Online Appendix: Local GMM estimation for nonparametric time-varying coefficient moment condition models

The online appendix is organized as follows. Section A provides proofs of Lemma 1, Theorem 1 and Corollaries 1 and 2. Section B presents auxiliary results and their proofs.

Notation: $\|\cdot\|$ is the Euclidean norm. $\|\cdot\|_p$ is the L_p norm. $\|\cdot\|_{sp}$ is the spectral norm of a matrix. $x_n = O(y_n)$ states that the deterministic sequence x_n is at most of order y_n. $x_n = O_p(y_n)$ states that the vector of random variables x_n is at most of order y_n in probability, and $x_n = o_p(y_n)$ is of smaller order in probability than y_n . The operator $\xrightarrow{p}$ denotes convergence in probability, and $\xrightarrow{d}$ denotes convergence in distribution. We use C for a generic positive (vector) of constant(s) when convenient. $\sigma(\mathcal{A})$ denotes the σ-algebra generated by a collection of sets $\mathcal{A}$.

A Proof of main results

A.1 Proof of Lemma 1

Proof of (i). By Triangular inequality,

$$\begin{aligned}\left\|\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\Big(g_j(\theta_t)-E(g_t(\theta_t))\Big)\right\| &\leqslant \left\|\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\Big(g_j(\theta_t)-E(g_j(\theta_t))\Big)\right\| \\ &\quad+\left\|\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\Big(E(g_j(\theta_t))-E(g_t(\theta_t))\Big)\right\| \\ &=\left\|M_{T,t}^{(1)}(\theta_t)\right\|+\left\|M_{T,t}^{(2)}(\theta_t)\right\|.\end{aligned}$$

(6) can be obtained using

$$\max_{1\leqslant t\leqslant T}\left\|\bar{g}_{T,t}(\theta_t)\right\| \leqslant \max_{1\leqslant t\leqslant T}\left\|M_{T,t}^{(1)}(\theta_t)\right\| + \max_{1\leqslant t\leqslant T}\left\|M_{T,t}^{(2)}(\theta_t)\right\|.$$

It follows immediately from Lemma B1(1b) that, for any $\varepsilon > 0$, $p > 2$,

$$\max_{1\leqslant t\leqslant T}\left\|M_{T,t}^{(1)}(\theta_t)\right\| = O_p((Tb)^{-1/2}\log^{1/2}T + (T^2b)^{1/p}(Tb)^{\varepsilon-1}).$$

Under Assumption 3.5(ii), for sufficiently small $\varepsilon > 0$, it holds $(T^2b)^{1/p}(Tb)^{\varepsilon-1} \leqslant (Tb)^{-1/2}$, this implies that

$$\max_{1\leqslant t\leqslant T}\left\|M_{T,t}^{(1)}(\theta_t)\right\| = O_p((Tb)^{-1/2}\sqrt{\log T}).$$

For $M^{(2)}_{T,t}(\theta_t)$, we have

$$\left\|M^{(2)}_{T,t}(\theta_t)\right\| \leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left\|E(g_j(\theta_t)) - E(g_t(\theta_t))\right\| \leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left\|E(g_j(\theta_t)) - E\left(g_t(\theta_t)\right)\right\|. \qquad \text{(A.1)}$$

By mean-value theorem, we have

$$g_j(\theta_t) = g_j(\theta_j) + \frac{\partial g_j(\overline{\theta}_t)}{\partial \theta'}\left(\theta_t - \theta_j\right),$$

where $\overline{\theta}_t$ lies between θ_t and θ_j. Then, by continuing from (A.1), we have

$$\left\|M^{(2)}_{T,t}(\theta_t)\right\| \leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left\|E\left(\frac{\partial g_j(\overline{\theta}_t)}{\partial \theta'}\right)\right\|\left\|\theta_t - \theta_j\right\| \leqslant C\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left(\frac{|j-t|}{T}\right) \approx Cb\int K(u)du = O(b).$$

This is because $\max_{1\leqslant t\leqslant T}\left\|E\left(\frac{\partial g_t(\overline{\theta}_t)}{\partial \theta'}\right)\right\| < \infty$ (Assumption 3.3) and the fact that continuously differentiable of $\theta(\cdot)$ on $[0,1]$ also implies that it is Lipschitz continuous. This completes the proof of (6).

(5) can be obtained similarly by noticing that

$$\max_{\theta\in\Theta}\left\|\overline{g}_{T,t}(\theta)\right\| \leqslant \max_{\theta\in\Theta}\left\|M^{(1)}_{T,t}(\theta)\right\| + \max_{\theta\in\Theta}\left\|M^{(2)}_{T,t}(\theta)\right\|.$$

Then, the results follow from Lemma B1(1a) and Lemma B1(2a).

Proof of (ii). Write

$$\begin{aligned}\frac{1}{\sqrt{Tb}}\sum_{j=1}^{T} k_{jt}g_j(\theta_t) &= \frac{1}{\sqrt{Tb}}\sum_{j=1}^{T} k_{jt}\Big(g_j(\theta_t) - g_j(\theta_j) + g_j(\theta_j)\Big)\\ &= \frac{1}{\sqrt{Tb}}\sum_{j=1}^{T} k_{jt}\left(g_j(\theta_t) - g_j(\theta_j)\right) + \frac{1}{\sqrt{Tb}}\sum_{j=1}^{T} k_{jt}g_j(\theta_j)\\ &:= B_{T,t} + V_{T,t}.\end{aligned}$$

We first show that

$$V_{T,t} \xrightarrow{d} \mathcal{N}\left(0, v_0 W_t\right),$$

where $W_t = \mathrm{Var}\left(g_t(\theta_t)\right)$[1]. We will proceed by assuming $m = 1$, since the case when $m > 1$ follows immediately from Cramér–Wold device.

