Online Appendix for

Testing Beta-Pricing Models Using Large Cross-Sections

Valentina Raponi, Cesare Robotti, and Paolo Zaffaroni

This Online Appendix is structured as follows. Section IA.1 contains the proofs of the propositions and theorems in the paper. In Section IA.2, we analyze how the assumption of random betas affects our results. In Section IA.3, we study the properties of alternative risk premia estimators in the traded-factor case. In Section IA.4, we present Monte Carlo simulation results. Section IA.5 is for the unbalanced panel case. Finally, Section IA.6 contains a set of figures that complements those in the paper. We refer to the paper for the notation used here. When we do not use the IA prefix, it means that we are referring to the main paper.

IA.1. Proofs

We start with a few preliminary lemmas and then use them to prove the propositions and theorems in the paper.

IA.1.1 Preliminary Lemmas

Lemma 1 Under assumptions 3–5,

$$\hat{\sigma}^2 - \sigma^2 = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{IA.1}$$

Proof. Rewrite $\hat{\sigma}^2 - \sigma^2$ as

$$\begin{aligned}\hat{\sigma}^2 - \sigma^2 &= \left(\hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 - \sigma^2\right) \\ &= \left(\hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2\right) + o\left(\frac{1}{\sqrt{N}}\right)\end{aligned} \tag{IA.2}$$

by assumption 5(i). Moreover,

$$\begin{aligned}\hat{\sigma}^2 - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 &= \frac{\operatorname{tr}(M\epsilon\epsilon')}{N(T-K-1)} - \frac{\operatorname{tr}(M)}{T-K-1}\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \\ &= \frac{\operatorname{tr}\left(P\left(\sum_{i=1}^{N}\sigma_i^2 I_T - \epsilon\epsilon'\right)\right)}{N(T-K-1)} + \frac{\operatorname{tr}(\epsilon\epsilon') - T\sum_{i=1}^{N}\sigma_i^2}{N(T-K-1)}.\end{aligned} \tag{IA.3}$$

As for the second term on the right-hand side of Equation (IA.3), we have

$$\begin{aligned}\frac{\operatorname{tr}(\epsilon\epsilon') - T\sum_{i=1}^{N}\sigma_i^2}{N(T-K-1)} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\epsilon_{it}^2 - \sigma_i^2\right)}{N(T-K-1)} \\ &= O_p\left(\frac{1}{\sqrt{N}}\frac{\sqrt{T}}{(T-K-1)}\right) = O_p\left(\frac{1}{\sqrt{N}}\right).\end{aligned} \tag{IA.4}$$

As for the first term on the right-hand side of Equation (IA.3), we have

$$\begin{aligned}
\frac{\operatorname{tr}\left(P\left(\sum_{i=1}^{N}\sigma_i^2 I_T - \epsilon\epsilon'\right)\right)}{N\left(T-K-1\right)} &= \frac{\sum_{t=1}^{T} d_t\left(D'D\right)^{-1}D'\left(\sum_{i=1}^{N}\sigma_i^2\imath_{t,T} - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)} \\
&= \frac{\sum_{t=1}^{T} p_t\left(\sum_{i=1}^{N}\sigma_i^2\imath_{t,T} - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)}, \qquad \text{(IA.5)}
\end{aligned}$$

where $\imath_{t,T}$ is a T-vector with one in the t-th position and zeros elsewhere, d_t is the t-th row of $D = [1_T,\ F]$, and $p_t = d_t\left(D'D\right)^{-1}D'$. Since Equation (IA.5) has a zero mean, we only need to consider its variance to determine the rate of convergence. We have

$$\begin{aligned}
&\operatorname{Var}\left(\frac{\sum_{t=1}^{T} p_t\left(\sum_{i=1}^{N}\sigma_i^2\imath_{t,T} - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)}\right) \\
=\ & \frac{1}{N^2(T-K-1)^2}E\left[\sum_{i,j=1}^{N}\sum_{t,s=1}^{T} p_t\left(\sigma_i^2\imath_{t,T} - \epsilon_i\epsilon_{it}\right)\left(\sigma_j^2\imath_{s,T} - \epsilon_j\epsilon_{js}\right)' p_s'\right] \\
=\ & \frac{1}{N^2(T-K-1)^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T} p_t E\left[\left(\sigma_i^2\imath_{t,T} - \epsilon_i\epsilon_{it}\right)\left(\sigma_j^2\imath_{s,T} - \epsilon_j\epsilon_{js}\right)'\right] p_s'. \qquad \text{(IA.6)}
\end{aligned}$$

Moreover, we have

$$E\left[\left(\sigma_i^2\imath_{t,T} - \epsilon_i\epsilon_{it}\right)\left(\sigma_j^2\imath_{s,T} - \epsilon_j\epsilon_{js}\right)'\right] = E\left[\sigma_i^2\sigma_j^2\imath_{t,T}\imath_{s,T}' + \epsilon_i\epsilon_j'\epsilon_{it}\epsilon_{js} - \sigma_i^2\imath_{t,T}\epsilon_j'\epsilon_{js} - \sigma_j^2\epsilon_{it}\epsilon_i\imath_{s,T}'\right]$$

$$= \begin{cases}
\mu_{4i}\imath_{t,T}\imath_{t,T}' + \sigma_i^4\left(I_T - 2\imath_{t,T}\imath_{t,T}'\right) & \text{if} \quad i = j,\ t = s \\
\kappa_{4,iijj}\imath_{t,T}\imath_{t,T}' + \sigma_{ij}^2(I_T + \imath_{t,T}\imath_{t,T}') & \text{if} \quad i \neq j,\ t = s \\
\sigma_i^4\imath_{s,T}\imath_{t,T}' & \text{if} \quad i = j,\ t \neq s \\
\sigma_{ij}^2\imath_{s,T}\imath_{t,T}' & \text{if} \quad i \neq j,\ t \neq s.
\end{cases} \qquad \text{(IA.7)}$$

It follows that

$$
\begin{aligned}
&\mathrm{Var}\left(\frac{\sum_{t=1}^{T} p_t\left(\sum_{i=1}^{N}\sigma_i^2 \imath_{t,T} - \sum_{i=1}^{N}\epsilon_i\epsilon_{it}\right)}{N(T-K-1)}\right) \\
= \;& \frac{1}{N^2(T-K-1)^2}\sum_{t=1}^{T}\sum_{i=1}^{N} p_t\left(\mu_{4i}\imath_{t,T}\imath_{t,T}' + \sigma_i^4\left(I_T - 2\imath_{t,T}\imath_{t,T}\right)\right)p_t' \\
& + \frac{1}{N^2(T-K-1)^2}\sum_{t=1}^{T}\sum_{i\neq j} p_t\left(\kappa_{4,iijj}\imath_{t,T}\imath_{t,T}' + \sigma_{ij}^2\left(I_T + \imath_{t,T}\imath_{t,T}'\right)\right)p_t' \\
& + \frac{1}{N^2(T-K-1)^2}\sum_{i=1}^{N}\sigma_i^4\sum_{t\neq s} p_t\imath_{s,T}\imath_{t,T}'p_s' \\
& + \frac{1}{N^2(T-K-1)^2}\sum_{i\neq j}\sigma_{ij}^2\sum_{t\neq s} p_t\imath_{s,T}\imath_{t,T}'p_s' \\
= \;& O\left(\frac{1}{N}\right)
\end{aligned}
\tag{IA.8}
$$

by assumptions 5(ii), 5(iii), 5(iv), and 5(viii), which implies that the first term on the right-hand side of Equation (IA.3) is $O_p\left(\frac{1}{\sqrt{N}}\right)$. Putting the pieces together concludes the proof. ■

Lemma 2 Let

$$\Lambda = \begin{bmatrix} 0 & 0_K' \\ 0_K & \sigma^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}. \tag{IA.9}$$

(i) Under assumptions 2–5,

$$\hat{X}'\hat{X} = O_p(N). \tag{IA.10}$$

In addition, under assumption 6,

(ii)

$$\hat{\Sigma}_X \xrightarrow{p} \Sigma_X + \Lambda, \tag{IA.11}$$

and

(iii)

$$\frac{(\hat{X}-X)'(\hat{X}-X)}{N} \xrightarrow{p} \Lambda. \tag{IA.12}$$

Proof.

(i) Consider

$$\hat{X}'\hat{X} = \begin{bmatrix} N & 1_N'\hat{B} \\ \hat{B}'1_N & \hat{B}'\hat{B} \end{bmatrix}. \tag{IA.13}$$

Then,

$$\hat{B}'1_N = \sum_{i=1}^{N} \hat{\beta}_i = \sum_{i=1}^{N} \beta_i + \mathcal{P}' \sum_{i=1}^{N} \epsilon_i. \tag{IA.14}$$

Under assumptions 4–5,

$$\begin{aligned} \operatorname{Var}\left(\sum_{t=1}^{T}\sum_{i=1}^{N}\epsilon_{it}(f_t-\bar{f})\right) &= \sum_{t,s=1}^{T}\sum_{i,j=1}^{N}(f_t-\bar{f})(f_s-\bar{f})'E[\epsilon_{it}\epsilon_{js}] \\ &\leq \sum_{t=1}^{T}\sum_{i,j=1}^{N}(f_t-\bar{f})(f_t-\bar{f})'|\sigma_{ij}| \\ &= O\left(N\sigma^2\sum_{t=1}^{T}(f_t-\bar{f})(f_t-\bar{f})'\right) = O(NT). \end{aligned} \tag{IA.15}$$

Using assumption 2, we have

$$\hat{B}'1_N = O_p\left(N + \left(\tfrac{N}{T}\right)^{\frac{1}{2}}\right) = O_p(N). \tag{IA.16}$$

Next, consider

$$\begin{aligned} \hat{B}'\hat{B} &= \sum_{i=1}^{N}\hat{\beta}_i\hat{\beta}_i' \\ &= \sum_{i=1}^{N}\left(\beta_i + \mathcal{P}'\epsilon_i\right)\left({\beta_i}' + {\epsilon_i}'\mathcal{P}\right) \\ &= \sum_{i=1}^{N}\beta_i{\beta_i}' + \mathcal{P}'\left(\sum_{i=1}^{N}\epsilon_i{\epsilon_i}'\right)\mathcal{P} \\ &\quad + \mathcal{P}'\left(\sum_{i=1}^{N}\epsilon_i{\beta_i}'\right) + \left(\sum_{i=1}^{N}\beta_i{\epsilon_i}'\right)\mathcal{P}. \end{aligned} \tag{IA.17}$$

By assumption 2,

$$\sum_{i=1}^{N}\beta_i{\beta_i}' = O(N). \tag{IA.18}$$

Using similar arguments as for Equation (IA.15),

$$\mathcal{P}'\left(\sum_{i=1}^{N}\epsilon_i{\beta_i}'\right) = O_p\left(\left(\frac{N}{T}\right)^{\frac{1}{2}}\right) \tag{IA.19}$$

and

$$\left(\sum_{i=1}^{N}\beta_i{\epsilon_i}'\right)\mathcal{P} = O_p\left(\left(\frac{N}{T}\right)^{\frac{1}{2}}\right). \tag{IA.20}$$

For $\mathcal{P}'\left(\sum_{i=1}^{N}\epsilon_i\epsilon_i{}'\right)\mathcal{P}$, consider its central part and take the norm of its expectation. Using assumptions 4–5,

$$\begin{aligned}
& \left\|E\left[\tilde{F}'\left(\sum_{i=1}^{N}\epsilon_i\epsilon_i{}'\right)\tilde{F}\right]\right\| \\
= & \left\|E\left[\sum_{t,s=1}^{T}\sum_{i=1}^{N}(f_t-\bar{f})(f_s-\bar{f})'\epsilon_{it}\epsilon_{is}\right]\right\| \\
\leq & \sum_{t,s=1}^{T}\sum_{i=1}^{N}\|(f_t-\bar{f})(f_s-\bar{f})'\||E\left[\epsilon_{it}\epsilon_{is}\right]| \\
= & \sum_{t=1}^{T}\sum_{i=1}^{N}\left\|\left(f_t-\bar{f}\right)\left(f_t-\bar{f}\right)'\right\|\sigma_i^2 \\
= & O\left(N\sigma^2\sum_{t=1}^{T}\|(f_t-\bar{f})(f_t-\bar{f})'\|\right)=O(NT).
\end{aligned} \tag{IA.21}$$

Then, we have

$$\mathcal{P}'\left(\sum_{i=1}^{N}\epsilon_i\epsilon_i{}'\right)\mathcal{P}=O_p\left(\frac{N}{T}\right) \tag{IA.22}$$

and

$$\hat{B}'\hat{B}=O_p\left(N+\left(\frac{N}{T}\right)^{\frac{1}{2}}+\frac{N}{T}\right)=O_p(N). \tag{IA.23}$$

This concludes the proof of part (i).

(ii) Using part (i) and under assumptions 3–6, we have

$$N^{-1}\hat{B}'1_N=\frac{1}{N}\sum_{i=1}^{N}\beta_i+O_p\left(\frac{1}{\sqrt{N}}\right) \tag{IA.24}$$

and

$$\begin{aligned}
N^{-1}\hat{B}'\hat{B} &= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i'+\mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\epsilon_i'\right)\mathcal{P}+\mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\beta_i'\right)+\left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\epsilon_i'\right)\mathcal{P} \\
&= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i'+\mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\epsilon_i'-\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2I_T+\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2I_T-\sigma^2I_T+\sigma^2I_T\right)\mathcal{P} \\
&\quad +\mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\beta_i'\right)+\left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\epsilon_i'\right)\mathcal{P} \\
&= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i'+\mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\left(\epsilon_i\epsilon_i'-\sigma_i^2I_T\right)\right)\mathcal{P}+\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i^2-\sigma^2\right)\mathcal{P}'\mathcal{P}+\sigma^2\mathcal{P}'\mathcal{P} \\
&\quad +\mathcal{P}'\left(\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\beta_i'\right)+\left(\frac{1}{N}\sum_{i=1}^{N}\beta_i\epsilon_i'\right)\mathcal{P} \\
&= \frac{1}{N}\sum_{i=1}^{N}\beta_i\beta_i'+\sigma^2\mathcal{P}'\mathcal{P}+O_p\left(\frac{1}{\sqrt{N}}\right)+o\left(\frac{1}{\sqrt{N}}\right)+O_p\left(\frac{1}{\sqrt{N}}\right)+O_p\left(\frac{1}{\sqrt{N}}\right).
\end{aligned} \tag{IA.25}$$

Assumption 2 concludes the proof of part (ii).

(iii) Note that

$$\frac{(\hat{X}-X)'(\hat{X}-X)}{N} = \frac{1}{N}\begin{bmatrix} 0_N' \\ (\hat{B}-B)' \end{bmatrix}[0_N,\ (\hat{B}-B)] = \begin{bmatrix} 0 & 0_K' \\ 0_K & \mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P} \end{bmatrix}. \tag{IA.26}$$

As in part (ii) we can write

$$\frac{\epsilon\epsilon'}{N} = \frac{1}{N}\sum_{i=1}^{N}\left(\epsilon_i\epsilon_i' - \sigma_i^2 I_T\right) + \left(\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i^2-\sigma^2\right)\right)I_T + \sigma^2 I_T. \tag{IA.27}$$

Assumptions 5(i) and 6(ii) conclude the proof since

$$\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P} = \sigma^2\mathcal{P}'\mathcal{P} + O_p\left(\frac{1}{\sqrt{N}}\right) + o\left(\frac{1}{\sqrt{N}}\right). \blacksquare \tag{IA.28}$$

Lemma 3 Under assumptions 2–5,

$$X'\bar{\epsilon} = O_p\left(\sqrt{N}\right). \tag{IA.29}$$

Proof. We have

$$X'\bar{\epsilon} = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} 1_N' \\ B' \end{bmatrix}\epsilon_t \tag{IA.30}$$

and

$$\begin{aligned}\mathrm{Var}\left(\frac{1}{T}\sum_{t=1}^{T}1_N'\epsilon_t\right) &= \frac{1}{T^2}\sum_{t,s=1}^{T}\sum_{i,j=1}^{N}E[\epsilon_{it}\epsilon_{js}] \\ &\leq \frac{1}{T^2}\sum_{t=1}^{T}\sum_{i,j=1}^{N}|\sigma_{ij}| \\ &= O\left(\frac{NT}{T^2}\sigma^2\right) = O(N).\end{aligned} \tag{IA.31}$$

Moreover, using assumptions 2 and 5(ii),

$$\begin{aligned}\mathrm{Var}\left(\frac{1}{T}\sum_{t=1}^{T}B'\epsilon_t\right) &= \frac{1}{T^2}\sum_{t,s=1}^{T}\sum_{i,j=1}^{N}E[\epsilon_{it}\epsilon_{js}]\beta_i\beta_j' \\ &\leq \frac{1}{T^2}\sum_{t=1}^{T}\sum_{i,j=1}^{N}|\beta_i\beta_j'||\sigma_{ij}| \\ &= O\left(\frac{NT}{T^2}\sigma^2\right) = O(N).\end{aligned} \tag{IA.32}$$

Putting the pieces together, $X'\bar{\epsilon} = O_p\left(\sqrt{N}\right)$. $\blacksquare$

Lemma 4 Under assumptions 3–5,

$$(\hat{X}-X)'X\Gamma^P = O_p\left(\sqrt{N}\right). \tag{IA.33}$$

Proof. We have

$$(\hat{X}-X)'X\Gamma^P = \begin{bmatrix} 0_N' \\ \mathcal{P}'\epsilon \end{bmatrix} X\Gamma^P. \tag{IA.34}$$

Using similar arguments to Equation (IA.15) concludes the proof. ■

Lemma 5 Under assumptions 3–5,

$$(\hat{X}-X)'\bar{\epsilon} = O_p\left(\sqrt{N}\right). \tag{IA.35}$$

Proof.

$$\begin{aligned}(\hat{X}-X)'\bar{\epsilon} &= \begin{bmatrix} 0 \\ \mathcal{P}'\epsilon\bar{\epsilon} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{P}'\epsilon\epsilon'\frac{1_T}{T} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \mathcal{P}'\left[\left(\epsilon\epsilon' - \sum_{i=1}^N \sigma_i^2 I_T\right) + \left(\sum_{i=1}^N \sigma_i^2 - N\sigma^2\right) I_T\right]\frac{1_T}{T} \end{bmatrix} = O_p(\sqrt{N})\end{aligned} \tag{IA.36}$$

by assumption 5. ■

Lemma 6 Under assumption 5 and the identification assumption $\kappa_4 = 0$, we have

$$\hat{\sigma}_4 \xrightarrow{p} \sigma_4. \tag{IA.37}$$

Proof. We need to show that (i) $E(\hat{\sigma}_4) \to \sigma_4$ and (ii) $\text{Var}(\hat{\sigma}_4) = O\left(\frac{1}{N}\right)$.

(i) By assumptions 5(iv), 5(vi), and 5(vii), we have

$$\begin{aligned} E\left[\frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \hat{\epsilon}_{it}^4\right] &= \frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N E\left[\hat{\epsilon}_{it}^4\right] \\ &= \frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \sum_{s_1,s_2,s_3,s_4=1}^T m_{ts_1}m_{ts_2}m_{ts_3}m_{ts_4}E\left[\epsilon_{is_1}\epsilon_{is_2}\epsilon_{is_3}\epsilon_{is_4}\right] \\ &= \frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \kappa_{4,iiii}\sum_{s=1}^T m_{ts}^4 + 3\frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N \sigma_i^4\left(\sum_{s=1}^T m_{ts}^2\right)^2 \\ &\to \kappa_4\sum_{t=1}^T\sum_{s=1}^T m_{ts}^4 + 3\sigma_4\sum_{t=1}^T\left(\sum_{s=1}^T m_{ts}^2\right)^2, \end{aligned} \tag{IA.38}$$

where $\hat{\epsilon}_{it} = \imath'_{t,T} M \epsilon_i$ and $M = [m_{ts}]$ for $t, s = 1, \ldots, T$. Note that

$$\begin{aligned}
\sum_{s=1}^{T} m_{ts}^2 &= \|m_t\|^2 \\
&= i_t' M i_t \\
&= i_t' \left(I_T - D(D'D)^{-1} D' \right) i_t \\
&= 1 - \mathrm{tr}\left(D(D'D)^{-1} D' i_t i_t' \right) \\
&= 1 - \mathrm{tr}\left(P i_t i_t' \right) \\
&= 1 - p_{tt} \\
&= m_{tt}, \qquad \text{(IA.39)}
\end{aligned}$$

where p_{tt} is the (t,t)-element of P. Then, we have

$$\sum_{t=1}^{T} \left(\sum_{s=1}^{T} m_{ts}^2 \right)^2 = \sum_{t=1}^{T} m_{tt}^2 = \mathrm{tr}\left(M^{(2)} \right). \qquad \text{(IA.40)}$$

By setting $\kappa_4 = 0$, it follows that

$$E[\hat{\sigma}_4] \rightarrow \sigma_4. \qquad \text{(IA.41)}$$

This concludes the proof of part (i).

(ii) As for the variance of $\hat{\sigma}_4$, we have

$$\begin{aligned}
\operatorname{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\epsilon}_{it}^4\right) &= \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\operatorname{Cov}\left(\hat{\epsilon}_{it}^4,\hat{\epsilon}_{js}^4\right) \\
&= \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\sum_{\substack{u_1,u_2,\\u_3,u_4=1}}^{T}\sum_{\substack{v_1,v_2,\\v_3,v_4=1}}^{T} m_{tu_1}m_{tu_2}m_{tu_3}m_{tu_4}m_{sv_1}m_{sv_2}m_{sv_3}m_{sv_4} \\
&\quad \times\operatorname{Cov}\left(\epsilon_{iu_1}\epsilon_{iu_2}\epsilon_{iu_3}\epsilon_{iu_4},\epsilon_{jv_1}\epsilon_{jv_2}\epsilon_{jv_3}\epsilon_{jv_4}\right) \\
&= \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t,s=1}^{T}\sum_{\substack{u_1,u_2,\\u_3,u_4=1}}^{T}\sum_{\substack{v_1,v_2,\\v_3,v_4=1}}^{T} m_{tu_1}m_{tu_2}m_{tu_3}m_{tu_4}m_{sv_1}m_{sv_2}m_{sv_3}m_{sv_4} \\
&\quad \times\Bigg(\kappa_8\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_1},\epsilon_{jv_2},\epsilon_{jv_3},\epsilon_{jv_4}\right) \\
&\quad +\sum^{(6,2)}\kappa_6\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_1},\epsilon_{jv_2}\right)\operatorname{Cov}\left(\epsilon_{jv_3},\epsilon_{jv_4}\right) \\
&\quad +\sum^{(4,4)}\kappa_4\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{jv_1},\epsilon_{jv_2}\right)\kappa_4\left(\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_3},\epsilon_{jv_4}\right) \\
&\quad +\sum^{(4,2,2)}\kappa_4\left(\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{jv_1},\epsilon_{jv_2}\right)\operatorname{Cov}\left(\epsilon_{iu_3},\epsilon_{iu_4}\right)\operatorname{Cov}\left(\epsilon_{jv_3},\epsilon_{jv_4}\right) \\
&\quad +\sum^{(2,2,2,2)}\operatorname{Cov}\left(\epsilon_{iu_1},\epsilon_{iu_2}\right)\operatorname{Cov}\left(\epsilon_{iu_3},\epsilon_{jv_1}\right)\operatorname{Cov}\left(\epsilon_{iu_4},\epsilon_{jv_2}\right)\operatorname{Cov}\left(\epsilon_{jv_3},\epsilon_{jv_4}\right)\Bigg),
\end{aligned} \tag{IA.42}$$

where $\kappa_4(\cdot)$, $\kappa_6(\cdot)$, and $\kappa_8(\cdot)$ denote the fourth-, sixth-, and eighth-order mixed cumulants, respectively. By $\sum^{(\nu_1,\nu_2,\ldots,\nu_k)}$ we denote the sum over all possible partitions of a group of K random variables into k subgroups of size $\nu_1,\nu_2,\ldots,\nu_k$, respectively. As an example, consider $\sum^{(6,2)}$. $\sum^{(6,2)}$ defines the sum over all possible partitions of the group of eight random variables $\{\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4},\epsilon_{jv_1},\epsilon_{jv_2},\epsilon_{jv_3},\epsilon_{jv_4}\}$ into two subgroups of size six and two, respectively. Moreover, since $E[\epsilon_{it}] = E\left[\epsilon_{it}^3\right] = 0$, we do not need to consider further partitions in the relation above.[1] Then, under assumptions 5(i), 5(ii), 5(v), and 5(viii), it follows that

$$\operatorname{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\epsilon}_{it}^4\right) = O\left(\frac{1}{N}\right) \tag{IA.43}$$

and $\operatorname{Var}(\hat{\sigma}_4) = O\left(\frac{1}{N}\right)$. This concludes the proof of part (ii). ■

[1] According to the theory on cumulants (see Brillinger 2001), evaluation of $\operatorname{Cov}\left(\epsilon_{iu_1}\epsilon_{iu_2}\epsilon_{iu_3}\epsilon_{iu_4},\epsilon_{jv_1}\epsilon_{jv_2}\epsilon_{jv_3}\epsilon_{jv_4}\right)$ requires considering the indecomposable partitions of the two sets, $\{\epsilon_{iu_1},\epsilon_{iu_2},\epsilon_{iu_3},\epsilon_{iu_4}\}$ and $\{\epsilon_{jv_1},\epsilon_{jv_2},\epsilon_{jv_3},\epsilon_{jv_4}\}$, meaning that there must be at least one subset that includes an element of both sets.

Lemma 7 Let $w = [w_1, \ldots, w_T]'$ and $s = [s_1, \ldots, s_T]'$ be two arbitrary T-vectors. Then, under Equation (64) and assumptions 2–7,

$$\frac{1}{N(T-K)} \sum_{i=1}^{N} \hat{\epsilon}_i' \hat{\epsilon}_i \sum_{k=1}^{T} w_k \epsilon_{ki} \sum_{r=1}^{T} s_r \epsilon_{ri} \xrightarrow{p} \frac{\text{tr}\left(M\left(S_1 + S_2\right)\right)}{(T-K)}, \tag{IA.44}$$

where $S_1 = \text{diag}\left[(s_1 w_1 \mu_4 + \sigma^4 \sum_{k\neq 1}^{T} w_k s_k), \ldots, (s_T w_T \mu_4 + \sigma^4 \sum_{k\neq T}^{T} w_k s_k)\right]$ and $S_2 = \sigma^4 \left(ws' + sw' - 2\text{diag}(w_1 s_1, \ldots, w_T s_T)\right)$.

Proof. Note that

$$\begin{aligned} & \frac{1}{N(T-K)} \sum_{i=1}^{N} \hat{\epsilon}_i' \hat{\epsilon}_i \sum_{k=1}^{T} \epsilon_{ki} \sum_{r=1}^{T} s_r \epsilon_{ri} = \\ = \; & \frac{1}{N(T-K)} \text{tr}\left(M\left(\sum_{i=1}^{N} \epsilon_i \epsilon_i' \left(\sum_{k=r=1}^{T} w_k s_k \epsilon_{ki}^2 + \sum_{r>k}^{T} w_k s_r \epsilon_{ik} \epsilon_{ir} + \sum_{r<k}^{T} w_k s_r \epsilon_{ik} \epsilon_{ir}\right)\right)\right). \end{aligned} \tag{IA.45}$$

For the first term of Equation (IA.45),

$$\begin{aligned} \frac{1}{N(T-K)} \text{tr}\left(M\left(\sum_{i=1}^{N} \epsilon_i \epsilon_i' \sum_{k=r=1}^{T} w_k s_k \epsilon_{ki}^2\right)\right) &= \frac{1}{(T-K)} \text{tr}\left(M\left(\frac{1}{N} \sum_{i=1}^{N} \sum_{k=r=1}^{T} \epsilon_i \epsilon_i' w_k s_k \epsilon_{ki}^2\right)\right) \\ &\xrightarrow{p} \frac{1}{(T-K)} \text{tr}\left(M S_1\right), \end{aligned} \tag{IA.46}$$

where

$$S_1 = \text{plim} \frac{1}{N} \sum_{i=1}^{N} \sum_{k=r=1}^{T} \epsilon_i \epsilon_i' w_k s_k \epsilon_{ki}^2 = \text{diag}\left[(s_1 w_1 \mu_4 + \sigma^4 \sum_{k\neq 1}^{T} w_k s_k), \ldots, (s_T w_T \mu_4 + \sigma^4 \sum_{k\neq T}^{T} w_k s_k)\right]. \tag{IA.47}$$

For the second and third terms of Equation (IA.45), we obtain

$$\begin{aligned} & \frac{1}{N(T-K)} \text{tr}\left(M\left(\sum_{i=1}^{N} \epsilon_i \epsilon_i' \left(\sum_{r>k}^{T} w_k s_r \epsilon_{ik} \epsilon_{ir} + \sum_{r<k}^{T} w_k s_r \epsilon_{ik} \epsilon_{ir}\right)\right)\right) \\ = \; & \frac{1}{(T-K)} \text{tr}\left(M\left(\frac{1}{N} \sum_{i=1}^{N} \sum_{r>k}^{T} \epsilon_i \epsilon_i' w_k s_r \epsilon_{ik} \epsilon_{ir} + \frac{1}{N} \sum_{i=1}^{N} \sum_{r<k}^{T} \epsilon_i \epsilon_i' w_k s_r \epsilon_{ik} \epsilon_{ir}\right)\right) \\ \xrightarrow{p} \; & \frac{1}{(T-K)} \text{tr}(M S_2), \end{aligned} \tag{IA.48}$$

where

$$\begin{aligned} S_2 &= \text{plim}\left(\frac{1}{N} \sum_{i=1}^{N} \sum_{r>k}^{T} \epsilon_i \epsilon_i' w_k s_r \epsilon_{ik} \epsilon_{ir} + \frac{1}{N} \sum_{i=1}^{N} \sum_{r<k}^{T} \epsilon_i \epsilon_i' w_k s_r \epsilon_{ik} \epsilon_{ir}\right) \\ &= \sigma^4 \left(ws' + sw' - 2\text{diag}(w_1 s_1, \ldots, w_T s_T)\right). \blacksquare \end{aligned} \tag{IA.49}$$

Lemma 8 Let $\hat{\tau}_\Phi = \frac{1}{N(T-K)} \sum_{i=1}^N \hat{\epsilon}_i' \hat{\epsilon}_i \hat{e}_i^P$. Then, under Equation (64) and assumptions 2–7,

$$\hat{\tau}_\Phi \xrightarrow{p} \tau_\Phi. \tag{IA.50}$$

Proof. Given

$$\begin{aligned} \hat{e}_i^P &= \bar{R}_i - \hat{X}_i' \hat{\Gamma}^* \\ &= X_i' \tilde{\Gamma}^P + e_i + \bar{\epsilon}_i - \hat{X}_i' \hat{\Gamma}^* \\ &= e_i + \bar{\epsilon}_i - (\hat{X}_i - X_i)' \tilde{\Gamma}^P - \hat{X}_i' (\hat{\Gamma}^* - \tilde{\Gamma}^P), \end{aligned} \tag{IA.51}$$

using the fact that $\hat{\epsilon}_i = M\epsilon_i$ and Equation (IA.51), we can write

$$\begin{aligned} \hat{\tau}_\Phi &= \frac{1}{N(T-K)} \sum_{i=1}^N \hat{\epsilon}_i' \hat{\epsilon}_i \hat{e}_i^P = \frac{1}{N(T-K)} \sum_{i=1}^N \epsilon_i' M M \epsilon_i \hat{e}_i \\ &= \frac{1}{N(T-K)} \sum_{i=1}^N \mathrm{tr}(M\epsilon_i' \epsilon_i)(e_i + \bar{\epsilon}_i - (\hat{X}_i - X_i)' \tilde{\Gamma}^P - \hat{X}_i' (\hat{\Gamma}^* - \tilde{\Gamma}^P)) \\ &= \frac{1}{N(T-K)} \sum_{i=1}^N \mathrm{tr}(M\epsilon_i' \epsilon_i e_i) + o_p(1) \xrightarrow{p} \frac{1}{(T-K)} \mathrm{tr}\left(M\tau_\Phi\right) = \tau_\Phi. \blacksquare \end{aligned} \tag{IA.52}$$

