Semiparametric Longitudinal Model with Irregular Time Autoregressive Error Process

Yang Bai[1,2], Jian Huang[3], Rui Li[1] and Jinhong You[1,2]

[1] *School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, P.R. China*
[2] *Key Laboratory of Mathematical Economics (SUFE), Ministry of Education of China*
[3] *Department of Statistics and Actuarial Science, and Department of Biostatistics, University of Iowa, Iowa City, Iowa 52242, USA*

Supplementary Material

The supplementary material here includes all the detailed proofs of Theorems 1-3 in the paper entitled "Semiparametric Longitudinal Model with Irregular Time Autoregressive Error Process" (SS-13-073), published in *Statistica Sinica*.

S1 Assumptions and Lemmas

We require the following regularity conditions for proving Theorems 1-3.

(A1) The observation times, $t_{i,j}$, are i.i.d from an unknown density function, $f(t)$, which is defined on the support $[0, T]$ and is uniformly bounded away from infinity and 0.

(A2) The functions g and $\eta_k, 1 \leq k \leq p$, have continuous second derivatives on $[0, T]$.

(A3) The numbers of measurements $m_i, 1 \leq i \leq n$ are uniformly bounded by a finite constant independent of n. for all n.

(A4) For every $1 \leq i \leq n$, $(\delta_{i,1}, \ldots, \delta_{i,m_i})$ are independent of $(e_{i,1}, \ldots, e_{i,m_i})$. In addition, $\max_{1\leq i\leq n} \sum_{j=1}^{m_i} E\|\delta_{i,j}\|^2 < \infty$.

(A5) The bandwidth h_N satisfies

$$Nh_N^8/(\log\log N)^{1/2} \to 0 \ \text{ and } Nh_N^2/(\log N)^2 \to \infty \text{ as } n \to \infty.$$

(A6) The bandwidth $h_N^* = cN^{-1/5}$ for some constant c and $h/h^* = o(1)$.

The conditions are reasonably mild. Condition (A1) is a standard assumption for nonparametric or semiparametric regression modeling, see, for example, Wang, Li and Huang (2008). The smoothness condition on $g(t)$ and $\eta_k(t)$ as given in (A2) determines the rate of convergence of the profile semiparametric least squares estimator of the parametric part and local polynomial estimator of the nonparametric component. Under (A3), the total sample size $N = \sum_{i=1}^{n} m_i$ is of the same order as the number of subjects n. It means that we have only local dependency in the sample. Condition (A4) is a technical assumption and is needed to establish the consistency of $\widehat{\beta}_N$. Condition (A5) is a standard regularity condition in nonparametric regressions. Condition (A6) is only needed for Theorem 3.

We first need three lemmas.

Lemma 1. *Suppose that assumptions (A1)-(A5) hold. Then,*

$$\sup_{t\in\mathcal{T}}\left|\frac{1}{Nh_N}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{j=1}^{m_i}K\left(\frac{t_{i,j}-t}{h_N}\right)\left(\frac{t_{i,j}-t}{h_N}\right)^k - f(t)\mu_k\right| = O_p\left\{h_N^2 + \left(\frac{\log N}{Nh_N}\right)^{1/2}\right\}$$

and

$$\sup_{t\in\mathcal{T}}\frac{1}{Nh_N}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{j=1}^{m_i}K\left(\frac{t_{i,j}-t}{h_N}\right)\left(\frac{t_{i,j}-t}{h_N}\right)^k \varepsilon_{i,j} = O_p\left\{\left(\frac{\log N}{Nh_N}\right)^{1/2}\right\}$$

where $k = 0, 1, 2, 4$, h_N *satisfies* $Nh_N^8/(\log\log N)^{1/2} \to 0$ *and* $Nh_N^2/(\log N)^2 \to \infty$ *as* $n \to \infty$.

Proof. Lemma 1 follows immediately from the results of Mack and Silverman (1982).

