Beta-Sorted Portfolios
Supplemental Appendix*

Matias D. Cattaneo† Richard K. Crump‡ Weining Wang§

November 8, 2024

Contents

*Cattaneo gratefully acknowledges financial support from the National Science Foundation (SES-1947662 and SES-2241575). Wang gratefully acknowledges financial support from the ESRC (Grant Reference: ES/T01573X/1).

†Department of Operations Research and Financial Engineering, Princeton University.
‡Macrofinance Studies, Federal Reserve Bank of New York.
§Faculty of Economics and Business, University of Groningen.

SA-1 First Step Estimator

Define $b_{it_0} = (\alpha_{it_0}, \beta_{it_0})^\top = \mathbb{E}(X_{t_0}X_{t_0}^\top|\mathcal{F}_{t_0-1})^{-1}\mathbb{E}(X_{t_0}R_{it_0}|\mathcal{F}_{t_0-1})$. We consider a slightly generalized first step estimator that allows for (one-sided) kernel weighting:

$$\widehat{b}_{it_0} = (\widehat{\alpha}_{it_0}, \widehat{\beta}_{it_0})^\top = \Big(\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)X_{t_0-s}X_{t_0-s}^\top\Big)^{-1}\Big(\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)X_{t_0-s}R_{i(t_0-s)}\Big),$$

where $H = \lfloor Th \rfloor$, $X_t = (1, f_t^\top)^\top$, and $K(.)$ is a kernel function satisfying the following assumption.

Assumption SA-1 (Kernel function)**.** $K(.) : [-1, 1] \mapsto \mathbb{R}_+$ *is Lipschitz continuous.*

SA-1.1 Preliminary Technical Lemma

For a $(m \times n)$-dimensional matrix $A = (a_{ij})_{1\le i\le m, 1\le j\le n}$, we define the induced matrix norms $|A|_1 = \max_{1\le j\le n} |\sum_{i=1}^m a_{i,j}|$, $|A|_2 = \max_{|v|=1} |Av|_2$, and $|A|_\infty = \max_{1\le i\le m} |\sum_{j=1}^n a_{i,j}|$.

To save notation, we define the following quantities:

$$A(t_0) = \frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)X_{t_0-s}X_{t_0-s}^\top, \qquad \tilde{A}(t_0) = \frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)\mathbb{E}[X_{t_0-s}X_{t_0-s}^\top|\mathcal{F}_{t_0-s-1}],$$

and

$$B_i(t_0) = \frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)X_{t_0-s}R_{i(t_0-s)}, \qquad \tilde{B}_i(t_0) = \frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)\mathbb{E}[X_{t_0-s}R_{it_0-s}|\mathcal{F}_{t_0-s-1}].$$

Lemma SA-1.1. *Suppose Assumptions 1–3 and SA-1 hold,* $\max_{1\le t\le T} n_t \lesssim n$, *and* $\log(nT)/(Th) \to 0$. *Then,*

$$\max_{\lfloor Th\rfloor+1\le t_0\le T} \big|A(t_0) - \mathbb{E}[A(t_0)]\big|_\infty \lesssim_{\mathbb{P}} \frac{T^{1/q}}{Th} + \sqrt{\frac{\log T}{Th}},$$

$$\max_{\lfloor Th\rfloor+1\le t_0\le T} \big|A(t_0) - \tilde{A}(t_0)\big|_\infty \lesssim_{\mathbb{P}} \frac{T^{1/q}}{Th} + \sqrt{\frac{\log T}{Th}},$$

$$\max_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}} \big|B_i(t_0) - \mathbb{E}[B_i(t_0)]\big|_\infty \lesssim_{\mathbb{P}} \frac{(nT)^{1/q}}{Th} + \sqrt{\frac{\log(nT)}{Th}},$$

$$\max_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}} \big|B_i(t_0) - \tilde{B}_i(t_0)\big|_\infty \lesssim_{\mathbb{P}} \frac{(nT)^{1/q}}{Th} + \sqrt{\frac{\log(nT)}{Th}}.$$

SA-1.2 Uniform Convergence Rate

The following theorem is a generalization of Theorem 3.1 in the paper.

Theorem SA-1.2. *Suppose the conditions of Lemma SA-1.1 hold, and $T^{2/q-1}n^{2/q}/h \to 0$. Then,*

$$\sup_{\lfloor Th \rfloor+1 \le t_0 \le T} \max_{1\le i\le n_{t_0}} \big|\widehat{b}_{it_0} - b_{it_0}\big|_\infty \lesssim_{\mathbb{P}} \sqrt{\frac{\log(nT)}{Th}} + h =: \mathsf{R}_{nT}.$$

The rate condition $T^{2/q-1}n^{2/q}/h \to 0$ implies that $\frac{T^{1/q}}{Th} \lesssim \sqrt{\frac{\log T}{Th}}$, and it favors higher moments. For example, it implies $\frac{n^{1/4}}{T^{3/4}h} \to 0$ when $q = 8$.

SA-1.3 Asymptotic Normality

We provide an asymptotic normality result for $\widehat{b}_{it_0}$, assuming conditional homoskedasticity (but possibly time-varying) for simplicity. This result is not used in the main paper, but is reported here for completeness.

Theorem SA-1.3. *Suppose the conditions of Theorem SA-1.2 hold,* $\mathbb{V}\mathrm{ar}[\varepsilon_{it_0}|\mathcal{F}_{t_0-s-1}] = \sigma^2_{\varepsilon,t_0}$ *is a non-random constant, and $Th^3 \to 0$. Then,*

$$\sqrt{Th}\Sigma_b(t_0)^{-1/2}\big(\widehat{b}_{it_0} - b_{it_0}\big) \to_{\mathcal{L}} \mathsf{N}(0, I), \qquad \Sigma_b(t_0) = \Sigma_A(t_0)^{-1}\Sigma_B(t_0)\Sigma_A(t_0)^{-1},$$

where $\Sigma_A(t_0) = \mathbb{E}[X_{t_0}X_{t_0}^\top]\int_{-1}^0 K(s)ds$ *and* $\Sigma_B(t_0) = \sigma^2_{\varepsilon,t_0}\mathbb{E}[X_{t_0}X_{t_0}^\top]\int_{-1}^0 K^2(s)ds$.

A plug-in consistent estimator of the asymptotic variance can be constructed using estimated residuals:

$$(\widehat{\sigma}^2_{\varepsilon,t_0}, \widehat{\varsigma}^2_{t_0})^\top = \arg\min_{c_0,c_1} \sum_{t=1}^{t_0-1}\sum_{i=1}^{n_t} K\Big(\frac{t-t_0}{Th}\Big)\big(\widehat{u}^2_{it} - c_0 - c_1(t-t_0)/T\big)^2,$$

where $\widehat{u}_{it}$ denote the estimated residuals.

SA-1.4 Proof of Lemma SA-1.1

We assume that $d = 1$ without loss of generality. For the first result, we have

$$\max_{\lfloor Th \rfloor+1 \le t_0 \le T} \big|A(t_0) - \mathbb{E}[A(t_0)]\big|_\infty \lesssim \mathfrak{R}_1 + \mathfrak{R}_2 + \mathfrak{R}_3,$$

where

$$\mathfrak{R}_1 = \max_{\lfloor Th \rfloor+1 \le t_0 \le T}\Big|\frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)\big(z^2_{t_0-s} - \mathbb{E}[z^2_{t_0-s}]\big)\Big|,$$

$$\mathfrak{R}_2 = \max_{\lfloor Th \rfloor+1 \le t_0 \le T}\Big|\frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big) z_{t_0-s}\tau\Big(\frac{t_0-s}{T}\Big)\Big|,$$

$$\mathfrak{R}_3 = \max_{\lfloor Th \rfloor+1 \le t_0 \le T}\Big|\frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big) z_{t_0-s}\Big|.$$

In addition, Assumption 2 implies that (i) $\{z_{t-s}^2 - \mathbb{E}[z_{t-s}^2]\}_s$, $\{z_{t-s}\tau\big(\frac{t-s}{T}\big)\}_s$, and z_t are mean zero and square-integrable, and (ii) $\Theta(z_\bullet^2; q, v) \lesssim 1$ with $v > 1/2 - 2/q$. Therefore, Lemma A.3 in Zhang and Wu (2012) implies that $\mathfrak{R}_1 + \mathfrak{R}_2 + \mathfrak{R}_3 \lesssim_{\mathbb{P}} \frac{T^{1/q}}{Th} + \sqrt{\frac{\log T}{Th}}$.

For the second result, we have

$$\max_{\lfloor Th \rfloor+1 \le t_0 \le T} \big|A(t_0) - \tilde{A}(t_0)\big|_\infty \lesssim \mathfrak{R}_4 + \mathfrak{R}_5 + \mathfrak{R}_6,$$

where

$$\begin{aligned}
\mathfrak{R}_4 &= \max_{\lfloor Th \rfloor+1 \le t_0 \le T} \Big|\frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)\big(z_{t_0-s}^2 - \mathbb{E}[z_{t_0-s}^2|\mathcal{F}_{t_0-s-1}]\big)\Big|,\\
\mathfrak{R}_5 &= \max_{\lfloor Th \rfloor+1 \le t_0 \le T} \Big|\frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)(z_{t_0-s} - \mathbb{E}[z_{t_0-s}|\mathcal{F}_{t_0-s-1}])\tau\Big(\frac{t_0-s}{T}\Big)\Big|,\\
\mathfrak{R}_6 &= \max_{\lfloor Th \rfloor+1 \le t_0 \le T} \Big|\frac{1}{Th}\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big)(z_{t_0-s} - \mathbb{E}[z_{t_0-s}|\mathcal{F}_{t_0-s-1}])\Big|.
\end{aligned}$$

These terms form a martingale difference sequence. Thus, we let u_s be equal to either $z_{t_0-s}^2 - \mathbb{E}[z_{t_0-s}^2]$, $(z_{t_0-s} - \mathbb{E}[z_{t_0-s}|\mathcal{F}_{t_0-s-1}])\tau\big(\frac{t_0-s}{T}\big)$, or $z_{t_0-s} - \mathbb{E}[z_{t_0-s}|\mathcal{F}_{t_0-s-1}]$, to save notation. Then, using summation by part

$$\begin{aligned}
&\max_{H+1 \le t_0 \le T} \Big|\sum_{s=1}^{H} K\Big(\frac{s}{Th}\Big) u_{t_0-s}\Big|\\
&\le \max_{H+1 \le t_0 \le T} \Big(\max_{t_0-H \le \ell \le t_0-1} \Big|\sum_{s=t_0-H}^{\ell} u_s\Big|\Big) \sum_{s=t_0-H+1}^{t_0-2} \Big|K\Big(\frac{t_0-s}{Th}\Big) - K\Big(\frac{t_0-s-1}{Th}\Big)\Big|\\
&\quad + \max_{H+1 \le t_0 \le T} \Big|\sum_{s=t_0-H}^{t_0-1} u_s\Big| K\Big(\frac{1}{Th}\Big)\\
&\lesssim \max_{H+1 \le t_0 \le T} \max_{t_0-H \le \ell \le t_0-1} \Big|\sum_{s=t_0-H}^{\ell} u_s\Big|,
\end{aligned}$$

because

$$\sum_{s=t_0-H+1}^{t_0-1} \Big|K\Big(\frac{t_0-t}{Th}\Big) - K\Big(\frac{t_0-s+1}{Th}\Big)\Big| \lesssim 1,$$

by the Lipschitz continuity of the kernel function. Let λ be a sufficient large positive constant. Then, Hall and Heyde (2014, Theorem 2.4) implies that

$$\mathbb{P}\Big(\max_{t_0-H \le \ell \le t_0-1} \Big|\sum_{t=t_0-H}^{\ell} u_t\Big| \ge 2\lambda\Big) \lesssim (2 \vee (\sqrt{Th} \max_{t_0-H \le t \le t_0-1} \|u_t\|_2/\lambda))\mathbb{P}\Big(\Big|\sum_{s=t_0-H}^{t_0-1} u_s\Big| \ge \lambda\Big).$$

Thus, it suffices to look at H blocks of observations. We have

$$\max_{T-H+1\leq t_0\leq T}\max_{t_0-H\leq \ell\leq t_0-1}\Big|\sum_{s=1}^{\ell}u_{t_0-s}\Big|\lesssim \max_{t_0\in 2H,3H,\cdots,T-H}\Big|\sum_{s=t_0-H}^{t_0}u_s\Big|,$$

and applying Freedman's inequality (Freedman, 1975) and the union bound, we can verify that $\mathfrak{R}_4+\mathfrak{R}_5+\mathfrak{R}_6\lesssim_{\mathbb{P}}\frac{T^{1/q}}{Th}+\sqrt{\frac{\log T}{Th}}$.

