+1} \right) - \frac{1}{\tau_{ij}^2} \left(Q_x \otimes Q_x + Q_x \otimes Q_x W_{K+1} \right)$ are $o(1)$ uniformly in i, j, and $w_i w_j \frac{\tau_i^2 \tau_j^2}{\tau_{ij}^2} \sigma_{ij}^2 = (1 + \lambda' \Sigma_f^{-1} \lambda)^{-2} \frac{\tau_i \tau_j}{\tau_{ij}^2} \frac{\sigma_{ij}^2}{\sigma_{ii} \sigma_{jj}}$. Then, point i) follows from

$\dfrac{1}{n}\sum_{i,j}\dfrac{\tau_i\tau_j}{\tau_{ij}^2}\dfrac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}} \to L$, P-a.s., where $L=\lim_{n\to\infty}E\left[\dfrac{1}{n}\sum_{i,j}\dfrac{\tau_i\tau_j}{\tau_{ij}^2}\dfrac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}}\right]$, which is proved by similar arguments as Lemma 13.

Let us now prove point ii). We have:

$$\begin{aligned}\frac{1}{J_n^{3/2}}\sum_m E\left[\|W_{m,nT}\|^3|\mathcal{H}\right] &\leq \frac{1}{n^{3/2}}\sum_m\left[\sum_{i\in I_m}w_i\tau_i^2\left(E\left[\|(Y_{i,T}\otimes Y_{i,T})\|^3|\mathcal{H}\right]^{1/3}+\left\|\widetilde{S}_{ii,T}\right\|\right)\right]^3\\ &\leq \frac{1}{n^{3/2}}\left(\sum_m\left(\sum_{i\in I_m}w_i\tau_i^2\right)^3\right)\left(\sup_i E\left[\|Y_{i,T}\otimes Y_{i,T}\|^3|\mathcal{H}\right]^{1/3}+\sup_i\left\|\widetilde{S}_{ii,T}\right\|\right)^3.\end{aligned}$$

Now,

$$\begin{aligned}E\left[\|Y_{i,T}\otimes Y_{i,T}\|^3|\mathcal{H}\right] &\leq E\left[\|Y_{i,T}\|^6|\mathcal{H}\right]=E\left[\left(Y_{i,T}'Y_{i,T}\right)^3|\mathcal{H}\right]\\ &= \frac{1}{T^3}\sum_{t_1,\dots,t_6}I_{i,t_1}\dots I_{i,t_6}E\left[\varepsilon_{i,t_1}\dots\varepsilon_{i,t_6}|\gamma_i\right]\left(x_{t_1}'x_{t_2}\right)\left(x_{t_3}'x_{t_4}\right)\left(x_{t_5}'x_{t_6}\right).\end{aligned}$$

By the independence property, the non-zero terms $E\left[\varepsilon_{i,t_1}\dots\varepsilon_{i,t_6}|\gamma_i\right]$ involve at most 3 different time indices, which implies together with BD.4 that $\sup_i E\left[\|Y_{i,T}\otimes Y_{i,T}\|^3|\mathcal{H}\right]=O(1)$, P-a.s. Similarly $\sup_i\left\|\widetilde{S}_{ii,T}\right\|=O(1)$, P-a.s. Thus, we get:

$$\frac{1}{J_n^{3/2}}\sum_{m=1}^{J_n}E\left[\|W_{m,nT}\|^3|\mathcal{H}\right]\leq C\frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i\in I_m\}\right)^3.$$

Then, point ii) follows from the next Lemma 14.

Lemma 14 *Under Assumptions BD.1-BD.4:* $\dfrac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i\in I_m\}\right)^3\to 0$, *$P$-a.s.*

STEP 2: We show that (58) implies the asymptotic normality condition in Assumption A.4. Indeed, from (58) we have:

$$\lim_{n,T\to\infty}P\left[\alpha'\Upsilon_{nT}\leq z|\mathcal{H}\right]=\Phi\left(\frac{z}{\sqrt{\alpha'\Omega\alpha}}\right),$$

for any $\alpha\in\mathbb{R}^{2(K+1)}$ and for any $z\in\mathbb{R}$, and P-a.s. We now apply the Lebesgue dominated convergence theorem, by using that the sequence of random variables $P\left[\alpha'\Upsilon_{nT}\leq z|\mathcal{H}\right]$ are such that $P\left[\alpha'\Upsilon_{nT}\leq z|\mathcal{H}\right]\leq 1$, uniformly in n and T. We conclude that, for any $\alpha\in\mathbb{R}^{2(K+1)}$, $z\in\mathbb{R}$:

$$\lim_{n,T\to\infty}P\left[\alpha'\Upsilon_{nT}\leq z\right]=\lim_{n,T\to\infty}E\left(P\left[\alpha'\Upsilon_{nT}\leq z|\mathcal{H}\right]\right)=\Phi\left(\frac{z}{\sqrt{\alpha'\Omega\alpha}}\right),$$

since $\Phi\left(\frac{z}{\sqrt{\alpha'\Omega\alpha}}\right)$ is independent of the information set $\mathcal{H}$. The conclusion follows.

A.5.11 Proof Lemma 12

Write:

$$B^{-1} - A^{-1} = \left[A\left(I - A^{-1}(A-B)\right)\right]^{-1} - A^{-1} = \left\{\left[I - A^{-1}(A-B)\right]^{-1} - I\right\} A^{-1},$$

and use that, for a square matrix C such that $\|C\| < 1$, we have

$$(I-C)^{-1} = I + C + C^2 + C^3 + ...$$

and

$$\left\|(I-C)^{-1} - I\right\| \quad \leq \|C\| + \|C\|^2 + ... \quad \leq \frac{\|C\|}{1-\|C\|}.$$

Thus, we get:

$$\begin{aligned} \left\|B^{-1} - A^{-1}\right\| &\leq \frac{\left\|A^{-1}(A-B)\right\|}{1-\left\|A^{-1}(A-B)\right\|}\left\|A^{-1}\right\| \\ &\leq \frac{\left\|A^{-1}\right\|^2\|A-B\|}{1-\left\|A^{-1}\right\|\|A-B\|} \\ &\leq 2\left\|A^{-1}\right\|^2\|A-B\|, \end{aligned}$$

if $\|A-B\| < \frac{1}{2}\|A^{-1}\|^{-1}$.

A.5.12 Proof of Lemma 13

Let us denote $\xi_{i,j} = \frac{1}{\tau_{ij}}\frac{\sigma_{ij}}{\sigma_{ii}\sigma_{jj}}b_i b_j' = w(\gamma_i, \gamma_j)$. We have $\frac{1}{n}\sum_{i,j}\xi_{i,j} = \frac{1}{n}\sum_i \xi_{ii} + \frac{1}{n}\sum_{i\neq j}\xi_{i,j}$. By the LLN we get $\frac{1}{n}\sum_i \xi_{ii} = \frac{1}{n}\sum_i \omega(\gamma_i) \to \int_0^1 \omega(\gamma)d\gamma$, P-a.s.. Let us now consider the double sum $\frac{1}{n}\sum_{i\neq j}\xi_{i,j}$. The proof proceeds in three steps.

STEP 1: We first prove that $\frac{1}{n}\sum_{i\neq j}\xi_{i,j} = L' + o_p(1)$, where $L' := \lim_{n\to\infty} n\sum_{m=1}^{J_n}\int_{I_m}\int_{I_m}\omega(\gamma,\gamma')d\gamma d\gamma'$. For this purpose, write $\frac{1}{n}\sum_{i\neq j}\xi_{i,j} = \sum_{m=1}^{J_n} X_m$, where $X_m := \frac{1}{n}\sum_{i\neq j}\omega(\gamma_i,\gamma_j)1\{\gamma_i,\gamma_j \in I_m\}$, by using block-dependence. Then, we have:

$$E[X_m] = \frac{1}{n}\sum_{i\neq j} E[\omega(\gamma_i,\gamma_j)1\{\gamma_i,\gamma_j \in I_m\}] = (n-1)\int_{I_m}\int_{I_m}\omega(\gamma,\gamma')d\gamma d\gamma' =: (n-1)\bar{\omega}_m,$$

which implies:

$$E\left[\frac{1}{n}\sum_{i\neq j}\xi_{i,j}\right] = (n-1)\sum_{m=1}^{J_n}\bar{\omega}_m \to L'.$$

Moreover:

$$\begin{aligned} V\left[X_m\right] &= \frac{1}{n^2}\sum_{i\neq j}\sum_{k\neq l} E\left[\omega(\gamma_i,\gamma_j)\omega(\gamma_k,\gamma_l)1\{\gamma_i,\gamma_j,\gamma_k,\gamma_l \in I_m\}\right] - E[X_m]^2 \\ &= \frac{1}{n^2}\left[n(n-1)(n-2)(n-3)\bar{\omega}_m^2 + O(n^3B_m^3) + O(n^2B_m^2)\right] - (n-1)^2\bar{\omega}_m^2 \\ &= O(nB_m^4) + O(nB_m^3) + O(B_m^2), \end{aligned}$$

and:

$$\begin{aligned} Cov(X_m,X_p) &= \frac{1}{n^2}\sum_{i\neq j}\sum_{k\neq l} E\left[\omega(\gamma_i,\gamma_j)\omega(\gamma_k,\gamma_l)1\{\gamma_i,\gamma_j \in I_m\}1\{\gamma_k,\gamma_l \in I_p\}\right] - E[X_m]E[X_p] \\ &= \frac{1}{n^2}\left[n(n-1)(n-2)(n-3)\bar{\omega}_m\bar{\omega}_p\right] - (n-1)^2\bar{\omega}_m\bar{\omega}_p = O(nB_m^2B_p^2), \end{aligned}$$

for $m \neq p$, which implies:

$$V\left[\frac{1}{n}\sum_{i\neq j}\xi_{i,j}\right] = \sum_{m=1}^{J_n} V[X_m] + \sum_{m,p=1,m\neq p}^{J_n} Cov(X_m,X_p) = o(1),$$

from BD.2. Then, Step 1 follows.