Lemma 9 Let

$$\hat{\tau}_\Omega = \frac{1}{N(T-K)} \sum_{i=1}^N \hat{\epsilon}_i' \hat{\epsilon}_i (\hat{e}_i^P)^2 - \frac{\sigma^4}{T} \left(1 + \frac{2\mathrm{tr}(M 1_T 1_T')}{T(T-K)}\right) - \frac{\mathrm{tr}(MS_F)}{(T-K)} + 2\frac{\mathrm{tr}(MC_F)}{T(T-K)}, \tag{IA.53}$$

where

$$S_F = \sigma^4 \begin{bmatrix} A'\left(3\tilde{f}_1\tilde{f}_1' + \sum_{t\neq 1}^T \tilde{f}_t \tilde{f}_t'\right) A & 2A'\tilde{f}_1\tilde{f}_2' A & \cdots & 2A'\tilde{f}_1\tilde{f}_T' A \\ 2A'\tilde{f}_2\tilde{f}_1' A & A'\left(3\tilde{f}_2\tilde{f}_2' + \sum_{t\neq 2}^T \tilde{f}_t \tilde{f}_t'\right) & \cdots & 2A'\tilde{f}_2\tilde{f}_T' A \\ \vdots & \vdots & \ddots & \vdots \\ 2A'\tilde{f}_T\tilde{f}_1' A & 2A'\tilde{f}_T\tilde{f}_2' A & \cdots & A'\left(3\tilde{f}_1\tilde{f}_1' + \sum_{t\neq T}^T \tilde{f}_T \tilde{f}_T'\right) A \end{bmatrix} \tag{IA.54}$$

and

$$C_F = \sigma^4 \begin{bmatrix} 3\tilde{f}_1' A + \sum_{t\neq 1}^T \tilde{f}_t' A & (\tilde{f}_1 + \tilde{f}_2)' A & \cdots & (\tilde{f}_1 + \tilde{f}_T)' A \\ (\tilde{f}_2 + \tilde{f}_1)' A & 3\tilde{f}_2' A + \sum_{t\neq 2}^T \tilde{f}_t' A & \cdots & (\tilde{f}_2 + \tilde{f}_T)' A \\ \vdots & \vdots & \ddots & \vdots \\ (\tilde{f}_T + \tilde{f}_1)' A & (\tilde{f}_T + \tilde{f}_2)' A & \cdots & 3\tilde{f}_T' A + \sum_{t\neq T}^T \tilde{f}_t' A \end{bmatrix}, \tag{IA.55}$$

with $A = (\tilde{F}'\tilde{F})^{-1}\tilde{\gamma}_1^P$. Then, under Equation (64) and assumptions 2–7,

$$\hat{\tau}_\Omega \xrightarrow{p} \tau_\Omega. \tag{IA.56}$$

Proof: By Equation (IA.51), we have

$$\begin{aligned}
(\hat{e}_i^P)^2 &= e_i^2 + \bar{\epsilon}_i^2 + \left((\hat{\beta}_i - \beta_i)'\tilde{\gamma}_1^P\right)^2 + \left([1, \hat{\beta}_i'](\hat{\Gamma}^* - \tilde{\Gamma}^P)\right)^2 \\
&+ 2e_i\left(\bar{\epsilon}_i - (\hat{\beta}_i - \beta_i)'\tilde{\gamma}_1^P - [1, \hat{\beta}_i'](\hat{\Gamma}^* - \tilde{\Gamma}^P)\right) \\
&+ 2\bar{\epsilon}_i\left(-(\hat{\beta}_i - \beta_i)'\tilde{\gamma}_1^P - [1, \hat{\beta}_i'](\hat{\Gamma}^* - \tilde{\Gamma}^P)\right) \\
&+ 2(\hat{\beta}_i - \beta_i)'\tilde{\gamma}_1^P[1, \hat{\beta}_i'](\hat{\Gamma}^* - \tilde{\Gamma}^P).
\end{aligned} \tag{IA.57}$$

Then,

$$\begin{aligned}
\hat{\tau}_\Omega &= \frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i(\hat{e}_i^P)^2 \\
&= \frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i e_i^2 + \frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i\bar{\epsilon}_i^2 + \frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i\left((\hat{\beta}_i - \beta_i)'\tilde{\gamma}_1^P\right)^2 \\
&- 2\frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i\bar{\epsilon}_i(\hat{\beta}_i - \beta_i)'\tilde{\gamma}_1^P + o_p(1),
\end{aligned} \tag{IA.58}$$

where all terms involving $(\hat{\Gamma}^* - \tilde{\Gamma}^P)$ are condensed into the $o_p(1)$ term. By assumption 7, the first term in Equation (IA.58) satisfies

$$\frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i e_i^2 = \frac{1}{(T-K)}\mathrm{tr}\left(M\frac{1}{N}\sum_{i=1}^N \epsilon_i\epsilon_i'.e_i^2\right) \xrightarrow{p} \frac{1}{(T-K)}\mathrm{tr}(M\tau_\Omega) = \tau_\Omega. \tag{IA.59}$$

For the second term in Equation (IA.58), we have

$$\frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i\bar{\epsilon}_i^2 = \frac{1}{T^2}\frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i\sum_{t=1}^T \epsilon_{it}\sum_{s=1}^T \epsilon_{is}. \tag{IA.60}$$

Then, applying Lemma 7 with $w = s = [1, \ldots, 1]'$, we have

$$\frac{1}{T^2}\frac{1}{N(T-K)}\sum_{i=1}^N \hat{\epsilon}_i'\hat{\epsilon}_i\sum_{t=1}^T \epsilon_{it}\sum_{s=1}^T \epsilon_{is} \xrightarrow{p} \frac{\sigma^4}{T}\left(1 + \frac{2\mathrm{tr}(M1_T1_T')}{T(T-K)}\right). \tag{IA.61}$$

For the third term in Equation (IA.58), we have

$$\frac{1}{N(T-K)}\sum_{i=1}^{N}\hat{\epsilon}_i'\hat{\epsilon}_i\left((\hat{\beta}_i-\beta_i)'\tilde{\gamma}_1^P\right)^2 = \frac{1}{N(T-K)}\sum_{i=1}^{N}\hat{\epsilon}_i'\hat{\epsilon}_i\sum_{t=1}^{T}\tilde{\gamma}_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}f_t\epsilon_{it}\sum_{s=1}^{T}\tilde{\gamma}_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}f_s\epsilon_{is}, \tag{IA.62}$$

and by Lemma 7 with $w = s = \left[\tilde{\gamma}_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}f_1,\ldots,\tilde{\gamma}_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}f_T\right]'$, one obtains

$$\frac{1}{N(T-K)}\sum_{i=1}^{N}\hat{\epsilon}_i'\hat{\epsilon}_i\left((\hat{\beta}_i-\beta_i)'\tilde{\gamma}_1^P\right)^2 \xrightarrow{p} \frac{\mathrm{tr}(MS_F)}{(T-K)}. \tag{IA.63}$$

Finally, for the fourth term in Equation (IA.58), rewriting it as

$$-2\frac{1}{N(T-K)}\sum_{i=1}^{N}\hat{\epsilon}_i'\hat{\epsilon}_i\bar{\epsilon}_i(\hat{\beta}_i-\beta_i)'\tilde{\gamma}_1^P = -2\frac{1}{NT(T-K)}\sum_{i=1}^{N}\hat{\epsilon}_i'\hat{\epsilon}_i\sum_{t=1}^{T}\epsilon_{it}\sum_{s=1}^{T}\epsilon_{is}\tilde{f}_s'(\tilde{F}'\tilde{F})^{-1}\tilde{\gamma}_1^P, \tag{IA.64}$$

and applying again Lemma 7 with $w = [1,\ldots,1]'$ and $s = [A'\tilde{f}_1,\ldots,A'\tilde{f}_T]'$, we obtain

$$-2\frac{1}{N(T-K)}\sum_{i=1}^{N}\hat{\epsilon}_i'\hat{\epsilon}_i\bar{\epsilon}_i(\hat{\beta}_i-\beta_i)'\tilde{\gamma}_1^P \xrightarrow{p} -2\frac{\mathrm{tr}(MC_F)}{T(T-K)}.\blacksquare \tag{IA.65}$$

IA.1.2 Proofs of Propositions and Theorems

Proof of Proposition 1. Consider the class of additive bias-adjusted estimators $\hat{\Gamma}^{bias-adj}$ for Γ^P:

$$\hat{\Gamma}^{bias-adj} = \hat{\Gamma} + (\frac{\hat{X}'\hat{X}}{N})^{-1}\hat{\Lambda}\hat{\Gamma}^{prelim} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R} + (\frac{\hat{X}'\hat{X}}{N})^{-1}\hat{\Lambda}\hat{\Gamma}^{prelim}, \tag{IA.66}$$

where $\hat{\Gamma}^{prelim}$ denotes any preliminary $\sqrt{N}$-consistent estimator of Γ^P. Setting $\hat{\Gamma}^{bias-adj} = \hat{\Gamma}^{prelim}$ and rearranging terms, we obtain

$$\left[I_{K+1} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\end{bmatrix}\right]\hat{\Gamma}^{bias-adj} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R}, \tag{IA.67}$$

which implies that

$$\hat{\Gamma}^{bias-adj} = \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\frac{\hat{X}'\bar{R}}{N} = \hat{\Gamma}^*.\blacksquare \tag{IA.68}$$

Proof of Proposition 2. By means of simple calculations, $\Sigma = \lambda\lambda' + \sigma_\eta^2 I_N$. Thus, $\sum_{i=1}^N\sigma_i^2/N = \sum_{i=1}^N(\lambda_i^2+\sigma_\eta^2)/N \to \sigma_\eta^2$ because $\sum_{i=1}^N\lambda_i^2 \le (\sum_{i=1}^N|\lambda_i|)^2 = O(N^{2\delta}) = o(N)$. Therefore, setting $\sigma^2 = \sigma_\eta^2$, one obtains $\sum_{i=1}^N(\sigma_i^2-\sigma^2)/N = \sum_{i=1}^N\lambda_i^2/N = (\lambda_1^2+\cdots+\lambda_q^2)/N + \sum_{i=q+1}^N\lambda_i^2/N = O(N^{2\delta-1}+N^{2\delta-1}) = o(1/\sqrt{N})$ since $\delta < 1/4$. It follows that assumption 5(i) is satisfied.

Next, given that $\sigma_{ij} = \lambda_i \lambda_j$ for $i \neq j$, we obtain $\sum_{i \neq j=1}^{N} |\sigma_{ij}| \leq (\sum_{i=1}^{N} |\lambda_i|)^2 = O(N^{2\delta}) = o(N)$, thus satisfying assumption 5(ii).

The maximum eigenvalue of Σ is bounded from below by the maximum eigenvalue of $\lambda\lambda'$, which equals $\lambda'\lambda$ (all the other $N-1$ eigenvalues of $\lambda\lambda'$ are zero), where $\lambda_1^2 + \cdots + \lambda_q^2 \leq \lambda'\lambda = O(N^{2\delta})$. Therefore, the maximum eigenvalue diverges at least at rate $o(\sqrt{N})$. ■

Proof of Proposition 3. The Fama and MacBeth (1973) standard errors with the Shanken (1992) correction are given by

$$SE_k^{FM} = \left((1+\hat{c})(\hat{W}_k - \mathbb{1}_{\{k>0\}}\hat{\sigma}_k^2) + \mathbb{1}_{\{k>0\}}\hat{\sigma}_k^2/T\right)^{\frac{1}{2}} \text{ and } SE_k^{FM,P} = \left((1+\hat{c})(W_k - \mathbb{1}_{\{k>0\}}\hat{\sigma}_k^2)\right)^{\frac{1}{2}}, \quad \text{(IA.69)}$$

for $k = 0, \ldots, K$, where $\hat{W}_k = \imath_{k+1,K+1}' \sum_{t=1}^{T} (\hat{\Gamma}_t - \bar{\hat{\Gamma}})(\hat{\Gamma}_t - \bar{\hat{\Gamma}})' \imath_{k+1,K+1}/(T-1)$, $\hat{\Gamma}_t = (\hat{X}'\hat{X})^{-1}\hat{X}'R_t$ with sample mean $\bar{\hat{\Gamma}}$, $\imath_{j,J}$ denotes the j-th column, for $j = 1, \ldots, J$, of the identity matrix I_J, $\hat{c} = \hat{\gamma}_1' \left(\tilde{F}'\tilde{F}/T\right)^{-1} \hat{\gamma}_1$, $\mathbb{1}_{\{\}}$ is the indicator function, and $\hat{\sigma}_k^2$ denotes the (k,k)-th element of $\tilde{F}'\tilde{F}/T$.

Consider the numerator of the t-ratios first. By Lemma 2(ii) and Lemmas 4 and 5, we obtain $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']' = (\Sigma_X + \Lambda)^{-1}\Sigma_X\Gamma^P + O_p\left(\frac{1}{\sqrt{N}}\right)$. By the blockwise formula of the inverse of a matrix (Magnus and Neudecker 2007, Section 1-11),

$$\begin{aligned}
(\Sigma_X + \Lambda)^{-1}\Sigma_X\Gamma^P &= \begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta + C \end{bmatrix}^{-1} \begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta \end{bmatrix} \Gamma^P \\
&= \begin{bmatrix} 1 + \mu_\beta' A^{-1}\mu_\beta & -\mu_\beta' A^{-1} \\ -A^{-1}\mu_\beta & A^{-1} \end{bmatrix} \begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta \end{bmatrix} \Gamma^P \\
&= \begin{bmatrix} 1 & \mu_\beta' - \mu_\beta' A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta') \\ 0 & A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta') \end{bmatrix} \Gamma^P. \qquad \text{(IA.70)}
\end{aligned}$$

Then,

$$\begin{aligned}
(\Sigma_X + \Lambda)^{-1}\Sigma_X\Gamma^P - \Gamma &= \begin{bmatrix} 1 & \mu_\beta' - \mu_\beta' A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta') \\ 0 & A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta') \end{bmatrix} \Gamma^P - \Gamma \\
&= \begin{bmatrix} 0 & \mu_\beta'(I_K - A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta')) \\ 0 & -(I_K - A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta')) \end{bmatrix} \Gamma \\
&+ \begin{bmatrix} 1 & \mu_\beta'(I_K - A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta')) \\ 0 & A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta') \end{bmatrix} \begin{bmatrix} 0 \\ \bar{f} - E[f_t] \end{bmatrix}. \qquad \text{(IA.71)}
\end{aligned}$$

Hence, $\text{plim}\ \hat{\gamma}_0 - \gamma_0 = \mu_\beta'(I_K - A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta'))\gamma_1^P = \mu_\beta' A^{-1} C \gamma_1^P$ and, for every $j = 1, \ldots, K$, $\text{plim}\ \hat{\gamma}_{1j} - \gamma_{1j} = -\imath_{j,K}'(I_K - A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta'))\gamma_1 + \imath_{j,K}' A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta')(\bar{f} - E[f_t])$ and $\text{plim}\ \hat{\gamma}_{1j} - \gamma_{1j}^P = -\imath_{j,K}'(I_K - A^{-1}(\Sigma_\beta - \mu_\beta\mu_\beta'))\gamma_1^P$. Consider now the behavior of the denominator of the t-ratios. It is easy to see that $\hat{W} = \frac{1}{T-1}\sum_{t=1}^{T}(\Gamma_t - \bar{\Gamma})(\Gamma_t - \bar{\Gamma})' = \hat{W}_a + \hat{W}_b + \hat{W}_c$, where

$$\hat{W}_a = (\hat{X}'\hat{X})^{-1}\hat{X}'\left[\frac{1}{T-1}\sum_{t=1}^{T}(\epsilon_t - \bar{\epsilon})(\epsilon_t - \bar{\epsilon})'\right]\hat{X}(\hat{X}'\hat{X})^{-1}, \tag{IA.72}$$

$$\hat{W}_b = (\hat{X}'\hat{X})^{-1}\hat{X}'B\left[(\frac{1}{T-1}\sum_{t=1}^{T}(f_t - \bar{f})(f_t - \bar{f})'\right]B'\hat{X}(\hat{X}'\hat{X})^{-1} \text{ and} \tag{IA.73}$$

$$\begin{aligned}\hat{W}_c &= (\hat{X}'\hat{X})^{-1}\hat{X}'\left[\frac{\sum_{t=1}^{T}(\epsilon_t - \bar{\epsilon})(f_t - \bar{f})'}{T-1}\right]B'\hat{X}(\hat{X}'\hat{X})^{-1} \\ &+(\hat{X}'\hat{X})^{-1}\hat{X}'B\left[\frac{\sum_{t=1}^{T}(f_t - \bar{f})(\epsilon_t - \bar{\epsilon})'}{T-1}\right]\hat{X}(\hat{X}'\hat{X})^{-1}.\end{aligned} \tag{IA.74}$$

Based on Lemmas 2-4 (details are available upon request), we obtain

$$\begin{aligned}\hat{W} \to_p W = W_a + W_b + W_c &\equiv (\Sigma_X + \Lambda)^{-1}\begin{bmatrix} 0 & 0_K' \\ 0_K & \frac{\sigma^4}{(T-1)}(\tilde{F}'\tilde{F})^{-1}\end{bmatrix}(\Sigma_X + \Lambda)^{-1} \\ &+ (\Sigma_X + \Lambda)^{-1}\begin{bmatrix}\mu_\beta' \\ \Sigma_\beta\end{bmatrix}\left[\frac{\tilde{F}'\tilde{F}}{T-1}\right]\left[\mu_\beta, \Sigma_\beta\right](\Sigma_X + \Lambda)^{-1} \\ &+ (\Sigma_X + \Lambda)^{-1}\frac{\sigma^2}{T-1}\begin{bmatrix} 0 & \mu_\beta' \\ \mu_\beta & 2\Sigma_\beta\end{bmatrix}(\Sigma_X + \Lambda)^{-1}.\end{aligned} \tag{IA.75}$$

$$W = \begin{bmatrix} 0 & 0_K' \\ 0_K & \frac{(\tilde{F}'\tilde{F})}{T-1}\end{bmatrix}. \tag{IA.76}$$

Therefore, since $\hat{W}_k = \imath_{k+1,K+1}'\hat{W}\imath_{k+1,K+1}$ for $k = 0, \ldots, K$, we have $(1+\hat{c})(\hat{W}_k - \mathbb{1}_{\{k>0\}}\hat{\sigma}_k^2) \to_p 0$ for any value of $\hat{c}$. It follows that $SE_k^{FM} \to_p \hat{\sigma}_k/\sqrt{T}$ and $SE_k^{FM,P} \to_p 0$. The proof of parts (i) and (ii) follows from dividing $\hat{\gamma}_0 - \gamma_0$, $\hat{\gamma}_{1k} - \gamma_{1k}$, and $\hat{\gamma}_{1k} - \gamma_{1k}^P$ by SE_k^{FM} and $SE_k^{FM,P}$, for the ex ante and ex post risk premia, respectively, and then taking the limit as $N \to \infty$. ■

Proof of Theorem 1. For part (i), starting from Equation (12), we have

$$\begin{aligned}\hat{\Gamma}^* &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\frac{\hat{X}'\bar{R}}{N} \\ &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\frac{\hat{X}'}{N}\left[\hat{X}\Gamma^P + \bar{\epsilon} - (\hat{X} - X)\Gamma^P\right] \\ &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'\hat{X}}{N}\Gamma^P + \frac{\hat{X}'}{N}\bar{\epsilon} - \frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P\right] \\ &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left(\frac{\hat{X}'\hat{X}}{N}\right)\left[\Gamma^P + \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\frac{\hat{X}'}{N}\bar{\epsilon} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P\right] \\ &= \left[I_{K+1} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\hat{\Lambda}\right]^{-1}\left[\Gamma^P + \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\frac{\hat{X}'}{N}\bar{\epsilon} - \left(\frac{\hat{X}'\hat{X}}{N}\right)^{-1}\frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P\right].\end{aligned} \tag{IA.77}$$

Hence,

$$
\begin{aligned}
\hat{\Gamma}^* - \Gamma^P &= \left(\frac{\hat{X}'\hat{X}}{N} - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'}{N}\bar{\epsilon} - \frac{\hat{X}'}{N}(\hat{X} - X)\Gamma^P + \hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'}{N}\bar{\epsilon} - \left(\frac{\hat{X}'}{N}(\hat{X} - X) - \hat{\Lambda}\right)\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'}{N}\bar{\epsilon} - \begin{bmatrix} 1_N'\frac{\epsilon'}{N}\mathcal{P}\gamma_1^P \\ \frac{B'\epsilon'}{N}\mathcal{P}\gamma_1^P + \mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P - \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\gamma_1^P \end{bmatrix}\right].
\end{aligned}
\tag{IA.78}
$$

By Lemmas 1 and 2(i), $\left(\hat{\Sigma}_X - \hat{\Lambda}\right) = O_p(1)$. In addition, Lemmas 3 and 5 imply that

$$
\begin{aligned}
\frac{\hat{X}'\bar{\epsilon}}{N} &= \frac{1}{N}(\hat{X} - X)'\bar{\epsilon} + \frac{1}{N}X'\bar{\epsilon} \\
&= O_p\left(\frac{1}{\sqrt{N}}\right),
\end{aligned}
\tag{IA.79}
$$

and assumption 6(i) implies that

$$
\mathcal{P}'\sum_{i=1}^{N}\epsilon_i = O_p\left(\sqrt{N}\right). \tag{IA.80}
$$

Note that

$$
\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P - \hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\gamma_1^P \tag{IA.81}
$$

can be rewritten as

$$
\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 I_T\right)\mathcal{P}\gamma_1^P - \left[(\hat{\sigma}^2 - \sigma^2) - \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 - \sigma^2\right)\right](\tilde{F}'\tilde{F})^{-1}\gamma_1^P. \tag{IA.82}
$$

Assumption 6(ii) implies that

$$
\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \frac{\sum_{i=1}^{N}\sigma_i^2}{N}I_T\right)\mathcal{P}\gamma_1^P = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{IA.83}
$$

Using Lemma 1 and assumption 5(i) concludes the proof of part (i) since $\hat{\sigma}^2 - \sigma^2 = O_p\left(\frac{1}{\sqrt{N}}\right)$ and $\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 - \sigma^2 = o\left(\frac{1}{\sqrt{N}}\right)$.

For part (ii), starting from (IA.78), we have

$$
\begin{aligned}
\sqrt{N}(\hat{\Gamma}^* - \Gamma^P) &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'\bar{\epsilon}}{\sqrt{N}} - \left(\frac{\hat{X}'}{\sqrt{N}}(\hat{X}-X)\Gamma^P\right) + \sqrt{N}\hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'\bar{\epsilon}}{\sqrt{N}} - \begin{bmatrix} 1_N' \\ \hat{B}' \end{bmatrix}\left[0_N,\ \frac{\epsilon'\mathcal{P}}{\sqrt{N}}\right]\Gamma^P + \sqrt{N}\hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\frac{X'\bar{\epsilon}}{\sqrt{N}} + \frac{1}{\sqrt{N}}\begin{bmatrix} 0_N' \\ \mathcal{P}'\epsilon \end{bmatrix}\frac{\epsilon' 1_T}{T} - \frac{1}{\sqrt{N}}\begin{bmatrix} 1_N'\epsilon'\mathcal{P} \\ \hat{B}'\epsilon'\mathcal{P} \end{bmatrix}\gamma_1^P + \sqrt{N}\hat{\Lambda}\Gamma^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\begin{bmatrix} 1_N' \\ B' \end{bmatrix}\frac{\epsilon' 1_T}{T\sqrt{N}} + \begin{bmatrix} -1_N'\frac{\epsilon'\mathcal{P}}{\sqrt{N}}\gamma_1^P \\ \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}\frac{1_T}{T} - B'\frac{\epsilon'\mathcal{P}}{\sqrt{N}}\gamma_1^P - \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}\mathcal{P}\gamma_1^P \end{bmatrix}\right. \\
&\qquad \left. + \sqrt{N}\hat{\sigma}^2(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\begin{bmatrix} \frac{1_N'}{\sqrt{N}}\epsilon'\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) \\ \frac{B'\epsilon'}{\sqrt{N}}\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) + \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) + \frac{\mathrm{tr}(M\epsilon\epsilon')}{\sqrt{N}(T-K-1)}\mathcal{P}'\mathcal{P}\gamma_1^P \end{bmatrix} \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left[\begin{bmatrix} \frac{1_N'\epsilon'}{\sqrt{N}}Q \\ \frac{B'\epsilon'}{\sqrt{N}}Q \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}Q + \frac{\mathrm{tr}(M\epsilon\epsilon')}{\sqrt{N}(T-K-1)}\mathcal{P}'\mathcal{P}\gamma_1^P \end{bmatrix}\right] \\
&= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}(I_1 + I_2).
\end{aligned} \tag{IA.84}
$$

Using Lemmas 1 and 2(ii), we have

$$
\left(\hat{\Sigma}_X - \hat{\Lambda}\right) \xrightarrow{p} \left(\begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta + \sigma^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix} - \begin{bmatrix} 0 & 0_K' \\ 0_K & \sigma^2(\tilde{F}'\tilde{F})^{-1} \end{bmatrix}\right) = \Sigma_X. \tag{IA.85}
$$

Consider now the terms I_1 and I_2. Both terms have a zero mean and, under assumption 5(vi), they are asymptotically uncorrelated. Assumptions 2, 5(i), 6(i), and 6(iii) imply that

$$
\begin{aligned}
\mathrm{Var}(I_1) &= E\begin{bmatrix} Q'\frac{1}{\sqrt{N}}\sum_{i=1}^N \epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j' Q & Q'\frac{1}{\sqrt{N}}\sum_{i=1}^N \epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j'(Q\otimes\beta_j') \\ \frac{1}{\sqrt{N}}\sum_{i=1}^N (Q'\otimes\beta_i)\epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j' Q & \frac{1}{\sqrt{N}}\sum_{i=1}^N (Q'\otimes\beta_i)\epsilon_i \frac{1}{\sqrt{N}}\sum_{j=1}^N \epsilon_j'(Q\otimes\beta_j') \end{bmatrix} \\
&= \begin{bmatrix} Q'\frac{1}{N}\sum_{i=1}^N E[\epsilon_i\epsilon_i']Q & Q'\frac{1}{N}\sum_{i=1}^N E[\epsilon_i\epsilon_i'](Q\otimes\beta_i') \\ \frac{1}{N}\sum_{i=1}^N (Q'\otimes\beta_i)E[\epsilon_i\epsilon_i']Q & \frac{1}{N}\sum_{i=1}^N (Q'\otimes\beta_i)E[\epsilon_i\epsilon_i'](Q\otimes\beta_i') \end{bmatrix} + o(1) \\
&\to \begin{bmatrix} \sigma^2 Q'Q & \sigma^2 Q'(Q\otimes\mu_\beta') \\ \sigma^2(Q'\otimes\mu_\beta)Q & \sigma^2(Q'Q\otimes\Sigma_\beta) \end{bmatrix} \\
&= \sigma^2 Q'Q\Sigma_X = \frac{\sigma^2}{T}\left[1 + \gamma_1^{P\prime}\left(\tilde{F}'\tilde{F}/T\right)^{-1}\gamma_1^P\right]\Sigma_X.
\end{aligned} \tag{IA.86}
$$

Next, consider I_2. Since $\mathcal{P}'\frac{1}{\sqrt{N}}\sum_{i=1}^N \sigma_i^2 Q + \frac{1}{T-K-1}\mathrm{tr}\left(M\frac{1}{\sqrt{N}}\sum_{i=1}^N \sigma_i^2\right)\mathcal{P}'\mathcal{P}\gamma_1^P = 0_K$, we have

$$
\begin{aligned}
I_2 &= \begin{bmatrix} 0 \\ (Q'\otimes P')\mathrm{vec}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N(\epsilon_i\epsilon_i' - \sigma_i^2 I_T)\right) + \frac{1}{T-K-1}\mathrm{tr}\left(M\frac{1}{\sqrt{N}}\sum_{i=1}^N(\epsilon_i\epsilon_i' - \sigma_i^2 I_T)\right)\mathcal{P}'\mathcal{P}\gamma_1^P \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ I_{22} \end{bmatrix}.
\end{aligned} \tag{IA.87}
$$

Therefore, $\mathrm{Var}(I_2)$ has the following form:

$$\mathrm{Var}(I_2) = \begin{bmatrix} 0 & 0_K' \\ 0_K & E\left[I_{22}I_{22}'\right] \end{bmatrix}. \tag{IA.88}$$

Under assumptions 5(i) and 6(ii), we have

$$\begin{aligned} E\left[I_{22}I_{22}'\right] &= E\Bigg[(Q'\otimes\mathcal{P}')\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'(Q\otimes\mathcal{P})\Bigg] \\ &\quad+E\Bigg[(Q'\otimes\mathcal{P}')\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'\frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\Bigg] \\ &\quad+E\Bigg[\mathcal{P}'\mathcal{P}\gamma_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'(Q\otimes\mathcal{P})\Bigg] \\ &\quad+E\Bigg[\mathcal{P}'\mathcal{P}\gamma_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathrm{vec}(\epsilon_i\epsilon_i'-\sigma_i^2 I_T)\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\mathrm{vec}(\epsilon_j\epsilon_j'-\sigma_j^2 I_T)'\frac{\mathrm{vec}(M)}{T-K-1} \\ &\quad\times\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\Bigg] \\ &\to \left[(Q'\otimes\mathcal{P}')+\mathcal{P}'\mathcal{P}\gamma_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\right]U_\epsilon\left[(Q\otimes\mathcal{P})+\frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\right]. \end{aligned} \tag{IA.89}$$

Defining $Z = \left[(Q\otimes\mathcal{P})+\frac{\mathrm{vec}(M)}{T-K-1}\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\right]$ concludes the proof of part (ii). ■

Proof of Theorem 2. By Theorem 1(i), $\hat{\gamma}_1^* \xrightarrow{p} \gamma_1^P$. Lemma 1 implies that $\hat{\Lambda}$ is a consistent estimator of Λ. Hence, using Lemma 2(ii), we have $\left(\hat{\Sigma}_X-\hat{\Lambda}\right)\xrightarrow{p}\Sigma_X$, which implies that $\hat{V}\xrightarrow{p}V$. A consistent estimator of W requires a consistent estimate of the matrix U_ϵ, which can be obtained using Lemma 6. This concludes the proof of Theorem 2. ■

Proof of Theorem 3. Writing

$$
\begin{aligned}
(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\frac{\hat{X}'R_t}{N} &= (\hat{\Sigma}_X-\hat{\Lambda})^{-1}\hat{\Sigma}_X\Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\hat{X}'\epsilon'\imath_{t,T}+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\hat{X}'(X-\hat{X})\Gamma_{t-1}^P\\
&= (\hat{\Sigma}_X-\hat{\Lambda})^{-1}(\hat{\Sigma}_X-\hat{\Lambda}+\hat{\Lambda})\Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\hat{X}'\epsilon'\imath_{t,T}+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\hat{X}'(X-\hat{X})\Gamma_{t-1}^P\\
&= \Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\left(\frac{\hat{X}'\epsilon'\imath_{t,T}}{N}+\frac{\hat{X}'(X-\hat{X})}{N}\Gamma_{t-1}^P+\hat{\Lambda}\Gamma_{t-1}^P\right)\\
&= \Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\left(\frac{X'\epsilon'\imath_{t,T}}{N}+\frac{(\hat{X}-X)'\epsilon'\imath_{t,T}}{N}+\frac{\hat{X}'(X-\hat{X})}{N}\Gamma_{t-1}^P+\hat{\Lambda}\Gamma_{t-1}^P\right)\\
&= \Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\left(\begin{bmatrix}1_N'\\ B'\end{bmatrix}\frac{\epsilon'\imath_{t,T}}{N}+\frac{1}{N}\begin{bmatrix}0_N'\\ \mathcal{P}'\epsilon\end{bmatrix}\epsilon'\imath_{t,T}\right.\\
&\left.+\frac{1}{N}\begin{bmatrix}-1_N'\epsilon'\mathcal{P}\gamma_{1,t-1}^P\\ -B'\epsilon'\mathcal{P}\gamma_{1,t-1}^P-\mathcal{P}'\epsilon\epsilon'\mathcal{P}\gamma_{1,t-1}^P\end{bmatrix}+\hat{\Lambda}\Gamma_{t-1}^P\right)\\
&= \Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\left(\begin{bmatrix}\frac{1_N'\epsilon'}{N}Q_{t-1}\\ \frac{B'\epsilon'}{N}Q_{t-1}\end{bmatrix}+\begin{bmatrix}0\\ \frac{P'\epsilon\epsilon'}{N}Q_{t-1}\end{bmatrix}+\hat{\Lambda}\Gamma_{t-1}^P\right)\\
&= \Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\left(\begin{bmatrix}\frac{1_N'\epsilon'}{N}Q_{t-1}\\ \frac{B'\epsilon'}{N}Q_{t-1}\end{bmatrix}+\begin{bmatrix}0\\ -\frac{P'\epsilon\epsilon'}{N}P\gamma_{1,t-1}^P\end{bmatrix}+\hat{\Lambda}\Gamma_{t-1}^P+\begin{bmatrix}0\\ \frac{P'\epsilon\epsilon'}{N}\imath_{t,T}\end{bmatrix}\right)
\end{aligned}
\tag{IA.90}
$$

with

$$
E\left(\begin{bmatrix}0\\ -\frac{\mathcal{P}'\epsilon\epsilon'}{N}\mathcal{P}\gamma_{1,t-1}^P\end{bmatrix}+\hat{\Lambda}\Gamma_{t-1}^P\right)=E\left(\begin{bmatrix}0\\ -\frac{\mathcal{P}'\epsilon\epsilon'}{N}\mathcal{P}\gamma_{1,t-1}^P\end{bmatrix}+\begin{bmatrix}0\\ \frac{\mathrm{tr}(M\epsilon\epsilon')}{N(T-K-1)}\mathcal{P}'\mathcal{P}\gamma_{1,t-1}^P\end{bmatrix}\right)=0_{K+1}
\tag{IA.91}
$$

and

$$
\begin{bmatrix}0\\ \frac{P'\epsilon\epsilon'}{N}\imath_{t,T}\end{bmatrix}\rightarrow_p\begin{bmatrix}0\\ \sigma^2P'\imath_{t,T}\end{bmatrix}=\begin{bmatrix}0\\ \sigma^2(\tilde{F}'\tilde{F})^{-1}\tilde{f}_t\end{bmatrix}
\tag{IA.92}
$$

yields part (i).