The last two conclusions follow analogously, using martingale methods and a union bound over i. For example, consider the fourth conclusion. We have

$$\begin{aligned}
&\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|B_i(t_0)-\tilde{B}_i(t_0)\big|_\infty\\
&\leq \max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\Big|\frac{1}{Th}\sum_{s=1}^{H}K\Big(\frac{s}{Th}\Big)X_{t_0-s}\varepsilon_{it_0-s}\Big|_\infty\\
&\quad+\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|\frac{1}{Th}\sum_{s=1}^{H}K\Big(\frac{s}{Th}\Big)(X_{t_0-s}X_{t_0-s}^\top-\mathbb{E}[X_{t_0-s}X_{t_0-s}^\top|\mathcal{F}_{t_0-s-1}])b_{it_0-s}\big|_\infty.
\end{aligned}$$

The term $\sum_{s=1}^{H}K\big(\frac{s}{Th}\big)X_{t_0-s}\varepsilon_{it_0-s}$ a martingale difference sequence, and therefore proceeding as above we verify the desired result. The second term of the upper bound in the preceding display is bounded similarly. □

SA-1.5 Proof of Theorem SA-1.2

We have $\widehat{b}_{it_0}-b_{it_0}=A(t_0)^{-1}B_i(t_0)-\tilde{A}(t_0)^{-1}B_i(t_0)+\tilde{A}(t_0)^{-1}B_i(t_0)-\tilde{A}(t_0)^{-1}\tilde{A}(t_0)b_{it_0}$, and therefore

$$\sup_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|\widehat{b}_{it_0}-b_{it_0}\big|_\infty\lesssim_{\mathbb{P}}\mathfrak{R}_1+\mathfrak{R}_2$$

where

$$\begin{aligned}
\mathfrak{R}_1&=\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|A(t_0)^{-1}B_i(t_0)-\tilde{A}(t_0)^{-1}\tilde{B}_i(t_0)\big|_\infty\\
\mathfrak{R}_2&=\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|\tilde{B}_i(t_0)-\tilde{A}(t_0)b_{it_0}\big|_\infty
\end{aligned}$$

because $\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\big|\tilde{A}(t_0)^{-1}\big|_\infty\lesssim_{\mathbb{P}}1$.

For the first term, we have

$$\begin{aligned}
\mathfrak{R}_1\leq&\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|A(t_0)^{-1}\big|_\infty\big|A(t_0)-\tilde{A}(t_0)\big|_\infty\big|\tilde{A}(t_0)^{-1}\big|_\infty\big|B_i(t_0)\big|_\infty\\
&+\max_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\big|\tilde{A}(t_0)^{-1}\big|_\infty\big|B_i(t_0)-\tilde{B}_i(t_0)\big|_\infty\\
\lesssim_{\mathbb{P}}&\frac{(nT)^{1/q}}{Th}+\sqrt{\frac{\log(nT)}{Th}}.
\end{aligned}$$

For the second term, we have

$$
\begin{aligned}
\mathfrak{R}_2 &= \max_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}\left|\frac{1}{Th}\sum_{s=1}^{H}K\Big(\frac{s}{Th}\Big)\mathbb{E}(X_{t_0-s}X_{t_0-t}^\top|\mathcal{F}_{t_0-s-1})(b_{i(t_0-s)}-b_{it_0})\right|\\
&\le \max_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}\left|b_{i(t_0-H)}-b_{it_0}\right|\left|\frac{1}{Th}\sum_{s=1}^{H}K\Big(\frac{s}{Th}\Big)\mathbb{E}(X_{t_0-s}X_{t_0-t}^\top|\mathcal{F}_{t_0-s-1})\right|\\
&\quad+\max_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}\left|\frac{1}{Th}\sum_{s=1}^{H-1}(b_{i(t_0-s)}-b_{i(t_0-s-1)})\sum_{j=1}^{s}K\Big(\frac{j}{Th}\Big)\mathbb{E}(X_{t_0-j}X_{t_0-j}^\top|\mathcal{F}_{t_0-j-1})\right|\\
&\lesssim_{\mathbb{P}} h + h\max_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}\max_{1\le s\le (H-1)}\left|\frac{1}{Th}\sum_{j=1}^{s}K\Big(\frac{j}{Th}\Big)\mathbb{E}(X_{t_0-j}X_{t_0-j}^\top|\mathcal{F}_{t_0-j-1})\right|\lesssim_{\mathbb{P}} h.
\end{aligned}
$$

This completes the proof. $\square$

SA-1.6 Proof of Theorem SA-1.3

From previous results, the bias is

$$\sup_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}\left|\tilde{A}^{-1}(\tilde{B}_i(t_0)-\tilde{A}(t_0)b_{it_0})\right|_\infty\lesssim_{\mathbb{P}} h,$$

and therefore we have the decomposition:

$$
\begin{aligned}
\widehat{b}_{it_0}-b_{it_0} &= A(t_0)^{-1}B_i(t_0)-\tilde{A}(t_0)^{-1}\tilde{B}_i(t_0)+O_{\mathbb{P}}(h)\\
&= A(t_0)^{-1}(A(t_0)-\tilde{A}(t_0))\tilde{A}(t_0)^{-1}B_i(t_0)+A(t_0)^{-1}(B_i(t_0)-\tilde{B}_i(t_0))+O_{\mathbb{P}}(h),\\
&= \mathfrak{R}_{1,it_0}-\mathfrak{R}_{2,it_0}+O_{\mathbb{P}}(h),
\end{aligned}
$$

with

$$
\begin{aligned}
\mathfrak{R}_{1,it_0} &= -A(t_0)^{-1}(A(t_0)-\tilde{A}(t_0))\tilde{A}(t_0)^{-1}(B_i(t_0)-\tilde{A}(t_0)b_{it_0}),\\
\mathfrak{R}_{2,it_0} &= -A(t_0)^{-1}\{(A(t_0)-\tilde{A}(t_0))b_{it_0}-(B_i(t_0)-\tilde{B}_i(t_0))\}.
\end{aligned}
$$

Similar to the proof of Theorem SA-1.2, we have

$$\sup_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}|\mathfrak{R}_{1,it_0}|_\infty\lesssim_{\mathbb{P}} h\left(\frac{(nT)^{1/q}}{Th}+\sqrt{\frac{\log(nT)}{Th}}\right)=o_{\mathbb{P}}\Big(\frac{1}{\sqrt{Th}}\Big),$$

under the rate conditions imposed.

Next, observe that

$$\sup_{\lfloor Th\rfloor+1\le t_0\le T}\max_{1\le i\le n_{t_0}}\left|\mathfrak{R}_{2,it_0}+A(t_0)^{-1}\sum_{s=1}^{H}K\Big(\frac{s}{Th}\Big)X_{t_0-s}\varepsilon_{it_0-s}\right|_\infty=o_{\mathbb{P}}\Big(\frac{1}{\sqrt{Th}}\Big),$$

because

$$\sup_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\Big|\frac{1}{Th}\sum_{s=1}^{H}K\Big(\frac{s}{Th}\Big)(X_{t_0-s}X_{t_0-s}^\top-\mathbb{E}[X_{t_0-s}X_{t_0-s}^\top|\mathcal{F}_{t_0-s-1}])b_{it_0-s}-(A(t_0)-\tilde{A}(t_0))b_{it_0}\Big|_\infty$$
$$\lesssim_{\mathbb{P}}\sup_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}|A(t_0)-\tilde{A}(t_0)|_\infty h=o_{\mathbb{P}}\Big(\frac{1}{\sqrt{Th}}\Big).$$

Therefore, putting the results above together and proceeding as before, we have

$$\sup_{\lfloor Th\rfloor+1\leq t_0\leq T}\max_{1\leq i\leq n_{t_0}}\Big|\sqrt{Th}(\widehat{b}_{it_0}-b_{it_0})+A(t_0)^{-1}\frac{1}{\sqrt{Th}}\sum_{s=1}^{H}K\Big(\frac{t}{Th}\Big)X_{t_0-t}\varepsilon_{it_0-t}\Big|_\infty=o_{\mathbb{P}}(1).$$

Furthermore, using prior results, it follows that

$$\sqrt{Th}\Sigma_b(t_0)^{-1/2}\Big(\widehat{b}_{it_0}-b_{it_0}\Big)=-\Sigma_b(t_0)^{-1/2}\Sigma_A(t_0)^{-1}\frac{1}{\sqrt{Th}}\sum_{s=1}^{H}K\Big(\frac{t}{Th}\Big)X_{t_0-t}\varepsilon_{it_0-t}+o_{\mathbb{P}}(1),$$

uniformly over t_0 and i. The stochastic linear approximation on the right-hand-side of the equal sign is a martingale difference sequence, and thus the proof can now be easily completed by applying Corollary 3.1 in Hall and Heyde (2014). □

SA-2 Second Step Estimators

This section presents the proofs of the main results reported in the paper. It also provides additional results that are either discussed heuristically in the paper or are not given there to streamline the presentation.

Remark SA-2.1. *For the results for the second step estimators we will condition on two events, namely, that* $\max_{H+1\leq t\leq T}\max_{1\leq i\leq n_t}\big|\widehat{\beta}_{it}\big|$ *is bounded and that* $(\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1}$ *exists and is finite uniformly in* t*:* (1) *When* $\mathsf{R}_{nT}\to 0$ *then* $\big|\widehat{\beta}_{it}\big|$ *is bounded with probability approaching one;* (2) *by Lemma SA-2.3 (below), if* $J^2\log(nT)/n+\mathsf{R}_{nT}\to 0$*, then* $\min_{H+1\leq t\leq T}\lambda_{\min}(\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)$ *is bounded away from zero with probability approaching one.*

SA-2.1 Preliminary Technical Lemmas

Let $k_{jt}=\lfloor n_tj/J_t\rfloor$ and $\kappa_{j,t}=j/J_t$. Recall

$$\widehat{\beta}_{(k_{jt}),t}=F_{\widehat{\beta},n,t}^{-1}(\kappa_{j,t}),\qquad F_{\widehat{\beta},n,t}(u)=\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(\widehat{\beta}_{it}\leq u),$$

and

$$\beta_{(k_{jt}),t}=F_{\beta,n,t}^{-1}(\kappa_{j,t}),\qquad F_{\beta,n,t}(u)=\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(\beta_{it}\leq u).$$

The following lemma presents some basic properties of the estimated quantiles and the resulting partitioning scheme.

Lemma SA-2.2. *Suppose the conditions of Theorem SA-1.2 hold, and $J^2\log(nT)/n \to 0$ and $\mathsf{R}_{nT} \to 0$. Then,*

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\max_{1\leq j\leq J_t}\left|\widehat{\beta}_{(k_{jt}),t}-\beta_{(k_{jt}),t}\right| \lesssim_{\mathbb{P}} \sqrt{\frac{\log(nT)}{n}}+\mathsf{R}_{nT},$$

$$\frac{1}{J}\lesssim_{\mathbb{P}}\min_{\lfloor Th\rfloor+1\leq t\leq T}\min_{1\leq j\leq J_t}|\beta_{(k_{jt}),t}-\beta_{(k_{(j-1)t}),t}| \leq \max_{\lfloor Th\rfloor+1\leq t\leq T}\max_{1\leq j\leq J_t}|\beta_{(k_{jt}),t}-\beta_{(k_{(j-1)t}),t}|\lesssim_{\mathbb{P}}\frac{1}{J},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\max_{1\leq j\leq J_t}\left|\widehat{\beta}_{(k_{jt}),t}-\beta_{(k_{jt}),t}-[\widehat{\beta}_{(k_{(j-1)t}),t}-\beta_{(k_{(j-1)t}),t}]\right| \lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{nJ}}+\frac{\mathsf{R}_{nT}\sqrt{\log(nT)}}{J}=:\mathsf{L}_{nT},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\max_{1\leq j\leq J_t}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(\widehat{\beta}_{it}\in\widehat{P}_{jt})-\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(\beta_{it}\in P_{jt})]\right|\lesssim_{\mathbb{P}}\mathsf{L}_{nT},$$

and

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\max_{1\leq j\leq J_t}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{1}(\beta_{it}\in P_{jt})-\mathbb{E}[\mathbb{1}(\beta_{it}\in P_{jt})]\right|\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{nJ}}.$$

The next lemma controls the convergence of the Gram matrix, score vector, and other related quantities underlying our estimator. Recall that Q_t is a diagonal matrix with elements $\{q_{jt}: j=1,\ldots,J_t\}$.