STEP 2: There exists a random variable $\tilde{L}$ such that $\frac{1}{n}\sum_{i\neq j}\xi_{i,j} \to \tilde{L}$, P-a.s.. To show this statement, we use that the event in which series $\frac{1}{n}\sum_{i\neq j}\xi_{i,j}$ converges is a tail event for the i.i.d. sequence (γ_i). Indeed, we have that $\frac{1}{n}\sum_{i\neq j}\xi_{i,j}$ converges if, and only if, $\frac{1}{n}\sum_{i,j\geq N, i\neq j}\xi_{i,j}$ converges, for any integer N. Then, by the Kolmogorov zero-one law, the event in which series $\frac{1}{n}\sum_{i\neq j}\xi_{i,j}$ converges has probability either 1 or 0. The latter case however is excluded by Step 1. Therefore, the sequence $\frac{1}{n}\sum_{i\neq j}\xi_{i,j}$ converges with probability 1, and Step 2 follows.

STEP 3: We have $\tilde{L} = L'$, with probability 1. Indeed, by Steps 1 and 2 it follows $\frac{1}{n}\sum_{i\neq j}\xi_{i,j} - L' = o_p(1)$ and $\frac{1}{n}\sum_{i\neq j}\xi_{i,j} - \tilde{L} = o_p(1)$. These equations imply that $\tilde{L} - L' = o_p(1)$, which holds if and only if $\tilde{L} = L$ with probability 1 (since $\tilde{L}$ and L' are independent of n).

A.5.13 Proof of Lemma 14

The proof is similar to the one of Lemma 13 and we give only the main steps. First, we prove that $\frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i \in I_m\}\right)^3 = o_p(1)$. Indeed, we have:

$$E\left[\frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i \in I_m\}\right)^3\right] = \frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\sum_{i,j,k} E\left[1\gamma_i, \gamma_j, \gamma_k \in I_m\right] = O\left(n^{3/2}\sum_{m=1}^{J_n} B_m^3\right) = o(1),$$

from Assumption BD.2, and we can show $V\left[\frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i \in I_m\}\right)^3\right] = o(1)$. Second, by using the monotone convergence theorem and the Kolmogorov zero-one law, we can show that sequence $\frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i \in I_m\}\right)^3$ converges with probability 1. Third, we conclude that the limit is 0 with probability 1.

Appendix 6: Monte-Carlo experiments

In this section, we perform simulation exercises on balanced and unbalanced panels in order to study the properties of our estimation and testing approaches. We pay particular attention to the interaction between panel dimensions n and T in finite samples since we face conditions like $n = o(T^3)$ for inference with $\hat{\nu}$, and $n = o(T^2)$ for inference with $\hat{Q}_e$ and $\hat{Q}_a$, in the theoretical results. The simulation design mimics the empirical features of our data. The balanced case serves as benchmark to understand when T is not sufficiently large w.r.t. n to apply the theory. The unbalanced case shows that we can exploit the guidelines found for the balanced case when we substitute the average of the sample sizes of the individual assets, i.e., a kind of operative sample size, for T. To summarize our Monte Carlo findings, we do not face any finite sample distortions for the inference with $\hat{\nu}$ when $n = 1,000$ and $T = 150$, and with $\hat{Q}_e$ and $\hat{Q}_a$ when $n = 1,000$ and $T = 350$. In light of these results, we do not expect to face significant inference bias in our empirical application.

A.6.1 Balanced panel

We simulate S datasets of excess returns from an unconditional one-factor model (CAPM), we estimate the parameter ν, and compute the test statistics. A simulated dataset includes: a vector of intercepts $a^s \in \mathbb{R}^n$, a vector of factor loadings $b^s \in \mathbb{R}^n$, and a variance-covariance matrix $\Omega^s \in \mathbb{R}^{n \times n}$. At each simulation $s = 1, ..., S$, we randomly draw $n \leq 9,904$ assets from the empirical sample that comprises $9,904$ individual stocks. The assets are listed by industrial sectors. We use the classification proposed by Ferson and Harvey (1999). The vector b^s is composed by the estimated factor loadings for the n randomly chosen assets. At each simulation, we build a block diagonal matrix Ω^s with blocks matching industrial sectors. The n elements of the main diagonal of Ω^s correspond to the variances of the estimated residuals of the individual assets. The off-diagonal elements of Ω^s are covariances computed by fixing correlations within a block equal to the average correlation of the industrial sector computed from the $9,904 \times 9,904$ thresholded variance-covariance matrix of estimated residuals. Hence we get a setting in line with the block dependence case developed in Appendix 4.

In order to study the size and power properties of our procedure, we set the values of the intercepts a_i^s according to four data generating processes:

DGP1: The true parameter is $\nu_0 = 0.00\%$ and the a_i^s are generated under the null hypothesis $\mathcal{H}_0 : \ a_i^s = 0$;

DGP2: The true parameter is the empirical estimate of ν, $\nu_0 = 2.57\%$, and the a_i^s are generated under the null hypothesis $\mathcal{H}_0 : \ a_i^s = b_i^s \nu_0$;

DGP3: The a_i^s are generated under the alternative hypothesis $\mathcal{H}_a : \ a_i^s = (0.5 b_i^s + 0.5)\, \nu_0$, where $\nu_0 = 2.57\%$;

DGP4: The a_i^s are generated under the three-factor alternative hypothesis: $\mathcal{H}_a : \ a_i^s = b_{i,(3)}^{s\prime} \nu_{0,(3)}$ where $b_{i,(3)}^s \in \mathbb{R}^3$ and $\nu_{0,(3)} = [2.92\%, \ -0.63\%, \ -9.96\%]'$ are estimates for the Fama-French model on the

CRSP dataset.

DGP1 and DGP2 match two different null hypotheses. The null hypothesis for DGP1 assumes that the factor comes from a tradable asset, and for DGP2 that it does not. DGP3 and DGP4 match two different alternative hypotheses as suggested by MacKinlay (1995). DGP3 is a "non risk-based alternative". It represents a deviation from CAPM, which is unrelated to risk: we take the one-factor model calibrated on the data with intercepts deviating from the no arbitrage restriction. DGP4 is a "risk-based alternative". It represents a deviation from CAPM, which comes from missing risk factors: we take intercepts from a three-factor model calibrated on the data, and then we estimate a one-factor model.

Let us define the simulated excess returns $R_{i,t}^s$ of asset i at time t as follows

$$R_{i,t}^s = a_i^s + b_i^s f_t + \varepsilon_{i,t}^s, \text{ for } i = 1, ..., n, \text{ and } t = 1, ..., T, \tag{59}$$

where f_t is the market excess return and $\varepsilon_{i,t}^s$ is the error term. The $n \times 1$ error vectors $\boldsymbol{\varepsilon}_t^s$ are independent across time and Gaussian with mean zero and variance-covariance matrix Ω^s. We apply our estimation approach on every simulated dataset of excess returns. We estimate the parameter ν and we compute the statistics described in Section 2.5 of the paper. Since the panel is balanced, we do not need to fix $\chi_{2,T}$. We only use $\chi_{1,T} = 15$. However, this trimming level does not affect the number of assets n in the simulations. In order to compute the thresholded estimator of the variance-covariance matrix of $\hat{\nu}$, namely $\tilde{\Sigma}_\nu$ (see Proposition 4 in the paper), and the thresholded variance estimator $\tilde{\Sigma}_\xi$ for the test statistics, we fix the parameter M equal to 0.0780, that is used in the empirical application. We define the parameter M using a cross-validation method as proposed in Bickel and Levina (2008). We build random subsamples from the CRSP sample. For each subsample, we minimize a risk function that exploits the difference between a thresholded variance-covariance matrix and a target variance-covariance matrix (see Bickel and Levina (2008) for details).

In order to understand how our estimation approach works for different finite samples, we perform exercises combining different values of the cross-sectional dimension n and the time dimension T. Table 5 reports estimation results for estimator $\hat{\nu}$, and for the bias-adjusted estimator $\hat{\nu}_B$, under DGP 1 and 2. The results include the bias of both estimators, the variance and the Root Mean Square Error (RMSE) of estimator $\hat{\nu}_B$, and the coverage of the 95% confidence interval for parameter ν based on Proposition 4. The bias of estimator $\hat{\nu}$ is decreasing in absolute value with time series size T and is rather stable w.r.t. cross-sectional size n. The analytical bias correction is rather effective, and the bias of estimator $\hat{\nu}_B$ is small. For instance, for sample sizes $T = 150$ and $n = 1000$, under DGP 2 the bias of estimator $\hat{\nu}_B$ is equal to -0.03, which in absolute value is about 1% of the true value of the parameter $\nu = 2.57$. The variance of estimator $\hat{\nu}_B$ is decreasing w.r.t. both time-series and cross-sectional sample sizes T and n. These features reflect the large sample distribution of the estimators derived in Proposition 3. The coverage of the confidence intervals is close to the nominal level 95% across the considered designs and sample sizes.