Next,

$$
\begin{aligned}
\hat{\Gamma}_{t-1}^* &= (\hat{\Sigma}_X-\hat{\Lambda})^{-1}\frac{\hat{X}'R_t}{N}-(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\begin{bmatrix}0\\ \hat{\sigma}^2P'\imath_{t,T}\end{bmatrix}\\
&= \Gamma_{t-1}^P+(\hat{\Sigma}_X-\hat{\Lambda})^{-1}\left(\begin{bmatrix}\frac{1_N'\epsilon'}{N}Q_{t-1}\\ \frac{B'\epsilon'}{N}Q_{t-1}\end{bmatrix}+\begin{bmatrix}0\\ \frac{P'\epsilon\epsilon'}{N}Q_{t-1}\end{bmatrix}-\begin{bmatrix}0\\ \hat{\sigma}^2P'Q_{t-1}\end{bmatrix}\right).
\end{aligned}
\tag{IA.93}
$$

The part of $\sqrt{N}(\hat{\Gamma}_{t-1}^*-\Gamma_{t-1}^P)$ that depends on $\epsilon\epsilon'$ can be written as

$$
\begin{aligned}
&(\hat{\Sigma}_X-\hat{\Lambda})^{-1}[(Q_{t-1}'\otimes P')-P'Q_{t-1}\mathrm{vec}(M)']\mathrm{vec}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N(\epsilon_i\epsilon_i'-\sigma_i^2I_T)\right)\\
&= (\hat{\Sigma}_X-\hat{\Lambda})^{-1}Z_{t-1}'\mathrm{vec}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N(\epsilon_i\epsilon_i'-\sigma_i^2I_T)\right),
\end{aligned}
\tag{IA.94}
$$

and the result follows along the proof of Theorem 1(ii). ■

Proof of Theorem 4. We first establish a simpler, asymptotically equivalent, expression for $\sqrt{N}\left(\frac{\hat{e}^{P\prime}\hat{e}^P}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}\right)$. Then, we derive the asymptotic distribution of this approximation. Consider the sample ex post pricing errors,

$$\hat{e}^P = \bar{R} - \hat{X}\hat{\Gamma}^*. \tag{IA.95}$$

Starting from $\bar{R} = \hat{X}\Gamma^P + \eta^P$ with $\eta^P = \bar{\epsilon} - (\hat{X} - X)\Gamma^P$, we have

$$\begin{aligned}\hat{e}^P &= \hat{X}\Gamma^P + \bar{\epsilon} - (\hat{X} - X)\Gamma^P - \hat{X}\hat{\Gamma}^* \\ &= \bar{\epsilon} - \hat{X}(\hat{\Gamma}^* - \Gamma^P) - (\hat{X} - X)\Gamma^P. \end{aligned}\tag{IA.96}$$

Then,

$$\begin{aligned}\hat{e}^{P\prime}\hat{e}^P &= \bar{\epsilon}'\bar{\epsilon} + \Gamma^{P\prime}(\hat{X} - X)'(\hat{X} - X)\Gamma^P - 2(\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\bar{\epsilon} - 2\Gamma^{P\prime}(\hat{X} - X)'\bar{\epsilon} \\ &\quad + 2\Gamma^{P\prime}(\hat{X} - X)'\hat{X}(\hat{\Gamma}^* - \Gamma^P) + (\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\hat{X}(\hat{\Gamma}^* - \Gamma^P).\end{aligned}$$

Note that

$$\frac{\bar{\epsilon}'\bar{\epsilon}}{N} = \frac{1}{T^2}1_T'\frac{\epsilon\epsilon'}{N}1_T \xrightarrow{p} \frac{\sigma^2}{T}, \tag{IA.97}$$

and, by Lemma 2(iii),

$$\Gamma^{P\prime}\frac{(\hat{X} - X)'(\hat{X} - X)}{N}\Gamma^P = \gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P \xrightarrow{p} \sigma^2\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P. \tag{IA.98}$$

Using Lemmas 3 and 5 and Theorem 1, we have

$$\frac{(\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\bar{\epsilon}}{N} = \frac{(\hat{\Gamma}^* - \Gamma^P)'(\hat{X} - X)'\bar{\epsilon}}{N} + \frac{(\hat{\Gamma}^* - \Gamma^P)'X'\bar{\epsilon}}{N} = O_p\left(\frac{1}{N}\right) \tag{IA.99}$$

and

$$\frac{\Gamma^{P\prime}(\hat{X} - X)'\bar{\epsilon}}{N} = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{IA.100}$$

In addition, using Lemmas 2(i), 2(iii), 4 and Theorem 1, we have

$$\begin{aligned}\frac{\Gamma^{P\prime}(\hat{X} - X)'\hat{X}(\hat{\Gamma}^* - \Gamma^P)}{N} &= \frac{\Gamma^{P\prime}(\hat{X} - X)'(\hat{X} - X)(\hat{\Gamma}^* - \Gamma^P)}{N} + \frac{\Gamma^{P\prime}(\hat{X} - X)'X(\hat{\Gamma}^* - \Gamma^P)}{N} \\ &= O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{N}\right)\end{aligned}\tag{IA.101}$$

and

$$\frac{(\hat{\Gamma}^* - \Gamma^P)'\hat{X}'\hat{X}(\hat{\Gamma}^* - \Gamma^P)}{N} = O_p\left(\frac{1}{N}\right). \tag{IA.102}$$

It follows that

$$\frac{\hat{e}^{P\prime}\hat{e}^P}{N} \xrightarrow{p} \frac{\sigma^2}{T} + \sigma^2\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P = \sigma^2 Q'Q. \tag{IA.103}$$

Collecting terms and rewriting explicitly only the ones that are $O_p\left(\frac{1}{\sqrt{N}}\right)$, we have

$$\frac{\hat{e}^{P\prime}\hat{e}^P}{N} = \frac{\bar{\epsilon}'\bar{\epsilon}}{N} \tag{IA.104}$$

$$+\frac{\Gamma^{P\prime}(\hat{X} - X)'(\hat{X} - X)\Gamma^P}{N} \tag{IA.105}$$

$$-2\,\frac{\Gamma^{P\prime}(\hat{X} - X)'\bar{\epsilon}}{N} \tag{IA.106}$$

$$+2\,\frac{\Gamma^{P\prime}(\hat{X} - X)'(\hat{X} - X)(\hat{\Gamma}^* - \Gamma^P)}{N} \tag{IA.107}$$

$$+O_p\left(\frac{1}{N}\right). \tag{IA.108}$$

Consider the sum of the three terms in Equations (IA.104)–(IA.106). Under assumption 5(i), we have

$$\begin{aligned}
& \frac{\bar{\epsilon}'\bar{\epsilon}}{N} + \frac{\Gamma^{P\prime}(\hat{X} - X)'(\hat{X} - X)\Gamma^P}{N} - 2\frac{\Gamma^{P\prime}(\hat{X} - X)'\bar{\epsilon}}{N} \\
= \; & \frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\frac{1_T}{T} + \gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P - 2\frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P \\
= \; & \frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) - \frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P + \gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P \\
= \; & \frac{1_T'}{T}\frac{\epsilon\epsilon'}{N}Q - Q'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P \\
= \; & Q'\frac{\epsilon\epsilon'}{N}\frac{1_T}{T} - Q'\frac{\epsilon\epsilon'}{N}\mathcal{P}\gamma_1^P \\
= \; & Q'\frac{\epsilon\epsilon'}{N}Q = Q'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)Q + \sigma^2 Q'Q + o\left(\frac{1}{\sqrt{N}}\right),
\end{aligned} \tag{IA.109}$$

where the $o\left(\frac{1}{\sqrt{N}}\right)$ term comes from $(\bar{\sigma}^2 - \sigma^2)Q'Q$. As for the term in Equation (IA.107), define

$$\left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1} = \begin{bmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{bmatrix}, \tag{IA.110}$$

where every block of $\left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}$ is $O_p(1)$ by the nonsingularity of Σ_X and Slutsky's theorem. Using the same arguments as for Theorem 2, we have

$$
\begin{aligned}
& 2\frac{\Gamma^{P\prime}(\hat{X}-X)'(\hat{X}-X)(\hat{\Gamma}^*-\Gamma^P)}{N} \\
= \;& 2\left[\gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\hat{\Sigma}_{21},\; \gamma_1^{P\prime}\mathcal{P}'\frac{\epsilon\epsilon'}{N}\mathcal{P}\hat{\Sigma}_{22}\right]\left[\begin{array}{c} \frac{1_N'\epsilon'Q}{N} \\ \frac{B'\epsilon'Q}{N} + Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) \end{array}\right] \\
= \;& 2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + o_p\left(\frac{1}{N}\right) \\
= \;& 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + o_p\left(\frac{1}{N}\right) + O_p\left(\frac{1}{N}\right),
\end{aligned} \tag{IA.111}
$$

where the two approximations on the right-hand side of the previous expression refer to

$$
\begin{aligned}
& 2(\bar{\sigma}^2 - \sigma^2)\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2(\bar{\sigma}^2 - \sigma^2)\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2(\bar{\sigma}^2 - \sigma^2)\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) = o_p\left(\frac{1}{N}\right)
\end{aligned} \tag{IA.112}
$$

and

$$
\begin{aligned}
& 2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\gamma_1^{P\prime}\mathcal{P}'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) = O_p\left(\frac{1}{N}\right),
\end{aligned} \tag{IA.113}
$$

respectively. Therefore, we have

$$
\begin{aligned}
\frac{\hat{e}^{P\prime}\hat{e}^P}{N} = \;& Q'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)Q + \sigma^2 Q'Q \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\
& +2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\mathrm{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + O_p\left(\frac{1}{N}\right) + o_p\left(\frac{1}{N}\right) + o\left(\frac{1}{\sqrt{N}}\right).
\end{aligned} \tag{IA.114}
$$

It follows that

$$\frac{\hat{e}^{P\prime}\hat{e}^{P}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q} = Q'\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right)Q - \left(\hat{\sigma}^2\hat{Q}'\hat{Q} - \sigma^2 Q'Q\right) + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\text{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + O_p\left(\frac{1}{N}\right) + o_p\left(\frac{1}{N}\right) + o\left(\frac{1}{\sqrt{N}}\right). \tag{IA.115}$$

Note that

$$\begin{aligned} & \hat{\sigma}^2\hat{Q}'\hat{Q} - \sigma^2 Q'Q \\ = \ & \frac{1}{T}(\hat{\sigma}^2 - \sigma^2) + \hat{\sigma}^2\hat{\gamma}_1^{*\prime}(\tilde{F}'\tilde{F})^{-1}\hat{\gamma}_1^* - \sigma^2\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P \\ = \ & \frac{1}{T}(\hat{\sigma}^2 - \sigma^2) + (\hat{\sigma}^2 - \sigma^2)\gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P + 2\sigma^2(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}\gamma_1^P + O_p\left(\frac{1}{N}\right) \\ = \ & (\hat{\sigma}^2 - \sigma^2)\left(\frac{1}{T} + \gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\right) + 2\sigma^2(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}\gamma_1^P + O_p\left(\frac{1}{N}\right) \\ = \ & (\hat{\sigma}^2 - \sigma^2)\left(\frac{1}{T} + \gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\right) + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{21}\frac{1_N'\epsilon'Q}{N} + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}\frac{B'\epsilon'Q}{N} \\ & + 2\sigma^2\gamma_1^{P\prime}\mathcal{P}'\mathcal{P}\hat{\Sigma}_{22}Z'\text{vec}\left(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\right) + O_p\left(\frac{1}{N}\right) + O_p\left(\frac{1}{N\sqrt{N}}\right), \end{aligned} \tag{IA.116}$$

where $\sigma^2(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}(\hat{\gamma}_1^* - \gamma_1^P) + 2\,(\hat{\sigma}^2 - \sigma^2)(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}\gamma_1^P = O_p\left(\frac{1}{N}\right)$ and $(\hat{\sigma}^2 - \sigma^2)(\hat{\gamma}_1^* - \gamma_1^P)'(\tilde{F}'\tilde{F})^{-1}(\hat{\gamma}_1^* - \gamma_1^P) = O_p\left(\frac{1}{N\sqrt{N}}\right)$. It follows that

$$\begin{aligned} & \frac{\hat{e}'\hat{e}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q} \\ = \ & Q'\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big)Q - (\hat{\sigma}^2 - \sigma^2)\Big(\frac{1}{T} + \gamma_1^{P\prime}(\tilde{F}'\tilde{F})^{-1}\gamma_1^P\Big) + O_p\Big(\frac{1}{N\sqrt{N}}\Big) + O_p\Big(\frac{1}{N}\Big) + o\Big(\frac{1}{\sqrt{N}}\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big) \\ = \ & \Big[\Big(Q' \otimes Q'\Big) - \frac{Q'Q}{T-K-1}\text{vec}(M)'\Big]\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big) \\ = \ & Z_Q'\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) + o_p\Big(\frac{1}{\sqrt{N}}\Big), \end{aligned} \tag{IA.117}$$

where, for simplicity, we have condensed $O_p\big(\frac{1}{N\sqrt{N}}\big) + O_p\big(\frac{1}{N}\big) + o\big(\frac{1}{\sqrt{N}}\big) + o_p\big(\frac{1}{\sqrt{N}}\big)$ into the single term $o_p\big(\frac{1}{\sqrt{N}}\big)$. Hence,

$$\sqrt{N}\Big(\frac{\hat{e}'\hat{e}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}\Big) = \sqrt{N}Z_Q'\text{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) + o_p(1), \tag{IA.118}$$

implying that the asymptotic distribution of $\sqrt{N}\Big(\frac{\hat{e}'\hat{e}}{N} - \hat{\sigma}^2\hat{Q}'\hat{Q}\Big)$ is equivalent to the asymptotic distribution of $\sqrt{N}Z_Q'\mathrm{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big)$. Finally, by assumption 6(ii), we have

$$\sqrt{N}Z_Q'\mathrm{vec}\Big(\frac{\epsilon\epsilon'}{N} - \bar{\sigma}^2 I_T\Big) \xrightarrow{d} \mathcal{N}\mathcal{N}\Big(0, Z_Q'U_\epsilon Z_Q\Big).\blacksquare \tag{IA.119}$$

Proof of Theorem 5. For part (i), in view of Equation (64), we obtain $\bar{R} = X\tilde{\Gamma}^P + e + \bar{\epsilon}$, where $\tilde{\Gamma}^P = \tilde{\Gamma} + \bar{f} - E[f_t]$. Using the same arguments as for Theorem 1,

$$\hat{\Gamma}^* - \tilde{\Gamma}^P = \left(\frac{\hat{X}'\hat{X}}{N} - \hat{\Lambda}\right)^{-1}\left[\frac{\hat{X}'\bar{\epsilon}}{N} - \left(\frac{\hat{X}'}{N}\left(\hat{X} - X\right) - \hat{\Lambda}\right)\tilde{\Gamma}^P + \frac{\hat{X}'e}{N}\right] \tag{IA.120}$$

with $\left(\frac{\hat{X}'\hat{X}}{N} - \hat{\Lambda}\right) = O_p(1)$, $\frac{\hat{X}'\bar{\epsilon}}{N} = O_p\left(\frac{1}{\sqrt{N}}\right)$, and $\left(\frac{\hat{X}'}{N}\left(\hat{X} - X\right) - \hat{\Lambda}\right) = O_p\left(\frac{1}{\sqrt{N}}\right)$. As for the term $\frac{\hat{X}'e}{N}$,

$$\begin{aligned}\frac{\hat{X}'e}{N} = \frac{X'e}{N} + \frac{(\hat{X} - X)'e}{N} &= 0_{K+1} + \frac{1}{N}\begin{bmatrix} 0 \\ \mathcal{P}'\epsilon\, e\end{bmatrix} \\ &= 0_{K+1} + O_p\left(\frac{1}{\sqrt{N}}\right)\end{aligned} \tag{IA.121}$$

since $\mathcal{P}'\epsilon\, e = O_p((\mathcal{P}'\sum_{i,j=1}^N \sigma_{ij}e_ie_j\mathcal{P})^{\frac{1}{2}}) = O_p(\sqrt{N})$ by assumption 7(i)-(ii). Next,

$$\begin{aligned}\sqrt{N}\left(\hat{\Gamma}^* - \tilde{\Gamma}^P\right) &= \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left(\begin{bmatrix}\frac{1_N'\epsilon'Q}{\sqrt{N}} \\ \frac{B'\epsilon'Q}{\sqrt{N}}\end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\mathcal{P}'\epsilon\epsilon'Q}{\sqrt{N}} + \frac{\mathrm{tr}(M\epsilon\epsilon')}{\sqrt{N}(T-K-1)}\mathcal{P}'\mathcal{P}\tilde{\gamma}_1^P\end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{P}'\frac{1}{\sqrt{N}}\sum_{i=1}^N \epsilon_i e_i\end{bmatrix}\right) \\ &\equiv \left(\hat{\Sigma}_X - \hat{\Lambda}\right)^{-1}\left(I_1 + I_2 + I_3\right).\end{aligned} \tag{IA.122}$$

As for terms I_1 and I_2, Theorem 1 applies, that is, $\left(\hat{\Sigma}_X - \hat{\Lambda}\right) \xrightarrow{p} \Sigma_X$, $\mathrm{Var}(I_1) = \frac{\sigma^2}{T}\left[1 + \tilde{\gamma}_1^{P\prime}\left(\frac{\tilde{F}'\tilde{F}}{T}\right)^{-1}\tilde{\gamma}_1^P\right]\Sigma_X$ and $\mathrm{Var}(I_2) = \begin{bmatrix} 0 & 0_K' \\ 0_K & E[I_{22}I_{22}']\end{bmatrix}$, with $E[I_{22}I_{22}'] = \left[(Q'\otimes\mathcal{P}') + \mathcal{P}'\mathcal{P}\tilde{\gamma}_1^P\frac{\mathrm{vec}(M)'}{T-K-1}\right]U_\epsilon\left[(Q\otimes\mathcal{P}) + \frac{\mathrm{vec}M}{T-K-1}\tilde{\gamma}_1^{P\prime}\mathcal{P}'\mathcal{P}\right]$, where $\mathrm{Cov}(I_1, I_2) = 0_{(K+1)\times(K+1)}$. Consider now the term I_3 and note that it has a zero mean. Its variance is equal to

$$\mathrm{Var}(I_3) = E\begin{bmatrix} 0 & 0_K' \\ 0_K & \mathcal{P}'\frac{1}{N}\sum_{i,j=1}^N \epsilon_i\epsilon_j'e_ie_j\mathcal{P}\end{bmatrix} \xrightarrow{p} \begin{bmatrix} 0 & 0_K' \\ 0_K & \tau_\Omega\mathcal{P}'\mathcal{P}\end{bmatrix} \equiv \Omega, \tag{IA.123}$$

and the covariance term satisfies

$$\mathrm{Cov}(I_1, I_3) = E\begin{bmatrix}\frac{1_N'\epsilon'Q}{\sqrt{N}} \\ \frac{B'\epsilon'Q}{\sqrt{N}}\end{bmatrix}\left[0, \frac{1}{\sqrt{N}}\sum_{i=1}^N e_i\epsilon_i'\mathcal{P}\right] \xrightarrow{p} \begin{bmatrix} 0 & \tau_\Phi Q'\mathcal{P} \\ 0_K & \tau_\Phi\left(Q'\otimes\mu_\beta\right)\mathcal{P}\end{bmatrix} \equiv \Phi, \tag{IA.124}$$

while $\text{Cov}(I_2, I_3) = 0_{(K+1)\times(K+1)}$ by the assumption of zero third moment of the error term. Using Lemmas 8 and 9, the proof of part (ii) becomes very similar to the proof of Theorem 2 and is omitted. ■

Proof of Theorem 6. For part (i), rewrite

$$\begin{bmatrix} \hat{\Gamma}^* \\ \hat{\delta}^* \end{bmatrix} = \begin{bmatrix} \Gamma^P \\ \delta \end{bmatrix} + \begin{bmatrix} \hat{X}'\hat{X} - \hat{\Lambda} & \hat{X}'C \\ C'\hat{X} & C'C \end{bmatrix}^{-1} \left[\begin{bmatrix} \hat{\Lambda}\Gamma^P \\ 0_{K_c} \end{bmatrix} + \begin{bmatrix} \hat{X}' \\ C' \end{bmatrix} (\bar{\epsilon} + (X - \hat{X})\Gamma^P) \right].$$

As for the bias associated with $\hat{\Gamma}^*$ (see the proof of Theorem 1), we have

$$\hat{\Lambda}\Gamma^P + \frac{1}{N}\hat{X}'(\bar{\epsilon} + (X - \hat{X})\Gamma^P) = O_p(N^{-1/2}). \tag{IA.125}$$

As for the bias associated with $\hat{\delta}^*$, we have

$$\frac{1}{N}C'(\bar{\epsilon} + (X - \hat{X})\Gamma^P) = \frac{1}{N}C'\epsilon'\left(\frac{1_T}{T} - \mathcal{P}\gamma_1^P\right) = \frac{1}{N}C'\epsilon'Q = O_p(N^{-1/2}) \tag{IA.126}$$

since $N^{-1}C'\epsilon' \to_p 0_{K_c \times T}$ and

$$\begin{aligned} \text{Var}\left(\frac{1}{N}C'\epsilon'Q\right) &= (Q' \otimes I_{K_c})\frac{1}{N^2}\sum_{i,j=1}^{N}\Sigma_{zz,ij}(Q \otimes I_{K_c}) = \frac{1}{N^2}(Q' \otimes I_{K_c})\sum_{i,j=1}^{N}\sigma_{ij}(I_T \otimes c_i c_j')(Q \otimes I_{K_c}) \\ &= \frac{1}{N^2}\sum_{i,j=1}^{N}\sigma_{ij}(Q'Qc_i c_j') = \frac{1}{N}\sigma^2(Q'Q\Sigma_{CC}) + o\left(\frac{1}{N}\right) \end{aligned} \tag{IA.127}$$

by assumption 8.

For part (ii), by straightforward generalizations of Lemmas 1 and 2(ii), we have

$$\frac{1}{N}\begin{bmatrix} \hat{X}'\hat{X} - N\hat{\Lambda} & \hat{X}'C \\ C'\hat{X} & C'C \end{bmatrix} \to_p \begin{bmatrix} \Sigma_X & \begin{bmatrix} \mu_C' \\ \Sigma_{CB}' \end{bmatrix} \\ \begin{bmatrix} \mu_C & \Sigma_{CB} \end{bmatrix} & \Sigma_{CC} \end{bmatrix} = L. \tag{IA.128}$$

We now prove that L is positive-definite. Using the blockwise formula for the inverse of a matrix, the invertibility of L follows from Σ_{CC} being positive-definite (see assumption 8(i)) and the invertibility of $\begin{bmatrix} 1 & \mu_\beta' \\ \mu_\beta & \Sigma_\beta \end{bmatrix} - \begin{bmatrix} \mu_C' \\ \Sigma_{CB}' \end{bmatrix} \Sigma_{CC}^{-1} \begin{bmatrix} \mu_C & \Sigma_{CB} \end{bmatrix}$. In turn, this holds if

$$D = \Sigma_\beta - \Sigma_{CB}'\Sigma_{CC}^{-1}\Sigma_{CB} \tag{IA.129}$$

is positive-definite and

$$1 - \mu_C'\Sigma_{CC}^{-1}\mu_C - (\mu_\beta' - \mu_C'\Sigma_{CC}^{-1}\Sigma_{CB})D^{-1}(\mu_\beta - \Sigma_{CB}'\Sigma_{CC}^{-1}\mu_C) \tag{IA.130}$$

is nonzero. The last equation can be rewritten as

$$1 - [\mu_C' \, \mu_\beta'] \begin{bmatrix} \Sigma_{CC} & \Sigma_{CB} \\ \Sigma_{CB}' & \Sigma_\beta \end{bmatrix}^{-1} \begin{bmatrix} \mu_C \\ \mu_\beta \end{bmatrix}. \tag{IA.131}$$

The positiveness of Equation (IA.131) and the positive-definiteness of D follow from assumption 8(i). Next, following the proof of Theorem 1,

$$\begin{aligned} \sqrt{N}\begin{bmatrix} \hat{\Gamma}^* - \Gamma^P \\ \hat{\delta}^* - \delta \end{bmatrix} &= \begin{bmatrix} \frac{\hat{X}'\hat{X}}{N} - \hat{\Lambda} & \frac{\hat{X}'C}{N} \\ \frac{C'\hat{X}}{N} & \frac{C'C}{N} \end{bmatrix}^{-1} \\ &\times \left(\begin{bmatrix} \frac{1_N'\epsilon'}{\sqrt{N}}Q \\ \frac{B'\epsilon'}{\sqrt{N}}Q \\ 0_{K_c} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{P}'\frac{\epsilon\epsilon'}{\sqrt{N}}Q + \frac{\mathrm{tr}(M\epsilon\epsilon')}{\sqrt{N}(T-K-1)}\mathcal{P}'\mathcal{P}\gamma_1^P \\ 0_{K_c} \end{bmatrix} + \begin{bmatrix} 0 \\ 0_K \\ \frac{C'\epsilon'}{\sqrt{N}}Q \end{bmatrix} \right) \\ &\equiv \begin{bmatrix} \frac{\hat{X}'\hat{X}}{N} - \hat{\Lambda} & \frac{\hat{X}'C}{N} \\ \frac{C'\hat{X}}{N} & \frac{C'C}{N} \end{bmatrix}^{-1} (I_1 + I_2 + I_3). \end{aligned} \tag{IA.132}$$

We now derive $\mathrm{Var}(I_3)$ and $\mathrm{Cov}(I_1, I_3')$ because the other terms can be directly obtained from Theorem 1 and $\mathrm{Cov}(I_2, I_3') = 0_{(K+K_c+1)\times(K+K_c+1)}$. We have

$$\mathrm{Var}(I_3) = \begin{bmatrix} 0_{(K+1)\times(K+1)}' & 0_{(K+1)\times K_c}' \\ 0_{K_c\times(K+1)} & \frac{Q'Q}{N}\sum_{i=1}^N \sigma_{ij}(c_i c_j') \end{bmatrix} \to \begin{bmatrix} 0_{(K+1)\times(K+1)}' & 0_{(K+1)\times K_c}' \\ 0_{K_c\times(K+1)} & \sigma^2 Q'Q\Sigma_{CC} \end{bmatrix} \tag{IA.133}$$

and, by Theorem 1,

$$\mathrm{Cov}(I_1, I_3') = \begin{bmatrix} 0_{(K+1)\times(K+1)} & \frac{Q'Q}{N}\sum_{i=1}^N \sigma_{ij}\left(\begin{bmatrix} 1 \\ \beta_i \end{bmatrix} c_j'\right) \\ 0_{K_c\times(K+1)} & 0_{K_c\times K_c} \end{bmatrix} \to \begin{bmatrix} 0_{(K+1)\times(K+1)} & \sigma^2 Q'Q \begin{bmatrix} \mu_C' \\ \Sigma_{CB}' \end{bmatrix} \\ 0_{K_c\times(K+1)} & 0_{K_c\times K_c} \end{bmatrix}. \tag{IA.134}$$

This concludes the proof. ■

IA.2. Random Betas

In this section, we discuss the modifications of the analysis that are necessary to accommodate random betas.[2] First, consider the case where the random betas are mutually independent of the innovations in individual asset returns. In this scenario, the asymptotic properties of the bias-adjusted Shanken (1992) estimator, $\hat{\Gamma}^*$, are not affected. The only changes involve assumptions 1

[2]Dealing with random betas requires a different specification of the sampling scheme. (See, for example, Gagliardini, Ossola, and Scaillet 2016.) In the interest of space, we do not provide the full details of the analysis here.

and 2. In particular, Equation (3) in assumption 1 must be replaced with $E[R_t|X] = X\Gamma$. Moreover, Equations (17) and (18) in assumption 2 must be stated in terms of convergence in probability, instead of conventional convergence, which is applicable to non-random sequences only. All the other assumptions remain unchanged, except that now Equation (38) involves random betas. As $N \to \infty$, we have

$$\lim \operatorname{Var}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(C_T' \otimes [1,\ \beta_i']'\right)\epsilon_i\right) = \lim \frac{1}{N}\sum_{i,j=1}^{N} E\left[\left(C_T' \otimes [1,\ \beta_i']'\right)\epsilon_i\epsilon_j'\left(C_T \otimes [1,\ \beta_j']\right)\right]$$
$$= (C_T'C_T)\lim\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 E\Big[[1,\ \beta_i']'[1,\ \beta_i']\Big] + (C_T'C_T)\lim\frac{1}{N}\sum_{i\neq j=1}^{N}\sigma_{ij}E\Big[\begin{bmatrix}1\\ \beta_i\end{bmatrix}\begin{bmatrix}1 & \beta_j'\end{bmatrix}\Big]$$
$$= (C_T'C_T)\sigma^2\Sigma_X. \tag{IA.135}$$

The second term in the second line of Equation (IA.135) converges to zero under our assumptions since $E\|\beta_i\beta_j'\| \leq E(\beta_i'\beta_i)^{\frac{1}{2}}E(\beta_j'\beta_j)^{\frac{1}{2}} \leq C < \infty$ and $\sum_{i\neq j=1}^{N}|\sigma_{ij}| = o(N)$. Equation (IA.135) coincides with the asymptotic covariance matrix in Equation (38), which holds for non-random β_i.

Consider now the case in which the β_i are potentially cross-sectionally correlated with the ϵ_i. When T is fixed, such covariance structure cannot be identified based on the OLS estimators $\hat{\beta}_i$ and $\hat{\epsilon}_i$ (either for a finite or an arbitrarily large N). Therefore, the possibility of cross-correlation between the β_i and the ϵ_i needs to be ruled out. By inspection of the proof of Theorem 1, the asymptotic covariance of $\sqrt{N}(\hat{\Gamma}^* - \Gamma^P)$ depends on, among other things, $N^{-\frac{1}{2}}\sum_{i=1}^{N}\beta_i\epsilon_i'Q$, where $Q = \frac{1_T}{T} - \mathcal{P}\gamma_1^P$. Letting the K-vector $\omega_i = E[\beta_i\epsilon_{it}] = \operatorname{Cov}(\beta_i, \epsilon_{it})$, we have

$$E[\beta_i\epsilon_i'] = \Omega_i = \omega_i 1_T', \tag{IA.136}$$

where the second equality follows from the i.i.d. assumption over time for the ϵ_{it} (see assumption 3). Then,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\beta_i\epsilon_i'Q = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\beta_i\epsilon_i' - \Omega_i)Q + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Omega_i Q. \tag{IA.137}$$

A straightforward generalization of assumption 6 (iii) implies that the first term of Equation (IA.137) converges to a normal distribution as $N \to \infty$. Given Equation (IA.136), the second term of Equation (IA.137) can be re-written as $\sqrt{N}^{-1}\sum_{i=1}^{N}\Omega_i Q = \sqrt{N}^{-1}\sum_{i=1}^{N}\omega_i 1_T'Q = \sqrt{N}^{-1}\sum_{i=1}^{N}\omega_i$ because $1_T'Q = 1$. As we show below, this latter term cannot be consistently estimated by OLS when T is fixed. Therefore, in order to avoid lack of identification in the asymptotic covariance of $\hat{\Gamma}^*$, the ω_i must satisfy the restriction $\sqrt{N}^{-1}\sum_{i=1}^{N}\omega_i = o(1)$, which contains $\omega_i = 0_K$ as a special case.

We now illustrate how this restriction is needed when considering the OLS estimator of the second term of Equation (IA.137). Starting with a fixed N, the OLS estimator of $\sqrt{N}^{-1}\sum_{i=1}^{N}\Omega_i Q = \sqrt{N}^{-1}\sum_{i=1}^{N}E[\beta_i\epsilon_i']Q$ is $\sqrt{N}^{-1}\sum_{i=1}^{N}\hat{\Omega}_i\hat{Q}$ with $\hat{\Omega}_i\hat{Q} = \hat{\beta}_i\hat{\epsilon}_i'\hat{Q}$. Since $\hat{\epsilon}_i$ and $\hat{Q}$ are orthogonal for any finite T and N, the estimated term $\sqrt{N}^{-1}\sum_{i=1}^{N}\hat{\Omega}_i\hat{Q}$ is a zero vector and $\sqrt{N}^{-1}\sum_{i=1}^{N}\Omega_i Q$ cannot be identified. Next, when N diverges, even without post-multiplying by $\hat{Q}$, it can be shown that $N^{-1}\sum_{i=1}^{N}\hat{\Omega}_i \to_p 0_{K\times T}$, and once again $\omega = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\omega_i$ cannot be identified.[3]

Therefore, under our fixed-T sampling scheme, the assumption $\omega_i = \mathrm{Cov}(\beta_i, \epsilon_{it}) = 0_K$ or, alternatively, the slightly more general assumption $\sqrt{N}^{-1}\sum_{i=1}^{N}\omega_i = o(1)$ is needed for identification purposes.