Lemma SA-2.3. *Suppose the conditions of Theorem SA-1.2 hold, and $J^2\log(nT)/n \to 0$ and $\mathsf{R}_{nT} \to 0$. Then,*

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\left|(\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1}-Q_t^{-1}\right|_\infty\lesssim_{\mathbb{P}}J^2\mathsf{L}_{nT}+J^2\sqrt{\frac{\log(nT)}{nJ}},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,t}\varepsilon_{it}\right|_\infty\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{nJ}},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}(\widehat{\Phi}_{i,t}-\Phi_{i,t})\varepsilon_{it}\right|_\infty\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)\mathsf{L}_{nT}}{n}},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,t}\beta_{it}(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\Big|_\infty\lesssim_{\mathbb{P}}\frac{\sqrt{\log(nT)}}{J}+\sqrt{\frac{\log(nT)}{nJ}},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}(\widehat{\Phi}_{i,t}-\Phi_{i,t})\beta_{it}(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\Big|_\infty\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)\mathsf{L}_{nT}}{n}}+\sqrt{\log(nT)}\mathsf{L}_{nT},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{k=1}^{n_t}\widehat{\Phi}_{i,t}\widehat{\Phi}_{k,t}^\top\beta_{it}\beta_{kt}\Big|_\infty\lesssim_{\mathbb{P}}\frac{1}{J^2},$$

$$\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{k=1}^{n_t}\big(\widehat{\Phi}_{i,t}-\Phi_{i,t}\big)\big(\widehat{\Phi}_{k,t}-\Phi_{k,t}^\circ\big)^\top\beta_{it}\beta_{kt}\Big|_\infty\lesssim_{\mathbb{P}}\mathsf{L}_{nT}^2.$$

Finally, the following lemma gives the cross-sectional convergence rate of $\widehat{\mu}_t(\beta)$ to $M_t(\beta)$.

Lemma SA-2.4. *Suppose the conditions of Theorem SA-1.2 hold, and $J^2\log(nT)/n\to 0$ and $\mathsf{R}_{nT}^2\log(nT)\to 0$. Then,*

$$\max_{H+1\leq t\leq T}\sup_{\beta\in\mathcal{B}}\big|\widehat{\mu}_t(\beta)-M_t(\beta)\big|\lesssim_{\mathbb{P}}\sqrt{\frac{J\log(nT)}{n}}+\frac{1}{J}+J\mathsf{L}_{nT}=o_{\mathbb{P}}(1).$$

SA-2.2 Proof of Lemma 4.1

We begin with the elementary decomposition

$$\begin{aligned}\widehat{\mu}(\beta)-\bar{\mu}_T(\beta;H)&=\frac{1}{T-H}\sum_{t=H+1}^{T}(\widehat{\mu}_t(\beta)-\mu_t(\beta))\\&=\frac{1}{T-H}\sum_{t=H+1}^{T}\Big(\widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\widehat{\Phi}_tR_t-\mu_t(\beta)\Big)\\&=\frac{1}{T-H}\sum_{t=H+1}^{T}\widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,t}\Big(\varepsilon_{it}+\beta_{it}(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\Big)\\&\quad+\mathscr{B}(\beta)+\mathscr{R}(\beta),\end{aligned}$$

where

$$\mathscr{B}(\beta)=\frac{1}{T-H}\sum_{t=H+1}^{T}\Big[\widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\big(\mu_t(\beta_{it})-\widehat{\Phi}_{i,t}^\top a_t^\circ\big)+\big(\widehat{p}_t(\beta)^\top a_t^\circ-\mu_t(\beta)\big)\Big],$$

and

$$\mathscr{R}(\beta) = \mathfrak{R}_1(\beta) + \mathfrak{R}_2(\beta) + \mathfrak{R}_3(\beta) + \mathfrak{R}_4(\beta),$$

with

$$\mathfrak{R}_1(\beta) = \frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top \Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big)\frac{1}{n_t}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\varepsilon_{it},$$
$$\mathfrak{R}_2(\beta) = \frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\big(\widehat{\Phi}_{i,t} - \Phi_{i,t}\big)\varepsilon_{it},$$
$$\mathfrak{R}_3(\beta) = \frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top \Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big)\frac{1}{n_t}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\beta_{it}(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}]),$$
$$\mathfrak{R}_4(\beta) = \frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\big(\widehat{\Phi}_{i,t} - \Phi_{i,t}\big)\beta_{it}(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}]),$$

where $a_t^\circ = (\mathbb{E}[\Phi_{i,t}\Phi_{i,t}^\top|\mathcal{G}_{t-1}])^{-1}\mathbb{E}[\Phi_{i,t}R_{it}|\mathcal{G}_{t-1}]$.

By previous results, the term $(\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1}$ exists and is finite with probability approaching one, and on that event, $\mathbb{E}[\mathfrak{R}_1(\beta)] = 0$. Thus, on that event, using the martingale structure,

$$\begin{aligned}
&\frac{1}{(T-H)^2}\sum_{t=H+1}^{T}\mathbb{V}\mathrm{ar}[\mathfrak{R}_1(\beta)|\mathcal{F}_{t-1}]\\
&\lesssim_{\mathbb{P}} \frac{1}{T}\max_{\lfloor Th\rfloor+1\le t\le T}\mathbb{V}\mathrm{ar}\Big[\widehat{p}_t(\beta)^\top\Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big)\frac{1}{n_t}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\varepsilon_{it}\Big|\mathcal{F}_{t-1}\Big]\\
&\lesssim_{\mathbb{P}} \frac{1}{T}\max_{\lfloor Th\rfloor+1\le t\le T}\widehat{p}_t(\beta)^\top\Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big)\frac{1}{n_t^2}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\widehat{\Phi}_{i,t}^\top\Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big)\widehat{p}_t(\beta)\\
&\lesssim_{\mathbb{P}} \frac{1}{nTJ}\Big(J^2\mathsf{L}_{nT} + J^2\sqrt{\frac{\log(nT)}{nJ}}\Big)^2 \lesssim \frac{J^3\mathsf{L}_{nT}^2}{nT} + \frac{J^2\log(nT)}{n^2T} = o_{\mathbb{P}}(\frac{J}{nT}).
\end{aligned}$$

Proceeding analogously, for the second term, we verify

$$\begin{aligned}
&\frac{1}{T}\max_{\lfloor Th\rfloor+1\le t\le T}\mathbb{V}\mathrm{ar}\Big[\widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\big(\widehat{\Phi}_{i,t} - \Phi_{i,t}\big)\varepsilon_{it}\Big|\mathcal{F}_{t-1}\Big]\\
&\lesssim_{\mathbb{P}} \frac{J^2}{T}\max_{\lfloor Th\rfloor+1\le t\le T}\frac{1}{n_t^2}\sum_{i=1}^{n_t}\big|(\widehat{\Phi}_{i,t} - \Phi_{i,t})(\widehat{\Phi}_{i,t} - \Phi_{i,t})^\top\big|_\infty \lesssim \frac{J^2\mathsf{L}_{nT}}{nT} = o_{\mathbb{P}}(\frac{J}{nT} + \frac{1}{TJ^2}),
\end{aligned}$$

because $J\mathsf{L}_{nT} \to 0$.

For the third term, first consider the case when $\beta \neq 0$. Proceeding as above, we have

$$\frac{1}{T}\max_{\lfloor Th\rfloor+1\le t\le T}\mathbb{V}\mathrm{ar}\Big[\widehat{p}_t(\beta)^\top\Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big)\frac{1}{n_t}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\beta_{it}(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\Big|\mathcal{F}_{t-1}\Big]$$

$$\lesssim_{\mathbb{P}} \frac{1}{T} \max_{\lfloor Th \rfloor+1 \le t \le T} \widehat{p}_t(\beta)^\top \Big((\widehat{\Phi}_t \widehat{\Phi}_t^\top / n_t)^{-1} - Q_t^{-1} \Big) \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{k=1}^{n_t} \widehat{\Phi}_{i,t} \widehat{\Phi}_{k,t}^\top \beta_{it} \beta_{kt}$$
$$\Big((\widehat{\Phi}_t \widehat{\Phi}_t^\top / n_t)^{-1} - Q_t^{-1} \Big) \widehat{p}_t(\beta)$$
$$\lesssim_{\mathbb{P}} \frac{1}{T} \Big(J^2 \mathsf{L}_{nT} + J^2 \sqrt{\frac{\log(nT)}{nJ}} \Big)^2 \max_{\lfloor Th \rfloor+1 \le t \le T} \Big| \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{k=1}^{n_t} \widehat{\Phi}_{i,t} \widehat{\Phi}_{k,t}^\top \beta_{it} \beta_{kt} \Big|_\infty$$
$$\lesssim_{\mathbb{P}} \frac{1}{T} \Big(J^2 \mathsf{L}_{nT} + J^2 \sqrt{\frac{\log(nT)}{nJ}} \Big)^2 (\frac{1}{J^2}) = o_{\mathbb{P}}(\frac{1}{T}).$$

When $\beta = 0$, we obtain the faster upper bound

$$\frac{1}{T} \max_{\lfloor Th \rfloor+1 \le t \le T} \mathbb{V}\mathrm{ar}\Big[\widehat{p}_t(\beta)^\top \Big((\widehat{\Phi}_t \widehat{\Phi}_t^\top / n_t)^{-1} - Q_t^{-1} \Big) \frac{1}{n_t} \sum_{i=1}^{n_t} \widehat{\Phi}_{i,t} \beta_{it} (f_t - \mathbb{E}[f_t | \mathcal{G}_{t-1}]) \Big| \mathcal{F}_{t-1} \Big]$$
$$\lesssim_{\mathbb{P}} \frac{1}{T} \Big(J^2 \mathsf{L}_{nT} + J^2 \sqrt{\frac{\log(nT)}{nJ}} \Big)^2 (\frac{1}{J^4}) = o_{\mathbb{P}}(\frac{J}{nT} + \frac{1}{TJ^2}),$$

because $J\mathsf{L}_{nT} \to 0$.

For the fourth term, first consider the case when $\beta \neq 0$. Using the same logic as before,

$$\frac{1}{T} \max_{\lfloor Th \rfloor+1 \le t \le T} \mathbb{V}\mathrm{ar}\Big[\widehat{p}_t(\beta)^\top Q_t^{-1} \frac{1}{n_t} \sum_{i=1}^{n_t} \big(\widehat{\Phi}_{i,t} - \Phi_{i,t} \big) \beta_{it} (f_t - \mathbb{E}[f_t | \mathcal{G}_{t-1}]) \Big| \mathcal{F}_{t-1} \Big]$$
$$\lesssim_{\mathbb{P}} \frac{J^2}{T} \max_{\lfloor Th \rfloor+1 \le t \le T} \Big| \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{k=1}^{n_t} \big(\widehat{\Phi}_{i,t} - \Phi_{i,t} \big) \big(\widehat{\Phi}_{k,t} - \Phi_{k,t} \big)^\top \beta_{it} \beta_{kt} \Big|_\infty$$
$$\lesssim_{\mathbb{P}} \frac{J^2}{T} \mathsf{L}_{nT}^2 = o_{\mathbb{P}}(T^{-1}),$$

since $J\mathsf{L}_{nT} \to 0$. Likewise, when $\beta = 0$,

$$\frac{1}{T} \max_{\lfloor Th \rfloor+1 \le t \le T} \mathbb{V}\mathrm{ar}\Big[\widehat{p}_t(\beta)^\top Q_t^{-1} \frac{1}{n_t} \sum_{i=1}^{n_t} \big(\widehat{\Phi}_{i,t} - \Phi_{i,t} \big) \beta_{it} (f_t - \mathbb{E}[f_t | \mathcal{G}_{t-1}]) \Big| \mathcal{F}_{t-1} \Big]$$
$$\lesssim_{\mathbb{P}} \frac{J^2}{T} \max_{\lfloor Th \rfloor+1 \le t \le T} \Big| \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{k=1}^{n_t} \big(\widehat{\Phi}_{i,t} - \Phi_{i,t} \big) \big(\widehat{\Phi}_{k,t} - \Phi_{k,t} \big)^\top \beta_{it} \beta_{kt} \Big|_\infty$$
$$\lesssim_{\mathbb{P}} \frac{1}{T} \mathsf{L}_{nT}^2 = o_{\mathbb{P}}(\frac{1}{TJ^2}),$$

because $J\mathsf{L}_{nT} \to 0$.

