SUPPLEMENTARY MATERIALS

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

Patrick Gagliardini, Elisa Ossola and Olivier Scaillet

These supplementary materials provide the derivation of Equations (9)-(12) (Appendix 3), the proofs of technical lemmas used in the paper (Appendix 4), the link of our no-arbitrage pricing restrictions with Chamberlain and Rothschild (1983) results (Appendix 5), the check that the high-level assumptions in the paper hold under block-dependence (Appendix 6), and the results of Monte-Carlo experiments that investigate the finite-sample properties of the estimators and test statistics (Appendix 7). We also present empirical results with long-only factors (Appendix 8), figures of estimated paths of $\hat{\nu}_t$ for the four-factor model estimated by using individual stocks and the 25FF portfolios (Appendix 9). We also provide additional figures for the 25FF portfolios (Appendix 10) and the industry portfolios (Appendix 11), the value-weighted estimates of risk premia (Appendix 12), as well as an empirical analysis of estimated time-varying betas (Appendix 13). We investigate the effects of model misspecification on risk premia estimation and give estimates of the pseudo-true values (Appendix 14). We also present the results of a preliminary analysis of idiosyncratic risk (Appendix 15), and provide some robustness checks for the empirical analysis (Appendix 16). Finally, we derive inference for the cost of equity and include some empirical results for Ford Motor, Disney Walt, Motorola and Sony (Appendix 17).

Appendix 3 Derivation of Equations (9)-(12)

A.3.1 Derivation of Equations (9) and (10)

From Equation (8) and by using $vec\,[ABC] = \left[C' \otimes A\right] vec\,[B]$ (MN Theorem 2, p. 35), we get $Z'_{t-1}B'_i f_t = vec\left[Z'_{t-1}B'_i f_t\right] = \left[f'_t \otimes Z'_{t-1}\right] vec\left[B'_i\right]$, and $Z'_{i,t-1}C'_i f_t = \left[f'_t \otimes Z'_{i,t-1}\right] vec\left[C'_i\right]$, which gives $Z'_{t-1}B'_i f_t + Z'_{i,t-1}C'_i f_t = x'_{2,i,t}\beta_{2,i}$.

Let us now consider the first two terms in the RHS of Equation (8).

a) By definition of matrix X_t in Section 3.1, we have

$$\begin{aligned} Z'_{t-1}B'_i\left(\Lambda - F\right)Z_{t-1} &= \frac{1}{2}Z'_{t-1}\left[B'_i\left(\Lambda - F\right) + \left(\Lambda - F\right)'B_i\right]Z_{t-1} \\ &= \frac{1}{2}vech\left[X_t\right]'vech\left[B'_i\left(\Lambda - F\right) + \left(\Lambda - F\right)'B_i\right]. \end{aligned}$$

By using the Moore-Penrose inverse of the duplication matrix D_p, we get

$$vech\left[B'_i\left(\Lambda - F\right) + \left(\Lambda - F\right)'B_i\right] = D_p^+\left[vec\left[B'_i\left(\Lambda - F\right)\right] + vec\left[\left(\Lambda - F\right)'B_i\right]\right].$$

Finally, by the properties of the vec operator and the commutation matrix W_p, and the definition of matrix N_p, we obtain

$$\begin{aligned} \frac{1}{2}D_p^+\left[vec\left[B'_i\left(\Lambda - F\right)\right] + vec\left[\left(\Lambda - F\right)'B_i\right]\right] &= \frac{1}{2}D_p^+(I_{p^2} + W_p)vec\left[B'_i\left(\Lambda - F\right)\right] \\ &= N_p\left[\left(\Lambda - F\right)' \otimes I_p\right]vec\left[B'_i\right]. \end{aligned}$$

b) By the properties of the tr and vec operators, we have

$$\begin{aligned} Z'_{i,t-1}C'_i\left(\Lambda - F\right)Z_{t-1} &= tr\left[Z_{t-1}Z'_{i,t-1}C'_i\left(\Lambda - F\right)\right] = vec\left[Z_{i,t-1}Z'_{t-1}\right]'vec\left[C'_i\left(\Lambda - F\right)\right] \\ &= \left(Z_{t-1} \otimes Z_{i,t-1}\right)'\left[\left(\Lambda - F\right)' \otimes I_q\right]vec\left[C'_i\right]. \end{aligned}$$

By combining a) and b), we get $Z'_{t-1}B'_i\left(\Lambda - F\right)Z_{t-1} + Z'_{i,t-1}C'_i\left(\Lambda - F\right)Z_{t-1} = x'_{1,i,t}\beta_{1,i}$ and $\beta_{1,i} = \left(\left(N_p\left[\left(\Lambda - F\right)' \otimes I_p\right]vec\left[B'_i\right]\right)', \left(\left[\left(\Lambda - F\right)' \otimes I_q\right]vec\left[C'_i\right]\right)'\right)'$.

A.3.2 Derivation of Equation (11)

We use $\beta_{1,i} = \left(\left(\frac{1}{2}D_p^+\left[vec\left[B'_i\left(\Lambda - F\right)\right] + vec\left[\left(\Lambda - F\right)'B_i\right]\right]\right)', \left(vec\left[C'_i\left(\Lambda - F\right)\right]\right)'\right)'$ from Section A.3.1. a) From the properties of the vec operator and the commutation matrix W_p, we get

$$vec\left[B'_i\left(\Lambda - F\right)\right] + vec\left[\left(\Lambda - F\right)'B_i\right] = (W_p + I_{p^2})vec\left[\left(\Lambda - F\right)'B_i\right] = (W_p + I_{p^2})\left(B'_i \otimes I_p\right)vec\left[\Lambda' - F'\right].$$

From $\nu = vec\left[\Lambda' - F'\right]$ we obtain

$$\frac{1}{2}D_p^+\left[vec\left[B'_i\left(\Lambda - F\right)\right] + vec\left[\left(\Lambda - F\right)'B_i\right]\right] = \frac{1}{2}D_p^+(I_{p^2} + W_p)\left(B'_i \otimes I_p\right)\nu = N_p\left(B'_i \otimes I_p\right)\nu.$$

b) From the properties of the vec operator and the commutation matrix $W_{p,q}$, we get

$$vec\left[C'_i\left(\Lambda - F\right)\right] = W_{p,q}vec[(\Lambda - F)'C_i] = W_{p,q}\left(C'_i \otimes I_p\right)\nu.$$

A.3.3 Derivation of Equation (12)

We use $vec\left[\beta'_{3,i}\right] = \left(vec\left[\{N_p\left(B'_i \otimes I_p\right)\}'\right]', vec\left[\{W_{p,q}\left(C'_i \otimes I_p\right)\}'\right]'\right)'$ from Equation (11).

a) By MN Theorem 2 p. 35 and Exercise 1 p. 56, and by writing $I_{pK} = I_K \otimes I_p$, we obtain

$$\begin{aligned} vec\left[N_p\left(B'_i \otimes I_p\right)\right] &= \left(I_{pK} \otimes N_p\right) vec\left[B'_i \otimes I_p\right] \\ &= \left(I_{pK} \otimes N_p\right)\{I_K \otimes [\left(W_p \otimes I_p\right)\left(I_p \otimes vec\left[I_p\right]\right)]\} vec\left[B'_i\right] \\ &= \{I_K \otimes [\left(I_p \otimes N_p\right)\left(W_p \otimes I_p\right)\left(I_p \otimes vec\left[I_p\right]\right)]\} vec\left[B'_i\right]. \end{aligned}$$

Moreover, $vec\left[\{N_p\left(B'_i \otimes I_p\right)\}'\right] = W_{p(p+1)/2,pK} vec\left[N_p\left(B'_i \otimes I_p\right)\right]$.

b) Similarly, $vec\left[W_{p,q}\left(C'_i \otimes I_p\right)\right] = \{I_K \otimes [\left(I_p \otimes W_{p,q}\right)\left(W_{p,q} \otimes I_p\right)\left(I_q \otimes vec\left[I_p\right]\right)]\} vec\left[C'_i\right]$ and $vec\left[\{W_{p,q}\left(C'_i \otimes I_p\right)\}'\right] = W_{pq,pK} vec\left[W_{p,q}\left(C'_i \otimes I_p\right)\right]$.

By combining a) and b) the conclusion follows.

Appendix 4 Proofs of statements and technical lemmas

A.4.1 Proof of Lemma 2

Let vector $(z_1, ..., z_n)$ be such that $\sum_i z_i^2 = 1$. From Equation (25), we have:

$$\sum_i \sum_j z_i [\Sigma_{\tilde{\varepsilon},1,n}]_{i,j} z_j = \sum_k \sum_l \sum_i \sum_j z^*_{k,i} z^*_{l,j} Cov(\varepsilon[G_k^{-1}(\gamma_i)], \varepsilon[G_l^{-1}(\gamma_j)] | \mathcal{F}_0), \tag{35}$$

where $z^*_{k,i} = w_k[G_k^{-1}(\gamma_i)] z_i$. Now, by the Cauchy-Schwarz inequality, we have:

$$\begin{aligned} &\sum_i \sum_j z^*_{k,i} z^*_{l,j} Cov(\varepsilon[G_k^{-1}(\gamma_i)], \varepsilon[G_l^{-1}(\gamma_j)] | \mathcal{F}_0) = Cov\left(\sum_i z^*_{k,i} \varepsilon[G_k^{-1}(\gamma_i)], \sum_j z^*_{l,j} \varepsilon[G_l^{-1}(\gamma_j)] | \mathcal{F}_0\right) \\ &\leq V\left(\sum_i z^*_{k,i} \varepsilon[G_k^{-1}(\gamma_i)] | \mathcal{F}_0\right)^{1/2} V\left(\sum_j z^*_{l,j} \varepsilon[G_l^{-1}(\gamma_j)] | \mathcal{F}_0\right)^{1/2} \\ &= \left(\sum_i \sum_j z^*_{k,i} z^*_{k,j} Cov(\varepsilon[G_k^{-1}(\gamma_i)], \varepsilon[G_k^{-1}(\gamma_j)] | \mathcal{F}_0)\right)^{1/2} \left(\sum_i \sum_j z^*_{l,i} z^*_{l,j} Cov(\varepsilon[G_l^{-1}(\gamma_i)], \varepsilon[G_l^{-1}(\gamma_j)] | \mathcal{F}_0)\right)^{1/2}. \end{aligned}$$

Moreover:

$$\begin{aligned}\sum_i \sum_j z^*_{k,i} z^*_{k,j} Cov(\varepsilon[G_k^{-1}(\gamma_i)], \varepsilon[G_k^{-1}(\gamma_j)]|\mathcal{F}_0) &\leq (\sum_i (z^*_{k,i})^2) eig_{max}(\Sigma_{\varepsilon,1,n}(G_k)) \\ &\leq \bar{w}_k^2 eig_{max}(\Sigma_{\varepsilon,1,n}(G_k)).\end{aligned}$$

Thus, for any vector $(z_1, ..., z_n)$ such that $\sum_i z_i^2 = 1$ we have:

$$\sum_i \sum_j z_i [\Sigma_{\tilde{\varepsilon},1,n}]_{i,j} z_j \leq \sum_k \sum_l \bar{w}_k \bar{w}_l eig_{max}(\Sigma_{\varepsilon,1,n}(G_k))^{1/2} eig_{max}(\Sigma_{\varepsilon,1,n}(G_l))^{1/2}.$$

Since the largest eigenvalue of a symmetric matrix is equal to the sup of the associated quadratic form w.r.t. vectors with unit length, the conclusion follows.

A.4.2 Proof of Lemma 3 (iii)

We have $\hat{w}_i - w_i = \mathbf{1}_i^\chi \left((diag[\hat{v}_i])^{-1} - (diag[v_i])^{-1}\right) + (\mathbf{1}_i^\chi - 1)(diag[v_i]^{-1})$ and $(diag[\hat{v}_i])^{-1} - (diag[v_i])^{-1}$ $= -(diag[\hat{v}_i])^{-1} diag[\hat{v}_i - v_i](diag[v_i])^{-1}$. Since $\|(diag[v_i])^{-1}\|$ 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 7. To prove that the first term is $o_p(1)$, it is sufficient to show:

$$\sup_i \mathbf{1}_i^\chi \|\hat{v}_i - v_i\| = o_p(1). \tag{36}$$

We use Equation (30). Since $\hat{\nu}_1 - \nu = O_p(T^{-c})$, for some $c > 0$ (by repeating the proof of Proposition 3 with known weights equal to 1), $\mathbf{1}_i^\chi \|\hat{Q}_{x,i}^{-1}\| \leq C\chi_{1,T}^2$, $\mathbf{1}_i^\chi \tau_{i,T} \leq \chi_{2,T}$, $\|S_{ii}\| \leq M$, and by using Assumption B.5, the uniform bound in (36) follows if we prove:

$$\sup_i \mathbf{1}_i^\chi \|\hat{S}_{ii} - S_{ii}\| = O_p(T^{-c}), \tag{37}$$

$$\sup_i \mathbf{1}_i^\chi \|\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1}\| = O_p(T^{-c}), \tag{38}$$

$$\sup_i \mathbf{1}_i^\chi |\tau_{i,T} - \tau_i| = O_p(T^{-c}), \tag{39}$$

for some $c > 0$. To prove the uniform bound (37), we use Equation (32). As in the proof of Lemma 3 (i), we have $\sup_i T^{-1/2}\|Y_{i,T}\| = O_{p,log}(T^{-\eta/2})$ from Assumption B.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 B.1 e) and f), respectively. Moreover, $\|\hat{Q}_{x,i}^{(4)}\| \leq M$, $\mathbf{1}_i^{\chi}\|\hat{Q}_{x,i}^{-1}\| \leq C\chi_{1,T}^2$ and $\mathbf{1}_i^{\chi}\tau_{i,T} \leq \chi_{2,T}$. Thus, from Assumption B.5, bound (37) follows. To prove (38), we use Equation (33) where $W_{i,T}$ is such that $\sup_i \|W_{i,T}\| = O_{p,log}(T^{-\eta/2})$ from Assumption B.1 b). Finally, (39) 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 B.1 d).

A.4.3 Proof of Lemma 4

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

$$
\begin{aligned}
Ab = vec\,[Ab] &= \left(b' \otimes A\right) vec\,[I_n] \\
&= vec\left[\left(b' \otimes A\right) vec\,[I_n]\right] \\
&= \left(vec\,[I_n]' \otimes I_m\right) vec\left[b' \otimes A\right] \\
&= \left(vec\,[I_n]' \otimes I_m\right) \left(I_n \otimes W_{n,1} \otimes I_m\right) \left(vec\left[b'\right] \otimes vec\,[A]\right) \\
&= \left(vec\,[I_n]' \otimes I_m\right) \left(I_{n^2} \otimes I_m\right) vec\left[vec\,[A]\,b'\right] \\
&= \left(vec\,[I_n]' \otimes I_m\right) vec\left[vec\,[A]\,b'\right].
\end{aligned}
$$

A.4.4 Proof of Lemma 6

A.4.4.1 Part i)

Let us write I_{131} as $I_{131} = (I_{d_1} \otimes E_2')\tilde{I}_{131}$ and:

$$\begin{aligned}
\tilde{I}_{131} &= \frac{1}{\sqrt{n}}\sum_i \tau_{i,T}^2 \left(\hat{w}_i \otimes \left[\hat{Q}_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\hat{Q}_{x,i}^{-1}\right]\right) \\
&= \frac{1}{\sqrt{n}}\sum_i \tau_{i,T}^2 \left(\hat{w}_i \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) \\
&\quad + \frac{1}{\sqrt{n}}\sum_i \tau_{i,T}^2 \left(\hat{w}_i \otimes \left[\left(\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1}\right)\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) \\
&\quad + \frac{1}{\sqrt{n}}\sum_i \tau_{i,T}^2 \left(\hat{w}_i \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\left(\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1}\right)\right]\right) \\
&\quad + \frac{1}{\sqrt{n}}\sum_i \tau_{i,T}^2 \left(\hat{w}_i \otimes \left[\left(\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1}\right)\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)\left(\hat{Q}_{x,i}^{-1} - Q_{x,i}^{-1}\right)\right]\right) \\
&=: I_{1311} + I_{1312} + I_{1312}' + I_{1313}.
\end{aligned}$$

We control the terms separately.

Proof that $I_{1311} = \frac{1}{\sqrt{n}}\sum_i \tau_i^2 \left(w_i \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) + O_{p,log}(\sqrt{n}/T)$ $= O_p(1) + O_{p,log}(\sqrt{n}/T)$. We use a decomposition similar to term I_{111} in the proof of Lemma 5:

$$\begin{aligned}
I_{1311} &= \frac{1}{\sqrt{n}}\sum_i \tau_i^2 \left(w_i \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) \\
&\quad + \frac{1}{\sqrt{n}}\sum_i \tau_i^2 \left(\mathbf{1}_i^{\chi} - 1\right)\left(w_i \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) \\
&\quad + \frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^{\chi}(\tau_{i,T}^2 - \tau_i^2)\left(w_i \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) \\
&\quad + \frac{1}{\sqrt{n}}\sum_i \mathbf{1}_i^{\chi}\tau_{i,T}^2 \left((diag[\hat{v}_i]^{-1} - diag[v_i]^{-1}) \otimes \left[Q_{x,i}^{-1}\left(Y_{i,T}Y_{i,T}' - S_{ii,T}\right)Q_{x,i}^{-1}\right]\right) \\
&=: I_{13111} + I_{13112} + I_{13113} + I_{13114}.
\end{aligned}$$

To simplify the notation, let us treat $x_{i,t}$ as a scalar. We first prove $I_{13111} = O_p(1)$. We have:

$$\begin{aligned}
E[I_{13111}^2|\mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}] &= \frac{1}{n}\sum_{i,j} w_i w_j \tau_i^2 \tau_j^2 Q_{x,i}^{-2} Q_{x,j}^{-2} cov\left(Y_{i,T}^2, Y_{j,T}^2 | \mathcal{F}_{\underline{T}}, I_{\underline{T}}(\gamma_i), I_{\underline{T}}(\gamma_j), \gamma_i, \gamma_j\right) \\
= \frac{1}{nT^2}\sum_{i,j}\sum_{t_1,t_2,t_3,t_4} & w_i w_j \tau_i^2 \tau_j^2 Q_{x,i}^{-2} Q_{x,j}^{-2} cov\left(\varepsilon_{i,t_1}\varepsilon_{i,t_2}, \varepsilon_{j,t_3}\varepsilon_{j,t_4} | \mathcal{F}_{\underline{T}}, \gamma_i, \gamma_j\right) I_{i,t_1} I_{i,t_2} I_{j,t_3} I_{j,t_4} x_{i,t_1} x_{i,t_2} x_{j,t_3} x_{j,t_4}.
\end{aligned}$$

From Assumptions B.3 b) and B.4, it follows $E[I_{13111}^2] = O(1)$. Hence, $I_{13111} = O_p(1)$. We can prove that $I_{13112} = o_p(1)$ and $I_{13113} = o_p(1)$ by using arguments similar to terms I_{1112} and I_{1113} in the proof of Lemma 5. Finally, let us prove that $I_{13114} = O_{p,log}(\sqrt{n}/T)$. Similarly to I_{1114} in the proof of Lemma 5, we use

$$\hat{v}_i^{-1} - v_i^{-1} = -v_i^{-2}\left(\hat{v}_i - v_i\right) + \hat{v}_i^{-1} v_i^{-2}\left(\hat{v}_i - v_i\right)^2, \tag{40}$$

and Equation (30). We focus on the term:

$$I_{131141} = -\frac{1}{\sqrt{n}} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^3 C_{\hat{\nu}_1}' \hat{Q}_{x,i}^{-1} \left(\hat{S}_{ii} - S_{ii}\right) \hat{Q}_{x,i}^{-1} C_{\hat{\nu}_1} Q_{x,i}^{-2} \left(Y_{i,T}^2 - S_{ii,T}\right),$$

the other contributions to I_{13114} can be controlled similarly. Now, we use Equation (32). We have:

$$\begin{aligned}
I_{131141} \quad = \quad & -\frac{1}{\sqrt{nT}} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^4 C_{\hat{\nu}_1}' \hat{Q}_{x,i}^{-1} W_{1,i,T} \hat{Q}_{x,i}^{-1} C_{\hat{\nu}_1} Q_{x,i}^{-2} \left(Y_{i,T}^2 - S_{ii,T}\right) \\
& -\frac{1}{\sqrt{nT}} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^4 C_{\hat{\nu}_1}' \hat{Q}_{x,i}^{-1} W_{2,i,T} \hat{Q}_{x,i}^{-1} C_{\hat{\nu}_1} Q_{x,i}^{-2} \left(Y_{i,T}^2 - S_{ii,T}\right) \\
& +2\frac{1}{\sqrt{n}T} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^5 C_{\hat{\nu}_1}' \hat{Q}_{x,i}^{-1} W_{3,i,T} \hat{Q}_{x,i}^{-1} Y_{i,T} \hat{Q}_{x,i}^{-1} C_{\hat{\nu}_1} Q_{x,i}^{-2} \left(Y_{i,T}^2 - S_{ii,T}\right) \\
& -\frac{1}{\sqrt{n}T} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^6 C_{\hat{\nu}_1}' \hat{Q}_{x,i}^{-1} \hat{Q}_{x,i}^{(4)} \hat{Q}_{x,i}^{-1} Y_{i,T}^2 \hat{Q}_{x,i}^{-2} C_{\hat{\nu}_1} Q_{x,i}^{-2} \left(Y_{i,T}^2 - S_{ii,T}\right) \\
=: \quad & -C_{\hat{\nu}_1}' (I_{1311411} + I_{1311412} + I_{13211413} + I_{1311414}) C_{\hat{\nu}_1}.
\end{aligned}$$

Let us focus on term $I_{1311411}$ and prove that it is $O_{p,log}(\sqrt{n}/T)$. We have:

$$\begin{aligned}
I_{1311411} \quad = \quad & \frac{1}{\sqrt{nT}} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^4 \hat{Q}_{x,i}^{-2} Q_{x,i}^{-2} W_{1,i,T} Y_{i,T}^2 - \frac{1}{\sqrt{nT}} \sum_i \mathbf{1}_i^{\chi} v_i^{-2} \tau_{i,T}^4 \hat{Q}_{x,i}^{-2} Q_{x,i}^{-2} W_{1,i,T} S_{ii,T} \\
=: \quad & I_{13114111} + I_{13114112}.
\end{aligned}$$

Term $I_{13114111}$ is such that:

$$|E[I_{13114111} | \mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}]| \leq \frac{C \chi_{1,T}^4 \chi_{2,T}^4}{\sqrt{n} T^2} \sum_i \sum_{t_1,t_2,t_3} |E[\eta_{i,t_1} \varepsilon_{i,t_2} \varepsilon_{i,t_3} | \mathcal{F}_{\underline{T}}, \gamma_i]|,$$

and

$$V[I_{13114111} | \mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}] \leq \frac{C \chi_{1,T}^8 \chi_{2,T}^8}{n T^4} \sum_{i,j} \sum_{t_1,\ldots,t_6} |cov(\eta_{i,t_1} \varepsilon_{i,t_2} \varepsilon_{i,t_3}, \eta_{j,t_4} \varepsilon_{j,t_5} \varepsilon_{j,t_6} | \mathcal{F}_{\underline{T}}, \gamma_i, \gamma_j)|.$$

From Assumptions B.2, B.3 f) and B.5, we get $E[I_{13114111}] = O_{log}(\sqrt{n}/T)$ and $V[I_{13114111}] = o(1)$, which implies $I_{13114111} = O_{p,log}(\sqrt{n}/T)$. The other terms making I_{13114} can be controlled similarly, and we get $I_{13114} = O_{p,log}(\sqrt{n}/T)$.

Proof that $I_{1312} = o_p(1)$. We have:

$$
\begin{aligned}
I_{1312} &= \frac{1}{\sqrt{n}} \sum_i \mathbf{1}_i^\chi \tau_{i,T}^2 \left(diag[v_i]^{-1} \otimes \left[\left(\hat{Q}_{x,i}^{-1} - \hat{Q}_{x,i}^{-1} \right) \left(Y_{i,T} Y_{i,T}' - S_{ii,T} \right) Q_{x,i}^{-1} \right] \right) \\
&\quad + \frac{1}{\sqrt{n}} \sum_i \mathbf{1}_i^\chi \tau_{i,T}^2 \left((diag[\hat{v}_i]^{-1} - diag[v_i]^{-1}) \otimes \left[\left(\hat{Q}_{x,i}^{-1} - \hat{Q}_{x,i}^{-1} \right) \left(Y_{i,T} Y_{i,T}' - S_{ii,T} \right) Q_{x,i}^{-1} \right] \right) \\
&=: I_{13121} + I_{13122}.
\end{aligned}
$$

We focus on term I_{13121}, use Equation (33), and treat $x_{i,t}$ as a scalar to ease notation. We have:

$$
I_{13121} = -\frac{1}{\sqrt{n}} \sum_i \mathbf{1}_i^\chi v_i^{-1} \tau_{i,T}^3 \hat{Q}_{x,i}^{-1} W_{i,T} Q_{x,i}^{-2} \left(Y_{i,T}^2 - S_{ii,T} \right).
$$

Then:

$$
E[\|I_{13121}\|^2 | \mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}] \leq \frac{C\chi_{1,T}^4 \chi_{2,T}^6}{nT^2} \sum_{i,j} \sum_{t_1,\ldots,t_4} \|W_{i,T}\| \|W_{j,T}\| |cov(\varepsilon_{i,t_1}\varepsilon_{i,t_2}, \varepsilon_{j,t_3}\varepsilon_{j,t_4} | \mathcal{F}_{\underline{T}}, \gamma_i, \gamma_j)|.
$$

By the Cauchy-Schwarz inequality, we get:

$$
\begin{aligned}
E[\|I_{13121}\|^2 | \{\gamma_i\}] &\leq C\chi_{1,T}^4 \chi_{2,T}^6 \sup_i E[\|W_{i,T}\|^4 | \gamma_i]^{1/2} \\
&\quad \frac{1}{nT^2} \sum_{i,j} \sum_{t_1,t_2,t_3,t_4} E\left[|cov(\varepsilon_{i,t_1}\varepsilon_{i,t_2}, \varepsilon_{j,t_3}\varepsilon_{j,t_4} | \mathcal{F}_{\underline{T}}, \gamma_i, \gamma_j)|^2 | \gamma_i, \gamma_j \right]^{1/2}.
\end{aligned}
$$

From Assumptions B.1 b), B.3 b), B.4 a), and B.5, we deduce $E[\|I_{13121}\|^2] = o(1)$, which implies $I_{13121} = o_p(1)$. Similar arguments can be used to prove that the other terms making I_{1312} are $o_p(1)$.

Proof that $I_{1313} = o_p(1)$. This step uses similar arguments as for I_{1312}.

A.4.4.2 Part (ii)

Let us treat $x_{i,t}$ as a scalar to ease notation. We have $I_{132} = (I_{d_1} \otimes E_2')\tilde{I}_{132}$ where $\tilde{I}_{132} = \frac{1}{\sqrt{nT}} \sum_i \hat{w}_i \tau_{i,T}^2 \hat{Q}_{x,i}^{-1} W_{1,i,T} \hat{Q}_{x,i}^{-1}$, and $W_{1,i,T}$ is as in Equation (32). Write:

$$
\tilde{I}_{132} = \frac{1}{\sqrt{nT}} \sum_i \mathbf{1}_i^\chi v_i^{-1} \tau_{i,T}^2 \hat{Q}_{x,i}^{-1} W_{1,i,T} \hat{Q}_{x,i}^{-1} + \frac{1}{\sqrt{nT}} \sum_i \mathbf{1}_i^\chi (\hat{v}_i^{-1} - v_i^{-1}) \tau_{i,T}^2 \hat{Q}_{x,i}^{-1} W_{1,i,T} \hat{Q}_{x,i}^{-1} =: I_{1321} + I_{1322}.
$$

Let us first consider I_{1321}. We have:

$$E[\|I_{1321}\|^2|\mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}] \leq C\chi_{1,T}^8\chi_{2,T}^4\frac{1}{nT^2}\sum_{i,j}\sum_{t_1,t_2}|cov(\eta_{i,t_1}, \eta_{j,t_2}|\mathcal{F}_{\underline{T}}, \gamma_i, \gamma_j)|.$$

From Assumptions B.3 a) and B.5, it follows $E[\|I_{1321}\|^2] = O_{log}(1/T)$, and thus $I_{1321} = O_{p,log}(1/\sqrt{T})$.

Let us now consider term I_{1322}. We use Equation (40), and plug in the decompositions (30) and (32). We focus on term $C_{\hat{\nu}_1}^2 I_{13221}$ of the resulting expansion, where:

$$I_{13221} = -\frac{1}{\sqrt{n}T}\sum_i \mathbf{1}_i^\chi v_i^{-2}\tau_{i,T}^4\hat{Q}_{x,i}^{-4}W_{1,i,T}^2.$$

The other terms can be treated similarly. We have:

$$E[I_{13221}|\mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}] \leq C\chi_{1,T}^8\chi_{2,T}^4\frac{1}{\sqrt{n}T^2}\sum_i\sum_{t_1,t_2}|cov(\varepsilon_{i,t_1}^2, \varepsilon_{i,t_2}^2|\mathcal{F}_{\underline{T}}, \gamma_i)|,$$

and

$$V[I_{13221}|\mathcal{F}_{\underline{T}}, \{I_{\underline{T}}(\gamma_i), \gamma_i\}] \leq C\chi_{1,T}^{16}\chi_{2,T}^8\frac{1}{nT^4}\sum_{i,j}\sum_{t_1,t_2,t_3,t_4}|cov(\eta_{i,t_1}\eta_{i,t_2}, \eta_{j,t_3}\eta_{j,t_4}|\mathcal{F}_{\underline{T}}, \gamma_i, \gamma_j)|.$$

From Assumptions B.3 a) and B.5, it follows $E[I_{13221}] = O_{log}(\sqrt{n}/T)$. By Assumptions B.3 d) and B.5 we can prove that $V[I_{13221}] = o(1)$, and it follows $I_{13221} = O_p(\sqrt{n}/T)$.

A.4.4.3 Part (iii)

We have $I_{133} = (I_{d_1} \otimes E_2')\tilde{I}_{133}$, where $I_{133} = -\frac{2}{\sqrt{n}T}\sum_i \hat{w}_i\tau_{i,T}^3\hat{Q}_{x,i}^{-3}W_{3,i,T}Y_{i,T}$ $+\frac{1}{\sqrt{n}T}\sum_i \hat{w}_i\tau_{i,T}^4\hat{Q}_{x,i}^{-4}\hat{Q}_{x,i}^{(4)}Y_{i,T}^2$ and $W_{3,i,T}$ and $\hat{Q}_{x,i}^{(4)}$ are as in Equation (32) and we treat $x_{i,t}$ as a scalar to ease notation. By similar arguments as in part (ii), we can prove that $I_{133} = O_{p,log}(\sqrt{n}/T)$.

A.4.4.4 Part (iv)

The statement follows from Lemma 3 (ii)-(iii), $\mathbf{1}_i^\chi\tau_{i,T} \leq \chi_{2,T}$, $\mathbf{1}_i^\chi\|\hat{Q}_{x,i}^{-1}\| \leq C\chi_{1,T}^2$, bound (37), $\|S_{ii}\| \leq M$ and Assumption B.5.

A.4.4.5 Part (v)

The statement follows from Equation (28), Lemma 3 (iv), $I_{11} = O_p(1)$, and $\frac{1}{n}\sum_i \hat{w}_i \tau_{i,T}^2 \hat{Q}_{x,i}^{-1} Y_{i,T} Y_{i,T}' \hat{Q}_{x,i}^{-1} = O_{p,log}(1)$.

A.4.5 Proof of Lemma 7

We have $P\left[\mathbf{1}_i^{\chi} = 0\right] \leq P\left[\tau_{i,T} \geq \chi_{2,T}\right] + P\left[CN\left(\hat{Q}_{x,i}\right) \geq \chi_{1,T}\right] =: P_{1,nT} + P_{2,nT}$. Let us first control $P_{1,nT}$. We have $P_{1,nT} \leq P\left[\frac{1}{T}\sum_t I_{i,t} \leq \chi_{2,T}^{-1}\right] \leq P\left[\frac{1}{T}\sum_t \left(I_{i,t} - \tau_i^{-1}\right) \leq \chi_{2,T}^{-1} - M^{-1}\right]$, where we use $\tau_i \leq M$ for all i (Assumption B.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 $P_{1,nT} \leq P\left[\left|\frac{1}{T}\sum_t \left(I_{i,t} - \tau_i^{-1}\right)\right| \geq \delta\right]$. By using that $\tau_i^{-1} = E[I_{i,t}|\gamma_i]$, and $P\left[\left|\frac{1}{T}\sum_t \left(I_{i,t} - \tau_i^{-1}\right)\right| \geq \delta\right] = E\left[P\left[\left|\frac{1}{T}\sum_t \left(I_{i,t} - E[I_{i,t}|\gamma_i]\right)\right| \geq \delta | \gamma_i\right]\right] \leq \sup_{\gamma \in [0,1]} P\left[\left|\frac{1}{T}\sum_t \left(I_t(\gamma) - E[I_t(\gamma)]\right)\right| \geq \delta\right]$, from Assumption B.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 B.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 $P_{2,nT} \leq P\left[eig_{min}(\hat{Q}_{x,i}) \leq M/\chi_{1,T}^2\right]$. By using that $eig_{min}(\hat{Q}_{x,i}) \geq eig_{min}(Q_{x,i}) - \|\hat{Q}_{x,i} - Q_{x,i}\|$, we get $P_{2,nT} \leq P\left[\|\hat{Q}_{x,i} - Q_{x,i}\| \geq eig_{min}(Q_{x,i}) - M/\chi_{1,T}^2\right]$. Now, let $\delta > 0$ be such that $eig_{min}(Q_{x,i}) - M/\chi_{1,T}^2 > \delta$ uniformly in i for large T (see Assumption B.4 d)). Then, by using $P\left[\|\hat{Q}_{x,i} - Q_{x,i}\| \geq \delta\right] \leq P\left[\left|\frac{1}{T}\sum_t I_{i,t}(x_{i,t}x_{i,t} - Q_{x,i})\right| \geq \sqrt{\delta}\right] + P\left[\tau_{i,T} \geq \sqrt{\delta}\right]$, we get $P_{2,nT} \leq P\left[\left|\frac{1}{T}\sum_t I_{i,t}(x_{i,t}x_{i,t} - Q_{x,i})\right| \geq \sqrt{\delta}\right] + O(T^{-\bar{b}})$. The first term in the RHS is $O(T^{-\bar{b}})$ by using $P\left[\left|\frac{1}{T}\sum_t I_{i,t}(x_{i,t}x_{i,t} - Q_{x,i})\right| \geq \sqrt{\delta}\right] \leq \sup_{\gamma \in [0,1]} P\left[\left|\frac{1}{T}\sum_t I_t(\gamma)(x_t(\gamma)x_t(\gamma)' - E[x_t(\gamma)x_t(\gamma)'])\right| \geq \sqrt{\delta}\right]$ and Assumption B.1 b). Then, $P_{2,nT} = O(T^{-\bar{b}})$, for any $\bar{b} > 0$.

A.4.6 Proof of Lemma 8

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

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

from Assumption B.1 d).

A.4.7 Proof of Lemma 9

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,

$$I_{31} = \frac{1}{n}\sum_{i,j} \|S_{ij}\| \mathbf{1}_{\{\|S_{ij}\|<\kappa\}} \leq \max_i \sum_j \|S_{ij}\|^{\bar q} \kappa^{1-\bar q} \leq \kappa^{1-\bar q} c_0(n) = O_p\left(\kappa^{1-\bar q} n^{\bar\delta}\right), \tag{41}$$

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

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:

$$I_{35} \leq \max_{i,j} \left\| \hat{S}_{ij} - S_{ij} \right\| \max_{i} \sum_{j} \left\| S_{ij} \right\|^{\bar{q}} \kappa^{-\bar{q}} = O_p \left(\psi_{nT} c_0 \left(n \right) \kappa^{-\bar{q}} \right). \tag{42}$$

Let us study I_{33}:

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

By Assumption A.4,

$$I_{37} \leq \kappa^{1-\bar{q}} c_0 \left(n \right). \tag{43}$$

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

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

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

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

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

$$I_{36} = O_p \left(\psi_{nT} n^2 \Psi_{nT} \left(\left(1 - v \right) \kappa \right) + \psi_{nT} c_0 \left(n \right) \left(v\kappa \right)^{-\bar{q}} \right). \tag{44}$$

Finally, we consider I_{34}. We have

$$\begin{aligned} I_{34} &\leq \max_{i} \sum_{j} \left(\left\| \hat{S}_{ij} - S_{ij} \right\| + \left\| \hat{S}_{ij} \right\| \right) \mathbf{1}_{\left\{ \left\| \hat{S}_{ij} \right\| < \kappa, \left\| S_{ij} \right\| \geq \kappa \right\}} \\ &\leq \max_{i,j} \left\| \hat{S}_{ij} - S_{ij} \right\| \max_{i} \sum_{j} \mathbf{1}_{\left\{ \left\| S_{ij} \right\| \geq \kappa \right\}} + \kappa \max_{i} \sum_{j} \mathbf{1}_{\left\{ \left\| S_{ij} \right\| \geq \kappa \right\}} \\ &= O_p \left(\psi_{nT} c_0 \left(n \right) \kappa^{-\bar{q}} + c_0 \left(n \right) \kappa^{1-\bar{q}} \right). \end{aligned} \tag{45}$$

Combining (41)-(45) the result follows.

A.4.8 Proof of Lemma 10

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

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

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 $\xi \geq 0$:

$$\begin{aligned}
\Psi_{nT}(\xi) &\leq \max_{i,j} P\left[\left\|\hat{S}^0_{ij} - S_{ij}\right\| \geq \frac{\xi}{4}\right] + \max_{i,j} P\left[\|A_{ij}\| \geq \frac{\xi}{4}\right] + \max_{i,j} P\left[\|B_{ij}\| \geq \frac{\xi}{4}\right] \\
&\quad + \max_{i,j} P\left[\|C_{ij}\| \geq \frac{\xi}{4}\right] = \Psi^0_{nT}(\xi/4) + 2P_{1,nT}(\xi/4) + P_{2,nT}(\xi/4), \qquad (46)
\end{aligned}$$

where $\Psi^0_{nT}(\xi/4) := \max_{i,j} P\left[\left\|\hat{S}^0_{ij} - S_{ij}\right\| \geq \frac{\xi}{4}\right]$, $P_{1,nT}(\xi/4) := \max_{i,j} P\left[\|A_{ij}\| \geq \frac{\xi}{4}\right]$, and $P_{2,nT}(\xi/4) := \max_{i,j} P\left[\|C_{ij}\| \geq \frac{\xi}{4}\right]$. Let us bound the three terms in the RHS of Inequality (46).