Since $g_t(\theta_t)$ is a martingale difference sequence (M.D.S.), by the central limit theorem (CLT)

[1] W_t depends on θ_t, but we write W_t for convenience, and change the notation to $W_t(\cdot)$ when needed.

for M.D.S. (e.g. Theorem 3.2 in Hall and Heyde (1980)), we need to verify that

$$\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 g_j^2(\theta_j) \xrightarrow{p} v_0 W_t, \tag{A.2}$$

$$\max_j \left| \frac{1}{\sqrt{Tb}} k_{jt} g_j(\theta_j) \right| \xrightarrow{p} 0, \;\; E\left(\max_j \; \frac{1}{Tb} k_{jt}^2 g_j^2(\theta_j) \right) < \infty. \tag{A.3}$$

Proof of (A.2). Write

$$\begin{aligned} \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 g_j^2(\theta_j) &= \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 \left(g_j^2(\theta_j) - E\left(g_j^2(\theta_j)\right)\right) + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 E\left(g_j^2(\theta_j)\right) \\ &:= j_T^{(1)} + j_T^{(2)}. \end{aligned}$$

Observe that $g_j^2(\theta_j) - E\left(g_j^2(\theta_j)\right)$ also satisfies Assumptions 3.1(ii) and 3.2. Then, by Lemma B1(i)(a), we have $\left|j_T^{(1)}\right| = O_p\left((Tb)^{-1/2}\right) = o_p(1)$. For $j_T^{(2)}$, we have

$$\begin{aligned} \left|j_T^{(2)}\right| &\leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 \left|E(g_j^2(\theta_j)) - E(g_t^2(\theta_j))\right| + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2\, E\left(g_t^2(\theta_j)\right) \\ &:= j_T^{(21)} + j_T^{(22)}. \end{aligned}$$

It follows by Assumption 3.4(i) that

$$j_T^{(21)} \leqslant C\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 \left(\frac{|j-t|}{T}\right) \approx Cb\int K^2(u)du = O(b),$$

where the approximation follows from the Riemann sum approximation of an integral. For $j_T^{(22)}$, by applying mean-value theorem, we have (as $Tb \to \infty$)

$$\begin{aligned} j_T^{(22)} &= \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2\, E\left(g_t^2(\theta_t)\right) + \frac{2}{Tb}\sum_{j=1}^{T} k_{jt}^2\, E\left(g_t(\bar\theta_t)\frac{\partial g_t(\bar\theta_t)}{\partial \theta'}\right)\left(\theta_j - \theta_t\right) \\ &= v_0 W_t + \frac{2}{Tb}\sum_{j=1}^{T} k_{jt}^2\, E\left(g_t(\bar\theta_t)\frac{\partial g_t(\bar\theta_t)}{\partial \theta'}\right)\theta_t^{(1)}\left(\frac{|j-t|}{T}\right) \\ &= v_0 W_t + O(b) = v_0 W_t + o(1), \end{aligned}$$

where the second equality follows from Taylor series expansion of θ_j around θ_t. This established (A.2).

Proof of (A.3). If Assumption 3.2 holds, by Theorem 12.10 in Davidson (1994), we have $E\left[g_j^2(\theta_j)\mathbb{1}\left(\left|g_j(\theta_j)\right| \geqslant \sqrt{Tb}\varepsilon\right)\right] \to 0$ as $Tb \to \infty$. Together with the fact that $(Tb)^{-1}\sum_{j=1}^{T} k_{jt}^2 =$

$O(1)$, we have, as $Tb \to \infty$,

$$P\left(\max_j \left|\frac{1}{\sqrt{Tb}} k_{jt} g_j(\theta_j)\right| \geqslant \varepsilon\right) \leqslant \varepsilon^{-2}(Tb)^{-1} \sum_{j=1}^{T} k_{jt}^2 E\left[g_j^2(\theta_j) \mathbb{1}\left(\left|g_j(\theta_j)\right| \geqslant \sqrt{Tb}\varepsilon\right)\right] \to 0,$$

for any $\varepsilon > 0$. This establishes (A.3).

We next move on to the analysis of $(Tb)^{-1/2} B_{T,t}$. By mean-value theorem, we have

$$\begin{aligned}
(Tb)^{-1/2} B_{T,t} &= \frac{1}{Tb} \sum_{j=1}^{T} k_{jt}\left(g_j(\theta_t) - g_j(\theta_j)\right) \\
&= \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} \frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'}\left(\theta_t - \theta_j\right) \\
&= \frac{1}{Tb} \sum_{j=1}^{T} k_{jt}\left(\frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'} - E\left(\frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'}\right)\right)\theta_t^{(1)}\left(\frac{|j-t|}{T}\right) + \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} E\left(\frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'}\right)\theta_t^{(1)}\left(\frac{|j-t|}{T}\right) \\
&:= B_{T,t}^{(1)} + B_{T,t}^{(2)},
\end{aligned}$$

where the third equality follows again from Taylor series expansion of θ_j at θ_t. For $B_{T,t}^{(1)}$, since $\frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'} - E\left(\frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'}\right)$ also satisfies Assumptions 3.1(ii) and 3.2, by Lemma B1(2)(a), we have $\left|B_{T,t}^{(1)}\right| = O_p(b) = o_p(1)$. Following similar analysis as for (A.2), we have

$$\begin{aligned}
B_{T,t}^{(2)} &= \frac{1}{Tb} \sum_{j=1}^{T} k_{jt}\left(E\left(\frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'}\right) - E\left(\frac{\partial g_t(\bar{\theta}_t)}{\partial \theta'}\right) + E\left(\frac{\partial g_t(\bar{\theta}_t)}{\partial \theta'}\right)\right)\theta_t^{(1)}\left(\frac{|j-t|}{T}\right) \\
&= O_p(b) + E\left(\frac{\partial g_t(\bar{\theta}_t)}{\partial \theta'}\right)\theta_t^{(1)} \frac{1}{Tb} \sum_{j=1}^{T} k_{jt}\left(\frac{|j-t|}{T}\right) = O_p(b).
\end{aligned}$$

This implies that $B_{T,t} = O_p\left(T^{1/2} b^{3/2}\right)$.