Finally, for the bias term, we verify that $\mathscr{B}(\beta) \lesssim_{\mathbb{P}} J^{-1}$. Define

$$\tilde{a}_t^\circ = (\Phi_t \Phi_t^\top)^{-1} \mathbb{E}(\Phi_t R_t | \mathcal{F}_{t-1}) = (\Phi_t \Phi_t^\top)^{-1} \Phi_t \mu_t(\beta_t),$$

where $\mu_t(\beta_t) = (\mu_t(\beta_{1t}), \dots, \mu_t(\beta_{n_t t}))'$. Then,

$$\max_{\lfloor Th \rfloor+1 \le t \le T} \big| a_t^\circ - \tilde{a}_t^\circ \big|_\infty$$

$$
\begin{aligned}
&= \max_{\lfloor Th \rfloor+1\le t\le T} \left|\left(\mathbb{E}[\Phi_t\Phi_t^\top|\mathcal{G}_{t-1}]\right)^{-1}\left(\Phi_t\Phi_t^\top - \mathbb{E}[\Phi_t\Phi_t^\top|\mathcal{G}_{t-1}]\right)\tilde{a}_t^\circ\right|_\infty \\
&\le \max_{\lfloor Th \rfloor+1\le t\le T} \left|\left(\mathbb{E}[\Phi_t\Phi_t^\top|\mathcal{G}_{t-1}]\right)^{-1}\right|_\infty \max_{\lfloor Th \rfloor+1\le t\le T} \left|\Phi_t\Phi_t^\top - \mathbb{E}[\Phi_t\Phi_t^\top|\mathcal{G}_{t-1}]\right|_\infty \max_{\lfloor Th \rfloor+1\le t\le T} \left|\tilde{a}_t^\circ\right|_\infty \\
&\lesssim_{\mathbb{P}} J^{-1},
\end{aligned}
$$

where the last line follows by Lemma SA-2.2, Bernstein's inequality and since $\left|\tilde{a}_t^\circ\right|_\infty$ is bounded on the event that $(\Phi_t\Phi_t^\top/n_t)^{-1}$ exists and is finite which occurs with probability approaching one. Therefore, $\mathscr{B}(\beta) \lesssim \mathscr{B}_1(\beta) + \mathscr{B}_2(\beta)$ where

$$
\begin{aligned}
\mathscr{B}_1(\beta) &= \frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\big(\mu_t(\beta_{it}) - \widehat{\Phi}_{i,t}^\top a_t^\circ\big), \\
\mathscr{B}_2(\beta) &= \frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\big(\widehat{\Phi}_{i,t}^\top a_t^\circ - \mu_t(\beta)\big).
\end{aligned}
$$

For the first term $\mathscr{B}_1(\beta)$ we have

$$
\begin{aligned}
|\mathscr{B}_1(\beta)| \lesssim &\left|\frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\mu_t(\beta_{it}) - \frac{1}{T-H}\sum_{t=H+1}^{T}\widehat{p}_t(\beta)^\top\tilde{a}_t^\circ\right| \\
&+ \left|\frac{1}{T-H}\sum_{t=H+1}^{T}\widehat{p}_t(\beta)^\top\left(\tilde{a}_t^\circ - a_t^\circ\right)\right|.
\end{aligned}
$$

The third term is $O_{\mathbb{P}}(J^{-1})$ by above calculations and because $|\widehat{p}_t(\beta)|_\infty = 1$. Next note that the first term is an average of $\mu_t(\beta_{it})$ in each $\widehat{P}_{jt}$ which contains β whereas the second term is an average of $\mu_t(\beta_{it})$ in each P_{jt} which contains β where P_{jt} are the portfolios constructed using the sample quantiles of β_{it}. Thus, the first two terms are bounded by

$$
\begin{aligned}
&\left|\frac{1}{T-H}\sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\mu_t(\beta_{it}) - \frac{1}{T-H}\sum_{t=H+1}^{T}\widehat{p}_t(\beta)^\top\tilde{a}_t^\circ\right| \\
&\lesssim_{\mathbb{P}} \max_{\lfloor Th \rfloor+1\le t\le T}\max_{1\le j\le J_t}\left|\big(\widehat{\beta}_{(k_{jt}),t}\vee\beta_{(k_{jt}),t}\big) - \big(\widehat{\beta}_{(k_{j(t-1)}),t}\wedge\beta_{(k_{j(t-1)}),t-1}\big)\right| \lesssim_{\mathbb{P}} J^{-1},
\end{aligned}
$$

by our smoothness assumptions on $\mu_t(\cdot)$ and by Lemma SA-2.2. The second term $\mathscr{B}_2(\beta)$ follows by similar steps. □

SA-2.3 Proof of Theorem 4.3

By Lemma 4.1, we have

$$
\frac{\widehat{\mu}(\beta) - \bar{\mu}_T(\beta;H) - \mathscr{B}(\beta)}{\sqrt{\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)]}} = \sum_{t=H+1}^{T}\eta_t(\beta) + \mathfrak{R}_0 + o_{\mathbb{P}}(1),
$$

where

$$\eta_t(\beta) := (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-1/2} \frac{1}{T-H} \widehat{p}_t(\beta)^\top Q_t^{-1} \frac{1}{n_t} \sum_{i=1}^{n_t} \big(\Phi_{i,t}\varepsilon_{it} + \mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\big),$$

forms a martingale difference sequence adapted to the filtration $\mathcal{F}_{t-1}$, and because

$$\begin{aligned}
\mathfrak{R}_0 &:= (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-1/2} \frac{1}{T-H} \sum_{t=H+1}^{T} \widehat{p}_t(\beta)^\top Q_t^{-1} \frac{1}{n_t} \sum_{i=1}^{n_t} (\Phi_{i,t}\beta_{it} - \mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}])(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}]) \\
&\lesssim_{\mathbb{P}} \sqrt{\frac{nT}{J}} \frac{1}{\sqrt{TnJ}} + \sqrt{\frac{J}{n}} = o_{\mathbb{P}}(1),
\end{aligned}$$

and

$$\frac{\mathscr{R}(\beta)}{\sqrt{\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)]}} = o_{\mathbb{P}}(1).$$

Thus, the proof is completed by employing the martingale central limit theorem of Hall and Heyde (2014, Corollary 3.1), whose conditions are implied by the following two conditions:

$$\sum_{t=H+1}^{T} \mathbb{E}[\eta_t(\beta)^4|\mathcal{F}_{t-1}] \to_{\mathbb{P}} 0 \tag{SA-2.1}$$

and

$$\mathbb{E}\Big| \sum_{t=H+1}^{T} \mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-1}] - 1 \Big|^2 \to 0. \tag{SA-2.2}$$

For the first condition (SA-2.1), we have

$$\sum_{t=H+1}^{T} \mathbb{E}[\eta_t(\beta)^4|\mathcal{F}_{t-1}] \lesssim \mathfrak{R}_1 + \mathfrak{R}_2,$$

where

$$\begin{aligned}
\mathfrak{R}_1 &:= (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-2} \frac{1}{T^4} \sum_{t=H+1}^{T} \mathbb{E}\Big[\Big(\widehat{p}_t(\beta)^\top Q_t^{-1} \frac{1}{n_t} \sum_{i=1}^{n_t} \Phi_{i,t}\varepsilon_{it}\Big)^4 \Big| \mathcal{F}_{t-1}\Big], \\
\mathfrak{R}_2 &:= (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-2} \frac{1}{T^4} \sum_{t=H+1}^{T} \mathbb{E}\Big[\Big(\widehat{p}_t(\beta)^\top Q_t^{-1} \mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\Big)^4 \Big| \mathcal{F}_{t-1}\Big].
\end{aligned}$$

For the first term we have

$$\begin{aligned}\mathfrak{R}_1 &\lesssim \min(\frac{n^2T^2}{J^2},T^2J^4)\frac{1}{T^4}\sum_{t=H+1}^{T}\frac{1}{n_t^4}\sum_{i=1}^{n_t}\mathbb{E}\Big[\Big(\widehat{p}_t(\beta)^\top Q_t^{-1}\Phi_{i,t}\varepsilon_{it}\Big)^4\Big|\mathcal{F}_{t-1}\Big]\\
&\quad+\min(\frac{n^2T^2}{J^2},T^2J^4)\frac{1}{T^4}\sum_{t=H+1}^{T}\frac{1}{n_t^4}\sum_{i=1}^{n_t}\sum_{k=1,k\neq i}^{n_t}\\
&\qquad\qquad\mathbb{E}\Big[\Big(\widehat{p}_t(\beta)^\top Q_t^{-1}\Phi_{i,t}\varepsilon_{it}\Big)^2\Big|\mathcal{F}_{t-1}\Big]\mathbb{E}\Big[\Big(\widehat{p}_t(\beta)^\top Q_t^{-1}\Phi_{k,t}\varepsilon_{kt}\Big)^2\Big|\mathcal{F}_{t-1}\Big]\\
&\lesssim_{\mathbb{P}}\frac{1}{T},\end{aligned}$$

and, similarly, $\mathfrak{R}_2 \lesssim_{\mathbb{P}} \frac{1}{T}$ for both cases of interest ($\beta\neq 0$ and $\beta=0$).

For the second condition (SA-2.1), first define

$$\mathsf{Z}_t(\beta)=\mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-1}]-\mathbb{E}[\eta_t(\beta)^2],\qquad \mathsf{M}_\ell(\beta)=\sum_{t=H+1}^{T}\big[\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell}]-\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell-1}]\big],$$

and because $\mathsf{Z}_t(\beta)$ is mean-zero and adapted to the filtration $\mathcal{F}_t$ and therefore

$$\mathsf{Z}_t(\beta)=\sum_{\ell=0}^{\infty}\big[\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell}]-\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell-1}]\big].$$

Then, we have

$$\begin{aligned}&\mathbb{E}\Big|\sum_{t=H+1}^{T}\mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-1}]-1\Big|^2\\
&=\sum_{t_2=H+1}^{T}\sum_{t_1=H+1}^{T}\mathbb{E}[\{\mathbb{E}[\eta_{t_1}(\beta)^2|\mathcal{F}_{t_1-1}]-\mathbb{E}[\eta_{t_1}(\beta)^2]\}\{\mathbb{E}[\eta_{t_2}(\beta)^2|\mathcal{F}_{t_2-1}]-\mathbb{E}[\eta_{t_2}(\beta)^2]\}]\\
&=\Bigg(\sqrt{\mathbb{E}\Big[\Big(\sum_{t=H+1}^{T}\mathsf{Z}_t(\beta)\Big)^2\Big]}\Bigg)^2=\Bigg(\sqrt{\mathbb{E}\Big[\Big(\sum_{t=H+1}^{T}\sum_{\ell=0}^{\infty}\big[\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell}]-\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell-1}]\big]\Big)^2\Big]}\Bigg)^2\\
&=\Bigg(\sqrt{\mathbb{E}\Big[\Big(\sum_{\ell=0}^{\infty}\mathsf{M}_\ell(\beta)\Big)^2\Big]}\Bigg)^2=\Big\|\sum_{\ell=0}^{\infty}\mathsf{M}_\ell(\beta)\Big\|^2\leq\Big(\sum_{\ell=0}^{\infty}\|\mathsf{M}_\ell(\beta)\|\Big)^2\\
&=\Bigg(\sum_{\ell=0}^{\infty}\sqrt{\sum_{t=H+1}^{T}\mathbb{E}\big[(\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell}]-\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell-1}])^2\big]}\Bigg)^2\\
&\lesssim\Bigg(\sum_{\ell=0}^{\infty}\sqrt{\mathfrak{R}_{3,\ell}(\beta)}\Bigg)^2+\Bigg(\sum_{\ell=0}^{\infty}\sqrt{\mathfrak{R}_{4,\ell}(\beta)}\Bigg)^2,\end{aligned}$$