IA.3. Nonparametric Estimation of Risk Premia on Traded Factors

Under Equation (48) and assuming that the factors are traded, it is well-known that the time-varying risk premia are given by

$$\gamma_{1,t} = E_t[f_{t+1}] - 1_K\gamma_{0,t}. \tag{IA.138}$$

Moreover, when a risk-free asset with one-period return $r_{f,t}$ is available for investment and we assume that the zero-beta rate is equal to the risk-free rate, the latter expression simplifies to $\gamma_{1,t} = E_t[f_{t+1}] - 1_K r_{f,t}$, that is, the risk premia coincide with the conditional expected excess factor returns. This suggests that any estimator of the conditional mean can be used for risk premia estimation. A popular estimator of $E_t[f_{t+1}]$ is the sample mean of T consecutive observations, that is,

$$\bar{f}_t = \frac{1}{T}\sum_{h=-T/2}^{T/2-1} f_{t+1+h} \qquad \text{or} \qquad \tilde{f}_t = \frac{1}{T}\sum_{h=1}^{T} f_{t+1-h}. \tag{IA.139}$$

Typically, the estimators in Equation (IA.139) are evaluated over consecutive rolling samples. In the absence of time variation, the risk premia are given by $\gamma_1 = E[f_{t+1}] - 1_K r_f$, and the unconditional mean $E[f_{t+1}]$ is consistently estimated (as T diverges) by the sample mean of f_{t+1} over the full sample.

We now summarize the statistical properties of the risk premia estimators in Equation (IA.139) as T diverges. It is convenient to simplify the exposition by setting $K = 1$ and assuming that the

[3] In particular, recalling that $M = I_T - D(D'D)^{-1}D'$ with $D = [1_T,\ F]$, under our assumptions $N^{-1}\sum_{i=1}^{N}\hat{\Omega}_i \to_p (\omega 1_T' + \sigma^2\mathcal{P}')M = 0_{K\times T}$ because M is orthogonal to both 1_T and $\mathcal{P}$.

risk-free rate is constant over time. In addition, assume that the realized factor return, f_{t+1}, can be written as

$$f_{t+1} - r_f = \gamma_{1,t} + u_{t+1} \text{ for some i.i.d. error } u_{t+1} \sim (0, \sigma^2). \tag{IA.140}$$

Finally, assume that the T observations used to compute the estimators above represent a subset of a possibly larger number of observations, $T_0 \geq T$. Then, we consider two alternative sampling schemes. First, we evaluate the estimators' behavior under the conventional scheme $T = T_0$, that is, using all the available data. Next, we consider a scheme where, even though T diverges with T_0, $T/T_0 \to 0$. The latter is the sampling scheme adopted in nonparametric kernel estimation, and it implies that the estimators $\bar{f}_t$ and $\tilde{f}_t$ are evaluated over a shrinking time interval (around period t) of relative length T/T_0 as T_0 diverges. Typically, samples of size T are rolled over the entire length T_0.

Theorem IA.1 Under Equation (IA.140),

(i) When $T = T_0$, $\bar{f}_t - r_f$ is an unbiased estimator of $\bar{\gamma}_{1,t} = T_0^{-1} \sum_{h=-T_0/2}^{T_0/2-1} \gamma_{1,t+h}$ for any T_0. Moreover, when $T_0 \to \infty$,

$$\bar{f}_t - r_f - \bar{\gamma}_{1,t} \quad = \quad O_p(T_0^{-\frac{1}{2}}), \tag{IA.141}$$

$$T_0^{\frac{1}{2}} (\bar{f}_t - r_f - \bar{\gamma}_{1,t}) \quad \to_d \quad \mathcal{N}(0, \sigma^2). \tag{IA.142}$$

Finally, σ^2 can be consistently estimated by means of $T_0^{-1} \sum_{h=-T_0/2}^{T_0/2-1} (f_{t+1+h} - \bar{f}_{t+h})^2$ under the smoothness assumption $T_0^{-1} \sum_{s=-T_0/2}^{T_0/2-1} (\gamma_{1,t+s} - \bar{\gamma}_{1,t+s})^2 = o_p(1)$.

The same properties apply to $\tilde{f}_t$ with respect to $\tilde{\gamma}_{1,t} = T_0^{-1} \sum_{h=1}^{T_0} \gamma_{1,t-h}$.

(ii) When $T < T_0$, the estimators in Equation (IA.139) are special cases of the kernel regression estimator

$$\hat{\gamma}_{1,t}^{\kappa} = \sum_{h=1}^{T_0} f_h w_h^{\kappa} - r_f, \text{with } w_h^{\kappa} = \frac{\kappa\left(\frac{(h-t)}{T-1}\right)}{\sum_{s=1}^{T_0} \kappa\left(\frac{(s-t)}{T-1}\right)}, \tag{IA.143}$$

where $\kappa(u) = \mathbb{1}_{\{\frac{1}{T-1} - \frac{0.5T}{T-1} \leq u \leq \frac{0.5T}{T-1}\}}$ and $\kappa(u) = \mathbb{1}_{\{-1 \leq u \leq 0\}}$ for $\bar{f}_t$ and $\tilde{f}_t$, respectively, and $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function.

When a) $T^{-1} + T_0^{1/2}(T/T_0)^{1.5} \to 0$; b) the kernel $\kappa(\cdot)$ integrates to unity and is differentiable, and both $\kappa(\cdot)$ and $\kappa'(\cdot)$ go to zero faster than $O((1+u^2)^{-1})$ for u large; and c) the true risk premium satisfies $\gamma_{1,t} = \gamma_1(\frac{T}{T_0})$ for a differentiable function $\gamma_1(\cdot)$, then

$$\hat{\gamma}_{1,t}^{\kappa} - \gamma_{1,t} = O_p(T^{-\frac{1}{2}}), \quad \text{(IA.144)}$$

$$T^{\frac{1}{2}}(\hat{\gamma}_{1,t}^{\kappa} - \gamma_{1,t}) \to_d \mathcal{N}\Big(0, \sigma^2 \int_{-\infty}^{\infty} \kappa^2(u) du\Big). \quad \text{(IA.145)}$$

Finally, σ^2 can be consistently estimated by means of $T_0^{-1}\sum_{h=1}^{T_0}(f_{h+1} - r_f - \hat{\gamma}_{1,h}^{\kappa})^2$ under the assumptions above.

The proof is available upon request. Part (i) follows from noting that

$$\bar{f}_t - r_f = \bar{\gamma}_{1,t} + T_0^{-1} \sum_{h=-T_0/2}^{T_0/2-1} u_{t+1+h}, \quad \text{(IA.146)}$$

and exploiting the properties of the sample mean of i.i.d. random variables. The proof of part (ii) follows from Robinson (1997), where the more general framework with non-i.i.d. innovations u_{t+1} is considered.

Part (i) of Theorem IA.1 indicates that the rolling estimators accurately estimate the average risk premia over a given time interval, but they will not converge to the true risk premia at some specific time t. In particular, note that the rolling estimators converge to the *integrated* risk premium $\gamma_1 = \lim_{T_0 \to \infty} \bar{\gamma}_{1,t} = \lim_{T_0 \to \infty} \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} \gamma_{1,s} ds$ (assuming that γ_1 is bounded). For inference, a smoothness condition that limits the degree of time variation in the true risk premia is required.

Part (ii) shows how the traditional rolling sample mean estimators in Equation (IA.139) can be obtained as special cases of the nonparametric Nadaraya-Watson estimator by suitably choosing the kernel function.[4] Although inference can be conducted for the t-th risk premium, the rate of convergence is slower than the usual "square-root" speed. For example, condition $T^{-1} + T_0^{1/2}(T/T_0)^{1.5} \to 0$ is satisfied when T is not larger than $T_0^{2/3}$. It follows that the rolling sample mean estimators will converge at rate $O(T^{-1/3})$. For instance, when $T_0 = 100$, only about $T = 20$ observations are available for inference on the time-varying risk premia. This small T implies that the confidence interval for the time-varying risk premium will be quite large. Therefore, not only T must diverge for the asymptotic theory to be valid, but T needs to be larger than what is required by the usual parametric rate.[5] Notice how the results in parts (i) and (ii) require additional smoothness assump-

[4] For the estimator $\bar{f}_t$ we provide the appropriate kernel function but a simpler, yet asymptotically equivalent, expression can be obtained by noting that $\mathbb{1}_{\frac{1}{T-1} - \frac{0.5T}{T-1} \le u \le \frac{0.5T}{T-1}} \sim \mathbb{1}_{|u| \le 0.5}$ for T large.

[5] As emphasized by Robinson (1997), the above results can be extended to non-i.i.d. errors by replacing σ^2 with $2\pi f_u(0)$, the spectral density of u_{t+1} at frequency zero (multiplied by 2π), and assuming boundedness of $f_u(\cdot)$.

tions on the form and degree of time variation in the true risk premia. Using high-frequency data, Ang and Kristensen (2012) rely on similar nonparametric techniques to develop tests of conditional beta-pricing models.

IA.4. Monte Carlo Simulations

In this section, we undertake a Monte Carlo simulation experiment to study the empirical rejection rates of the specification test and t-ratios of the bias-adjusted estimator of Shanken (1992). The return-generating process under the null of a correctly specified asset-pricing model is given by

$$R_t = \gamma_0 1_N + B\left(\gamma_1 + f_t - E[f_t]\right) + \epsilon_t, \tag{IA.147}$$

where $\epsilon_t \sim \mathcal{N}(0,\ \Sigma)$. To study the power of the specification test, we generate the returns on the test assets as in Equation (2), that is, we do not impose the asset-pricing restriction.

In all of our simulation experiments, we consider balanced panels with a time-series dimension of $T = 36$ and $T = 72$ observations. Specifically, f_t in Equation (IA.147) is the excess market return (from Kenneth French's website) from January 2008 to December 2010 for $T = 36$, and the excess market return from January 2008 to December 2013 for $T = 72$. In our simulation designs, the factor realizations are taken as given and kept fixed throughout. This is consistent with the fact that our analysis of the ex post risk premia is conditional on the realizations of the factors. In addition, $E[f_t]$ in Equation (IA.147) is set equal to the time-series mean of f_t over the 2008–2010 sample when performing the analysis for $T = 36$ and to the time-series mean of f_t over the 2008–2013 sample when performing the analysis for $T = 72$. To obtain representative values for the parameters γ_0, γ_1, B, and Σ in Equation (IA.147) and Equation (2), we employ a cross-section of 3,000 stocks from CRSP in addition to the excess market return. Based on this balanced panel of 3,000 stock returns and the excess market return, for each time-series sample size, we compute the OLS estimates of B, γ_0, and γ_1. Then, we set the B, γ_0, and γ_1 parameters in Equation (IA.147) and in Equation (2) equal to these OLS estimates. The calibration of Σ is a more delicate task and is described below. In the simulations, we consider cross-sections of $N = 100$, 500, 1,000, and 3,000 stocks. All results are based on 10,000 Monte Carlo replications. Our econometric approach, designed for large N and fixed T, should be able to handle this large number of assets over relative short time spans. The rejection rates of the various tests are computed using our asymptotic results in the paper.

IA.4.1 Percentage errors and root mean squared errors of the estimates

We start from the case in which Σ is a spherical matrix, that is, $\Sigma = \sigma^2 I_T$. In the simulations, we set σ^2 equal to the cross-sectional average (over the 3,000 stocks) of the σ_i^2 estimated from the data. Table IA.1 reports the percentage error (Bias) and root mean squared error (RMSE), all in percent, of the OLS estimator and of the bias-adjusted estimator of Shanken (1992). Panels A and B are for $T = 36$ and $T = 72$, respectively.

Table IA.1 about here

Panel A shows that the bias of the OLS estimator is substantial. For $\hat{\gamma}_0$, the bias ranges from 28.8% for $N = 100$ to 22.9% for N=3,000, while for $\hat{\gamma}_1$ the bias ranges from −24.8% for $N = 100$ to −17.8% for N=3,000. For $\hat{\Gamma}^*$, the bias is small for $N = 100$ (−2.3% for $\hat{\gamma}_0^*$ and 1.8% for $\hat{\gamma}_1^*$) and becomes negligible for $N \geq 500$. As for the RMSE, the typical bias-variance trade-off emerges up to $N = 500$, with the OLS estimator exhibiting a smaller RMSE than the OLS bias-adjusted estimator. When $N > 500$, the RMSE of the bias-adjusted estimator of Shanken (1992) becomes substantially smaller than the one of the OLS estimator. Panel B for $T = 72$ conveys a similar message. As expected from the theoretical analysis, the larger time-series dimension helps in reducing the bias and RMSE associated with the OLS estimator. However, the bias for the OLS estimator is still substantial and ranges from −18.5% for $N = 100$ to −11.7% for N=3,000. For the bias-adjusted estimator, the bias becomes negligible, even for $N = 100$ when $T = 72$.

Next, we consider the case in which the Σ matrix is either diagonal or full. As emphasized above, our theoretical results hinge upon the assumption that the model disturbances are weakly cross-sectionally correlated. In order to generate shocks under a weak factor structure, we consider the following data-generating process (DGP). Define

$$\epsilon^{(1)} = \eta \left(\frac{\sqrt{\theta}}{N^{\delta}} \right) c' + \sqrt{1-\theta} Z, \qquad \text{(IA.148)}$$

where η and c are T and N-vectors of i.i.d. standard normal random variables, respectively, Z is a $T \times N$ matrix of i.i.d. standard normal random variables, $0 \leq \theta \leq 1$ is a shrinkage parameter that controls the weight assigned to the diagonal and extra-diagonal elements of Σ, and δ is a parameter that controls the strength of the cross-sectional dependence of the shocks (the bigger δ is, the weaker the dependence). Our $T \times N$ matrix of shocks is then generated as

$$\epsilon = \epsilon^{(1)} \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_N^2 \end{bmatrix}^{0.5} \begin{bmatrix} \frac{\theta}{N^{2\delta}} c_1^2 + (1-\theta) & & & \\ & \frac{\theta}{N^{2\delta}} c_2^2 + (1-\theta) & & \\ & & \ddots & \\ & & & \frac{\theta}{N^{2\delta}} c_N^2 + (1-\theta) \end{bmatrix}^{-0.5}, \tag{IA.149}$$

where c_i is the i-th element of c. Given this specification for the shocks, for our theoretical results to hold, we require $\delta > 0$.

As discussed in the paper, the factor structure in Equations (IA.148)–(IA.149) induces a substantial degree of cross-correlation between the ϵ_{it}. We demonstrate this by means of a simple numerical example. For each Monte Carlo replication, we compute the following quantity for the data generating process above:

$$A(\delta, N) \equiv \frac{1}{N} \sum_{j=1}^{N} \left(\frac{\sum_{\substack{i=1 \\ i \neq j}}^{N} |\sigma_{ij}|}{\sigma_j^2 + \sum_{\substack{i=1 \\ i \neq j}}^{N} |\sigma_{ij}|} \right). \tag{IA.150}$$

For the case of zero cross-correlation, that is, $\sigma_{ij} = 0$ when $i \neq j$, $A(\delta, N) = 0$. In contrast, when the cross-correlations become big, $A(\delta, N)$ approaches 1. As we vary δ and N (average across 1,000 Monte Carlo iterations), we obtain

$A(\delta, N)$	value
A(0.25, 100)	0.8585
A 0.25, 500)	0.9268
A(0.25, 3,000)	0.9628
A(0.255, 100)	0.8560
A(0.255, 500)	0.9235
A(0.255, 3,000)	0.9606
A(0.5, 100)	0.5418
A(0.5, 500)	0.5680
A(0.5, 3,000)	0.5612

This simple numerical example shows that for every value of δ (which measures the degree of dependence), the sum of the cross-covariances tends to increase with N in relative terms.

In Table IA.2, we report results for the diagonal case, that is, we set $\theta = 0$ in the above data-generating process. To obtain representative values of the shock variances, while accounting for the fact that $\hat{\Sigma}$ is ill conditioned when T is small and N is large, we first estimate the residual variances from the historical data. Then, at each Monte Carlo iteration, we generate a string of Beta(p, q)-distributed random variables with the p and q parameters calibrated to the cross-sectional mean and variance of the $\hat{\sigma}_i^2$. This resampling procedure is used to minimize the impact of an ill-conditioned $\hat{\Sigma}$ on the simulation results.

Table IA.2 about here

Overall, we find that the OLS estimator exhibits a slightly higher bias compared to the spherical Σ case. The bias-adjusted estimator of Shanken (1992) continues to perform very well in terms of bias for all the time-series and cross-sectional dimensions considered. The RMSEs of both estimators are now a bit higher than in the spherical case, and the bias-adjusted estimator still outperforms the OLS estimator for $N \geq 500$.

Finally, in Tables IA.3 and IA.4, we allow for weak cross-sectional dependence of the model disturbances by setting $\theta = 0.5$ in the above DGP.

Tables IA.3 and IA.4 about here

In Table IA.3, we consider the situation in which δ, the parameter that regulates the strength of the cross-sectional dependence, is equal to 0.5. Consistent with our theoretical results, the bias-adjusted estimator continues to perform very well in this scenario. Setting $\delta = 0.25$ in Table IA.4 has only a modest effect on the bias and RMSEs of the two estimators. Overall, the first four tables reveal a superiority of the bias-adjusted estimator of Shanken (1992) over the OLS estimator, not only in terms of bias, but also in terms of RMSE when $N > 500$. Furthermore, the bias-adjusted estimator shows little sensitivity to changes in the length of the time-series, consistent with the idea that this estimator should perform well for any given T.

IA.4.2 Rejection rates of the t-tests

In Tables IA.5 through IA.8, we consider the empirical rejection rates of the centered t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of time-series and cross-sectional observations using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with covariance matrix calibrated as in Tables IA.1 through IA.4. The t-statistics are compared with the critical values from a standard normal distribution. We consider three t-statistics. For the OLS estimator of the ex post risk premia, the first t-statistic is the one that uses the traditional Fama and MacBeth (1973) standard error (t_{FM}), while the second t-statistic (t_{EIV}) is the one that uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992). Both of these t-statistics were developed in a large-T and fixed-N framework. We report them here to determine how misleading inference can be when using these t-statistics in a large-N and fixed-T setup. Finally, the third t-statistic is the one associated with the bias-adjusted estimator of Shanken (1992) and is based on the asymptotic distribution in part (ii) of our Theorem 1.

Table IA.5 about here

Starting from the spherical Σ case, Table IA.5 shows that the t-statistics associated with the OLS estimator only slightly overreject the null hypothesis for $N = 100$. However, as N increases, the performance of these t-statistics substantially deteriorates. For example, when N=3,000, the rejections rate of the Fama and MacBeth (1973) t-statistic associated with $\hat{\gamma}_1$ is either 41.6% for $T = 36$ or 33.3% for $T = 72$ at the 5% nominal level. The strong size distortions of the Fama and MacBeth (1973) t-test do not show any improvement when accounting for the EIV bias, due to the estimation of the betas in the first stage. In contrast, our proposed t-statistic, based on Theorems 1 and 2, performs extremely well for all T and N. A similar picture emerges in the Σ full case (Tables IA.6 and IA.7), with the rejection rates of our proposed t-test being always aligned with the critical values from a standard normal distribution.

Tables IA.6 and IA.7 about here

In Table IA.8, we increase the strength of the cross-sectional dependence of the residuals by setting $\delta = 0.25$.

Table IA.8 about here

In this situation, we start to notice some slight over-rejections for the t-test associated with the bias-adjusted estimator of Shanken (1992). For example, when T = 36 and N=3,000, the rejection rate for the t-test associated with $\hat{\gamma}_1^*$ is 6.8% at the 5% level, and when T = 72 and N=3,000, the rejection rate for the t-test associated with $\hat{\gamma}_1^*$ is 5.8% at the 5% level. Overall, these results suggest that our proposed t-test is relatively well behaved even when moving toward a fairly strong factor structure in the residuals. Furthermore, using the standard tools that were developed in a large-T and fixed-N framework can lead to strong over-rejections of the null hypothesis, with the likely consequence that a factor will be found to be priced even when it does not help in explaining the cross-sectional variation in individual stock returns.

IA.4.3 Rejection rates of the specification test

In Tables IA.9 and IA.10, we investigate the size and power properties of our specification test $\mathcal{S}^*$ based on the results in Theorem 4. Table IA.9 refers to T = 36, while Table IA.10 is for T = 72.

Tables IA.9 and IA.10 about here

Since our test statistic $\mathcal{S}^*$ has a standard normal distribution, we consider two-sided p-values in the computation of the rejection rates. The results in the two tables suggest that the rejection rates of our test under the null that the model is correctly specified are excellent for the spherical and diagonal cases. When simulating with Σ full, the specification test is very well sized when δ = 0.5 but it over-rejects a bit too much when δ = 0.25. The power properties of our specification test are fairly good when N = 100 and excellent when $N \geq 500$. As expected, power increases when the number of assets becomes large and the rejection rates are similar across time-series sample sizes. Overall, these simulation results suggest that our test $\mathcal{S}^*$ should be fairly reliable for the time-series and cross-sectional dimensions encountered in our empirical work.

Table IA.1
Bias and root mean squared error of the OLS and bias-adjusted Shanken (1992) estimators in a one-factor model ($\boldsymbol{\Sigma}$ spherical)

Statistics	N = 100	N = 500	N=1,000	N=3,000
		Panel A: T = 36		
Bias($\hat{\gamma}_0$)	28.8%	26.2%	24.6%	22.9%
Bias($\hat{\gamma}_0^*$)	–2.3%	–0.3%	0.3%	–0.2%
RMSE($\hat{\gamma}_0$)	0.3675	0.1875	0.1427	0.1066
RMSE($\hat{\gamma}_0^*$)	0.4509	0.1892	0.1255	0.0699
Bias($\hat{\gamma}_1$)	–24.8%	–20.0%	–18.8%	–17.8%
Bias($\hat{\gamma}_1^*$)	1.8%	0.1%	–0.2%	0.2%
RMSE($\hat{\gamma}_1$)	0.3539	0.1642	0.1277	0.1,000
RMSE($\hat{\gamma}_1^*$)	0.4529	0.1655	0.1098	0.0609
		Panel B: T = 72		
Bias($\hat{\gamma}_0$)	11.6%	9.8%	8.7%	7.9%
Bias($\hat{\gamma}_0^*$)	–0.8%	–0.0%	–0.0%	–0.1%
RMSE($\hat{\gamma}_0$)	0.2504	0.1198	0.0877	0.0628
RMSE($\hat{\gamma}_0^*$)	0.2881	0.1165	0.0766	0.0426
Bias($\hat{\gamma}_1$)	–18.5%	–14.1%	–12.4%	–11.7%
Bias($\hat{\gamma}_1^*$)	1.0%	–0.0%	0.2%	0.1%
RMSE($\hat{\gamma}_1$)	0.2437	0.1063	0.0787	0.0597
RMSE($\hat{\gamma}_1^*$)	0.2868	0.1026	0.0674	0.0379

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets, for the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$ and the bias-adjusted estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Table IA.2
Bias and root mean squared error of the OLS and bias-adjusted Shanken (1992) estimators in a one-factor model ($\boldsymbol{\Sigma}$ diagonal)

Statistics	N = 100	N = 500	N=1,000	N=3,000
		Panel A: T = 36		
Bias($\hat{\gamma}_0$)	30.1%	25.8%	24.8%	23.0%
Bias($\hat{\gamma}_0^*$)	–0.7%	–0.8%	0.4%	–0.1%
RMSE($\hat{\gamma}_0$)	0.4047	0.1976	0.1495	0.1100
RMSE($\hat{\gamma}_0^*$)	0.5027	0.2054	0.1364	0.0763
Bias($\hat{\gamma}_1$)	–25.5%	–19.6%	–18.7%	–17.9%
Bias($\hat{\gamma}_1^*$)	0.9%	0.6%	–0.1%	0.1%
RMSE($\hat{\gamma}_1$)	0.3949	0.1733	0.1339	0.1033
RMSE($\hat{\gamma}_1^*$)	0.5104	0.1815	0.1208	0.0681
		Panel B: T = 72		
Bias($\hat{\gamma}_0$)	11.2%	10.0%	8.6%	8.0%
Bias($\hat{\gamma}_0^*$)	–1.2%	0.2%	–0.1%	0.0%
RMSE($\hat{\gamma}_0$)	0.2673	0.1246	0.0899	0.0643
RMSE($\hat{\gamma}_0^*$)	0.3116	0.1223	0.0804	0.0446
Bias($\hat{\gamma}_1$)	–18.1%	–14.3%	–12.3%	–11.8%
Bias($\hat{\gamma}_1^*$)	1.5%	–0.3%	0.3%	–0.0%
RMSE($\hat{\gamma}_1$)	0.2621	0.1112	0.0809	0.0612
RMSE($\hat{\gamma}_1^*$)	0.3120	0.1087	0.0711	0.0400

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets, for the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$ and the bias-adjusted estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Table IA.3
Bias and root mean squared error of the OLS and bias-adjusted Shanken (1992) estimators in a one-factor model ($\boldsymbol{\Sigma}$ full, $\boldsymbol{\delta} = 0.5$)

Statistics	$N = 100$	$N = 500$	N=1,000	N=3,000
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	28.8%	26.0%	24.6%	22.7%
Bias($\hat{\gamma}_0^*$)	−2.6%	−0.6%	0.3%	−0.4%
RMSE($\hat{\gamma}_0$)	0.4065	0.1960	0.1506	0.1089
RMSE($\hat{\gamma}_0^*$)	0.5081	0.2031	0.1385	0.0760
Bias($\hat{\gamma}_1$)	−24.2%	−19.6%	−18.9%	−17.7%
Bias($\hat{\gamma}_1^*$)	2.7%	0.7%	−0.3%	0.3%
RMSE($\hat{\gamma}_1$)	0.3963	0.1727	0.1352	0.1028
RMSE($\hat{\gamma}_1^*$)	0.5159	0.1806	0.1220	0.0681
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	11.8%	9.4%	8.6%	8.0%
Bias($\hat{\gamma}_0^*$)	−0.5%	−0.5%	−0.1%	−0.0%
RMSE($\hat{\gamma}_0$)	0.2671	0.1227	0.0910	0.0642
RMSE($\hat{\gamma}_0^*$)	0.3099	0.1225	0.0820	0.0447
Bias($\hat{\gamma}_1$)	−19.0%	−13.6%	−12.4%	−11.7%
Bias($\hat{\gamma}_1^*$)	0.5%	0.6%	0.1%	0.1%
RMSE($\hat{\gamma}_1$)	0.2614	0.1104	0.0819	0.0611
RMSE($\hat{\gamma}_1^*$)	0.3096	0.1100	0.0720	0.0405

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets, for the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$ and the bias-adjusted estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Table IA.4
Bias and root mean squared error of the OLS and bias-adjusted Shanken (1992) estimators in a one-factor model ($\boldsymbol{\Sigma}$ full, $\boldsymbol{\delta} = 0.25$)

Statistics	$N = 100$	$N = 500$	N=1,000	N=3,000
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	28.8%	26.6%	24.2%	23.5%
Bias($\hat{\gamma}_0^*$)	−2.5%	0.1%	−0.3%	0.5%
RMSE($\hat{\gamma}_0$)	0.4191	0.2053	0.1536	0.1135
RMSE($\hat{\gamma}_0^*$)	0.5254	0.2152	0.1450	0.0809
Bias($\hat{\gamma}_1$)	−24.8%	−19.9%	−18.5%	−18.3%
Bias($\hat{\gamma}_1^*$)	2.0%	0.2%	0.2%	−0.4%
RMSE($\hat{\gamma}_1$)	0.4116	0.1824	0.1380	0.1072
RMSE($\hat{\gamma}_1^*$)	0.5355	0.1935	0.1288	0.0731
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	12.2%	9.7%	8.8%	7.9%
Bias($\hat{\gamma}_0^*$)	−0.1%	−0.2%	0.1%	−0.1%
RMSE($\hat{\gamma}_0$)	0.2795	0.1287	0.0939	0.0645
RMSE($\hat{\gamma}_0^*$)	0.3252	0.1292	0.0853	0.0459
Bias($\hat{\gamma}_1$)	−19.3%	−13.9%	−12.6%	−11.7%
Bias($\hat{\gamma}_1^*$)	0.0%	0.2%	−0.1%	0.2%
RMSE($\hat{\gamma}_1$)	0.2761	0.1155	0.0854	0.0615
RMSE($\hat{\gamma}_1^*$)	0.3279	0.1158	0.0763	0.0416

The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets, for the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$ and the bias-adjusted estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$. The model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12.