A.2 Proof of Theorem 1

The local CU-GMM estimator is defined by (3):

$$\hat{\theta}_t = \arg\min_{\theta \in \Theta} Q_{t,T},$$

where the criteria function $Q_{t,T}$ is given by

$$Q_{t,T}(\theta) = \overline{g}'_{T,t}(\theta)\overline{W}_{T,t}^{-1}(\theta)\overline{g}_{T,t}(\theta).$$

We first prove the consistency of the estimator. Let

$$Q_t(\theta) = \Big(E\big[g_t(\theta)\big]\Big)' (v_0 W_t(\theta))^{-1} \Big(E\big[g_t(\theta)\big]\Big),$$

where $v_0 = \int K^2(u)du$ and $W_t(\theta) = \mathrm{Var}(g_t(\theta))$. In view of Theorem 2.1 in Newey and McFadden (1994), it is sufficient to verify that

(i) Θ is compact (assumed in Assumption 3.3);

(ii) $Q_t(\theta)$ is uniquely minimized at θ_t (implies by Assumption 3.3);

(iii) $Q_t(\theta)$ is continuous in Θ (implied in Assumption 3.2));

(iv) Uniform consistency:

$$\max_{\theta\in\Theta} \left|Q_{t,T}(\theta) - Q_t(\theta)\right| \xrightarrow{p} 0.$$

Thus, it remains to show (iv), which follows from

$$\max_{\theta\in\Theta} \left\|\overline{g}_{T,t}(\theta) - E(g_t(\theta))\right\| \xrightarrow{p} 0, \tag{A.4}$$

$$\max_{\theta\in\Theta} \left\|\overline{W}_{T,t}^{-1}(\theta) - v_0^{-1} W_t^{-1}(\theta)\right\| \xrightarrow{p} 0. \tag{A.5}$$

(A.4) is exactly (5) in Lemma 1(i). For (A.5), notice that

$$\max_{\theta\in\Theta} \left\|\overline{W}_{T,t}^{-1}(\theta) - v_0^{-1} W_t^{-1}(\theta)\right\| \leqslant \max_{\theta\in\Theta} \|W_t(\theta)\|_{sp}^{-1} \max_{\theta\in\Theta} \left\|v_0 W_t(\theta) - \overline{W}_{T,t}(\theta)\right\|_{sp} v_0^{-1} \max_{\theta\in\Theta} \left\|\overline{W}_{T,t}(\theta)\right\|^{-1}.$$

Recall that, for each $\theta \in \Theta$,

$$\begin{aligned}\overline{W}_{T,t}(\theta) &= \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 \left(g_j(\theta)g_j'(\theta) - E\left(g_j(\theta)g_j'(\theta)\right)\right) + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 E\left(g_j(\theta)g_j'(\theta)\right) \\ &= \overline{W}_{T,t}^{(1)}(\theta) + \overline{W}_{T,t}^{(2)}(\theta).\end{aligned}$$

Following same arguments as used for (A.2), we have $\max_{\theta\in\Theta} \left\|\overline{W}_{T,t}^{(1)}(\theta)\right\| = O_p\left((Tb)^{-1/2}\right)$, $\overline{W}_{T,t}^{(2)}(\theta) = v_0 W_t(\theta) + o(1)$. This further implies that $\max_{\theta\in\Theta} \left\|v_0 W_t(\theta) - \overline{W}_{T,t}(\theta)\right\|_{sp} = O_p\left((Tb)^{-1/2}\right) = o_p(1)$. Together with Assumption 3.2 (which implies that both $\max_{\theta\in\Theta} \|W_t(\theta)\|_{sp}^{-1}$ and $\max_{\theta\in\Theta} \left\|\overline{W}_{T,t}(\theta)\right\|^{-1}$ are $O_p(1)$), we have

$$\max_{\theta\in\Theta} \left\|\overline{W}_{T,t}^{-1}(\theta) - v_0^{-1} W_t^{-1}(\theta)\right\| = o_p(1), \tag{A.6}$$

which establish (A.5).

By expanding the first-order condition of $\frac{\partial Q_{t,T}(\hat{\theta}_t)}{\partial\theta}=0$ around θ_t, we have

$$\frac{\partial Q_{t,T}(\theta_t)}{\partial\theta}+\frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'}(\hat{\theta}_t-\theta_t)=0,$$

where $\overline{\theta}_t$ lies between $\hat{\theta}_t$ and θ_t. By rearranging terms, we have

$$\begin{aligned}\hat{\theta}_t-\theta_t&=-\Big(\frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'}\Big)^{-1}\frac{\partial Q_{t,T}(\theta_t)}{\partial\theta}\\&=-\Big(\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'}\Big)^{-1}\frac{\partial Q_{t,T}(\theta_t)}{\partial\theta}+\Big[\Big(\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'}\Big)^{-1}-\Big(\frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'}\Big)^{-1}\Big]\frac{\partial Q_{t,T}(\theta_t)}{\partial\theta}.\end{aligned}\tag{A.7}$$

We need to show that

$$\max_{1\leqslant t\leqslant T}\left\|\left(\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'}\right)^{-1}-\left(\frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'}\right)^{-1}\right\|_{sp}=o_p(1),\tag{A.8}$$

$$\max_{1\leqslant t\leqslant T}\left\|\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'}\right\|_{sp}=O_p(1).\tag{A.9}$$

Then, uniform consistency rate and asymptotic normality are determined by $\frac{\partial Q_{t,T}(\theta_t)}{\partial\theta}$. Thus, we need a detailed expansion for the first and second order derivatives for the criteria function $Q_{t,T}$.

Let us first compute the score:

$$\begin{aligned}\frac{\partial Q_{t,T}(\theta_t)}{\partial\theta}&=2\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\frac{\partial g_j(\theta_t)}{\partial\theta'}\Big]'\overline{W}_{T,t}^{-1}(\theta_t)\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)\Big]\\&\quad+\Big(A_{2,1,t}(\theta_t),\cdots,A_{2,d,t}(\theta_t)\Big)'\\&=A_{1,t}(\theta_t)+A_{2,t}(\theta_t).\end{aligned}$$

The ℓ_1th elements in $A_{2,t}(\theta_t)$ is given by

$$A_{2,\ell_1,t}(\theta_t)=\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)\Big]'\frac{\partial\overline{W}_{T,t}^{-1}(\theta_t)}{\partial\theta_{\ell_1}}\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)\Big],$$

where

$$\frac{\partial\overline{W}_{T,t}^{-1}(\theta_t)}{\partial\theta_{\ell_1}}=-\overline{W}_{T,t}^{-1}(\theta_t)\frac{\partial\overline{W}_{T,t}(\theta_t)}{\partial\theta_{\ell_1}}\overline{W}_{T,t}^{-1}(\theta_t).\tag{A.10}$$

We will show that

$$\max_{1\leqslant t\leqslant T}\left\|A_{1,t}(\theta_t)\right\| = O_p\Big(b + (Tb)^{-1/2}\sqrt{\log T}\Big), \tag{A.11}$$

$$\max_{1\leqslant t\leqslant T}\left|A_{2,\ell_1,t}(\theta_t)\right| = o_p(1), \quad \text{for } \ \ell_1 = 1,2,\cdots,d, \tag{A.12}$$

which implies that the dominating term is $A_{1,t}(\theta_t)$, while $A_{2,t}(\theta_t)$ is smaller order term.