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

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

We deduce:

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

for small ξ. We use

$$
P_{3,nT} \leq \sup_{\gamma,\tilde{\gamma}\in[0,1]} P\left[\left\|\frac{1}{T}\sum_t I_t(\gamma)I_t(\tilde{\gamma})\left(\varepsilon_t(\gamma)\varepsilon_t(\tilde{\gamma})x_t(\gamma)x_t(\tilde{\gamma})' - E\left[\varepsilon_t(\gamma)\varepsilon_t(\tilde{\gamma})x_t(\gamma)x_t(\tilde{\gamma})'\right]\right)\right\| \geq \frac{\xi}{8M}\right]
$$

and Assumption B.1 e) to get $P_{3,nT} \leq C_1 T\exp\left\{-C_2^*\xi^2T^{\eta}\right\} + C_3^*\xi^{-1}\exp\left\{-C_4T^{\bar{\eta}}\right\}$, for some constants $C_1, C_2^*, C_3^*, C_4 > 0$. To bound $P_{4,nT}$, we use $\tau_{ij} \leq M$ and $|\tau_{ij,T}-\tau_{ij}| \leq \tau_{ij}\tau_{ij,T}|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}| \leq \tau_{ij}\dfrac{|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}|}{\tau^{-1}_{ij}-|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}|} \leq 2M^2|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}|$, if $|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}| \leq M^{-1}/2$. Thus, we have $P_{4,nT} \leq 2\max_{i,j} P\left[|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}| \geq \dfrac{1}{2M^2}\sqrt{\dfrac{\xi}{8}}\right]$, for small ξ. By using $\tau^{-1}_{ij,T} = \dfrac{1}{T}\sum_t I_{ij,t}$ and $\tau^{-1}_{ij} = E[I_{ij,t}|\gamma_i,\gamma_j]$, from Assumption B.1 d) we get:

$$
\begin{aligned}
\max_{i,j} P\left[|\tau^{-1}_{ij,T}-\tau^{-1}_{ij}| \geq \frac{1}{2M^2}\sqrt{\frac{\xi}{8}}\right] \quad &\leq \quad \sup_{\gamma,\tilde{\gamma}\in[0,1]} P\left[\left|\frac{1}{T}\sum_t (I_t(\gamma)I_t(\tilde{\gamma}) - E[I_t(\gamma)I_t(\tilde{\gamma})])\right| \geq \frac{1}{2M^2}\sqrt{\frac{\xi}{8}}\right] \\
&\leq \quad C_1T\exp\left\{-C_2^*\xi T^{\eta}\right\} + C_3^*\xi^{-1/2}\exp\left\{-C_4T^{\bar{\eta}}\right\}.
\end{aligned}
$$

We deduce:

$$
\Psi^0_{nT}(\xi/4) \leq C_1^*T\exp\left\{-C_2^*\xi^2T^{\eta}\right\} + C_3^*\xi^{-1}\exp\left\{-C_4T^{\bar{\eta}}\right\}. \tag{47}
$$

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

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

Let $\chi_{3,T} = (\log T)^a$, for $a > 0$. From a similar argument as in the proof of Lemma 7, and Assumption B.1 d), we have $\max_{i,j} P\left[\tau_{ij,T} \geq \chi_{3,T}\right] = O(T^{-\bar{b}})$, for any $\bar{b} > 0$. Thus,

$$
\begin{aligned}
& P_{1,nT}\left(\xi/4\right) \\
\leq\; & \max_{i,j} P\left[\tau_{ij,T} \max_{k,l,m}\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{i,t,k}x_{i,t,l}x_{j,t,m}\right|\left\|\hat{\beta}_j - \beta_j\right\| \geq \frac{\xi}{4C}\right] \\
\leq\; & \max_{i,j} P\left[\tau_{ij,T} \geq \chi_{3,T}\right] + \max_{i,j} P\left[\max_{k,l,m}\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{i,t,k}x_{i,t,l}x_{j,t,m}\right| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}} \text{ and } \tau_{ij,T} \leq \chi_{3,T}\right] \\
& + \max_{i,j} P\left[\left\|\hat{\beta}_j - \beta_j\right\| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}} \text{ and } \tau_{ij,T} \leq \chi_{3,T}\right] \\
\leq\; & d^3 \max_{i,j}\max_{k,l,m} P\left[\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{i,t,k}x_{i,t,l}x_{j,t,m}\right| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}}\right] \\
& + P\left[\left\|\hat{\beta}_j - \beta_j\right\| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}} \text{ and } \tau_{j,T} \leq \chi_{3,T}\right] + O(T^{-\bar{b}}).
\end{aligned} \tag{48}
$$

By Assumption B.1 f),

$$
\begin{aligned}
\max_{i,j}\max_{k,l,m} P\left[\left|\frac{1}{T}\sum_t I_{ij,t}\varepsilon_{i,t}x_{i,t,k}x_{i,t,l}x_{j,t,m}\right| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}}\right] \quad \leq \quad & C_1 T \exp\left\{-\frac{C_2^*\xi}{\chi_{3,T}}T^{\eta}\right\} \\
& + C_3^*\sqrt{\frac{\chi_{3,T}}{\xi}}\exp\left\{-C_4 T^{\bar{\eta}}\right\}.
\end{aligned} \tag{49}
$$

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

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

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

$$
\begin{aligned}
& P\left[\left\|\hat{\beta}_j - \beta_j\right\| \geq \sqrt{\frac{\xi}{4\chi_{3,T}C}} \text{ and } \tau_{j,T} \leq \chi_{3,T}\right] \\
\leq\ & P\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_{j,t}\varepsilon_{j,t}\right\| \geq \frac{1}{2}\sqrt{\frac{\xi}{4\chi_{3,T}C}}\chi_{3,T}^{-1}\left\|Q_{x,j}^{-1}\right\|^{-1}\right] \\
& + P\left[\left\|\hat{Q}_{x,j}^{-1} - Q_{x,j}^{-1}\right\|\left\|\frac{1}{T}\sum_t I_{j,t}x_{j,t}\varepsilon_{j,t}\right\| \geq \frac{1}{2}\sqrt{\frac{\xi}{4\chi_{3,T}C}}\chi_{3,T}^{-1}\right] \\
\leq\ & P\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_{j,t}\varepsilon_{j,t}\right\| \geq \sqrt{\frac{\xi}{16\chi_{3,T}^3C}}\left\|Q_{x,j}^{-1}\right\|^{-1}\right] \\
& + P\left[\left\|\hat{Q}_{x,j}^{-1} - Q_{x,j}^{-1}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3C}\right)^{1/4}\right] + P\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_{j,t}\varepsilon_{j,t}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3C}\right)^{1/4}\right] \\
\leq\ & 2P\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_{j,t}\varepsilon_{j,t}\right\| \geq \sqrt{\frac{\xi}{16\chi_{3,T}^3C}}\left\|Q_{x,j}^{-1}\right\|^{-1}\right] + P\left[\left\|\hat{Q}_{x,j}^{-1} - Q_{x,j}^{-1}\right\| \geq \left(\frac{\xi}{16\chi_{3,T}^3C}\right)^{1/4}\right], \quad (50)
\end{aligned}
$$

for small ξ. From Assumption B.4 d), $\left\|Q_{x,j}^{-1}\right\|$ is bounded uniformly in j. Then, from Assumption B.1c), the first probability in the RHS of Inequality (50) is such that:

$$
P\left[\left\|\frac{1}{T}\sum_t I_{j,t}x_{,jt}\varepsilon_{j,t}\right\| \geq \sqrt{\frac{\xi}{16\chi_{3,T}^3C}}\left\|Q_{x,j}^{-1}\right\|^{-1}\right] \leq C_1 T\exp\left\{-\frac{C_2^*\xi}{\chi_{3,T}^3}T^{\eta}\right\} + C_3^*\sqrt{\frac{\chi_{3,T}^3}{\xi}}\exp\left\{-C_4 T^{\bar{\eta}}\right\}. \quad (51)
$$

To bound the second probability in the RHS of Inequality (50) we use the next Lemma.

Lemma 13 *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 13, we get:

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

for small $\xi > 0$. From Assumptions B.1b) and B.1c),

$$P\left[\left\|\hat{Q}_{x,j}-Q_{x,j}\right\| \geq \frac{1}{2}\left(\frac{\xi}{16\chi_{3,T}^{3}C}\right)^{1/4}\|Q_{x,j}^{-1}\|^{-2}\right] \leq C_1 T\exp\left\{-C_2^*\sqrt{\frac{\xi}{\chi_{3,T}^3}}T^{\eta}\right\} +2C_3^*\left(\frac{\chi_{3,T}^3}{\xi}\right)^{1/4}\exp\left\{-C_4T^{\bar{\eta}}\right\}. \quad (52)$$

Then, from (48)-(52) we get:

$$P_{1,nT}\left(\xi/4\right) \leq C_1^*T\exp\left\{-C_2^*\xi T^{\eta}/\chi_{3,T}^3\right\}+\frac{C_3^*\chi_{3,T}^{3/2}}{\sqrt{\xi}}\exp\left\{-C_4T^{\bar{\eta}}\right\}+O(T^{-\bar{b}}), \quad (53)$$

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

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

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

Thus, we have:

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

By the same arguments as above, we get:

$$P_{2,nT}\left(\xi/4\right) \leq C_1^*T\exp\left\{-C_2^*\xi T^{\eta}/\chi_{3,T}^3\right\}+\frac{C_3^*\chi_{3,T}^{3/2}}{\sqrt{\xi}}\exp\left\{-C_4T^{\bar{\eta}}\right\}, \quad (54)$$

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

d) Conclusion. From inequalities (46), (47), (53) and (54), we deduce:

$$\Psi_{nT}\left(\xi\right) \leq C_1^*T\exp\left\{-C_2^*\xi_T^2T^{\eta}\right\}+\frac{C_3^*}{\xi_T}\exp\left\{-C_4T^{\bar{\eta}}\right\}+O(T^{-\bar{b}}),$$

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

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

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

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

for large ϵ. The conclusion follows.

A.4.9 Proof of Lemma 11

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

$$
\begin{aligned}
\hat{e}_i &= \beta_{1,i} - \beta_{3,i}\hat{\nu} + C_{\hat{\nu}}'\left(\hat{\beta}_i - \beta_i\right) \\
&= \beta_{1,i} - \beta_{3,i}\nu + C_{\hat{\nu}}'\left(\hat{\beta}_i - \beta_i\right) - \beta_{3,i}\left(\hat{\nu} - \nu\right) \qquad (55) \\
&= C_{\hat{\nu}}'\left(\hat{\beta}_i - \beta_i\right) - \beta_{3,i}\left(\hat{\nu} - \nu\right).
\end{aligned}
$$

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

$$
\begin{aligned}
\hat{Q}_e &= \frac{1}{n}\sum_i \left(\hat{\beta}_i - \beta_i\right)' C_{\hat{\nu}}\hat{w}_i C_{\hat{\nu}}'\left(\hat{\beta}_i - \beta_i\right) - 2\left(\hat{\nu} - \nu\right)'\frac{1}{n}\sum_i \beta_{3,i}'\hat{w}_i C_{\hat{\nu}}'\left(\hat{\beta}_i - \beta_i\right) \\
&\quad + \left(\hat{\nu} - \nu\right)'\frac{1}{n}\sum_i \beta_{3,i}'\hat{w}_i\beta_{3,i}\left(\hat{\nu} - \nu\right) \\
&=: \frac{1}{n}\sum_i \left(\hat{\beta}_i - \beta_i\right)' C_{\hat{\nu}}\hat{w}_i C_{\hat{\nu}}'\left(\hat{\beta}_i - \beta_i\right) - 2I_{71} + I_{72}.
\end{aligned}
$$

Let us study the second term in the RHS:

$$
I_{71} = \frac{1}{\sqrt{nT}}\left(\hat{\nu} - \nu\right)'\frac{1}{\sqrt{n}}\sum_i \tau_{i,T}\beta_{3,i}'\hat{w}_i C_{\hat{\nu}}'\hat{Q}_{x,i}^{-1}Y_{i,T} =: \frac{1}{\sqrt{nT}}(\hat{\nu} - \nu)' I_{711},
$$

where $I_{711} = O_p(1)$ by the same arguments used to control term I_{11} in the proof of Proposition 4. We have $\hat{\nu} - \nu = O_{p,log}\left(\frac{1}{\sqrt{nT}} + \frac{1}{T}\right)$ and $C_{\hat{\nu}} = O_p\left(1\right)$ by Lemma 6 (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 3 (ii)-(iii) and Lemma 6 (v), we have $I_{72} = O_{p,log}\left(\frac{1}{nT} + \frac{1}{T^2}\right)$. The conclusion follows.

A.4.10 Proof of Lemma 12

Under $\mathcal{H}_1$, and using Equation (55), we have $\hat{e}_i = e_i + C'_{\hat{\nu}}\left(\hat{\beta}_i - \beta_i\right) - \beta_{3,i}\left(\hat{\nu} - \nu_\infty\right)$. By definition of $\hat{Q}_e$, it follows:

$$
\begin{aligned}
\hat{Q}_e &= \frac{1}{n}\sum_i e_i'\hat{w}_i e_i + 2\frac{1}{n}\sum_i \left(\hat{\beta}_i - \beta_i\right)' C_{\hat{\nu}}\hat{w}_i e_i - 2\left(\hat{\nu} - \nu_\infty\right)' \frac{1}{n}\sum_i \beta_{3,i}'\hat{w}_i e_i \\
&\quad + \frac{1}{n}\sum_i \left(\hat{\beta}_i - \beta_i\right)' C_{\hat{\nu}}\hat{w}_i C'_{\hat{\nu}}\left(\hat{\beta}_i - \beta_i\right) - 2\left(\hat{\nu} - \nu_\infty\right)' \frac{1}{n}\sum_i \beta_{3,i}'\hat{w}_i C'_{\hat{\nu}}\left(\hat{\beta}_i - \beta_i\right) \\
&\quad + \left(\hat{\nu} - \nu_\infty\right)' \frac{1}{n}\sum_i \beta_{3,i}'\hat{w}_i\beta_{3,i}\left(\hat{\nu} - \nu_\infty\right) =: I_{81} + I_{82} + I_{83} + I_{84} + I_{85} + I_{86}. \qquad (56)
\end{aligned}
$$

From Equations (30) and (32) and similar arguments as in Section A.2.4 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_{11} in the proof of Proposition 4, we have $I_{82} = \frac{2}{\sqrt{nT}}\left(\frac{1}{\sqrt{n}}\sum_i \tau_{i,T} Y'_{i,T}\hat{Q}_{x,i}^{-1} C_{\hat{\nu}}\hat{w}_i e_i\right) = O_p\left(\frac{1}{\sqrt{nT}}\right)$. By using $\frac{1}{n}\sum_i \beta_{3,i}'\hat{w}_i e_i = \frac{1}{n}\sum_i \beta_{3,i}' w_i e_i + O_{p,log}\left(\frac{1}{\sqrt{T}}\right) = O_p\left(\frac{1}{\sqrt{n}}\right) + O_{p,log}\left(\frac{1}{\sqrt{T}}\right)$ 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.4.11 Proof of Lemma 13

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

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

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

Appendix 5 Link to Chamberlain and Rothschild (1983)

In this appendix, we establish the link between the no-arbitrage conditions and asset pricing restrictions in CR on the one hand, and the asset pricing restriction (3) in the other hand. As in Appendix A.2.1, for any sequence (γ_i) in Γ let $\mathcal{P}_n$ be the set of portfolios investing in the n assets $\gamma_1, \gamma_2, ..., \gamma_n$ with $\mathcal{F}_0$-measurable shares. By assuming that the shares are finite P-a.s., we have $E\left[p_n^2|\mathcal{F}_0\right]<\infty$, P-a.s., and we can build on the framework of Hansen and Richard (1987) with conditionally square integrable payoffs. Moreover, we denote by $\mathcal{P}=\bigcup_{n=1}^{\infty}\mathcal{P}_n$ the set of finite portfolios with conditionally square integrable payoff.

Let $\mathcal{J}^* \subset \Gamma$ be the set of countable collections of assets (γ_i) such that Conditions (i) and (ii) hold for any portfolio sequence $(p_n)\in\mathcal{P}$, where Conditions (i) and (ii) are: (i) If $V\left[p_n|\mathcal{F}_0\right] \overset{a.s}{\to} 0$ and $C(p_n) \overset{a.s}{\to} 0$, then $E\left[p_n|\mathcal{F}_0\right] \overset{a.s}{\to} 0$; (ii) If $V[p_n|\mathcal{F}_0] \overset{a.s}{\to} 0$, $C(p_n)\geq 0$, P-a.s., $\limsup_{n\to\infty}|C(p_n)|\geq\epsilon$ on a set of nonzero measure, for a constant $\epsilon>0$, and $E[p_n|\mathcal{F}_0] \overset{a.s}{\to} \bar{\delta}$, for a constant $\bar{\delta}$, then $\bar{\delta}>0$. Condition (i) means that, if the conditional variability and cost vanish, so does the conditional expected return. Condition (ii) means that, if the conditional variability vanishes and the cost is positive, the conditional expected return is positive. They correspond to Conditions A.1 (i) and (ii) in CR written conditionally on $\mathcal{F}_0$ and for a given countable collection of assets (γ_i). Hence, the set $\mathcal{J}^*$ is the set permitting no asymptotic arbitrage opportunities in the sense of CR in a conditional setting (see also Chamberlain (1983)). We use the convergence of conditional expectations as in Hansen and Richard (1987), and focus on $a.s.$ convergence as opposed to convergence in probability (see Hansen and Richard (1987), footnote 5 on p. 594) since this helps when defining the extension of the cost function $C(\cdot)$ to the completion of set $\mathcal{P}$. Let $\mathcal{J}^{**}\subset\Gamma$ be the set of sequences (γ_i)

such that $\inf_{\nu \in \mathbb{R}^K} \sum_{i=1}^{\infty} [a(\gamma_i) - b(\gamma_i)'\nu]^2 < \infty$, P-a.s. These sequences met the summability condition of CR in a conditional setting. In the proof of the following proposition, we assume that β is bounded on $[0,1] \times \Omega$ and $E[f_1|\mathcal{F}_0]$ is bounded on Ω.

Proposition APR: *Under Assumptions APR.1-APR.3, and (i)* $\inf_{n \geq 1} eig_{min}(\Sigma_{\varepsilon,t,n}) > 0$, P-a.s., *for a.e.* (γ_i) *in* Γ, *(ii)* $eig_{min}(V[f_t|\mathcal{F}_{t-1}]) > 0$, P-a.s., *we have: either* $\bar{\mu}_\Gamma(\mathcal{J}^*) = \bar{\mu}_\Gamma(\mathcal{J}^{**}) = 1$, *or* $\bar{\mu}_\Gamma(\mathcal{J}^*) = \bar{\mu}_\Gamma(\mathcal{J}^{**}) = 0$. *The former case occurs if, and only if, the asset pricing restriction (3) holds.*

When we condition on $\mathcal{F}_0$, the fact that the set of sequences such that $\inf_{\nu \in \mathbb{R}^K} \sum_{i=1}^{\infty} [a(\gamma_i) - b(\gamma_i)'\nu]^2 < \infty$ has μ_Γ-measure equal to either 1, or 0, is a consequence of the Kolmogorov zero-one law (e.g., Billingsley (1995)). Indeed, $\inf_{\nu \in \mathbb{R}^K} \sum_{i=1}^{\infty} [a(\gamma_i) - b(\gamma_i)'\nu]^2 < \infty$ if, and only if, $\inf_{\nu \in \mathbb{R}^K} \sum_{i=n}^{\infty} [a(\gamma_i) - b(\gamma_i)'\nu]^2 < \infty$, for any $n \in \mathbb{N}$. Thus, the zero-one law applies since the event $\inf_{\nu \in \mathbb{R}^K} \sum_{i=1}^{\infty} [a(\gamma_i) - b(\gamma_i)'\nu]^2 < \infty$ belongs to the tail sigma-field $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(\gamma_i, i = n, n+1, \ldots)$, and the variables γ_i are i.i.d. under measure μ_Γ. Proposition APR shows that this zero-one measure property applies also for the set $\mathcal{J}^{**}$. Proposition APR shows that the asset pricing (3) characterizes the functions $\beta = (a, b')'$ defined on $[0,1] \times \Omega$ that are compatible with absence of asymptotic arbitrage opportunities in the continuum economy under the definitions of arbitrage used in CR and in Hansen and Richard (1987). Moreover, Proposition APR also provides a reverse implication compared to Proposition 1: when the asset pricing restriction (3) does not hold, asymptotic arbitrage in the sense of Assumption APR.4, or of Assumptions A.1 i) and ii) of CR, exists for $\bar{\mu}_\Gamma$-almost any countable collection of assets.

Proof of Proposition APR: The proof involves four steps.

STEP 1: If the asset pricing restriction (3) holds, then $\bar{\mu}_\Gamma(\mathcal{J}^{**}) = 1$. Indeed, if the asset pricing restriction (3) holds for some $\mathcal{F}_0$-measurable function ν, we have for a.e. $\omega \in \Omega$: $a(\gamma, \omega) - b(\gamma, \omega)'\nu(\omega) = 0$ for a.e. $\gamma \in [0,1]$. Since functions a and b are jointly measurable on $[0,1] \times \Omega$, this implies that for a.e. $\gamma \in [0,1]$: $a(\gamma, \omega) - b(\gamma, \omega)'\nu(\omega) = 0$ for a.e. $\omega \in \Omega$. Then, the set $\left\{ (\gamma_i) \in \Gamma : \sum_{i=1}^{\infty} [a(\gamma_i) - b(\gamma_i)'\nu]^2 = 0,\ P\text{-a.s.} \right\} = \bigcap_{i=1}^{\infty} \left\{ (\gamma_i) \in \Gamma : a(\gamma_i, \omega) - b(\gamma_i, \omega)'\nu(\omega) = 0, \text{for a.e. } \omega \in \Omega \right\}$ has μ_Γ-measure 1. Since this set is a subset of $\mathcal{J}^{**}$, it follows $\bar{\mu}_\Gamma(\mathcal{J}^{**}) = 1$.

STEP 2: If the asset pricing restriction (3) does not hold, then $\bar{\mu}_\Gamma\left(\mathcal{J}^{**}\right)=0$. If the asset pricing restriction (3) does not hold, the quantity $\delta=\inf_{\nu\in\mathbb{R}^K}\int[a(\gamma)-b(\gamma)'\nu]^2d\gamma$ is such that $\delta(\omega)\geq\underline{\delta}$ for all $\omega\in A$, for a set $A\in\mathcal{F}_0$ with $P(A)>0$ and a scalar $\underline{\delta}>0$. To prove $\bar{\mu}_\Gamma\left(\mathcal{J}^{**}\right)=0$, we show $\mathcal{J}_1\cap\mathcal{J}^{**}=\emptyset$, where $\mathcal{J}_1$ is the set with μ_Γ-measure 1 defined in Lemma 1. Indeed, $\mathcal{J}_1\cap\mathcal{J}^{**}=\emptyset$ implies that $\mathcal{J}^{**}\subset\mathcal{J}_1^c$ is a negligible set under measure μ_Γ, and thus has $\bar{\mu}_\Gamma$ measure 0. The proof of $\mathcal{J}_1\cap\mathcal{J}^{**}=\emptyset$ is by contradiction. Let us assume that sequence (γ_i) is in $\mathcal{J}_1\cap\mathcal{J}^{**}$, and let $\xi_n:=\inf_{\nu\in\mathbb{R}^K}\frac{1}{n}\sum_{i=1}^n[a(\gamma_i)-b(\gamma_i)'\nu]^2$. Since $(\gamma_i)\in\mathcal{J}_1$, from Inequality (19), we have $\xi_n 1_{A\cap S_n^*}\geq 2^{-1}\underline{\delta}1_{A\cap S_n^*}$, where the set S_n^* defined in the proof of Proposition 1 is such that $P(S_n^*)\to 1$ as $n\to\infty$. This implies that $E[\xi_n^2]\geq E[\xi_n^2|1_{A\cap S_n^*}=1]P(A\cap S_n^*)\geq(\underline{\delta}^2/4)P(A\cap S_n^*)$ $\to(\underline{\delta}^2/4)P(A)$, and thus:

$$\liminf_{n\to\infty}E[\xi_n^2]>0. \tag{57}$$

Since $(\gamma_i)\in\mathcal{J}^{**}$, we have $\xi_n\to 0$, P-a.s.. Moreover, since function β is bounded, we have $|\xi_n|\leq C$, P-a.s., for some constant C. Then, by the Lebesgue dominated convergence theorem, it follows that $E[\xi_n^2]\to 0$. This is impossible, if (57) holds.

STEP 3: If the asset pricing restriction (3) holds, then $\bar{\mu}_\Gamma(\mathcal{J}^*)=1$. If (3) holds, it follows that $\mu_n=B_n\lambda$, P-a.s., for all n, for μ_Γ-almost all sequences (γ_i), where $\lambda=\nu+E\left[f_1|\mathcal{F}_0\right]$. Then, for any portfolio sequence (p_n), we get $E[p_n|\mathcal{F}_0]=R_0C(p_n)+\alpha_n'B_n\lambda$. From Assumption APR.2 (iv) and boundedness of $E\left[f_1|\mathcal{F}_0\right]$, it follows that λ is bounded on Ω. Moreover, we have: $V\left[p_n|\mathcal{F}_0\right]=(B_n'\alpha_n)'V[f_1|\mathcal{F}_0](B_n'\alpha_n)+\alpha_n'\Sigma_{\varepsilon,1,n}\alpha_n\geq eig_{\min}(V[f_1|\mathcal{F}_0])\left\|B_n'\alpha_n\right\|^2$, where $eig_{\min}(V[f_1|\mathcal{F}_0])>0$, P-a.s.. Then, Conditions (i) and (ii) in the definition of set $\mathcal{J}^*$ follow, for μ_Γ-almost any sequence (γ_i), that is, $\mu_\Gamma(\mathcal{J}^*)=\bar{\mu}_\Gamma(\mathcal{J}^*)=1$.

STEP 4: If the asset pricing restriction (3) does not hold, then $\bar{\mu}_\Gamma(\mathcal{J}^*)=0$. To prove that $\bar{\mu}_\Gamma(\mathcal{J}^*)=0$, we show that $\mathcal{J}^*\cap\mathcal{J}\cap\mathcal{J}_1=\emptyset$, where $\mathcal{J}$ and $\mathcal{J}_1$ are the sets with μ_Γ-measure 1 defined in Assumption APR.3 and in Lemma 1, respectively. The proof is by contradiction. Let us assume that sequence (γ_i) is in set $\mathcal{J}^*\cap\mathcal{J}\cap\mathcal{J}_1$. By following the same arguments as in CR on p. 1292 and 1295, we have:

$$\mu_n'\Sigma_n^{-1}\mu_n=\sup_{p_n\in\mathcal{P}_n:C(p_n)=0}E[p_n|\mathcal{F}_0]^2/V[p_n|\mathcal{F}_0], \tag{58}$$

$$\Sigma_n^{-1}\geq eig_{\max}(\Sigma_{\varepsilon,1,n})^{-1}[I_n-B_n(B_n'B_n)^{-1}B_n'], \tag{59}$$

P-a.s.. Let us prove that the RHS of (58) is upper bounded uniformly in n. We use Hilbert space methods as in Hansen and Richard (1987) applied to the conditional economy generated by the countable collection of assets (γ_i). Let $\langle p, q\rangle_{\mathcal{F}_0} = E[pq|\mathcal{F}_0]$ and $\|p\|_{\mathcal{F}_0} = \langle p, p\rangle^{1/2}_{\mathcal{F}_0}$ be the conditional scalar product and norm in the linear space of $\mathcal{F}_1$-measurable random variables, which are square integrable conditionally to $\mathcal{F}_0$. Conditional convergence of (p_n) to p is defined as $\|p_n - p\|_{\mathcal{F}_0} \overset{a.s.}{\to} 0$ for $n \to \infty$. Conditional Cauchy sequences are defined similarly. Since $(\gamma_i) \in \mathcal{J}^*$, Condition (ii) is satisfied for any portfolio sequence in $\mathcal{P}$. This implies that Condition (iii): If $E[p_n^2|\mathcal{F}_0] \overset{a.s.}{\to} 0$, then $C(p_n) \overset{a.s.}{\to} 0$, holds for any portfolio sequence (p_n) in $\mathcal{P}$. Indeed, suppose that (p_n) is such that $E[p_n^2|\mathcal{F}_0] \overset{a.s.}{\to} 0$ but $C(p_n)$ does not converge to 0 $a.s.$. Define the new portfolio sequence (p_n'), such that $p_n' = p_n$ if $C(p_n) \geq 0$, and $p_n' = -p_n$ otherwise. Then, portfolio sequence (p_n') violates Condition (ii), which is impossible. Condition (iii) implies conditional continuity of function $C(\cdot)$ at the zero payoff in $\mathcal{P}$, and corresponds to Assumption 2.3 in Hansen and Richard (1987). Now, by using Condition (iii), we can extend the cost function $C(\cdot)$ to the linear space $\bar{\mathcal{P}}$, that is the conditional completion of $\mathcal{P}$ w.r.t. the limits of conditional Cauchy sequences. Indeed, let $p \in \bar{\mathcal{P}}$, and let (p_n) be a conditional Cauchy sequence in $\mathcal{P}$ converging conditionally to p. Then, $C(p_n)$ is a Cauchy sequence in $\mathbb{R}$, P-a.s.. By the completeness property of $\mathbb{R}$, this Cauchy sequence converges to a unique value, P-a.s., which we define as $C(p)$. For any $p \in \bar{\mathcal{P}}$, random variable $C(p)$ is $\mathcal{F}_0$-measurable by Theorem 20.A in Halmos (1950). This extension of the function $C(\cdot)$ on $\bar{\mathcal{P}}$ is conditionally linear and conditionally continuous at the zero payoff. By Theorem 2.1 in Hansen and Richard (1987), there exists a $\mathcal{F}_1$-measurable random variable c such that $E[c^2|\mathcal{F}_0] < \infty$ and $C(p) = E[cp|\mathcal{F}_0]$, P-a.s., for any portfolio $p \in \bar{\mathcal{P}}$. This property is the conditional analogue of the Riesz Representation Theorem. Any portfolio $p \in \bar{\mathcal{P}}$ can be written as $p = \pi_0 + \pi_1 c + \tilde{p}$, where π_0 and π_1 are $\mathcal{F}_0$-measurable, and $\tilde{p}$ is conditionally orthogonal to 1 and c, namely, $E[\tilde{p}|\mathcal{F}_0] = E[c\tilde{p}|\mathcal{F}_0] = 0$. If the portfolio p has zero cost, i.e., $C(p) = 0$, then $p = \pi_0\left(1 - E[c|\mathcal{F}_0]E[c^2|\mathcal{F}_0]^{-1}c\right) + \tilde{p} =: \pi_0 p^* + \tilde{p}$. The payoff p^* is the residual of the conditional projection of the constant payoff 1 on the payoff c. Since the component $\tilde{p}$ contributes to the conditional variance of portfolio p but not to its conditional mean, we deduce that for any portfolio $p \in \bar{\mathcal{P}}$ such that $C(p) = 0$, we get:

$$E[p|\mathcal{F}_0]^2/V[p|\mathcal{F}_0] \leq E[p^*|\mathcal{F}_0]^2/V[p^*|\mathcal{F}_0] =: \rho^2 < \infty, \tag{60}$$

P-a.s. (see CR, Corollary 1, for a similar result in their unconditional framework). From (58), (59), and (60), we get: $\rho^2 eig_{\max}(\Sigma_{\varepsilon,1,n}) \geq \mu_n'\left(I_n - B_n(B_n'B_n)^{-1}B_n'\right)\mu_n = \min_{\lambda\in\mathbb{R}^K}\|\mu_n - B_n\lambda\|^2 = \min_{\nu\in\mathbb{R}^K}\|A_n - B_n\nu\|^2 =$

$\min_{\nu \in \mathbb{R}^K} \sum_{i=1}^{n} [a(\gamma_i) - b(\gamma_i)'\nu]^2$, for any $n \in \mathbb{N}$, P-a.s.. Hence, we deduce that $\xi_n = \min_{\nu \in \mathbb{R}^K} \frac{1}{n} \sum_{i=1}^{n} [a(\gamma_i) - b(\gamma_i)'\nu]^2$ is such that: $\xi_n \leq \rho^2 \frac{1}{n} eig_{\max}(\Sigma_{\varepsilon,1,n})$, for any n, P-a.s. Since $(\gamma_i) \in \mathcal{J}$, from Assumption APR.3, the RHS converges in L^2 to 0. Then, we get $E[\xi_n^2] \to 0$ as $n \to \infty$. However, since the asset pricing restriction (3) does not hold and $(\gamma_i) \in \mathcal{J}_1$, we know from Inequality (57) that $E[\xi_n^2]$ is bounded away from 0, and we get a contradiction.

Appendix 6 Check of assumptions under block dependence

In this appendix, we verify that the eigenvalue condition in Assumption APR.3, and the cross-sectional/time-series dependence and CLT conditions in Assumptions A.1-A.5, are satisfied under a block-dependence structure in a time-invariant and serially i.i.d. framework. We start by providing the main result (Section A.6.1), we prove it (Section A.6.2), and then prove two auxiliary lemmas (Sections A.6.3 and A.6.4).

A.6.1 Main result

Let us assume that:

BD.1 The errors $\varepsilon_t(\gamma)$ are i.i.d. over time with $E[\varepsilon_t(\gamma)] = 0$ and $E[\varepsilon_t(\gamma)^3] = 0$, for all $\gamma \in [0, 1]$. For any n, there exists a partition of the interval $[0, 1]$ into $J_n \leq n$ subintervals $I_1, ..., I_{J_n}$, such that $\varepsilon_t(\gamma)$ and $\varepsilon_t(\gamma')$ are independent if γ and γ' belong to different subintervals, and $J_n \to \infty$ as $n \to \infty$.

BD.2 The blocks are such that $n \sum_{m=1}^{J_n} B_m^2 = O(1)$, $n^{3/2} \sum_{m=1}^{J_n} B_m^3 = o(1)$, where $B_m = \int_{I_m} dG(\gamma)$.

BD.3 The factors (f_t) and the indicators $(I_t(\gamma))$, $\gamma \in [0, 1]$, are i.i.d. over time, mutually independent, and independent of the errors $(\varepsilon_t(\gamma))$, $\gamma \in [0, 1]$.

BD.4 There exists a constant M such that $\|f_t\| \leq M$, P-a.s.. Moreover, $\sup_{\gamma \in [0,1]} E[|\varepsilon_t(\gamma)|^6] < \infty$, $\sup_{\gamma \in [0,1]} \|\beta(\gamma)\| < \infty$ and $\inf_{\gamma \in [0,1]} E[I_t(\gamma)] > 0$.

The block-dependence structure as in Assumption BD.1 is satisfied for instance when there are unobserved industry-specific factors independent among industries and over time, as in Ang, Liu, and Schwarz (2008). In empirical applications, blocks can match industrial sectors. Then, the number J_n of blocks amounts to a couple of dozens, and the number of assets n amounts to a couple of thousands. There are approximately nB_m assets in block m, when n is large. In the asymptotic analysis, Assumption BD.2 on block sizes and block number requires that the largest block size shrinks with n and that there are not too many large blocks, i.e., the partition in independent blocks is sufficiently fine grained asymptotically. Within blocks, covariances do not need to vanish asymptotically.

Lemma 14 *Let Assumptions BD.1-4 on block dependence and Assumptions SC.1-SC.2 on random sampling hold. Then, Assumptions APR.3, A.1, A.2, A.3, A.4 (with any $\bar{q} \in (0,1)$ and $\bar{\delta} \in (1/2,1)$) and A.5 are satisfied.*

The proof of Lemma 14 uses a result on almost sure convergence in Stout (1974), a large deviation theorem based on the Hoeffding inequality in Bosq (1998), and CLTs for martingale difference arrays in Davidson (1994) and White (2001).

Instead of a block structure, we can also assume that the covariance matrix is full, but with off-diagonal elements vanishing asymptotically. We could also accommodate weak serial dependence and conditioning information. In those settings, we can carry out similar checks, although at the cost of increased notational complexity.

A.6.2 Proof of Lemma 14

A.6.2.1 Assumption APR.3

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

$$eig_{\max}(\Sigma_{\varepsilon,1,n}) \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |Cov[\varepsilon_t(\gamma_i), \varepsilon_t(\gamma_j)]| \leq C \max_{m=1,\dots,J_n} \sum_{j=1}^{n} 1\{\gamma_j \in I_m\} \tag{61}$$

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

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

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.6.2.2 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.6.2.3 Assumption A.2

Let us consider condition a). In the time-invariant case under BD.1 and BD.3, we have $S_{ij} = \sigma_{ij} Q_x$ and $v_3 = w_i b_i$, where $Q_x = E\left[x_t^2\right]$. Then, Assumption A.2 a) is equivalent to $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} w_i \tau_i Y_{i,T} \otimes b_i \Rightarrow N(0, S_b)$, where $S_b := \lim_{n\to\infty} E\left[\frac{1}{n}\sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} \sigma_{ij}(Q_x \otimes b_i b_j')\right]$. 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) $\sup_{t=1,\ldots,T} E[\|\xi_{n,t}\|^{2+\delta}] = O(1)$, for some $\delta > 0$.

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

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

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:

$$V[\zeta_{n,T}] = \frac{1}{T^2}\sum_t V[\xi_{n,t}^2] + \frac{1}{T^2}\sum_{t\neq s} Cov\left(\xi_{n,t}^2, \xi_{n,s}^2\right), \tag{62}$$

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 $\alpha_{ij} = w_i w_j \frac{\tau_i\tau_j}{\tau_{ij}} b_i b_j Q_x$. 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 $\sum_{i,j,l} E\left[1\{\gamma_i,\gamma_j \in I_m\}\right] E\left[1\{\gamma_i,\gamma_l \in I_p\}\right] = O\left(nB_mB_p + n^2(B_m^2 B_p + B_m B_p^2) + n^3 B_m^2 B_p^2\right)\right)$, 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:

$$Cov(\xi_{n,t}^2, \xi_{n,s}^2) = o(1), \tag{63}$$

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:

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

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 $\sum_{i,j,k,l}\sum_{m=1}^{J_n} E[1\{\gamma_i,\gamma_j,\gamma_k,\gamma_l \in I_m\}] = O\left(\sum_{m=1}^{J_n}(nB_m + n^2B_m^2 + n^3B_m^3 + n^4B_m^4)\right)$ and $\sum_{i,j,k,l}\sum_{m,p=1}^{J_n} E[1\{\gamma_i,\gamma_j \in I_m\}]E[1\{\gamma_k,\gamma_l \in I_p\}] = O\left(\sum_{m,p=1}^{J_n}(n^2B_mB_p + n^3B_m^2B_p + n^4B_m^2B_p^2)\right)$, 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, $n \max_{m=1,\ldots,n} B_m^2 = O(1)$, and we get $E\left[V(\xi_{n,t}^2|\mathcal{G}_n)\right] = O(1)$.

Thus, we have shown:

$$
V(\xi_{n,t}^2) = O(1), \tag{64}
$$

uniformly in t.

From (62), (63) and (64), we get $V[\zeta_{nT}] = o(1)$, and condition (ii) follows. From (64) and by using $E[\xi_{n,t}^2] = O(1)$, condition (iii) follows for $\delta = 2$. Finally, condition (iv) follows from $\frac{1}{n}\sum_{i,j} w_iw_j\frac{\tau_i\tau_j}{\tau_{ij}}\sigma_{ij}b_ib_j' = (1+\lambda'V[f_t]\lambda)^{-2}\frac{1}{n}\sum_{i,j}\frac{1}{\tau_{ij}}\frac{\sigma_{ij}}{\sigma_{ii}\sigma_{jj}}b_ib_j'$ and the next Lemma 15.

Lemma 15 *Under Assumptions BD.1-BD.4:* $\frac{1}{n}\sum_{i,j}\frac{1}{\tau_{ij}}\frac{\sigma_{ij}}{\sigma_{ii}\sigma_{jj}}b_ib_j' \to L$, *$P$-a.s., where:*

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

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.6.2.4 Assumption A.3

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

A.6.2.5 Assumption A.4

We have to show that $\max_i \sum_j \|S_{ij}\|^{\bar{q}} = O_p(n^{\bar{\delta}})$, for any $\bar{q} \in (0,1)$ and $\bar{\delta} > 1/2$. From $S_{ij} = \sigma_{ij} Q_x$, and an argument similar to (61):

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

for any $\bar{q} > 0$. 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 $\varepsilon_n := n^{\bar{\delta}}$, with $\bar{\delta} > 1/2$. Then:

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

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

A.6.2.6 Assumption A.5

In the time-invariant $i.i.d.$ case we have $S_{ii,T} = \sigma_{ii} \hat{Q}_{x,i}$ and $S_{ij} = \sigma_{ij} Q_x$. Then, Assumption A.5 boils down to $\Upsilon_{nT} := \frac{1}{\sqrt{n}} \sum_i w_i \tau_i^2 \left[Y_{i,T} \otimes Y_{i,T} - \tilde{S}_{ii,T} \right] \Rightarrow N(0, \Omega)$, as $n, T \to \infty$, where $\widetilde{S}_{ii,T} = \sigma_{ii} vec(\hat{Q}_{x,i})$ and $\Omega = \lim_{n\to\infty} E\left[\frac{1}{n} \sum_{i,j} w_i w_j \frac{\tau_i^2 \tau_j^2}{\tau_{ij}^2} \sigma_{ij}^2 \right] \left[Q_x \otimes Q_x + (Q_x \otimes Q_x) W_{K+1} \right]$. Let us denote by $\mathcal{H} = \sigma\left((f_t), (I_t(\gamma)), \gamma \in [0,1], \gamma_i, i = 1, 2, \dots\right)$ 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

$$\Upsilon_{nT} \Rightarrow N(0, \Omega), \quad n, T \to \infty, \tag{65}$$

P-a.s.. 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) $\lim\limits_{n,T\to\infty} \frac{1}{J_n} \sum\limits_m V[W_{m,nT}|\mathcal{H}] = \Omega$, P-a.s, and

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

To show (i), we use:

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

where $\sum\limits_{i,j \in I_m}$ 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:

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

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

$\frac{1}{n}\sum_{i,j}\frac{\tau_i\tau_j}{\tau_{ij}^2}\frac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}} \to L$, P-a.s., where $L = \lim_{n\to\infty} E\left[\frac{1}{n}\sum_{i,j}\frac{\tau_i\tau_j}{\tau_{ij}^2}\frac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}}\right]$, which is proved by similar arguments as Lemma 15.

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,

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

By the independence property, the non-zero terms $E\left[\varepsilon_{i,t_1}\ldots\varepsilon_{i,t_6}|\gamma_i\right]$ 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 16.

Lemma 16 *Under Assumptions BD.1-BD.4:* $\frac{1}{n^{3/2}}\sum_{m=1}^{J_n}\left(\sum_i 1\{\gamma_i\in I_m\}\right)^3 \to 0$, *$P$-a.s.*

STEP 2: We show that (65) implies the asymptotic normality condition in Assumption A.4. Indeed, from (65) 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.6.3 Proof of Lemma 15

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:

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

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.6.4 Proof of Lemma 16

The proof is similar to the one of Lemma 15 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 7 Monte-Carlo experiments

In this appendix, we report the results of Monte Carlo experiments to investigate the finite sample behaviour of our estimators and test statistics (Section A.7.1) and the accuracy of the CLT asymptotic approximations underlying Assumption A.2a) (Section A.7.2).

A.7.1 Finite sample behaviour of estimators and test statistics

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.7.1.1 Balanced panel

We simulate S datasets of excess returns from a time-invariant 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 sample of our empirical analysis 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 6.

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

$$R_{i,t}^s = a_i^s + b_i^s f_t + \varepsilon_{i,t}^s, \text{ for } i = 1, ..., n, \text{ and } t = 1, ..., T, \tag{66}$$

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 3.5 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 5 in the paper), and the thresholded variance estimator $\tilde{\Sigma}_\xi$ (see Proposition 6) 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 4 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 5. 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 4. The coverage of the confidence intervals is close to the nominal level 95% across the considered designs and sample sizes.