In Table 6, we display the rejection rates for the test of the null hypothesis $\nu = 0$ (tradable factor). This null hypothesis is satisfied in DGP 1, and the rejection rates are rather close to the nominal size 5% of the test, with a slight overrejection. In DGP 2, parameter ν is different from zero, and the test features a power equal to 100%.

Tables 7 and 8 report the results for the tests of the null hypotheses $\mathcal{H}_0 : a(\gamma) = 0$ and $\mathcal{H}_0 : a(\gamma) = b(\gamma)' \nu$, respectively. The test statistics are based on $\hat{Q}_a$ and $\hat{Q}_e$ as defined in Proposition 5. DGP 1 satisfies the null hypothesis for both tests. For $T = 150$, we observe an oversize, that is increasing w.r.t. cross-sectional size n. The time series dimension $T = 150$ is likely too small compared to cross-sectional size $n = 1000$ and this combination does not reflect the condition $n = o(T^2)$ for the validity of the asymptotic Gaussian approximation of the statistics. For $T = 500$ instead, the rejection rates of the tests are quite close to the nominal size. DGP 2 satisfies the null hypothesis of the test based on $\hat{Q}_e$, but corresponds to an alternative hypothesis for the test based on $\hat{Q}_a$. The former statistic features a similar behaviour as under DGP 1, while the power of the latter statistic is increasing w.r.t. n. Finally, the power of both statistics under the "non risk-based"' and "risk-based" alternatives in DGP 3 and DGP 4 is very large, with rejection rates close to 100% for all considered combinations of sample sizes n and T.

Table 5: Estimation of ν, balanced case

$T = 150$	DGP 1				DGP 2			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Bias($\hat{\nu}$)	-0.0742	-0.0567	-0.0585	-0.0586	-0.1630	-0.1472	-0.1484	-0.1493
Bias($\hat{\nu}_B$)	-0.0244	-0.0063	-0.0082	-0.0083	-0.0319	-0.0156	-0.0169	-0.0178
Var($\hat{\nu}_B$)	0.1167	0.0394	0.0179	0.0121	0.1140	0.0401	0.0189	0.0121
RMSE($\hat{\nu}_B$)	0.3423	0.1985	0.1340	0.1102	0.3390	0.2007	0.1383	0.1114
Coverage	0.9320	0.9290	0.9350	0.9370	0.9370	0.9290	0.9320	0.9360
$T = 500$	DGP 1				DGP 2			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Bias ($\hat{\nu}$)	-0.0587	-0.0640	-0.0687	-0.0654	-0.0847	-0.0926	-0.0972	-0.0937
Bias($\hat{\nu}_B$)	-0.0002	-0.0063	-0.0110	-0.0077	-0.0025	-0.0074	-0.0120	-0.0085
Var($\hat{\nu}_B$)	0.0343	0.0113	0.0060	0.0040	0.0341	0.0114	0.0061	0.0041
RMSE($\hat{\nu}_B$)	0.1851	0.1066	0.0781	0.0634	0.1846	0.1068	0.0788	0.0642
Coverage	0.9370	0.9340	0.9370	0.9390	0.9430	0.9370	0.9360	0.9320

Table 6: Test of $\nu = 0$, balanced case

$T = 150$	DGP 1				DGP 2			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Rejection rate	0.0680	0.0710	0.0650	0.0630	1.0000	1.0000	1.0000	1.0000
$T = 500$	DGP 1				DGP 2			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Rejection rate	0.0630	0.0660	0.0630	0.0610	1.0000	1.0000	1.0000	1.0000

Table 7: Test of the null hypothesis $\mathcal{H}_0: \ a(\gamma) = 0$, balanced case

$T = 150$	**DGP 1**			**DGP 2**			**DGP 3**			**DGP 4**		
n	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500
Size/Power	0.1180	0.1400	0.1500	0.3850	0.5720	0.7170	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$T = 500$	**DGP 1**			**DGP 2**			**DGP 3**			**DGP 4**		
n	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500
Size/Power	0.0730	0.0610	0.0740	0.9240	0.9920	0.9970	0.9990	1.0000	1.0000	0.9990	1.0000	1.0000

Table 8: Test of the null hypothesis $\mathcal{H}_0: \ a(\gamma) = b(\gamma)' \nu$, balanced case

$T = 150$	**DGP 1**			**DGP 2**			**DGP 3**			**DGP 4**		
n	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500
Size/Power	0.1110	0.1340	0.1460	0.1070	0.1360	0.1420	0.9970	1.0000	1.0000	1.0000	1.0000	1.0000
$T = 500$	**DGP 1**			**DGP 2**			**DGP 3**			**DGP 4**		
n	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500	500	1,000	1,500
Size/Power	0.0710	0.0570	0.0730	0.0730	0.0690	0.0750	0.9990	1.0000	1.0000	0.9990	1.0000	1.0000

A.6.2 Unbalanced panel

Let us repeat similar exercises as in the previous section, but with unbalanced characteristics for the simulated datasets. We introduce these characteristics through a matrix of observability indicators $I^s \in \mathbb{R}^{n \times T}$. The matrix gathers the indicator vectors for the n randomly chosen assets. We fix the maximal sample size $T = 546$ as in the empirical application. In the unbalanced setting, the excess returns $R^s_{i,t}$ of asset i at time t is:

$$R^s_{i,t} = a^s_i + b^s_i f_t + \varepsilon^s_{i,t}, \text{ if } I^s_{i,t} = 1, \text{ for } i = 1, ..., n, \text{ and } t = 1, ..., T, \tag{60}$$

where $I^s_{i,t}$ is the observability indicator of asset i at time t.

In Tables 9 and 10, we provide the operative cross-sectional and time-series sample sizes in the Monte-Carlo repetitions for trimming $\chi_{1,T} = 15$ and four different levels of trimming $\chi_{2,T}$. More precisely, in Table 9 we report the average number $\bar{n}^{\chi}$ of retained assets across simulations, as well as the minimum $\min(n^{\chi})$ and the maximum $\max(n^{\chi})$ across simulations. For the lowest level of trimming $\chi_{2,T} = T/12$, all assets are kept in all simulations, while for the level of trimming $\chi_{2,T} = T/60$ on average we keep about two thirds of the assets. In Table 10, we report the average across assets of the $\bar{T}_i$, that are the average time-series size T_i for asset i across simulations, as well as the min and the max of the $\bar{T}_i$. Since the distribution of T_i for an asset i is right-skewed, we also report the average across assets of the median T_i. For trimming level $\chi_{2,T} = T/60$, the average mean time-series size is about 180 months, while the average median time-series size is 140 months.

In Table 11, we display the results for estimators $\hat{\nu}$ and $\hat{\nu}_B$. The bias adjustment reduces substantially the bias for estimation of parameter ν. For trimming level $\chi_{2,T} = T/60$, the coverage of the confidence interval is close to the nominal size 95% for all considered cross-sectional sizes, while for $\chi_{2,T} = T/12$ the coverage deteriorates with increasing cross-sectional size. In comparison with Table 5, the bias and variance of estimator $\hat{\nu}_B$ are larger than the ones obtained in the balanced case with time-series size $T = 500$. However, for trimming level $\chi_{2,T} = T/60$, the results are similar to the ones with $T = 150$ in Table 5. In fact, this time-series size of the balanced panel reflects the operative sample sizes for that trimming level observed in Table 10. Similar comments apply for Table 12, where we report the results for the test of the hypothesis $\nu = 0$. For trimming level $\chi_{2,T} = T/60$, the size of the test is close to the nominal level 5% under DGP 1, and the the power is 100% under DGP 2.

In Tables 13 and 14, we display the results for the tests based on $\hat{Q}_a$ and $\hat{Q}_e$, respectively. For trimming level $\chi_{2,T} = T/120$, we observe an oversize, that increases with the cross-sectional dimension. We get a similar behaviour with more severe oversize with lower trimming levels (not reported). We expect these findings from the results in the previous section. Indeed, for trimming level $\chi_{2,T} = T/120$, the operative time-series sample size in Table 10 is around 200 months, and in Tables 7 and 8, for a balanced panel with $T = 150$, the statistics are oversized. For trimming level $\chi_{2,T} = T/240$ with operative size of about 350 months, the oversize of the statistics is moderate. Finally, the power of the statistics is very large also in the unbalanced case, and close to 100%.