Table IA.5
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ spherical)

Panel A: T = 36

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.128	0.074	0.021	0.141	0.078	0.022
500	0.186	0.113	0.040	0.213	0.132	0.047
1,000	0.243	0.156	0.059	0.290	0.197	0.075
3,000	0.438	0.324	0.153	0.538	0.416	0.219
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.127	0.073	0.020	0.140	0.077	0.022
500	0.185	0.113	0.039	0.211	0.132	0.047
1,000	0.243	0.156	0.059	0.289	0.197	0.075
3,000	0.437	0.323	0.152	0.537	0.415	0.218
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.097	0.051	0.010	0.100	0.048	0.010
500	0.105	0.053	0.011	0.107	0.055	0.012
1,000	0.103	0.052	0.010	0.105	0.054	0.011
3,000	0.098	0.051	0.011	0.100	0.049	0.010

Table IA.5 (Continued)
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ spherical)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.123	0.063	0.016	0.124	0.066	0.016
500	0.167	0.099	0.030	0.181	0.109	0.033
1,000	0.211	0.133	0.041	0.237	0.154	0.053
3,000	0.378	0.263	0.109	0.449	0.333	0.150
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.122	0.063	0.015	0.123	0.065	0.016
500	0.166	0.099	0.030	0.181	0.108	0.033
1,000	0.210	0.132	0.040	0.236	0.153	0.052
3,000	0.377	0.261	0.108	0.448	0.331	0.149
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.096	0.047	0.009	0.100	0.048	0.009
500	0.097	0.049	0.010	0.098	0.049	0.010
1,000	0.100	0.047	0.009	0.103	0.048	0.009
3,000	0.103	0.054	0.010	0.106	0.054	0.010

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$, which uses the traditional Fama and MacBeth (1973) standard error, $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$, which uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992) and $t(\cdot)$ denotes the t-statistic associated with the Shanken estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$, which uses the standard error formulae of Theorem 2. Finally, the rejection rates for the t-test associated with the bias-adjusted estimator of Shanken (1992) are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Table IA.6
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ diagonal)

Panel A: T = 36

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.122	0.066	0.019	0.125	0.072	0.018
500	0.163	0.104	0.033	0.179	0.112	0.036
1,000	0.226	0.141	0.050	0.248	0.166	0.060
3,000	0.398	0.292	0.128	0.474	0.362	0.174
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.120	0.065	0.018	0.124	0.070	0.017
500	0.163	0.103	0.033	0.179	0.111	0.036
1,000	0.225	0.141	0.050	0.247	0.165	0.060
3,000	0.397	0.291	0.127	0.473	0.362	0.173
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.093	0.045	0.011	0.091	0.044	0.010
500	0.102	0.051	0.010	0.096	0.049	0.011
1,000	0.099	0.048	0.009	0.101	0.051	0.009
3,000	0.099	0.053	0.012	0.099	0.051	0.010

Table IA.6 (Continued)
Size of t-tests in a one-Factor model (Σ diagonal)

Panel B: T = 72

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.115	0.060	0.015	0.121	0.064	0.015
500	0.157	0.089	0.027	0.165	0.096	0.030
1,000	0.199	0.121	0.036	0.219	0.137	0.044
3,000	0.353	0.250	0.103	0.416	0.302	0.134
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.114	0.059	0.014	0.119	0.063	0.015
500	0.157	0.089	0.027	0.163	0.096	0.029
1,000	0.198	0.120	0.036	0.218	0.136	0.044
3,000	0.351	0.248	0.102	0.414	0.301	0.132
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.097	0.048	0.010	0.096	0.048	0.007
500	0.095	0.046	0.010	0.093	0.047	0.010
1,000	0.097	0.049	0.011	0.095	0.049	0.010
3,000	0.103	0.052	0.010	0.102	0.051	0.010

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$, which uses the traditional Fama and MacBeth (1973) standard error, $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$, which uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992) and $t(\cdot)$ denotes the t-statistic associated with the Shanken estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$, which uses the standard error formulae of Theorem 2. Finally, the rejection rates for the t-test associated with the bias-adjusted estimator of Shanken (1992) are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Table IA.7
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ full, $\boldsymbol{\delta}$ = 0.5)

Panel A: T = 36

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.126	0.069	0.020	0.125	0.070	0.021
500	0.166	0.097	0.030	0.181	0.109	0.034
1,000	0.227	0.143	0.049	0.258	0.170	0.063
3,000	0.393	0.282	0.123	0.472	0.354	0.168
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.124	0.068	0.019	0.123	0.068	0.021
500	0.166	0.096	0.030	0.180	0.109	0.034
1,000	0.227	0.142	0.049	0.257	0.170	0.063
3,000	0.392	0.281	0.122	0.470	0.353	0.167
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.097	0.045	0.012	0.094	0.046	0.011
500	0.094	0.045	0.009	0.095	0.045	0.010
1,000	0.106	0.051	0.011	0.102	0.050	0.010
3,000	0.100	0.051	0.011	0.100	0.053	0.011

Table IA.7 (Continued)
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ full, $\boldsymbol{\delta} = 0.5$)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.113	0.062	0.014	0.119	0.061	0.014
500	0.150	0.086	0.025	0.165	0.096	0.029
1,000	0.202	0.127	0.041	0.228	0.141	0.047
3,000	0.353	0.246	0.102	0.417	0.302	0.137
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.112	0.062	0.014	0.117	0.060	0.014
500	0.149	0.085	0.025	0.164	0.096	0.029
1,000	0.201	0.126	0.041	0.227	0.141	0.047
3,000	0.352	0.244	0.100	0.415	0.301	0.136
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.094	0.046	0.010	0.091	0.044	0.009
500	0.095	0.047	0.010	0.094	0.050	0.011
1,000	0.105	0.052	0.011	0.102	0.052	0.010
3,000	0.102	0.052	0.012	0.102	0.053	0.013

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$, which uses the traditional Fama and MacBeth (1973) standard error, $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0,\ \hat{\gamma}_1']'$, which uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992) and $t(\cdot)$ denotes the t-statistic associated with the Shanken estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*,\ \hat{\gamma}_1^{*\prime}]'$, which uses the standard error formulae of Theorem 2. Finally, the rejection rates for the t-test associated with the bias-adjusted estimator of Shanken (1992) are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Table IA.8
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ full, $\boldsymbol{\delta} = 0.25$)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.125	0.068	0.017	0.124	0.068	0.018
500	0.163	0.095	0.034	0.174	0.109	0.039
1,000	0.215	0.131	0.046	0.241	0.155	0.057
3,000	0.389	0.280	0.125	0.459	0.343	0.164
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.123	0.067	0.017	0.123	0.067	0.017
500	0.162	0.095	0.033	0.174	0.109	0.039
1,000	0.214	0.130	0.046	0.240	0.155	0.057
3,000	0.388	0.278	0.124	0.458	0.341	0.163
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.109	0.060	0.015	0.112	0.059	0.015
500	0.115	0.062	0.018	0.117	0.064	0.019
1,000	0.122	0.065	0.016	0.119	0.066	0.017
3,000	0.121	0.069	0.018	0.124	0.068	0.018

Table IA.8 (Continued)
Size of t-tests in a one-factor model ($\boldsymbol{\Sigma}$ full, $\boldsymbol{\delta} = 0.25$)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
	$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$		
100	0.119	0.060	0.014	0.123	0.066	0.015
500	0.155	0.091	0.025	0.163	0.098	0.030
1,000	0.199	0.126	0.042	0.222	0.138	0.050
3,000	0.334	0.229	0.092	0.390	0.280	0.124
	$t_{EIV}(\hat{\gamma}_0)$			$t_{EIV}(\hat{\gamma}_1)$		
100	0.117	0.059	0.014	0.122	0.065	0.015
500	0.155	0.090	0.025	0.162	0.098	0.030
1,000	0.198	0.125	0.042	0.222	0.138	0.049
3,000	0.333	0.228	0.091	0.388	0.278	0.123
	$t(\hat{\gamma}_0^*)$			$t(\hat{\gamma}_1^*)$		
100	0.108	0.057	0.012	0.110	0.059	0.015
500	0.114	0.062	0.015	0.119	0.065	0.015
1,000	0.121	0.063	0.015	0.122	0.067	0.016
3,000	0.111	0.057	0.012	0.114	0.058	0.014

The table presents the size properties of t-tests of statistical significance. The null hypothesis is that the parameter of interest is equal to its true value. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12. $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0, \ \hat{\gamma}_1']'$, which uses the traditional Fama and MacBeth (1973) standard error, $t_{EIV}(\cdot)$ denotes the t-statistic associated with the OLS estimator $\hat{\Gamma} = [\hat{\gamma}_0, \ \hat{\gamma}_1']'$, which uses the EIV-adjusted standard error in Theorem 1(ii) of Shanken (1992) and $t(\cdot)$ denotes the t-statistic associated with the Shanken estimator $\hat{\Gamma}^* = [\hat{\gamma}_0^*, \ \hat{\gamma}_1^{*\prime}]'$, which uses the standard error formulae of Theorem 2. Finally, the rejection rates for the t-test associated with the bias-adjusted estimator of Shanken (1992) are based on the asymptotic distribution in part (ii) of Theorem 1. The t-statistics are compared with the critical values from a standard normal distribution.

Table IA.9
Rejection rates of the specification test in a one-factor model ($T = 36$)

	Size			Power		
N	10%	5%	1%	10%	5%	1%
			Panel A: Σ spherical			
100	0.103	0.049	0.009	0.882	0.823	0.675
500	0.098	0.050	0.009	1.000	1.000	0.998
1,000	0.101	0.052	0.011	1.000	1.000	1.000
3,000	0.101	0.050	0.009	1.000	1.000	1.000
			Panel B: Σ diagonal			
100	0.085	0.037	0.010	0.634	0.529	0.340
500	0.093	0.046	0.010	0.983	0.967	0.894
1,000	0.099	0.050	0.009	1.000	1.000	0.996
3,000	0.097	0.046	0.011	1.000	1.000	1.000
			Panel C: Σ full ($\delta = 0.5$)			
100	0.084	0.040	0.011	0.639	0.534	0.332
500	0.101	0.050	0.012	0.982	0.965	0.887
1,000	0.095	0.049	0.011	1.000	1.000	0.997
3,000	0.108	0.056	0.011	1.000	1.000	1.000
			Panel D: Σ full ($\delta = 0.25$)			
100	0.110	0.060	0.021	0.621	0.522	0.336
500	0.145	0.084	0.029	0.977	0.956	0.874
1,000	0.145	0.088	0.029	1.000	0.999	0.993
3,000	0.146	0.087	0.030	1.000	1.000	1.000

The table presents the size and power properties of our test $\mathcal{S}^*$ of correct model specification. The null hypothesis is that the model is correctly specified. The alternative hypothesis is that the model is misspecified. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2010:12 ($T = 36$). Finally, the rejection rates for the specification test are based on the asymptotic distribution in Theorem 4. The rejection rates of the test are based on two-sided p-values.

Table IA.10
Rejection rates of the specification test in a one-factor model (T = 72)

	Size			Power		
N	10%	5%	1%	10%	5%	1%
			Panel A: Σ spherical			
100	0.095	0.045	0.009	0.929	0.891	0.781
500	0.101	0.047	0.009	1.000	1.000	1.000
1,000	0.104	0.055	0.010	1.000	1.000	1.000
3,000	0.099	0.048	0.010	1.000	1.000	1.000
			Panel B: Σ diagonal			
100	0.085	0.041	0.010	0.771	0.676	0.480
500	0.098	0.046	0.010	1.000	1.000	0.997
1,000	0.101	0.049	0.012	1.000	1.000	1.000
3,000	0.102	0.051	0.011	1.000	1.000	1.000
			Panel C: Σ full (δ = 0.5)			
100	0.085	0.039	0.011	0.770	0.681	0.482
500	0.092	0.046	0.009	1.000	0.999	0.996
1,000	0.094	0.049	0.010	1.000	1.000	1.000
3,000	0.097	0.047	0.010	1.000	1.000	1.000
			Panel D: Σ full (δ = 0.25)			
100	0.120	0.063	0.023	0.749	0.660	0.470
500	0.140	0.083	0.029	1.000	0.999	0.994
1,000	0.149	0.086	0.030	1.000	1.000	1.000
3,000	0.153	0.093	0.034	1.000	1.000	1.000

The table presents the size and power properties of our test $\mathcal{S}^*$ of correct model specification presented in Theorem 3. The null hypothesis is that the model is correctly specified. The alternative hypothesis is that the model is misspecified. The results are reported for different levels of significance (10%, 5%, and 1%) and for different values of the number of stocks (N) using 10,000 simulations, assuming that the model disturbances are generated from a multivariate normal distribution with a covariance matrix calibrated to 3,000 NYSE-AMEX-NASDAQ individual stock returns over the period 2008:1–2013:12 (T = 72). Finally, the rejection rates for the specification test are based on the asymptotic distribution in Theorem 4. The rejection rates of the test are based on two-sided p-values.

IA.5. Unbalanced Panels

In this section, we extend our methodology to the case of an unbalanced panel, focusing for simplicity on the base case of correctly specified models with constant risk premia. Following Gagliardini, Ossola, and Scaillet (2016), we assume a missing at random design (see, for example, Rubin 1976), that is, independence between unobservability and return generating process. This allows us to keep the factor structure linear. In the following analysis, we explicitly account for the randomness of T_i, the time-series sample size for asset i. Define the following $T \times T$ matrix

$$J_i = \text{diag}(J_{i1} \cdots J_{it} \cdots J_{iT}) \qquad i = 1, \ldots, N, \tag{IA.151}$$

where $\sum_{t=1}^{T} J_{it} = T_i$ and $J_{it} = 1$ if the return on asset i is observed by the econometrician at date t, and zero otherwise. In addition, let $R_{i,u} = J_i R_i$, $F_{i,u} = J_i F$, and $\epsilon_{i,u} = J_i \epsilon_i$, and assume that asset returns are governed by the multifactor model

$$J_{it} R_{it} \quad = \quad J_{it}\alpha_i + J_{it} f_t' \beta_i + J_{it}\epsilon_{it}, \tag{IA.152}$$

that is, the same data generating process of Section 1 pre-multiplied by J_{it}. Let $\bar{R}_{i,u} = \frac{1}{T_i}\sum_{t=1}^{T} J_{it} R_{it}$, $\bar{f}_{i,u} = \frac{1}{T_i}\sum_{t=1}^{T} J_{it} f_t$, and $\bar{\epsilon}_{i,u} = \frac{1}{T_i}\sum_{t=1}^{T} J_{it}\epsilon_{it}$. Averaging Equation (IA.152) over time, imposing the asset-pricing restriction, and noting that $E[R_{it}] = \alpha_i + \beta_i' E[f_t]$ yields

$$\bar{R}_{i,u} = \gamma_0 + \hat{\beta}_{i,u}' \gamma_{1i,u}^P + \eta_{i,u}^P, \tag{IA.153}$$

where $\gamma_{1i,u}^P = \gamma_1 + \bar{f}_{i,u} - E(f_t)$, $\eta_{i,u}^P = \bar{\epsilon}_{i,u} - (\hat{\beta}_{i,u} - \beta_i)'\gamma_{1i,u}^P$, $\hat{\beta}_{i,u} = \beta_i + \mathcal{P}_{i,u}'\epsilon_i$, $\mathcal{P}_{i,u} = \tilde{F}_{i,u}(\tilde{F}_{i,u}'\tilde{F}_{i,u})^{-1}$, and $\tilde{F}_{i,u} = F_{i,u} - J_i 1_T \bar{f}_{i,u}'$. Since the panel is unbalanced, there is now a sequence of ex post risk premia, one for each asset i.

In matrix form, we have

$$\bar{R}_u = \gamma_0 1_N + \begin{bmatrix} \hat{\beta}_{1,u}' & & 0_{K\times(N-1)}' \\ \vdots & \ddots & \vdots \\ 0_{K\times(N-1)}' & & \hat{\beta}_{N,u}' \end{bmatrix} \begin{bmatrix} \gamma_{11,u}^P \\ \vdots \\ \gamma_{1N,u}^P \end{bmatrix} + \begin{bmatrix} \eta_{1,u}^P \\ \vdots \\ \eta_{N,u}^P \end{bmatrix}, \tag{IA.154}$$

where $\bar{R}_u = (\bar{R}_{1,u}, \ldots, \bar{R}_{N,u})'$. Define the $N \times K$ matrix $\hat{X}_u = [1_N,\ \hat{B}_u]$, where $\hat{B}_u = (\hat{\beta}_{1,u}, \ldots, \hat{\beta}_{N,u})'$. Denote by $\hat{\epsilon}_{i,u}$ the T-vector of residuals from the first-pass (unbalanced) OLS regressions in

$$R_{i,u} = \alpha_i J_i 1_T + F_{i,u}\beta_i + \epsilon_{i,u}, \quad i = 1, \ldots, N. \tag{IA.155}$$

The modified estimator of the ex post risk premia in the unbalanced panel case is

$$\hat{\Gamma}_u^* = \begin{bmatrix} \hat{\gamma}_{0,u}^* \\ \hat{\gamma}_{1,u}^* \end{bmatrix} = \left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1} \frac{\hat{X}_u' \bar{R}_u}{N}, \tag{IA.156}$$

where $\hat{\Sigma}_{X,u} = \frac{\hat{X}_u' \hat{X}_u}{N}$, $\hat{\Lambda}_u = \begin{bmatrix} 0 & 0_K' \\ 0_K & \hat{\sigma}_u^2 \hat{\mathcal{F}}_u \end{bmatrix}$ with $\hat{\sigma}_u^2 = \frac{1}{N}\sum_{i=1}^N \left(\frac{1}{T_i - K - 1} \mathrm{tr}\left(\hat{\epsilon}_{i,u} \hat{\epsilon}_{i,u}'\right)\right)$ and $\hat{\mathcal{F}}_u = \frac{1}{N}\sum_{i=1}^N \left(\tilde{F}_{i,u}' \tilde{F}_{i,u}\right)^{-1}$.

The estimator $\hat{\Gamma}_u^*$ in Equation (IA.156) generalizes the bias-adjusted estimator of Shanken (1992) to the unbalanced panel case and coincides with the Shanken's estimator when the panel is balanced. Let $\Sigma_{X,i} = \begin{bmatrix} 1 & \beta_i' \\ \beta_i & \beta_i \beta_i' \end{bmatrix}$, $\Sigma_{F\beta} = \mathrm{plim} \frac{1}{N} \sum_{i=1}^N \beta_i' F' F \beta_i \Sigma_{X,i}$, $\mathcal{F}_u = \mathrm{plim} \frac{1}{N} \sum_{i=1}^N \mathcal{P}_{i,u}' \mathcal{P}_{i,u}$, and $Q_{i,u} = \frac{J_i 1_T}{T_i} - \mathcal{P}_{i,u} \gamma_1^P$. Finally, define $Z_{i,u} = \left[\left(Q_{i,u} \otimes \mathcal{P}_{i,u}\right) + \frac{\mathrm{vec}\left(M_{i,u}\right)}{T_i - K - 1} \gamma_1^{P\prime} \mathcal{P}_{i,u}' \mathcal{P}_{i,u}\right]$ and $M_{i,u} = \left[I_T - J_i D (D' J_i D)^{-1} D' J_i\right] J_i$.

The following additional assumptions are required for the asymptotic analysis in the unbalanced panel case.

Assumption IA.1

$$\sup_i \|\beta_i\| \le C < \infty, \tag{IA.157}$$

where C is a generic constant.

Assumption IA.2 J_{it} is i.i.d. across i and t. Let C be a generic constant, and assume that $T_i > K + 1$, for every $i = 1, \ldots, N$. Then, we have

(i)

$$\tau = E\left[\frac{1}{T_i}\right] \le C < \infty \quad \text{and} \quad \theta = \mathrm{Var}\left(\frac{J_{it}}{T_i}\right) \le C < \infty. \tag{IA.158}$$

(ii)

$$\left|E\left[\frac{J_{it}}{T_i^4}\right]\right| \le C < \infty. \tag{IA.159}$$

(iii)

$$E\left[\frac{J_{it}}{(T_i - K - 1)^2}\right] \le C < \infty. \tag{IA.160}$$

(iv) Let d_t be the t-th row of D and define $p_{it,u} = J_{it} d_t \left(D' J_i D\right)^{-1} D' J_i$. Then, we have

$$E\left[\frac{p_{it,u} p_{it,u}'}{(T_i - K - 1)^2}\right] \leq C < \infty. \tag{IA.161}$$

(v) Let $\tilde{F}_{i,u} = F_{i,u} - J_i 1_T \frac{1_T' J_i F}{T_i}$, where $F_{i,u} = J_i F$. Then,

$$E\left[(\tilde{F}_{i,u}' \tilde{F}_{i,u})^{-1}\right] \leq C < \infty \tag{IA.162}$$

and

$$E\left\|(\tilde{F}_{i,u}' \tilde{F}_{i,u})^{-1}\right\|^4 \leq C < \infty. \tag{IA.163}$$

(vi) Let $m_{ts,i}^u$ be the (t, s)-th element of the matrix $M_{i,u} = (I_T - J_i D(D' J_i D)^{-1} D' J_i) J_i$, and define $M_{i,u}^{(2)} = M_{i,u} \odot M_{i,u}$, where $\odot$ denotes the Hadamard product operator. Then,

$$\sup_i E\left[|m_{ts,i}^u|^8\right] \leq C < \infty \tag{IA.164}$$

and

$$\frac{1}{N} \sum_{i=1}^{N} \operatorname{tr}\left(M_{i,u}^{(2)}\right) \xrightarrow{p} C > 0. \tag{IA.165}$$

Assumption IA.1 is a boundedness assumption. In assumption IA.2, we assume a missing at random design, that is, independence between unobservability and return generating process. Assumptions IA.2(i)–(iv) are ruling out that the distribution of the T_i is too concentrated around zero. Assumption IA.2(v) is essentially extending the non-singularity of the covariance matrix of the factors to the missing-at-random design, and assumption IA.2(vi) is technical in nature. The consistency and asymptotic normality of the proposed estimator are provided in the following theorem. The proofs of the theorems in this section are available upon request.

Theorem IA.2 Under assumptions 1–6 and assumptions IA.1–IA.2, as $N \to \infty$, we have

(i)

$$\hat{\Gamma}_u^* - \Gamma^P = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{IA.166}$$

(ii)

$$\sqrt{N}\left(\hat{\Gamma}_u^* - \Gamma^P\right) \xrightarrow{d} \mathcal{N}\left(0_{K+1}, V_u + \Sigma_X^{-1}(W_u + \Theta)\Sigma_X^{-1}\right), \tag{IA.167}$$

where

$$V_u = \sigma^2\left(\tau + \gamma_1^{P\prime}\mathcal{F}_u\gamma_1^P\right)\Sigma_X^{-1}, \qquad W_u = \begin{bmatrix} 0 & 0_K' \\ 0_K & \text{plim}\frac{1}{N}\sum_{i=1}^N Z_{i,u}'U_\epsilon Z_{i,u} \end{bmatrix}, \qquad \Theta = \theta\Sigma_{F\beta} - \sigma^2\Psi,$$

with $\Psi = \begin{bmatrix} 0 & \gamma_1^{P\prime}\mathcal{F}_\gamma \\ \mathcal{F}_\gamma'\gamma_1^P & \mathcal{F}_{\gamma\beta} \end{bmatrix}$, $\mathcal{F}_\gamma = \text{plim}\frac{1}{N}\sum_{i=1}^N \mathcal{P}_{i,u}'\mathcal{P}_{i,u}(\bar{f}_{i,u} - \bar{f})'\beta_i$, and $\mathcal{F}_{\gamma\beta} = \text{plim}\frac{1}{N}\sum_{i=1}^N(\beta_i\beta_i'(\bar{f}_{i,u} - \bar{f})\gamma_1^{P\prime}\mathcal{P}_{i,u}'\mathcal{P}_{i,u} + \mathcal{P}_{i,u}'\mathcal{P}_{i,u}\gamma_1^P(\bar{f}_{i,u} - \bar{f})'\beta_i\beta_i' - (\bar{f}_{i,u} - \bar{f})'\beta_i\beta_i'(\bar{f}_{i,u} - \bar{f})\mathcal{P}_{i,u}'\mathcal{P}_{i,u})$.

It should be noted that the asymptotic covariance matrix in Theorem IA.2 is similar to the one for the balanced panel case provided in Theorem 1. The additional terms in part (ii) of Theorem IA.2 account for the randomness of the sample size T_i. When the panel is balanced, Theorem IA.2 reduces to Theorem 1 since $T_i = T$, $J_{it} = 1$, $\bar{f}_{i,u} = \bar{f}$, which implies that $\tau = 1/T$, $\theta = 0$, $\Psi = \Theta = 0_{(K+1)\times(K+1)}$, and all the relevant quantities do not depend on i anymore.

For statistical inference, we need a consistent estimator of the asymptotic covariance matrix of $\hat{\Gamma}_u^*$, as illustrated in the next theorem. Let $\hat{\tau} = \frac{1}{N}\sum_{i=1}^N \frac{1}{T_i}$, $\hat{\Sigma}_{X,i}^a = \begin{bmatrix} 1 & \hat{\beta}_{i,u}' \\ \hat{\beta}_{i,u} & \hat{\Sigma}_{\hat{\beta}_{i,u}}^a \end{bmatrix}$, where $\hat{\Sigma}_{\hat{\beta}_{i,u}}^a = \hat{\beta}_{i,u}\hat{\beta}_{i,u}' - \hat{\sigma}_u^2\mathcal{P}_{i,u}'\mathcal{P}_{i,u}$, $\hat{b}_i = \text{tr}(F'F\hat{\Sigma}_{\hat{\beta}_{i,u}}^a)$, and $A_i = \mathcal{P}_{i,u}'\mathcal{P}_{i,u}F'F$. Also, let $\hat{\mathcal{U}}_i = \sum_{t=1}^T(\mathcal{P}_{i,u}' \otimes f_t'\mathcal{P}_{i,u}')\hat{U}_\epsilon(\mathcal{P}_{i,u}\otimes\mathcal{P}_{i,u}f_t)$, where $\hat{U}_\epsilon$ (as in the balanced panel case) is a plug-in estimator of U_ϵ that depends only on $\hat{\sigma}_{4,u} = \frac{\frac{1}{N}\sum_{t=1}^T\sum_{i=1}^N\hat{\epsilon}_{it,u}^4}{3\frac{1}{N}\sum_{i=1}^N\text{tr}\left(M_{i,u}^{(2)}\right)}$, with $\hat{\epsilon}_{it,u}$ being the t-th element of $\hat{\epsilon}_{i,u}$ and $M_{i,u}^{(2)} = M_{i,u}\odot M_{i,u}$. Finally, let $\hat{\Sigma}_{F\beta} = \frac{1}{N}\sum_{i=1}^N\hat{b}_i\hat{\Sigma}_{X,i}^a - \hat{\Upsilon}$, where $\hat{\Upsilon} = \frac{1}{N}\sum_{i=1}^N\begin{bmatrix} 0 & 2\hat{\sigma}_u^2\hat{\beta}_{i,u}'A_i' \\ 2\hat{\sigma}_u^2A_i\hat{\beta}_{i,u} & 2\hat{\sigma}_u^2(A_i\hat{\Sigma}_{\hat{\beta}_{i,u}}^a + \hat{\Sigma}_{\hat{\beta}_{i,u}}^a A_i') + \hat{\mathcal{U}}_i \end{bmatrix}$, $\hat{\theta} = \frac{1}{NT}\sum_{t=1}^T\sum_{i=1}^N\frac{J_{it}}{T_i^2} - \frac{1}{T^2}$, and $\hat{Z}_{i,u} = \left[\left(\hat{Q}_{i,u}\otimes\mathcal{P}_{i,u}\right) + \frac{\text{vec}(M_{i,u})}{T_i-K-1}\hat{\gamma}_{1,u}^{*\prime}\mathcal{P}_{i,u}'\mathcal{P}_{i,u}\right]$, where $\hat{Q}_{i,u} = \frac{J_i1_T}{T_i} - \mathcal{P}_{i,u}\hat{\gamma}_{1,u}^*$.

Theorem IA.3 Under assumptions 2–6 and assumptions IA.1–IA.2, setting $\kappa_4 = 0$, as $N\to\infty$, we have

$$\hat{V}_u + \left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1}(\hat{W}_u + \hat{\Theta})\left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1} \xrightarrow{p} V_u + \Sigma_X^{-1}(W_u + \Theta)\Sigma_X^{-1}, \tag{IA.168}$$

where

$$\hat{V}_u = \left[\hat{\sigma}_u^2\left(\hat{\tau} + \hat{\gamma}_{1,u}^{*\prime}\hat{\mathcal{F}}_u\hat{\gamma}_{1,u}^*\right)\right]\left(\hat{\Sigma}_{X,u} - \hat{\Lambda}_u\right)^{-1}, \qquad \hat{W}_u = \begin{bmatrix} 0 & 0_K' \\ 0_K & \frac{1}{N}\sum_{i=1}^N\hat{Z}_{i,u}'\hat{U}_\epsilon\hat{Z}_{i,u} \end{bmatrix}, \qquad \hat{\Theta} = \hat{\theta}\hat{\Sigma}_{F\beta} - \hat{\sigma}_u^2\hat{\Psi},$$

with $\hat{\Psi} = \begin{bmatrix} 0 & \hat{\gamma}_{1,u}^{*\prime}\hat{\mathcal{F}}_{\gamma} \\ \hat{\mathcal{F}}_{\gamma}^{\prime}\hat{\gamma}_{1,u}^{*} & \hat{\mathcal{F}}_{\gamma\beta} \end{bmatrix}$, $\hat{\mathcal{F}}_{\gamma} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{P}_{i,u}^{\prime}\mathcal{P}_{i,u}(\bar{f}_{i,u}-\bar{f})^{\prime}\hat{\beta}_{i,u}$, and $\hat{\mathcal{F}}_{\gamma\beta} = \frac{1}{N}\sum_{i=1}^{N}\hat{\Sigma}^{a}_{\hat{\beta}_{i,u}}(\bar{f}_{i,u}-\bar{f})\hat{\gamma}_{1,u}^{*\prime}\mathcal{P}_{i,u}^{\prime}\mathcal{P}_{i,u} + \frac{1}{N}\sum_{i=1}^{N}\mathcal{P}_{i,u}^{\prime}\mathcal{P}_{i,u}\hat{\gamma}_{1,u}^{*}(\bar{f}_{i,u}-\bar{f})^{\prime}\hat{\Sigma}^{a}_{\hat{\beta}_{i,u}} - \frac{1}{N}\sum_{i=1}^{N}(\bar{f}_{i,u}-\bar{f})^{\prime}\hat{\Sigma}^{a}_{\hat{\beta}_{i,u}}(\bar{f}_{i,u}-\bar{f})\mathcal{P}_{i,u}^{\prime}\mathcal{P}_{i,u}$.

Turning to specification testing, let

$$\hat{e}_{u}^{P} = \bar{R}_{u} - \hat{X}_{u}\hat{\Gamma}_{u}^{*} \tag{IA.169}$$

be the N-vector of ex post sample pricing errors. Define $\hat{\mathcal{Q}}_{u} = \frac{\hat{e}_{u}^{P\prime}\hat{e}_{u}^{P}}{N}$ as the sum of squared ex post sample pricing errors and denote by $\hat{\Sigma}^{a}_{\hat{\beta}_{u}} = \left(\frac{\hat{B}_{u}^{\prime}\hat{B}_{u}}{N} - \hat{\sigma}_{u}^{2}\hat{\mathcal{F}}_{u}\right)$, $\hat{b} = \mathrm{tr}(F^{\prime}F\hat{\Sigma}^{a}_{\hat{\beta}_{u}})$,
$\omega_{N} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{J_{it}}{T_{i}} - \frac{1}{T}\right)^{2}\mathrm{tr}\left(\mathcal{P}_{i,u}f_{t}f_{t}^{\prime}\mathcal{P}_{i,u}^{\prime}\right)$, and $Z_{Q_{i,u}} = \left[\left(Q_{i,u}^{\prime}\otimes Q_{i,u}^{\prime}\right) - \frac{Q_{i,u}^{\prime}Q_{i,u}\mathrm{vec}(M_{i,u})^{\prime}}{T_{i}-K-1}\right]^{\prime}$. Finally, consider the centered statistic

$$\mathcal{S}_{u} = \sqrt{N}\left(\hat{\mathcal{Q}}_{u} - \hat{\sigma}_{u}^{2}(\hat{\tau} + \hat{\gamma}_{1,u}^{*\prime}\hat{\mathcal{F}}_{u}\hat{\gamma}_{1,u}^{*}) - \hat{\theta}\,\hat{b}\right). \tag{IA.170}$$

Theorem IA.4 Under assumptions 2–6 and assumptions IA.1–IA.2, as $N \to \infty$, we have

$$\mathcal{S}_{u} \xrightarrow{d} \mathcal{N}\left(0, \mathcal{V}_{u} + \mathcal{W}_{u}\right), \tag{IA.171}$$

where $\mathcal{V}_{u} = \mathrm{plim}\frac{1}{N}\sum_{i=1}^{N}\tilde{Z}_{Q_{i,u}}^{\prime}U_{\epsilon}\tilde{Z}_{Q_{i,u}}$ and $\mathcal{W}_{u} = 4\sigma^{2}\mathrm{plim}\frac{1}{N}\sum_{i=1}^{N}W_{i}^{\prime}W_{i}$, with

$$\tilde{Z}_{Q_{i,u}} = Z_{Q_{i,u}} + \left(\omega_{N}\left(\frac{\mathrm{vec}(M_{i,u})}{T_{i}-K-1}\right) - \sum_{t=1}^{T}\left(\frac{J_{it}}{T_{i}} - \frac{1}{T}\right)^{2}\mathrm{vec}\left(\mathcal{P}_{i,u}f_{t}f_{t}^{\prime}\mathcal{P}_{i,u}^{\prime}\right)\right)$$

and

$$W_{i} = \left[(\gamma_{1i,u}^{P} - \gamma_{1}^{P})^{\prime}\beta_{i}Q_{i,u}^{\prime} - \sum_{t=1}^{T}\left(\frac{J_{it}}{T_{i}} - \frac{1}{T}\right)^{2}\beta_{i}^{\prime}f_{t}f_{t}^{\prime}\mathcal{P}_{i,u}^{\prime}\right]^{\prime}.$$

Note that when the panel is balanced, Theorem IA.4 reduces to Theorem 4 since $\frac{J_{it}}{T_{i}} = \frac{1}{T}$ and $\bar{f}_{i,u} = \bar{f}$, which implies that $\mathcal{W}_{u} = 0$, $Q_{i,u} = Q$, and $\tilde{Z}_{Q_{i,u}} = Z_{Q_{i,u}} = Z_{Q}$. This variance can be consistently estimated. Let $\hat{Z}_{Q_{i,u}} = \left[\left(\hat{Q}_{i,u}^{\prime}\otimes\hat{Q}_{i,u}^{\prime}\right) - \frac{\hat{Q}_{i,u}^{\prime}\hat{Q}_{i,u}\mathrm{vec}(M_{i,u})^{\prime}}{T_{i}-K-1}\right]^{\prime}$ and $\hat{\tilde{Z}}_{Q_{i,u}} = \hat{Z}_{Q_{i,u}} + \left(\omega_{N}\left(\frac{\mathrm{vec}(M_{i,u})}{T_{i}-K-1}\right) - \sum_{t=1}^{T}\left(\frac{J_{it}}{T_{i}} - \frac{1}{T}\right)^{2}\mathrm{vec}\left(\mathcal{P}_{i,u}f_{t}f_{t}^{\prime}\mathcal{P}_{i,u}^{\prime}\right)\right)$. Then, the estimators of $\mathcal{V}_{u}$ and $\mathcal{W}_{u}$ are given by

$$\hat{\mathcal{V}}_{u} = \frac{1}{N}\sum_{i=1}^{N}\hat{\tilde{Z}}_{Q_{i,u}}^{\prime}\hat{U}_{\epsilon,u}\hat{\tilde{Z}}_{Q_{i,u}} \tag{IA.172}$$

and

$$
\begin{aligned}
\hat{\mathcal{W}}_u &= 4\hat{\sigma}_u^2 \frac{1}{N}\sum_{i=1}^{N}\Bigg(\hat{Q}_{i,u}'\hat{Q}_{i,u}(\bar{f}_{i,u}-\bar{f})'\hat{\Sigma}^a_{\hat{\beta}_{i,u}}(\bar{f}_{i,u}-\bar{f})\Bigg) \\
&\quad +4\hat{\sigma}_u^2 \frac{1}{N}\sum_{i=1}^{N}\Bigg(\sum_{t=1}^{T}\left(\frac{J_{it}}{T_i}-\frac{1}{T}\right)^4 \operatorname{tr}\Bigg(f_t f_t' \mathcal{P}_{i,u}' \mathcal{P}_{i,u} f_t f_t' \hat{\Sigma}^a_{\hat{\beta}_{i,u}}\Bigg) \\
&\quad -2\hat{Q}_{i,u}'\mathcal{P}_{i,u}\sum_{t=1}^{T}\left(\frac{J_{it}}{T_i}-\frac{1}{T}\right)^2 f_t f_t' \hat{\Sigma}^a_{\hat{\beta}_{i,u}}(\bar{f}_{i,u}-\bar{f})\Bigg).
\end{aligned}
\tag{IA.173}
$$

IA.6. Empirical Application: CAPM, Fama and French (1993) Three-Factor Model, and Fama and French (2015) Five-Factor Model

This section contain several figures for the CAPM and the Fama and French (1993) three-factor model (FF3). We also report further results for the Fama and French (2015) five-factor model (FF5). We first consider specification testing. Then, we present the risk and characteristic premia estimates for these three beta-pricing models. All figures are formatted as in the paper.