In Table 5, 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 6 and 7 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 6. 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 4: 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 5: 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 6: 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 7: 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.7.1.2 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:

$$R^s_{i,t} = a^s_i + b^s_i f_t + \varepsilon^s_{i,t}, \text{ if } I^s_{i,t} = 1, \text{ for } i = 1, ..., n, \text{ and } t = 1, ..., T, \tag{67}$$

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

In Tables 8 and 9, 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 8 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 (rounded). 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 9, 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 10, 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 4, 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 9. Similar comments apply for Table 11, 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 12 and 13, 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 9 is around 200 months, and in Tables 6 and 7, 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 8: 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 9: Operative time-series sample size

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

Table 10: 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 11: 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 12: Test of the null hypothesis $\mathcal{H}_0 :\ a(\gamma) = 0$, 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 13: Test of the null hypothesis $\mathcal{H}_0 :\ a(\gamma) = b(\gamma)\nu$, 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

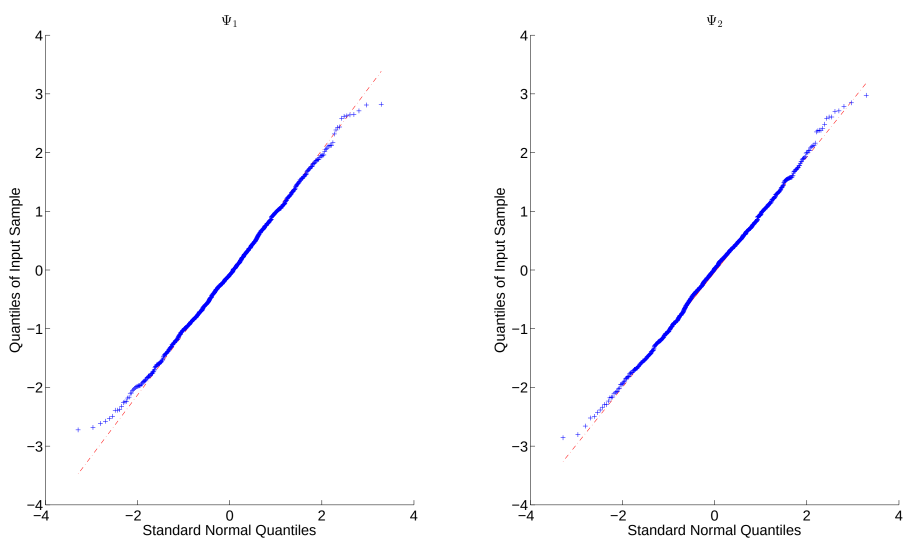
A.7.2 The CLT in Assumption A.2a)

In this section, we provide simulation exercises to assess the empirical validity of the CLT in Assumption A.2a). We simulate S datasets of error terms $\varepsilon_{i,t}$ from a time-invariant one-factor model (CAPM). At each simulation $s = 1, ..., S$, we randomly draw $n \leq 9,904$ assets from the sample of our empirical analysis, and we build a block diagonal matrix Ω^s as described in the previous section. For each s, 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 perform the exercise for the unbalanced case. We fix the maximal sample size $T = 546$ as in the empirical application. In the time-invariant one-factor framework, the statistic in Assumption A.2a) reduces to $\frac{1}{\sqrt{n}} \sum_i w_i \tau_i Q_{x,i}^{-1} Y_{i,T} b_i$ with asymptotic variance $S_{v_3} = \lim_{n \to \infty} E\left[\frac{1}{n} \sum_{i,j} w_i w_j \frac{\tau_i \tau_j}{\tau_{ij}} S_{Q,ij} b_i b_j\right]$.

At each simulation, we compute the 2×1 vector $\Psi^s = \left(S^s_{v3}\right)^{-1/2} \frac{1}{\sqrt{n}} \sum_i w^s_i \tau^s_i \left(Q^s_{x,i}\right)^{-1} Y^s_{i,T} b^s_i$ with $Y^s_{i,T} = \frac{1}{\sqrt{T}} \sum_t I^s_{i,t} x_t \varepsilon^s_{i,t}$ and $S^s_{v3} = \frac{1}{n} \sum_{i,j} w^s_i w^s_j \frac{\tau^s_i \tau^s_j}{\tau^s_{ij}} S^s_{Q,ij} b^s_i b^s_j$, where scalars w^s_i, τ^s_i, τ^s_{ij}, b^s_i, matrices $Q^s_{x,i}$, $S^s_{Q,ij}$, and indicator processes $\left(I^s_{i,t}\right)$ for draw s are those estimated for assets i and j in the empirical analysis.

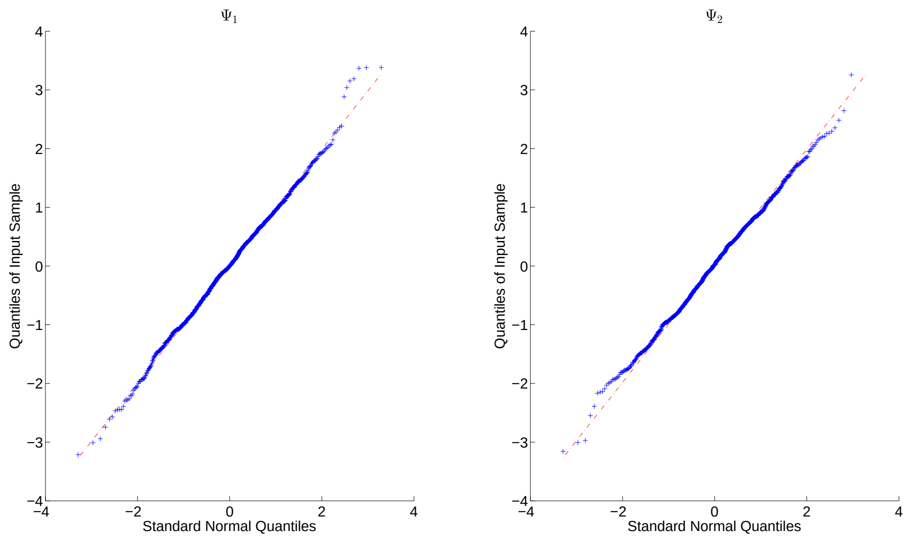
Figures 3 and 4 compare the univariate distributions of the two components of simulated vectors $\Psi^s = [\Psi^s_1, \Psi^s_2]' \in \mathbb{R}^2$, $s = 1, ..., 1,000$, with the standard normal distribution through Q-Q plots. The cross-sectional size is $n = 1,000$ in Figure 3, and $n = 3,000$ in Figure 4. Figures 3 and 4 show that the finite sample distributions are well approximated by the asymptotic Gaussian distributions already for $n = 1,000$. This finding suggests that the possible heavy tails in the cross sectional distribution of asset characteristics should not affect the validity of our CLT assumptions.

Figure 3: Q-Q plots of the simulated components of Ψ for $n = 1,000$



The figure compares the finite-sample distributions of the two components of vector Ψ (right panel and left panel) with the standard normal distribution. We estimate the finite-sample distributions with an unbalanced panel of $n = 1,000$ individual stocks in the Monte-Carlo exercise.

Figure 4: Q-Q plots of the simulated components of Ψ for $n = 3,000$



The figure compares the finite-sample distributions of the two components of vector Ψ (right panel and left panel) with the standard normal distribution. We estimate the finite-sample distributions with an unbalanced panel of $n = 3,000$ individual stocks in the Monte-Carlo exercise.

Appendix 8 Long-only factors

In this section, we estimate a time-invariant model with long-only factors derived from the FF methodology. We use the market factor (denoted by m) available on Ken French's website, then we build the long-only factors from the six FF research portfolios available on Ken French's website. The "Small" factor (denoted by s) is the average excess return of the three small portfolios, and the "Value" factor (denoted by h) is the average excess return of the two value portfolios. The long-only factors should be more immune to market imperfections (e.g., transaction costs). We estimate the time-invariant three-factor model using the individual stocks, the 25 FF and the 44 Indu. portfolios as base assets. The annualized estimates of the risk premia and the components of ν are reported in Table 14. The estimated risk premium for the market factor is positive across the three universes of assets, albeit not statistically significant at the 5% level for the 25 FF portfolios. The small long-only factor is positively and significantly remunerated for the individual stocks (9.24%) and the 25 FF portfolios (9.12%). It is not significantly remunerated for the 44 Indu. portfolios. The risk premium on the value factor is positive and not significant for the individual stocks and the 44 Indu. portfolios. We observe that the estimates of risk premia for the 25 FF portfolios are less accurate than the estimates that we get in Table 1. Moreover, we get close to zero estimates for the components of vector ν when we use the 25 FF portfolios as base assets. On the contrary, we get non-zero estimates when we use the individual stocks and the 44 Indu. portfolios. In particular, these datasets yield negative and significant estimates of ν_h (-4.06% and -4.37%). Thus, the estimates of the time-invariant risk premia are close to the average of the factors only for the 25 FF portfolios. Market imperfections due to rebalancing and short selling are probably not the key drivers in the explanation of why we get non-zero estimated ν in Section 4.2.

Table 14: Estimated annualized risk premia and ν for the time-invariant three-factor model with long-only factors

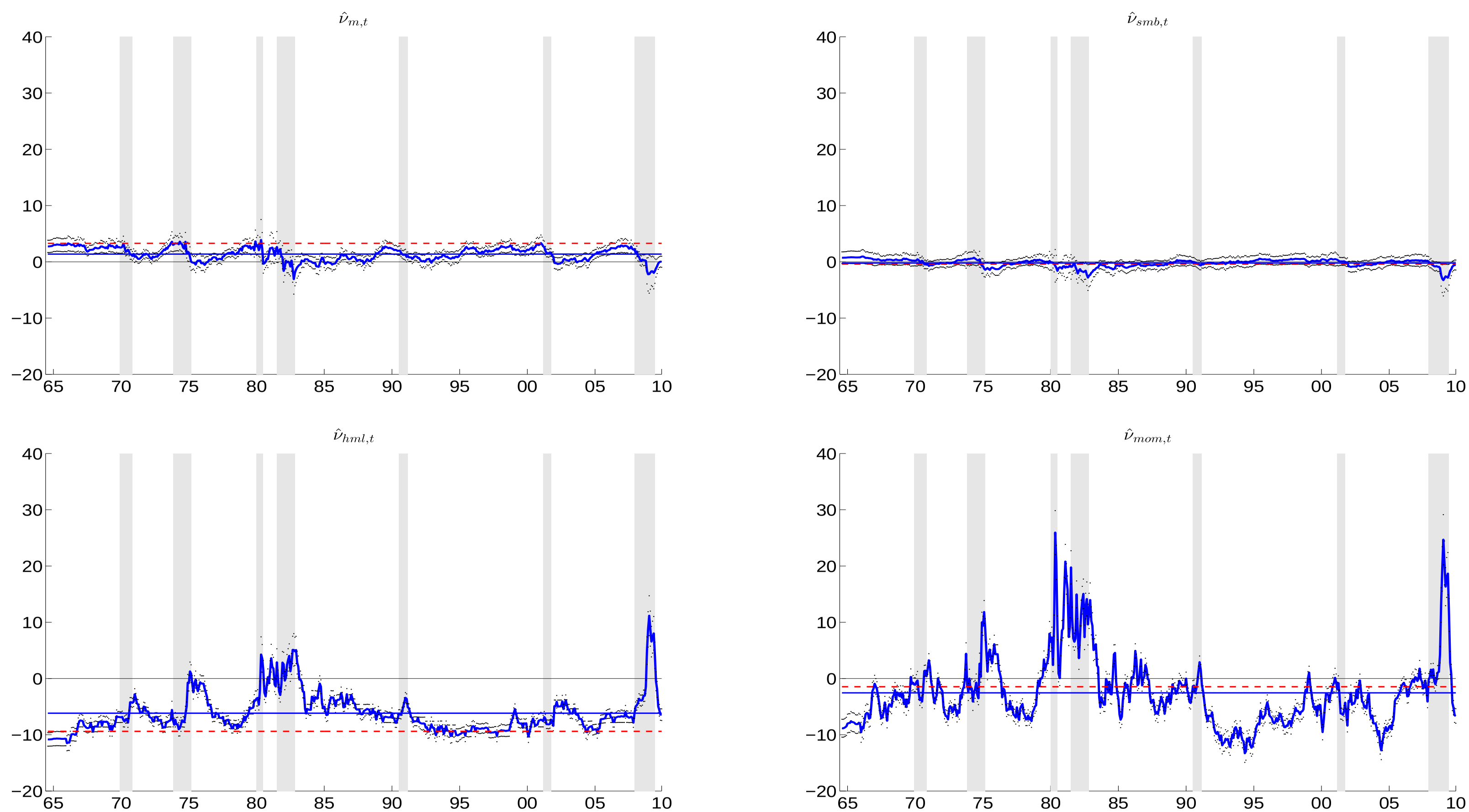
	Stocks ($n = 9,936$)		FF Portfolios ($n = 25$)		Indu. Portfolios ($n = 44$)	
	bias corrected estimate (%)	95% conf. interval	point estimate (%)	95% conf. interval	point estimate (%)	95% conf. interval
($n^{\chi} = 9,846$)						
λ_m	7.49	(2.61, 12.37)	4.72	(-0.22, 9.66)	6.57	(1.60, 11.54)
λ_s	9.24	(2.66, 15.82)	9.12	(2.54, 15.71)	4.69	(-2.27, 11.65)
λ_h	5.46	(-0.09, 11.02)	10.30	(4.70, 15.90)	5.16	(-0.92, 11.23)
ν_m	2.64	(2.14, 3.13)	-0.14	(-0.90, 0.62)	1.72	(0.79, 2.65)
ν_s	0.30	(-0.27, 0.88)	0.19	(-0.08, 0.45)	-4.25	(-6.25, -1.96)
ν_h	-4.06	(-4.50, -3.63)	0.77	(0.04, 1.51)	-4.37	(-6.83, -1.91)

The table contains the estimates of the annualized risk premia and of the components of vector ν for the market (λ_m, ν_m), small (λ_s, ν_s), and value (λ_h, ν_h) long-only factors. We report the bias corrected estimates $\hat{\lambda}_B$ of λ, and $\hat{\nu}_B$ of ν for individual stocks ($n = 9,936$, $n^{\chi} = 9,846$). In order to build the confidence intervals, we use the HAC estimators $\hat{\Sigma}_f$ defined in Section 3.4 and $\tilde{\Sigma}_\nu$ defined in Proposition 5 for $n = 9,936$. When we consider the 25 FF and 44 Indu. portfolios as base assets, we compute the estimates of the variance-covariance matrices $\Sigma_{\lambda,n}$ and $\Sigma_{\nu,n}$ defined in Section 3.3.

Appendix 9 Paths of ν_t for the four-factor model estimated from individual stocks and the 25 FF portfolios

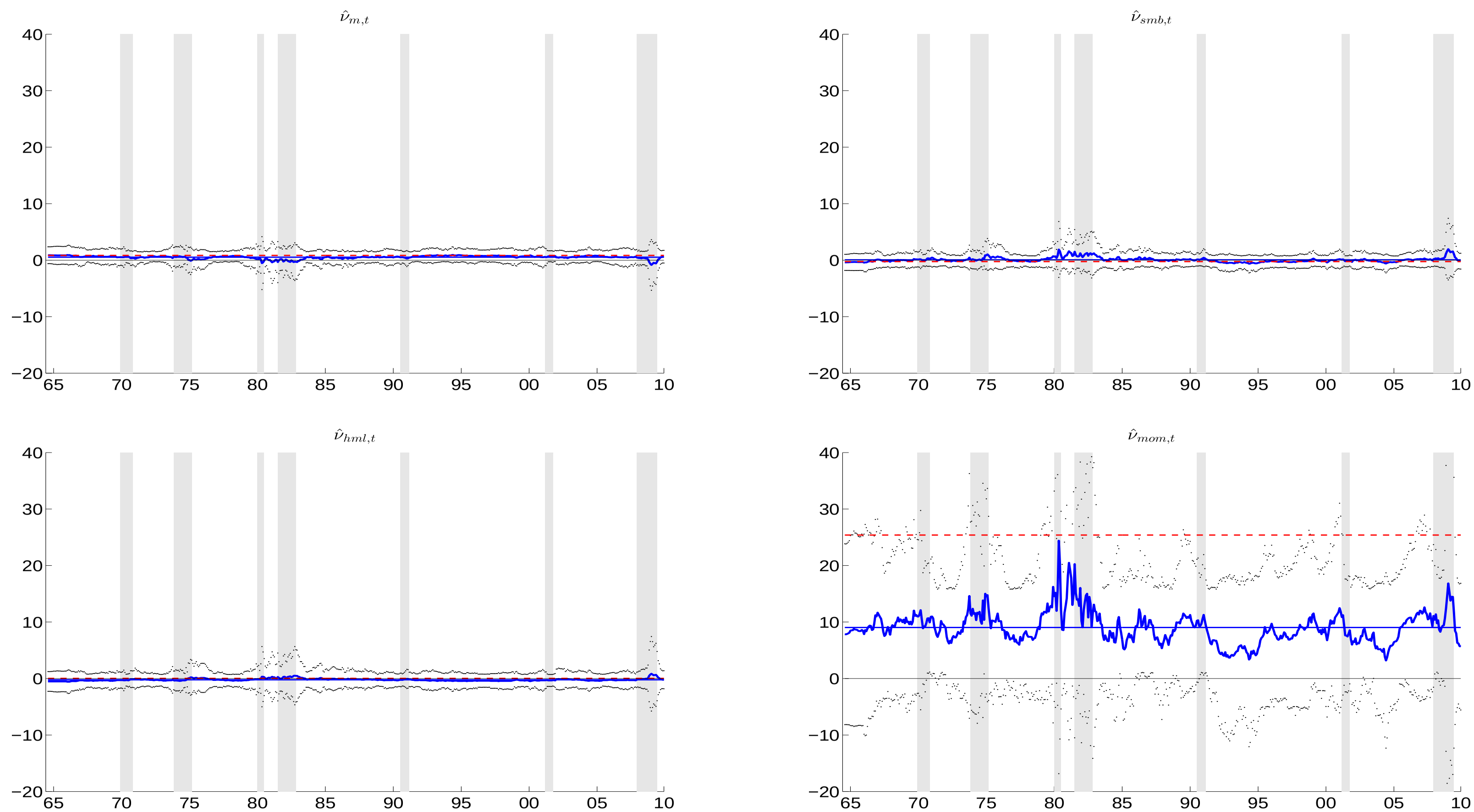
In this appendix, we provide the time-varying paths of $\hat{\nu}_t$ in Figures 5 and 6 for the four factor model, estimated from individual stocks and the 25 FF portfolios. On Figure 6, we see that the path for the momentum factor is not centered around zero and is very imprecisely estimated on the 25 FF portfolios. A first explanation might be the misspecification induced by the ad hoc portfolio aggregation based on size and value sorting and the time-varying specification for the momentum factor sensitivity (see also the theory in A.14.1). A second explanation of the statistical inaccuracy might be the tight factor structure observed by Lewellen, Nagel, and Shanken (2010). The paths for the other combinations of models and base assets move also a lot and are not centered on zero as in Figure 5.

Figure 5: Path of estimated annualized ν_t with $n = 9,936$ in the four-factor model



The figure plots the path of estimated annualized $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$, $\hat{\nu}_{hml,t}$, and $\hat{\nu}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 3,900$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Figure 6: Path of estimated annualized ν_t with $n = 25$ in the four-factor model

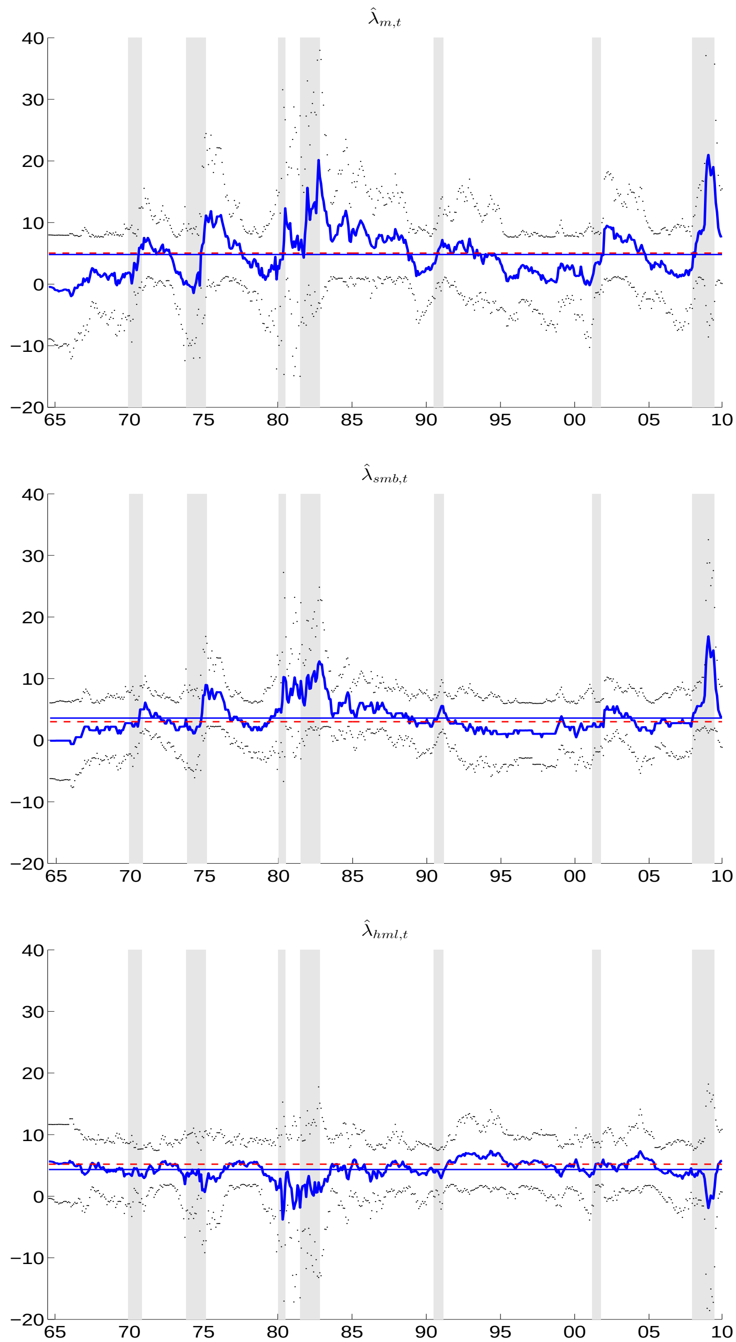


The figure plots the path of estimated annualized $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$, $\hat{\nu}_{hml,t}$, and $\hat{\nu}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. We use the returns of the 25 FF portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).

Appendix 10 Additional figures for the 25 FF portfolios

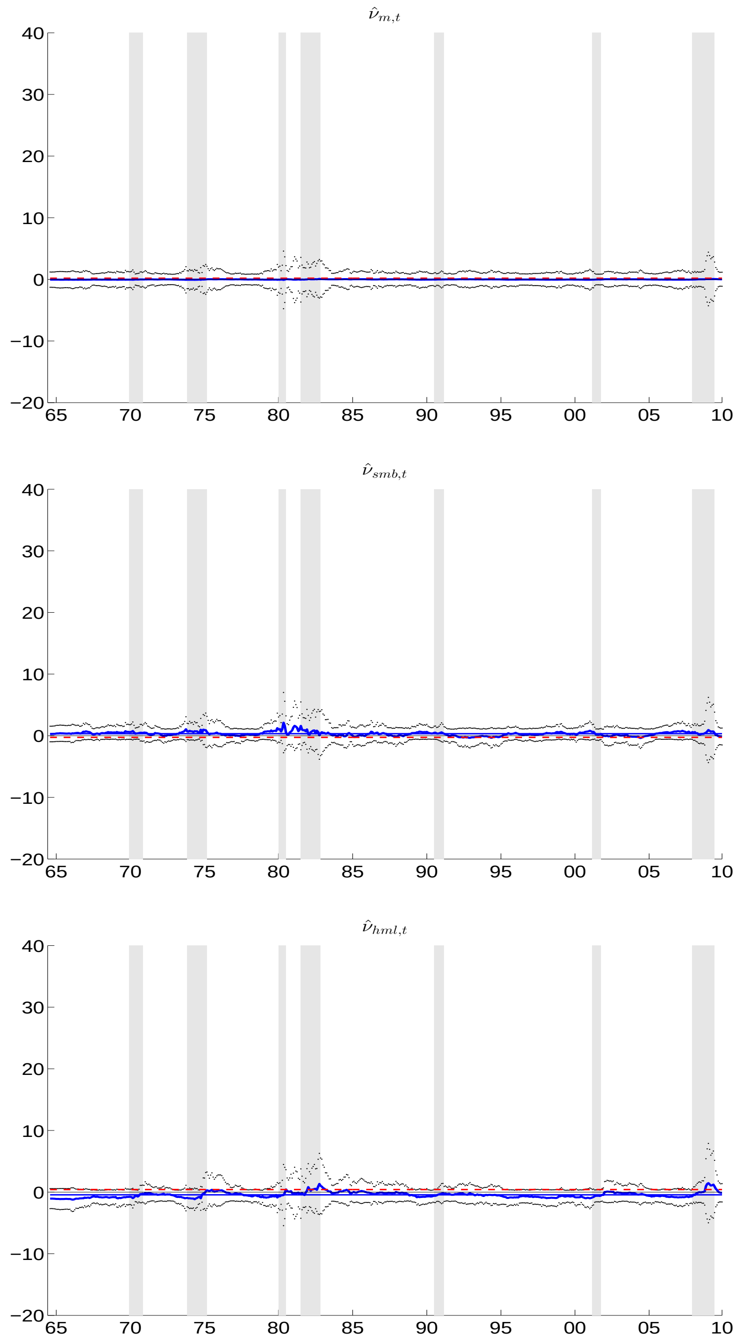
In this appendix, we provide the time-varying paths of $\hat{\lambda}_t$ and $\hat{\nu}_t$ in Figures 7 and 8 for the Fama-French model, estimated from the 25 FF portfolios. The paths of risk premia in the Fama-French model look similar to the corresponding estimates for the four-factor model in Figure 2. In Figure 8, we get paths of $\hat{\nu}_t$ close to zero for the market, size and value factors. The estimates are almost constant and centered on zero, consistent with a time-invariant model and tradable factors, as revealed by the parametric test results in Section 4 of the paper.

Figure 7: Path of estimated annualized risk premia with $n = 25$ in the Fama-French model



The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, and $\hat{\lambda}_{hml,t}$ and their pointwise confidence intervals at 95% probability level in the Fama-French model. We use the returns of the 25 FF portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).

Figure 8: Path of estimated annualized ν_t with $n = 25$ in the Fama-French model

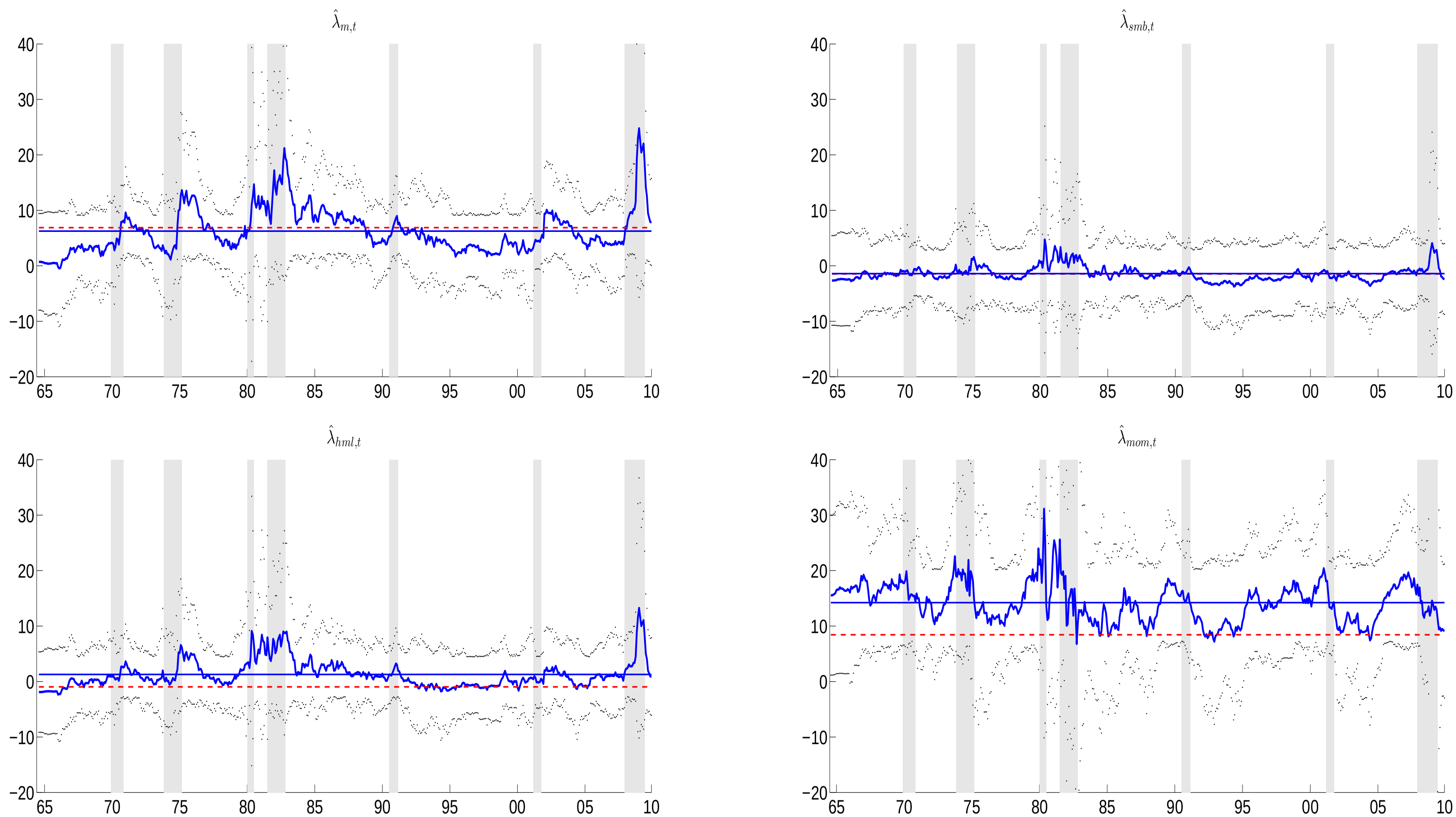


The figure plots the path of estimated annualized risk premia $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$ and $\hat{\nu}_{hml,t}$ and their pointwise confidence intervals at 95% probability level in the Fama-French model. We use the returns of the 25 FF portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).

Appendix 11 Additional figures for the industry portfolios

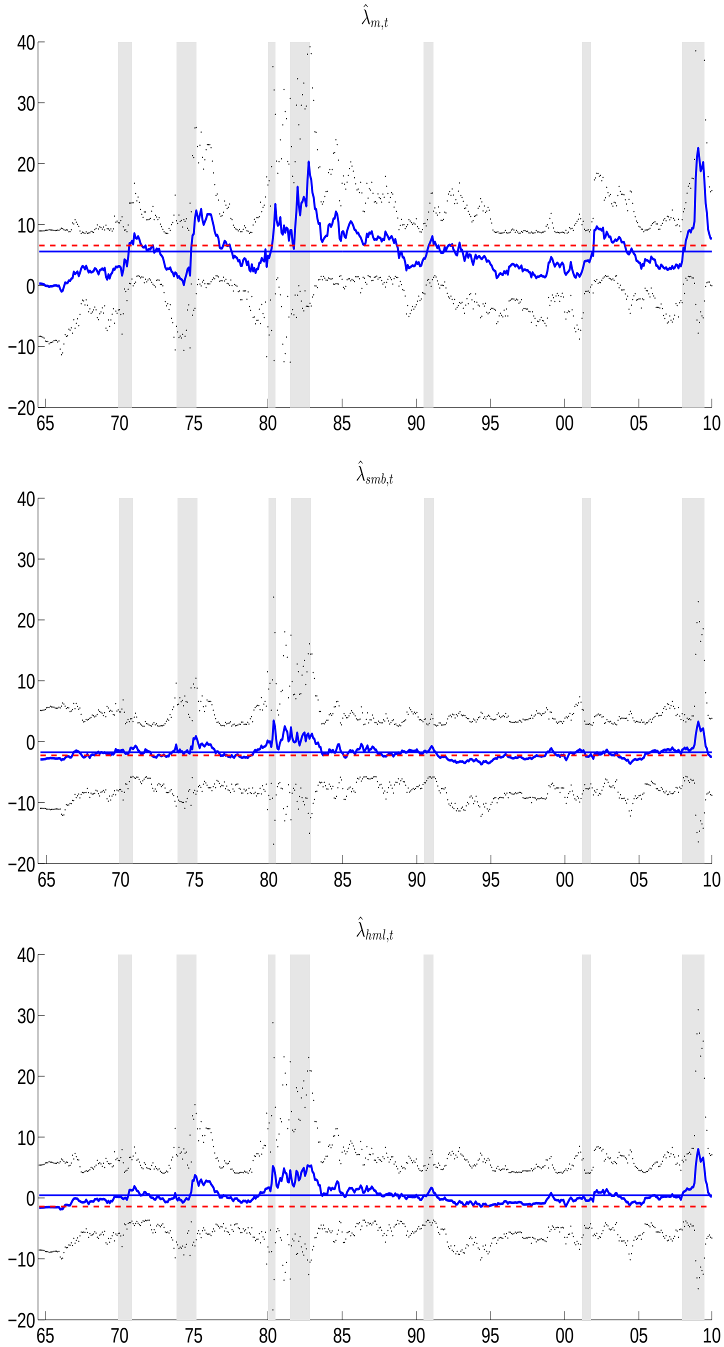
In this appendix, we provide the time-varying paths of $\hat{\lambda}_t$ and $\hat{\nu}_t$ in Figures 9-12 for the four-factor model, and the Fama-French model, estimated from the 44 Indu. portfolios. The factors are as in Section 4 of the paper. The paths look very similar for the market, size and value factors between the two asset pricing models. In Figures 9 and 10, the risk premia for the market, size and value factors feature a counter-cyclical pattern, and they are similar to the paths of risk premia obtained using the individual stocks as base assets (see Figures 1, 5, 46, and 49). The risk premium for the momentum factor is pro-cyclical and similar to that obtained from the individual stocks. In Figures 11 and 12, we get paths of $\hat{\nu}_t$ close to zero only for the market factor. The time-varying results with the 44 Indu. portfolios differ from those with the 25 FF portfolios (see Figures 2, 7, 8, and 6). This finding is similar to the estimation results for time-invariant specifications in Section 4.2.

Figure 9: Path of estimated annualized risk premia with $n = 44$ in the four-factor model



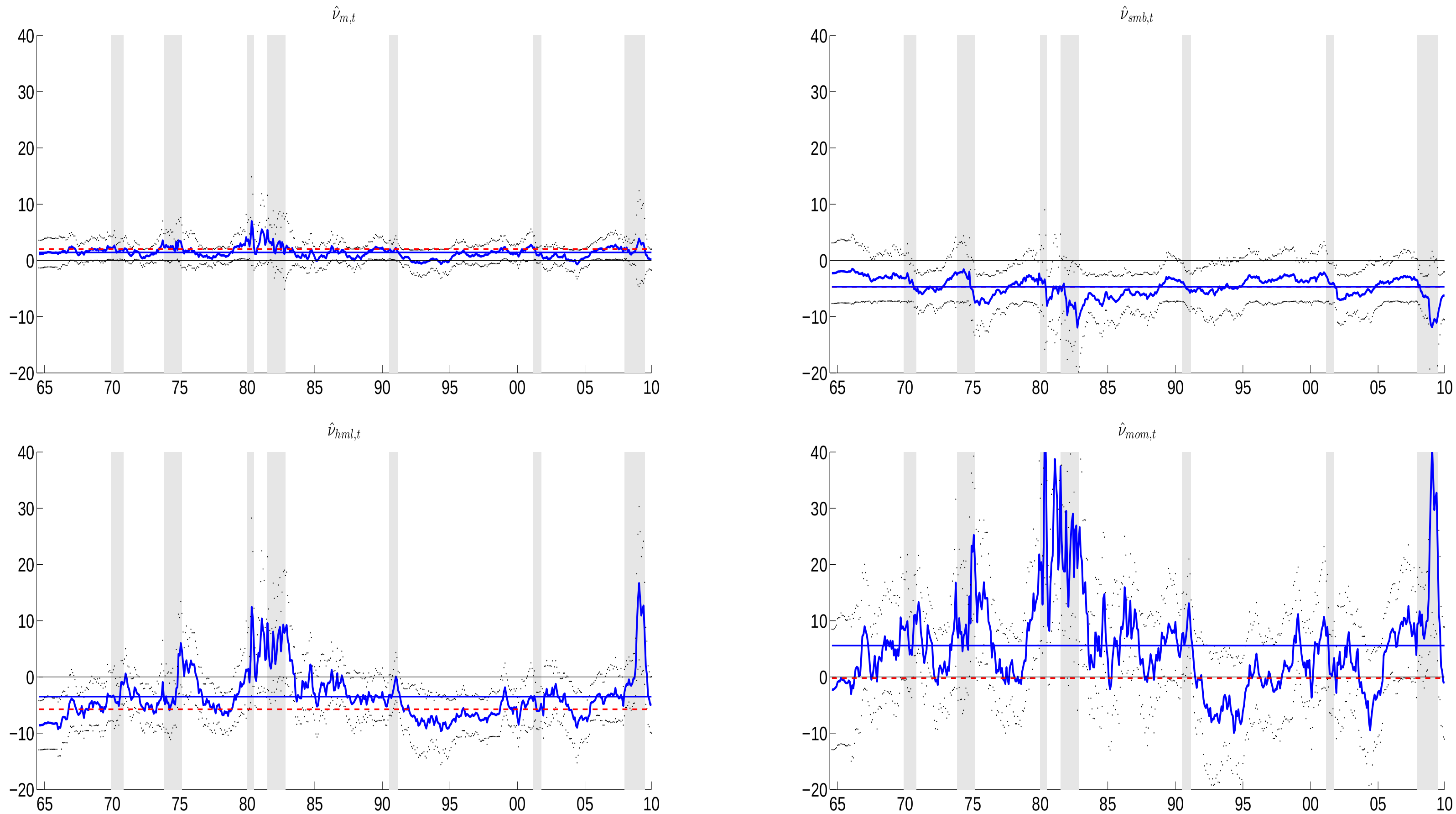
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$, and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider the 44 Indu. portfolios. The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Figure 10: Path of estimated annualized risk premia with $n = 44$ in the Fama-French model



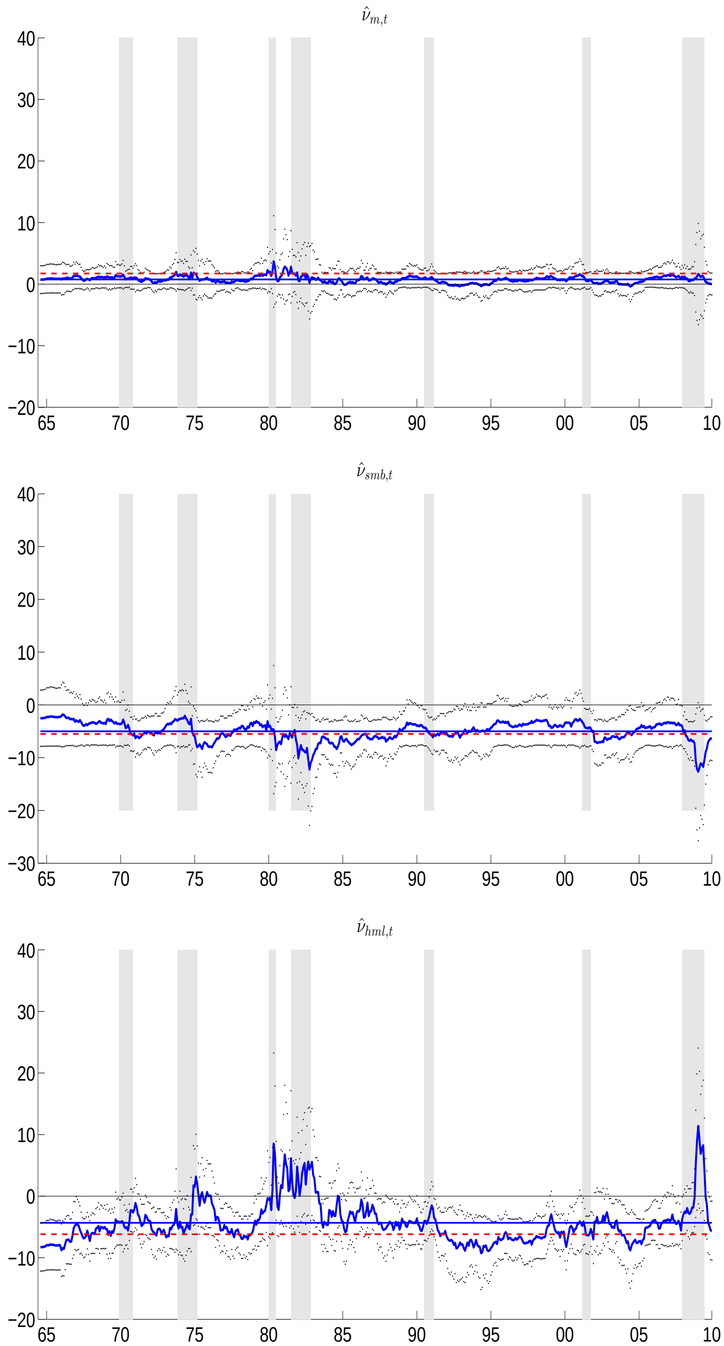
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, and $\hat{\lambda}_{hml,t}$ and their pointwise confidence intervals at 95% probability level in the Fama-French model. We use the returns of the 44 Indu. portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).