Table 9: Operative cross-sectional sample size

trimming level	$\chi_{2,T} = \frac{T}{12}$				$\chi_{2,T} = \frac{T}{60}$			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
$\bar{n}^{\chi}$	1,000	3,000	6,000	9,000	660	2,000	4,000	6,000
$\min(n^{\chi})$	1,000	3,000	6,000	9,000	600	1,900	3,900	5,900
$\max(n^{\chi})$	1,000	3,000	6,000	9,000	700	2,100	4,100	6,100
trimming level	$\chi_{2,T} = \frac{T}{120}$				$\chi_{2,T} = \frac{T}{240}$			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
$\bar{n}^{\chi}$	400	1,250	2,400	3,600	140	430	850	1,250
$\min(n^{\chi})$	350	1,100	2,300	3,500	100	370	800	1,200
$\max(n^{\chi})$	440	1,300	2,500	3,650	170	470	900	1,300

Table 10: Operative time-series sample size

trimming level	$\chi_{2,T} = \frac{T}{12}$	$\chi_{2,T} = \frac{T}{60}$	$\chi_{2,T} = \frac{T}{120}$	$\chi_{2,T} = \frac{T}{240}$
$\text{mean}(\bar{T}_i)$	130	180	240	360
$\min(\bar{T}_i)$	110	160	210	350
$\max(\bar{T}_i)$	140	190	260	380
$\text{mean}(\text{median}(T_i))$	90	140	197	330

Table 11: Estimation of ν, unbalanced case

trimming level: $\chi_{2,T} = \frac{T}{12}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Bias($\hat{\nu}$)	-0.3059	-0.3119	-0.3047	-0.3021	-0.4211	-0.4324	-0.4202	-0.4201
Bias($\hat{\nu}_B$)	-0.0893	-0.0954	-0.0880	-0.0854	-0.1127	-0.1233	-0.1113	-0.1113
Var($\hat{\nu}_B$)	0.1207	0.0409	0.0214	0.0124	0.1222	0.0405	0.0218	0.0124
RMSE($\hat{\nu}_B$)	0.3586	0.2235	0.1706	0.1402	0.3671	0.2360	0.1848	0.1574
Coverage	0.9230	0.9010	0.8740	0.8750	0.9180	0.8880	0.8410	0.8320

trimming level: $\chi_{2,T} = \frac{T}{60}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Bias($\hat{\nu}$)	-0.1703	-0.1738	-0.1675	-0.1596	-0.2454	-0.2478	-0.0411	-0.2329
Bias($\hat{\nu}_B$)	-0.0349	-0.0381	-0.0318	-0.0238	-0.0453	-0.0474	-0.0411	-0.0325
Var($\hat{\nu}_B$)	0.1294	0.0436	0.0231	0.0141	0.1281	0.0438	0.0232	0.0144
RMSE($\hat{\nu}_B$)	0.3613	0.2122	0.1551	0.1212	0.3606	0.2145	0.1578	0.1241
Coverage	0.9360	0.9310	0.9240	0.9350	0.9430	0.9310	0.9200	0.9300

Table 12: Test of $\nu = 0$, unbalanced case

trimming level: $\chi_{2,T} = \frac{T}{12}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Rejection rate	0.0770	0.0990	0.1260	0.1250	1.0000	1.0000	1.0000	1.0000
trimming level: $\chi_{2,T} = \frac{T}{60}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Rejection rate	0.0640	0.0690	0.0760	0.0650	1.0000	1.0000	1.0000	1.0000

Table 13: Test of the null hypothesis $\mathcal{H}_0 : \ \beta_1 (\gamma) = 0$, unbalanced case

trimming level: $\chi_{2,T} = \frac{T}{120}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	0.1180	0.1710	0.2420	0.3030	0.6010	0.9410	0.9980	1.000
	DGP 3				**DGP 4**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	1.0000	1.0000	1.0000	1.0000	0.9990	1.0000	1.0000	1.0000
trimming level: $\chi_{2,T} = \frac{T}{240}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	0.0880	0.0860	0.1020	0.1310	0.5320	0.8730	0.9920	1.0000
	DGP 3				**DGP 4**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	1.0000	1.0000	1.0000	1.0000	0.9740	1.0000	1.0000	1.0000

Table 14: Test of the null hypothesis $\mathcal{H}_0: \ \beta_1(\gamma) = \beta_3(\gamma)\nu$, unbalanced case

trimming level: $\chi_{2,T} = \frac{T}{120}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	0.1130	0.1670	0.2370	0.3010	0.0940	0.2190	0.2590	0.3740
	DGP 3				**DGP 4**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	1.0000	1.0000	1.0000	1.0000	0.9990	1.0000	1.0000	1.0000

trimming level: $\chi_{2,T} = \frac{T}{240}$								
	DGP 1				**DGP 2**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	0.0800	0.0790	0.1000	0.1290	0.0790	0.0870	0.1080	0.1440
	DGP 3				**DGP 4**			
n	1,000	3,000	6,000	9,000	1,000	3,000	6,000	9,000
Size/Power	0.9990	1.0000	1.0000	1.0000	0.9690	1.0000	1.0000	1.0000

Appendix 7: Cost of equity

We can use the results in Section 3 for estimation and inference on the cost of equity in conditional factor models. We can estimate the time varying cost of equity $CE_{i,t} = r_{f,t} + b_{i,t}'\lambda_t$ of firm i with $\widehat{CE}_{i,t} = r_{f,t} + \hat{b}_{i,t}'\hat{\lambda}_t$, where $r_{f,t}$ is the risk-free rate. We have (see Appendix 7.1)

$$\begin{aligned}\sqrt{T}\left(\widehat{CE}_{i,t} - CE_{i,t}\right) &= \psi_{i,t}'E_2'\sqrt{T}\left(\hat{\beta}_i - \beta_i\right) \\ &\quad + \left(Z_{t-1}' \otimes b_{i,t}'\right)W_{p,K}\sqrt{T}vec\left[\hat{\Lambda}' - \Lambda'\right] + o_p(1), \qquad (61)\end{aligned}$$

where $\psi_{i,t} = \left(\lambda_t' \otimes Z_{t-1}', \lambda_t' \otimes Z_{i,t-1}'\right)'$. Standard results on OLS imply that estimator $\hat{\beta}_i$ is asymptotically normal, $\sqrt{T}\left(\hat{\beta}_i - \beta_i\right) \Rightarrow N\left(0, \tau_i Q_{x,i}^{-1} S_{ii} Q_{x,i}^{-1}\right)$, and independent of estimator $\hat{\Lambda}$. Then, from Proposition

9, we deduce that $\sqrt{T}\left(\widehat{CE}_{i,t} - CE_{i,t}\right) \Rightarrow N\left(0, \Sigma_{CE_{i,t}}\right)$, conditionally on Z_{t-1}, where

$$\Sigma_{CE_{i,t}} = \tau_i \psi_{i,t}' E_2' Q_{x,i}^{-1} S_{ii} Q_{x,i}^{-1} E_2 \psi_{i,t} + \left(Z_{t-1}' \otimes b_{i,t}'\right) W_{p,K} \Sigma_\Lambda W_{K,p} \left(Z_{t-1} \otimes b_{i,t}\right).$$

Figure 4 plots the path of the estimated annualized costs of equity for Ford Motor, Disney, Motorola and Sony. The cost of equity has risen tremendously during the recent subprime crisis.

A.7.1 Proof of Equation (61)

We have:

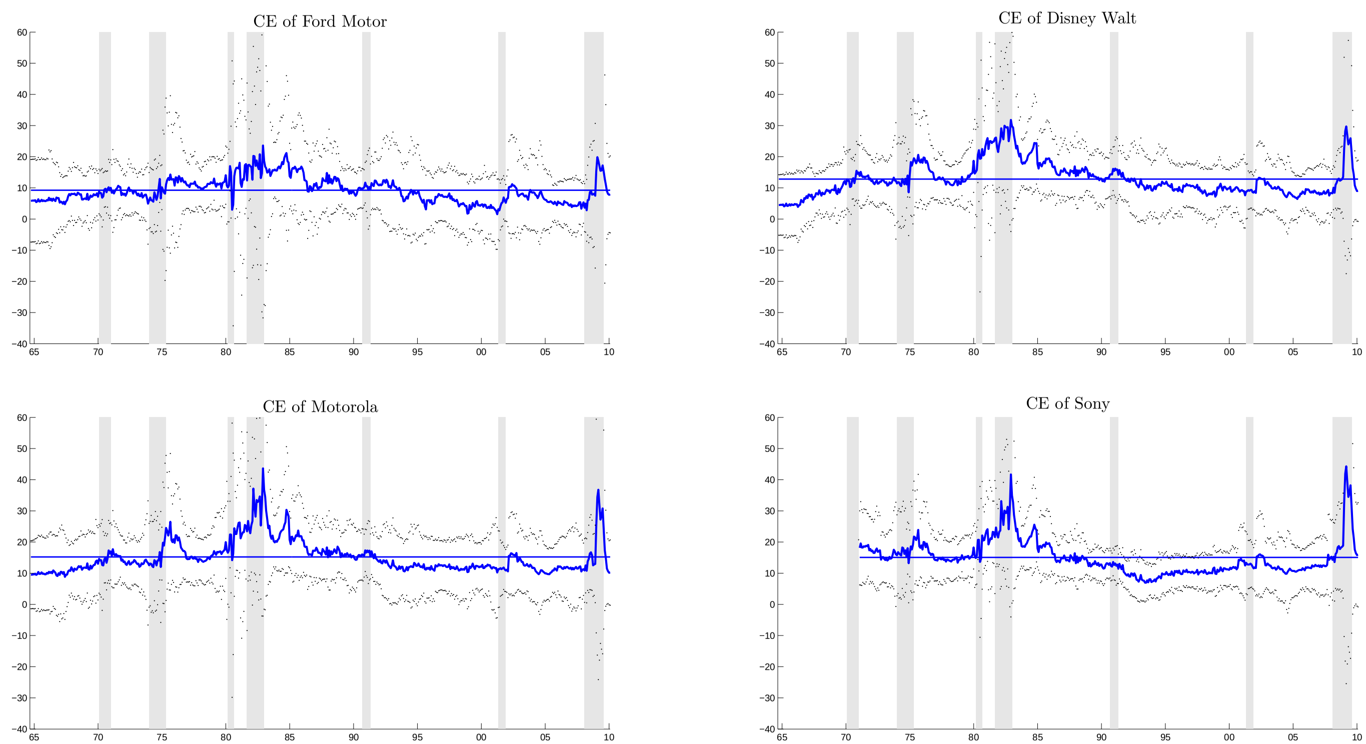
$$\hat{b}_{i,t}' \hat{\lambda}_t = tr\left[Z_{t-1} Z_{t-1}' \hat{B}_i' \hat{\Lambda}\right] + tr\left[Z_{t-1} Z_{i,t-1}' \hat{C}_i' \hat{\Lambda}\right] = \left(Z_{t-1}' \otimes Z_{t-1}'\right) vec\left[\hat{B}_i' \hat{\Lambda}\right] + \left(Z_{t-1}' \otimes Z_{i,t-1}'\right) vec\left[\hat{C}_i' \hat{\Lambda}\right].$$