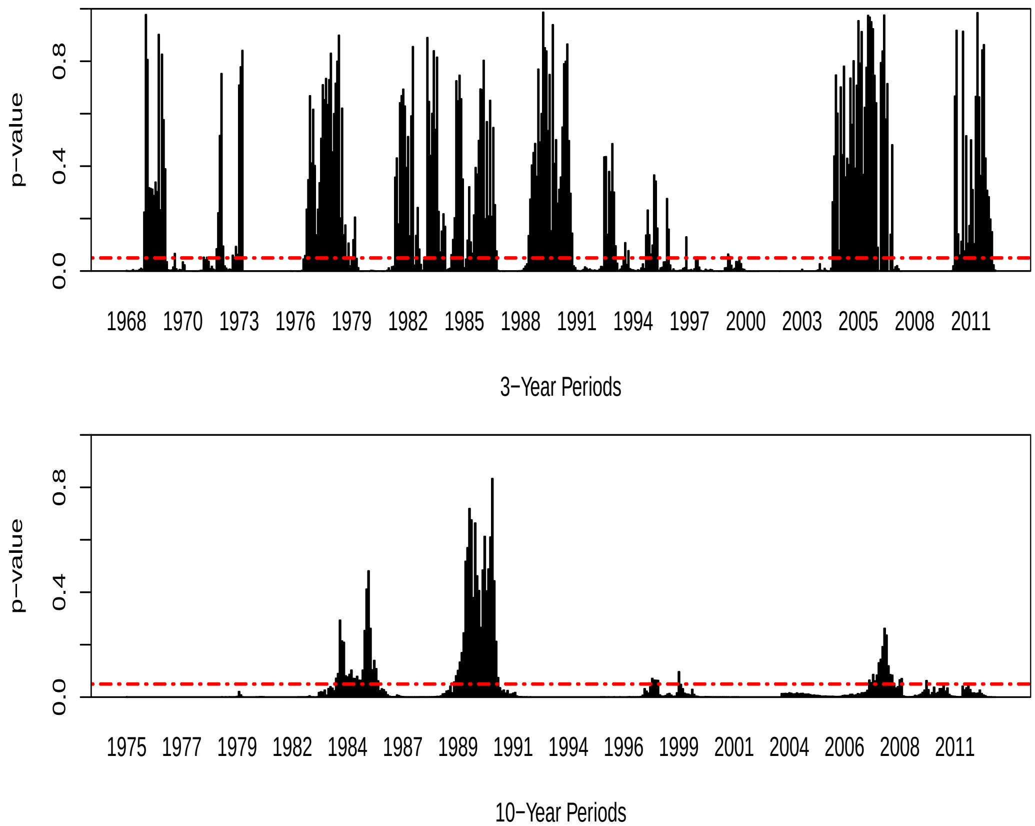


Figure IA.1
Specification testing for CAPM

The figure presents the time series of p-values (black line) of $\mathcal{S}^*$ for CAPM. Rolling time windows of three (top panel) and 10 years (bottom panel) are used. The dashed dotted red line denotes the 5% significance level of the test. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

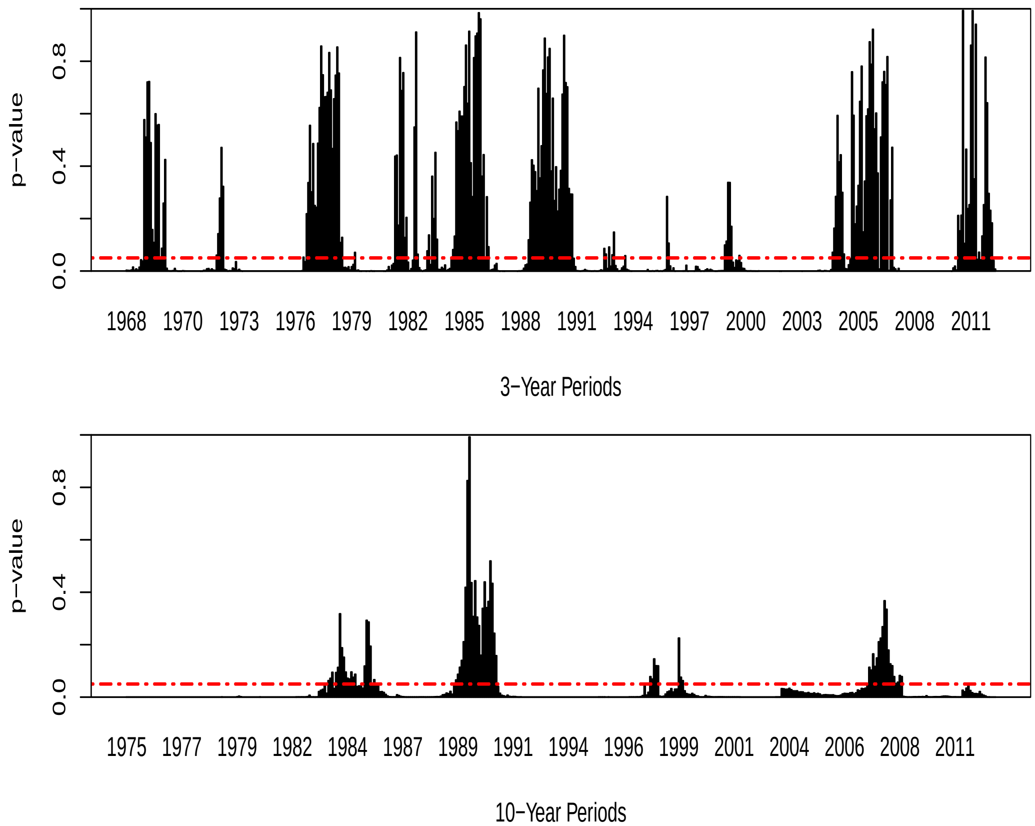


Figure IA.2
Specification testing for the liquidity-augmented CAPM

The figure presents the time series of p-values (black line) of $\mathcal{S}^*$ for the liquidity-augmented CAPM. Rolling time windows of three (top panel) and 10 years (bottom panel) are used. The dashed dotted red line denotes the 5% significance level of the test. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013.

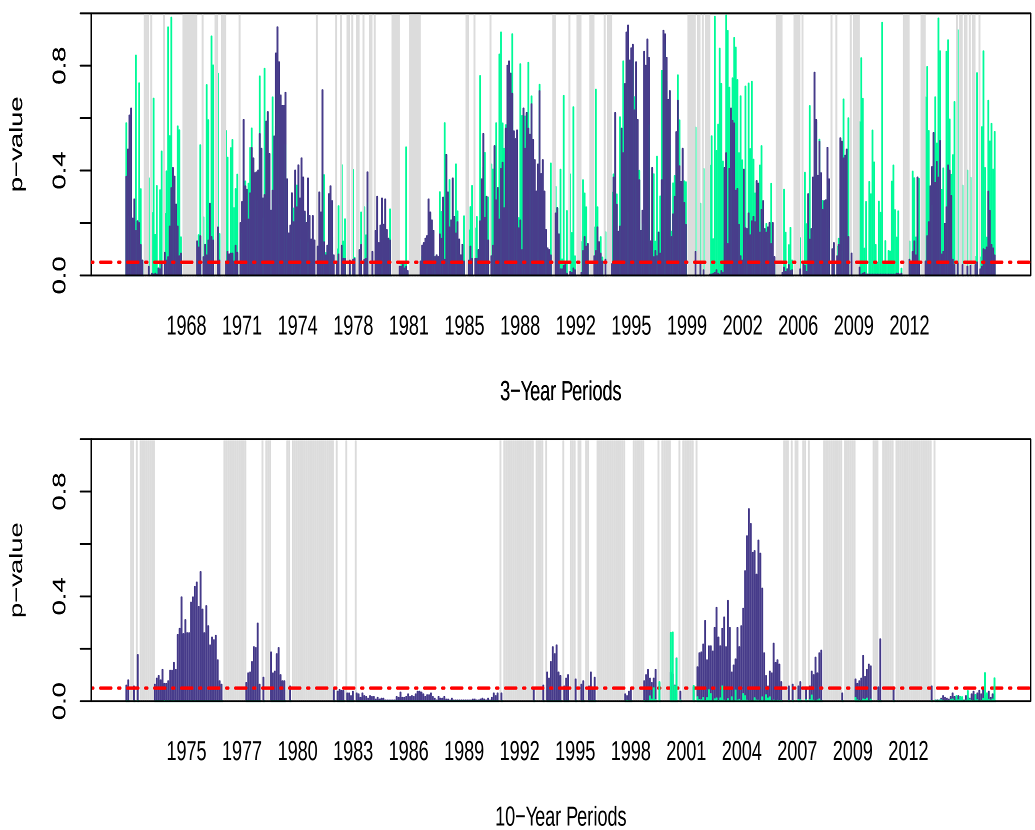


Figure IA.3
Specification testing for CAPM using the Gibbons, Ross, and Shanken (1989) and Gungor and Luger (2016) tests

The figure presents the time series of p-values of the GRS (blue line) and GL (green line) tests for CAPM. Rolling time windows of three (top panel) and 10 years (bottom panel) are used. The dashed dotted red line denotes the 5% significance level of the tests. The gray bars are for the periods in which the GL test is inconclusive. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

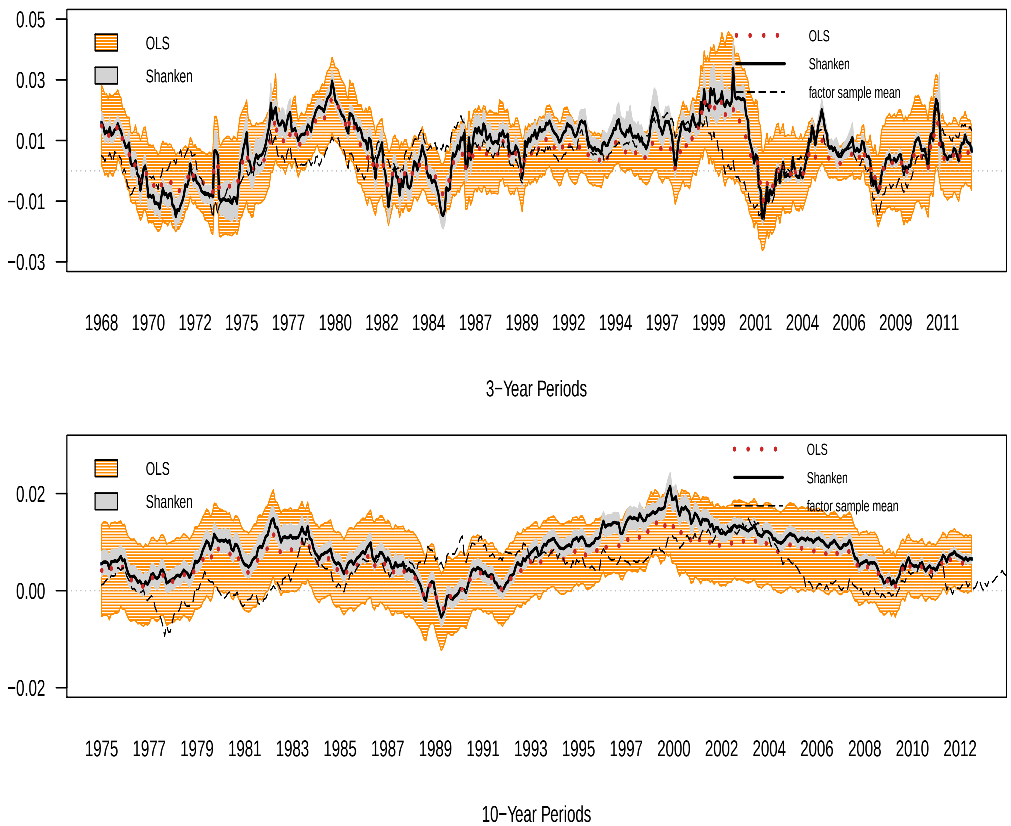


Figure IA.4
Estimates and confidence intervals for the market risk premium in CAPM
The figure presents the estimates and the associated confidence intervals for the market risk premium in CAPM. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

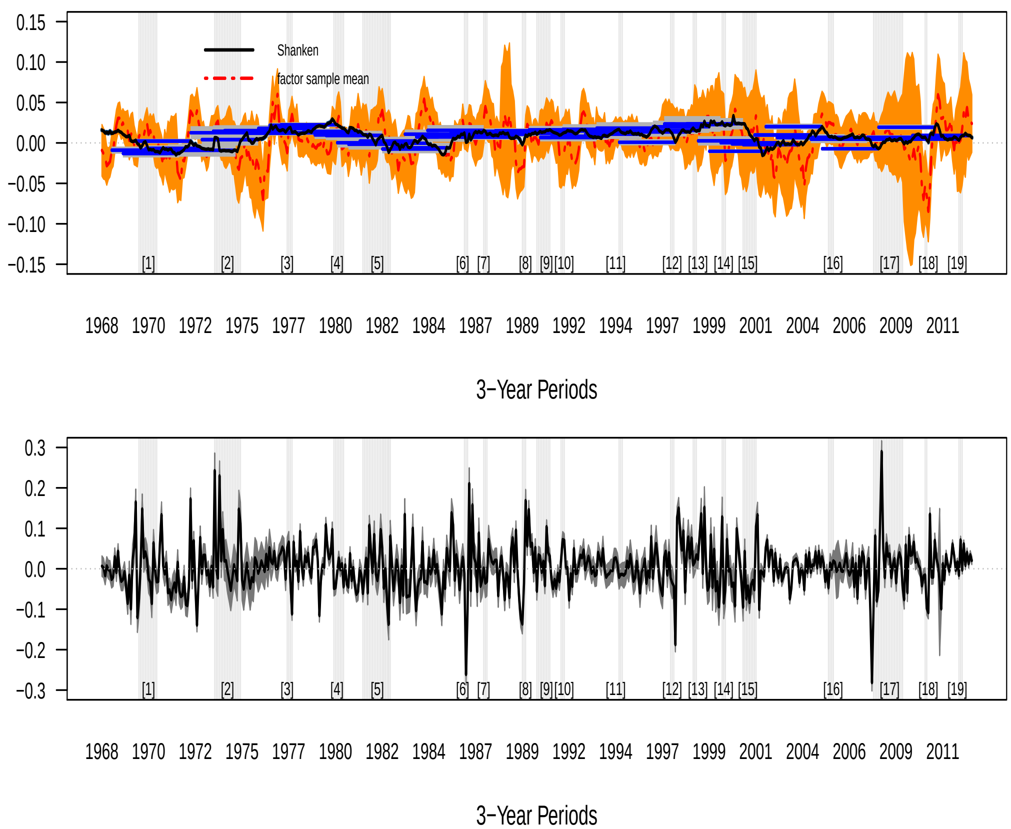


Figure IA.5
Estimates and confidence intervals for the time-varying market risk premium in CAPM

The figure presents the estimates and the associated confidence intervals for the time-varying market risk premium in CAPM. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the market excess return (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the market excess return is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

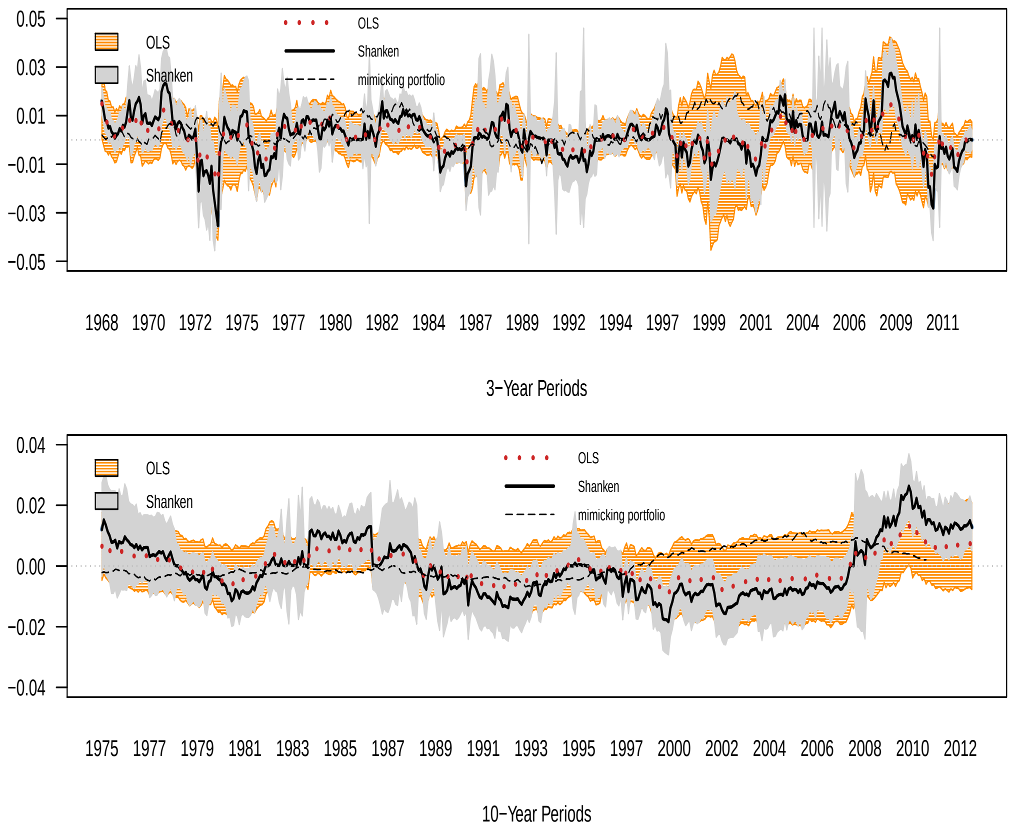
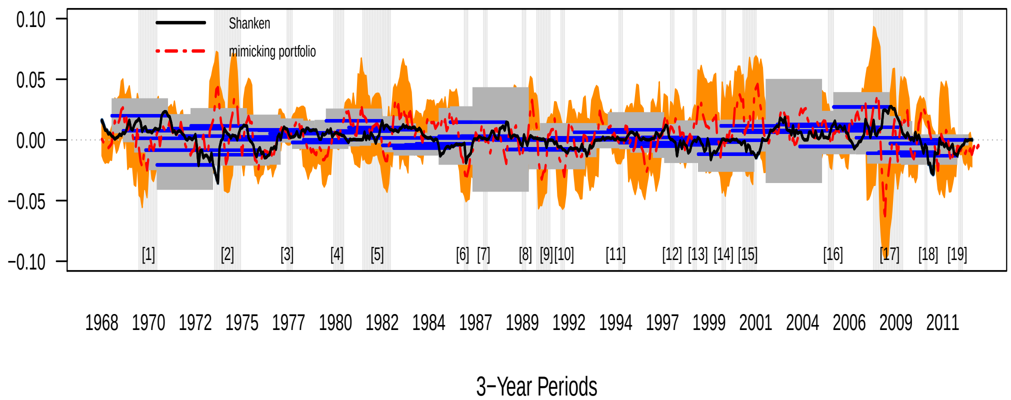


Figure IA.6
Estimates and confidence intervals for the liquidity risk premium in the liquidity-augmented CAPM

The figure presents the estimates and the associated confidence intervals for the liquidity risk premium in the liquidity-augmented CAPM. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the mimicking portfolio rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013.



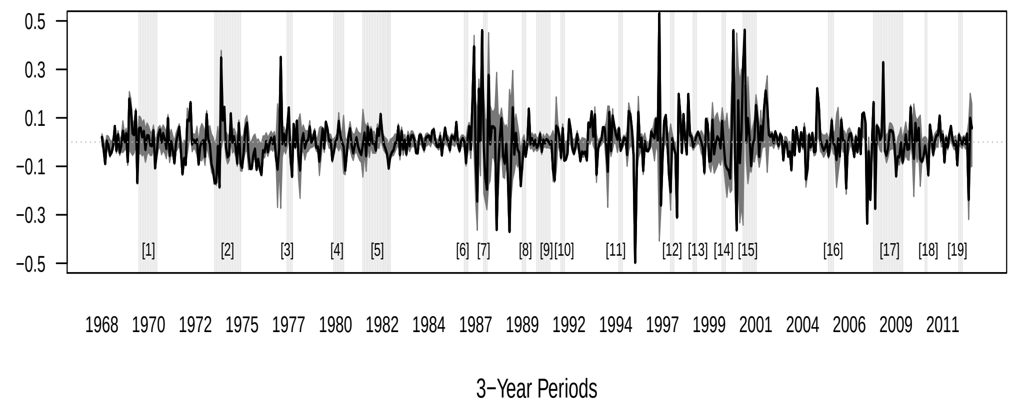


Figure IA.7
Estimates and confidence intervals for the time-varying liquidity risk premium in the liquidity-augmented CAPM

The figure presents the estimates and the associated confidence intervals for the time-varying liquidity risk premium in the liquidity-augmented CAPM. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months) of the corresponding mimicking portfolio excess return (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

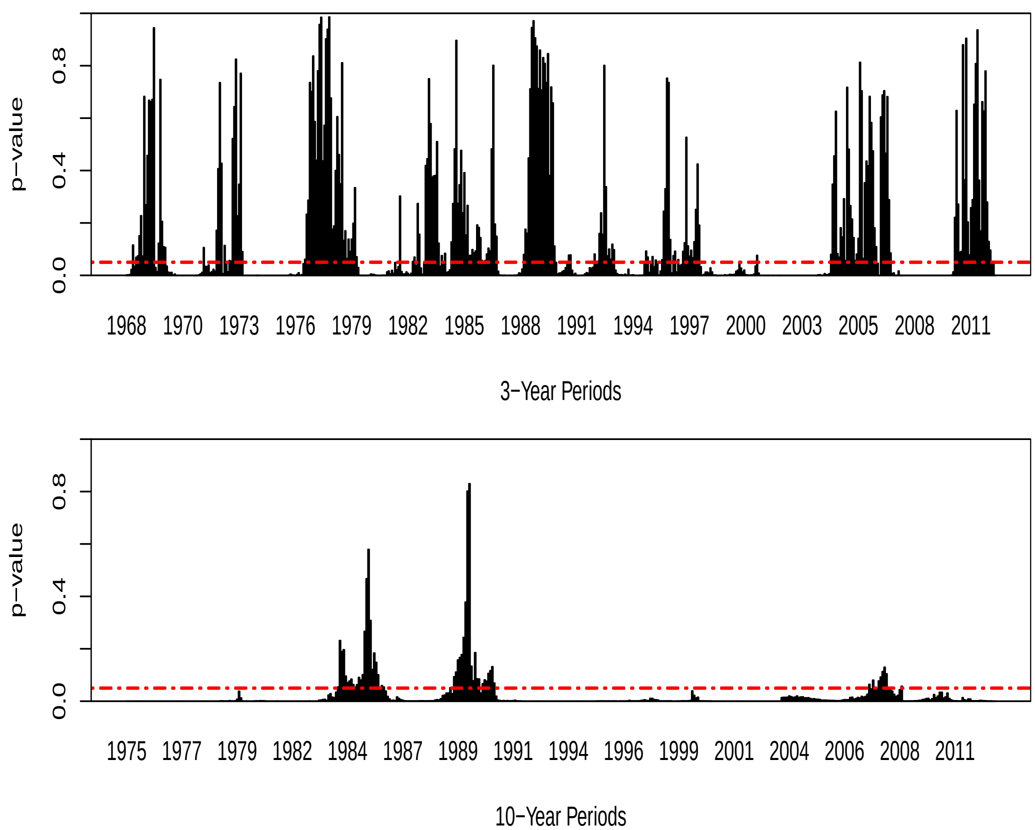


Figure IA.8
Specification testing for the Fama and French (1993) (FF3) three-factor model

The figure presents the time series of p-values (black line) of $\mathcal{S}^*$ for FF3. Rolling time windows of three (top panel) and 10 years (bottom panel) are used. The dashed dotted red line denotes the 5% significance level of the test. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

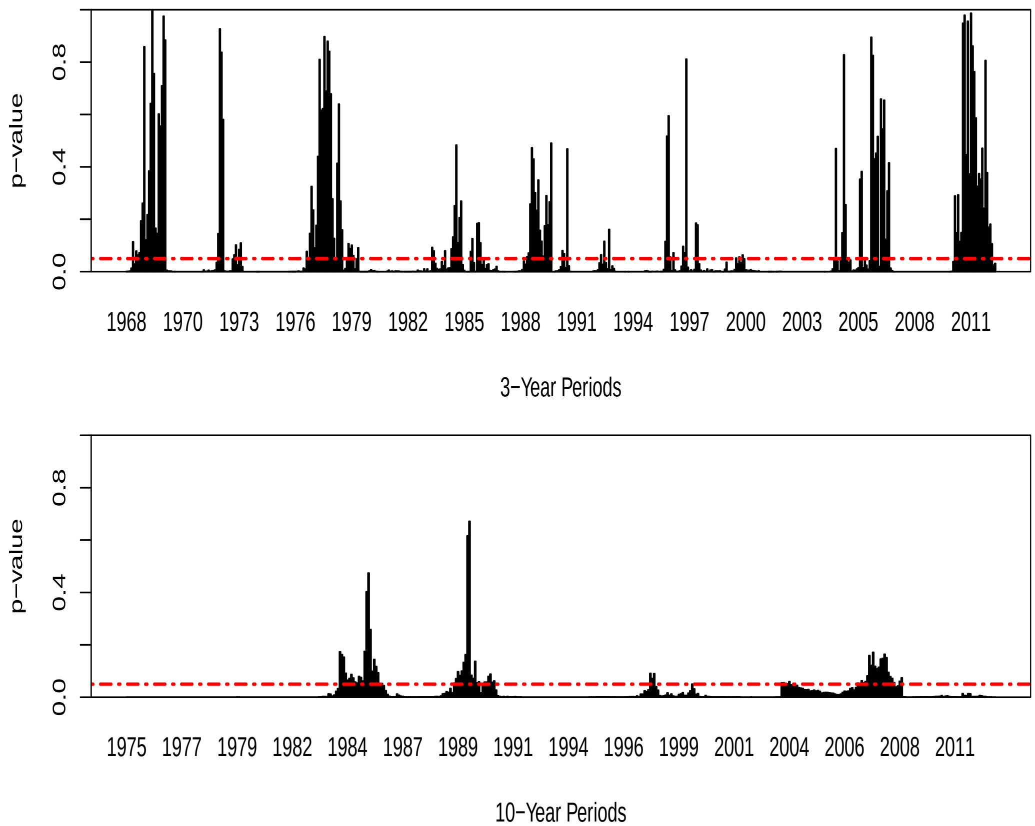


Figure IA.9
Specification testing for the liquidity-augmented Fama and French (1993) (FF3) three-factor model

The figure presents the time series of p-values (black line) of $\mathcal{S}^*$ for the liquidity-augmented FF3 model. Rolling time windows of three (top panel) and 10 years (bottom panel) are used. The dashed dotted red line denotes the 5% significance level of the test. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013.

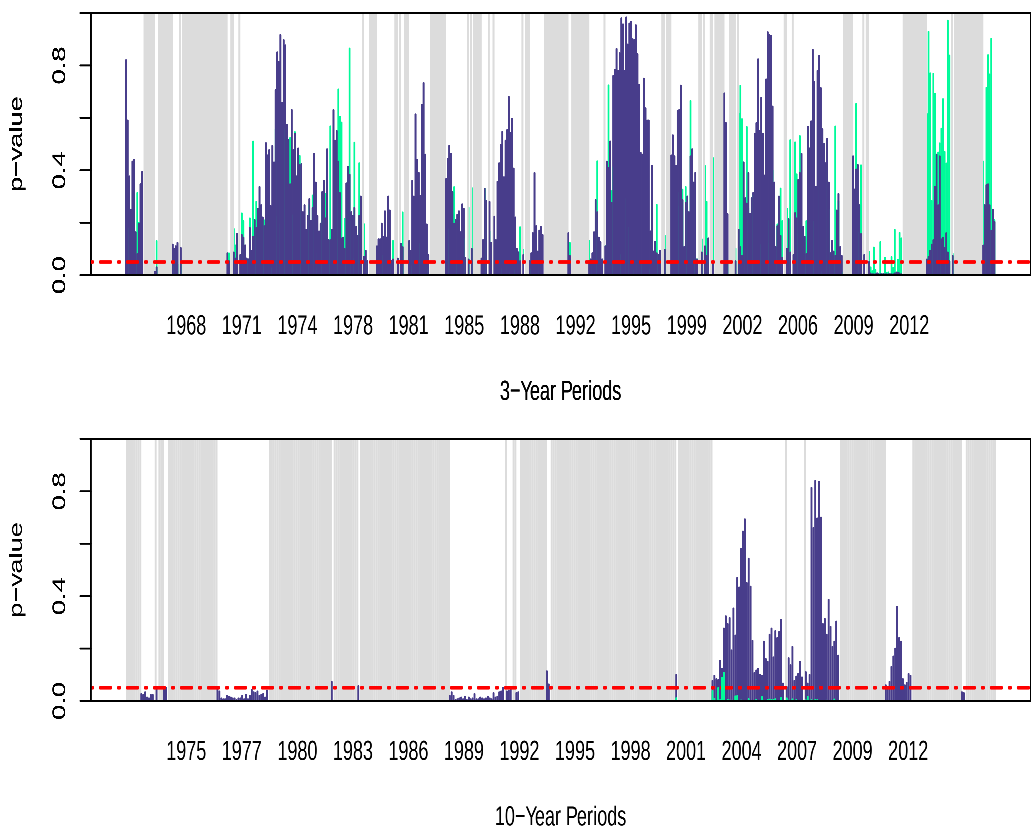


Figure IA.10
Specification testing for the Fama and French (1993) (FF3) three-factor model using the Gibbons, Ross, and Shanken (1989) and Gungor and Luger (2016) tests

The figure presents the time series of *p*-values of the GRS (blue line) and GL (green line) tests for FF3. Rolling time windows of three (top panel) and 10 years (bottom panel) are used. The dashed dotted red line denotes the 5% significance level of the tests. The gray bars are for the periods in which the GL test is inconclusive. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

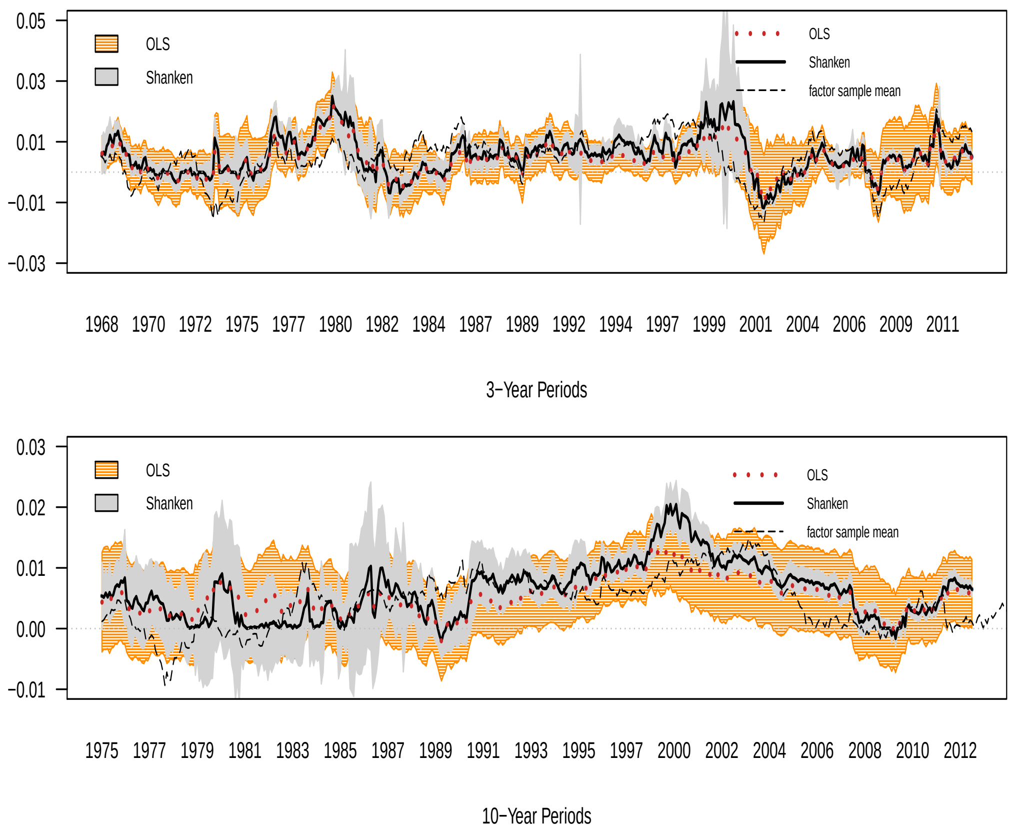


Figure IA.11
Estimates and confidence intervals for the market risk premium in the Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the market risk premium in FF3. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