because, for $k \geq 0$, we have

$$
\begin{aligned}
&\mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-\ell-k}] \\
&= (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-1} \frac{1}{(T-H)^2} \frac{1}{n_t^2} \sum_{i=1}^{n_t} \mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{i,t} \varepsilon_{it}\big)^2 \Big| \mathcal{F}_{t-\ell-k} \Big] \\
&\qquad + (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-1} \frac{1}{(T-H)^2} \mathbb{E}\Big[\Big(\widehat{p}_t(\beta)^\top Q_t^{-1} \mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\Big)^2 \Big| \mathcal{F}_{t-\ell-k} \Big],
\end{aligned}
$$

and therefore

$$
\begin{aligned}
&\sum_{t=H+1}^{T} \mathbb{E}\big[(\mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell}] - \mathbb{E}[\mathsf{Z}_t(\beta)|\mathcal{F}_{t-\ell-1}])^2 \big] \\
&= \sum_{t=H+1}^{T} \mathbb{E}\big[(\mathbb{E}[\mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-1}]|\mathcal{F}_{t-\ell}] - \mathbb{E}[\mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-1}]|\mathcal{F}_{t-\ell-1}])^2 \big] \\
&= \sum_{t=H+1}^{T} \mathbb{E}\big[(\mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-\ell}] - \mathbb{E}[\eta_t(\beta)^2|\mathcal{F}_{t-\ell-1}])^2 \big] \lesssim \mathfrak{R}_{3,\ell}(\beta) + \mathfrak{R}_{4,\ell}(\beta),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathfrak{R}_{3,\ell}(\beta) &= (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-2} \frac{1}{(T-H)^4} \sum_{t=H+1}^{T} \frac{1}{n_t^4} \sum_{i=1}^{n_t} \sum_{k=1}^{n_t} \\
&\quad \mathbb{E}\Bigg[\Bigg(\Big(\mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{i,t}\varepsilon_{it}\big)^2 \Big| \mathcal{F}_{t-\ell} \Big] - \mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{i,t}\varepsilon_{it}\big)^2 \Big| \mathcal{F}_{t-\ell-1} \Big] \Big) \Bigg) \\
&\qquad \times \Bigg(\Big(\mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{k,t}\varepsilon_{kt}\big)^2 \Big| \mathcal{F}_{t-\ell} \Big] - \mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{k,t}\varepsilon_{kt}\big)^2 \Big| \mathcal{F}_{t-\ell-1} \Big] \Big) \Bigg) \Bigg] \\
&\lesssim (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-2} \frac{1}{T^3 n^2} \max_{H+1 \leq t \leq T} \max_{1 \leq i \leq n_t} \\
&\qquad \mathbb{E}\Bigg[\Big(\mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{i,t}\varepsilon_{it}\big)^2 \Big| \mathcal{F}_{t-\ell} \Big] - \mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \Phi_{i,t}\varepsilon_{it}\big)^2 \Big| \mathcal{F}_{t-\ell-1} \Big] \Big)^2 \Bigg],
\end{aligned}
$$

and

$$
\begin{aligned}
\mathfrak{R}_{4,\ell}(\beta) &\lesssim (\mathbb{E}[\sigma_\varepsilon^2(\beta) + \sigma_f^2(\beta)])^{-2} \frac{1}{T^3} \\
&\quad \max_{H+1 \leq t \leq T} \mathbb{E}\Bigg[\Bigg(\Big(\mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\big)^2 \Big| \mathcal{F}_{t-\ell} \Big] \\
&\qquad - \mathbb{E}\Big[\big(\widehat{p}_t(\beta)^\top Q_t^{-1} \mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\big)^2 \Big| \mathcal{F}_{t-\ell-1} \Big] \Big) \Bigg)^2 \Bigg].
\end{aligned}
$$

To complete the proof, observe that under our imposed assumptions,

$$
\begin{aligned}
&\sum_{\ell=0}^{\infty}\sqrt{\mathfrak{R}_{3,\ell}}\\
&=(\mathbb{E}[\sigma_{\varepsilon}^2(\beta)+\sigma_f^2(\beta)])^{-1}\frac{1}{T^{3/2}n}\max_{H+1\le t\le T}\max_{1\le i\le n_t}\\
&\qquad\sum_{\ell=0}^{\infty}\sqrt{\mathbb{E}\left[\left(\mathbb{E}\left[\left(\widehat{p}_t(\beta)^\top Q_t^{-1}\Phi_{i,t}\varepsilon_{it}\right)^2\Big|\mathcal{F}_{t-\ell}\right]-\mathbb{E}\left[\left(\widehat{p}_t(\beta)^\top Q_t^{-1}\Phi_{i,t}\varepsilon_{it}\right)^2\Big|\mathcal{F}_{t-\ell-1}\right]\right)^2\right]}\\
&\lesssim \min(\frac{nT}{J},TJ^2)\frac{1}{T^{3/2}n}\Theta_{n,T}((\widehat{p}_\bullet(\beta)^\top Q_\bullet^{-1}\Phi_{i,\bullet}\varepsilon_{i\bullet})^2;q,v)=o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{\ell=0}^{\infty}\sqrt{\mathfrak{R}_{4,\ell}}\\
&=(\mathbb{E}[\sigma_{\varepsilon}^2(\beta)+\sigma_f^2(\beta)])^{-1}\frac{1}{T^{3/2}}\max_{H+1\le t\le T}\\
&\qquad\sum_{\ell=0}^{\infty}\Bigg(\mathbb{E}\Bigg[\Bigg(\left(\mathbb{E}\left[\left(\widehat{p}_t(\beta)^\top Q_t^{-1}\mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\right)^2\Big|\mathcal{F}_{t-\ell}\right]\right.\\
&\qquad\qquad\left.-\mathbb{E}\left[\left(\widehat{p}_t(\beta)^\top Q_t^{-1}\mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}](f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\right)^2\Big|\mathcal{F}_{t-\ell-1}\right]\right)\Bigg)^2\Bigg]\Bigg)^{1/2}\\
&\lesssim \min(\frac{n}{J},J^2)\frac{1}{T^{1/2}}\Theta_{n,T}((\widehat{p}_\bullet(\beta)^\top Q_\bullet^{-1}\mathbb{E}[\Phi_{i,\bullet}\beta_{i\bullet}|\mathcal{G}_{t-1}]f_\bullet)^2;q,v)=o(1).
\end{aligned}
$$

Then the conclusion follows. □

SA-2.4 Proof of Theorem 4.5(i)

It will be convenient to define the rate $r_{n,T}(\beta)=\frac{J}{Tn}+\frac{1}{TJ^2}$ for $\beta=0$ and $r_{n,T}(\beta)=\frac{1}{T}$ for $\beta\neq 0$. We start with the result when $\beta\neq 0$,

$$
|\widehat{\sigma}^2_{\mathtt{FM}}(\beta)-\sigma_{\varepsilon}^2(\beta)-\sigma_f^2(\beta)-\sigma_{\mu}^2(\beta)|\lesssim \mathfrak{R}_1+\mathfrak{R}_2+\mathfrak{R}_3,
$$

where

$$
\begin{aligned}
\mathfrak{R}_1&=\Big|\frac{1}{(T-H)^2}\sum_{t=H+1}^{T}(\widehat{\mu}_t(\beta)-\mu_t(\beta))^2-\sigma_{\varepsilon}^2(\beta)-\sigma_f^2(\beta)\Big|,\\
\mathfrak{R}_2&=\Big|\frac{1}{(T-H)}(\widehat{\mu}(\beta)-\bar{\mu}_T(\beta;H))^2\Big|,\\
\mathfrak{R}_3&=\Big|\frac{2}{(T-H)^2}\sum_{t=H+1}^{T}(\widehat{\mu}_t(\beta)-\mu_t(\beta))(\mu_t(\beta)-\bar{\mu}_T(\beta;H))\Big|.
\end{aligned}
$$

For $\mathfrak{R}_1$, we can follow similar steps as for the proof of Lemma 4.1 to obtain that

$$\mathfrak{R}_1 = O_{\mathbb{P}}\Big(\frac{1}{TJ^2}\Big) + o_{\mathbb{P}}(r_{n,T}(\beta)) = o_{\mathbb{P}}\Big(\frac{1}{T}\Big).$$

Next, note that $\mathfrak{R}_2 = o_{\mathbb{P}}(\frac{1}{TJ} + \frac{1}{T^2})$ by Lemma 4.1 and Theorem 4.3 so that $\mathfrak{R}_2 = o_{\mathbb{P}}\Big(\frac{1}{T}\Big)$. Finally, for $\mathfrak{R}_3$ we have that

$$\mathfrak{R}_3 \lesssim \Big|\frac{1}{(T-H)^2}\sum_{t=H+1}^{T}\left(\widehat{\mu}_t(\beta) - \mu_t(\beta)\right)\mu_t(\beta)\Big| + \Big|\frac{1}{(T-H)}\bar{\mu}_T(\beta;H)\left(\widehat{\mu}(\beta) - \bar{\mu}_T(\beta;H)\right)\Big|.$$

By similar steps as in the proof of Lemma 4.1, and using the fact that $|\mu_t(\beta)|$ is bounded, the first term is $O_{\mathbb{P}}(\frac{1}{TJ} + \frac{\sqrt{r_{n,T}(\beta)}}{T})$ and by Lemma 4.1 and Theorem 4.3 the second term is $O_{\mathbb{P}}(\frac{\sqrt{r_{n,T}(\beta)}}{T})$. Thus, $\mathfrak{R}_3 = o_{\mathbb{P}}\Big(\frac{1}{T}\Big)$, and the result for $\beta \neq 0$ follows.

Now consider the $\beta = 0$ case. We have that,

$$\left|\widetilde{\sigma}^2_{\mathtt{FM}}(\beta) - \sigma^2_{\varepsilon}(\beta) - \sigma^2_f(\beta) - \sigma^2_{\mu}(\beta)\right| \lesssim \mathfrak{R}_1 + \mathfrak{R}_2 + \mathfrak{R}_3,$$

where

$$\begin{aligned}
\mathfrak{R}_1 &= \Big|\frac{1}{(T-H)^2}\sum_{t=H+1}^{T}\left(\widehat{\mu}_t(\beta) - \mathscr{B}_t(\beta) - \mu_t(\beta)\right)^2 - \sigma^2_{\varepsilon}(\beta) - \sigma^2_f(\beta)\Big|,\\
\mathfrak{R}_2 &= \Big|\frac{1}{(T-H)}\left(\widehat{\mu}(\beta) - \mathscr{B}(\beta) - \bar{\mu}_T(\beta;H)\right)^2\Big|,\\
\mathfrak{R}_3 &= \Big|\frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\left(\widehat{\mu}_t(\beta) - \mathscr{B}_t(\beta) - \mu_t(\beta)\right)\left(\mu_t(\beta) - \bar{\mu}_T(\beta;H)\right)\Big|.
\end{aligned}$$

For $\mathfrak{R}_1$, we can follow similar steps as for the proof of Lemma 4.1 to obtain that $\mathfrak{R}_1 = o_{\mathbb{P}}(r_{n,T}(\beta))$. Next, note that $\mathfrak{R}_2 = o_{\mathbb{P}}(r_{n,T}(\beta))$ directly by Lemma 4.1 and Theorem 4.3. Finally, for $\mathfrak{R}_3$, using the Cauchy–Schwarz inequality, we have

$$\mathfrak{R}_3^2 \lesssim \Big|\frac{1}{(T-H)^2}\sum_{t=H+1}^{T}\left(\widehat{\mu}_t(\beta) - \mathscr{B}_t(\beta) - \mu_t(\beta)\right)^2\Big| \cdot \sigma^2_{\mu}(\beta).$$

By the same steps as for $\mathfrak{R}_1$ we have that the first factor is $O_{\mathbb{P}}(r_{n,T}(\beta))$. By assumption, we have that $\sigma^2_{\mu}(\beta) = o_{\mathbb{P}}(r_{n,T}(\beta))$ and so $\mathfrak{R}_3 = o_{\mathbb{P}}(r_{n,T}(\beta))$. □

SA-2.5 Proof of Theorem 4.5(ii)