Figure 11: Path of estimated annualized ν_t with $n = 44$ in the four-factor model



The figure plots the path of estimated annualized $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$, $\hat{\nu}_{hml,t}$, and $\hat{\nu}_{mom,t}$ and their confidence intervals at 95% probability level in the four-factor model. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider the 44 Indu. portfolios. The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Figure 12: Path of estimated annualized ν_t with $n = 44$ in the Fama-French model



The figure plots the path of estimated annualized $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$ and $\hat{\nu}_{hml,t}$ and their pointwise confidence intervals at 95% probability level in the Fama-French model. We use the returns of the 44 Indu. portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).

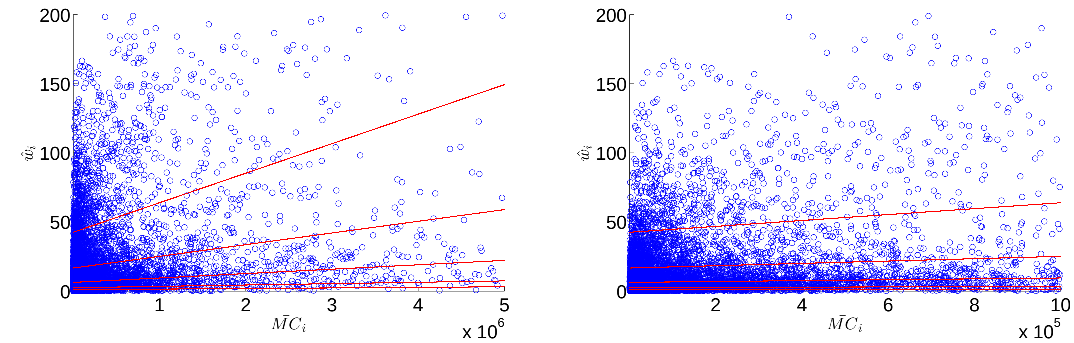
Appendix 12 Value-weighted estimation of time-invariant and time-varying specifications

In this appendix, we show that the weights for the WLS estimator of ν defined in Equation (13) are increasing w.r.t. the size of asset i. We also report some results for value-weighted estimation of time-invariant and time-varying specifications. Here the weights directly account for the size characteristic of the assets through their time-average market capitalisation. These results provide robustness checks that our results are not entirely driven by small stocks. We start the section by an empirical analysis of the weights $\hat{w}_i$. Then, we describe the two-pass methodology for value-weighted estimators before examining the empirical results and comparing them with WLS and OLS outputs. In this appendix, $mc_{i,t}$ denotes the market capitalisation of firm i at month t.

A.12.1 Empirical analysis of the weights $\hat{w}_i$

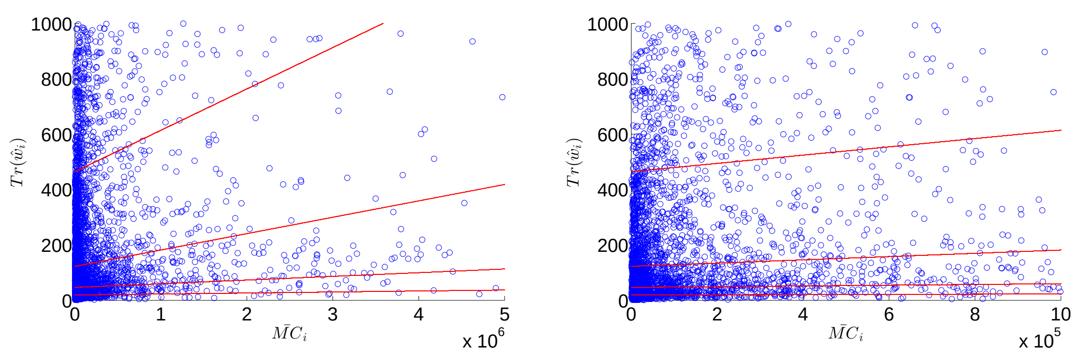
Figure 13 plots the averages over time of size, $\bar{MC}_i = \frac{1}{T_i}\sum_t I_{i,t} mc_{i,t}$, versus weights $\hat{w}_i = \mathbf{1}_i^{\chi}\hat{v}_i^{-1}$ for the time-invariant four-factor model estimated on the $n^{\chi} = 9,902$ stocks. Figure 14 plots the averages over time of size, $\bar{MC}_i$, versus $Tr\left[\hat{w}_i\right]$, where $\hat{w}_i = \mathbf{1}_i^{\chi}\left(diag\left[\hat{v}_i\right]\right)^{-1}$ is estimated by assuming a time-varying four-factor model on the $n^{\chi} = 3,900$ individual stocks. In Figures 13 and 14, we also report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%. All results show that the weights $\hat{w}_i$ are positively related with size. Thus, larger stocks receive larger weights in the second pass of the procedure to get the WLS estimator of ν. We find a similar positive association between the second-pass weights and the average size for the 25 FF and the 44 Indu. portfolios.

Figure 13: $\bar{MC}_i$ **vs** $\hat{w}_i$ **for the time-invariant four-factor model**



The left panel plots the averages over time of the market capitalisation $\bar{MC}_i$ w.r.t. the estimated weights $\hat{w}_i$ defined in Section 3.2 and computed on the time-invariant four-factor model for the $n^{\chi} = 9,902$ individual stocks. The right panel is a zoom for the average market capitalisation below 1×10^6. We report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%.

Figure 14: $\bar{MC}_i$ **vs** $Tr\,[\hat{w}_i]$ **for the time-varying four-factor model**



The left panel plots the averages over time of the market capitalisation $\bar{MC}_i$ w.r.t. $Tr\,[\hat{w}_i]$, where $\hat{w}_i$ are the estimated weights defined in Section 3.2 and computed on the time-varying four-factor model for the $n^{\chi} = 3,900$ individual stocks. The right panel is a zoom for the average market capitalisation below 1×10^6. We report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%.

A.12.2 Value-weighted estimators

We consider the two-pass approach introduced in Section 3.2, but with a value-weighted version of the risk premia estimator thanks to a different estimator for ν. The first pass of the estimation approach and the trimming device remain unchanged. The second pass consists in computing a cross-sectional estimator of ν using a multivariate WLS approach. There, the weights account for the size characteristic of asset i, through its time-average market capitalisation, $\bar{MC}_i = \frac{1}{T_i}\sum_t I_{i,t} mc_{i,t}$. The value-weighted (VW) estimator is

$$\hat{\nu}_{VW} = \hat{Q}_{\beta_3}^{-1}\frac{1}{n}\sum_i w_i \hat{\beta}_{3,i}' \hat{\beta}_{1,i}, \tag{68}$$

where $\hat{Q}_{\beta_3} = \frac{1}{n}\sum_i w_i \hat{\beta}_{3,i}' \hat{\beta}_{3,i}$ and $w_i = \frac{\bar{MC}_i}{\sum_i \bar{MC}_i}$. The weight w_i is a scalar and does not require a first-step estimator because it involves observable data only. The final estimator of the risk premia corresponds to the estimator $\hat{\lambda}_t$ introduced in Section 3.2 replacing Equation (13) with Equation (68). The asymptotic properties of risk premia and cross-sectional estimators remain unchanged w.r.t. the asymptotic results in Section 3.3 up to replacement of the weights in the bias correction terms and asymptotic variances.

A.12.3 Estimation results for time-invariant specifications

Tables 15 and 16 present the time-invariant risk premia estimates with the value-weighted estimator and the corresponding estimates of the components of ν obtained with the cross-sectional estimator $\hat{\nu}_{VW}$ in Equation (68). For comparison purposes, we also present the time-invariant estimates obtained with the equally-weighted (OLS) estimator of risk premia, $\hat{\lambda}_1 = \hat{\nu}_1 + \frac{1}{T}\sum_t f_t$, with the cross-sectional OLS estimator $\hat{\nu}_1$. We consider the $n = 9,936$ individual stocks. The VW results in Tables 15 and 16 are similar to those obtained with the WLS estimator in Table 1. Moreover, the signs of the value-weighted risk premia estimates are the same as those of the OLS estimates. In particular, the estimate of the VW market risk premium (11.84%) is a bit larger than the estimate obtained with the WLS estimator. The OLS estimated market risk premium (2.91%) is smaller than the VW estimated one, and is not statistically significant. The VW estimated value risk premium is negative (-8.99%), and statistically significant. The OLS estimate of the value risk premium is negative as well, albeit not statistically significant. We find a statistically significant negative estimate of ν_{hml} (-13.77%) for the VW estimator in line with the estimate obtained from the WLS estimator (-9.38%), and the OLS estimator (-6.01%). The confidence intervals for the VW and OLS estimators of the components of ν are wider than the confidence intervals for the WLS estimators. This finding is not surprising, given that the weights used in the WLS estimator are optimal at least under cross-sectional independence of the error terms. The block-dependence structure with small correlation within blocks that we find in the data after thresholding the estimated covariance matrix of residuals is not too far from such an exact factor structure.

A.12.4 Estimation results for time-varying specifications

Figures 15 and 16 plot the value-weighted (VW) and equally-weighted (OLS) estimated time-varying paths of risk premia for individual stocks ($n^{\chi} = 3,900$). The VW estimated risk premium for the market features a counter-cyclical pattern, and is more volatile than the OLS estimates. The paths of $\hat{\lambda}_t$ for the size factor look similar in Figures 15 and 16, but their pro-cyclical pattern differs from the WLS estimate (see Figure 1). The VW estimates of the value risk premium are negative and more stable over time than the WLS and OLS estimates (see Figures 1 and 16). The confidence intervals for the VW and OLS estimators are wider than the confidence intervals for the WLS estimators. Figures 17 and 18 plot the value-weighted and equally-weighted estimated paths of ν_t. The paths are away from zero over time, especially for the OLS estimates. We report the value-weighted and equally-weighted estimates of the components of vector ν in Table 17. The estimates of ν differ from the WLS estimates of Table 2. This explains the differences between the VW and OLS estimated paths shown in Figures 15-18 compared to the WLS estimated paths. For instance, the large negative VW and OLS estimates for the impact of default spread on the size factor coefficient (-7.3882 and -7.5468) imply the pro-cyclical pattern observed in Figures 15 and 16. However, the confidence intervals for the VW and OLS estimates of the components of ν are large and in particular much larger than the confidence intervals of the WLS estimates of Table 2. Thus, the observed differences between the estimates obtained with the different weighting schemes may be due to the statistical inaccuracy of the VW and OLS estimates.

Table 15: Value-weighted and equally-weighted estimates of annualized risk premia for the time-invariant models

	Stocks ($n = 9,936$)			
	Value-weighted estimate (VW)		Equally-weighted estimate (OLS)	
	bias corrected estimate (%)	95% conf. interval	bias corrected estimate (%)	95% conf. interval
	Four-factor model			
($n^{\chi} = 9,902$)				
λ_m	11.84	(6.96, 16.72)	2.91	(-1.98, 7.79)
λ_{smb}	3.51	(0.16, 6.87)	4.03	(0.67, 7.38)
λ_{hml}	-8.99	(-12.45, -5.53)	-1.23	(-4.69, 2.24)
λ_{mom}	13.25	(8.65, 17.85)	5.99	(1.39, 10.59)
	Fama-French model			
($n^{\chi} = 9,904$)				
λ_m	10.86	(5.98, 15.30)	2.92	(-1.96, 7.80)
λ_{smb}	3.36	(0.01, 6.72)	3.90	(0.54, 7.25)
λ_{hml}	-8.96	(-12.42, -5.49)	-1.20	(-4.67, 2.26)
	CAPM			
($n^{\chi} = 9,904$)				
λ_m	12.62	(7.74, 17.50)	3.76	(-1.12, 8.64)

The table contains the value-weighted and equally-weighted estimates of annualized risk premia for the market (λ_m), size (λ_{smb}), book-to-market (λ_{hml}), and momentum (λ_{mom}) factors. We report the bias corrected estimates $\hat{\lambda}_B$ of λ for individual stocks ($n = 9,936$). In order to build the confidence intervals for $n = 9,936$, we use the HAC estimator $\hat{\Sigma}_f$ defined in Section 3.4.

Table 16: Value-weighted and equally-weighted estimates of annualized ν for the time-invariant models

	Stocks ($n = 9,936$)			
	Value-weighted estimate (VW)		Equally-weighted estimate (OLS)	
	bias corrected estimate (%)	95% conf. interval	bias corrected estimate (%)	95% conf. interval
	Four-factor model			
($n^{\chi} = 9,902$)				
ν_m	6.99	(5.55, 8.43)	-1.95	(-3.85, -0.05)
ν_{smb}	0.24	(-2.33, 2.81)	0.76	(-1.06, 2.57)
ν_{hml}	-13.77	(-16.26, -11.29)	-6.01	(-7.88, -4.15)
ν_{mom}	4.62	(-0.30, 9.53)	-2.64	(-5.91, 0.64)
	Fama-French model			
($n^{\chi} = 9,904$)				
ν_m	6.01	(4.71, 7.30)	-1.93	(-4.18, -0.31)
ν_{smb}	0.09	(-1.88, 2.06)	0.63	(-1.04, 2.29)
ν_{hml}	-13.74	(-15.91, -11.57)	-5.99	(-8.00, -3.97)
	CAPM			
($n^{\chi} = 9,904$)				
ν_m	7.77	(6.14, 9.40)	-1.10	(-3.95, 1.77)

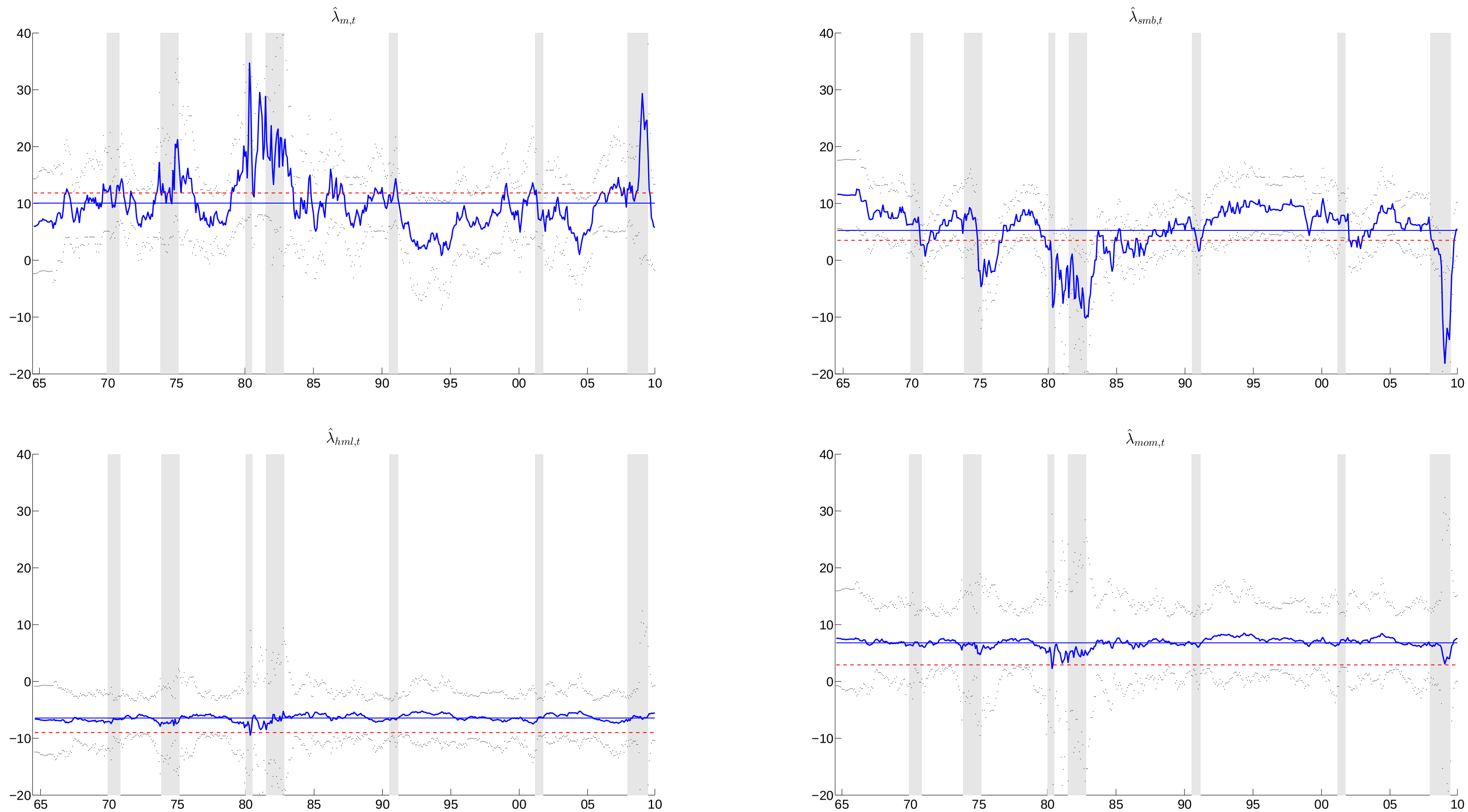
The table contains the value-weighted and equally-weighted annualized estimates of the components of vector ν for the market (ν_m), size (ν_{smb}), book-to-market (ν_{hml}), and momentum (ν_{mom}) factors. We report the bias corrected estimates $\hat{\nu}_B$ of ν for individual stocks ($n = 9,936$). In order to build the confidence intervals, we compute $\tilde{\Sigma}_\nu$ in Proposition 5 for $n = 9,936$.

Table 17: Estimated annualized components of ν for the time-varying four-factor model

		value-weighted estimate (VW) of ν	equally-weighted estimate (OLS) of ν
m	$const$	5.2206	1.9634
		$(3.2894,\ 7.1517)$	$(0.5951,\ 3.3316)$
	ds_{t-1}	1.3546	-3.8687
		$(-2.3692,\ 5.0784)$	$(-8.5946,\ 0.8573)$
	ts_{t-1}	-5.1721	-1.2790
		$(-6.8086,\ -3.5356)$	$(-3.2226,\ 0.6647)$
smb	$const$	-1.9799	1.1089
		$(-0.6870,\ 4.6468)$	$(-0.1936,\ 2.4113)$
	ds_{t-1}	-7.3882	-7.5468
		$(-11.2165,\ -3.6162)$	$(-12.2291,\ -2.8645)$
	ts_{t-1}	0.1753	-0.3858
		$(-2.0315,\ 2.3820)$	$(-1.2706,\ 2.0422)$
hml	$const$	-11.2165	-5.6595
		$(-14.0983,\ -8.3347)$	$(-7.2660,\ -4.0529)$
	ds_{t-1}	1.6581	3.3613
		$(-2.1038,\ 5.4199)$	$(-3.4080,\ 10.1305)$
	ts_{t-1}	-0.2633	-5.6149
		$(-2.4270,\ 1.9004)$	$(-7.4538,\ -3.7760)$
mom	$const$	-0.5102	-6.6244
		$(-3.6884,\ 2.6680)$	$(-8.8359,\ -4.4129)$
	ds_{t-1}	1.9627	24.0970
		$(-3.8599,\ 7.7852)$	$(16.6218,\ 31.5721)$
	ts_{t-1}	-1.7040	-3.3855
		$(-4.3746,\ 0.9666)$	$(-6.1543,\ -0.6167)$

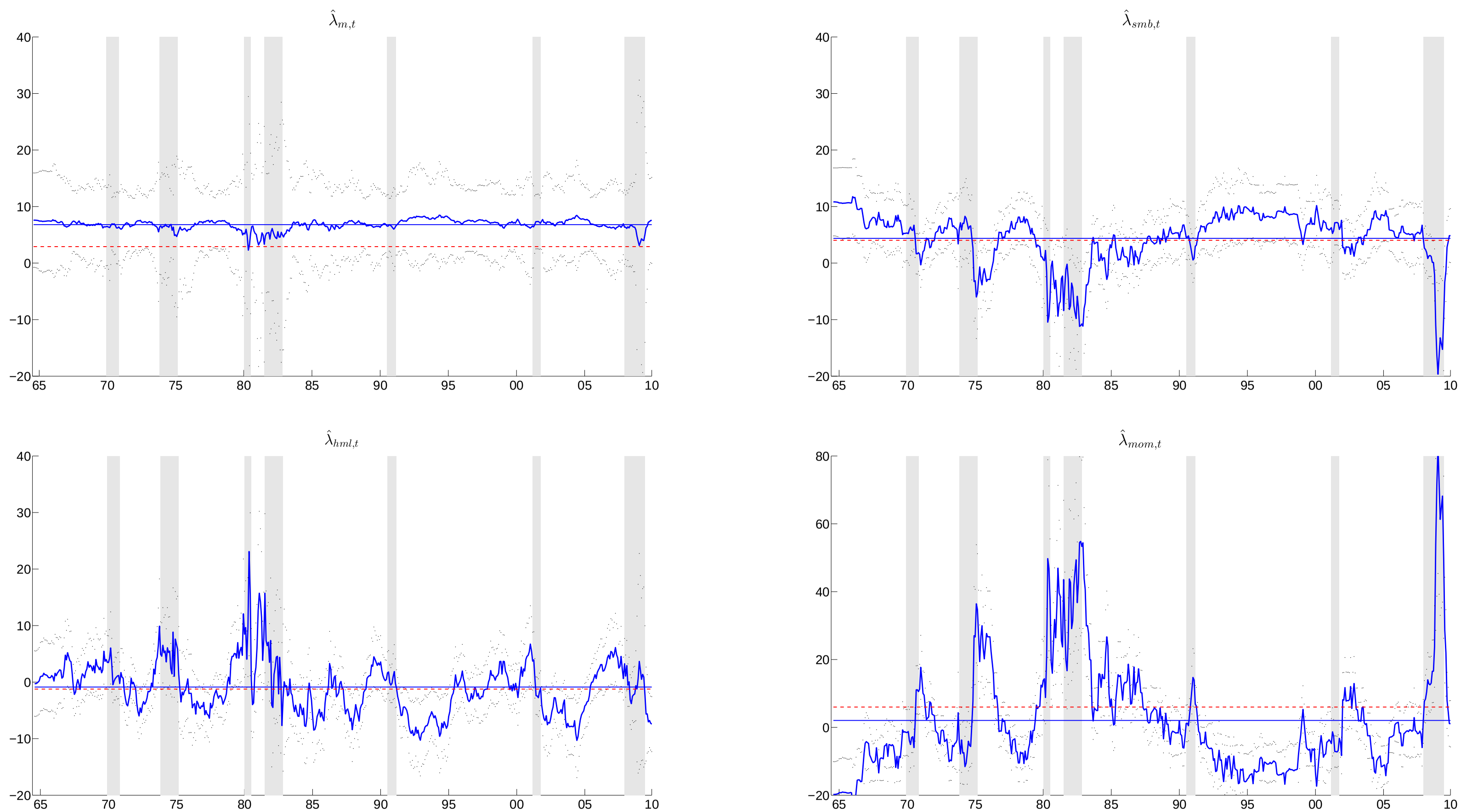
The table contains the value-weighted and equally-weighted estimated annualized components of vector ν, and their confidence intervals at 95% probability level for the individual stocks ($n = 9,936$ and $n^{\chi} = 3,900$). We report the bias corrected estimates $\hat{\nu}_B$ of ν. In order to build the confidence intervals for ν, we use the thresholded variance-covariance matrix of Proposition 5. The default spread ds_{t-1} and the term spread ts_{t-1} are centered and standardized.

Figure 15: Path of value-weighted estimates of annualized risk premia with $n = 9,936$



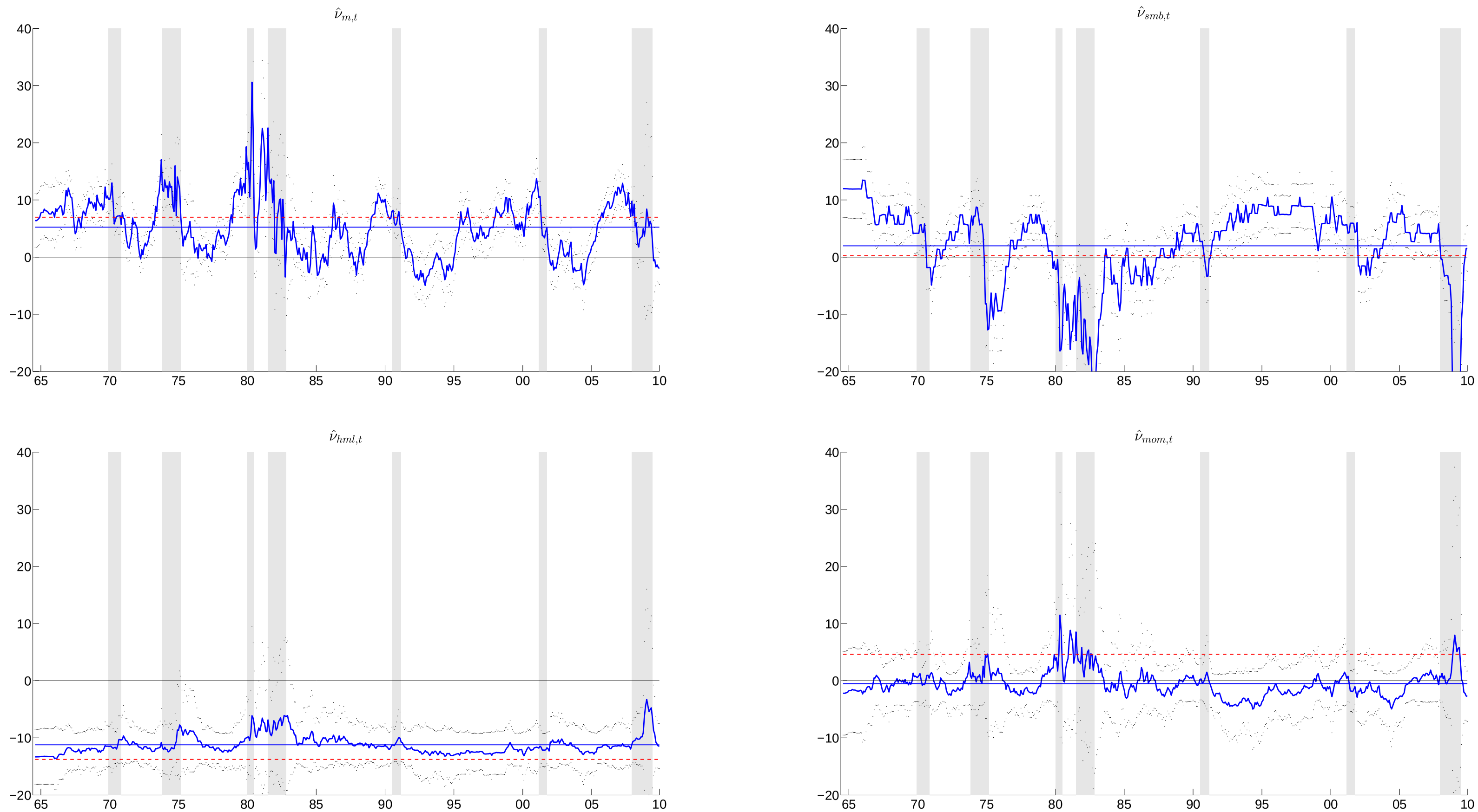
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$, and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. Estimates use a value-weighting (VW) scheme. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 3,900$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Figure 16: Path of equally-weighted estimates of annualized risk premia with $n = 9,936$



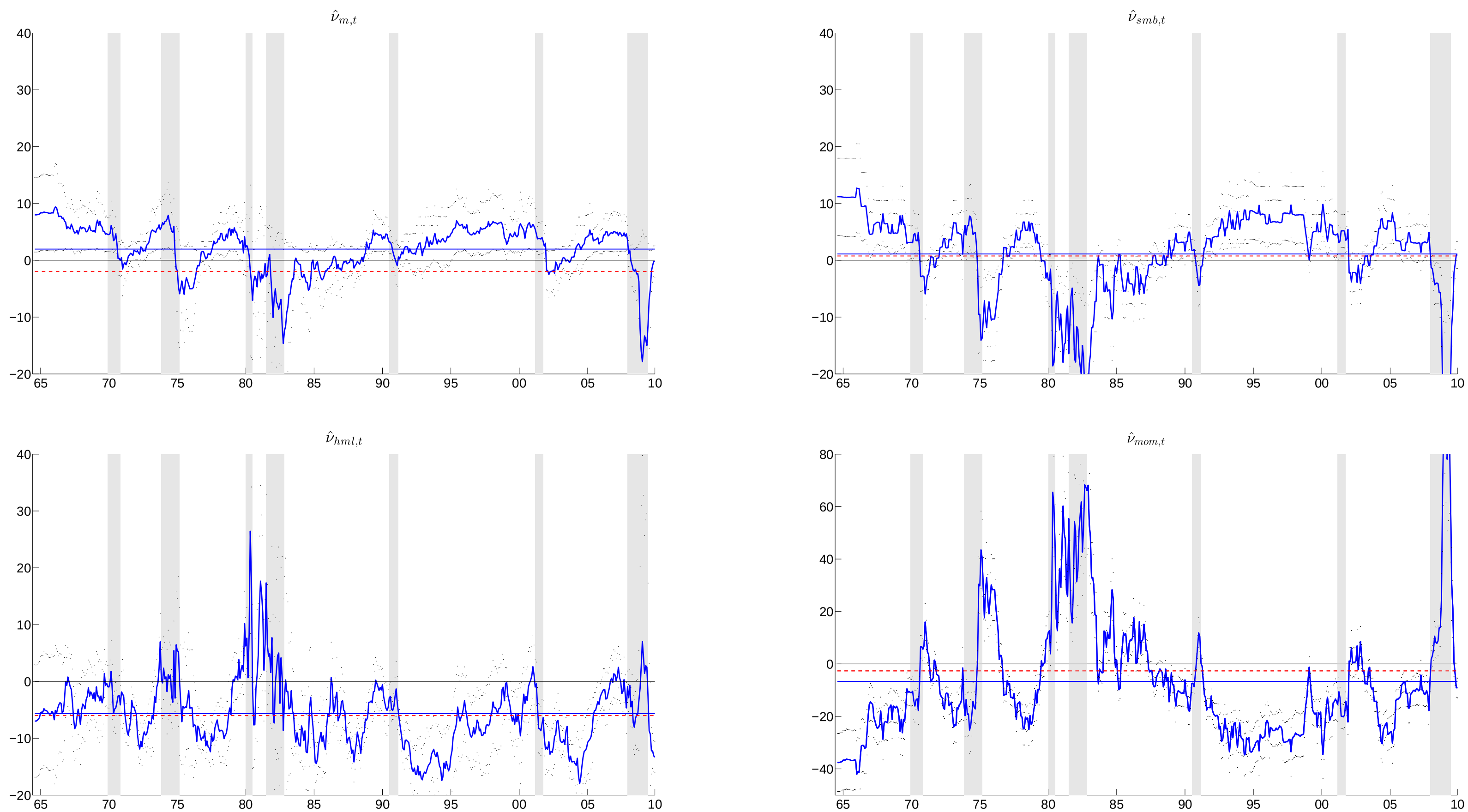
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$, and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. Estimates use an equally-weighting (OLS) scheme. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 3,900$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Figure 17: Path of value-weighted estimates of annualized ν_t with $n = 9,936$



The figure plots the path of value-weighted estimated annualized $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$, $\hat{\nu}_{hml,t}$, and $\hat{\nu}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. Estimates use a value-weighting (VW) scheme. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 3,900$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Figure 18: Path of equally-weighted estimates of annualized ν_t with $n = 9,936$



The figure plots the path of equally-weighted estimated annualized $\hat{\nu}_{m,t}$, $\hat{\nu}_{smb,t}$, $\hat{\nu}_{hml,t}$, and $\hat{\nu}_{mom,t}$ and their pointwise confidence intervals at 95% probability level in the four-factor model. Estimates use an equally-weighting (OLS) scheme. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 3,900$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER). The recessions start at the peak of a business cycle and end at the trough.

Appendix 13 Empirical analysis of estimated time-varying betas

In this appendix, we provide an empirical comparison between the estimated time-varying betas for the portfolios and the individual stocks. The aim is to show that individual stock betas and Indu. portfolio betas move over time substantially while the FF portfolio betas are much more stable. This means that the time-invariant models for the individual stocks and Indu. portfolios in Section 4.2 are likely misspecified as discussed theoretically in Section A.14.1. The first subsection A.13.1 looks at the distributional characteristics of the estimated time-varying betas. The second subsection A.13.2 presents test results for time invariance of betas similar to the ones developed for λ_t and ν_t in Section 4.3. The third subsection A.13.3 gives examples of estimated paths of factor sensitivities.

A.13.1 Distribution of variability of estimated time-varying betas

The time-varying factor loadings are $b_{i,t} = Q_{i,t}\beta_{2,i}$, with $Q_{i,t} = \left(I_K \otimes Z'_{t-1}, I_K \otimes Z'_{i,t-1}\right) \in \mathbb{R}^{K\times d_2}$, from the functional specification in Assumption FS.1. In this section, we compute the estimates $\hat{b}_{i,t} = Q_{i,t}\hat{\beta}_{2,i}$ if $I_{i,t} = 1$, and we study their stability over time. Table 18 reports the cross-sectional summary statistics of three measures of dispersion over time of the $\hat{b}_{i,t}$: (i) the standard deviation $std\left(\hat{b}_{k,i}\right) = \sqrt{V\left[\hat{b}_{k,i}\right]}$, with $V\left[\hat{b}_{k,i}\right] = \frac{1}{T_i}\sum_t I_{i,t}\left(\hat{b}_{k,i,t} - \bar{b}_{k,i}\right)$ and $\bar{b}_{k,i} = \frac{1}{T_i}\sum_t I_{i,t}\hat{b}_{k,i,t}$ for $k = 1, ..., K$; (ii) the coefficient of variation $cv\left(\hat{b}_{k,i}\right) = \frac{std\left(\hat{b}_{k,i}\right)}{\left|\bar{b}_{k,i}\right|}$; (iii) the quartile coefficient of dispersion $qc\left(\hat{b}_{k,i}\right) = \frac{Q_3\left(\hat{b}_{k,i}\right) - Q_1\left(\hat{b}_{k,i}\right)}{\left|Q_2\left(\hat{b}_{k,i}\right)\right|}$, where $Q_1\left(\hat{b}_{k,i}\right)$, $Q_2\left(\hat{b}_{k,i}\right)$, and $Q_3\left(\hat{b}_{k,i}\right)$ are the lower, median, and upper quartiles of the $\hat{b}_{k,i,t}$ over time. Figures 19-21 plot the cross-sectional distributions of the three measures of time variation of betas for individual stocks and portfolios.

The individual stocks have the most pronounced time variation in the sensitivities of the four factors. For instance, the median standard deviation of betas is about 10 times larger for the individual stocks than for the 25 FF portfolios uniformly across factors. A similar comment applies for the other measures of time variability of betas. The cross-sectional distributions of the time variability measures for individual stocks are strongly right skewed, and the largest values fall outside the support $[0, 15]$ displayed in Figures 20 and

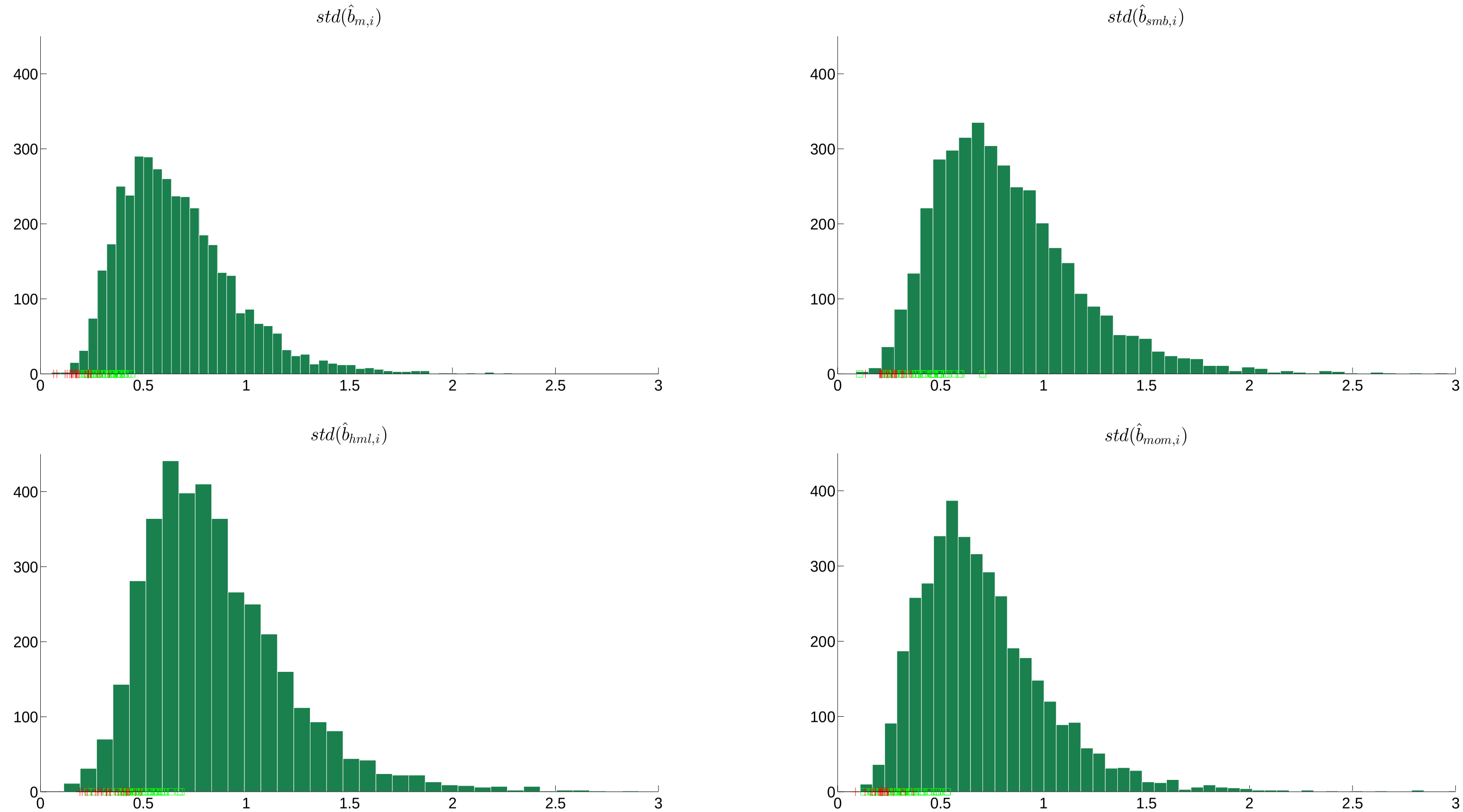
21. In Table 18, the median quartile coefficient of dispersion of the momentum betas is almost equal for the individual stocks and the 25 FF portfolios, and the mean value is larger for the 25 FF portfolios. However this result is driven by the large quartile coefficients of dispersion of two portfolios. Excluding those portfolios, the mean coefficient of dispersion of the momentum beta for the remaining 23 FF portfolios is 2.2074, an order of magnitude smaller than for individual stocks.