Proof of (A.12). We first establish a bound for (A.10). First, recall that

$$\overline{W}_{T,t}(\theta_t) = \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 g_j(\theta_t) g_j'(\theta_t).$$

Then, using similar steps as in the proof of (A.5) and applying Lemma B1(1)(b), we obtain

$$\max_{1\leqslant t\leqslant T}\left\|\overline{W}_{T,t}^{-1}(\theta_t) - v_0^{-1}W_t^{-1}(\theta_t)\right\| = O_p\Big((Tb)^{-1/2}\sqrt{\log T} + b\Big) = o_p(1). \tag{A.13}$$

Next, we consider

$$\frac{\partial \overline{W}_{T,t}(\theta_t)}{\partial\theta_{\ell_1}} = \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2\Big(g_j(\theta_t)\big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)' + \big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)g_j'(\theta_t)\Big).$$

Write

$$\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 g_j(\theta_t)\left(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\right)' = \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2\left(g_j(\theta_t)\left(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\right)' - E\left(g_j(\theta_t)\left(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\right)'\right)\right) + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 E\left(g_j(\theta_t)\left(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\right)'\right).$$

Observe that any elements in $g_j(\theta_t)\big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)' - E\left(g_j(\theta_t)\left(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\right)'\right)$ satisfy Assumptions B1-B2, by Lemma B1(1b), we obtain

$$\max_{1\leqslant t\leqslant T}\left\|\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2\Big(g_j(\theta_t)\big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)' - E\Big[g_j(\theta_t)\big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)'\Big]\Big)\right\| = O_p((Tb)^{-1/2}\sqrt{\log T}) = o_p(1).$$

Next, notice that

$$\begin{aligned}
\max_{1\leqslant t\leqslant T}\left\|W_{t,d_1}\right\| &= \max_{1\leqslant t\leqslant T}\left\|\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 E\Big[g_j(\theta_t)\big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)'\Big]\Big)\right\| \\
&\leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 \max_{1\leqslant t\leqslant T} E\left\|g_j(\theta_t)\big(\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\big)'\right\| \\
&\leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2\Big\{\max_{1\leqslant t\leqslant T} E\left\|g_j(\theta_t)\right\|\Big\}^{1/2}\Big\{\max_{1\leqslant t\leqslant T} E\left\|\frac{\partial g_j(\theta_t)}{\partial\theta_{\ell_1}}\right\|\Big\}^{1/2} < \infty,
\end{aligned}$$

which follows from Assumption 3.2. This implies that

$$\frac{\partial \overline{W}_{T,t}(\theta_t)}{\partial \theta_{\ell_1}} = W_{t,d_1} + W'_{t,d_1} + o_p(1),$$

which holds uniformly over t. Thus, we have, by continuing from (A.10),

$$\max_{1\leqslant t\leqslant T} \left\| \frac{\partial \overline{W}^{-1}_{T,t}(\theta_t)}{\partial \theta_{\ell_1}} \right\| \leqslant v_0^{-2} \max_{1\leqslant t\leqslant T} \|W_t(\theta_t)\|_{sp}^{-1} \max_{1\leqslant t\leqslant T} \left\| \frac{\partial \overline{W}_{T,t}(\theta_t)}{\partial \theta_{\ell_1}} \right\|_{sp} \max_{1\leqslant t\leqslant T} \|W_t(\theta_t)\|^{-1} + o_p(1) = O_p(1).$$

By Lemma 1(i), we have

$$\max_{1\leqslant t\leqslant T} \left\| \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} g_j(\theta_t) \right\| = O_p\Big(b + (Tb)^{-1/2} \sqrt{\log T}\Big) = o_p(1).$$

This implies (A.12)

$$\max_{1\leqslant t\leqslant T} \left|A_{2,\ell_1,t}(\theta_t)\right| \leqslant \max_{1\leqslant t\leqslant T} \left\| \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} g_j(\theta_t) \right\|_{sp} \max_{1\leqslant t\leqslant T} \left\| \frac{\partial \overline{W}^{-1}_{T,t}(\theta_t)}{\partial \theta_{\ell_1}} \right\| \max_{1\leqslant t\leqslant T} \left\| \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} g_j(\theta_t) \right\| + o_p(1) = o_p(1).$$

Proof of (A.11). Define

$$G_{D,t} = \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} \left(\frac{\partial g_j(\theta_t)}{\partial \theta'} - E\Big[\frac{\partial g_j(\theta_t)}{\partial \theta'}\Big] \right),$$

$$W_{D,t} = \overline{W}^{-1}_{T,t}(\theta_t) - v_0^{-1} W_t^{-1}(\theta_t).$$

We have shown in (A.13) that $\max_{1\leqslant t\leqslant T} \left\|W_{D,t}\right\| = o_p(1)$. Similarly, observe that any (a,b)th elements in $\frac{\partial g_j(\theta_t)}{\partial \theta'}$ also satisfy Assumptions B1-B2, by applying Lemma B1(1b), we obtain

$$\max_{1\leqslant t\leqslant T} \left\|G_{D,t}\right\| = O_p((Tb)^{-1/2} \sqrt{\log T}) = o_p(1).$$

Following similar steps in the proof of (A.2), we could show that

$$\frac{1}{Tb} \sum_{j=1}^{T} k_{jt} E\Big[\frac{\partial g_j(\theta_t)}{\partial \theta'}\Big] = E\Big[\frac{\partial g_t(\theta_t)}{\partial \theta'}\Big] + o(1).$$

Define $G_t = E\left[\frac{\partial g_t(\theta_t)}{\partial \theta'}\right]$. Let us rewrite $A_{1,t}(\theta_t)$:

$$A_{1,t}(\theta_t) = v_0^{-1} G_t' W_t^{-1} \Big(\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\Big) + G_{D,t}' W_{D,t} \Big(\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\Big)$$
$$+ v_0^{-1} G_{D,t}' W_t^{-1} \Big(\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\Big) + G_t' W_{D,t} \Big(\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\Big).$$