Recall that we define the rate $r_{n,T}(\beta) = \frac{J}{Tn} + \frac{1}{TJ^2}$ for $\beta = 0$ and $r_{n,T}(\beta) = \frac{1}{T}$ for $\beta \neq 0$ from the Proof of Theorem 4.5(i). We first decompose $\widehat{\sigma}^2_{f,\mathtt{PI}}(\beta)$ as

$$\widehat{\sigma}^2_{f,\mathtt{PI}}(\beta) - \widetilde{\sigma}^2_{f,\mathtt{PI}}(\beta) - s_{nT}(\beta) = \mathfrak{R}(\beta),$$

where

$$s_{nT}(\beta) = \frac{1}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\left(\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})\widehat{\beta}_{it}\right)^2(\mathbb{E}[f_t|\mathcal{G}_{t-1}] - \widehat{\mathbb{E}[f_t|\mathcal{G}_{t-1}]})^2,$$

and

$$\mathfrak{R}(\beta) := \frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\left(\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})\widehat{\beta}_{it}\right)^2(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])(\mathbb{E}[f_t|\mathcal{G}_{t-1}] - \widehat{\mathbb{E}[f_t|\mathcal{G}_{t-1}]}).$$

Clearly, $s_{nT}(\beta) \geq 0$ for all $\beta \in \mathcal{B}$. For $\mathfrak{R}(\beta)$, note that the summands form a martingale difference sequence with respect to $\mathcal{F}_{t-1}$ so that, when $\beta \neq 0$,

$$\begin{aligned}
\mathbb{E}[\mathfrak{R}^2(\beta)] &= \frac{2}{(T-H)^4}\mathbb{E}\Big[\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^4\widehat{q}_{jt}^4}\left(\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})\widehat{\beta}_{it}\right)^4(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])^2(\mathbb{E}[f_t|\mathcal{G}_{t-1}] - \widehat{\mathbb{E}[f_t|\mathcal{G}_{t-1}]})^2\Big] \\
&\lesssim \frac{1}{(T-H)^4}\sum_{t=H+1}^{T}\mathbb{E}\Big[(\mathbb{E}[f_t|\mathcal{G}_{t-1}] - \widehat{\mathbb{E}[f_t|\mathcal{G}_{t-1}]})^2\Big], \\
&\lesssim r_{n,T}(\beta)^2.
\end{aligned}$$

When $\beta = 0$, we can follow similar steps and obtain that $\mathbb{V}\mathrm{ar}[\mathfrak{R}(\beta)] = o(r_{n,T}(\beta)^2)$ using Lemma SA-2.2. Finally, we need only show that $\left|\widetilde{\sigma}^2_{f,\mathtt{PI}}(\beta) - \sigma^2_f(\beta)\right| = o_{\mathbb{P}}(r_{n,T}(\beta))$ which holds under the conditions of Theorem SA-1.2 and given the results in Lemma SA-2.2 to Lemma SA-2.3.

We next prove that $|\widehat{\sigma}^2_{\varepsilon,\mathtt{PI}}(\beta) - \widetilde{\sigma}^2_{\varepsilon,\mathtt{PI}}(\beta)| = o_{\mathbb{P}}(r_{n,T}(\beta))$. We have that,

$$|\widehat{\sigma}^2_{\varepsilon,\mathtt{PI}}(\beta) - \widetilde{\sigma}^2_{\varepsilon,\mathtt{PI}}(\beta)| \leq \mathfrak{R}_4 + \mathfrak{R}_5 + \mathfrak{R}_6 + \mathfrak{R}_7,$$

where

$$\begin{aligned}
\mathfrak{R}_4 &= \frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})(M_t(\beta_{it}) - M_t(\widehat{\beta}_{it}))^2 \\
\mathfrak{R}_5 &= \frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})(M_t(\widehat{\beta}_{it}) - \widehat{\mu}_t(\widehat{\beta}_{it}))^2 \\
\mathfrak{R}_6 &= \frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})|\varepsilon_{it}||M_t(\beta_{it}) - M_t(\widehat{\beta}_{it})| \\
\mathfrak{R}_7 &= \frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})|\varepsilon_{it}||M_t(\widehat{\beta}_{it}) - \widehat{\mu}_t(\widehat{\beta}_{it})|.
\end{aligned}$$

For the first term,

$$
\begin{aligned}
\mathfrak{R}_4 &\leq \frac{4}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})(\mu_t(\beta_{it})-\mu_t(\widehat{\beta}_{it}))^2\\
&+\frac{4}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})(\beta_{it}-\widehat{\beta}_{it})^2(f_t-\mathbb{E}[f_t|\mathcal{F}_t]))^2\\
&\lesssim_{\mathbb{P}} \frac{J}{nT}\mathsf{R}_{nT}^2 = o_{\mathbb{P}}(\frac{J}{nT}).
\end{aligned}
$$

For the second term,

$$
\begin{aligned}
\mathfrak{R}_5 &= \frac{2}{(T-H)^2}\sum_{t=H+1}^{T}\sum_{j=1}^{J_t}\frac{1}{n_t^2\widehat{q}_{jt}^2}\sum_{i=1}^{n_t}\widehat{p}_{j,t}(\beta)\widehat{p}_{j,t}(\widehat{\beta}_{it})(M_t(\widehat{\beta}_{it})-\widehat{\mu}_t(\widehat{\beta}_{it}))^2\\
&= o_{\mathbb{P}}(\frac{J}{nT}).
\end{aligned}
$$

For the last two terms we have

$$
\mathfrak{R}_6 \lesssim_{\mathbb{P}} \frac{J}{nT}\mathsf{R}_{nT} = o_{\mathbb{P}}(\frac{J}{nT}), \qquad \mathfrak{R}_7 \lesssim_{\mathbb{P}} \frac{J}{nT}\sqrt{\frac{J\log(nT)}{n}+\frac{1}{J^2}} = o_{\mathbb{P}}(\frac{J}{nT}).
$$

Finally, we need only show that $\left|\widetilde{\sigma}_{\varepsilon,\mathtt{PI}}^2(\beta)-\sigma_{\varepsilon}^2(\beta)\right| = o_{\mathbb{P}}(r_{n,T}(\beta))$. The above statement holds under the conditions of Theorem SA-1.2 and given the results in Lemma SA-2.2 to Lemma SA-2.3. This completes the proof. □

SA-2.6 Proof of Lemma SA-2.2

For the first result, note that by Theorem SA-1.2 and the assumptions imposed,

$$
\max_{1\leq j\leq J_t-1}\left|\widehat{\beta}_{(k_{jt}),t}-\beta_{(k_{jt}),t}\right| = \max_{1\leq j\leq J_t}\left|F_{\widehat{\beta},n,t}^{-1}(\kappa_{j,t})-F_{\beta,n,t}^{-1}(\kappa_{j,t})\right| \lesssim \mathfrak{R}_{1,t}+\mathfrak{R}_{2,t}+\mathfrak{R}_{3,t},
$$

where

$$
\begin{aligned}
\mathfrak{R}_{1,t} &= \max_{1\leq j\leq J_t}\left|F_{\beta,n,t}^{-1}(\kappa_{j,t})-F_{\beta,t}^{-1}(\kappa_{j,t})\right|,\\
\mathfrak{R}_{2,t} &= \max_{1\leq j\leq J_t}\left|F_{\widehat{\beta},t}^{-1}(\kappa_{j,t})-F_{\beta,t}^{-1}(\kappa_{j,t})\right|,\\
\mathfrak{R}_{3,t} &= \max_{1\leq j\leq J_t}\left|F_{\widehat{\beta},n,t}^{-1}(\kappa_{j,t})-F_{\widehat{\beta},t}^{-1}(\kappa_{j,t})\right|.
\end{aligned}
$$

Note that $\mathfrak{R}_{1,t} \lesssim \mathfrak{R}_{11,t}+\mathfrak{R}_{12,t}$ with

$$
\begin{aligned}
\mathfrak{R}_{11,t} &= \max_{1\leq j\leq J_t}\left|F_{\beta,n,t}^{-1}(\kappa_{j,t})-F_{\beta,t}^{-1}(\kappa_{j,t})-\left[F_{\beta,n,t}(F_{\beta,t}^{-1}(\kappa_{j,t}))-F_{\beta,t}(F_{\beta,t}^{-1}(\kappa_{j,t}))\right]/f_{\beta,t}(F_{\beta,t}^{-1}(\kappa_{j,t}))\right|,\\
\mathfrak{R}_{12,t} &= \max_{1\leq j\leq J_t}\left|F_{\beta,n,t}(F_{\beta,t}^{-1}(\kappa_{j,t}))-F_{\beta,t}(F_{\beta,t}^{-1}(\kappa_{j,t}))\right]\Big|,
\end{aligned}
$$

and it follows by standard arguments that

$$\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{1,t}\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{n}}.$$

Next, we have

$$\begin{aligned}\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{2,t}&=\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\left|F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})-F^{-1}_{\beta,t}(\kappa_{j,t})\right|\\&\lesssim_{\mathbb{P}}\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\sup_{|v|\lesssim\mathsf{R}_{nT}}\left|F^{-1}_{\beta,t}(\kappa_{j,t}+v)-F^{-1}_{\beta,t}(\kappa_{j,t})\right|\lesssim\mathsf{R}_{nT}.\end{aligned}$$

Finally, for $\mathfrak{R}_{3,t}$, we proceed as for $\mathfrak{R}_{1,t}$ but taking into account the first-step estimation. Thus, $\mathfrak{R}_{3,t}\lesssim\mathfrak{R}_{31,t}+\mathfrak{R}_{32,t}+\mathfrak{R}_{33,t}$ with

$$\begin{aligned}\mathfrak{R}_{31,t}&=\max_{1\le j\le J_t}\left|F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t})-F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})-\left[F_{\widehat{\beta},n,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))-F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))\right]/f_{\beta,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))\right|,\\\mathfrak{R}_{32,t}&=\sup_{u\in\mathcal{B},|v|\lesssim\mathsf{R}_{nT}}\left|F_{\beta,n,t}(u+v)-F_{\beta,n,t}(u)-F_{\beta,t}(u+v)+F_{\beta,t}(u)\right|,\\\mathfrak{R}_{33,t}&=\sup_{u\in\mathcal{B}}\left|F_{\beta,n,t}(u)-F_{\beta,t}(u)\right|,\end{aligned}$$

with arbitrary high probability for n and T large enough, uniformly in t, and where $\mathcal{B}=[\beta_l,\beta_u]$. We set $f_{\beta,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))=f_{\beta,t}(\beta_l)$ if $F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})<\beta_l$, and $f_{\beta,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))=f_{\beta,t}(\beta_u)$ if $F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})>\beta_u$. Recall also that we set $\widehat{\beta}_{(0)t}=\beta_l$ and $\widehat{\beta}_{(n_t)t}=\beta_u$ for simplicity.