Table 18: Summary statistics of $std\left(\hat{b}_{k,i}\right)$, $cv\left(\hat{b}_{k,i}\right)$ and $qc\left(\hat{b}_{k,i}\right)$ for the four-factor model

	$n = 9,936$	$n = 25$	$n = 44$	$n = 9,936$	$n = 25$	$n = 44$	$n = 9,936$	$n = 25$	$n = 44$
b_m		std			cv			qc	
Min	0.0029	0.0041	0.0365	0.0013	0.0038	0.0354	0.0021	0.0055	0.0395
Median	0.3858	0.0311	0.1089	0.4000	0.0298	0.1031	0.5493	0.0430	0.1319
Mean	0.5235	0.0390	0.1077	1.0494	0.0376	0.1076	1.6491	0.0473	0.1355
Max	5.2495	0.1105	0.1964	208.2981	0.1045	0.2924	505.5851	0.0995	0.3250
Std	0.4821	0.0252	0.0417	5.0411	0.0233	0.0503	11.3116	0.0280	0.0647
b_{smb}		std			cv			qc	
Min	0.0078	0.0182	0.0115	0.0000	0.0484	0.0983	0.0000	0.0472	0.1409
Median	0.5764	0.0701	0.1916	0.8000	0.1005	0.6089	1.0000	0.1402	0.7833
Mean	0.7979	0.0713	0.1861	3.6000	0.1903	0.8565	3.7000	0.2402	1.2567
Max	10.3200	0.1285	0.4968	2,131.0000	0.5966	4.8580	2,392.0000	0.8739	8.0974
Std	0.7686	0.0278	0.0975	393.0000	0.1528	0.8496	41.40000	0.2110	1.5337
b_{hml}		std			cv			qc	
Min	0.0128	0.0366	0.0532	0.0000	0.0486	0.2898	0.0000	0.0496	0.2879
Median	0.6275	0.1031	0.2247	1.3000	0.2810	0.8727	1.7000	0.4048	1.0380
Mean	0.8562	0.1176	0.2436	8.7000	0.4316	1.3741	14.3000	0.6345	1.7489
Max	16.8142	0.2373	0.4649	7,117.8000	1.5348	4.8387	8,355.80	2.3510	9.5961
Std	0.8416	0.0561	0.0945	13.9600	0.4286	1.1810	19.2100	0.6470	1.8855
b_{mom}		std			cv			qc	
Min	0.0120	0.0075	0.0164	0.0000	0.0813	0.2191	0.0000	0.6997	0.3196
Median	0.4231	0.0462	0.1019	2.0000	1.5244	1.6877	2.0000	1.9840	2.0369
Mean	0.6002	0.0521	0.1212	12.0000	2.7166	2.8861	17.0000	141.7774	10.8291
Max	9.5135	0.1352	0.2831	11,885.0000	21.0727	28.0131	10,847.0000	3,343.4000	300.2741
Std	0.6339	0.0276	0.0721	214.0000	4.1031	4.4284	270.0000	667.6600	45.4179

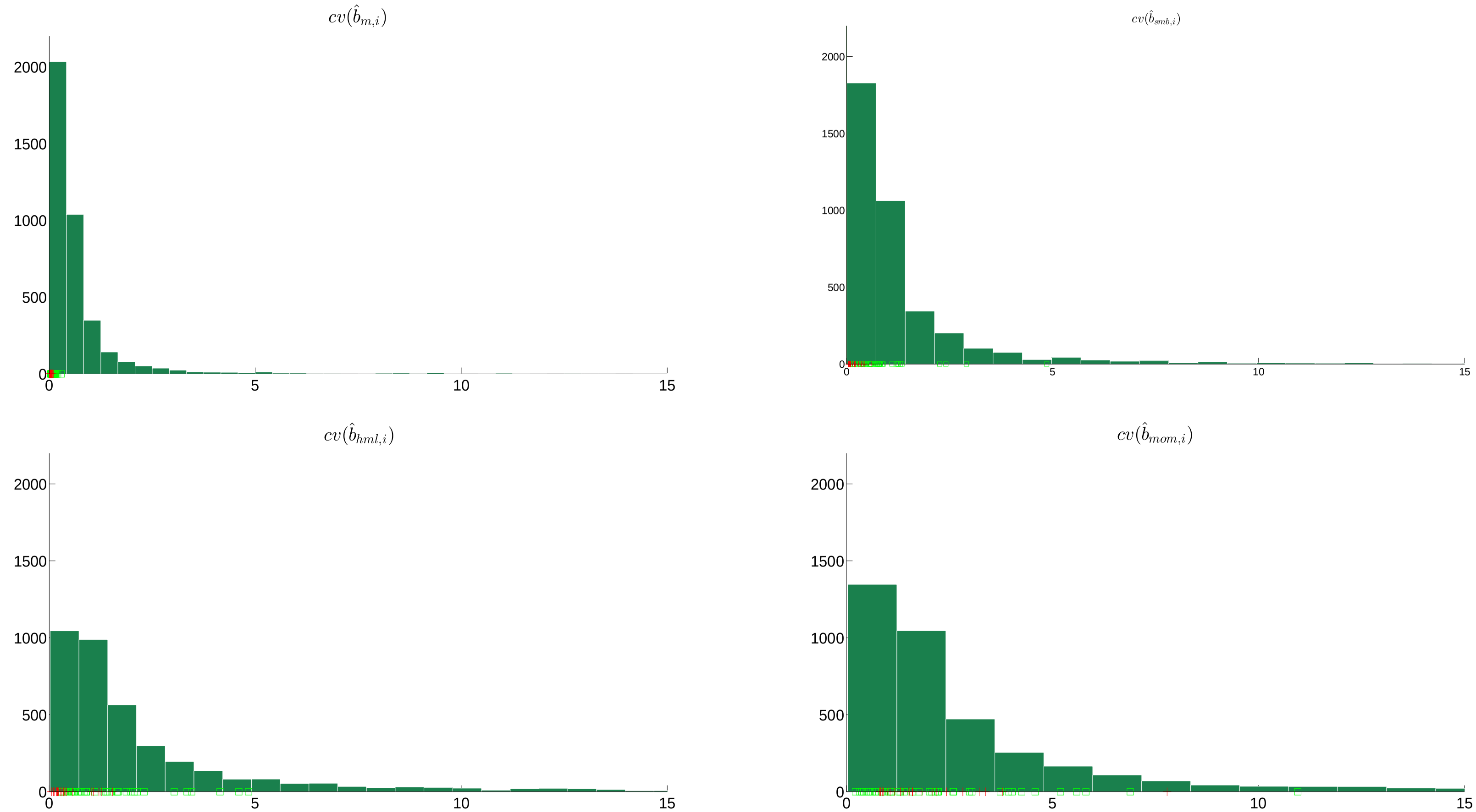
The table contains the descriptive statistics (cross-sectional minimum, median, mean, maximum and standard deviation) of the standard deviation (std), the coefficient of variation (cv) and the quartile coefficient of dispersion (qc) over time of the time-varying estimated loadings for the market (b_m), size (b_{smb}), book-to-market (b_{hml}) and momentum (b_{hml}) factors. We report the results for individual stocks ($n = 9,936$, $n^{\chi} = 3,900$), for the 25 FF ($n = 25$) and 44 Indu. ($n = 44$) portfolios.

Figure 19: Cross-sectional distributions of the standard deviations of $\hat{b}_{k,i,t}$ over time, $std\left(\hat{b}_{k,i}\right)$, for the four-factor model



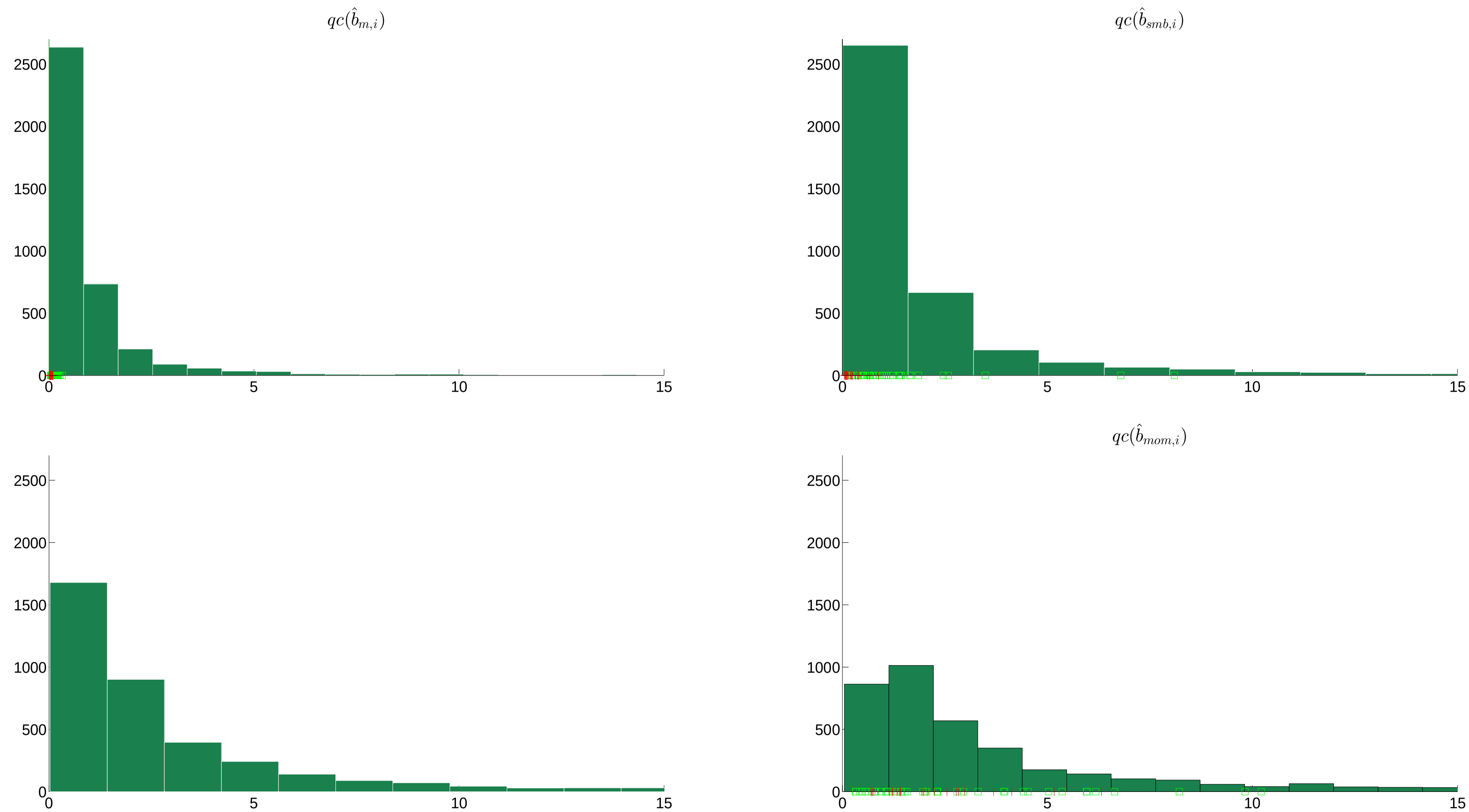
The figure plots the cross-sectional distributions of the standard deviations of estimated time-varying betas for the four-factor model, $\hat{b}_{m,i,t}, \hat{b}_{smb,i,t}, \hat{b}_{hml,i,t}$, and $\hat{b}_{mom,i,t}$, on individual stocks (histograms, $n = 9,936$ and $n^{\chi} = 3,900$), on the 25 FF portfolios (red crosses) and the 44 Indu. portfolios (green squares).

Figure 20: Cross-sectional distributions of the coefficient of variation of $\hat{b}_{k,i,t}$ over time, $cv\left(\hat{b}_{k,i}\right)$, for the four-factor model



The figure plots the cross-sectional distributions of the coefficient of variation of estimated time-varying betas for the four-factor model, $\hat{b}_{m,i,t}, \hat{b}_{smb,i,t}, \hat{b}_{hml,i,t}$, and $\hat{b}_{mom,i,t}$, on individual stocks (histograms, $n = 9,936$ and $n^{\chi} = 3,900$), on the 25 FF portfolios (red crosses) , and the 44 Indu. portfolios (green squares).

Figure 21: Cross-sectional distributions of the quartile coefficient of disperion of $\hat{b}_{k,i,t}$ over time, $qc\left(\hat{b}_{k,i}\right)$, for the four-factor model



The figure plots the cross-sectional distributions of the quantile coefficient of dispersion of estimated time-varying betas for the four-factor model, $\hat{b}_{m,i,t}, \hat{b}_{smb,i,t}, \hat{b}_{hml,i,t}$, and $\hat{b}_{mom,i,t}$, on individual stocks (histograms, $n = 9,936$ and $n^{\chi} = 3,900$), on the 25 FF portfolios (red crosses) , and the 44 Indu. portfolios (green squares).

A.13.2 Test results for time invariance of betas

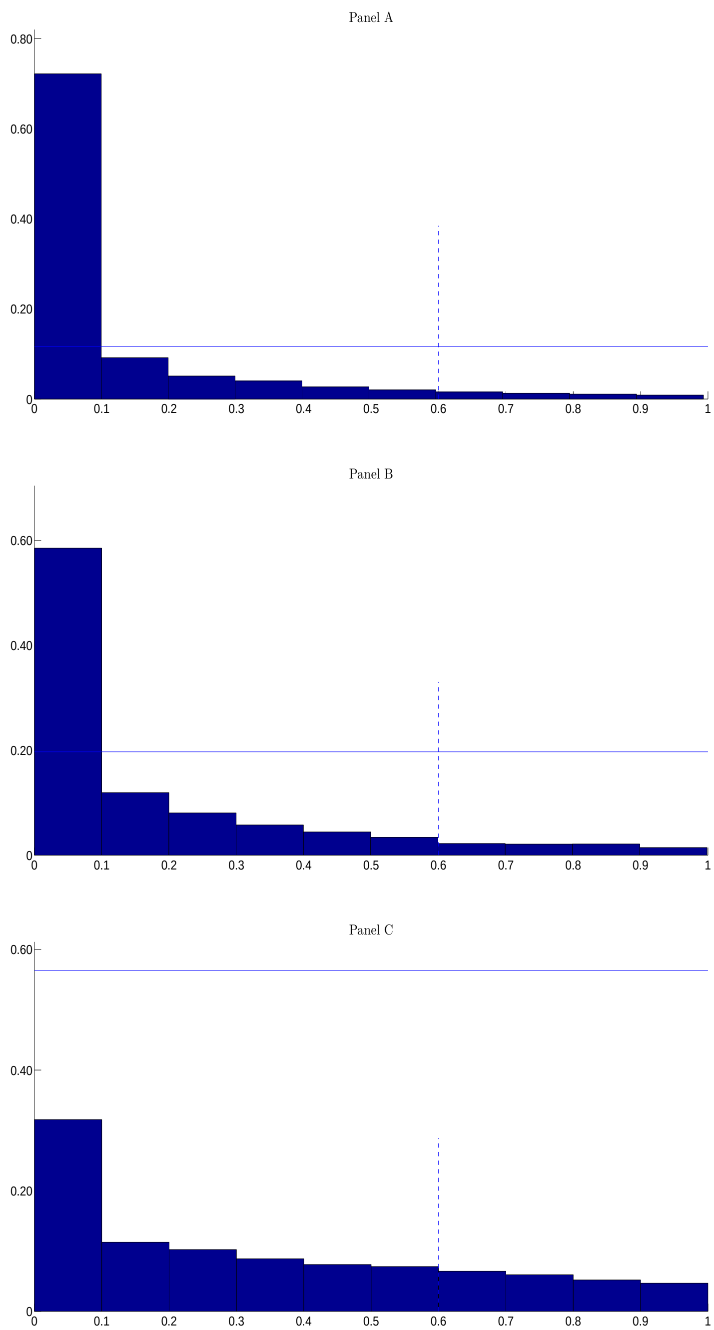
Time variation of factor loadings $b_{i,t}$ goes through the coefficients of parameter $\beta_{2,i}$ which load on the instruments. For each asset i, the null hypothesis is $\mathcal{H}_0^{\beta_{2,i}}$: $A\beta_{2,i} = 0$, where the matrix $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ with diagonal blocks $A_1 = I_K \otimes \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $A_2 = I_K \otimes I_q$ selects the instrument coefficients when $p = 3$ as in our empirical specification. We build the standard chi-square statistic with $K(p-1+q)$ degrees of freedom by using

$$\sqrt{T}\left(\hat{\beta}_{2,i} - \beta_{2,i}\right) = \sqrt{T}E_2'\left(\hat{\beta}_i - \beta_i\right) = \tau_{i,T}E_2'\hat{Q}_{x,i}^{-1}Y_{i,T}. \tag{69}$$

Standard results on OLS imply that the estimator $\hat{\beta}_i$ is asymptotically normal, $\sqrt{T}\left(\hat{\beta}_i - \beta_i\right) \Rightarrow N\left(0, \tau_i Q_{x,i}^{-1} S_{ii} Q_{x,i}^{-1}\right)$.

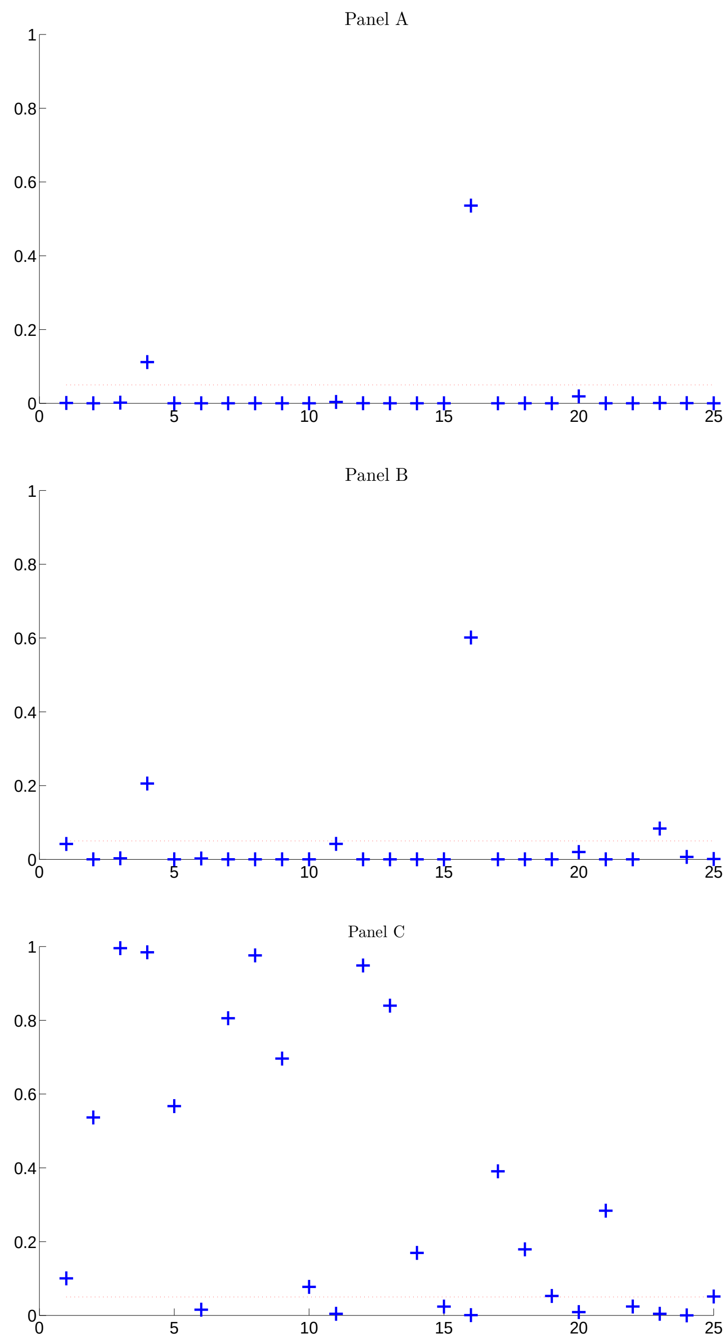
We compute the statistic for the four-factor model, the Fama-French model, and the CAPM. For individual stocks, we plot the histogram of the p-values of $n^{\chi} = 3,900$ statistics in Figure 22. We estimate the proportion of assets with time-invariant betas, denoted $\hat{\pi}_0$, as in Barras, Scaillet and Wermers (2010) (see also Bajgrowicz and Scaillet (2012)). We fix the threshold of their approach at $\lambda^* = 0.6$, and we get $\hat{\pi}_0 = 11.67\%$ for the four-factor model. We get $\hat{\pi}_0 = 19.69\%$ for the Fama-French model, and $\hat{\pi}_0 = 56.51\%$ for the CAPM. For the 25 FF and 44 Indu. portfolios, we show the p-values of the statistics in Figures 23 and 24, respectively. Moreover, for the 25 FF portfolios, we compute the $K(p-1+q)$ t-statistics of components of vector $A\beta_{2,i}$. We perform this exercise assuming the four-factor model, the Fama-French model, and the CAPM. Tables 19-21 report the p-values of the t-statistics for each portfolio. The null hypothesis of time-invariant betas in the four-factor model is not rejected at the 5% level for 2 out of the 25 FF portfolios, and for 6 out of the 44 Indu. portfolios. For both the individual stocks and the portfolios, the proportion of assets, for which the time-invariance hypothesis is not rejected, increases when we switch from the four-factor model to the Fama-French model and the CAPM. The misspecification due to omitted factors can mask the time variation in betas. Moreover, the number of constraints from time-invariance of betas increases with the number of factors. Overall, the findings in Tables 19-21 and Figures 19-24 suggest that (i) there is statistical evidence to reject the hypothesis of time-invariant betas for the majority of individual stocks and portfolios, but (ii) time variation in betas is much more pronounced for the individual stocks than for portfolios.

Figure 22: Distribution of p-values of statistics for the null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ on individual stocks



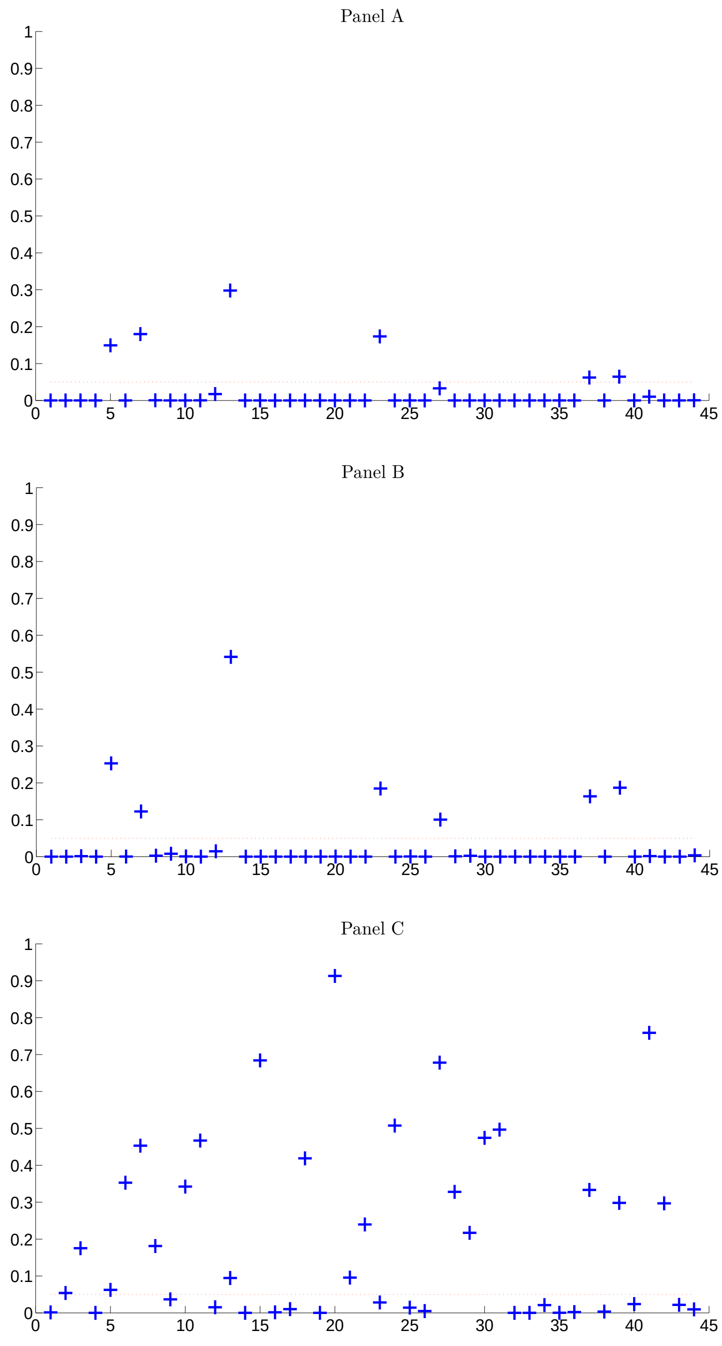
Panel A represents the p-value histogram of the $n^{\chi} = 3,900$ statistics for testing the null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ on individual stocks computed by assuming a four-factor model. Panel B plots the distribution of $n^{\chi} = 4,545$ p-values using the Fama-French model. Panel C plots the p-values histogram for the $n^{\chi} = 5,225$ stocks computed by assuming the CAPM. We also display the threshold λ^* (dashed vertical line) and the estimated proportion of assets with time-invariant betas, $\hat{\pi}_0$ (solid horizontal line).

Figure 23: p-values of statistics for the null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ on the 25 FF portfolios



The figure represents the p-values of the $n = 25$ statistics for testing the null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ computed on the 25 FF portfolios. Panel A, B, and C represent the p-values computed by using the four-factor model, the Fama-French model, and the CAPM, respectively. We also plot the probability level at 5% (dashed red line).

Figure 24: p-values of statistics for the null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ on the 44 Indu. portfolios



The figure represents the p-values of the $n = 44$ statistics for testing the null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ computed on the 44 Indu. portfolios. Panel A, B, and C represent the p-values computed by using the four-factor model, the Fama-French model, and the CAPM, respectively. We also plot the probability level at 5% (dashed red line).

Table 19: p-values of the t-statistics for the components of $A\hat{\beta}_{2,i}$ for the 25 FF portfolios and the four-factor model

	Portfolio	1	2	3	4	5	6	7	8	9	10	11	12	13
	ds_{t-1}	0.4599	0.6185	0.4356	0.4548	0.4432	0.1831	0.1135	0.5233	0.7541	0.5114	0.7382	0.0647	0.8481
m	ts_{t-1}	0.6584	0.6517	0.6834	0.7228	0.2680	0.3131	0.7852	0.1950	0.1156	0.0005	0.8073	0.1768	0.0071
	$bm_{i,t-1}$	0.0976	0.0055	0.3975	0.1640	0.1038	0.0007	0.4470	0.2099	0.4447	0.0151	0.2022	0.2644	0.4862
	ds_{t-1}	0.1223	0.0859	0.0064	0.2465	0.4586	0.1683	0.2447	0.8990	0.0395	0.0289	0.9005	0.4365	0.5095
smb	ts_{t-1}	0.0189	0.1501	0.0938	0.2820	0.0171	0.6009	0.0732	0.8610	0.0408	0.0557	0.0077	0.6464	0.0706
	$bm_{i,t-1}$	0.0514	0.2139	0.0173	0.2119	0.0451	0.0252	0.0000	0.0000	0.0015	0.8296	0.0037	0.0000	0.0000
	ds_{t-1}	0.9678	0.4848	0.4141	0.1766	0.5357	0.2775	0.9541	0.1474	0.2168	0.0556	0.3575	0.7250	0.0168
hml	ts_{t-1}	0.3009	0.2786	0.2627	0.0251	0.2717	0.0194	0.7075	0.9352	0.0519	0.8536	0.2403	0.5425	0.5304
	$bm_{i,t-1}$	0.2019	0.1397	0.5248	0.8637	0.3888	0.2263	0.3365	0.9560	0.9280	0.0006	0.8066	0.4886	0.0910
	ds_{t-1}	0.6860	0.9257	0.6865	0.5461	0.0100	0.0301	0.2530	0.6400	0.1534	0.9856	0.3619	0.3835	0.5017
mom	ts_{t-1}	0.3482	0.5326	0.4638	0.8555	0.4688	0.1311	0.1423	0.0036	0.3800	0.4977	0.1016	0.0395	0.433
	$bm_{i,t-1}$	0.0962	0.4054	0.1342	0.4952	0.6964	0.0035	0.0272	0.8324	0.1495	0.5121	0.0134	0.1114	0.6725
p-values for $\mathcal{H}_0^{\beta_{2,i}}$		0.0009	0.0000	0.0017	0.1117	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0037	0.0003	0.0000

	Portfolio	14	15	16	17	18	19	20	21	22	23	24	25
	ds_{t-1}	0.1303	0.0013	0.9973	0.400	0.2704	0.6714	0.1186	0.0459	0.2473	0.1775	0.5812	0.0037
m	ts_{t-1}	0.6402	0.0025	0.8609	0.2055	0.0189	0.2668	0.6195	0.0454	0.6947	0.1387	0.0353	0.1105
	$bm_{i,t-1}$	0.7233	0.4948	0.2548	0.1706	0.1649	0.1101	0.1108	0.0000	0.2302	0.0503	0.6767	0.7626
	ds_{t-1}	0.0898	0.1962	0.1013	0.0107	0.1603	0.1882	0.1090	0.1470	0.5581	0.1293	0.6939	0.4463
smb	ts_{t-1}	0.0000	0.8924	0.8245	0.0000	0.0002	0.0044	0.2585	0.0024	0.0417	0.0083	0.7002	0.0091
	$bm_{i,t-1}$	0.0007	0.2947	0.6232	0.0001	0.0000	0.0001	0.0312	0.2663	0.0289	0.0219	0.0000	0.4522
	ds_{t-1}	0.8898	0.0023	0.7911	0.2648	0.0125	0.3582	0.2970	0.7678	0.0490	0.6217	0.0755	0.5504
hml	ts_{t-1}	0.1770	0.5773	0.8343	0.6240	0.2653	0.8440	0.2591	0.2443	0.7116	0.3477	0.1381	0.0004
	$bm_{i,t-1}$	0.7110	0.0337	0.8939	0.1758	0.3545	0.4109	0.0074	0.0081	0.7634	0.0376	0.9627	0.0415
	ds_{t-1}	0.7466	0.0209	0.4460	0.4506	0.9760	0.3856	0.6150	0.6161	0.2053	0.5055	0.6394	0.7532
mom	ts_{t-1}	0.7259	0.4885	0.1521	0.5367	0.8791	0.0034	0.9300	0.0220	0.0044	0.8331	0.8990	0.0101
	$bm_{i,t-1}$	0.5076	0.5047	0.7720	0.1908	0.1468	0.0453	0.6284	0.0005	0.1625	0.0004	0.1031	0.9169
p-values for $\mathcal{H}_0^{\beta_{2,i}}$		0.0000	0.0000	0.5360	0.0000	0.0000	0.0000	0.0189	0.0000	0.0000	0.0009	0.0005	0.0000

The table reports the p-values of the t-statistics computed on each component of vector $A\hat{\beta}_{2,i}$, estimated on the 25 FF portfolios by assuming the four-factor model. For comparison purposes, we also report the p-values for the joint null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ (see Figure 23, Panel A).

Table 20: p-values of the t-statistics for the components of $A\hat{\beta}_{2,i}$ for the 25 FF portfolios and the Fama-French model

	Portfolio	1	2	3	4	5	6	7	8	9	10	11	12	13
	ds_{t-1}	0.0733	0.1738	0.3635	0.3456	0.5784	0.7323	0.3457	0.2645	0.5706	0.3156	0.7191	0.1367	0.6048
m	ts_{t-1}	0.4409	0.6374	0.7057	0.6129	0.2835	0.2048	0.9056	0.2105	0.1070	0.0005	0.8593	0.1847	0.0108
	$bm_{i,t-1}$	0.1324	0.0081	0.5098	0.1878	0.0429	0.0021	0.2179	0.1962	0.4834	0.0168	0.3305	0.1914	0.3862
	ds_{t-1}	0.1459	0.0861	0.0123	0.2738	0.3186	0.3676	0.5540	0.8204	0.0490	0.0440	0.7496	0.5272	0.7128
smb	ts_{t-1}	0.0248	0.3091	0.1622	0.3762	0.0071	0.4027	0.0883	0.4520	0.0023	0.2589	0.0171	0.8453	0.0095
	$bm_{i,t-1}$	0.0158	0.1726	0.0187	0.1428	0.0455	0.1364	0.0000	0.0000	0.0012	0.9687	0.0085	0.0000	0.0000
	ds_{t-1}	0.0523	0.0499	0.4741	0.7366	0.4097	0.0334	0.7426	0.9243	0.7577	0.0004	0.7209	0.3340	0.0633
hml	ts_{t-1}	0.3982	0.7438	0.9607	0.4064	0.0082	0.0199	0.3157	0.5476	0.3598	0.8374	0.5712	0.3042	0.4313
	$bm_{i,t-1}$	0.4678	0.5325	0.7249	0.7649	0.6699	0.0671	0.0904	0.0067	0.3447	0.4400	0.1479	0.0353	0.4496
p-values for $\mathcal{H}_0^{\beta_{2,i}}$		0.0417	0.0000	0.0028	0.2055	0.0000	0.0025	0.0000	0.0000	0.0000	0.0000	0.0419	0.0000	0.0000

	Portfolio	14	15	16	17	18	19	20	21	22	23	24	25
	ds_{t-1}	0.3016	0.0000	0.8815	0.3852	0.9783	0.5598	0.4009	0.0120	0.0596	0.1627	0.3556	0.0218
m	ts_{t-1}	0.7432	0.0012	0.8280	0.2201	0.0172	0.2321	0.4984	0.0824	0.5970	0.1949	0.0855	0.0753
	$bm_{i,t-1}$	0.5189	0.3695	0.2622	0.0778	0.0672	0.2529	0.0447	0.0000	0.2093	0.0418	0.8899	0.5907
	ds_{t-1}	0.1529	0.0377	0.1048	0.0295	0.2915	0.3308	0.0633	0.1957	0.6493	0.1268	0.9920	0.4404
smb	ts_{t-1}	0.0000	0.2151	0.9222	0.0000	0.0231	0.0226	0.7312	0.0043	0.0045	0.0277	0.3240	0.0103
	$bm_{i,t-1}$	0.0008	0.4977	0.5754	0.0001	0.0000	0.0000	0.0400	0.8390	0.0716	0.1191	0.0002	0.7700
	ds_{t-1}	0.9560	0.0161	0.8782	0.0970	0.1201	0.4852	0.0161	0.0214	0.8588	0.0774	0.6542	0.0219
hml	ts_{t-1}	0.9727	0.0133	0.3853	0.5180	0.8978	0.1470	0.3894	0.1204	0.0757	0.4976	0.5606	0.6095
	$bm_{i,t-1}$	0.7079	0.4841	0.1070	0.4380	0.4987	0.0030	0.7976	0.1272	0.0038	0.5805	0.6000	0.0174
p-values for $\mathcal{H}_0^{\beta_{2,i}}$		0.0000	0.0000	0.6013	0.0000	0.0000	0.0000	0.0200	0.0000	0.0000	0.0835	0.0067	0.0010

The table reports the p-values of the t-statistics computed on each component of vector $A\hat{\beta}_{2,i}$, estimated on the 25 FF portfolios by assuming the Fama-French model. For comparison purposes, we also report the p-values for the joint null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ (see Figure 23, Panel B).

Table 21: p-values of the t-statistics for the components of $A\beta_{2,i}$ for the 25 FF portfolios and the CAPM

	Portfolio	1	2	3	4	5	6	7	8	9	10	11	12	13
	ds_{t-1}	0.2822	0.4219	0.8277	0.776	0.1996	0.0034	0.3452	0.7554	0.3288	0.2308	0.0032	0.5610	0.4998
m	ts_{t-1}	0.1260	0.6239	0.8390	0.7438	0.9200	0.9148	0.6101	0.8916	0.8077	0.2202	0.2291	0.7162	0.7482
	$bm_{i,t-1}$	0.4183	0.4919	0.9931	0.8983	0.756	0.5648	0.7754	0.9376	0.9035	0.3168	0.2999	0.9809	0.6946
p-values for $\mathcal{H}_0^{\beta_{2,i}}$		0.1004	0.5365	0.9955	0.9843	0.5672	0.0155	0.8056	0.9758	0.6962	0.0773	0.0043	0.9486	0.8396

	Portfolio	14	15	16	17	18	19	20	21	22	23	24	25
	ds_{t-1}	0.6779	0.5691	0.0007	0.7009	0.2842	0.1467	0.0618	0.0673	0.0643	0.0254	0.0002	0.7756
m	ts_{t-1}	0.0800	0.0819	0.2468	0.1490	0.1606	0.1712	0.1964	0.9456	0.0056	0.1262	0.2399	0.0064
	$bm_{i,t-1}$	0.7112	0.0795	0.3907	0.8753	0.7991	0.7391	0.1331	0.3524	0.1547	0.5142	0.8446	0.5030
p-values for $\mathcal{H}_0^{\beta_{2,i}}$		0.1694	0.0240	0.0008	0.3907	0.1792	0.0526	0.0090	0.2838	0.0240	0.0043	0.0000	0.0512

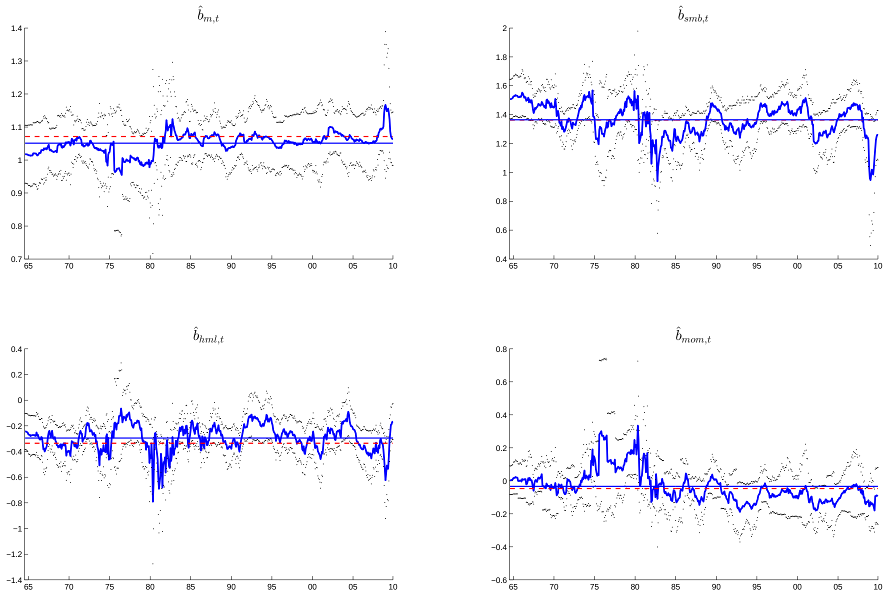
The table reports the p-values of the t-statistics computed on each component of vector $A\hat{\beta}_{2,i}$, estimated on the 25 FF portfolios by assuming the CAPM. For comparison purposes, we also report the p-values for the joint null hypothesis $\mathcal{H}_0^{\beta_{2,i}}$ (see Figure 23, Panel C).

A.13.3 Examples of estimated beta paths

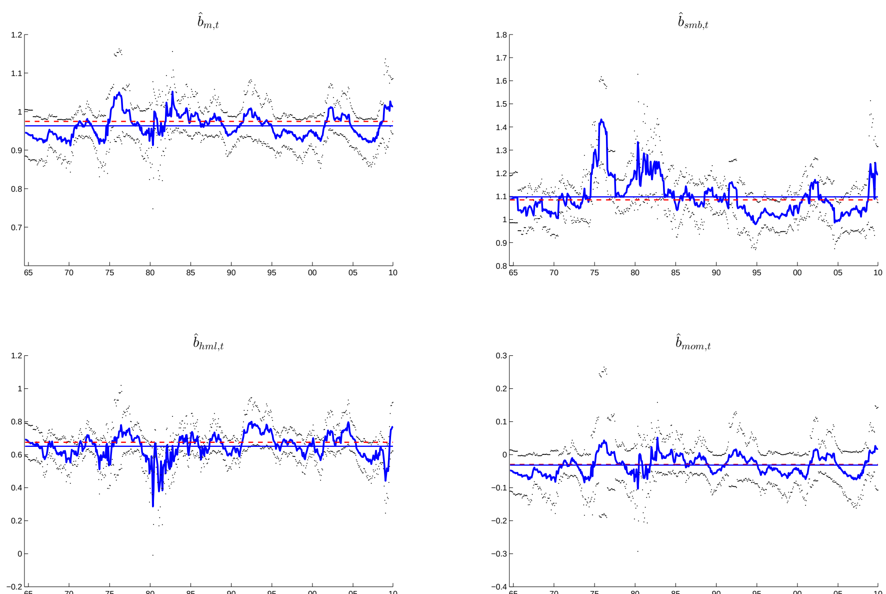
In this section, we plot the estimated path of time-varying betas for some of the 25 FF and 44 Indu. portfolios. In order to build confidence intervals for $\hat{b}_{i,t} = Q_{i,t}\hat{\beta}_{2,i}$, we use Equation (69) and deduce that $\sqrt{T}\left(\hat{b}_{i,t} - b_{i,t}\right) \Rightarrow N\left(0, \tau_i Q_{i,t} E_2' Q_{x,i}^{-1} S_{ii} Q_{x,i}^{-1} E_2 Q_{i,t}'\right)$. The examples of estimated paths show that the betas for the 25 FF portfolios are more stable than for the 44 Indu. portfolios. As expected, portfolios of small firms have overall a large value of beta w.r.t the smb factor, and portfolios of value firms have a large beta w.r.t. the hml factor (see Figures 25 and 26). We observe both cyclical and countercyclical variation in the FF portfolios betas. Some Indu. portfolios, such as those of the Agriculture, Fabricated Products, and Petroleum and Natural Gas sectors, show substantial countercyclical variation in their beta w.r.t. the value factor (see Figures 28 and 29).

Figure 25: Paths of the components of $\hat{b}_{i,t}$ with $n = 25$

Panel A: $i = 1$ (Small size, Low value)



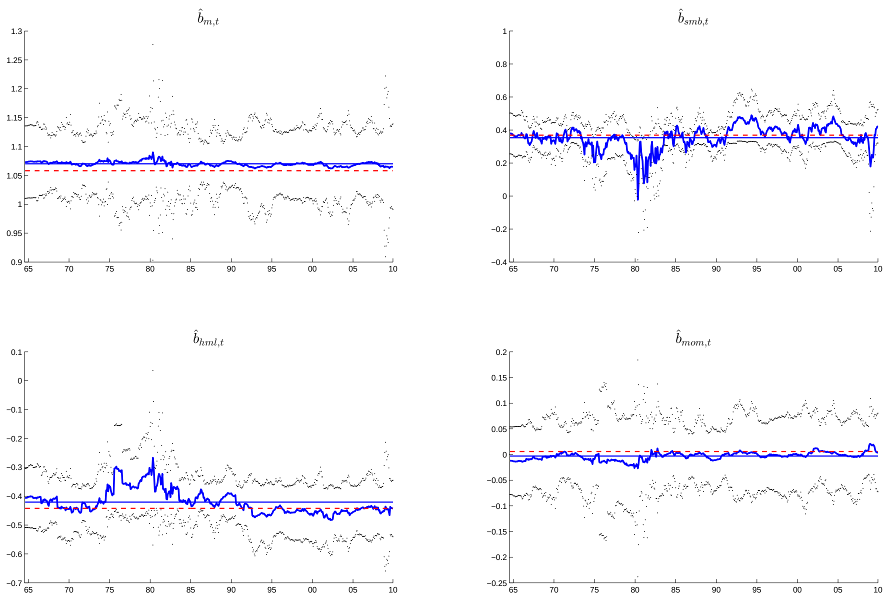
Panel B: $i = 5$ (Small Size, High value)



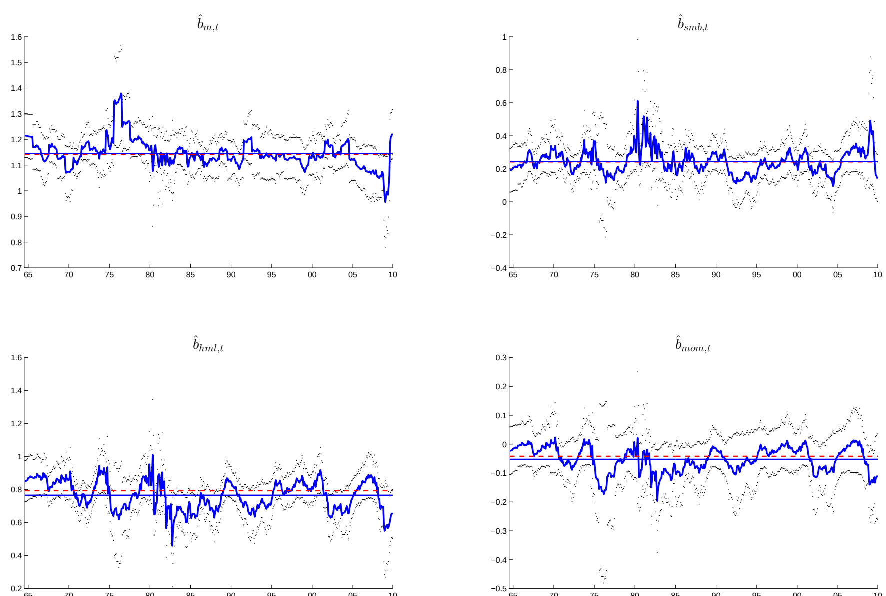
The figure plots the path of estimated time-varying betas $\hat{b}_{m,t}$, $\hat{b}_{smb,t}$, $\hat{b}_{hml,t}$ and $\hat{b}_{mom,t}$ and their pointwise confidence intervals at 95% probability level for two small size FF portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line).