Clearly, the dominating term is the first one. Then, we have

$$\max_{1\leqslant t\leqslant T} \left\|A_{1,t}(\theta_t)\right\| \leqslant \Big(\max_{1\leqslant t\leqslant T} \|G_t\|_{sp}\Big)\, v_0^{-1} \max_{1\leqslant t\leqslant T} \|W_t(\theta_t)\|_{sp}^{-1} \left(\max_{1\leqslant t\leqslant T} \left\|\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\right\|\right) + o_p(1) = O_p\Big(b + (Tb)^{-1/2}\sqrt{\log T}\Big).$$

Consider now the second order derivatives of the criteria function:

$$\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'} = \begin{bmatrix} \frac{\partial A_{1,t}(\theta_t)}{\partial\theta_1} & \cdots & \frac{\partial A_{1,t}(\theta_t)}{\partial\theta_d} \end{bmatrix}_{d\times d} + \begin{bmatrix} \frac{\partial A_{2,1,t}(\theta_t)}{\partial\theta'} \\ \vdots \\ \frac{\partial A_{2,d,t}(\theta_t)}{\partial\theta'} \end{bmatrix}_{d\times d}.$$

We will show that

$$\max_{1\leqslant t\leqslant T} \left\|\frac{\partial A_1(\theta_t)}{\partial\theta_{\ell_2}}\right\| = O_p(1), \quad \ell_2 = 1, \cdots, d, \tag{A.14}$$
$$\max_{1\leqslant t\leqslant T} \left\|\frac{\partial A_{2,\ell_2}(\theta_t)}{\partial\theta'}\right\| = o_p(1), \quad \ell_2 = 1, \cdots, d. \tag{A.15}$$

Proof of (A.14). Consider

$$\frac{\partial A_{1,t}(\theta_t)}{\partial\theta_{\ell_2}} = \left[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right]' \overline{W}_{T,t}^{-1}(\theta_t)\left[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\right]$$
$$+ \left[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right]' \frac{\partial \overline{W}_{T,t}^{-1}(\theta_t)}{\partial\theta_{\ell_2}}\left[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt} g_j(\theta_t)\right]$$
$$+ \left[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right]' \overline{W}_{T,t}^{-1}(\theta_t)\left[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right]$$
$$= B_{11,t}(\theta_t) + B_{12,t}(\theta_t) + B_{13,t}(\theta_t).$$

We need to find bounds for the above three terms. First, we write

$$\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'} = \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left(\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'} - E\left(\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right)\right) + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt} E\left(\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right).$$

Following again similar steps as in the proof of either Lemma 1(i), we have, $\forall \ell_2$,

$$\max_{1\leqslant t\leqslant T}\left\|\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\left(\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}-E\left(\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right)\right)\right\| = O_p\left((Tb)^{-1/2}\sqrt{\log T}\right) = o_p(1),$$

and

$$\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}E\left(\frac{\partial^2 g_j(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right) = E\left(\frac{\partial^2 g_t(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right) + o(1).$$

Finally, observe that both $B_{11,t}(\theta_t)$ and $B_{12,t}(\theta_t)$ involve $\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)$, following the arguments used to establish (A.11) and (A.12), it is straightforward to verify that

$$\max_{1\leqslant t\leqslant T}\left\|B_{11,t}(\theta_t)\right\| = o_p(1), \quad \max_{1\leqslant t\leqslant T}\left\|B_{12,t}(\theta_t)\right\| = o_p(1).$$

Clearly, the dominating term is $B_{13,t}(\theta_t)$: $\forall \ell_2$,

$$\max_{1\leqslant t\leqslant T}\left\|B_{13,t}(\theta_t)\right\| \leqslant \max_{1\leqslant t\leqslant T}\left\|E\left(\frac{\partial^2 g_t(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right)\right\|_{sp} v_0^{-1} \max_{1\leqslant t\leqslant T}\|W_t(\theta_t)\|_{sp}^{-1} \max_{1\leqslant t\leqslant T}\left\|E\left(\frac{\partial^2 g_t(\theta_t)}{\partial\theta_{\ell_2}\partial\theta'}\right)\right\| + o_p(1) = O_p(1).$$

Summing up, we get: $\forall \ell_2$,

$$\max_{1\leqslant t\leqslant T}\left\|\frac{\partial A_{1,t}(\theta_t)}{\partial\theta_{\ell_2}}\right\| = O_p(1).$$

Proof of (A.15). Consider

$$\frac{\partial A_{2,\ell_2,t}(\theta_t)}{\partial\theta} = 2\left[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)\right]'\frac{\partial\overline{W}_{T,t}^{-1}(\theta_t)}{\partial\theta_{\ell_2}}\left[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\frac{\partial g_j(\theta_t)}{\partial\theta'}\right] + \left[A_{2,1,1,t}(\theta_t)\cdots A_{2,d,1,t}(\theta_t)\right]_{1\times d},$$

where a typical element $A_{2,\ell_4,1,t}(\theta_t)$, $\ell_4 = 1, 2, \cdots, d$ is given by

$$A_{2,\ell_4,1,t}(\theta_t) = \left[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)\right]'\frac{\partial^2\overline{W}_{T,t}(\theta_t)}{\partial\theta_{\ell_1}\partial\theta_{\ell_4}}\left[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)\right].$$

Since both elements above involves $\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}g_j(\theta_t)$, similar arguments as above leads to (A.15), which concludes the claim. Again, by triangular inequality, we establish (A.9):

$$\max_{1\leqslant t\leqslant T}\left\|\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'}\right\|_{sp} \leqslant \max_{1\leqslant t\leqslant T}\left\|B_{1,t}(\theta_t)\right\|_{sp} + \max_{1\leqslant t\leqslant T}\left\|B_{2,t}(\theta_t)\right\|_{sp} = O_p(1).$$

We now move to (A.8):

$$\left\| \Big(\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'}\Big)^{-1} - \Big(\frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'}\Big)^{-1} \right\|_{sp} \leqslant \left\| \frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'} \right\|_{sp}^{-1} \left\| \frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'} - \frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'} \right\|_{sp} \left\| \frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'} \right\|^{-1}.$$

We need to show:

$$\max_{1\leqslant t\leqslant T} \left\| \frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'} - \frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'} \right\|_{sp} = o_p(1),$$
$$\max_{1\leqslant t\leqslant T} \left\| \frac{\partial^2 Q_{t,T}(\overline{\theta}_t)}{\partial\theta\partial\theta'} \right\|^{-1} = O_p(1).$$

These bounds follow immediately by letting $\overline{\theta}_t \xrightarrow{p} \theta_t$ uniformly over t (by the uniform consistency of $\hat{\theta}_t$) and (A.9).