For $\mathfrak{R}_{31,t}$, using standard results, for any $c\in\mathbb{R}$, we have

$$\begin{aligned}&\mathbb{P}\Big(\sqrt{n}\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{31,t}>v\Big)\le\sum_{t=\lfloor Th\rfloor+1}^{T}\mathbb{P}\Big(\sqrt{n}\mathfrak{R}_{31,t}>v\Big)\\&\le\sum_{t=\lfloor Th\rfloor+1}^{T}\mathbb{P}\Big(\sup_{|u|\le\mathsf{R}_{nT},y\in\mathcal{B}}|Z_{n,\beta,t}(y+c/\sqrt{n_t}+u)-Z_{n,\beta,t}(y+u)|\gtrsim v\Big)\\&\quad+\sum_{t=\lfloor Th\rfloor+1}^{T}\mathbb{P}\Big(\max_{1\le j\le J_t}|\sqrt{n_t}(F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})+c/\sqrt{n_t})-F_{\widehat{\beta},n,t}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t}))/f_{\beta,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))-c|\gtrsim v\Big),\end{aligned}$$

where $Z_{n,\beta,t}(y)=\sqrt{n_t}(F_{\beta,n,t}(y)-F_{\beta,t}(y))$. The first term in the upper bound vanishes due to the modulus of continuity of the empirical distribution function (Stute, 1982), the uniform inequality in Massart (1990) with the union bound, and our imposed rate conditions. For the second term in the upper bound, first note that

$$\begin{aligned}F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})+c/\sqrt{n_t})&=\mathbb{P}(\widehat{\beta}_{it}\le F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})+c/\sqrt{n_t}|\mathcal{G}_{t-1})\\&=\mathbb{P}(\beta_{it}\le\beta_{it}-\widehat{\beta}_{it}+F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})+c/\sqrt{n_t}|\mathcal{G}_{t-1})\\&=F_{\beta,t}(\beta_{it}-\widehat{\beta}_{it}+F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})+c/\sqrt{n_t})\end{aligned}$$

$$= F_{\beta,t}(\beta_{it} - \widehat{\beta}_{it} + F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})) + f_{\beta,t}(\tilde{c})\frac{c}{\sqrt{n_t}},$$

where $\tilde{c}$ is some point between $\beta_{it} - \widehat{\beta}_{it} + F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}) + c/\sqrt{n_t}$ and $\beta_{it} - \widehat{\beta}_{it} + F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})$. Therefore,

$$\begin{aligned}
&\max_{1\le j\le J_t} |\sqrt{n_t}(F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}) + c/\sqrt{n_t}) - F_{\widehat{\beta},n,t}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t}))/f_{\beta,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})) - c| \\
&\lesssim \max_{1\le j\le J_t} |\sqrt{n_t}(F_{\beta,t}(\beta_{it} - \widehat{\beta}_{it} + F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})) - F_{\beta,n,t}(\beta_{it} - \widehat{\beta}_{it} + F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t}))| \\
&\quad + \max_{1\le j\le J_t} \Big|\frac{f_{\beta,t}(\tilde{c})}{f_{\beta,t}(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}))}c - c\Big|.
\end{aligned}$$

Therefore, we have

$$\max_{1\le t\le \lfloor Th\rfloor+1}\max_{\lfloor Th\rfloor+1\le t\le T} \mathfrak{R}_{31,t} \lesssim_{\mathbb{P}} \sqrt{\frac{\mathsf{R}_{nT}\log(nT)}{n}} + \frac{\mathsf{R}_{nT}}{\sqrt{n}}.$$

Similarly, for $\mathfrak{R}_{2,t}$ and $\mathfrak{R}_{3,t}$, by the modulus of continuity of the empirical distribution function Stute (1982, Lemma 2.3), Assumption 6, and the uniform inequality in Massart (1990) with the union bound, we obtain

$$\max_{\lfloor Th\rfloor+1\le t\le T} \mathfrak{R}_{32,t} \lesssim_{\mathbb{P}} \sqrt{\frac{\mathsf{R}_{nT}\log(nT)}{n}},$$

and

$$\max_{\lfloor Th\rfloor+1\le t\le T} \mathfrak{R}_{33,t} \lesssim_{\mathbb{P}} \sqrt{\frac{\log(nT)}{n}}.$$

For the upper bound of the second result, define $U_{(k_{(j-1)t}),t} = F_{\beta,t}(\beta_{(k_{(j-1)t}),t})$, and note that $U_{(k_{jt}),t} - U_{(k_{j-1t}),t}$ follows a Beta distribution with parameter $(k_{jt} - k_{j-1t}, n_t + 1 - (k_{jt} - k_{j-1t}))$ conditional on $\mathcal{G}_{t-1}$, and thus $\mathbb{E}[U_{(k_{jt}),t} - U_{(k_{j-1t}),t}|\mathcal{G}_{t-1}] = (k_{jt} - k_{j-1t})/(n_t + 1)$. Employing a Taylor series expansion of $F^{-1}_{\beta,t}(b)$ and following B.10 in Bobkov, Gentil, and Ledoux (2001), we verify

$$\begin{aligned}
&\mathbb{P}\Big(\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t} |\beta_{(k_{jt}),t} - \beta_{(k_{(j-1)t}),t}| \gtrsim J^{-1}\Big) \\
&\lesssim \sum_{t=\lfloor Th\rfloor+1}^{T}\sum_{j=1}^{J_t} \mathbb{E}\Big[\mathbb{P}\big(|\beta_{(k_{jt}),t} - \beta_{(k_{(j-1)t}),t}| \gtrsim J^{-1}|\mathcal{G}_{t-1}\big)\Big] \\
&\lesssim \sum_{t=\lfloor Th\rfloor+1}^{T}\sum_{j=1}^{J_t} \mathbb{E}\Big[\mathbb{P}\big(|U_{(k_{jt}),t} - U_{(k_{j-1t}),t}| \gtrsim J^{-1}|\mathcal{G}_{t-1}\big)\Big] \\
&\lesssim TJ\max_{\lfloor Th\rfloor+1\le t\le T} \exp\Big(-C(n_t+1)(J_t^{-1}(1-(n_t+1)^{-1}))^2\Big),
\end{aligned}$$

for a positive constant C. It follows that the upper bound in the last display goes to 0 under the rate conditions imposed. The lower bound in the second result of the lemma is proven similarly using the results in Skorski (2023).

For the third results, it follows the same steps as for $\max_{1\le j\le J_t-1}\left|\widehat{\beta}_{(k_{jt}),t}-\beta_{(k_{jt}),t}\right|$ except for a additional differencing step:

$$\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\left|\widehat{\beta}_{(k_{jt}),t}-\beta_{(k_{jt}),t}-[\widehat{\beta}_{(k_{(j-1)t}),t}-\beta_{(k_{(j-1)t}),t}]\right|\lesssim\sqrt{\frac{\log(nT)}{nJ}}+\frac{\mathsf{R}_{nT}}{J}.$$

We shall break to the following term,

$$\begin{aligned}
\mathfrak{R}_{1,t}&=\max_{1\le j\le J_t}\left|F^{-1}_{\beta,n,t}(\kappa_{j,t})-F^{-1}_{\beta,t}(\kappa_{j,t})-(F^{-1}_{\beta,n,t}(\kappa_{j-1,t})-F^{-1}_{\beta,t}(\kappa_{j-1,t}))\right|,\\
\mathfrak{R}_{2,t}&=\max_{1\le j\le J_t}\left|F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})-F^{-1}_{\beta,t}(\kappa_{j,t})-(F^{-1}_{\widehat{\beta},t}(\kappa_{j-1,t})-F^{-1}_{\beta,t}(\kappa_{j-1,t}))\right|,\\
\mathfrak{R}_{3,t}&=\max_{1\le j\le J_t}\left|F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t})-F^{-1}_{\widehat{\beta},t}(\kappa_{j,t})-(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j-1,t})-F^{-1}_{\widehat{\beta},t}(\kappa_{j-1,t}))\right|.
\end{aligned}$$

Similar to the previous step, we have $\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{1,t}\lesssim_p\sqrt{\frac{\log(nT)}{nJ}}$.

Next,

$$\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{2,t}\lesssim\mathsf{R}_{nT}/J.$$

Then,

$$\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{3,t}\lesssim_{\mathbb{P}}\sqrt{\frac{\mathsf{R}_{nT}\log(nT)}{nJ}}+\frac{\mathsf{R}_{nT}}{J\sqrt{n}}.$$

Thus the conclusion holds.

For the next result,

$$\begin{aligned}
&\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\left[\mathbb{1}(\widehat{\beta}_{it}\in\widehat{P}_{jt})-\mathbb{1}(\beta_{it}\in P_{jt})\right]\right|\\
&=\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\left[\mathbb{1}\left(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t})<\widehat{\beta}_{it}\le F^{-1}_{\widehat{\beta},n,t}(\kappa_{j+1,t})\right)-\mathbb{1}\left(F^{-1}_{\beta,t}(\kappa_{j,t})<\beta_{it}\le F^{-1}_{\beta,t}(\kappa_{j+1,t})\right)\right]\right|\\
&\le\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{4,t}+\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{5,t},
\end{aligned}$$

where

$$\begin{aligned}
\mathfrak{R}_{4,t}&=\max_{1\le j\le J_t}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\left[\mathbb{1}\left(F^{-1}_{\beta,n,t}(\kappa_{j,t})<\beta_{it}\le F^{-1}_{\beta,n,t}(\kappa_{j+1,t})\right)-\mathbb{1}\left(F^{-1}_{\beta,t}(\kappa_{j,t})<\beta_{it}\le F^{-1}_{\beta,t}(\kappa_{j+1,t})\right)\right]\right|\\
\mathfrak{R}_{5,t}&=\max_{1\le j\le J_t}\left|\frac{1}{n_t}\sum_{i=1}^{n_t}\left[\mathbb{1}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t})<\widehat{\beta}_{it}\le F^{-1}_{\widehat{\beta},n,t}(\kappa_{j+1,t}))-\mathbb{1}\left(F^{-1}_{\beta,n,t}(\kappa_{j,t})<\beta_{it}\le F^{-1}_{\beta,n,t}(\kappa_{j+1,t})\right)\right]\right|.
\end{aligned}$$

For the term $\mathfrak{R}_{4,t}$,

$$\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{4,t}\le \max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{41,t}+\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{42,t}\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{nJ}},$$

where

$$\mathfrak{R}_{41,t}=\max_{1\le j\le J_t}\Big|F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j+1,t}))-F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j,t}))-\big[F_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j+1,t}))-F_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j,t}))\big]\Big|,$$

and

$$\begin{aligned}\mathfrak{R}_{42,t}=\max_{1\le j\le J_t}\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}&\big[\mathbb{1}\big(F^{-1}_{\beta,n,t}(\kappa_{j,t})<\beta_{it}\le F^{-1}_{\beta,n,t}(\kappa_{j+1,t})\big)-\mathbb{1}\big(F^{-1}_{\beta,t}(\kappa_{j,t})<\beta_{it}\le F^{-1}_{\beta,t}(\kappa_{j+1,t})\big)\big]\\&-\big[F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j+1,t}))-F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j,t}))-\big[F_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j+1,t}))-F_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j,t}))\big]\Big|.\end{aligned}$$

To see the above result, notice that

$$\begin{aligned}&\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{41,t}\\&\lesssim \max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\Big|f_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j+1,t}))[F^{-1}_{\beta,n,t}(\kappa_{j+1,t})-F^{-1}_{\beta,t}(\kappa_{j+1,t})]-f_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j,t}))[F^{-1}_{\beta,n,t}(\kappa_{j,t}))-F^{-1}_{\beta,t}(\kappa_{j,t})]\Big|,\\&\quad+\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\Big|F^{-1}_{\beta,n,t}(\kappa_{j+1,t})-F^{-1}_{\beta,t}(\kappa_{j+1,t})\Big|^2\\&\lesssim \max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\Big|f_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j+1,t}))-f_{\beta,t}(F^{-1}_{\beta,t}(\kappa_{j,t}))\Big|\Big|F^{-1}_{\beta,n,t}(\kappa_{j+1,t})-F^{-1}_{\beta,t}(\kappa_{j+1,t})\Big|\\&\quad+\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\Big|F^{-1}_{\beta,n,t}(\kappa_{j+1,t})-F^{-1}_{\beta,t}(\kappa_{j+1,t})]-[F^{-1}_{\beta,n,t}(\kappa_{j,t}))-F^{-1}_{\beta,t}(\kappa_{j,t})]\Big|\\&\quad+\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\Big|F^{-1}_{\beta,n,t}(\kappa_{j+1,t})-F^{-1}_{\beta,t}(\kappa_{j+1,t})\Big|^2\\&\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{nJ}},\end{aligned}$$

and employing standard empirical process theory we also verify that

$$\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{42,t}\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{n^{3/2}J}}.$$

Finally, for the term $\mathfrak{R}_{5,t}$, we have

$$\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{5,t}\le \max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{51,t}+\max_{\lfloor Th\rfloor+1\le t\le T}\mathfrak{R}_{52,t}\lesssim_{\mathbb{P}}\frac{\mathsf{R}_{nT}}{J}\sqrt{\log(nT)}+\sqrt{\frac{\mathsf{R}_{nT}}{nJ}}\sqrt{\log(nT)},$$

where

$$\mathfrak{R}_{51,t}=\max_{1\le j\le J_t}\Big|F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j+1,t}))-F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t}))-\big[F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j+1,t}))-F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j,t}))\big]\Big|$$

and

$$\mathfrak{R}_{52,t} = \max_{1\le j\le J_t}\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}\big[\mathbb{1}\big(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t}) < \widehat{\beta}_{it} \le F^{-1}_{\widehat{\beta},n,t}(\kappa_{j+1,t})\big) - \mathbb{1}\big(F^{-1}_{\widehat{\beta},t}(\kappa_{j,t}) < \widehat{\beta}_{it} \le F^{-1}_{\widehat{\beta},t}(\kappa_{j+1,t})\big)\big]$$

$$- \big[F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j+1,t})) - F_{\widehat{\beta},t}(F^{-1}_{\widehat{\beta},n,t}(\kappa_{j,t})) - \big[F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j+1,t})) - F_{\beta,t}(F^{-1}_{\beta,n,t}(\kappa_{j,t}))\big]\Big|.$$

and the proof is completed using the same logic as before.