Figure 26: Paths of the components of $\hat{b}_{i,t}$ with $n = 25$

Panel A: $i = 16$ (Size 4, Low value)



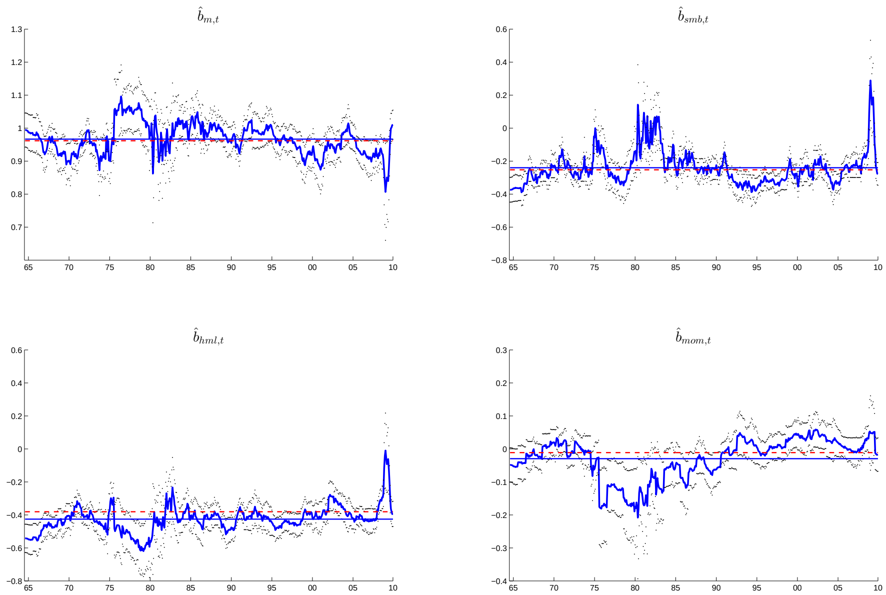
Panel B: $i = 20$ (Size 4, High value)



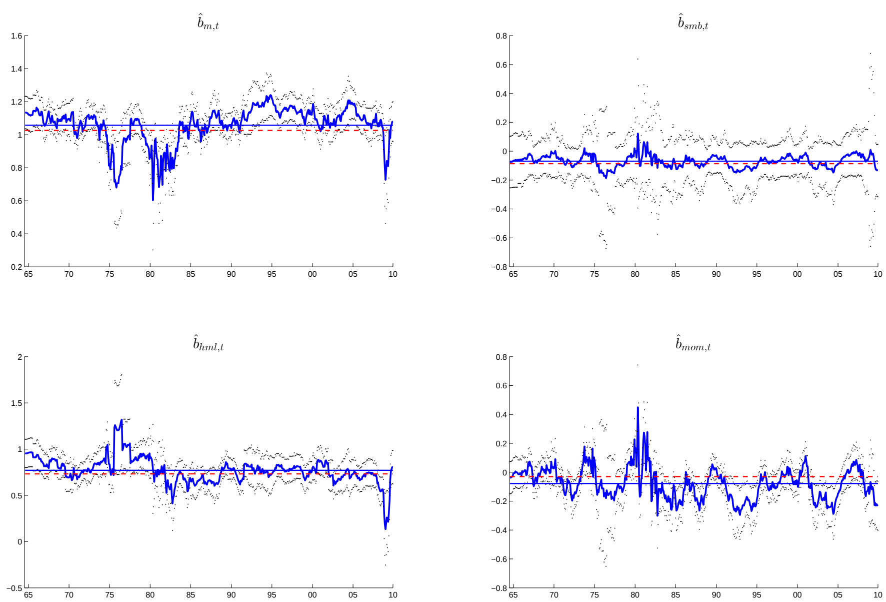
The figure plots the path of estimated time-varying betas $\hat{b}_{m,t}$, $\hat{b}_{smb,t}$, $\hat{b}_{hml,t}$ and $\hat{b}_{mom,t}$ and their pointwise confidence intervals at 95% probability level for two FF portfolios in the fourth size quintile. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line).

Figure 27: Paths of the components of $\hat{b}_{i,t}$ with $n = 25$

Panel A: $i = 21$ (Big Size, Low value)



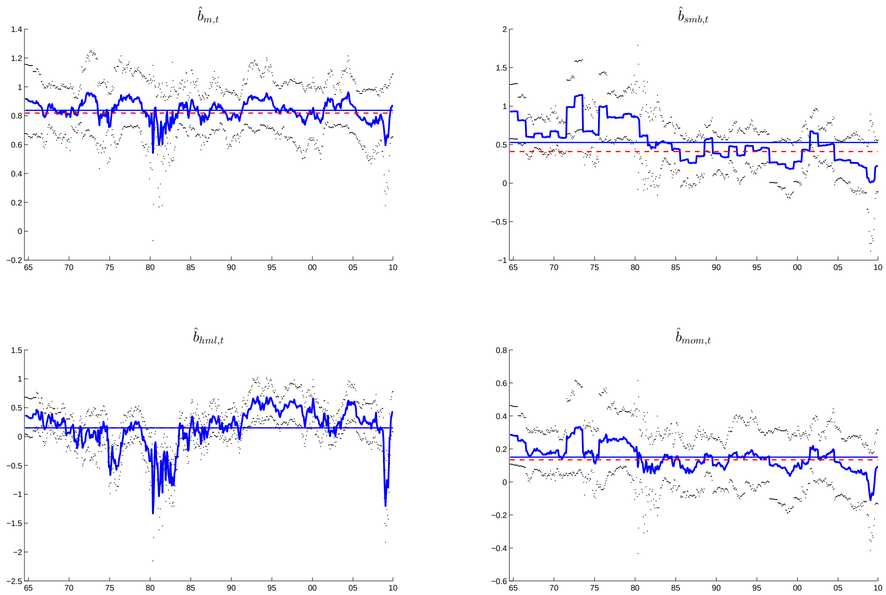
Panel B: $i = 25$ (Big size, High value)



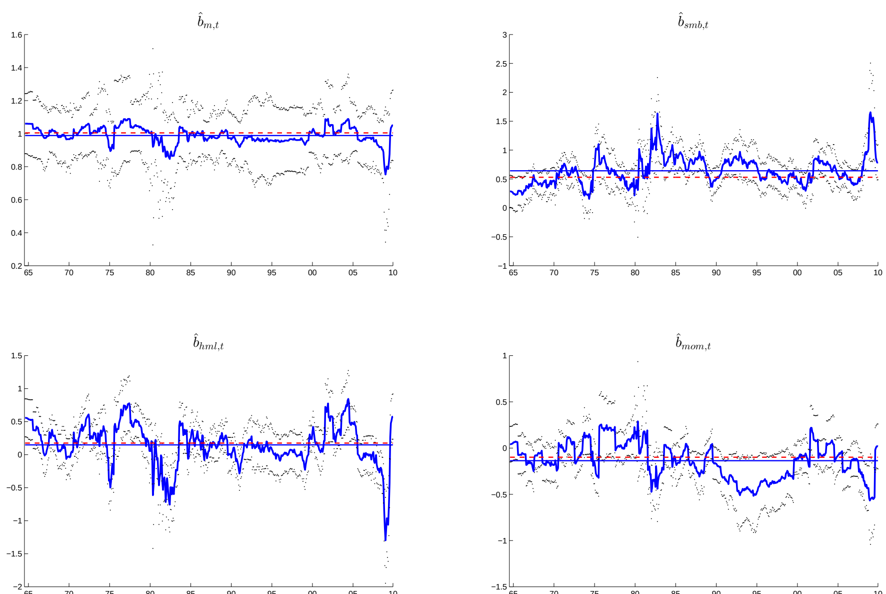
The figure plots the path of estimated time-varying betas $\hat{b}_{m,t}$, $\hat{b}_{smb,t}$, $\hat{b}_{hml,t}$ and $\hat{b}_{mom,t}$ and their pointwise confidence intervals at 95% probability level for two big size FF portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line).

Figure 28: Paths of the components of $\hat{b}_{i,t}$ with $n = 44$

Panel A: $i = 1$ (Agriculture)



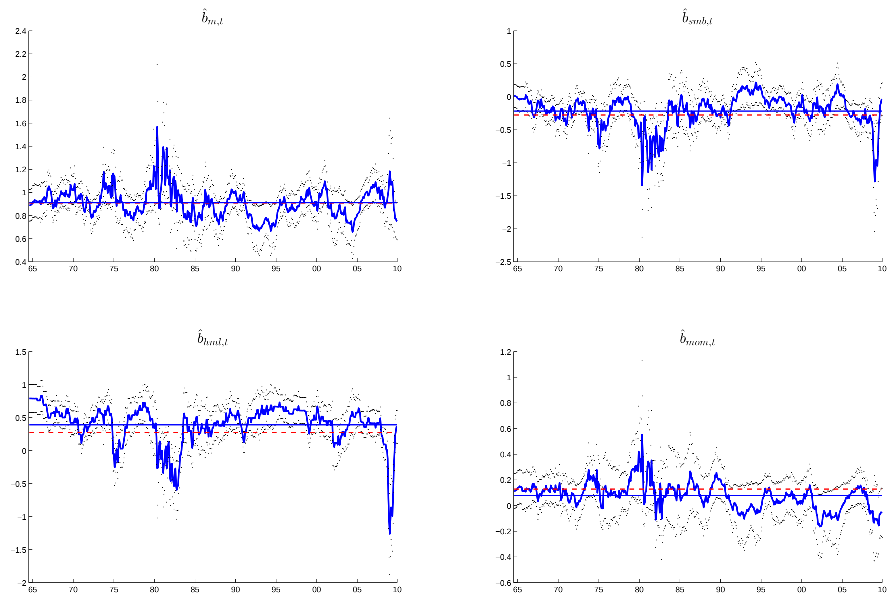
Panel B: $i = 20$ (Fabricated Products)



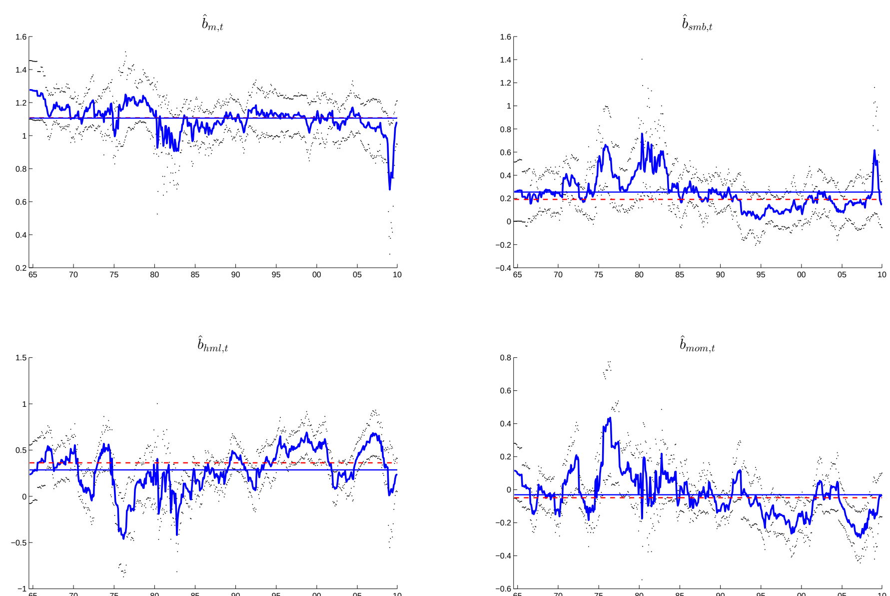
The figure plots the path of estimated time-varying betas $\hat{b}_{m,t}$, $\hat{b}_{smb,t}$, $\hat{b}_{hml,t}$ and $\hat{b}_{mom,t}$ and their pointwise confidence intervals at 95% probability level for two Indu. portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line).

Figure 29: Paths of the components of $\hat{b}_{i,t}$ with $n = 44$

Panel A: $i = 30$ (Petroleum and Natural Gas)



Panel B: $i = 40$ (Transportation)



The figure plots the path of estimated time-varying betas $\hat{b}_{m,t}$, $\hat{b}_{smb,t}$, $\hat{b}_{hml,t}$ and $\hat{b}_{mom,t}$ and their pointwise confidence intervals at 95% probability level for two Indu. portfolios. We also report the time-invariant (dashed horizontal line) and the average conditional estimate (solid horizontal line).

Appendix 14 Misspecification analysis

In this appendix, we first present theoretical results on the role of misspecification and aggregation (Section A.14.1) and we derive the pseudo-true value of the risk premia parameter when we estimate a potentially misspecified time-invariant model using either individual assets (Section A.14.2) or portfolios (Section A.14.3). Then, we estimate these pseudo-true values using our dataset (Section A.14.4).

A.14.1 The role of misspecification and aggregation

A potential explanation of the differences between the results on individual stocks and portfolios, as well as between sets of portfolios, is the uneven degree of misspecification of a given model across universes of assets. Using mimicking portfolio returns as observable factors and aggregating assets into portfolios may induce misspecification in the functional form of the beta dynamics. Risk premia estimated by the two-pass methodology from misspecified models converge to pseudo-true values. Estimation from individual stocks and portfolios may yield different pseudo-true values. In this section, we present theoretical and empirical arguments to support the plausibility of these claims for explaining the findings in Sections 4.2 and 4.3 of the paper.

Suppose that the Data Generating Process (DGP) for the excess returns in the continuum economy is:

$$R_t(\gamma) = c_t(\gamma) + d_t(\gamma)' h_t + \varepsilon_t(\gamma), \tag{70}$$

where h_t is a $r \times 1$ vector of “structural", or “economic", unknown factors with time-varying loadings $d_t(\gamma)$. The intercepts are $c_t(\gamma) = d_t(\gamma)'\mu_t$ for some stochastic vector μ_t because of the no-arbitrage restriction. We have $\mu_t = 0$ for tradable factors. In applying the two pass methodology, we approximate the unobservable factors by the excess returns of some mimicking portfolios. The market, Fama-French, and momentum factors are standard examples.

Let us formalize the concept of mimicking portfolio construction. Take a weighting function $w(\gamma, \omega)$, which is $\mathcal{F}_0$-measurable w.r.t. $\omega \in \Omega$ for a.e. $\gamma \in [0, 1]$, and Lebesgue measurable w.r.t. $\gamma \in [0, 1]$ for a.e. $\omega \in \Omega$, such that $\int w(\gamma, \omega) d\gamma = 1$ for a.e. $\omega \in \Omega$. Quantities $w_t(\gamma, \omega) = w[\gamma, S^{t-1}(\omega)]$, for γ varying, yield the portfolio weights $w_t(\gamma_i)/n_t^w$ at time t, where $n_t^w = \sum_i w_t(\gamma_i)$ is the weighted number of the n sampled assets included in the portfolio w at time t. The excess return of the portfolio w is $R_t^w = \frac{1}{n_t^w}\sum_i w_t(\gamma_i)R_t(\gamma_i)$. From Equation (70), we have:

$$R_t^w = (d_t^w)'(h_t + \mu_t) + \varepsilon_t^w, \tag{71}$$

with factor sensitivities $d_t^w = \frac{1}{n_t^w}\sum_i w_t(\gamma_i)d_t(\gamma_i)$ and an error term $\varepsilon_t^w = \frac{1}{n_t^w}\sum_i w_t(\gamma_i)\varepsilon_t(\gamma_i)$. We have that ε_t^w is close to zero for large n if the error terms of the individual assets feature weak cross-sectional dependence and the portfolio is sufficiently diversified. Thus, the $k \times 1$ vector f_t of excess returns from k diversified portfolios is close to $D_t(h_t + \mu_t)$, for some $k \times r$ matrix D_t which is measurable w.r.t. the information $\mathcal{F}_{t-1}$. To focus this section on specification analysis (see the next Appendix 15 for discussion on missing factor impact), we assume $k = r$, namely that the number of observable factors corresponds to the number of unknown factors, and we neglect approximation errors. Then, we have:

$$h_t + \mu_t = D_t^{-1} f_t. \tag{72}$$

for non redundant observable factors. Replacing Equation (72) into model (70) shows that the asset returns satisfy model (1) with factors f_t and sensitivities $b_t(\gamma) = (D_t^{-1})'d_t(\gamma)$. By construction, we get $\nu_t = 0$ because the factors f_t are returns of tradable portfolios. Thus, model (1) is correctly specified as long as we set the correct number of factors, even if the observable factors f_t do not correspond to the unknown factors h_t. Indeed, the vector f_t dynamically spans the true factor space. However, a constrained parametric model for the economic factor sensitivities, instead of a generic unconstrained $d_t(\gamma)$, does not necessarily transmit to the observable factor sensitivities. For instance, if the economic factor sensitivities are linear functions of some instruments, the observable factor sensitivities are not necessarily linear functions of these instruments. Choosing mimicking portfolio returns as observable factors jointly with a constrained parametrization can lead to a first source of misspecification.

A second potential source of misspecification comes from the aggregation of assets into portfolios. Let w^j for $j = 1, ..., m$ be a set of portfolios. We use the index j and the cardinality m for portfolios in order to distinguish them from the corresponding i and n for the fundamental assets. Under model (1) for the individual assets, the asset pricing restrictions yield the portfolio returns:

$$R_t^j = a_t^j + (b_t^j)' f_t + \varepsilon_t^j, \tag{73}$$

with factor sensitivities:

$$b_t^j = \frac{1}{n_t^j} \sum_i w_t^j(\gamma_i) b_t(\gamma_i), \tag{74}$$

intercepts $a_t^j = (b_t^j)'\nu_t$, and error terms $\varepsilon_t^j = \frac{1}{n_t^j} \sum_i w_t^j(\gamma_i)\varepsilon_t(\gamma_i)$. Model (73) is a factor model with the same factors as the original model for the individual assets, and time-varying alphas and betas. Hence, as observed in Section 2.2 for repackaging, we have robustness w.r.t. portfolio aggregation. However, if we choose a constrained parametric specification for the coefficients of a time-varying model, that parametric choice does not transmit easily under portfolio aggregation. First, the dynamics of the portfolio betas result from a combination of the dynamics of the individual stock betas and of the portfolio weights. Second, even with time-invariant portfolio weights, the aggregation of the asset specific instruments is complex, and results in models with portfolio specific instruments which involve unknown model parameters. For instance, let us consider the linear beta specification $b_{i,t} = B_i Z_{t-1} + C_i Z_{i,t-1}$ with a scalar stock specific instrument estimated in our empirical analysis, and equally-weighted portfolios, i.e. $w_t^j = 1/|A^j|$ for $\gamma \in A^j$, and 0 otherwise, for all j and t, where $A^j \subset [0,1]$ is a measurable set with non-zero measure $|A^j|$. Then, from (74), the portfolio betas are $b_t^j = B^j Z_{t-1} + C^j Z_{t-1}^j$, where the portfolio coefficients $B^j = \frac{1}{n^j} \sum_{i:\gamma_i \in A^j} B_i$ and $C^j = \frac{1}{n^j} \sum_{i:\gamma_i \in A^j} C_i$ are averages of the individual coefficients, n_j is the number of indices i with $\gamma_i \in A^j$, and the portfolio specific instrument $Z_{t-1}^j = \sum_{i:\gamma_i \in A^j} C_i Z_{i,t-1} \Big/ \sum_{i:\gamma_i \in A^j} C_i$ is a weighted average of the assets specific instruments, with weights involving the unknown coefficients C_i. If we use an ad-hoc aggregation scheme to define the portfolio specific instruments, the resulting model is in general misspecified. If we try to replace the unknown C_i with estimates to get a proxy for the Z_{t-1}^j, we need first to estimate the model for the individual assets and face an EIV problem. For the FF portfolios, misspecification of the beta dynamics may result from the time-varying portfolio weights and the ad-hoc aggregation scheme used to construct the portfolio specific instrument, namely the book-to-market equity of the portfolio as in Section 4.3 of the paper.

Under misspecification, the two-pass methodology may yield different pseudo-true values for the risk premia depending on the selected universe of assets. Let us assume that the DGP for the individual stock returns is given by model (1)-(3), with possibly time-varying betas and risk premia, but the researcher estimates a time-invariant model. For expository purposes, we focus on the OLS estimator in the second pass. We show in Section A.14.2 that the pseudo true value of parameter ν using individual stock returns is $\nu^* = \left(\int b^*(\gamma)b^*(\gamma)'dG(\gamma)\right)^{-1}\int b^*(\gamma)a^*(\gamma)dG(\gamma)$, where the pseudo-true values of sensitivities and intercepts are:

$$\begin{aligned} b^*(\gamma) &= \left[I_K + V[f_t]^{-1}Cov(f_t,\nu_t)\right]E[b_t(\gamma)] + E\left[\xi_t(b_t(\gamma) - E[b_t(\gamma)])\right], \\ a^*(\gamma) &= E\left[\nu_t - Cov(\nu_t,f_t)V[f_t]^{-1}f_t\right]'E[b_t(\gamma)] - E\left[\eta_t'(b_t(\gamma) - E[b_t(\gamma)])\right], \end{aligned}$$

and the matrix and vector processes ξ_t and η_t are defined by $\xi_t = V[f_t]^{-1}(f_t - E[f_t])(\nu_t + f_t)'$ and $\eta_t = \left(E[f_t]'V[f_t]^{-1}(f_t - E[f_t]) - 1\right)(\nu_t + f_t)$. Expectations, variances, and covariances are w.r.t. the DGP. The pseudo-true value ν^* is equal to the unconditional expectation $E[\nu_t]$ if the individual betas are uncorrelated with the conditional expectations of f_t and ν_t given $\mathcal{F}_{t-1}$, and process ν_t is uncorrelated with f_t. Then the pseudo-true risk premia vector is $\lambda^* = \nu^* + E[f_t] = E[\lambda_t]$. Here, even if the model is misspecified, there is no effect on the time-averaged risk premia. However, in general, time-variation distorts risk premia estimates. Even if the factors f_t are tradable, i.e., $\nu_t = 0$, we may have $\nu^* \neq 0$. The factors may appear as nontradable because of a misspecified time-invariant model as it is likely in Section 4.2.

If we estimate the time-invariant model using the returns on m portfolios w^j, with $j = 1, ..., m$, the pseudo-true value of ν becomes $\nu^{**} = \left(\sum_j b_j^* b_j^{*\prime}\right)^{-1}\sum_j b_j^* a_j^*$, where (see Section A.14.3)

$$b_j^* = \int E[w_t^j(\gamma)]b^*(\gamma)dG(\gamma) + \int Cov\left(\xi_t b_t(\gamma), w_t^j(\gamma)\right)dG(\gamma),$$

and

$$a_j^* = \int E[w_t^j(\gamma)]a^*(\gamma)dG(\gamma) - \int Cov\left(\eta_t' b_t(\gamma), w_t^j(\gamma)\right)dG(\gamma).$$

The pseudo-true portfolio loadings b_j^* are the sum of two components. The first one is an aggregate of the pseudo-true individual loadings $b^*(\gamma)$ weighted by the time-averaged portfolio weights $E[w_t^j(\gamma)]$. The second component is induced by the time-variation of the portfolio weights and its interaction with f_t, ν_t, and factor sensitivities. A similar comment applies to the pseudo-true portfolio intercepts a_j^*. If the portfolio weights are time-invariant, building portfolios corresponds to aggregating the individual pseudo-true alphas and betas. The portfolio aggregation effect is more complex if portfolio weights are time-varying. In general, the pseudo-true value ν^{**} depends on the number m of chosen portfolios and the weights $w_t^j(\gamma)$ they are built on, and we expect the pseudo-true values ν^{**} and ν^* not to be equal as the different estimated $\hat{\nu}$ in Table 1 Panel B may indicate. Besides, even if we observe that the portfolio betas are more stable over time, this does not imply that ν^{**} will be closer to zero than ν^*, when $\nu_t = 0$. We give a simple estimation exercise (see Section A.14.4) to check whether the numerical values for these pseudo-true values and their differences are compatible with the order of magnitude observed in Table 1 Panel B, including values close to zero in some cases. For the value factor, time variation in the portfolio weights can explain the large discrepancy between the pseudo-true values computed on the 25 FF portfolios and the individual stocks.

The above discussion concentrates on the impact of misspecification when the econometrician estimates a time-invariant model. Similar computations and remarks apply for estimation of misspecified time-varying models.

A.14.2 Pseudo-true value using individual assets

The pseudo true values of the regression coefficients are $\beta^*(\gamma) = (a^*(\gamma), b(\gamma)^{*\prime})' = Q_x^{-1}E[x_t R_t(\gamma)]$, for all $\gamma \in [0,1]$, where the expectation is w.r.t. the DGP. Let $\beta_i^* = \beta^*(\gamma_i)$. If the OLS estimator is used in the second pass, and matrix $E[b_i^* b_i^{*\prime}]$ is positive definite, the pseudo-true value of parameter ν is $\nu_1^* = E[b_i^* b_i^{*\prime}]^{-1}E[b_i^* a_i^*]$. The pseudo-true weights are $w_i^* = (v_i^*)^{-1}$ with $v_i^* = \tau_i c_{\nu_1^*}' Q_x^{-1} S_{ii}^* Q_x^{-1} c_{\nu_1^*}$, where $S_{ii}^* = E[(\varepsilon_{i,t}^*)^2 x_t x_t | \gamma_i]$ and $\varepsilon_{i,t}^* = R_{i,t} - x_t' \beta_i^*$. If the WLS estimator is used in the second pass, and matrix $E[w_i^* b_i^* b_i^{*\prime}]$ is positive definite, the pseudo-true value of parameter ν is

$$\nu^* = E[w_i^* b_i^* b_i^{*\prime}]^{-1} E[w_i^* b_i^* a_i^*]. \tag{75}$$

Then, the pseudo-true value of the risk premia vector is $\lambda^* = \nu^* + E[f_t]$.

Let $\hat{\nu}$ be the estimator defined in Equation (14) of the paper, using the first-pass estimators $\hat{\beta}_i$ and the weights $\hat{w}_i$ for the second pass. The next lemma states that the estimators converge to the corresponding pseudo-true values and is proved at the end of this subsection.

Lemma 17 *Suppose Assumptions A.1b), SC.1-SC.2, B.1, B.4, B.5 hold. Moreover, let* $\sup_{\gamma \in [0,1]} P\left[\left\| \frac{1}{T} \sum_t I_t(\gamma) x_t \varepsilon_t^*(\gamma) \right\| \geq \delta \right]$ *satisfy the large deviation bound in Assumption B.1, for any* $\delta > 0$ *and* $T \in \mathbb{N}$, *where* $\varepsilon_t^*(\gamma) = R_t(\gamma) - x_t' \beta^*(\gamma)$ *is the pseudo-true error. Then, as* $n, T \to \infty$ *such that*

$n = O(T^{\bar{\gamma}})$ *for* $\bar{\gamma} > 0$*, we have: (i)* $\sup_i \mathbf{1}_i^{\chi}\|\hat{\beta}_i - \beta_i^*\| = o_p(1)$*; (ii)* $\frac{1}{n}\sum_i \|\hat{w}_i - w_i^*\| = o_p(1)$*; (iii)* $\hat{\nu} = \nu^* + o_p(1)$.

Let us now derive more explicit expressions for the components $a^*(\gamma)$ and $b^*(\gamma)$ of the pseudo-true coefficients vector. We have:

$$b^*(\gamma) = V[f_t]^{-1}Cov(f_t, R_t(\gamma)), \quad a^*(\gamma) = E[R_t(\gamma)] - E[f_t]'b^*(\gamma), \tag{76}$$

for all $\gamma \in [0,1]$. From $R_t(\gamma) = (f_t + \nu_t)'b_t(\gamma) + \varepsilon_t(\gamma)$, we have:

$$\begin{aligned} E[R_t(\gamma)] &= E[(f_t + \nu_t)'b_t(\gamma)] \\ &= E[\nu_t]'E[b_t(\gamma)] + E[f_t]'E[b_t(\gamma)] + E[(f_t + \nu_t)'(b_t(\gamma) - E[b_t(\gamma)])], \end{aligned}$$

and:

$$\begin{aligned} Cov(f_t, R_t(\gamma)) &= Cov(f_t, (f_t + \nu_t)'b_t(\gamma)) \\ &= (V[f_t] + Cov(f_t, \nu_t))\, E[b_t(\gamma)] + Cov(f_t, (f_t + \nu_t)'(b_t(\gamma) - E[b_t(\gamma)])) \\ &= (V[f_t] + Cov(f_t, \nu_t))\, E[b_t(\gamma)] + E[(f_t - E[f_t])(f_t + \nu_t)'(b_t(\gamma) - E[b_t(\gamma)])]. \end{aligned}$$

Then, by replacing into (76) and rearranging terms, we get:

$$b^*(\gamma) = \left[I_K + V[f_t]^{-1}Cov(f_t, \nu_t)\right] E[b_t(\gamma)] + E\left[\xi_t(b_t(\gamma) - E[b_t(\gamma)])\right], \tag{77}$$

$$a^*(\gamma) = E\left[\nu_t - Cov(\nu_t, f_t)V[f_t]^{-1}f_t\right]' E[b_t(\gamma)] - E\left[\eta_t'(b_t(\gamma) - E[b_t(\gamma)])\right], \tag{78}$$

for all $\gamma \in [0,1]$, where $\xi_t = V[f_t]^{-1}(f_t - E[f_t])(\nu_t + f_t)'$ and $\eta_t = \left(E[f_t]'V[f_t]^{-1}(f_t - E[f_t]) - 1\right)(\nu_t + f_t)$.

Proof of Lemma 17: We have $\hat{\beta}_i - \beta_i^* = \tau_{i,T}\hat{Q}_{x,i}^{-1}\frac{1}{T}\sum_t I_{i,t}x_t\varepsilon_{i,t}^*$. Then part (i) follows by similar arguments as in the proof of Lemma 3 (i) for a well-specified time-invariant model. The proof of part (ii) is similar to the proof of Lemma 3 (iii) and is omitted. Finally, using parts (i)-(ii) of this lemma, Assumption SC.2 and the LLN, we have:

$$\frac{1}{n}\sum_i \hat{w}_i\hat{b}_i\hat{b}_i' = \frac{1}{n}\sum_i w_i^*b_i^*b_i^{*\prime} + o_p(1) = E[w_i^*b_i^*b_i^{*\prime}] + o_p(1),$$

and

$$\frac{1}{n}\sum_i \hat{w}_i\hat{b}_i\hat{a}_i = \frac{1}{n}\sum_i w_i^*b_i^*a_i^* + o_p(1) = E[w_i^*b_i^*a_i^*] + o_p(1).$$

Since matrix $E[w_i^*b_i^*b_i^{*\prime}]$ is invertible, part (iii) follows.

A.14.3 Pseudo-true value using portfolios

Let us now assume that we estimate the time-invariant model on a set of m portfolios w^j, with $j = 1, ..., m$. If the portfolios are well diversified, and the number of underlying assets n tends to infinity, the idiosyncratic error terms ε_t^j vanish in Equation (73). Then, the portfolio returns are $R_t^j = (b_t^j)'(f_t + \nu_t)$, where the portfolio sensitivities are:

$$b_t^j = \int w_t^j(\gamma) b_t(\gamma) dG(\gamma). \tag{79}$$

Then, the pseudo-true values of the regression coefficients are obtained along the lines of Section A.14.2 replacing $R_t(\gamma)$ with R_t^j, and $b_t(\gamma)$ with b_t^j. We get $\beta^{*j} = (a^{*j}, (b^{*j})')'$ where:

$$\begin{aligned} b^{*j} &= \left[I_K + V[f_t]^{-1} Cov(f_t, \nu_t)\right] E[b_t^j] + E\left[\xi_t (b_t^j - E[b_t^j])\right], & (80) \\ a^{*j} &= E\left[\nu_t - Cov(\nu_t, f_t) V[f_t]^{-1} f_t\right]' E[b_t^j] - E\left[\eta_t' (b_t^j - E[b_t^j])\right], & (81) \end{aligned}$$

for all $j = 1, ..., m$. Then, when the OLS estimator is used in the second pass, the pseudo-true value of parameter ν is $\nu_1^* = \left(\sum_j b^{*j}(b^{*j})'\right)^{-1} \sum_j b^{*j} a^{*j}$. When the WLS estimator is used, the pseudo-true value of parameter ν is $\nu^* = \left(\sum_j (v^{*j})^{-1} b^{*j}(b^{*j})'\right)^{-1} \sum_j (v^{*j})^{-1} b^{*j} a^{*j}$, where the reciprocal of the pseudo-true weights are $v^{*j} = c'_{\nu_1^*} Q_x^{-1} S^{*j} Q_x^{-1} c_{\nu_1^*}$, with $S^{*j} = E[(\varepsilon_t^{*j})^2 x_t x_t]$ and $\varepsilon_t^{*j} = R_t^j - x_t' \beta^{*j}$.

Let us now derive the expressions of the pseudo-true regression coefficients given in Section A.14.1. From (79), we have:

$$E[b_t^j] = \int E[w_t^j(\gamma)] E[b_t(\gamma)] dG(\gamma) + \int Cov\left(b_t(\gamma), w_t^j(\gamma)\right) dG(\gamma),$$

and:

$$\begin{aligned} b_t^j - E[b_t^j] &= \int E[w_t^j(\gamma)] \left(b_t(\gamma) - E[b_t(\gamma)]\right) dG(\gamma) \\ &\quad + \int \left(w_t^j(\gamma) - E[w_t^j(\gamma)]\right) b_t(\gamma) dG(\gamma) - \int Cov\left(b_t(\gamma), w_t^j(\gamma)\right) dG(\gamma). \end{aligned}$$

By replacing into (80), we get:

$$
\begin{aligned}
b^{*j} &= \int E[w_t^j(\gamma)]b^*(\gamma)dG(\gamma) \\
&\quad + \left[I_K + V[f_t]^{-1}Cov(f_t,\nu_t) - E[\xi_t]\right]\int Cov\left(b_t(\gamma), w_t^j(\gamma)\right)dG(\gamma) \\
&\quad + \int Cov\left(\xi_t b_t(\gamma), w_t^j(\gamma)\right)dG(\gamma).
\end{aligned}
$$

Since $E[\xi_t] = I_K + V[f_t]^{-1}Cov(f_t,\nu_t)$, the second term in the RHS vanishes, and we get the expression of b^{*j} given in Section A.14.1. The proof of the expression of a^{*j} is similar, by using $E[\eta_t] = -E\left[\nu_t - Cov(\nu_t, f_t)V[f_t]^{-1}f_t\right]$.

A.14.4 Empirical pseudo-true values

In Table 22, we report the estimates of the pseudo-true values of parameter vector ν in a time-invariant four-factor model obtained with the individual stocks, the 25 FF portfolios, and the 44 Indu. portfolios. We get the estimates by replacing the expectations in Equations (75), (77)-(78), and (80)-(81) with sample averages. To assess the contributions of misspecifications along different directions, we consider several alternative assumptions on the DGP for process ν_t and factor sensitivities $b_t(\gamma)$ of the individual stocks. Specifically, we assume that the vector ν_t is either (i) time-invariant and equal to zero, or (ii) time-invariant and equal to the time-average $\bar{\nu} = [1.3772,\ -0.2122,\ -6.1630,\ -2.5507]'$ of the estimates $\hat{\nu_t}$ obtained with the time-varying model applied on individual stocks in Section 4.3, or (iii) time-varying and equal to the estimates $\hat{\nu_t}$. Furthermore, we assume that the betas of the $n^{\chi} = 3,900$ individual stocks after trimming are either (a) time-invariant and equal to the time averages of the estimates $\hat{b}_{i,t}$ obtained with the time-varying model in Section 4.3, or (b) time-varying and equal to the estimates $\hat{b}_{i,t}$. The combination of (i)-(iii) and (a)-(b) yields six alternative (empirical) DGPs. We compute the portfolio betas by aggregating the betas of the $3,900$ individual stocks using weights $\hat{w}_{i,t}^j$. These weights are obtained by following the methodology underlying the FF and Indu. portfolios applied to the $3,900$ assets of our trimmed sample. To assess the contribution of time-varying portfolio weights, we also compute the pseudo-true values using the returns of 25 and 44 portfolios with time-invariant weights equal to the time-averages of the corresponding weights $\hat{w}_{i,t}^j$. Thus, the pseudo-true values are computed for five different universes of assets.

For the DGPs with time constant $b_{i,t}$ and ν_t, the time-invariant model is correctly specified on individual stocks. This explains why the (pseudo-)true values of ν with individual stocks, and with time constant portfolio weights, coincide in the first and third subpanels. Moreover, Equations (77)-(78) and (80)-(81) imply that these pseudo-true values of ν coincide also when ν_t is time-varying but the individual stocks betas are constant, as observed in the fifth subpanel. Instead, the pseudo-true values with time-varying portfolio weights differ from the pseudo-true values with individual stocks for all DGPs. The largest differences across universes of assets are observed for the value and momentum factors. We get a substantial difference between $\nu^*_{hml} = -6.1636$ on the individual stocks and $\nu^{**}_{hml} = -3.0085$ on the 25 FF portfolios (with time-varying weights) already for the DGP with constant $\nu_t = \bar{\nu}$ and constant $b_{i,t}$. The five pseudo-true values for ν_{hml} do not change a lot when we move to DGPs with time-variation in ν_t and/or $b_{i,t}$. Moreover, the estimate of ν_{hml} on the 25FF portfolios with time-varying weights are asymptotically larger than the estimates with constant weights. These findings suggest that, for the value factor, the difference between the results with the individual stocks and the FF portfolios is due mainly to time variation in the portfolio weights. For the momentum factor, the largest discrepancies between individual stocks and FF portfolios are observed for the DGPs with time-varying betas and weights. The pseudo-true values for the 44 Indu. portfolios are more similar to the the pseudo-true values for individual stocks.

Table 22: Estimated pseudo-true values of parameter ν for the four-factor model

		$n = 9,936$	$n = 25$		$n = 44$	
			CW	TVW	CW	TVW
$\nu_t = 0$, $b_{i,t}$ constant	ν^*_m	0.0000	0.0000	-0.3427	0.0000	-0.0801
	ν^*_{smb}	0.0000	0.0000	0.6167	0.0000	0.1843
	ν^*_{hml}	0.0000	0.0000	1.1304	0.0000	-0.4866
	ν^*_{mom}	0.0000	0.0000	0.8850	0.0000	-2.3739
$\nu_t = 0$, $b_{i,t}$ time-varying	ν^*_m	-0.0251	1.5815	-0.0349	0.5738	-0.3040
	ν^*_{smb}	0.6486	0.7998	0.8877	0.6075	1.2729
	ν^*_{hml}	-1.1835	-4.9452	0.6012	-0.6365	-0.8209
	ν^*_{mom}	-4.5639	-1.0871	-1.4821	-3.3692	-5.5221
$\nu_t = \bar{\nu}$, $b_{i,t}$ constant	ν^*_m	1.3772	1.3772	0.4453	1.3772	1.0312
	ν^*_{smb}	-0.2122	-0.2122	0.4779	-0.2122	0.0657
	ν^*_{hml}	-6.1636	-6.1636	-3.0085	-6.1636	-5.8395
	ν^*_{mom}	-2.5507	-2.5507	-0.7216	-2.5507	-4.5657
$\nu_t = \bar{\nu}$, $b_{i,t}$ time-varying	ν^*_m	1.3406	2.6374	0.6123	1.6079	0.9199
	ν^*_{smb}	0.1490	0.1940	0.7492	0.1824	0.8432
	ν^*_{hml}	-6.5468	-9.8461	-3.4016	-6.1935	-6.4573
	ν^*_{mom}	-6.6899	-3.5831	-2.6132	-5.4675	-8.0675
$\nu_t = \hat{\nu}_t$, $b_{i,t}$ constant	ν^*_m	1.3788	1.3788	0.8521	1.3788	1.0816
	ν^*_{smb}	-0.2158	-0.2158	0.4970	-0.2158	0.1172
	ν^*_{hml}	-6.1291	-6.1291	-3.9565	-6.1291	-5.9395
	ν^*_{mom}	-2.4741	-2.4741	-0.9824	-2.4741	-4.2506
$\nu_t = \hat{\nu}_t$, $b_{i,t}$ time-varying	ν^*_m	1.0201	1.5269	-0.0080	1.4433	0.6526
	ν^*_{smb}	0.1678	0.1870	0.8511	-0.3721	0.6996
	ν^*_{hml}	-6.0848	-8.1776	-2.6871	-6.6668	-6.5043
	ν^*_{mom}	-4.8815	-3.9304	-1.6555	-6.0449	-7.4999

The table contains the annualized estimates of the pseudo-true values of parameter ν for the market (ν^*_m), size (ν^*_{smb}), book-to-market (ν^*_{hml}) and momentum (ν^*_{mom}) factors. We report the estimates ν^* for individual stocks ($n = 9,936$, $n^\chi = 3,900$), the 25 FF and 44 Indu. portfolios as base assets for several DGPs. For portfolios, we report both the estimates with time-varying portfolio weights (TVW) and the estimates obtained assuming time constant weights (CW).