<u>Uniform consistency rate</u>. By continuing from (A.7), we obtain the consistency results:

$$\max_{1\leqslant t\leqslant T} \left\|\hat{\theta}_t - \theta_t\right\| \leqslant \max_{1\leqslant t\leqslant T} \left\| \frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\partial\theta'} \right\|_{sp}^{-1} \max_{1\leqslant t\leqslant T} \left\| \frac{\partial Q_{t,T}(\theta_t)}{\partial\theta} \right\| + o_p(1)$$
$$= O_p\Big(b + (Tb)^{-1/2}\sqrt{\log T}\Big).$$

<u>CLT</u>. Based on the above analysis, we can rewrite the estimator as

$$\sqrt{Tb}\,\left(\hat{\theta}_t - \theta_t\right) = -\left(\frac{\partial^2 Q_{t,T}(\theta_t)}{\partial\theta\theta'}\right)^{-1} \frac{\partial Q_{t,T}(\theta_t)}{\partial\theta} + o_p(1)$$
$$= -(G_t'\,(v_0 W_t)^{-1}\,G_t)^{-1} G_t'\,(v_0 W_t)^{-1}\,\frac{1}{\sqrt{Tb}}\sum_{j=1}^{T} k_{jt} g_j(\theta_t) + o_p(1).$$

Then, we have

$$\sqrt{Tb}\left(\hat{\theta}_t - \theta_t + (G_t'\,(v_0 W_t)^{-1}\,G_t)^{-1} G_t'\,(v_0 W_t)^{-1}\,\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left(g_j(\theta_j) - g_j(\theta_t)\right)\right)$$
$$= -(G_t'\,(v_0 W_t)^{-1}\,G_t)^{-1} G_t'\,(v_0 W_t)^{-1}\,\frac{1}{\sqrt{Tb}}\sum_{j=1}^{T} k_{jt} g_j(\theta_j) + o_p(1).$$

By Lemma 1(ii), together with Slutsky's theorem, we obtain

$$\sqrt{Tb}\,\left(\hat{\theta}_t - \theta_t - b\mu_1\theta_t^{(1)}\right) \xrightarrow{d} \mathcal{N}\left(0, v_0\left(G_t' W_t^{-1} G_t\right)^{-1}\right),$$

where $\mu_1 = \int uK(u)du$ and $v_0 = \int K^2(u)du$. $\theta_t^{(1)}$ is the first order derivative of θ_t. G_t and W_t are given by

$$G_t = E\left[\frac{\partial g_t(\theta_t)}{\partial \theta'}\right], \quad W_t = \mathrm{Var}\left(g_t(\theta_t)\right).$$

This completes the proof.

A.3 Proof of Corollary 1

By triangular inequality,

$$\begin{aligned}\left\|\hat{G}_{T,t} - G_t\right\| &\leqslant \left\|\hat{G}_{T,t} - \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial g_j(\theta_t)}{\partial \theta'}\right\| + \left\|\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\frac{\partial g_j(\theta_t)}{\partial \theta'} - G_t\right\| \\ &:= G_{T,t,1} + G_{T,t,2}.\end{aligned}$$

In the previous section, we have shown that $\left\|G_{T,t,2}\right\| = o_p(1)$. For $G_{T,t,1}$, notice that, by mean-value theorem,

$$\left\|G_{T,t,1}\right\| \leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left\|\frac{\partial g_j(\hat{\theta}_t)}{\partial \theta'} - \frac{\partial g_j(\theta_t)}{\partial \theta'}\right\| \leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left\|\frac{\partial^2 g_j(\bar{\theta}_t)}{\partial \theta_\ell \partial \theta'}\right\|\left\|\hat{\theta}_t - \theta_t\right\|,$$

which holds for all $\ell = 1, 2, \cdots, d$. Since $\max_{\theta\in\Theta}\ \max_{1\leqslant t\leqslant T}\left\|\frac{\partial^2 g_t(\theta)}{\partial \theta_\ell \partial \theta'}\right\|_{sp} < \infty$ and by uniform consistency we have that $\max_{1\leqslant t\leqslant T}\left\|\hat{\theta}_t - \theta_t\right\| = o_p(1)$. This completes the proof.

For $\hat{W}_{T,t}$, by triangular inequality,

$$\begin{aligned}\left\|\hat{W}_{T,t} - v_0 W_t\right\| &\leqslant \left\|\hat{W}_{T,t} - \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 g_j(\theta_t) g_j'(\theta_t)\right\| + \left\|\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}^2 g_j(\theta_t) g_j'(\theta_t) - v_0 W_t\right\| \\ &= W_{T,t,1} + W_{T,t,2}.\end{aligned}$$

In the previous section, we have shown that $\left\|W_{T,t,2}\right\| = o_p(1)$. Following similar analysis as for $G_{T,t,1}$, we could show that $\left\|W_{T,t,1}\right\| = o_p(1)$. This completes the proof.

A.4 Proof of Corollary 2

Consider the following decomposition of $V_{T,t}$:

$$
\begin{aligned}
V_{T,t} &= \left(\hat{W}_{T,t}^{-1/2} - v_0^{-1/2} W_t^{-1/2} + v_0^{-1/2} W_t^{-1/2}\right) \frac{1}{\sqrt{Tb}} \sum_{j=1}^{T} k_{jt} g_j(\hat{\theta}_t) \\
&= v_0^{-1/2} W_t^{-1/2} \frac{1}{\sqrt{Tb}} \sum_{j=1}^{T} k_{jt} g_j(\theta_t) + W_t^{-1/2} G_t \sqrt{Tb} \Big(-(G_t' W_t^{-1} G_t)^{-1} G_t' W_t^{-1} \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} g_j(\theta_t) \Big) + o_p(1) \\
&= v_0^{-1/2} W_t^{-1/2} \Big(I_m - G_t (G_t' W_t^{-1} G_t)^{-1} G_t' W_t^{-1} \Big) \frac{1}{\sqrt{Tb}} \sum_{j=1}^{T} k_{jt} g_j(\theta_t) + o_p(1) \\
&= v_0^{-1/2} W_t^{-1/2} \Big(I_m - G_t (G_t' W_t^{-1} G_t)^{-1} G_t' W_t^{-1} \Big) \frac{1}{\sqrt{Tb}} \sum_{j=1}^{T} k_{jt} \Big(g_j(\theta_j) + \frac{\partial g_j(\bar{\theta}_t)}{\partial \theta'} (\theta_t - \theta_j) \Big) + o_p(1) \\
&= v_0^{-1/2} W_t^{-1/2} \Big(I_m - G_t (G_t' W_t^{-1} G_t)^{-1} G_t' W_t^{-1} \Big) \frac{1}{\sqrt{Tb}} \sum_{j=1}^{T} k_{jt} g_j(\theta_j) + o_p(1),
\end{aligned}
$$