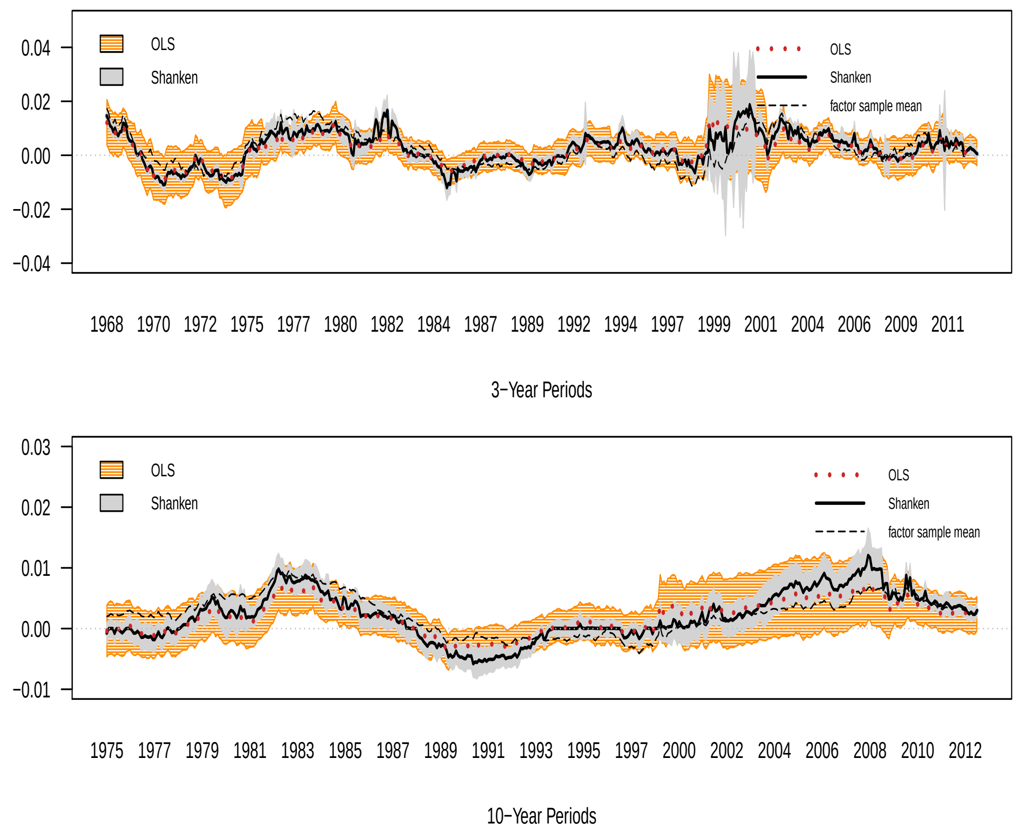


Figure IA.12
Estimates and confidence intervals for the size risk premium in the Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the size risk premium in FF3. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

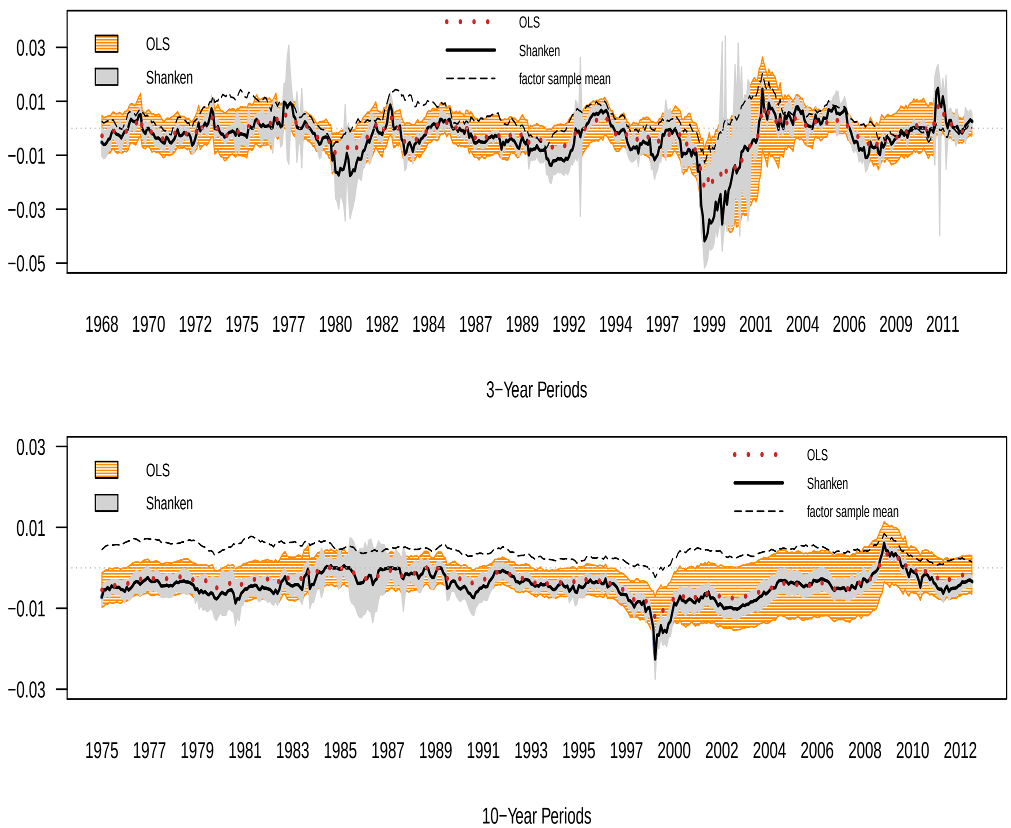


Figure IA.13
Estimates and confidence intervals for the value risk premium in the Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the value risk premium in FF3. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

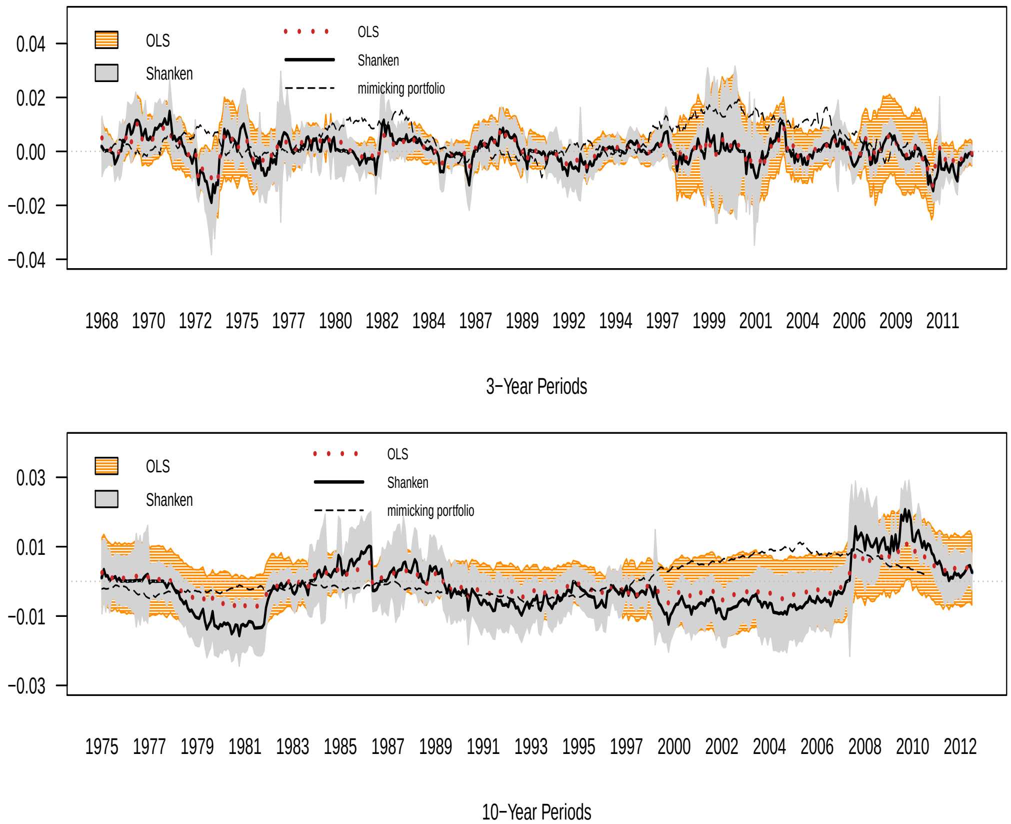


Figure IA.14
Estimates and confidence intervals for the liquidity risk premium in the liquidity-augmented Fama and French (1993) (FF3) three-factor model
The figure presents the estimates and the associated confidence intervals for the liquidity risk premium in the liquidity-augmented FF3 model. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the mimicking portfolio rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013.

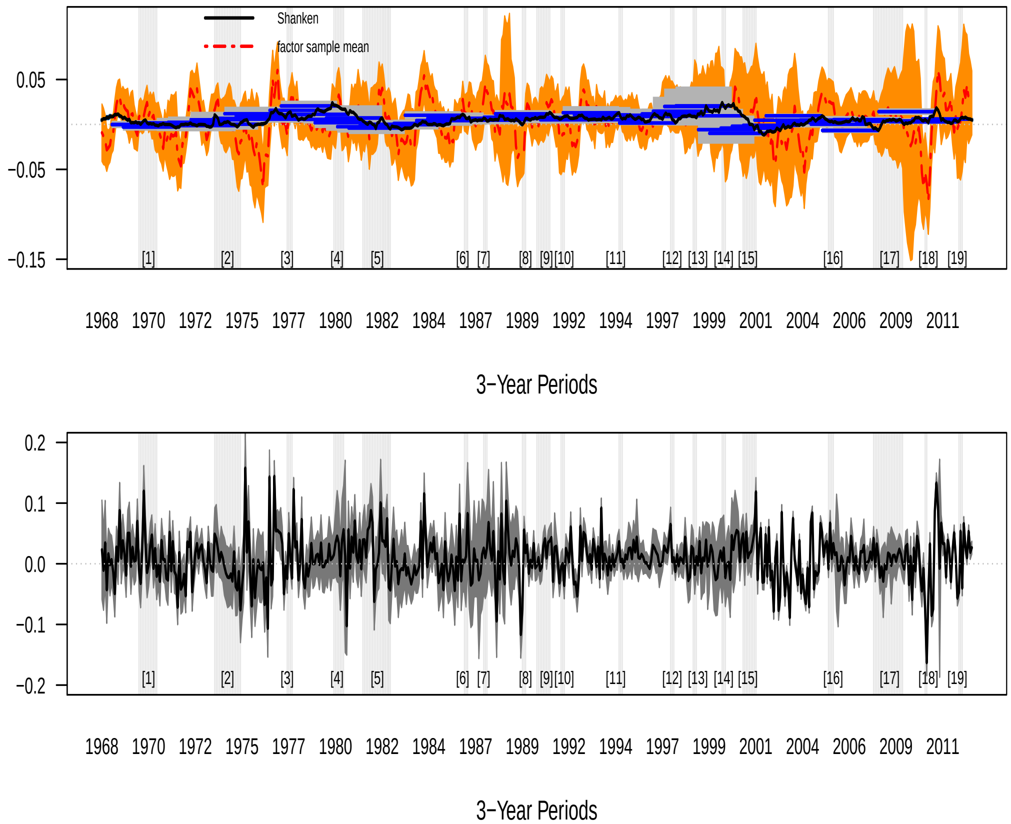
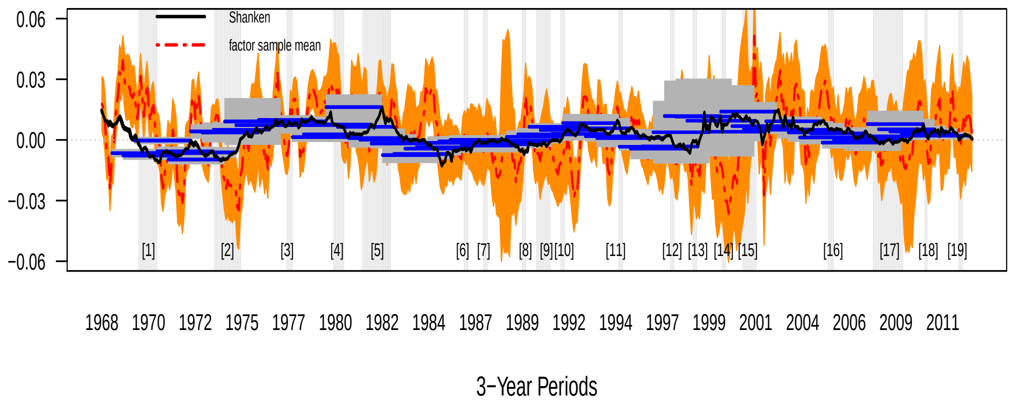


Figure IA.15
Estimates and confidence intervals for the time-varying market risk premium in the Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying market risk premium in FF3. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the market excess return (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the market excess return is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.



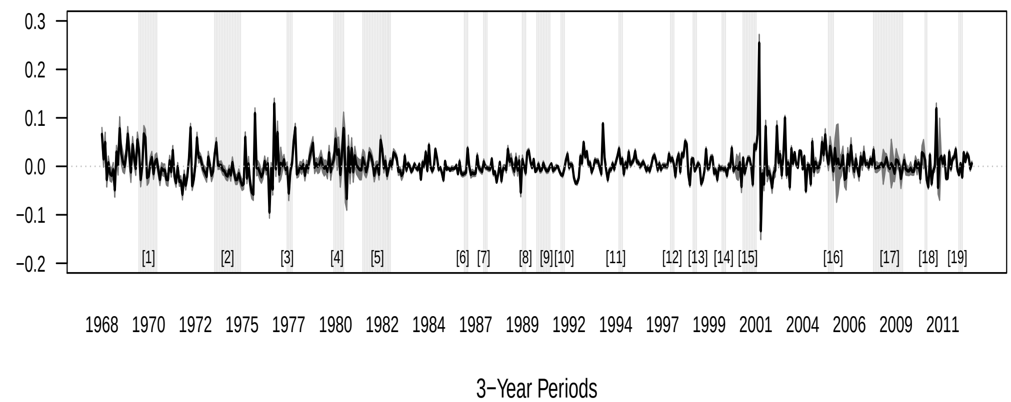


Figure IA.16
Estimates and confidence intervals for the time-varying size risk premium in the Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying size risk premium in FF3. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the size factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the size factor return spread is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

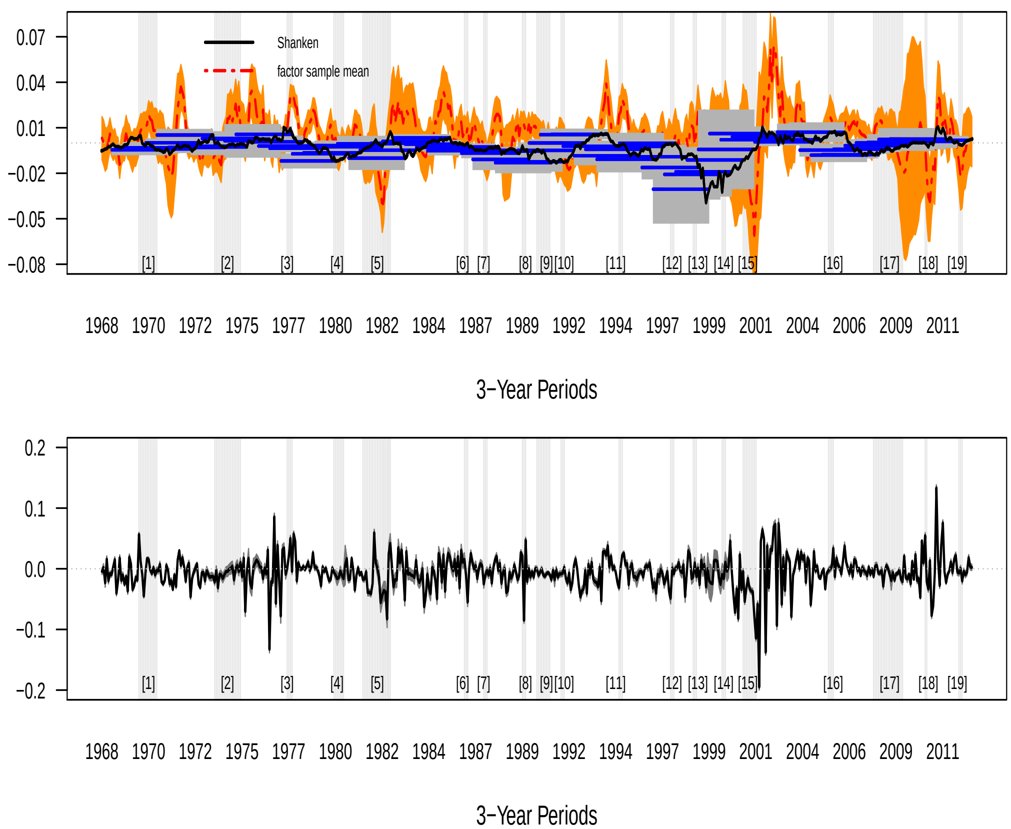


Figure IA.17
Estimates and confidence intervals for the time-varying value risk premium in the Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying value risk premium in FF3. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the value factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the value factor return spread is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

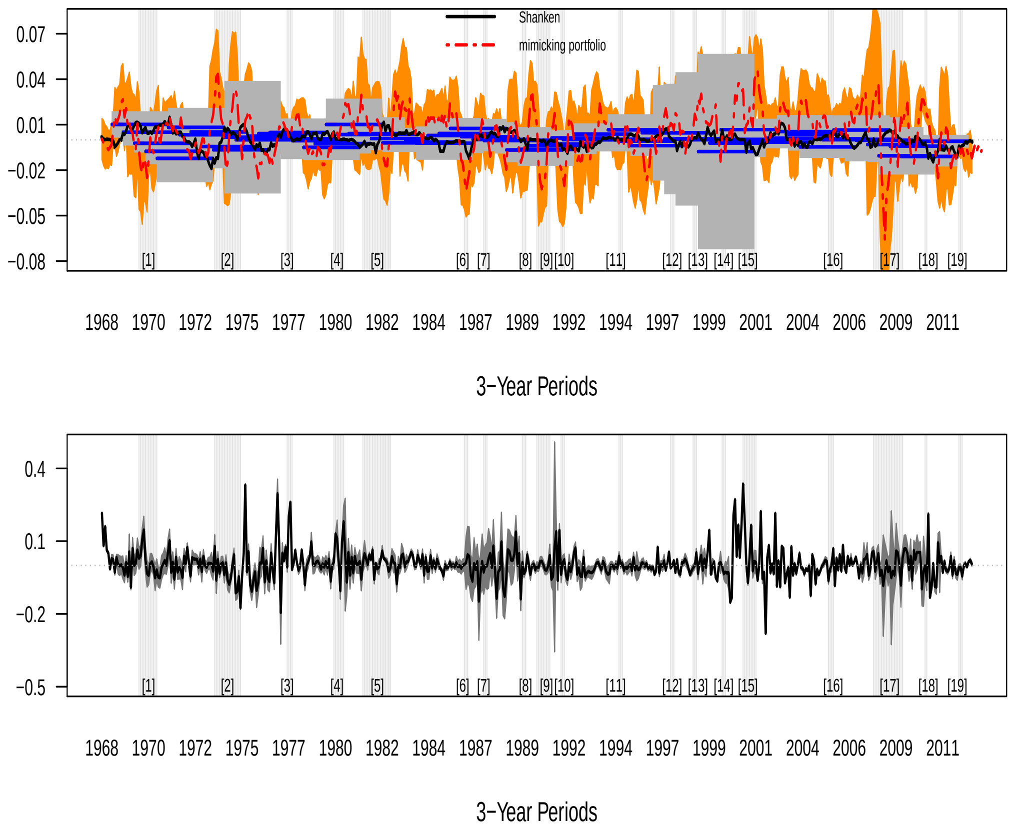


Figure IA.18
Estimates and confidence intervals for the time-varying liquidity risk premium in the liquidity-augmented Fama and French (1993) (FF3) three-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying liquidity risk premium in the liquidity-augmented FF3 model. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months) of the corresponding mimicking portfolio excess return (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

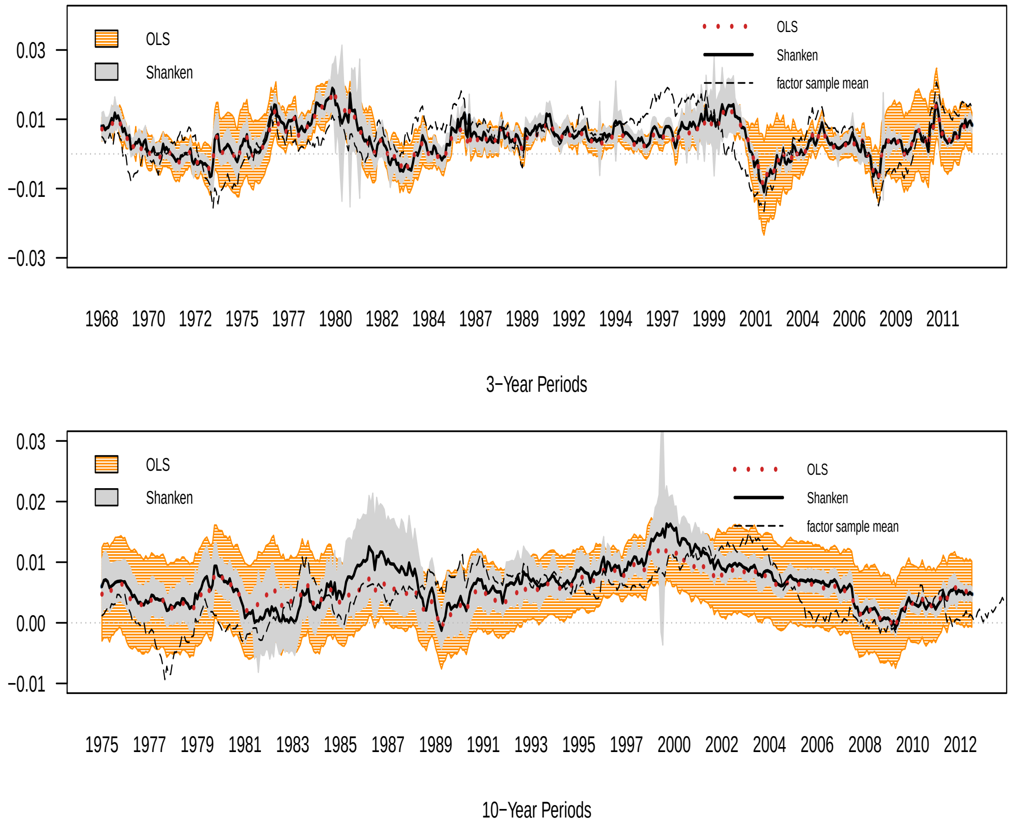
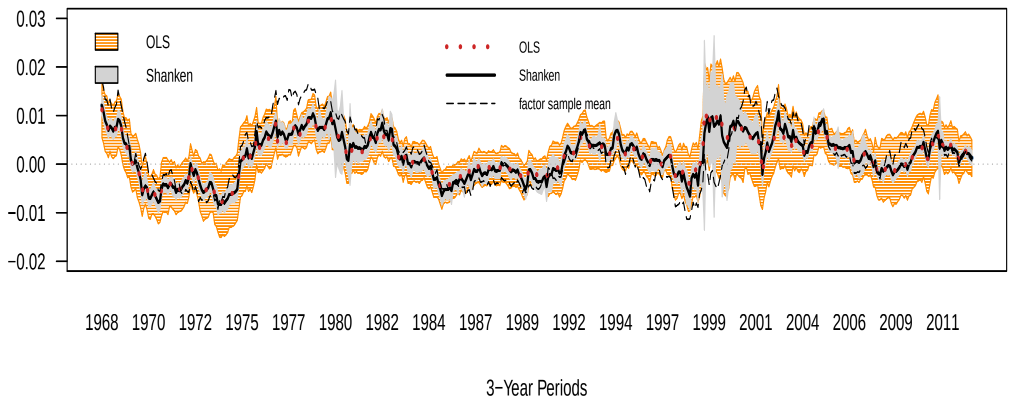


Figure IA.19
Estimates and confidence intervals for the market risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the market risk premium in FF5. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.



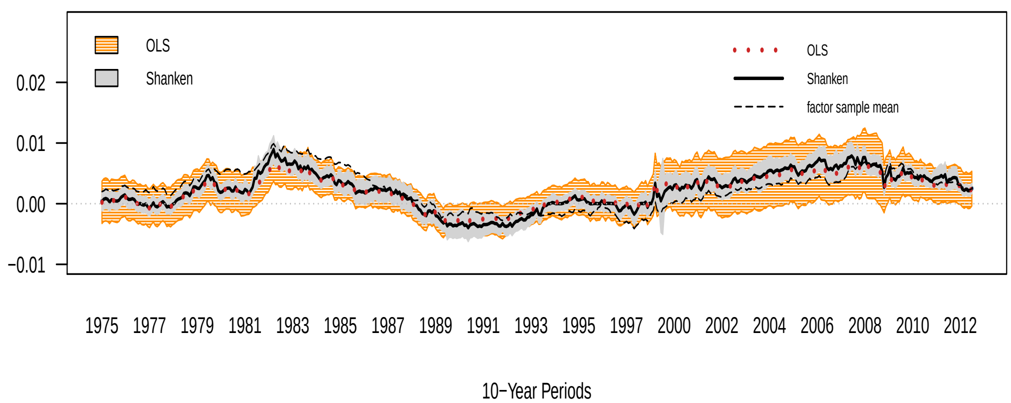


Figure IA.20
Estimates and confidence intervals for the size risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the size risk premium in FF5. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

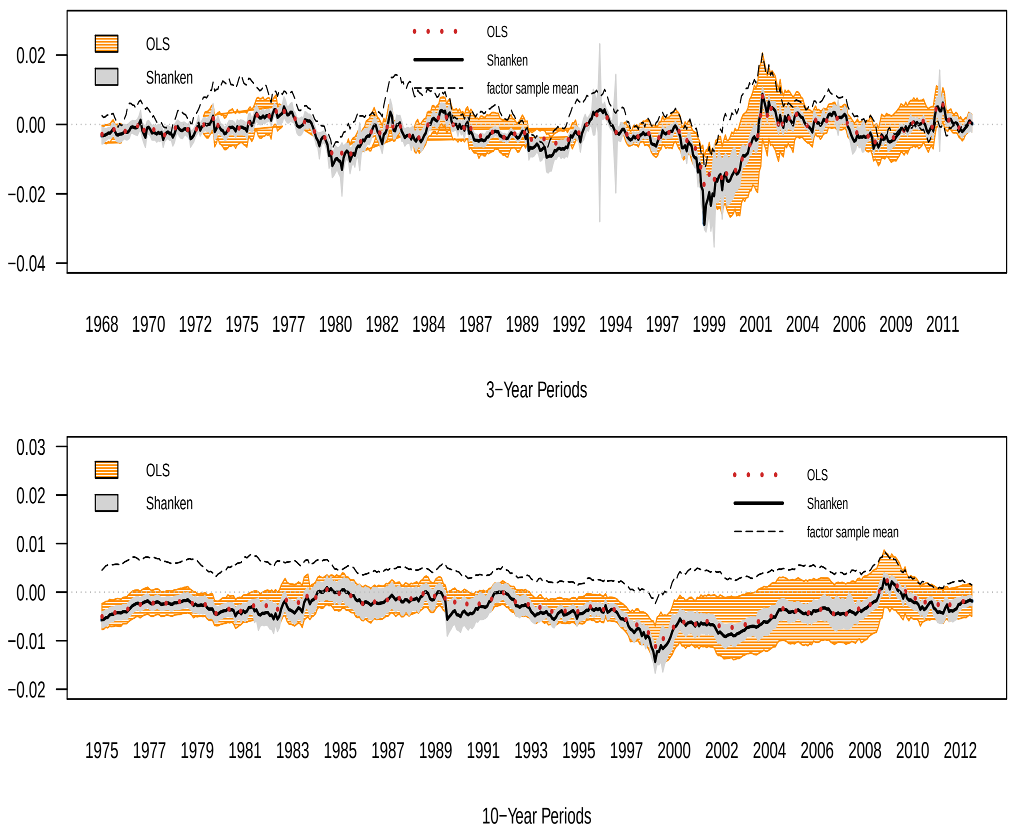


Figure IA.21
Estimates and confidence intervals for the value risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the value risk premium in FF5. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

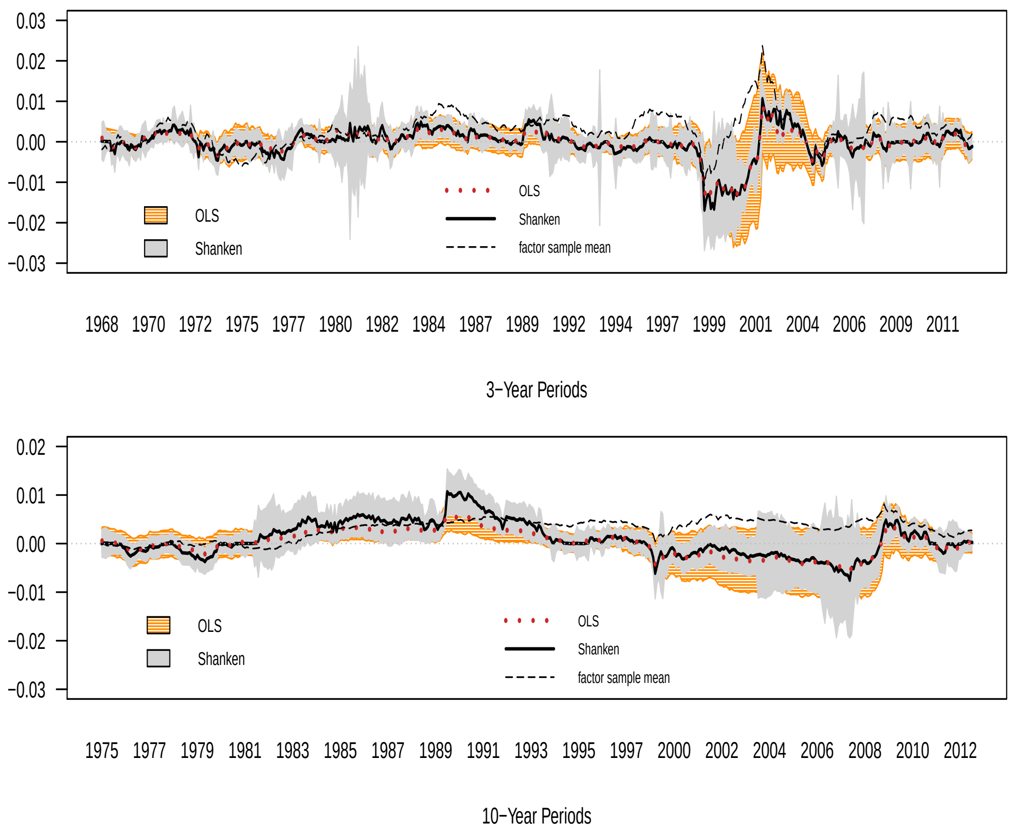


Figure IA.22
Estimates and confidence intervals for the profitability risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the profitability risk premium in FF5. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

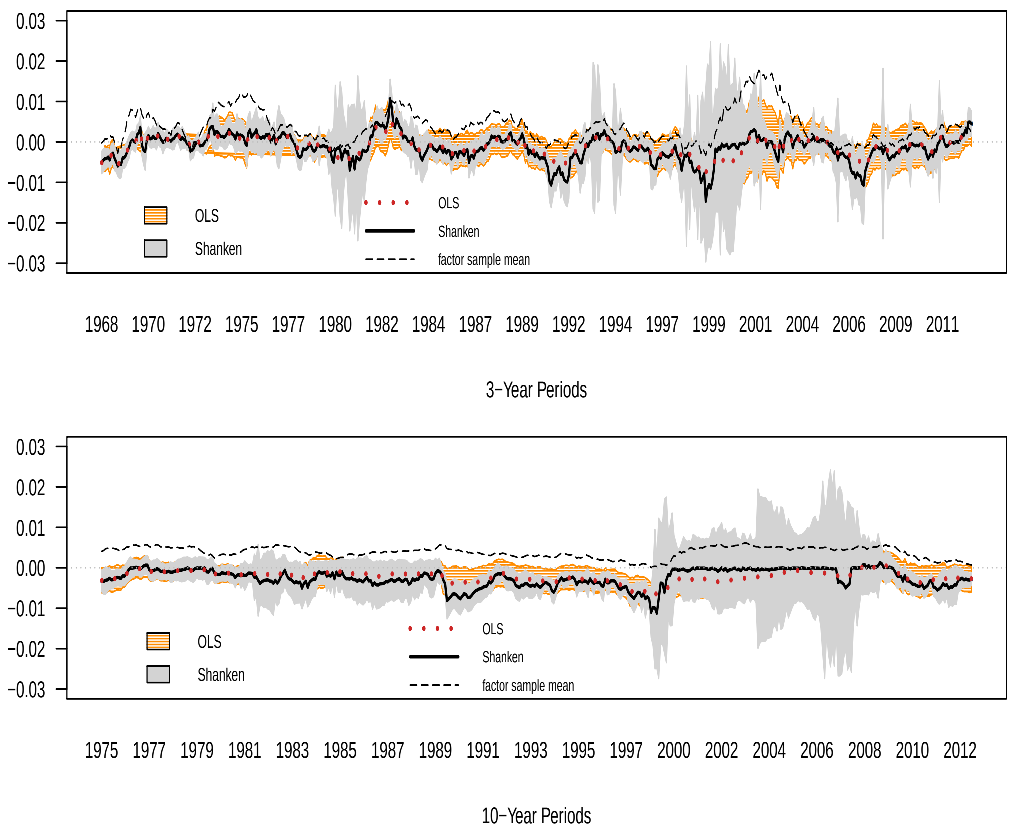


Figure IA.23
Estimates and confidence intervals for the investment risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the investment risk premium in FF5. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013.