Lemma 2. *Suppose that assumptions (A1)-(A5) hold. Then,*

$$\max_{1\le i\le n, 1\le j\le m_i}|\tilde{g}(t_{i,j})| = O_p(h_n^{-2}) \quad \text{and} \quad \max_{1\le i\le n, 1\le j\le m_i}|\tilde{\varepsilon}_{i,j}| = O_p(1/\sqrt{nh_n})$$

where $(\tilde{g}(t_{1,1}), \ldots, \tilde{g}(t_{1,m_1}), \tilde{g}(t_{n,m_n}))^\top = (I-S)(g(t_{1,1}), \ldots, g(t_{1,m_1}), g(t_{n,m_n}))^\top$ *and* $(\tilde{\varepsilon}_{1,1}, \ldots, \tilde{\varepsilon}_{1,m_1}, \tilde{\varepsilon}_{n,m_n})^\top = S(\varepsilon_{1,1}, \ldots, \varepsilon_{1,m_1}, \varepsilon_{n,m_n})^\top$.

Proof. By the definition of $\tilde{g}(t_{i,j})$ and $\tilde{\varepsilon}_{i,j}$ we have

$$\tilde{g}(t_{i,j}) = g(t_{i,j}) - \sum_{i_1=1}^{n}\sum_{j_1=1}^{m_i}\omega_{i_1,j_1}(t_{i,j})g(t_{i_1,j_1}) \quad \text{and} \quad \tilde{\varepsilon}_{i,j} = \sum_{i_1=1}^{n}\sum_{j_1=1}^{m_i}\omega_{i_1,j_1}(t_{i,j})\varepsilon_{i_1,j_1}$$

with

$$\omega_{i_1,j_1}(t_{i,j}) = K((t_{i_1,j_1} - t_{i,j})/h_N)\{D_2(t_{i,j}) - (t_{i_1,j_1} - t_{i,j})\}/\{D_2(t_{i,j})D_0(t_{i,j}) - D_1^2(t_{i,j})\}$$

being the local linear weights and $D_s(t) = \sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}(t_{i_1,j_1} - t)^s K((t_{i_1,j_1} - t)/h_N)$. Applying Lemma 1 and the the fact that $\omega_{i,j}(t) = (f(t_{i,j}))^{-1}K((t_{i,j} - t)/h_N)/(Nh_N) + O_p\{(Nh_N)^{-2}\}$ uniformly over $(h_N, 1 - h_N)$, the lemma follows.

Lemma 3. *Suppose the conditions (A1)-(A5) satisfy, we have*

$$\frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\tilde{g}(t_{i,j})\varepsilon_{i,j} = o_p(1/\sqrt{N}) \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\eta_{i,j}\tilde{\varepsilon}_{i,j} = o_p(1/\sqrt{N}),$$

where $\tilde{g}(t_{i,j})$ *and* $\tilde{\varepsilon}_{i,j}$ *are defined in Lemma 2.*

Proof. Lemma 2 together with Lemmas A.5 and A.6 in Liang, Härdle and Carroll (1999) entail Lemma 3.

For simplicity of notation, below denote $d_{i,j,k}(a,b) = a + bd_{i,j,k}$. In particular, we write $d^0_{i,j,k} = d_{i,j,k}(a_{0k}, b_{0k})$. Let (β_0, a_0, b_0) be the true value of (β, a, b), respectively.

S2 Detailed Proofs

Proof of Theorem 1. Let $\theta_0 = (\beta_0^\top, a_0^\top, b_0^\top)^\top$ be the true value of the parameters and write $\theta = (u^\top, v^\top, w^\top)^\top$ with $u = (u_1, \cdots, u_p)^\top$, $v = (v_1, \cdots, v_d)^\top$ and $w = (w_1, \cdots, w_d)^\top$. Let $\widehat{R}_{i,j}(\beta_0) = \widehat{Y}_{i,j} - \widehat{X}_{i,j}^\top \beta_0$. Some calculation shows that

$$Q(\theta_0 + n^{-\frac{1}{2}}\theta) - Q(\theta_0) = J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7 + J_8,$$