Finally, the last result follows by Bernstein's inequality and standard calculations. □

SA-2.7 Proof of Lemma SA-2.3

For the first result, we have

$$\begin{aligned}
&\max_{\lfloor Th\rfloor+1\le t\le T}\Big|(\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1} - Q_t^{-1}\Big|_\infty \\
&\le \max_{\lfloor Th\rfloor+1\le t\le T}|(\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1}|_\infty \max_{1\le j\le J_t}\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}(\mathbb{1}(\widehat{\beta}_{it}\in\widehat{P}_{jt}) - \mathbb{1}(\beta_{it}\in P_{jt}))\Big||(\Phi_t\Phi_t^\top/n_t)^{-1}|_\infty \\
&\quad + \max_{\lfloor Th\rfloor+1\le t\le T}|(\Phi_t\Phi_t^\top/n_t)^{-1} - Q_t^{-1}|_\infty \\
&\lesssim_{\mathbb{P}} J^2\mathsf{L}_{nT} + J^2\sqrt{\frac{\log(nT)}{nJ}}.
\end{aligned}$$

For the second result, using the union bound, Markov's inequality, the conditional on $\mathcal{F}_{t-1}, f_t$ i.i.d. property of ε_{it}, and Bernstein's inequality,

$$\begin{aligned}
&\mathbb{P}\Big(\max_{\lfloor Th\rfloor+1\le t\le T}\max_{1\le j\le J_t}\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,j,t}\varepsilon_{it}\Big| > u\Big) \\
&\lesssim \sum_{t=\lfloor Th\rfloor+1}^{T}\sum_{j=1}^{J_t}\mathbb{P}\Big(\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,j,t}(\varepsilon_{it}\mathbb{1}(|\varepsilon_{it}|>M) - \mathbb{E}[\varepsilon_{it}\mathbb{1}(|\varepsilon_{it}|>M)|\mathcal{F}_{t-1},f_t])\Big| > u\Big) \\
&\quad + \sum_{t=\lfloor Th\rfloor+1}^{T}\sum_{j=1}^{J_t}\mathbb{P}\Big(\Big|\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,j,t}(\varepsilon_{it}\mathbb{1}(|\varepsilon_{it}|\le M) - \mathbb{E}[\varepsilon_{it}\mathbb{1}(|\varepsilon_{it}|\le M)|\mathcal{F}_{t-1},f_t])\Big| > u\Big) \\
&\lesssim \frac{T}{u^2M^{q-2}n} + \sum_{t=\lfloor Th\rfloor+1}^{T}\sum_{j=1}^{J_t}\mathbb{E}\Big[\exp\Big(-\frac{nu^2/2}{\vartheta_{j,t}+Mu/3}\Big)\Big],
\end{aligned}$$

where $\vartheta_{j,t} = \max_{1\le i\le n_t}\mathbb{E}[\Phi^2_{i,j,t}\varepsilon^2_{it}|\mathcal{F}_{t-1},f_t] \lesssim_{\mathbb{P}} J^{-1}$. Thus, setting $M = \sqrt{nJ^{-1}/\log(nT)}$ and $u = C\sqrt{\frac{\log(nT)}{nJ}}$ for C large enough, the result follows.

The third result is proven similarly using $\max_{1\le i\le n_t}\mathbb{E}[(\widehat{\Phi}_{i,j,t}-\Phi_{i,j,t})^2\varepsilon^2_{it}|\mathcal{F}_{t-1},f_t] \lesssim_{\mathbb{P}} \mathsf{L}_{nT}$.

For the fourth result,

$$\max_{\lfloor Th\rfloor+1\le t\le T}\Big|(f_t - \mathbb{E}[f_t|\mathcal{G}_{t-1}])\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,t}\beta_{it}\Big|_\infty$$

$$
\begin{aligned}
&\leq \max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\frac{1}{n_t}\sum_{i=1}^{n_t}\big(\Phi_{i,t}\beta_{it}-\mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}]\big)\Big|_\infty\\
&\quad+\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}]\Big|_\infty\\
&\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)}{nJ}}+\frac{\sqrt{\log(nT)}}{J},
\end{aligned}
$$

using Bernstein's inequality conditional on $\mathcal{G}_{t-1}$.

The fifth result follows similarly to the third result.

$$
\begin{aligned}
&\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\frac{1}{n_t}\sum_{i=1}^{n_t}(\widehat{\Phi}_{i,t}-\Phi_{i,t})\beta_{it}\Big|_\infty\\
&\leq \max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\frac{1}{n_t}\sum_{i=1}^{n_t}(\widehat{\Phi}_{i,t}-\Phi_{i,t})\beta_{it}-\mathbb{E}[(\widehat{\Phi}_{i,t}-\Phi_{i,t})\beta_{it}|\mathcal{G}_{t-1}]\big)\Big|_\infty\\
&\quad+\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\frac{1}{n_t}\sum_{i=1}^{n_t}\mathbb{E}[(\widehat{\Phi}_{i,t}-\Phi_{i,t})\beta_{it}|\mathcal{G}_{t-1}]\Big|_\infty\\
&\lesssim_{\mathbb{P}}\sqrt{\frac{\log(nT)\mathsf{L}_{nT}}{n}}+\sqrt{\log(nT)}\mathsf{L}_{nT}.
\end{aligned}
$$

For the sixth result,

$$
\begin{aligned}
&\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\widehat{\Phi}_{i,t}\widehat{\Phi}_{j,t}^\top\beta_{it}\beta_{jt}\Big|_\infty\\
&\lesssim \max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\widehat{\Phi}_{i,t}\widehat{\Phi}_{j,t}^\top\beta_{it}\beta_{jt}-\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\Phi_{i,t}\Phi_{j,t}^\top\beta_{it}\beta_{jt}\Big|_\infty\\
&+\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\{\Phi_{i,t}\beta_{it}-\mathbb{E}[\Phi_{i,t}\beta_{it}|\mathcal{G}_{t-1}]\}\{\Phi_{j,t}^\top\beta_{jt}-\mathbb{E}[\Phi_{j,t}^\top\beta_{it}|\mathcal{G}_{t-1}]\}\Big|_\infty\\
&+\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\beta_{it}\Phi_{i,t}\mathbb{E}[\Phi_{j,t}^\top\beta_{jt}|\mathcal{G}_{t-1}]\Big|_\infty+\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\mathbb{E}[\beta_{it}\Phi_{i,t}|\mathcal{G}_{t-1}]\Phi_{j,t}^\top\beta_{jt}\Big|_\infty\\
&+\max_{\lfloor Th\rfloor+1\leq t\leq T}\Big|\frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\mathbb{E}[\beta_{it}\Phi_{i,t}|\mathcal{G}_{t-1}]\mathbb{E}[\Phi_{j,t}^\top\beta_{jt}|\mathcal{G}_{t-1}]\Big|_\infty\\
&\lesssim_{\mathbb{P}}\frac{\mathsf{R}_{nT}}{J^2}+\frac{1}{J}\sqrt{\frac{\log(nT)}{nJ}}+\frac{1}{nJ}+\frac{1}{J^2}\lesssim\frac{1}{J^2}.
\end{aligned}
$$

The last result follows similarly to the proof of the previous terms. □

SA-2.8 Proof of Lemma SA-2.4

We have

$$
\begin{aligned}
&\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}\big|\widehat{\mu}_t(\beta)-M_t(\beta)\big|\\
&=\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}\big|\widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\widehat{\Phi}_tR_t-\mu_t(\beta)-\beta(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}])\big|\\
&\le\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}\Big|\widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,t}\varepsilon_{it}\Big|+\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{B}_t(\beta)|+\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{R}_t(\beta)|
\end{aligned}
$$

where

$$
\mathscr{B}_t(\beta)=\widehat{p}_t(\beta)^\top(\widehat{\Phi}_t\widehat{\Phi}_t^\top)^{-1}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\big(\mu_t(\beta_{it})-\widehat{\Phi}_{i,t}^\top a_t^\circ\big)+\big(\widehat{p}_t(\beta)^\top a_t^\circ-\mu_t(\beta)\big),
$$

with $a_t^\circ=(\mathbb{E}[\Phi_{i,t}\Phi_{i,t}^\top|\mathcal{G}_{t-1}])^{-1}\mathbb{E}[\Phi_{i,t}R_{it}|\mathcal{G}_{t-1}]$, and $\mathscr{R}_t(\beta)=\mathfrak{R}_{1t}(\beta)+\mathfrak{R}_{2t}(\beta)+\mathfrak{R}_{3t}(\beta)+\mathfrak{R}_{4t}(\beta)$ with

$$
\begin{aligned}
\mathfrak{R}_{1t}(\beta)&=\widehat{p}_t(\beta)^\top\Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1}-Q_t^{-1}\Big)\frac{1}{n_t}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\varepsilon_{it},\\
\mathfrak{R}_{2t}(\beta)&=\widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\big(\widehat{\Phi}_{i,t}-\Phi_{i,t}\big)\varepsilon_{it},\\
\mathfrak{R}_{3t}(\beta)&=\widehat{p}_t(\beta)^\top\Big((\widehat{\Phi}_t\widehat{\Phi}_t^\top/n_t)^{-1}-Q_t^{-1}\Big)\frac{1}{n_t}\sum_{i=1}^{n_t}\widehat{\Phi}_{i,t}\beta_{it}(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}]),\\
\mathfrak{R}_{4t}(\beta)&=\widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\big(\widehat{\Phi}_{i,t}-\Phi_{i,t}\big)\beta_{it}(f_t-\mathbb{E}[f_t|\mathcal{G}_{t-1}]),
\end{aligned}
$$

Proceeding as in the proof of Lemma 4.1,

$$
\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}\Big|\widehat{p}_t(\beta)^\top Q_t^{-1}\frac{1}{n_t}\sum_{i=1}^{n_t}\Phi_{i,t}\varepsilon_{it}\Big|\lesssim_{\mathbb{P}}\sqrt{\frac{J\log(nT)}{n}},
$$

$$
\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{B}_t(\beta)|\lesssim_{\mathbb{P}}\frac{1}{J},
$$

$$
\begin{aligned}
\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{R}_{1t}(\beta)|&\lesssim_{\mathbb{P}}\Big(J^2\mathsf{L}_{nT}+J^2\sqrt{\frac{\log(nT)}{nJ}}\Big)\Big(J\sqrt{\frac{\log(nT)\mathsf{L}_{nT}}{n}}+\sqrt{\frac{\log(nT)}{nJ}}\Big)\\
&=o\Big(\sqrt{\frac{J\log(nT)}{n}}\Big),
\end{aligned}
$$

$$\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{R}_{2t}(\beta)| \lesssim_{\mathbb{P}} J\sqrt{\frac{\log(nT)\mathsf{L}_{nT}}{n}},$$

$$\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{R}_{3t}(\beta)| \lesssim_{\mathbb{P}} \mathsf{J}\mathsf{L}_{nT} + J\sqrt{\frac{\log(nT)}{nJ}},$$

and

$$\max_{H+1\le t\le T}\sup_{\beta\in\mathcal{B}}|\mathscr{R}_{4t}(\beta)| \lesssim_{\mathbb{P}} \mathsf{J}\mathsf{L}_{nT}.$$

This completes the proof. □

SA-3 References

Bobkov, S. G., I. Gentil, and M. Ledoux (2001): "Hypercontractivity of Hamilton-Jacobi equations," *Journal de Mathématiques Pures et Appliquées*, 80(7), 669–696.

Freedman, D. A. (1975): "On tail probabilities for martingales," *Annals of Probability*, 3(1), 100–118.

Hall, P., and C. C. Heyde (2014): *Martingale limit theory and its application*. Academic press.

Massart, P. (1990): "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality," *Annals of Probability*, pp. 1269–1283.

Skorski, M. (2023): "Bernstein-type bounds for beta distribution," *Modern Stochastics: Theory and Applications*, 10(2), 211–228.

Stute, W. (1982): "The oscillation behavior of empirical processes," *Annals of Probability*, pp. 86–107.

Zhang, T., and W. B. Wu (2012): "Inference of time-varying regression models," *Annals of Statistics*, 40(3), 1376–1402.