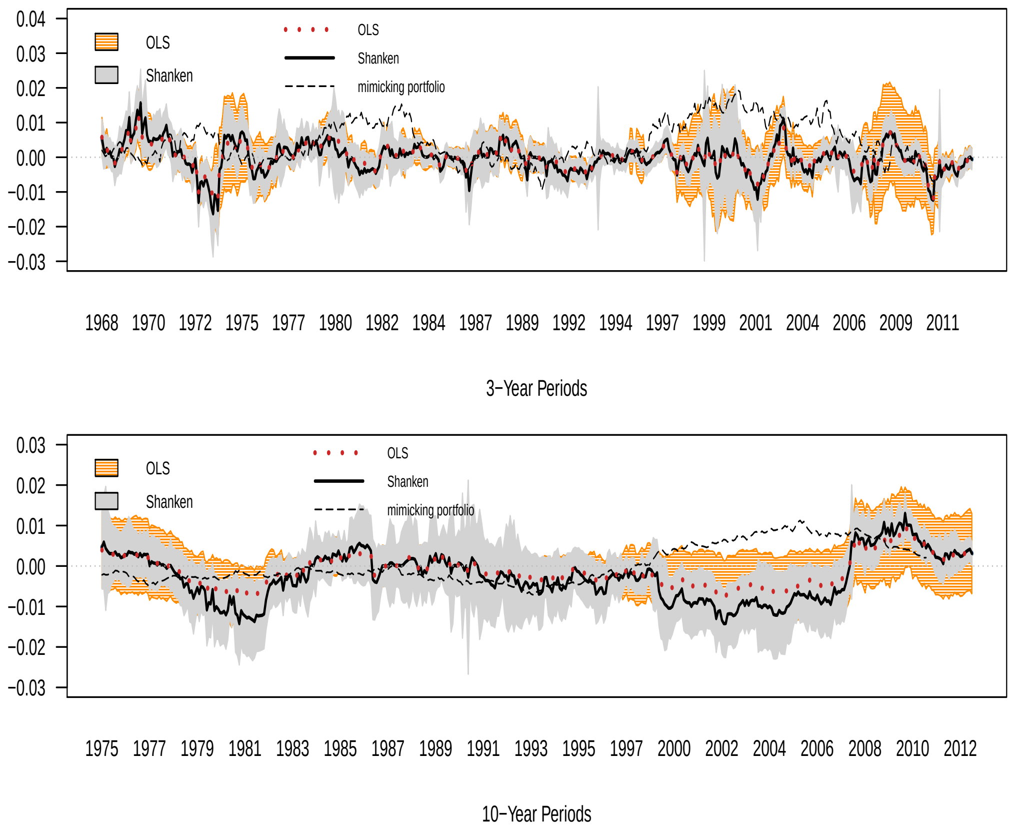


Figure IA.24
Estimates and confidence intervals for the liquidity risk premium in the liquidity-augmented Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the liquidity risk premium in the liquidity-augmented FF5 model. The bold black line is for the Shanken (1992) estimator. The corresponding gray band represents the 95% confidence intervals based on the large-N standard errors of Theorem 5. We also report the OLS CSR estimator (dotted red line) and the corresponding 95% confidence interval (striped orange band) based on the traditional large-T standard errors. Finally, the dashed black line is for the mimicking portfolio rolling factor sample mean. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013.

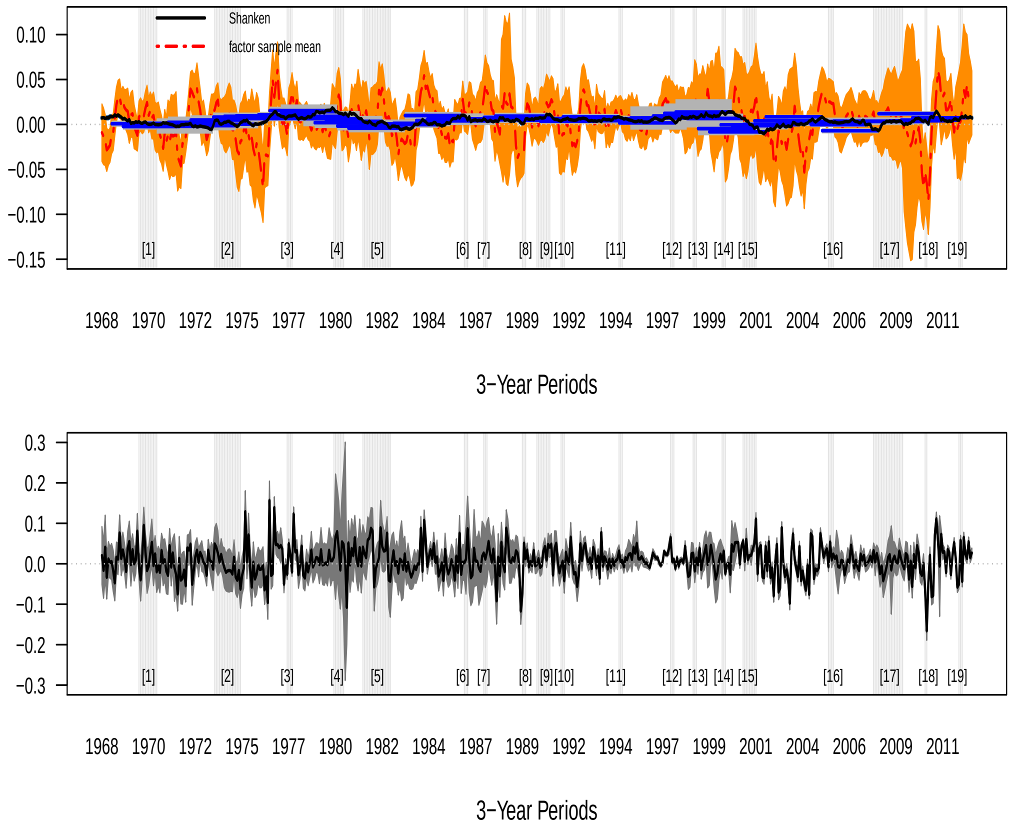


Figure IA.25
Estimates and confidence intervals for the time-varying market risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying market risk premium in FF5. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the size factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the market excess return is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

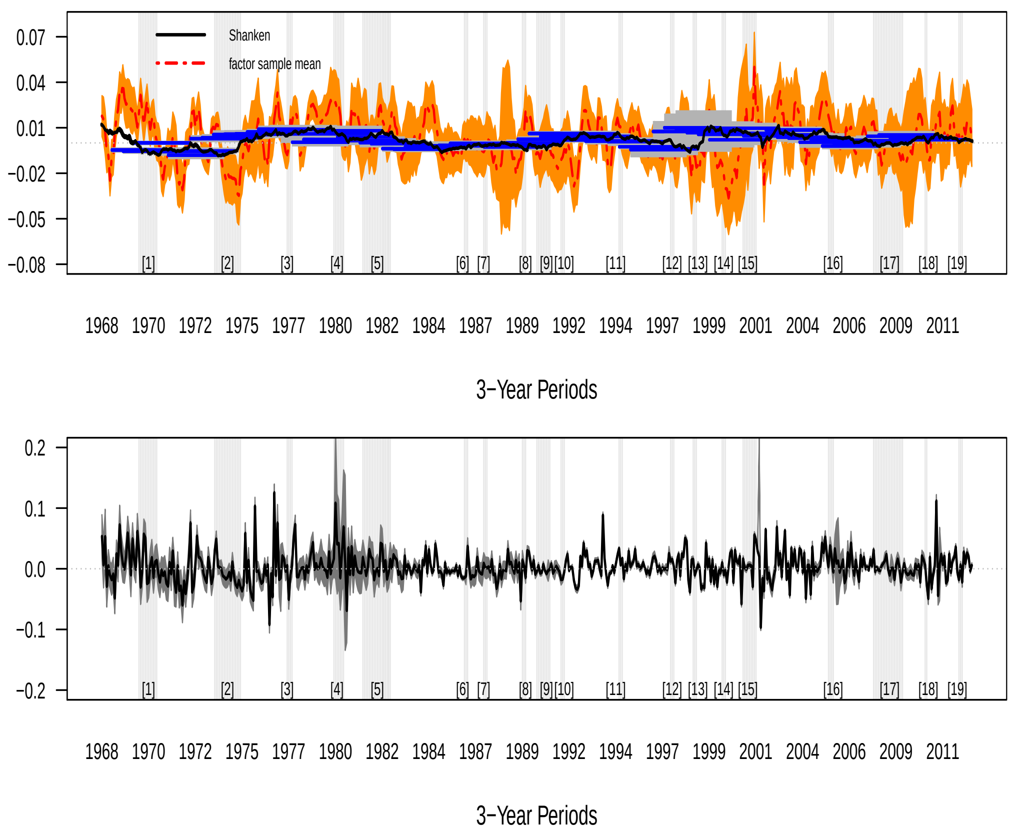


Figure IA.26
Estimates and confidence intervals for the time-varying size risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying size risk premium in FF5. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the size factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the size factor return spread is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

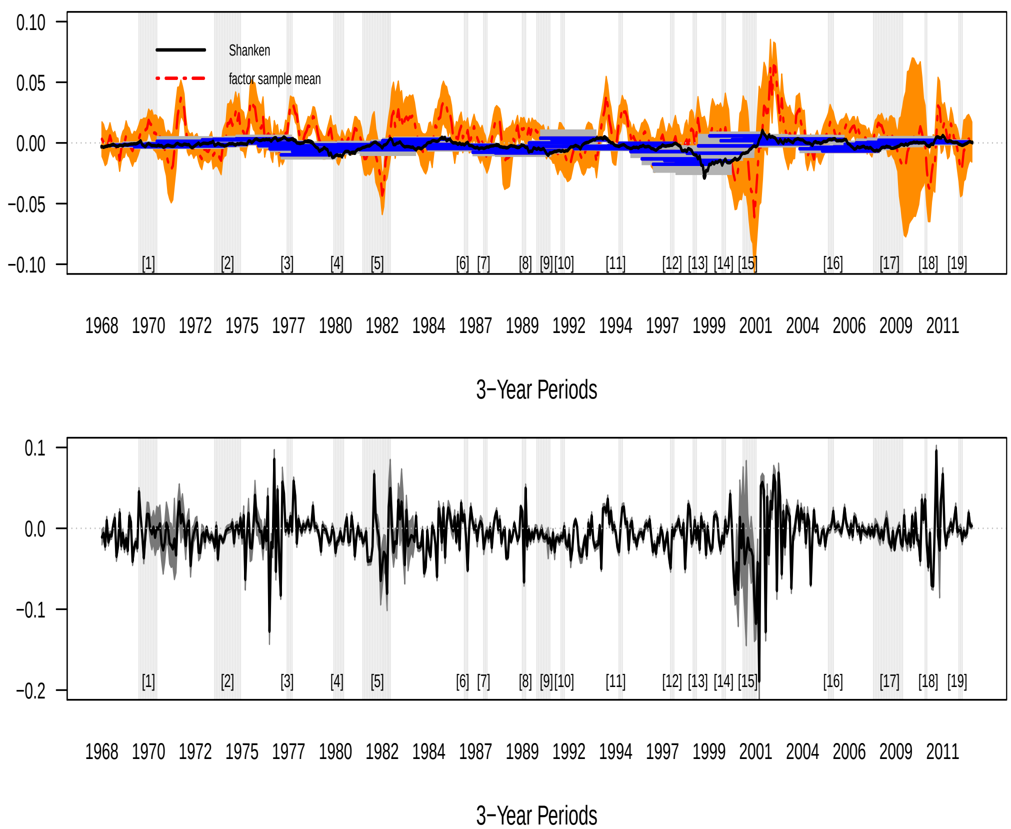


Figure IA.27
Estimates and confidence intervals for the time-varying value risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying value risk premium in FF5. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length T = 36 observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the value factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the value factor return spread is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

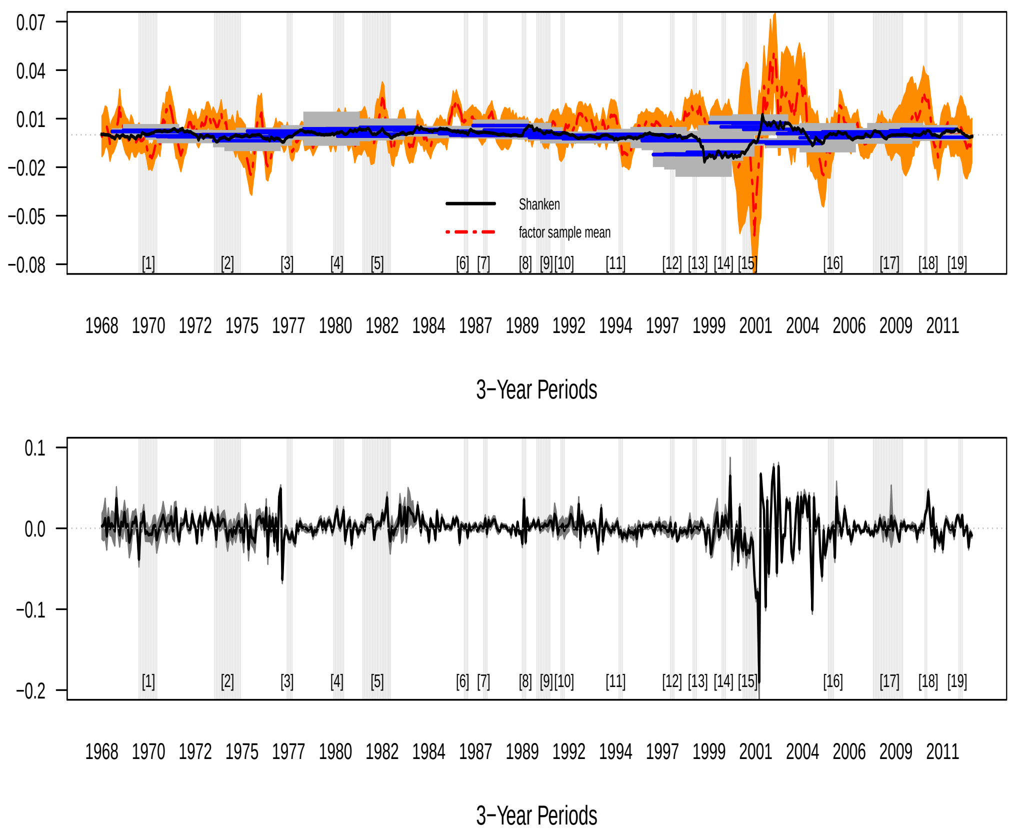


Figure IA.28
Estimates and confidence intervals for the time-varying profitability risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying profitability risk premium in FF5. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the profitability factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the profitability factor return spread is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

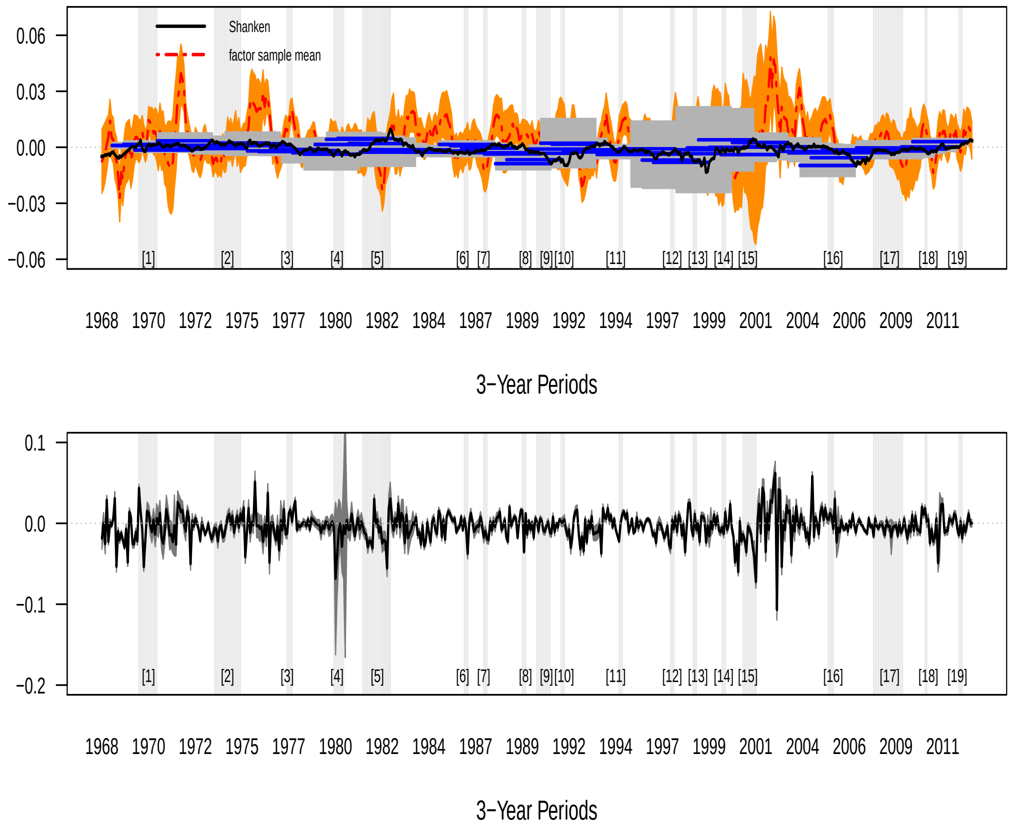
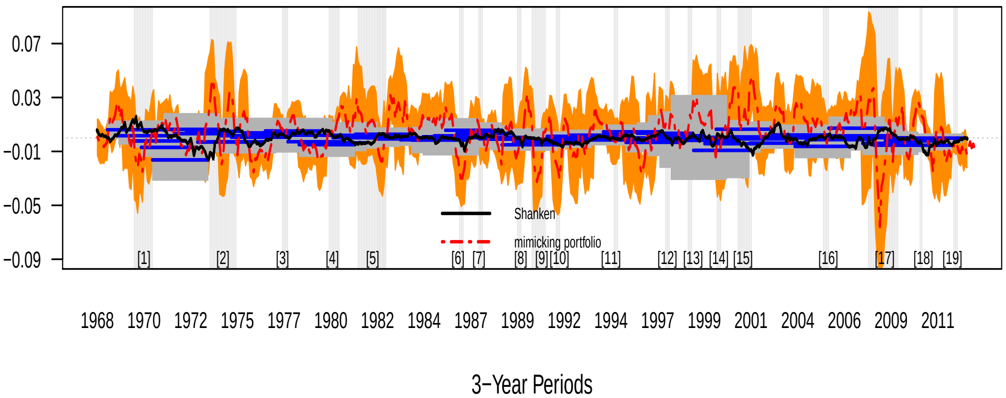


Figure IA.29
Estimates and confidence intervals for the time-varying investment risk premium in the Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying investment risk premium in FF5. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months of daily data) of the investment factor return spread (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's website from January 1966 to December 2013. The daily data on the investment factor return spread is from Kenneth French's website. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.



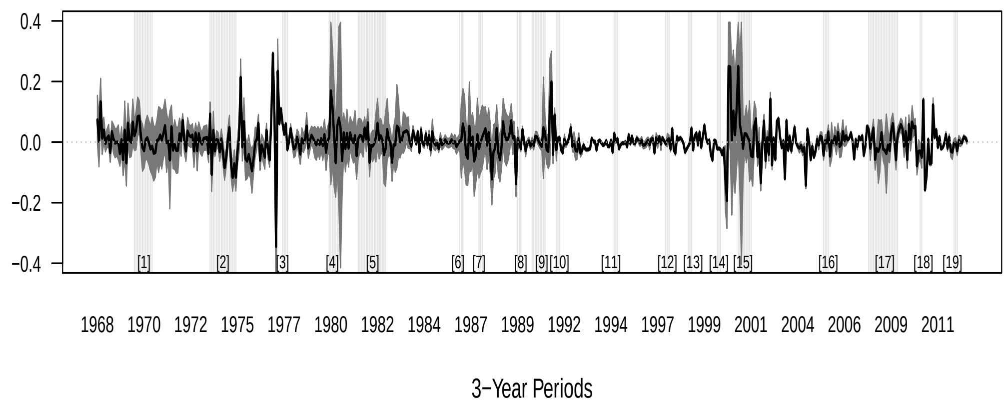


Figure IA.30

Estimates and confidence intervals for the time-varying liquidity risk premium in the liquidity-augmented Fama and French (2015) (FF5) five-factor model

The figure presents the estimates and the associated confidence intervals for the time-varying liquidity risk premium in the liquidity-augmented FF5 model. The top panel reports the Shanken (1992) large-N estimates, expressed in terms of a single line (black line) and in terms of horizontal bars of length $T = 36$ observations (blue line), with the corresponding 95% confidence intervals based on the large-N standard errors of Theorem 5 (gray band). We also report the rolling sample mean (using fixed rolling windows of six months) of the corresponding mimicking portfolio excess return (dashed dotted red line) and the corresponding 95% confidence interval (orange band). The bottom panel reports the modified Shanken (1992) estimator (black line) and the corresponding 95% confidence interval (gray band) based on the large-N standard errors of part (ii) of Theorem 3. We use monthly return data on individual stocks from CRSP and factor data from Kenneth French's and Ľuboš Pástor's websites from January 1966 to December 2013. The light gray bands correspond to the NBER recession dates and to various economic and financial crises. They are numbered as follows: [1] 1969:10-1970:11, [2] 1973:11-1975:3, [3] 1977:8-1977:11, [4] 1980:1-1980:7, [5] 1981:7-1982:11, [6] 1986:10-1986:12, [7] 1987:9-1987:11, [8] 1989:9-1989:12, [9] 1990:7-1991:3, [10] 1991:8-1992:12, [11] 1994:7-1994:10, [12] 1997:5-1997:9, [13] 1998:8-1998:10, [14] 2000:2-2000:4, [15] 2001:3-2001:11, [16] 2005:8-2005:11, [17] 2007:12-2009:6, [18] 2010:8-2010:10, [19] 2012:5-2012:7.

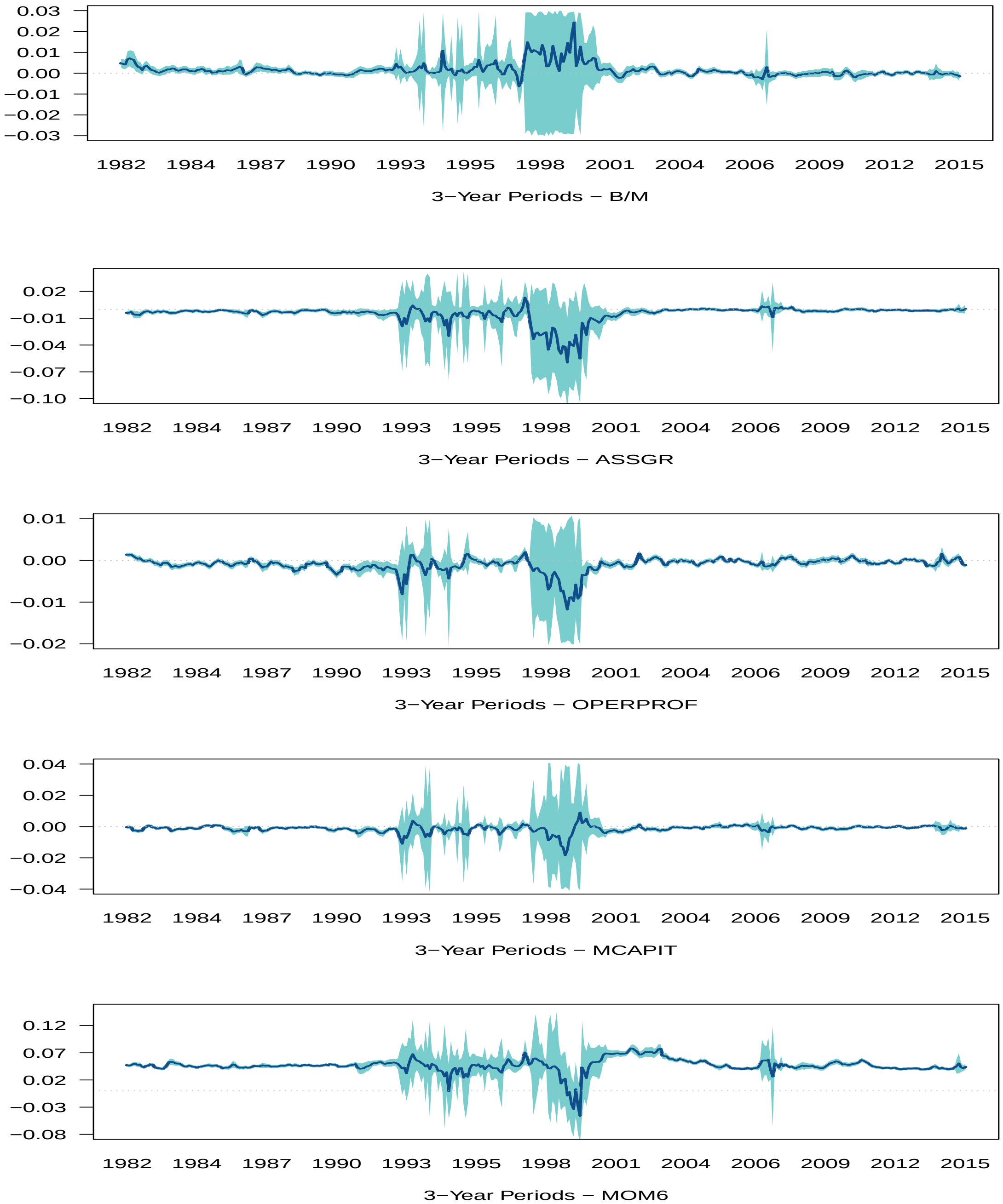


Figure IA.31
Estimates and confidence intervals for the characteristic premia in CAPM

The figure presents estimates (blue line) of the characteristic premia on the book-to-market ratio (B/M), asset growth (ASSGR), operating profitability (OPERPROF), market capitalization (MCAPIT), and six-month momentum (MOM6), and the associated confidence intervals based on Theorem 7 (light blue band), for the CAPM. The data is from DeMiguel et al. (forthcoming) and Kenneth French's website (from January 1980 to December 2015).

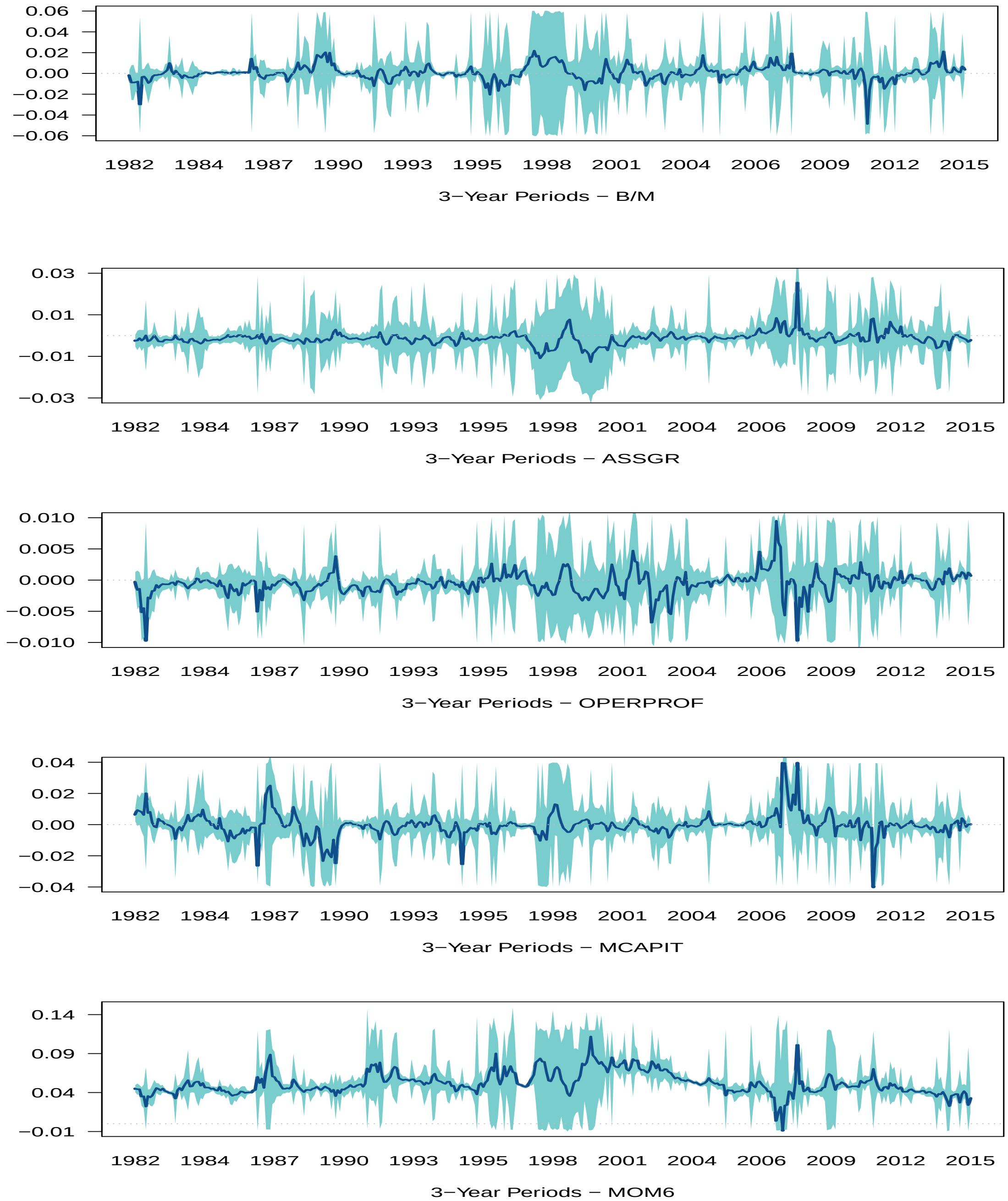


Figure IA.32
Estimates and confidence intervals for the characteristic premia in the Fama and French (1993) (FF3) three-factor model

The figure presents estimates (blue line) of the characteristic premia on the book-to-market ratio (B/M), asset growth (ASSGR), operating profitability (OPERPROF), market capitalization (MCAPIT), and six-month momentum (MOM6), and the associated confidence intervals based on Theorem 7 (light blue band), for the FF3 model. The data is from DeMiguel et al. (forthcoming) and Kenneth French's website (from January 1980 to December 2015).

References

Ang, A., and D. Kristensen. 2012. Testing conditional factor models. *Journal of Financial Economics* 106:132–56.

Brillinger, D. R. 2001. *Time series: Data analysis and theory*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal. Forthcoming. A transaction-cost perspective on the multitude of firm characteristics. *Review of Financial Studies*.

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

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

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

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

Gibbons, M. R., S. A. Ross, and J. Shanken. 1989. A test of the efficiency of a given portfolio. *Econometrica* 57:1121–52.

Gungor, S., and R. Luger. 2016. Multivariate tests of mean-variance efficiency and spanning with a large number of assets and time-varying covariances. *Journal of Business & Economic Statistics* 34:161–75.

Magnus, J. R., and H. Neudecker. 2007. *Matrix differential calculus with applications in statistics and econometrics*. Chichester, United Kingdom: Wiley.

Robinson, P. 1997. Large-sample inference for nonparametric regression with dependent errors. *Annals of Statistics* 25:2054–83.

Rubin, D. B. 1976. Inference and missing data. *Biometrika* 63:581–92.

Shanken, J. 1992. On the estimation of beta-pricing models. *Review of Financial Studies* 5:1–33.