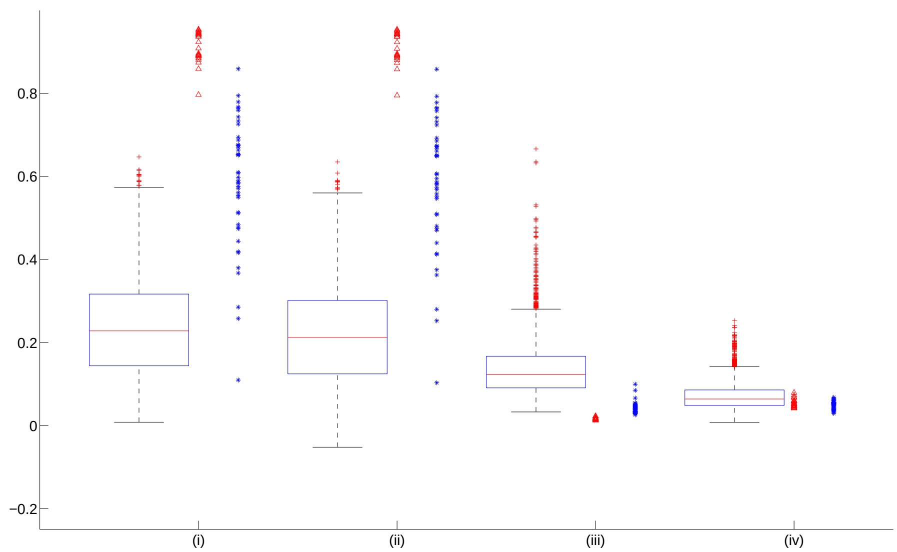
Appendix 15 Empirical analysis of idiosyncratic risks

In this appendix, we compare the cross-sectional distributions of $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$, the idiosyncratic risk (square root of residual variance), and the estimated time-series coefficient of determination $\hat{\rho}_i^2$ (ratio of explained variance and total variance) for the time-varying specifications assuming the four-factor model for the excess returns. We can view those estimates as measures of limits-to-arbitrage and missing factor impact. We relate these measures to stock characteristics such as size, book-to-market, and sample size. For each asset (either stock, or portfolio) i, we compute four measures: (i) the estimated time-series coefficient of determination $\hat{\rho}_i^2 = \frac{ESS_i}{TSS_i}$, where $ESS_i = \sum_t I_{i,t}\left(\hat{R}_{i,t} - \bar{\hat{R}}_i\right)^2$, with $\hat{R}_{i,t} = \hat{\beta}_i' x_{i,t}$ and $\bar{\hat{R}}_i = \frac{1}{T_i}\sum_t I_{i,t}\hat{R}_{i,t}$, and $TSS_i = \sum_t I_{i,t}\left(R_{i,t} - \bar{R}_i\right)^2$, with $\bar{R}_i = \frac{1}{T_i}\sum_t I_{i,t}R_{i,t}$; (ii) the estimated adjusted R^2 defined by $\hat{\rho}_{ad,i}^2 = 1 - \frac{(T_i-1)}{(T_i-d)}\left(1-\hat{\rho}_i^2\right)$; (iii) the idiosyncratic risk $IdiVol_i = \sqrt{\frac{RSS_i}{T_i}}$, with $RSS_i = \sum_t I_{i,t}\hat{\varepsilon}_{i,t}^2$; (iv) the systematic risk $SysRisk_i = \sqrt{\frac{ESS_i}{T_i}}$.

Figures 30 and 31 compare the cross-sectional distributions of the four measures (i)-(iv) computed on the time-invariant and time-varying four-factor models using the individual stocks, 25 FF and 44 Indu. portfolios as base assets. For comparison purposes, the cross-sectional distributions for individual stocks in both figures refer to the $n^{\chi} = 3,900$ stocks used in the estimation of the time-varying specification after trimming. The time-series (adjusted) $\hat{\rho}_i^2$ of the 25 FF portfolios are all larger than 0.80. The estimates $\hat{\rho}_i^2$ of the individual stocks are typically much smaller, with a median below 0.30. As expected, the excess returns of individual stocks also have larger idiosyncratic volatilities. The time-series adjusted $\hat{\rho}_i^2$ of individual stocks tend to be a bit larger in the time-varying model than in the time-invariant one, as a result of the explanatory power that we gain by allowing for beta dynamics. Figures 30 and 31 show that the use of the FF portfolios also shrinks the dispersion of $\hat{\rho}_i^2$, $IdiVol_i$, and $SysRisk_i$, by a large amount. The distributions for the individual stocks and the 44 Indu. portfolios are comparable and share a wide support. Figure 32 plots the cross-sectional distributions of $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$ for the three universes of assets. We observe a huge heterogeneity in $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$ for the individual stocks in Figure 32, similar to the one observed on $IdiVol_i$ in Figure 31. We may face the presence of limits-to-arbitrage and missing factors in that case. On the contrary, the estimates $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$ are concentrated close to zero for the 25 FF and 44 Indu. portfolios.

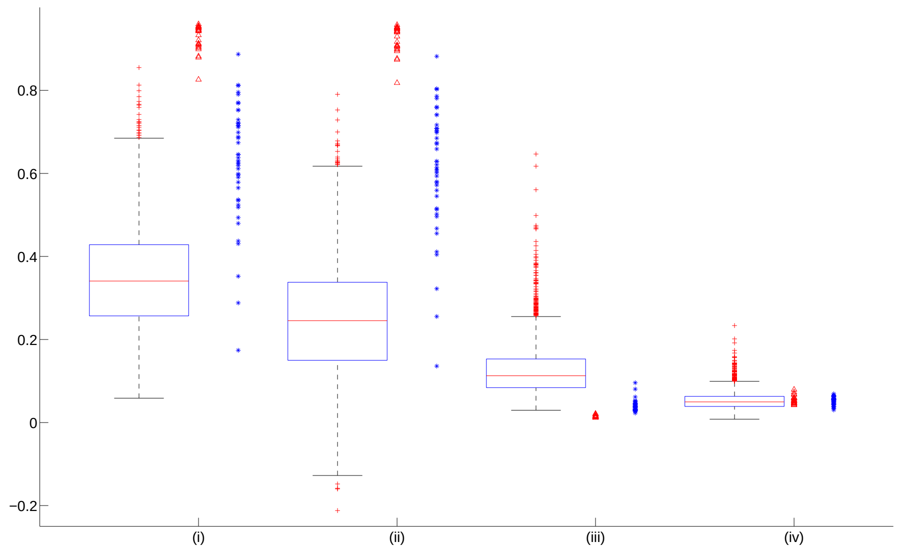
Figures 33-35 plot the idiosyncratic risks $IdiVol_i$ and systematic risks $SysRisk_i$ versus the estimated coefficients of determination $\hat{\rho}_i^2$. These three quantities are related by $\hat{\rho}_i^2 = \dfrac{SysRisk_i^2}{SysRisk_i^2 + IdiVol_i^2}$. Figures 36-44 plot the averages over time of size, $\bar{MC}_i = \dfrac{1}{T_i}\sum_t I_{i,t} mc_{i,t}$, and book-to-market $\bar{BM}_i = \dfrac{1}{T_i}\sum_t I_{i,t} bm_{i,t}$, versus $\hat{\rho}_i^2$, $IdiVol_i$ and $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$. Figure 45 plots the relation between the inverse standardized time-series dimension $\tau_{i,T} = \dfrac{T}{T_i}$ and the idiosyncratic and systematic risks $IdiVol_i$ and $SysRisk_i$, the estimated coefficients of determination $\hat{\rho}_i^2$, the estimated $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$, and the time averaged size and book-to-market $\bar{MC}_i$ and $\bar{BM}_i$. For the individual stocks, we report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10% on Figures 33, 36, 39, 42 and 45. Figures 34-35, 37-38, 40-41, 43-44 do not seem to deliver a clear relationship between $\left(\hat{\rho}_i^2, IdiVol_i\right)$, $\left(\hat{\rho}_i^2, SysRisk_i\right)$, $\left(\bar{MC}_i, \hat{\rho}_i^2\right)$, $\left(\bar{BM}_i, \hat{\rho}_i^2\right)$, $\left(\bar{MC}_i, IdiVol_i\right)$, $\left(\bar{BM}_i, IdiVol_i\right)$, $\left(\bar{MC}_i, \hat{\beta}_{1,i}'\hat{\beta}_{1,i}\right)$, and $\left(\bar{BM}_i, \hat{\beta}_{1,i}'\hat{\beta}_{1,i}\right)$ when we examine the 25 FF and 44 Indu. portfolios, except perhaps a negative association between $\hat{\rho}_i^2$ and $IdiVol_i$ for the 44 Indu. portfolios. This lack of conclusive evidence is probably due to the small number of points and aggregation effects. On the contrary, for the individual stocks, the linear quantile regressions show a positive relationship for $\left(\hat{\rho}_i^2, SysRisk_i\right)$, $\left(\bar{MC}_i, \hat{\rho}_i^2\right)$, and a negative relationship for $\left(\hat{\rho}_i^2, IdiVol_i\right)$, $\left(\bar{MC}_i, IdiVol_i\right)$, and $\left(\bar{MC}_i, \hat{\beta}_{1,i}'\hat{\beta}_{1,i}\right)$. Moreover, the linear quantile regressions at 50% or higher show a negative relationship for $\left(\bar{BM}_i, IdiVol_i\right)$ and $\left(\bar{BM}_i, \hat{\beta}_{1,i}'\hat{\beta}_{1,i}\right)$ in Figures 39 and 42. In Figure 45, we observe a positive relationship for $(\tau_{i,T}, IdiVol_i)$, $(\tau_{i,T}, SysRisk_i)$, $\left(\tau_{i,T}, \hat{\beta}_{1,i}'\hat{\beta}_{1,i}\right)$, and a negative relationship for $\left(\tau_{i,T}, \bar{MC}_i\right)$. Our preliminary results based on linear quantile regressions reveal that stocks with small size tend to yield large $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$, large idiosyncratic risks, and small estimated $\hat{\rho}_i^2$. We also find that firms with short observation periods (i.e., large $\tau_{i,T}$) tend to be associated with large values of both idiosyncratic and systematic risks (with a larger proportion of systematic risk to total risk), large $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$, as well as small market capitalisation. The results observed for $IdiVol_i$ and $\hat{\beta}_{1,i}'\hat{\beta}_{1,i}$ are similar as expected from their interpretation as measures of limits-to-arbitrage and missing factor impact. Measuring and understanding limits-to-arbitrage and missing factor impact on individual stocks certainly awaits more work.

Figure 30: Cross-sectional distributions of $\hat{\rho}_i^2$, $\hat{\rho}_{ad,i}^2$, $IdiVol_i$, and $SysRisk_i$ for the time-invariant four-factor model



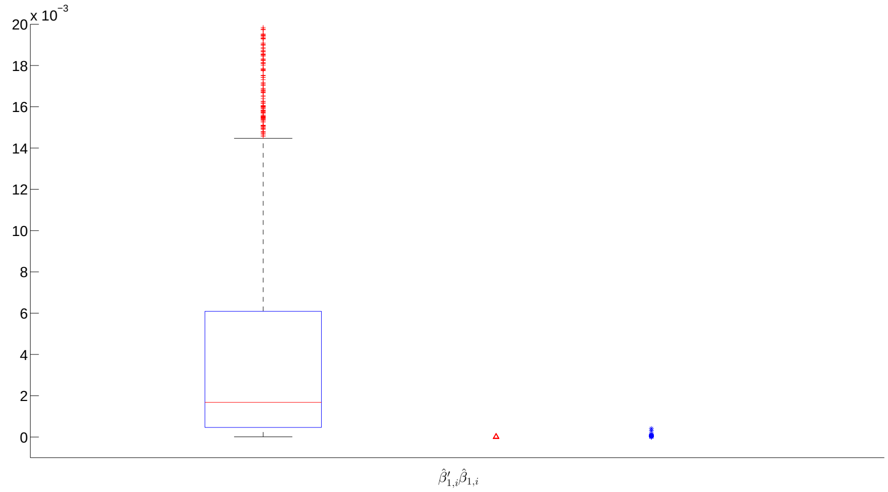
The figure displays the cross-sectional distributions of (i) the estimated coefficients of determination $\hat{\rho}_i^2$, (ii) the estimated adjusted coefficients of determination $\hat{\rho}_{ad,i}^2$, (iii) the idiosyncratic risks $IdiVol_i$, and (iv) the systematic risks $SysRisk_i$ for the individual stocks (box-plots), the 25 FF portfolios (red triangles) and the 44 Indu. portfolios (blue stars). Estimates are for the time-invariant four-factor model. For comparison purposes, the cross-sectional distribution for individual stocks refers to the $n^{\chi} = 3,900$ stocks that are used in the estimation of the time-varying model after trimming.

Figure 31: Cross-sectional distributions of $\hat{\rho}_i^2$, $\hat{\rho}_{ad,i}^2$, $IdiVol_i$, and $SysRisk_i$ for the time-varying four-factor model



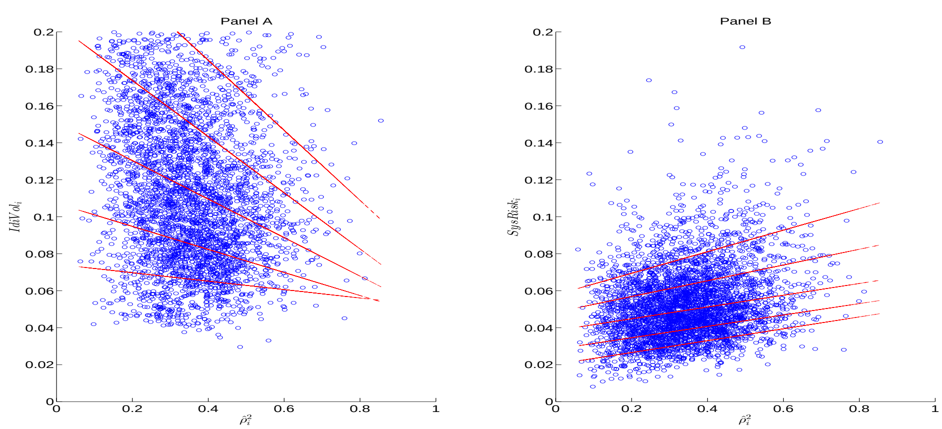
The figure displays the cross-sectional distributions of (i) the estimated coefficients of determination $\hat{\rho}_i^2$, (ii) the estimated adjusted coefficients of determination $\hat{\rho}_{ad,i}^2$, (iii) the idiosyncratic risks $IdiVol_i$, and (iv) the systematic risks $SysRisk_i$ for the $n^{\chi} = 3,900$ individual stocks (box-plots), the 25 FF portfolios (red triangles) and the 44 Indu. portfolios (blue stars). Estimates are for the time-varying four-factor model.

Figure 32: Cross-sectional distributions of $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ for the time-varying four-factor model



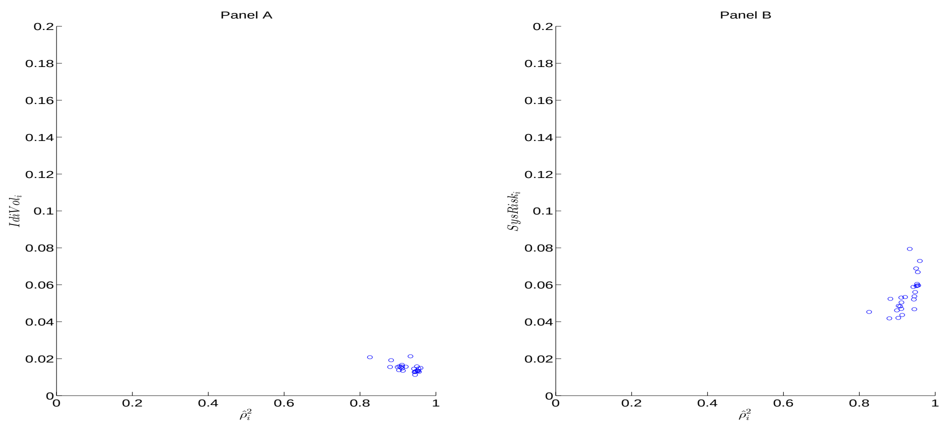
The figure plots the cross-sectional distributions of $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ for the $n^{\chi} = 3,900$ individual stocks (box-plot), the 25 FF portfolios (red triangles) and the 44 Indu. portfolios (blue stars). Estimated $\hat{\beta}_{1,i}$ are for the time-varying four-factor model.

Figure 33: $\hat{\rho}_i^2$ vs $IdiVol_i$ and $SysRisk_i$ for the $n^{\chi} = 3,900$ stocks



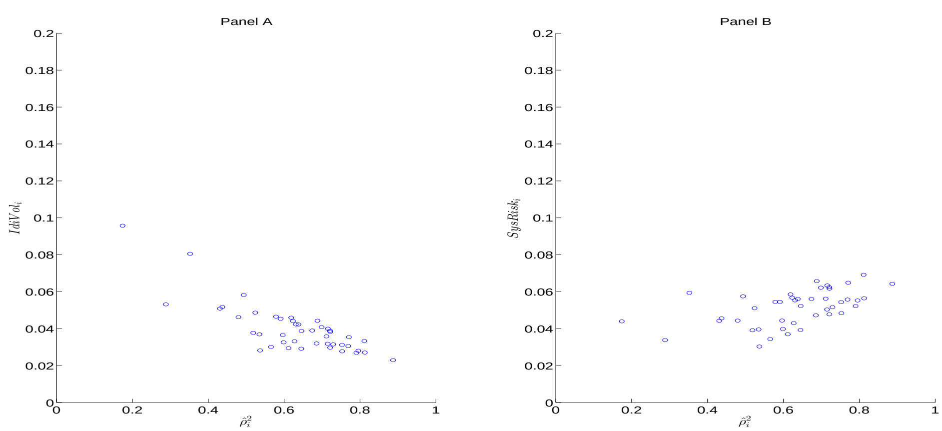
The figure plots the estimated coefficients of determination $\hat{\rho}_i^2$ w.r.t. the idiosyncratic risks $IdiVol_i$ (Panel A) and the systematic risks $SysRisk_i$ (Panel B) computed on the time-varying four-factor model using the $n^{\chi} = 3,900$ individual stocks. We report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%.

Figure 34: $\hat{\rho}_i^2$ vs $IdiVol_i$ and $SysRisk_i$ for the 25 FF portfolios



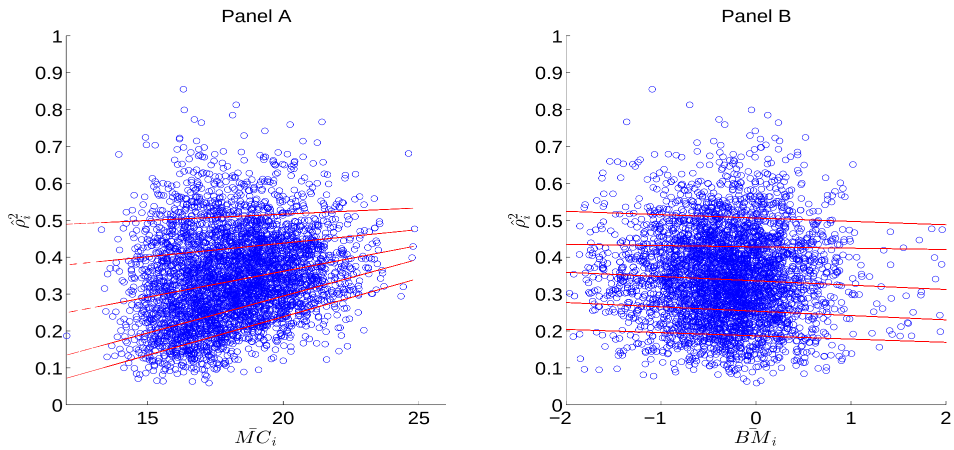
The figure plots the estimated coefficients of determination $\hat{\rho}_i^2$ w.r.t. the idiosyncratic risks $IdiVol_i$ (Panel A) and the systematic risks $SysRisk_i$ (Panel B) computed on the time-varying four-factor model for the 25 FF portfolios.

Figure 35: $\hat{\rho}_i^2$ vs $IdiVol_i$ and $SysRisk_i$ for the 44 Indu. portfolios



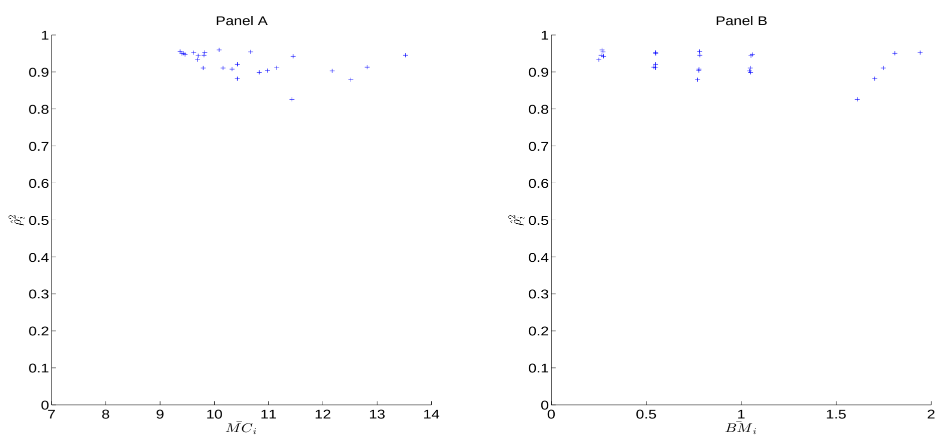
The figure plots the coefficients of determination $\hat{\rho}_i^2$ w.r.t. the idiosyncratic risks $IdiVol_i$ (Panel A) and the systematic risks $SysRisk_i$ (Panel B) computed on the time-varying four-factor model for the 44 Indu. portfolios.

Figure 36: $\bar{MC}_i$ and $\bar{BM}_i$ vs $\hat{\rho}_i^2$ for the $n^\chi = 3,900$ individual stocks



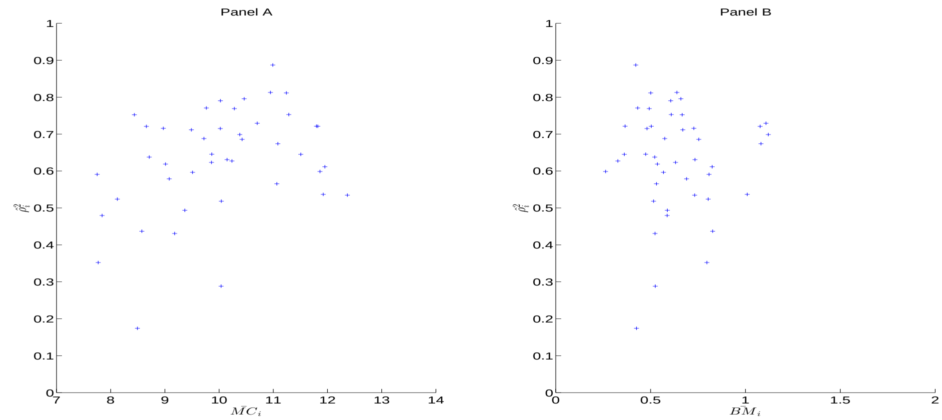
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t. the estimated coefficients of determination $\hat{\rho}_i^2$ computed on the time-varying four-factor model for the $n^\chi = 3,900$ individual stocks. We report the estimated linear quantile regressions for probability levels 90%, 75%, 50%, 25%, and 10%.

Figure 37: $\bar{MC}_i$ and $\bar{BM}_i$ vs $\hat{\rho}_i^2$ for the 25 FF portfolios



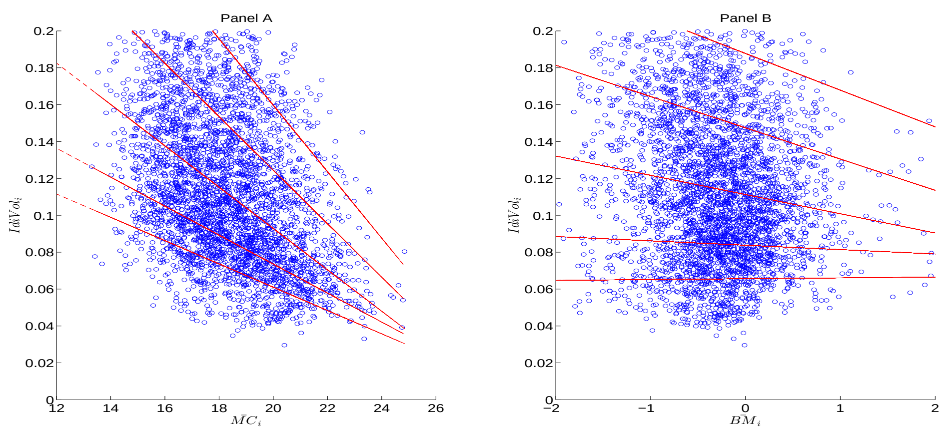
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t. the estimated coefficients of determination $\hat{\rho}_i^2$ computed on the time-varying four-factor model for the 25 FF portfolios.

Figure 38: $\bar{MC}_i$ **and** $\bar{BM}_i$ **vs** $\hat{\rho}_i^2$ **for the 44 Indu. portfolios**



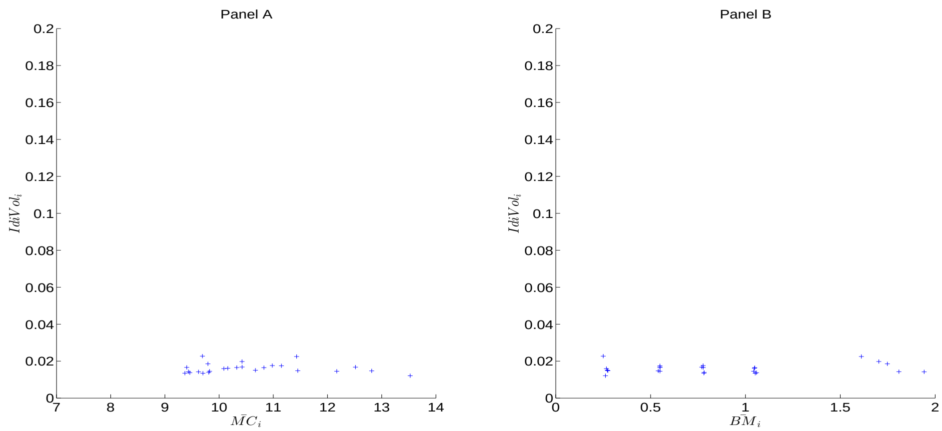
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t. the estimated coefficients of determination $\hat{\rho}_i^2$ computed on the time-varying four-factor model for the 44 Indu. portfolios.

Figure 39: $\bar{MC}_i$ **and** $\bar{BM}_i$ **vs** $IdiVol_i$ **for the** $n^\chi = 3,900$ **individual stocks**



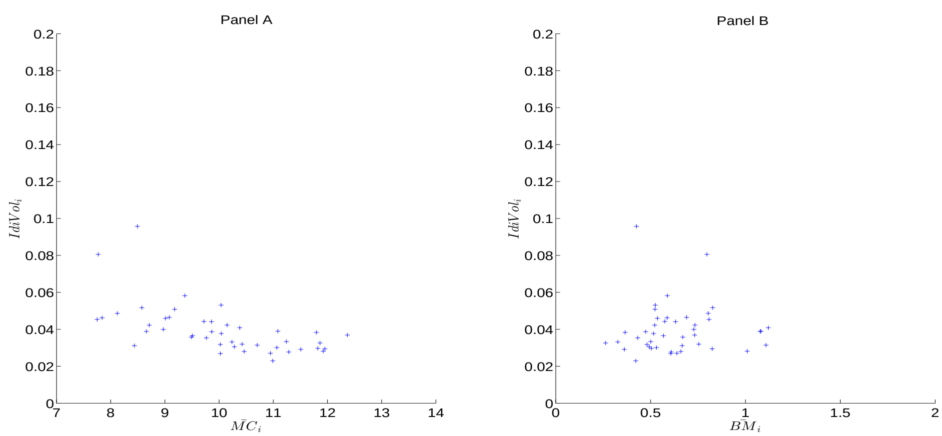
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t. the idiosyncratic risks $IdiVol_i$ computed on the time-varying four-factor model for the $n^\chi = 3,900$ individual stocks. We report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%.

Figure 40: $\bar{MC}_i$ and $\bar{BM}_i$ vs $IdiVol_i$ for the 25 FF portfolios



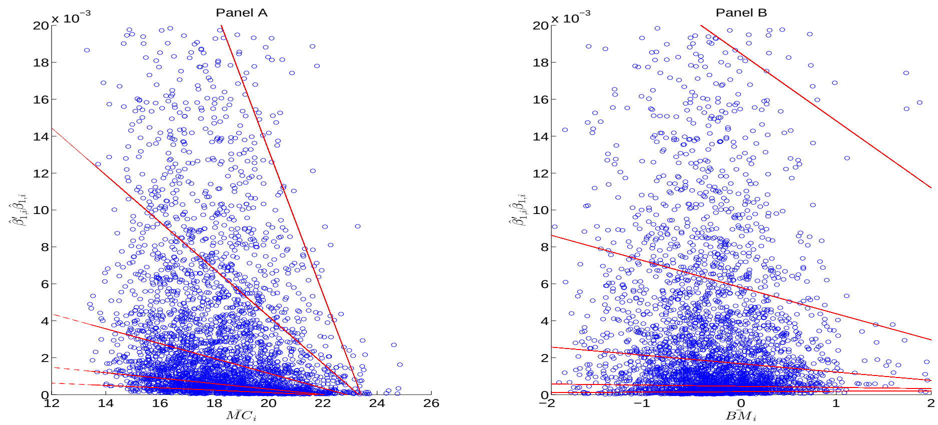
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t the idiosyncratic risks $IdiVol_i$ computed on the time-varying four-factor model for the 25 FF portfolios.

Figure 41: $\bar{MC}_i$ and $\bar{BM}_i$ vs $IdiVol_i$ for the 44 Indu. portfolios



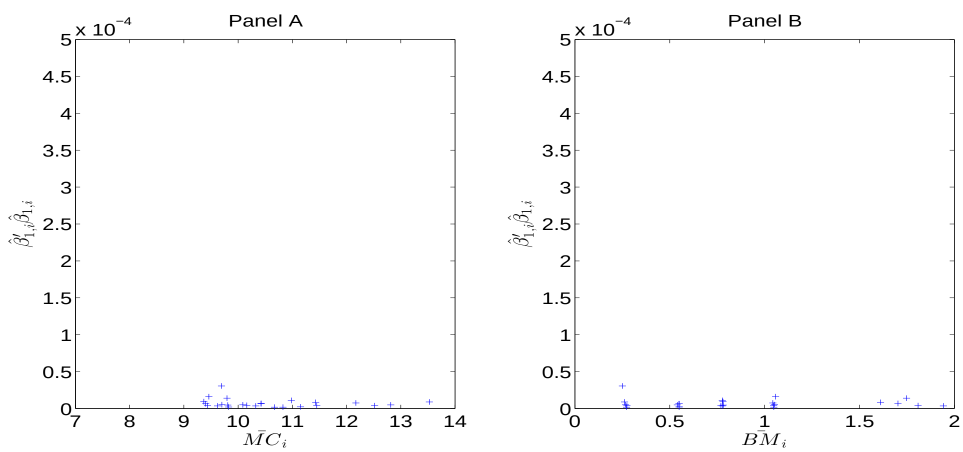
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t. the idiosyncratic risks $IdiVol_i$ computed on the time-varying four-factor model for the 44 Indu. portfolios.

Figure 42: $\bar{MC}_i$ and $\bar{BM}_i$ vs $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ for the $n^\chi = 3,900$ individual stocks



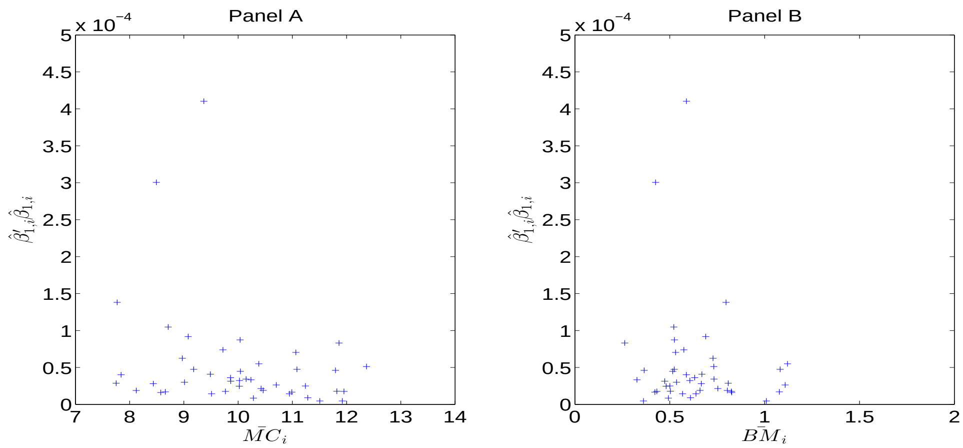
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t. the $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$. Estimated $\hat{\beta}_{1,i}$ are for the time-varying four-factor model using the $n^\chi = 3,900$ individual stocks as base assets. We report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%.

Figure 43: $\bar{MC}_i$ and $\bar{BM}_i$ vs $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ for the 25 FF portfolios



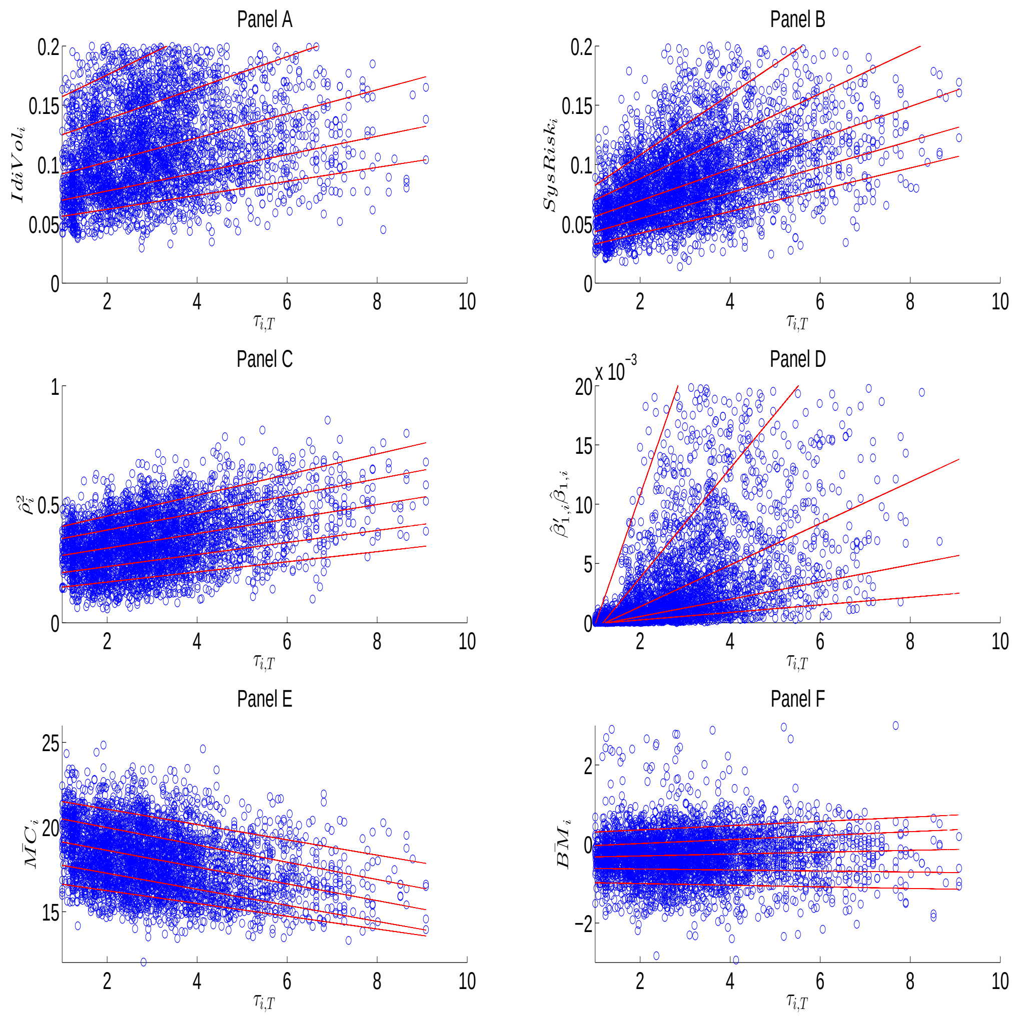
The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $\bar{BM}_i$ (Panel B) w.r.t the $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$. Estimated $\hat{\beta}_{1,i}$ are for the time-varying four-factor model using the 25 FF portfolios as base assets.

Figure 44: $\bar{MC}_i$ and $B\bar{M}_i$ vs $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ for the 44 Indu. portfolios



The figure plots the averages over time of the market capitalisation $\bar{MC}_i$ (Panel A) and the book-to-market $B\bar{M}_i$ (Panel B) w.r.t. $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$. Estimated $\hat{\beta}_{1,i}$ are for the time-varying four-factor model using the 44 Indu. portfolios as base assets.

Figure 45: $\tau_{i,T}$ **vs** $\hat{\rho}_i^2,\ IdiVol_i$**,** $\bar{MC}_i$ **and** $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ **for the** $n^{\chi} = 3,900$ **individual stocks**



The figure plots the inverse standardized sample size $\tau_{i,T}$ w.r.t. the idiosyncratic risks $IdiVol_i$ (panel A), the systematic risks $SysRisk_i$ (panel B), the estimated coefficients of determination $\hat{\rho}_i^2$ (panel C), the $\hat{\beta}'_{1,i}\hat{\beta}_{1,i}$ (panel D), the averages over time of the market capitalisation $\bar{MC}_i$ (Panel E), and the averages over time of book-to-market $B\bar{M}_i$ (Panel F), computed on the time-varying four-factor model for the $n^{\chi} = 3,900$ individual stocks. We report the estimated linear quantile regressions for probability levels $90\%,\ 75\%,\ 50\%,\ 25\%$, and 10%.

Appendix 16 Robustness checks on the beta specification for individual stocks

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 23, 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 capitalisation, $bm_{i,t}$ the book-to-market, and $ind_{i,t}$ the return of the corresponding industry portfolio. For each exercise, when not explicitly indicated in Table 23, 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 23 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 46 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. Figure 47 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. Figures 48 and 49 compares the paths of estimated annualized $\hat{\nu}_t$. The paths look similar through the asset pricing models. 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. Overall, the conditional estimates of the risk premia and coefficients vector ν are stable with respect to the asset pricing model that is assumed for the excess returns.

Figures 50 and 51 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 50, we get the estimates by setting the scalar $Z_{i,t-1}$ equal to the market capitalisation of firm i. In Figure 51, 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 24, upper panel. The test statistics reject the null hypotheses at 5% level.

The time-varying paths of the risk premia showed in Figures 52 and 53 are computed by modifying the set of common instruments $Z_{t-1} = \left[1,\ Z_{t-1}^{*\prime}\right]'$ (Exercise c). In Figure 52, 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 53, 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 as an effect of the large number of parameters to estimate in the first pass. The confidence bands in Figure 53 are wider than in Figure 1. The test statistics reject the null hypothesis at 5% level (see

Table 24), 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 modeled via common instruments $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$ as usual. In Figure 54, $Z_{i,t-1}$ is a bivariate vector that includes the constant and the book-to-market equity of firm i. In Figure 55, vector $Z_{i,t-1}$ includes the constant and the return of the industry portfolio as asset-specific instrument. When $b_{i,t}$ does not depend on Z_{t-1}, the vector $Z_{i,t-1}$ contains the element 1 to include the constant in the beta specification. The paths of the risk premia for the four factors in Figure 54 look more volatile w.r.t. the paths in Figure 1. The risk premia for market, size and value factors in Figure 55 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 24, lower panel, the test statistic does not reject the asset pricing restriction $\mathcal{H}_0 :\ \beta_1(\gamma) = \beta_3(\gamma)\,\nu$ for the conditional specification with time-varying betas depending on book-to-market equity.