Thus, we get:

$$\begin{aligned}
& \sqrt{T}\left(\widehat{CE}_{i,t} - CE_{i,t}\right) \\
= \; & \left(Z_{t-1}' \otimes Z_{t-1}'\right) \sqrt{T}\left(vec\left[\hat{B}_i' \hat{\Lambda}\right] - vec\left[B_i' \Lambda\right]\right) + \left(Z_{t-1}' \otimes Z_{i,t-1}'\right) \sqrt{T}\left(vec\left[\hat{C}_i' \hat{\Lambda}\right] - vec\left[C_i' \Lambda\right]\right) \\
= \; & \left(Z_{t-1}' \otimes Z_{t-1}'\right) \left[\left(\hat{\Lambda}' \otimes I_p\right) \sqrt{T} vec\left[\hat{B}_i' - B_i'\right] + \left(I_p \otimes B_i'\right) \sqrt{T} vec\left[\hat{\Lambda} - \Lambda\right]\right] \\
& + \left(Z_{t-1}' \otimes Z_{i,t-1}'\right) \left[\left(\hat{\Lambda}' \otimes I_q\right) \sqrt{T} vec\left[\hat{C}_i' - C_i'\right] + \left(I_p \otimes C_i'\right) \sqrt{T} vec\left[\hat{\Lambda} - \Lambda\right]\right].
\end{aligned}$$

By using that $\hat{\Lambda} = \Lambda + o_p(1)$ and $vec\left[\hat{\Lambda} - \Lambda\right] = W_{p,K} vec\left[\hat{\Lambda}' - \Lambda'\right]$, Equation (61) follows.

Figure 4: Path of estimated annualized costs of equity



The figure plots the path of estimated annualized costs of equity for Ford Motor, Disney Walt, Motorola and Sony and their pointwise confidence intervals at 95% probability level. We also report the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determinated by the National Bureau of Economic Research (NBER).

Appendix 8: Robustness checks

In this section, we perform several checks to evaluate the robustness of the empirical results reported in the paper. In particular, we estimate the paths of the time-varying risk premia and we compute the test statistics by:

a. Assuming several asset pricing models as baseline specification;

b. Using several sets of asset-specific instruments $Z_{i,t-1}$;

c. Using several sets of common instruments Z_{t-1};

d. Assuming that the time-varying betas $b_{i,t}$ depend only on the asset-specific instruments.

In Table 15, we provide the details of the conditional specifications for the four exercises. We use the following abbreviations. For common instruments, we denote by ts_t the term spread, ds_t the default spread, and $divY_t$ the dividend yield. The dividend yield is provided by CRSP. For asset-specific instruments, we denote by $mc_{i,t}$ the market capitalization, $bm_{i,t}$ the book-to-market, and $ind_{i,t}$ the return of the corresponding industry portfolio. For each exercise, when not explicitly indicated in Table 15, the specification is the four-factor model, the vector of common instruments is $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$ and the asset-specific instrument is the scalar $Z_{i,t-1} = bm_{i,t-1}$. Table 15 reports the operative trimmed population of individual stocks and the number of regressors in the first-pass time series regression for each exercise that we implement. Indeed, the population of individual stocks changes depending on the asset pricing model (Exercise a) as an effect of the trimming conditions: the number of assets decreases as the number K of factors increases. Moreover, by using the four-factor model as baseline and modifying the sets of instruments, the number of assets decreases as the number of regressors in the first pass increases (see Exercise c) .

We first present conditional estimates of risk premia by using several asset pricing models as baseline (Exercise a). Panel A of Figure 5 compares the estimated time-varying paths of market risk premia when we assume the four-factor model (shown in Section 4) and the CAPM. Panel B compares the estimates $\hat{\lambda}_{m,t}$ for the four-factor model and the Fama-French model. The paths look very similar. The discrepancy between the estimates of the CAPM and the four-factor model is explained by the three factors (size, value and momentum factor) that we introduce in the four-factor model. Figure 6 plots the estimated time-varying paths of risk premia for the size and value factors computed on the four-factor model and on the Fama-French model. The risk premium for the size factor is very similar for the two models. The value risk premium for the Fama-French model takes slightly smaller values than that for the four-factor model and it exhibits a counter-cyclical path. Overall, the conditional estimates of the risk premia are stable with respect to the asset pricing model that is assumed for the excess returns.

Figures 7 and 8 plot the estimates of the risk premia by adopting several sets of asset-specific instruments $Z_{i,t-1}$ (Exercise b). We do not modify the set of common instruments Z_{t-1} compared to Section 4 of the paper. In Figure 7, we get the estimates by setting the scalar $Z_{i,t-1}$ equal to the market capitalization of firm i. In Figure 8, we set $Z_{i,t-1}$ equal to the monthly returns of the industry portfolio for the industry asset i belongs to. We use the 48 Fama-French industry portfolios. The risk premia paths look very similar to the results in Section 4. The results for the tests of the asset pricing restrictions for the conditional specifications in Exercise b are reported in Table 16, upper panel. The test statistics reject the null hypotheses at 5% level.

The time-varying paths of the risk premia showed in Figures 9 and 10 are computed by modifying the set of common instruments $Z_{t-1} = \left[1,\ Z_{t-1}^{*\prime}\right]'$ (Exercise c). In Figure 9, Z_t^* is a bivariate vector that includes the default spread and the dividend yield. The paths of the risk premia for market, value and momentum factors look similar to the results in Section 4. However, the risk premium for the size factor features a very stable pattern that does not correspond to the unconditional estimate. In Figure 10, vector Z_t^* includes the term spread, the default spread, and the dividend yield. The paths of the risk premia look similar to the results in Section 4. Introducing the dividend yield increases the discrepancy between the unconditional estimates and the average over time of conditional estimates for the size and momentum factors w.r.t. the results shown in Figure 1. On the contrary, this discrepancy is smaller for the value premium. Moreover, the risk premium of the momentum factor takes larger values than that in Figure 1. We also notice that including the dividend yield among the common instruments decreases the number of stocks after trimming. The test statistics reject the null hypothesis at 5% level (see Table 16), middle panel.

Finally, we consider conditional specifications in which the time-varying betas are linear functions of asset specific instruments $Z_{i,t-1}$ only (Exercise d). The risk premia are modelled via common instruments $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$ as usual. In Figure 11, $Z_{i,t-1}$ is a bivariate vector that includes the constant and the book-to-market equity of firm i. In Figure 12, vector $Z_{i,t-1}$ includes the constant and the return of the industry portfolio as asset-specific instrument. The paths of the risk premia for the four factors in Figure 11 look more volatile w.r.t. the paths in Figure 1. The risk premia for market, size and value factors in Figure 12 look similar to the results in Section 4. The risk premium for the momentum factor features a less stable pattern, albeit its confidence intervals look similar to that in Figure 1. In Table 16, lower panel, the test statistic does not reject the asset pricing restrinction $\mathcal{H}_0 : \beta_1(\gamma) = \beta_3(\gamma)\nu$ for the conditional specification with time-varying betas depending on book-to-market equity.