where $v_0 = \int K^2(u)du$. The second equality follows first from the fact that $\hat{W}_{T,t}^{-1/2}$ is a consistent estimator of $W_t^{-1/2}$ and the expansion of each $g_j(\hat{\theta}_t)$ around true $g_j(\theta_t)$. The fourth equality follows from the expansion of each $g_j(\theta_t)$ around $g_j(\theta_j)$. The last equality follows by the assumption $T^{1/2} b^{3/2} \to 0$ so that the smoothing bias vanishes asymptotically.

Recall that $v_0^{-1/2} W_t^{-1/2} \frac{1}{\sqrt{Tb}} \sum_{j=1}^{T} k_{jt} g_j(\theta_j)$ converges to the standard normal distribution and the fact that $I_m - G_t (G_t' W_t^{-1} G_t)^{-1} G_t' W_t^{-1}$ is idempotent of rank $m - d$. Then, the results follow immediately from Rao et al. (1973)(p.186).

B Auxiliary results

Definition B1. *The random function $f(x, \theta) : \mathbb{R} \times \Theta \to \mathbb{R}$ satisfies the standard measurability and differentiability conditions on $\mathbb{R} \times \Theta \to \mathbb{R}$ if*

(1) for each $\theta \in \Theta$, $f(\cdot, \theta)$ is measurable;

(2) for each $x \in \mathbb{R}$, $f(x, \cdot)$ is twice continuously differentiable on Θ.

We shall obtain the uniform bounds for sums

$$
S_{T,t}(\theta) := \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} (f_j(\theta) - E f_j(\theta)), \tag{B.1}
$$

$$
\Delta_{T,t}(\theta) := \frac{1}{Tb} \sum_{j=1}^{T} k_{jt} (E f_j^r(\theta) - E f_t^r(\theta)), \tag{B.2}
$$

for $r = 1, 2$.

Assumption B1. *(i)* Θ *is compact;*

(ii) The stochastic process x_t *is an* α*-mixing (but not necessarily stationary) sequence with the mixing coefficients* $\alpha(j)$ *satisfying* $\alpha(j) \leqslant c\phi^j$ *with* $0 < \phi < 1$ *and* $c > 0$*;*

(iii) $f(x_t, \theta) = f_t(\theta)$ *satisfies the standard measurability and differentiability conditions as in Definition B1 and*

$$\max_{\theta\in\Theta} \max_{1\leqslant t\leqslant T} |f_t(\theta)|_p < \infty, \quad \max_{\theta\in\Theta} \max_{1\leqslant t\leqslant T} \left|\frac{\partial f_t(\theta)}{\partial \theta'}\right|_p < \infty,$$

for some $p > 2$*;*

(iv) For any $\theta \in \Theta$, $E(f_t(\theta))^r = \mu^r(t/T)$ *satisfies the following*

$$|\mu^r(j/T) - \mu^r(t/T)| \leqslant C\left(\frac{|j-t|}{T}\right), \quad j, t = 1, 2, \cdots, T,$$

for $r = 1, 2$ *and the positive constant* C *does not depend on* j, t *and* T*.*

Assumption B2. *The weights* k_{jt} *are computed with a kernel function*

$$k_{jt} = K(\frac{j-t}{Tb}),$$

where $b \to 0$, $Tb \to \infty$. $K(u)$, $u \in \mathbb{R}$, *is a non-negative continuous function satisfying*

$$K(u) \leqslant C(1+u^\nu)^{-1}, \ |(d/du)K(u)| \leqslant C(1+u^\nu)^{-1},$$

for some $C > 0$ *and* $\nu > 3$.

Lemma B1. *Under Assumptions B1-B2, we have*

(1) (a) For any sequence $1 \leqslant t = t_T \leqslant T$*, as* $b \to 0$, $Tb \to \infty$*,*

$$\max_{\theta\in\Theta} \left|S_{T,t}(\theta)\right| = O_p((Tb)^{-1/2});$$

(b) If $c_1 T^{\frac{2}{p}+\delta-1} \leqslant b \leqslant c_2 T^{-\delta}$ *for some* $\delta > 0$, $c_1, c_2 > 0$, $p > 2$ *as in Assumption B1(iii), then for any* $\varepsilon > 0$, $p > 2$*,*

$$\max_{1\leqslant t\leqslant T} \left|S_{T,t}(\theta)\right| = O_p((Tb)^{-1/2}\log^{1/2} T + (T^2 b)^{1/p}(Tb)^{\varepsilon-1}).$$

(2) (a) For any sequence $1 \leqslant t = t_T \leqslant T$*, as* $T \to \infty$*,*

$$\max_{\theta\in\Theta} \left|\Delta_{T,t}(\theta)\right| = O_p(b);$$

(b) If $c_1 T^{\frac{2}{p}+\delta-1} \leqslant b \leqslant c_2 T^{-\delta}$ *for some* $\delta > 0$, $c_1, c_2 > 0$, $p > 2$ *as in Assumption B1(iii), then for any* $\varepsilon > 0$, $p > 2$,

$$\max_{1\leqslant t\leqslant T} \left|\Delta_{T,t}(\theta)\right| = O_p(b).$$

Proof. (1) (b) is (51) in Dendramis et al. (2021). For a given θ, (a) is (48) in Dendramis et al. (2021)[2]. In the next step, we show that, results in (48) from Dendramis et al. (2021) hold uniformly over θ. We follow the steps in Wooldridge (1994). Let $\delta > 0$. Since Θ is compact, there exists a finite covering of Θ, $\Theta \subset \cup_{j=1}^{K} \Theta_j$, where $\Theta_j = \Theta_\delta(\theta_j)$ is the sphere of radius δ about θ_j and $K \equiv K(\delta)$. It follows that, for each $\varepsilon > 0$,

$$\begin{aligned} P\Big[\max_{\theta\in\Theta} \left|S_{T,t}(\theta)\right| > (Tb)^{-1/2}\varepsilon\Big] &\leqslant P\Big[\max_{1\leqslant j\leqslant K}\max_{\theta\in\Theta_j} \left|S_{T,t}(\theta)\right| > (Tb)^{-1/2}\varepsilon\Big] \\ &\leqslant \sum_{j=1}^{K} P\Big[\max_{\theta\in\Theta_j} \left|S_{T,t}(\theta)\right| > (Tb)^{-1/2}\varepsilon\Big]. \end{aligned}$$