Table 23: Operative cross-sectional sample size (n^{χ}), number of factors (K) and instruments (p and q) and first-pass regressors (d) in the four exercises of robustness checks

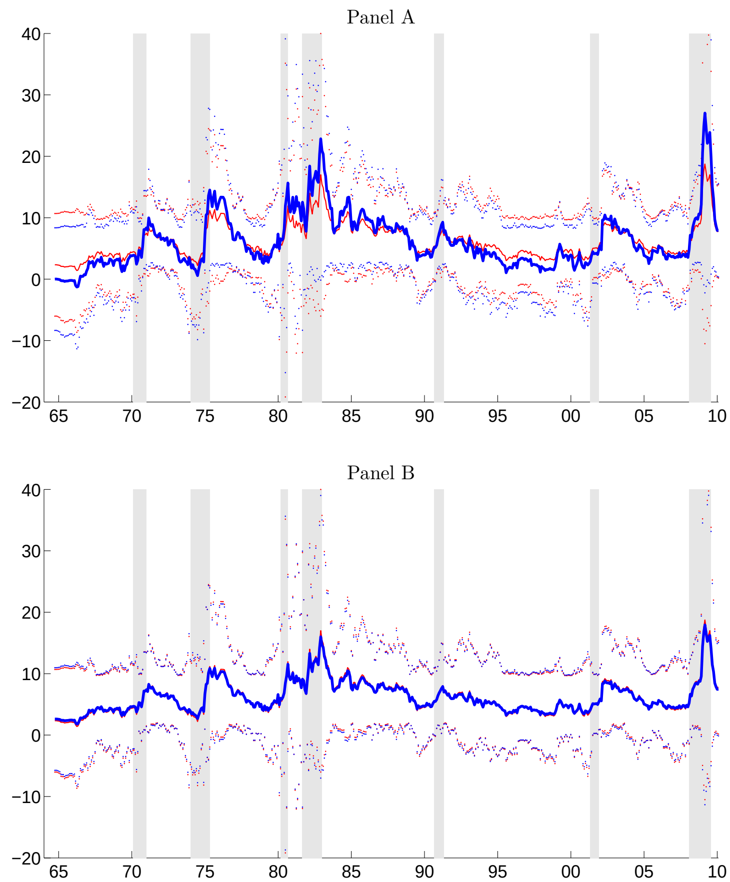
	n^{χ}	K	p	q	d		n^{χ}	K	p	q	d
Exercise a.						Exercise c.					
CAPM	5,225	1	3	1	13	$Z_{t-1} = [1, ds_{t-1}, divY_{t-1}]'$	1,107	4	3	1	25
Fama-French model	4,545	3	3	1	21	$Z_{t-1} = [1, ds_{t-1}, ts_{t-1}, divY_{t-1}]'$	667	4	4	1	34
Exercise b.						Exercise d.					
$Z_{i,t-1} = mc_{i,t-1}$	3,835	4	3	1	25	$Z_{i,t-1} = [1, bm_{i,t-1}]'$	6,135	4	3	2	14
$Z_{i,t-1} = ind_{i,t-1}$	4,816	4	3	1	25	$Z_{i,t-1} = [1, ind_{i,t-1}]'$	6,515	4	3	2	14

Table 24: 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.				
	$Z_{i,t-1} = mc_{i,t-1}$	$Z_{i,t-1} = ind_{i,t-1}$	$Z_{i,t-1} = mc_{i,t-1}$	$Z_{i,t-1} = ind_{i,t-1}$
	$(n^{\chi} = 1,383)$	$(n^{\chi} = 1,396)$	$(n^{\chi} = 1,383)$	$(n^{\chi} = 1,396)$
Test statistic	8.0493	5.7373	8.7126	6.4544
p-value	0.0000	0.0000	0.0000	0.0000
Exercise c.				
	$Z_{t-1} = [1, ds_{t-1}, divY_{t-1}]'$	$Z_{t-1} = [1, ds_{t-1}, ts_{t-1}, divY_{t-1}]'$	$Z_{t-1} = [1, ds_{t-1}, divY_{t-1}]'$	$Z_{t-1} = [1, ds_{t-1}, ts_{t-1}, divY_{t-1}]'$
	$(n^{\chi} = 818)$	$(n^{\chi} = 638)$	$(n^{\chi} = 818)$	$(n^{\chi} = 638)$
Test statistic	2.7954	2.2463	3.6335	2.8434
p-value	0.0026	0.0123	0.0000	0.0022
Exercise d.				
	$Z_{i,t-1} = [1, bm_{i,t-1}]'$	$Z_{i,t-1} = [1, ind_{i,t-1}]'$	$Z_{i,t-1} = [1, bm_{i,t-1}]'$	$Z_{i,t-1} = [1, ind_{i,t-1}]'$
	$(n^{\chi} = 1,395)$	$(n^{\chi} = 1,396)$	$(n^{\chi} = 1,395)$	$(n^{\chi} = 1,396)$
Test statistic	1.0950	3.1826	2.8564	5.0002
p-value	0.1368	0.0007	0.0021	0.0000

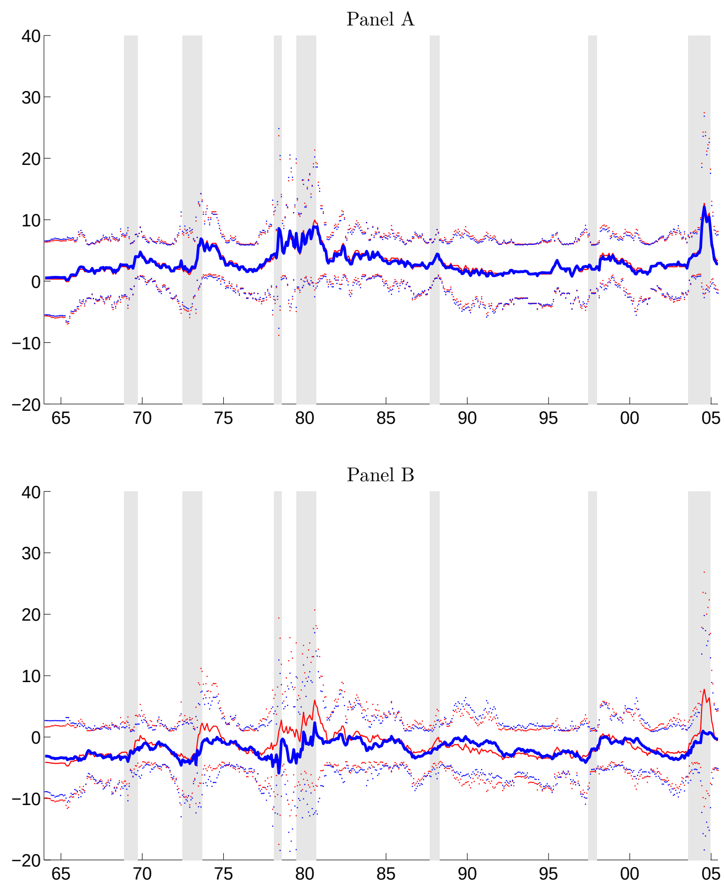
We compute the statistics $\tilde{\Sigma}_{\xi}^{-1/2}\hat{\xi}_{nT}$ based on $\hat{Q}_e$ and $\hat{Q}_a$ defined in Proposition 6 for n^{χ} individual stocks to test the null hypotheses $\mathcal{H}_0 :\ \beta_1(\gamma) = \beta_3(\gamma)\nu$ and $\mathcal{H}_0 :\ \beta_1(\gamma) = 0$. The table reports 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 46: Path of estimated annualized risk premia for the market factor



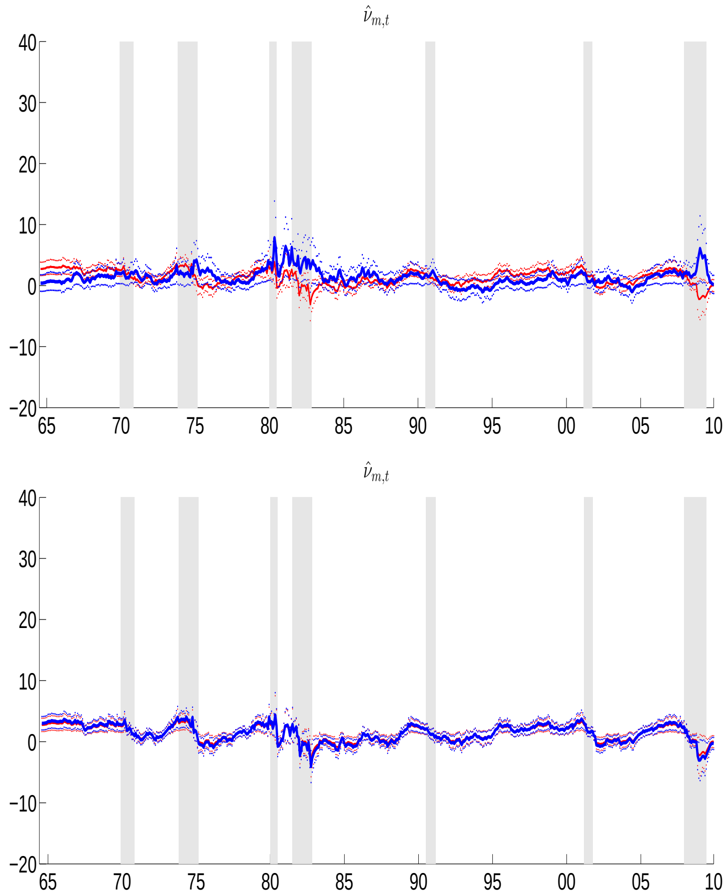
Panel A plots the paths of estimated annualized market risk premia $\hat{\lambda}_{m,t}$ computed by using the four-factor model (thin red line) and the CAPM (thick blue line). Panel B plots the paths of market risk premia $\hat{\lambda}_{m,t}$ estimated by assuming the four-factor model (thin red line) 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 Research (NBER).

Figure 47: Path of estimated annualized risk premia for the size and value factors



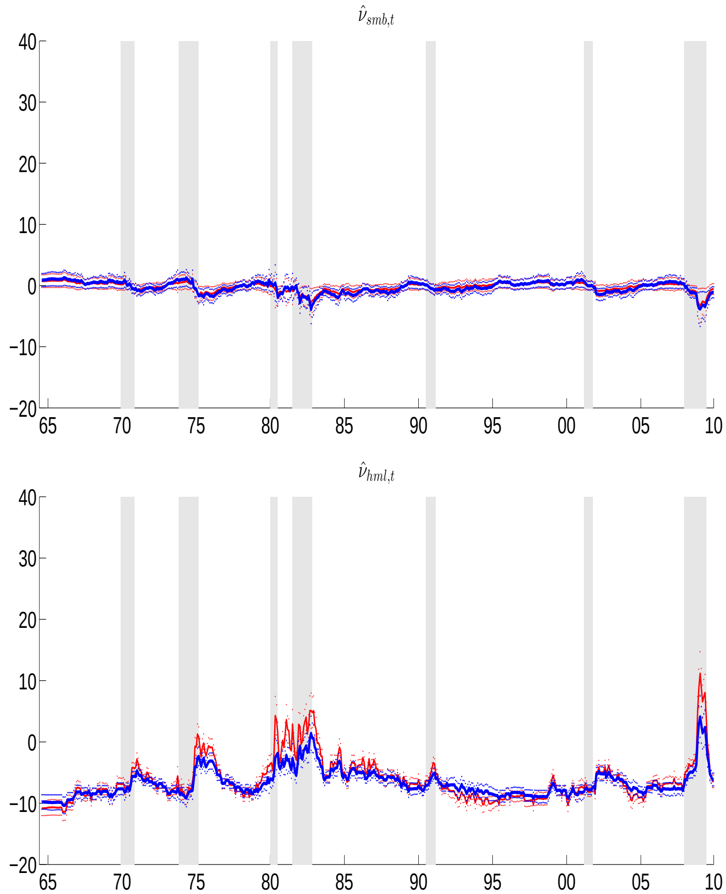
The figure plots the paths of estimated annualized risk premia $\hat{\lambda}_{smb,t}$ (Panel A) and $\hat{\lambda}_{hml,t}$ (Panel B) computed by using the four-factor model (thin red line) and the Fama-French model (thick blue line). The pointwise confidence intervals at 95% level are also displayed.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

Figure 48: Path of estimated annualized ν_t for the market factor



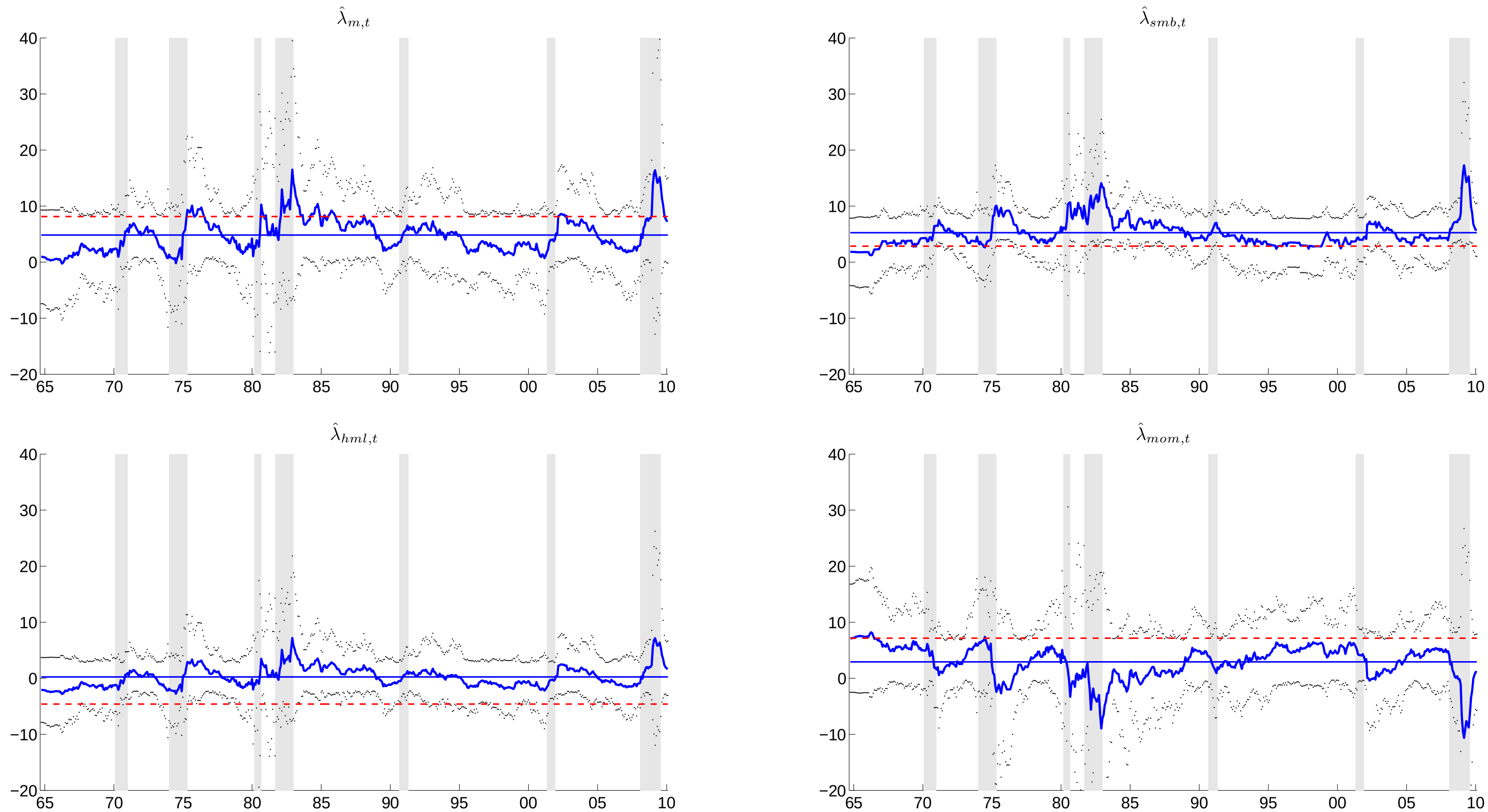
Panel A plots the paths of estimated annualized $\hat{\nu}_{m,t}$ computed by using the four-factor model (thin red line) and the CAPM (thick blue line). Panel B plots the paths of $\hat{\nu}_{m,t}$ estimated by assuming the four-factor model (thin red line) and the Fama-French model (thick blue line).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

Figure 49: Path of estimated annualized ν_t for the size and value factors



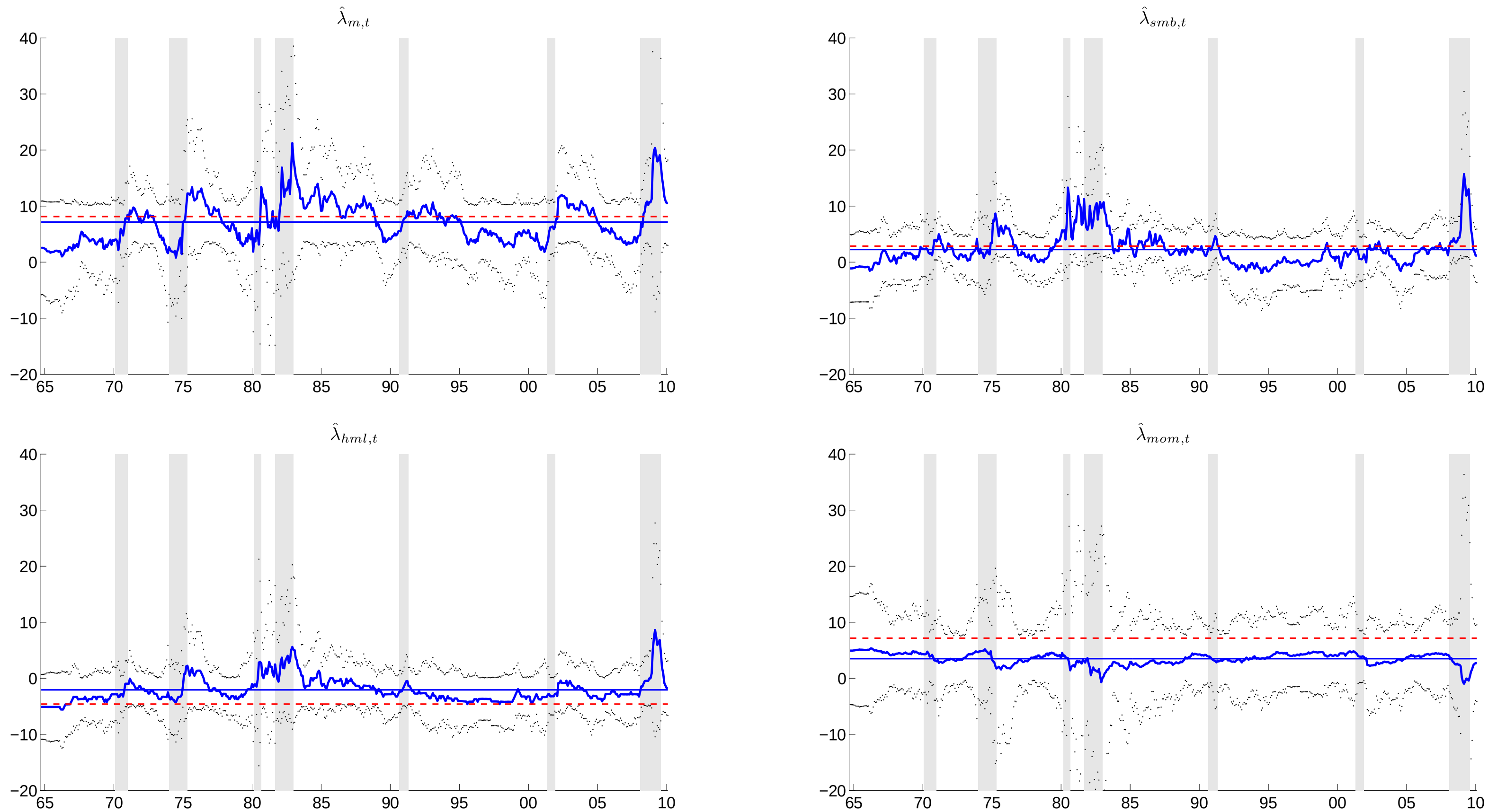
The figure plots the paths of estimated annualized $\hat{\nu}_{smb,t}$ (Panel A) and $\hat{\nu}_{hml,t}$ (Panel B) computed by using the four-factor model (thin red line) and the Fama-French model (thick blue line). The pointwise confidence intervals at 95% level are also displayed. The vertical shaded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

Figure 50: Path of estimated annualized risk premia computed using $Z_{i,t-1} = mc_{i,t-1}$



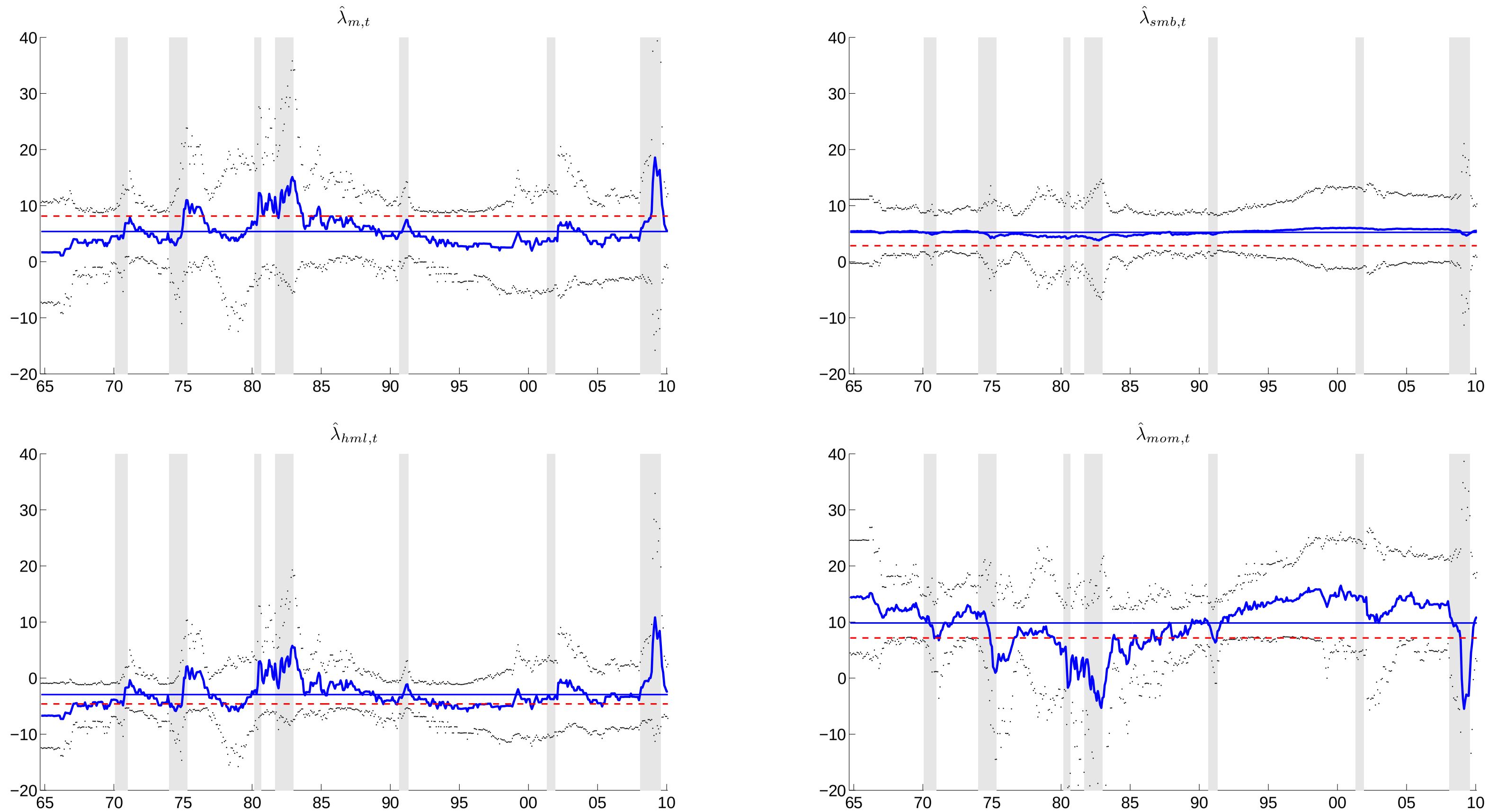
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when market capitalisation is used as asset-specific instrument. The vector of common instruments is $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$. We also display the time-invariant 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 Research (NBER).

Figure 51: Path of estimated annualized risk premia computed using $Z_{i,t-1} = ind_{i,t-1}$



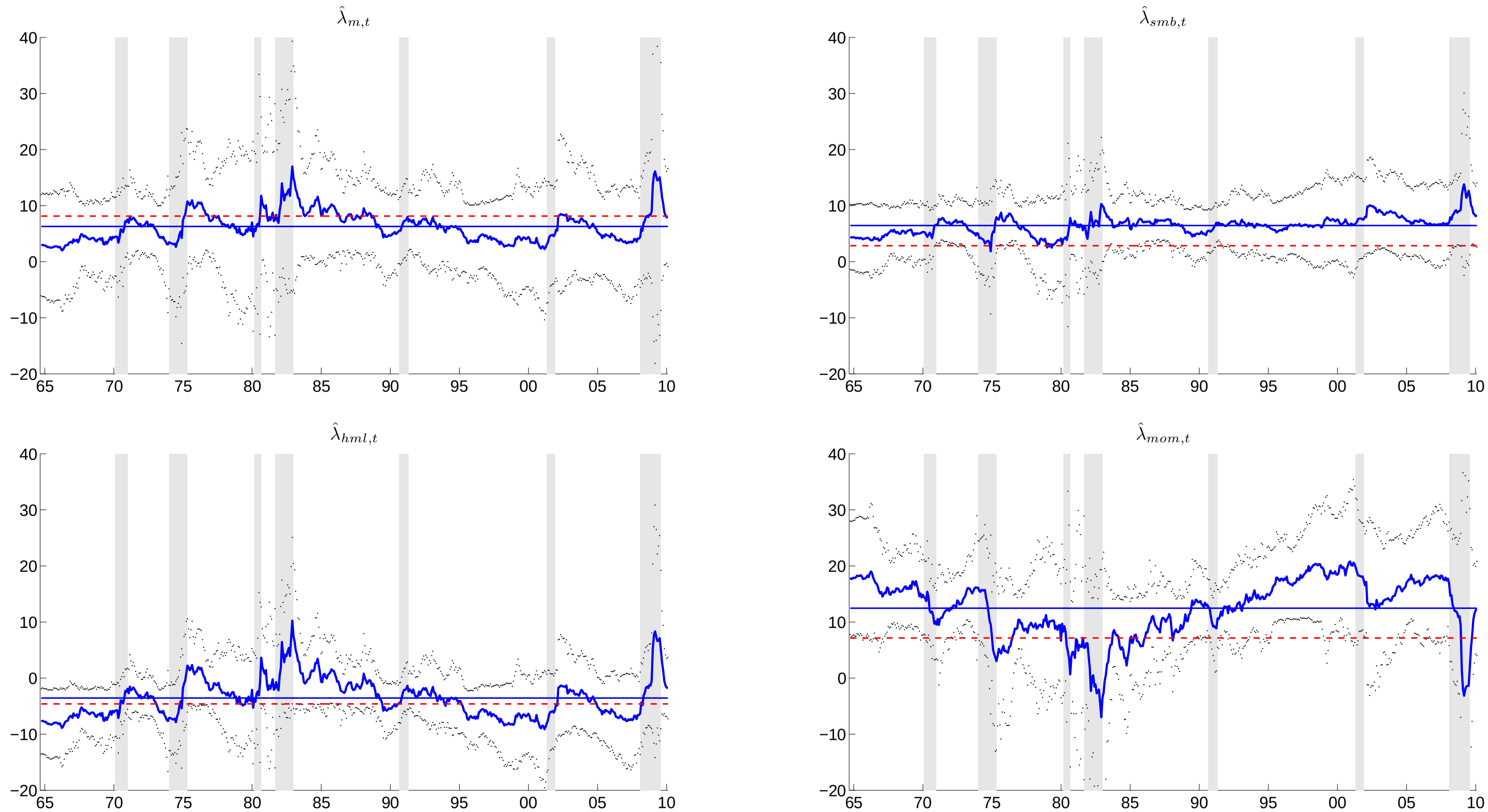
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when the returns of industry portfolios are used as asset-specific instrument. The vector of common instruments is $Z_{t-1} = [1,\ ts_{t-1},\ ds_{t-1}]'$. We also display the time-invariant estimate (dashed horizontal line) and the average conditional estimate (solid horizontal line). We consider all stocks as base assets ($n = 9,936$ and $n^{\chi} = 4,861$). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).

Figure 52: Path of estimated annualized risk premia computed using $Z_{t-1} = [1,\ ds_{t-1},\ divY_{t-1}]'$



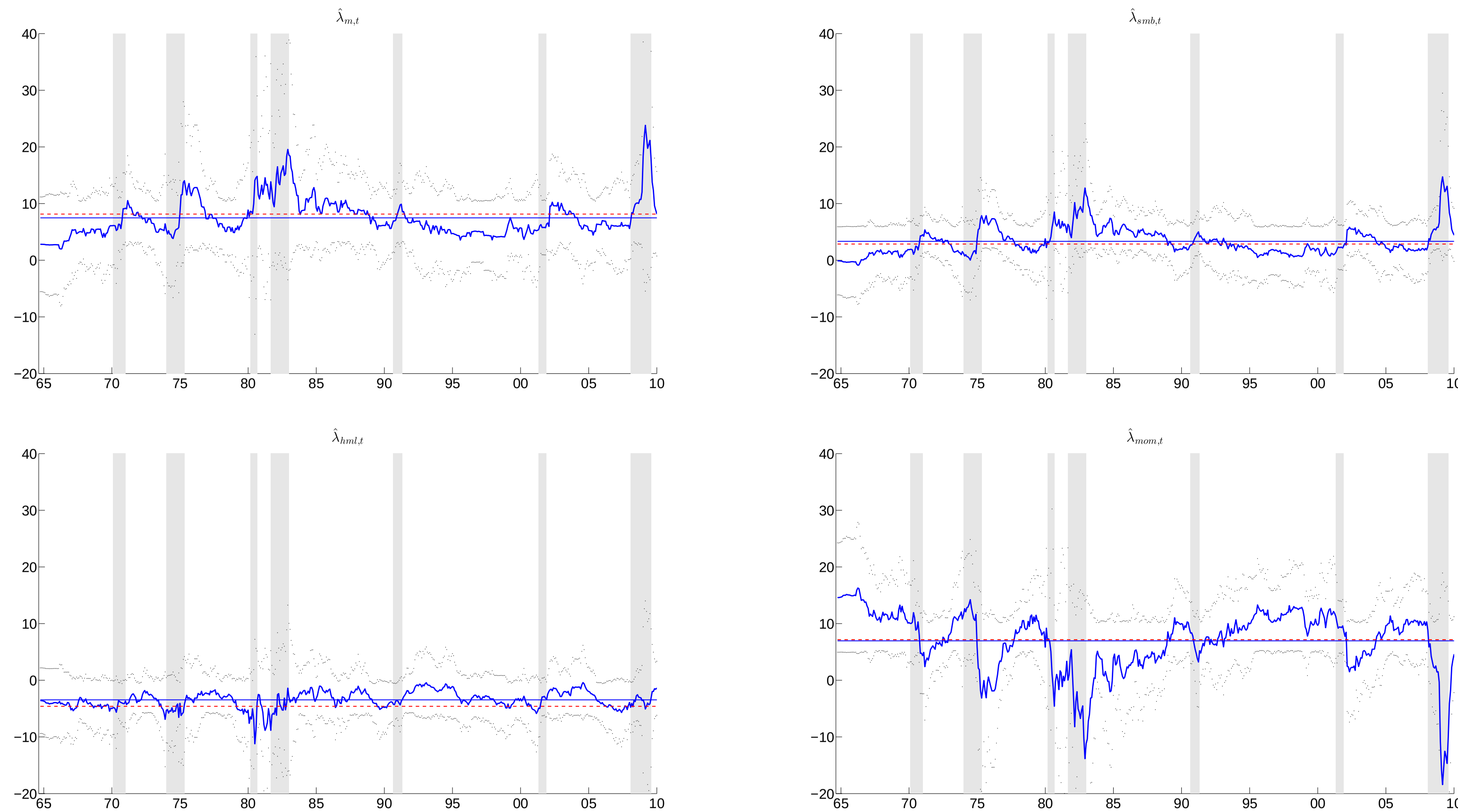
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when default spread and dividend yield are used as common instruments. The stock specific instrument is book-to-market equity. We also display the time-invariant 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 Research (NBER).

Figure 53: Path of estimated annualized risk premia computed using $Z_{t-1} = [1,\ ds_{t-1},\ ts_{t-1},\ divY_{t-1}]'$



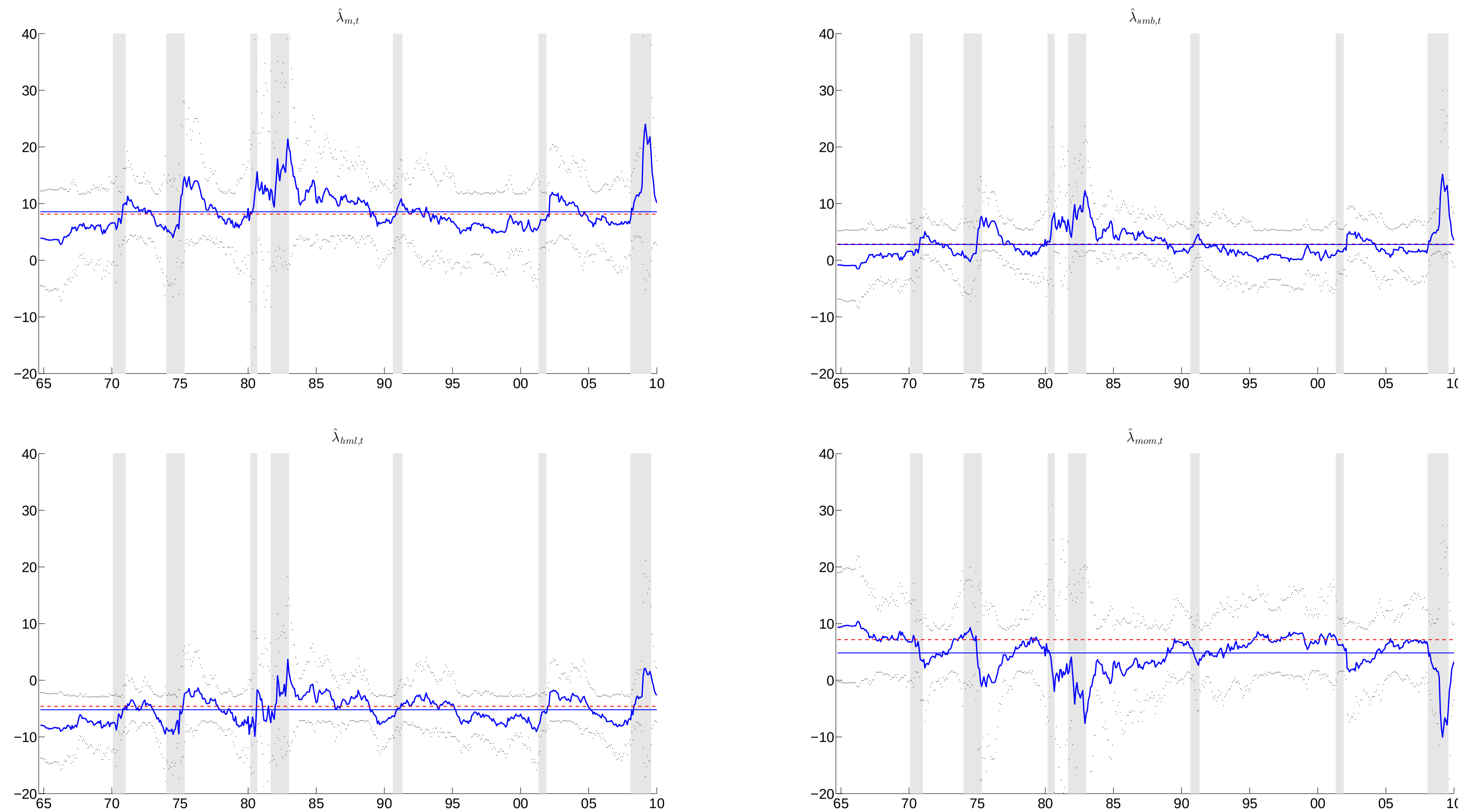
The figure plots the path of estimated annualized risk premia $\hat{\lambda}_{m,t}$, $\hat{\lambda}_{smb,t}$, $\hat{\lambda}_{hml,t}$ and $\hat{\lambda}_{mom,t}$ and their pointwise confidence intervals at 95% level when default spread, term spread and dividend yield are used as common instruments. The stock specific instrument is book-to-market equity. We also display the time-invariant 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 Research (NBER).

Figure 54: Path of estimated annualized risk premia with time-varying betas modeled 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 time-invariant 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 Research (NBER).

Figure 55: Path of estimated annualized risk premia with time-varying betas modeled 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 time-invariant 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 Research (NBER).

Appendix 17 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 A.17.1)

$$\begin{aligned}\sqrt{T}\left(\widehat{CE}_{i,t} - CE_{i,t}\right) &= \psi_{i,t}'E_2'\sqrt{T}\left(\hat{\beta}_i - \beta_i\right) \\ &\quad + \left(Z_{t-1}' \otimes b_{i,t}'\right) W_{p,K}\sqrt{T}vec\left[\hat{\Lambda}' - \Lambda'\right] + o_p(1), \end{aligned} \tag{82}$$

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 4, 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 56 plots the path of the estimated annualized costs of equity for Ford Motor, Disney, Motorola and Sony. We use the time-varying four-factor model estimated on individual stocks ($n = 9,936,\ n^\chi = 3,900$). The cost of equity has risen tremendously during the recent subprime crisis.

A.17.1 Proof of Equation (82)

We have:

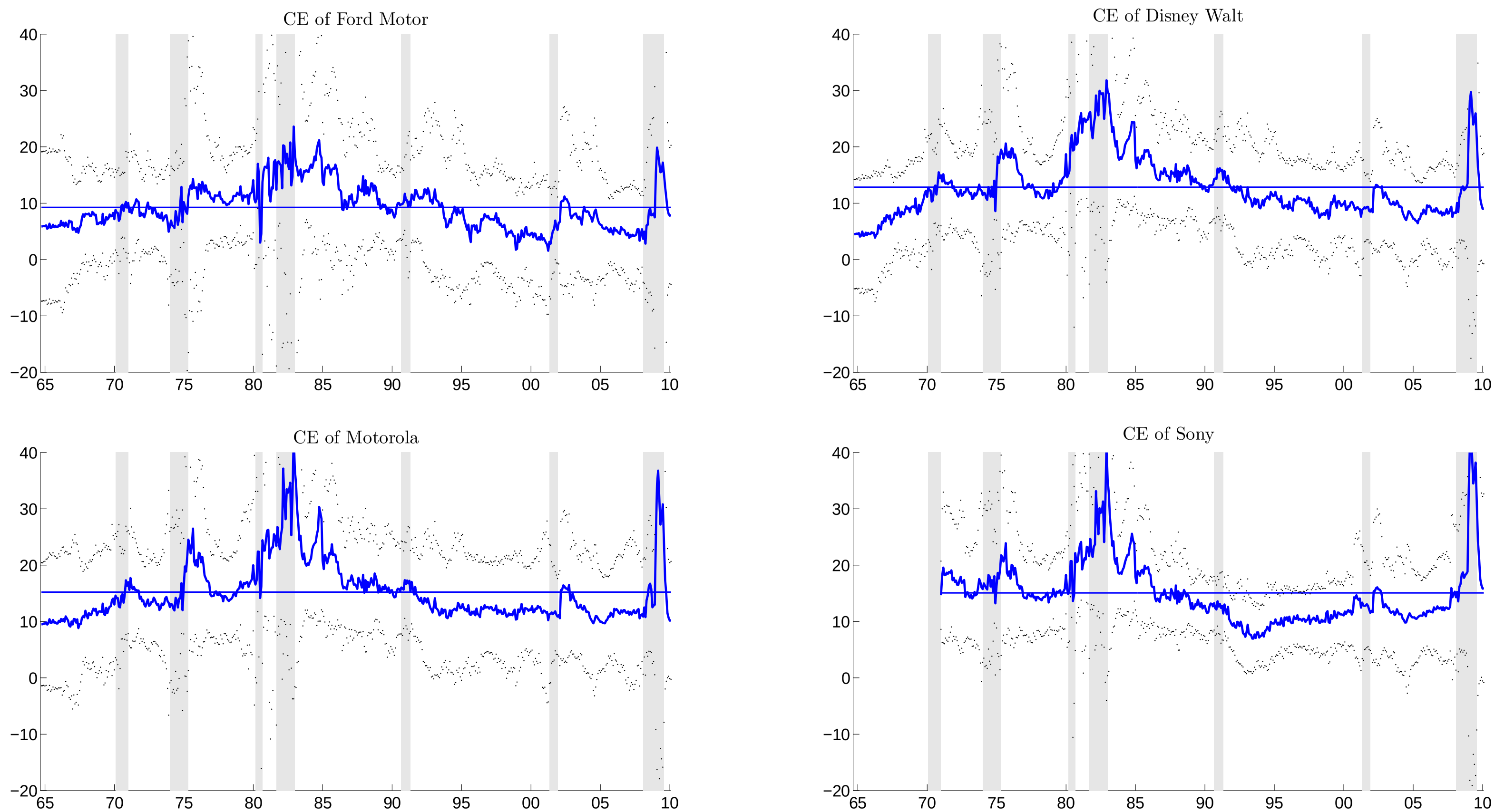
$$\hat{b}_{i,t}'\hat{\lambda}_t = tr\left[Z_{t-1}Z_{t-1}'\hat{B}_i'\hat{\Lambda}\right] + tr\left[Z_{t-1}Z_{i,t-1}'\hat{C}_i'\hat{\Lambda}\right] = \left(Z_{t-1}' \otimes Z_{t-1}'\right) vec\left[\hat{B}_i'\hat{\Lambda}\right] + \left(Z_{t-1}' \otimes Z_{i,t-1}'\right) vec\left[\hat{C}_i'\hat{\Lambda}\right].$$

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 (82) follows.

Figure 56: 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 use the time-varying four-factor model estimated on individual stocks ($n = 9,936,\ n^{\chi} = 3,900$) We also report the average conditional estimate (solid horizontal line). The vertical shaded areas denote recessions determined by the National Bureau of Economic Research (NBER).