Table 15: Operative cross-sectional sample size (n^χ), number of factors (K) and instruments (q and p) and first-pass regressors (d) in the four exercises of robustness checks

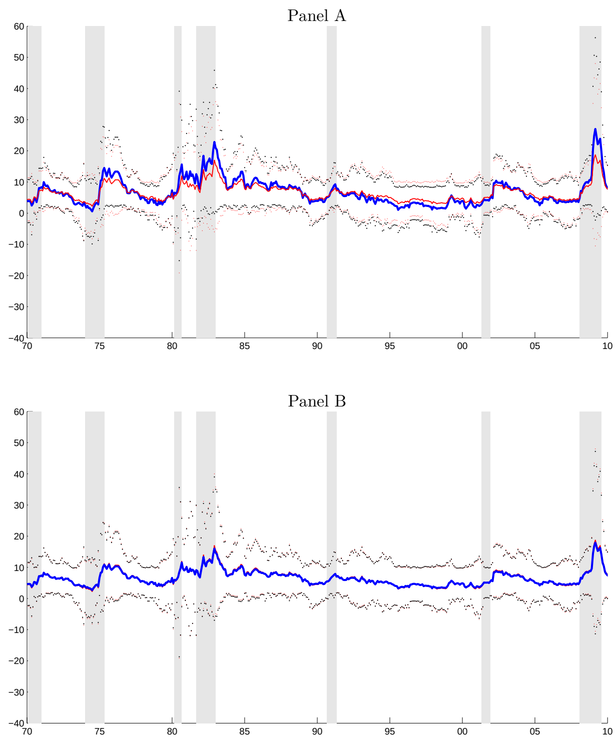
	n^χ	K	p	q	d		n^χ	K	p	q	d
Exercise a.						Exercise c.					
CAPM	5,225	1	3	1	13	$Z_{t-1} = [1, ds_{t-1}, divY_{t-1}]'$	1,107	4	3	1	25
Fama-French model	4,545	3	3	1	21	$Z_{t-1} = [1, ds_{t-1}, ts_{t-1}, divY_{t-1}]'$	667	4	4	1	34
Exercise b.						Exercise d.					
$Z_{i,t-1} = mc_{i,t-1}$	3,835	4	3	1	25	$Z_{i,t-1} = [1, bm_{i,t-1}]'$	6,208	4	3	2	8
$Z_{i,t-1} = ind_{i,t-1}$	4,748	4	3	1	25	$Z_{i,t-1} = [1, ind_{i,t-1}]'$	6,430	4	3	2	8

Table 16: Test results for asset pricing restrictions

	Test of the null hypothesis $\mathcal{H}_0 :\ \beta_1(\gamma) = \beta_3(\gamma)\nu$		Test of the null hypothesis $\mathcal{H}_0 :\ \beta_1(\gamma) = 0$	
Exercise b.				
	$Z_{i,t-1} = mc_{i,t-1}$	$Z_{i,t-1} = ind_{i,t-1}$	$Z_{i,t-1} = mc_{i,t-1}$	$Z_{i,t-1} = ind_{i,t-1}$
	$(n^{\chi} = 3,835)$	$(n^{\chi} = 4,748)$	$(n^{\chi} = 3,835)$	$(n^{\chi} = 4,748)$
Test statistic	8.0493	5.7601	8.7126	6.4357
p-value	0.0000	0.0000	0.0000	0.0000
Exercise c.				
	$Z_{t-1} = [1, ds_{t-1}, divY_{t-1}]'$	$Z_{t-1} = [1, ds_{t-1}, ts_{t-1}, divY_{t-1}]'$	$Z_{t-1} = [1, ds_{t-1}, divY_{t-1}]'$	$Z_{t-1} = [1, ds_{t-1}, ts_{t-1}, divY_{t-1}]'$
	$(n^{\chi} = 1,107)$	$(n^{\chi} = 667)$	$(n^{\chi} = 1,107)$	$(n^{\chi} = 667)$
Test statistic	2.7954	2.2463	3.6335	2.8434
p-value	0.0026	0.0123	0.0000	0.0022
Exercise d.				
	$Z_{i,t-1} = [1, bm_{i,t-1}]'$	$Z_{i,t-1} = [1, ind_{i,t-1}]'$	$Z_{i,t-1} = [1, bm_{i,t-1}]'$	$Z_{i,t-1} = [1, ind_{i,t-1}]'$
	$(n^{\chi} = 6,208)$	$(n^{\chi} = 6,430)$	$(n^{\chi} = 6,208)$	$(n^{\chi} = 6,430)$
Test statistic	1.3394	5.4079	3.5772	8.3430
p-value	0.0902	0.0000	0.0000	0.0000

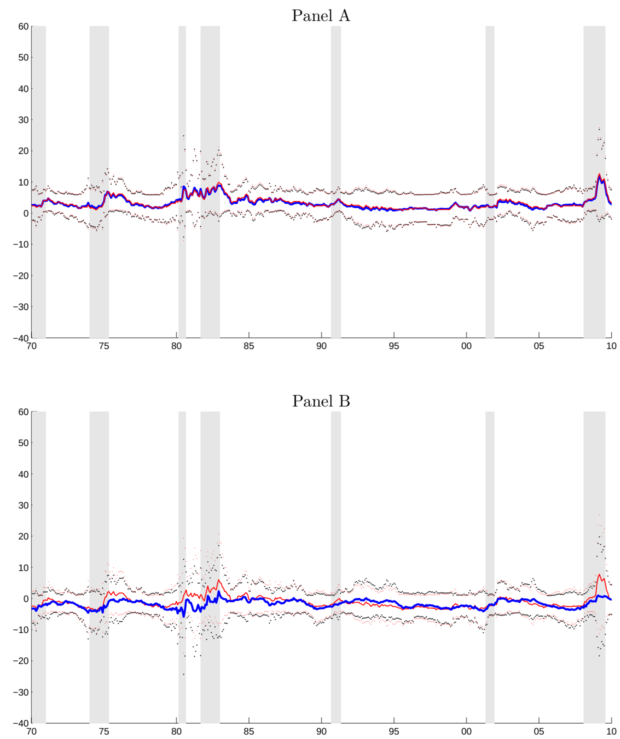
We compute the statistics $\tilde{\Sigma}_{\xi}^{-1/2}\hat{\xi}_{nT}$ based on $\hat{Q}_e$ and $\hat{Q}_a$ defined in Proposition 5 for n^{χ} individual stocks to test the null hypotheses $\mathcal{H}_0 :\ \beta_1(\gamma) = \beta_3(\gamma)\nu$ and $\mathcal{H}_0 :\ \beta_1(\gamma) = 0$. The table reports the statistics and their p-values when we use several sets of asset-specific instruments $Z_{i,t-1}$ (Exercise b) and common instruments Z_{t-1} (Exercise c), and when time-varying betas are functions of the asset-specific instruments only (Exercise d).

Figure 5: Path of estimated annualized risk premia for the market factor



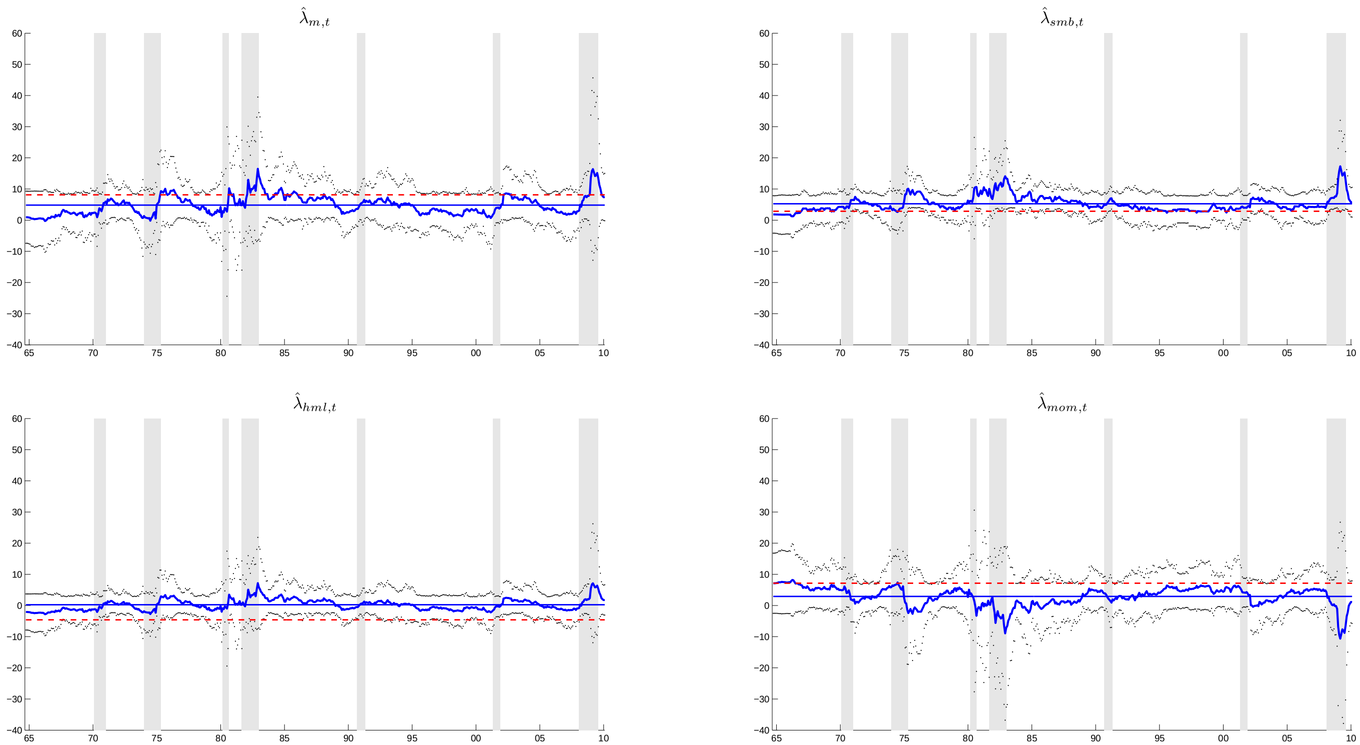
Panel A plots the paths of estimated annualized market risk premia $\hat{\lambda}_{m,t}$ computed by using the four-factor model (thin red line) and the CAPM (thick blue line). Panel B plot the paths of market risk premia $\hat{\lambda}_{m,t}$ estimated by assuming the four-factor model (thin red line) and the Fama-French model (thick blue line). The pointwise confidence intervals at 95% level are also displayed. The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 6: Path of estimated annualized risk premia for the size and value factors