We will bound each probability in the above summand. For $\theta \in \Theta_j$, by triangular inequality,

$$\begin{aligned} \left|S_{T,t}(\theta)\right| &= \left|\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\big(f_j(\theta) - f_j(\theta_j) + f_j(\theta_j) - Ef_j(\theta_j) + Ef_j(\theta_j) - Ef_j(\theta)\big)\right| \\ &\leqslant \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left|f_j(\theta) - f_j(\theta_j)\right| + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left|f_j(\theta_j) - Ef_j(\theta_j)\right| + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left|Ef_j(\theta_j) - Ef_j(\theta)\right|. \end{aligned}$$

Observe that $f_t(\cdot)$ is differentiable, by mean-value theorem, we have

$$\left|f_j(\theta) - f_j(\theta_j)\right| \leqslant c_j\left|\theta - \theta_j\right| \leqslant \delta c_j, \quad \left|Ef_j(\theta_j) - Ef_j(\theta)\right| \leqslant \overline{c}_j\left|\theta - \theta_j\right| \leqslant \delta\overline{c}_j,$$

where

$$c_j = \frac{\partial f_j(\theta^*)}{\partial\theta'}, \quad \overline{c}_j = E\Big[\frac{\partial f_j(\theta^{**})}{\partial\theta'}\Big],$$

for some θ^*, θ^{**} lie between θ and θ_j. Thus, we have

$$\begin{aligned} \max_{\theta\in\Theta_j} \left|S_{T,t}(\theta)\right| &\leqslant \delta\Big[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}(c_j - \overline{c}_j)\Big] + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left|f_j(\theta_j) - Ef_j(\theta_j)\right| + 2\delta\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}c_j \\ &\leqslant \delta\Big[\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}(c_j - \overline{c}_j)\Big] + \frac{1}{Tb}\sum_{j=1}^{T} k_{jt}\left|f_j(\theta_j) - Ef_j(\theta_j)\right| + 2\delta\overline{C}, \end{aligned}$$

[2] The results presented in Dendramis et al. (2021) are expressed in terms of $H = Tb$.

where $\frac{1}{Tb}\sum_{j=1}^{T} k_{jt}c_j \leqslant \overline{C}$, which is implies by Assumption B1(iii). It follows that

$$P\Big[\max_{\theta\in\Theta_j}\big|S_{T,t}(\theta)\big| > (Tb)^{-1/2}\varepsilon\Big] \leqslant P\Bigg[\delta\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}(c_j-\overline{c}_j)\Big] + \frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\big|f_j(\theta_j)-Ef_j(\theta_j)\big| > (Tb)^{-1/2}\varepsilon - 2\delta\overline{C}\Bigg]$$

$$\leqslant P\Bigg[\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}(c_j-\overline{c}_j)\Big] + \frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\big|f_j(\theta_j)-Ef_j(\theta_j)\big| > (Tb)^{-1/2}\frac{\varepsilon}{2}\Bigg],$$

where the second inequality follows by letting $\delta \leqslant 1$ such that $(Tb)^{-1/2}\varepsilon - 2\delta\overline{C} < (Tb)^{-1/2}\frac{\varepsilon}{2}$. Letting $\theta^* = \theta^{**}$, by applying (48) in Dendramis et al. (2021), we have

$$\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}(c_j-\overline{c}_j) = O_p((Tb)^{-1/2}), \quad \frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\big|f_j(\theta_j)-Ef_j(\theta_j)\big| = O_p((Tb)^{-1/2}).$$

Then, since $K = K(\delta)$ is finite, we can choose T_0, such that

$$P\Bigg[\Big[\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}(c_j-\overline{c}_j)\Big] + \frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\big|f_j(\theta_j)-Ef_j(\theta_j)\big| > (Tb)^{-1/2}\frac{\varepsilon}{2}\Bigg] \leqslant \frac{\varepsilon}{K}$$

holds for all $T \geqslant T_0$. Then

$$P\Big[\max_{\theta\in\Theta}\big|S_{T,t}(\theta)\big| > (Tb)^{-1/2}\varepsilon\Big] \leqslant \varepsilon,$$

which establishes the results.

(2) Notice that, when t is at the interior point,

$$\big|\Delta_{T,t}(\theta)\big| \leqslant C\frac{1}{Tb}\sum_{j=1}^{T}k_{jt}\left(\frac{|j-t|}{T}\right) \approx Cb\int uK(u)du = O(b),$$

where the approximation follows from Riemann sum approximation of an integral. The results hold for all t. The proof of (2)(a) follows similar as in (1) by utilizing the compactness of Θ, so we omit. The case when t is at the boundary point is also similar.

□

References

Davidson, J., 1994. Stochastic limit theory: An introduction for econometricians. OUP Oxford.

Dendramis, Y., Giraitis, L., Kapetanios, G., 2021. Estimation of time-varying covariance matrices for large datasets. Econometric Theory 37, 1100–1134.

Hall, P., Heyde, C.C., 1980. Martingale limit theory and its application. Academic press.

Newey, W.K., McFadden, D., 1994. Large sample estimation and hypothesis testing. Handbook of econometrics 4, 2111–2245.

Rao, C.R., Rao, C.R., Statistiker, M., Rao, C.R., Rao, C.R., 1973. Linear statistical inference and its applications. volume 2. Wiley New York.

Wooldridge, J.M., 1994. Estimation and inference for dependent processes. Handbook of econometrics 4, 2639–2738.