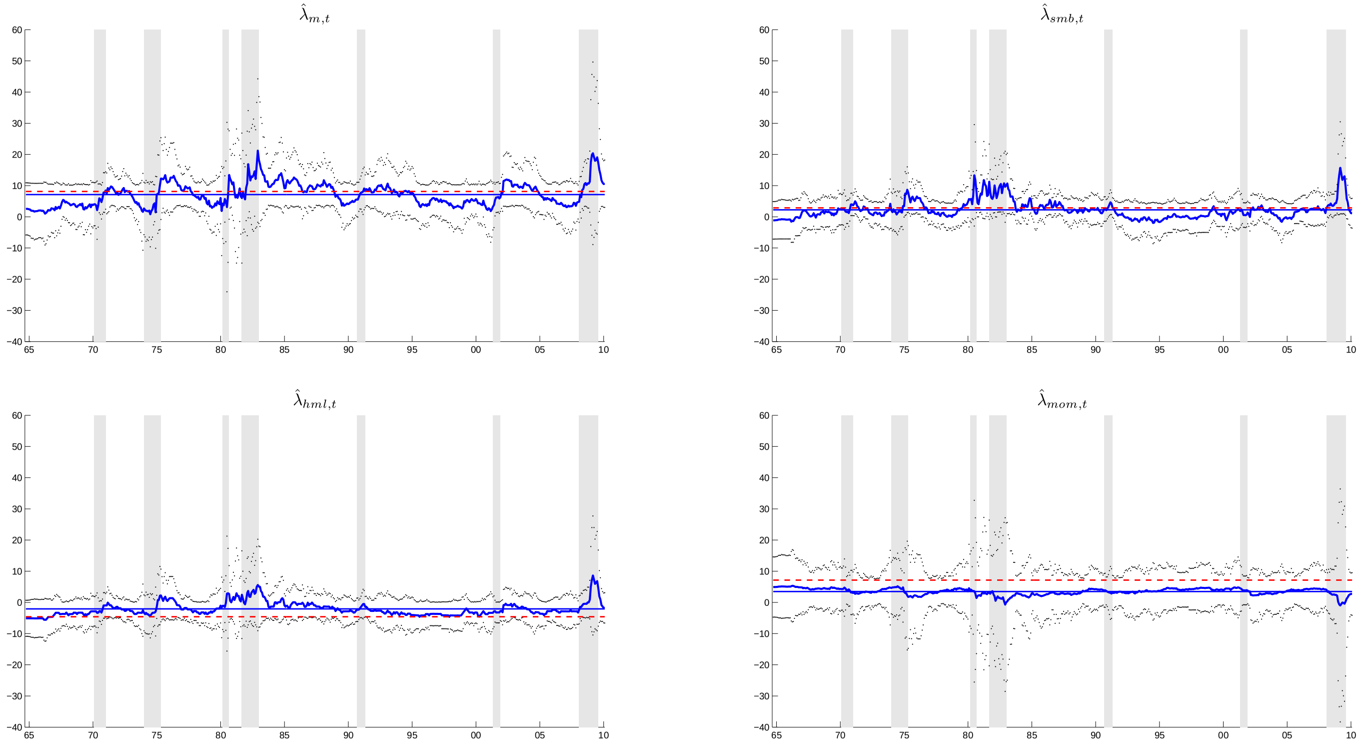
The figure plots the paths of estimated annualized risk premia $\hat{\lambda}_{smb,t}$ (Panel A) and $\hat{\lambda}_{hml,t}$ (Panel B) computed by using the four-factor model (thin red line) and the Fama-French model (thick blue line). The pointwise confidence intervals at 95% level are also displayed. The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 7: Path of estimated annualized risk premia computed using $Z_{i,t-1} = mc_{i,t-1}$



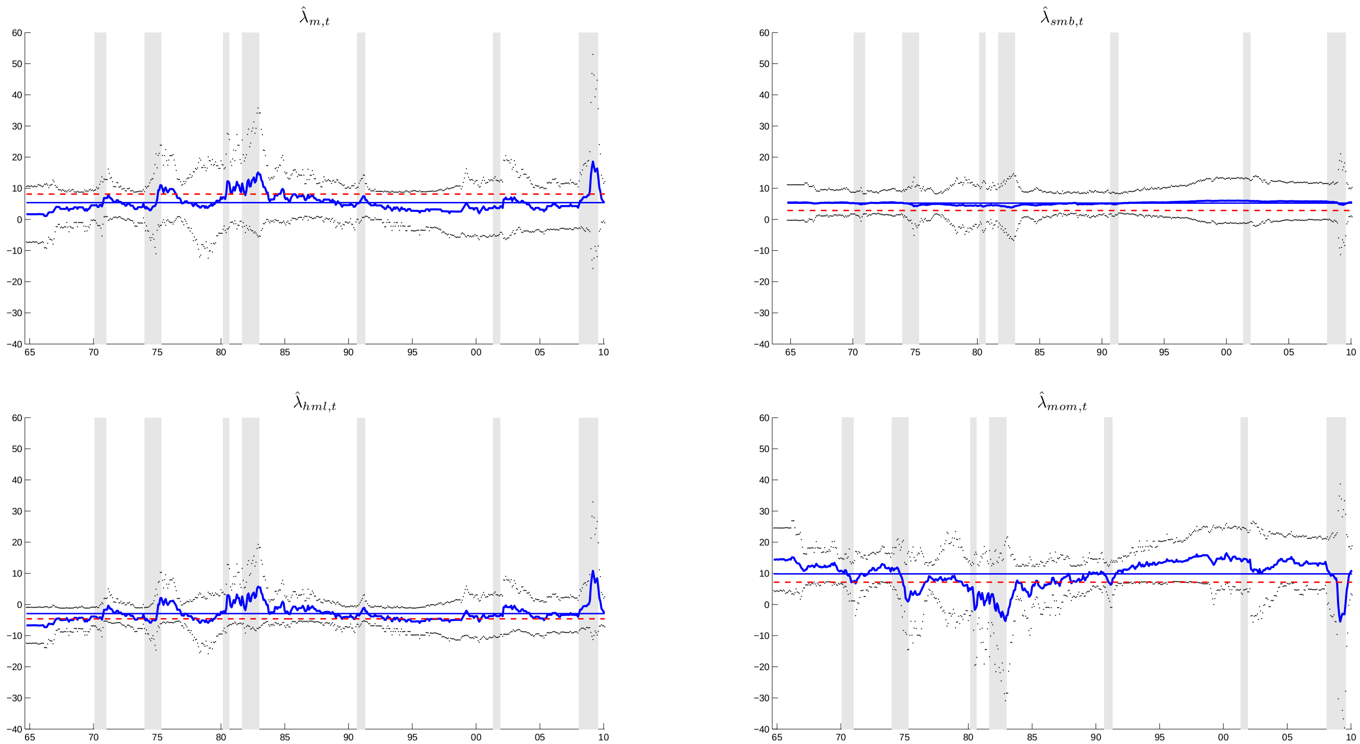
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when market capitalization is used as asset-specific instrument. The vector of common instruments is $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$. We also display the unconditional estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 3,835$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 8: Path of estimated annualized risk premia computed using $Z_{i,t-1} = ind_{i,t-1}$



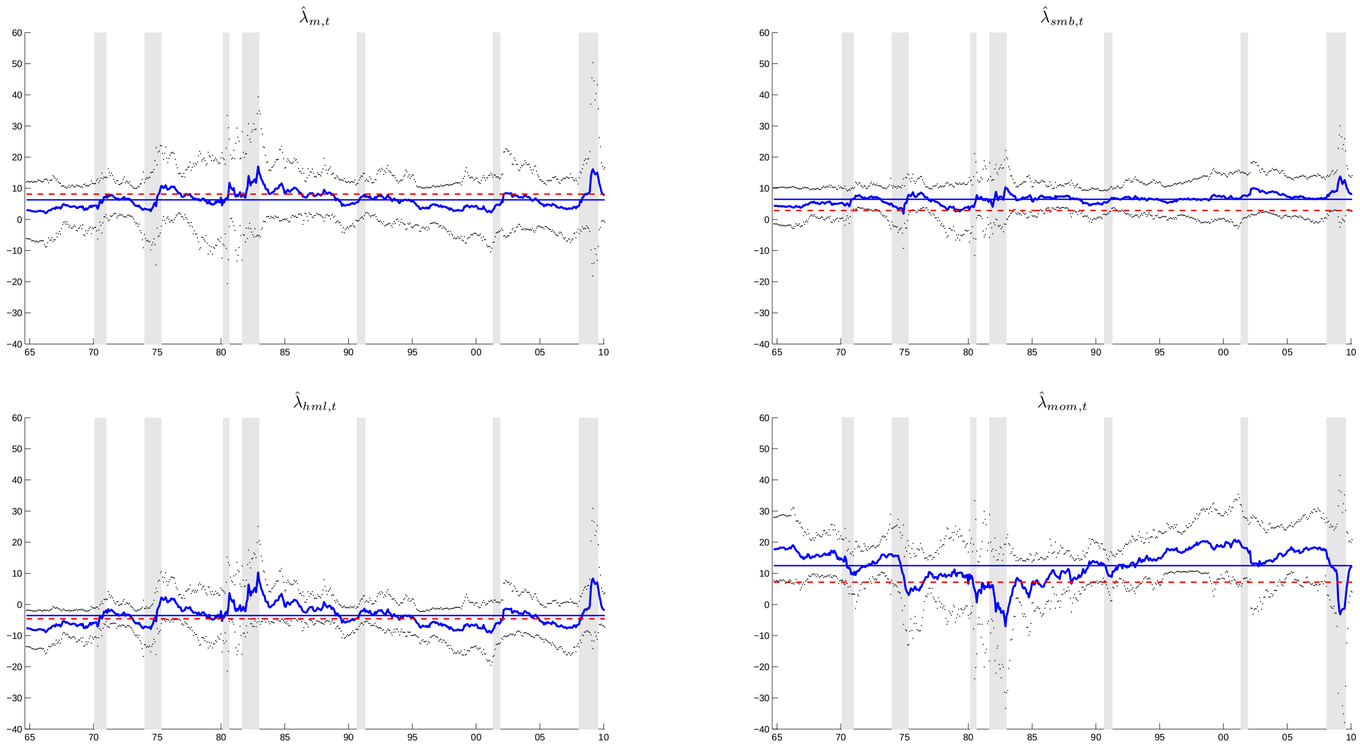
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when the returns of industry portfolios are used as asset-specific instrument. The vector of common instruments is $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$. We also display the unconditional estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 4,748$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 9: Path of estimated annualized risk premia computed using $Z_{t-1} = [1,\ ds_{t-1},\ divY_{t-1}]'$



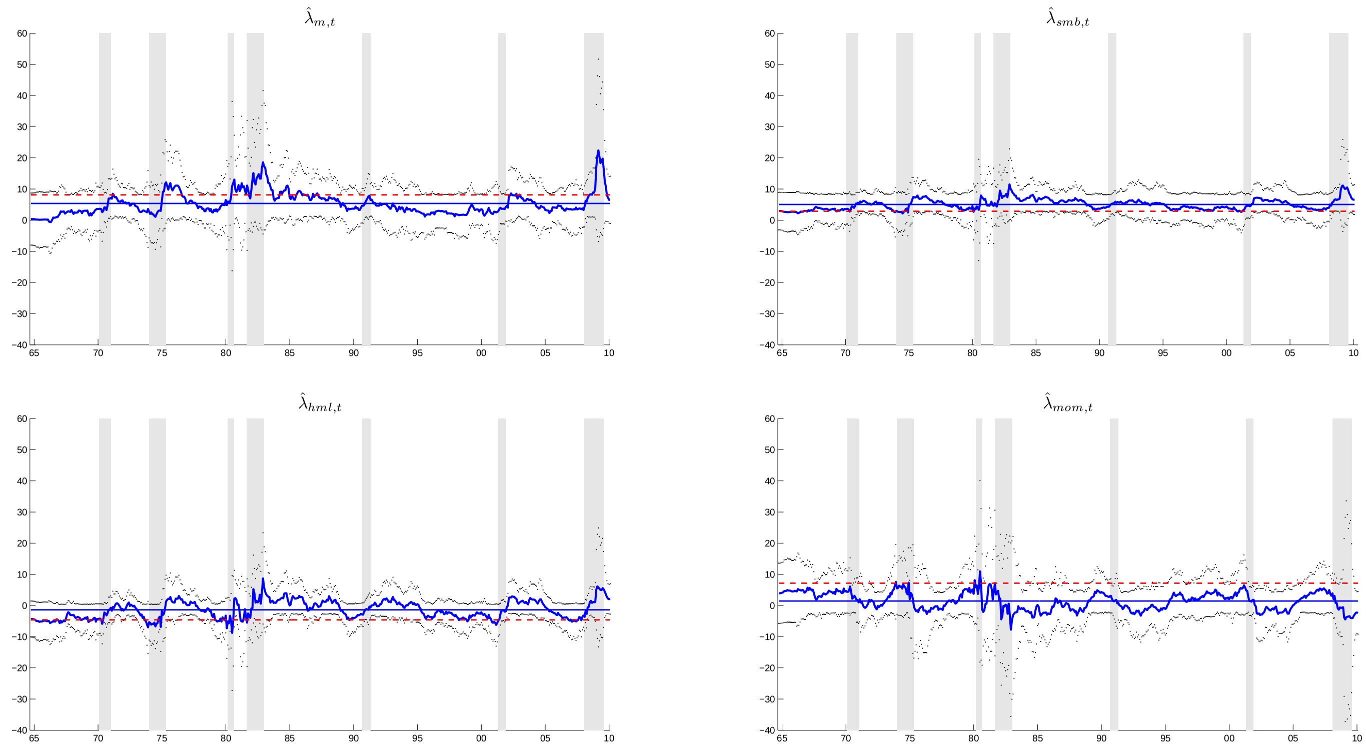
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when default spread and dividend yield are used as common instruments. The stock specific instrument is book-to-market equity. We also display the unconditional estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 1,107$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 10: Path of estimated annualized risk premia computed using $Z_{t-1} = [1,\ ds_{t-1},\ ts_{t-1},\ divY_{t-1}]'$



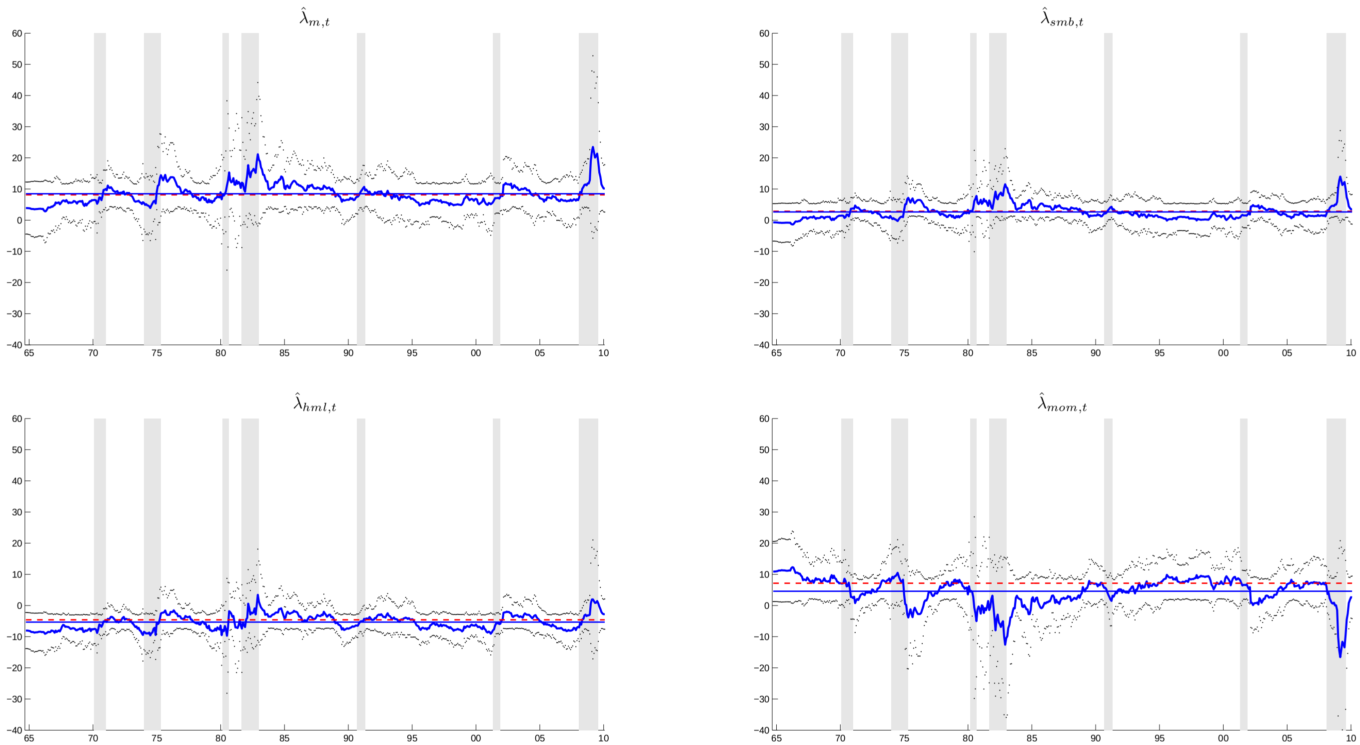
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when default spread, term spread and dividend yield are used as common instruments. The stock specific instrument is book-to-market equity. We also display the unconditional estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 667$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 11: Path of estimated annualized risk premia with time-varying betas modelled via $Z_{i,t-1} = [1,\ bm_{i,t-1}]'$



The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when time-varying betas are linear functions of the book-to-market instrument only. The risk premia vector involves the common instruments $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$. We also report the unconditional estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks ($n = 9,936$ and $n^{\chi} = 6,208$) as base assets. The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).

Figure 12: Path of estimated annualized risk premia with time-varying betas modelled via $Z_{i,t-1} = [1,\ ind_{i,t-1}]'$



The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when time-varying betas are linear functions of industry portfolio returns. The risk premia vector involves the common instruments $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$. We also report the unconditional estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks ($n = 9,936$ and $n^{\chi} = 6,430$) as base assets. The vertical shaded areas denote recessions determined by the National Bureau of Economic Reasearch (NBER).