where

$$\begin{aligned}
J_1 &= n^{-1} \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} \Big[\Big\{ \widehat{X}_{i,j}^\top u - \sum_{k=1}^{q} d^0_{i,j,k} \widehat{X}_{i,j-k}^\top u \Big\}^2 + \Big\{ \sum_{k=1}^{q} d_{i,j,k}(w_k, v_k) \widehat{R}_{i,j-k}(\beta_0) \Big\}^2 \Big], \\
J_2 &= -2n^{-\frac{1}{2}} u^\top \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} \Big\{ \widehat{R}_{i,j}(\beta_0) - \sum_{k=1}^{q} d^0_{i,j,k} \widehat{R}_{i,j-k}(\beta_0) \Big\} \Big\{ \widehat{X}_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k} \widehat{X}_{i,j-k} \Big\}, \\
J_3 &= -2n^{-\frac{1}{2}} \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} \Big\{ \widehat{R}_{i,j}(\beta_0) - \sum_{k=1}^{q} d^0_{i,j,k} (\widehat{R}_{i,j-k}(\beta_0)) \Big\} \sum_{k=1}^{q} d_{i,j,k}(v_k, w_k) \widehat{R}_{i,j-k}(\beta_0), \\
J_4 &= 2n^{-1} u^\top \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} \Big\{ \widehat{X}_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k} \widehat{X}_{i,j-k}) \Big\} \sum_{k=1}^{q} d_{i,j,k}(v_k, w_k) \widehat{R}_{i,j-k}(\beta_0), \\
J_5 &= 2n^{-1} u^\top \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} \Big[\Big\{ \widehat{R}_{i,j}(\beta_0) - \sum_{k=1}^{q} d^0_{i,j,k} \widehat{R}_{i,j-k}(\beta_0) - n^{-\frac{1}{2}} u^\top \Big(\widehat{X}_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k} \widehat{X}_{i,j-k} \Big) \\
&\quad - n^{-\frac{1}{2}} \sum_{k=1}^{q} d_{i,j,k}(v_k, w_k) \Big\} \widehat{R}_{i,j-k}(\beta_0) \sum_{k=1}^{q} d_{i,j,k}(v_k, w_k) \widehat{X}_{i,j-k} \Big], \\
J_6 &= n^{-2} u^\top \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} \Big\{ \sum_{k=1}^{q} d_{i,j,k}(v_k, w_k) \widehat{X}_{i,j-k} \Big\} \Big\{ \sum_{k=1}^{q} d_{i,j,k}(v_k, w_k) \widehat{X}_{i,j-k} \Big\}^\top u, \\
J_7 &= -2n^{-\frac{1}{2}} u^\top \sum_{i=1}^{n} \sum_{j=1}^{q} \widehat{R}_{i,j}(\beta_0) \widehat{X}_{i,j} \text{ and } J_8 = n^{-1} u^\top \sum_{i=1}^{n} \sum_{j=1}^{q} \widehat{X}_{i,j} \widehat{X}_{i,j}^\top u.
\end{aligned}$$

For J_1, by Lemma 2 we have

$$\begin{aligned}
J_1 &= n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\left\{\widehat{X}_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k}\widehat{X}_{i,j-k}\right\}\left\{\widehat{X}_{i,j} - \sum_{k=1}^{q}(a_{0k}+b_{0k}d_{i,j,k})\widehat{X}_{i,j-k}\right\}^\top u \\
&+ n^{-1}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\left\{\sum_{k=1}^{q} d_{i,j,k}(v_k,w_k)(\tilde{g}(t_{i,j-k})+\varepsilon_{i,j-k}+\tilde{\varepsilon}_{i,j-k})\right\}^2 \\
&= n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\left\{\delta_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k}\delta_{i,j-k}\right\}\left\{\delta_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k}\delta_{i,j-k}\right\}^\top u \\
&+ n^{-1}(v^\top, w^\top)\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}(A_{ij}^\top B_{ij}B_{ij}^\top A_{ij})(v^\top,w^\top)^\top + O_p\left\{h_N^2 + \log N (Nh_N)^{-1/2}\right\} \\
&= J_{1,1} + J_{1,2} + O_p\left\{h_N^2 + \log N (Nh_N)^{-1/2}\right\},\ \text{say,}
\end{aligned}$$

where

$$A_{ij} = (I_d, \mathrm{diag}(d_{i,j,1},\cdots,d_{i,j,q})), B_{ij} = (\varepsilon_{i,j-1},\cdots,\varepsilon_{i,j-q})^\top.$$

It is easy to see that

$$J_{1,1} \to_p u^\top \lim_{n\to\infty}\frac{N}{n}\frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\delta^*_{i,j}\delta^{*\top}_{i,j}u \text{ and } J_{1,2} \to_p \lim_{n\to\infty}\frac{N-dn}{n}(v^\top,w^\top)\Lambda(v^\top,w^\top)^\top.$$

For J_2, based on Lemmas 2 and 3, it can be shown that

$$J_2 = -2n^{-\frac{1}{2}}u^\top \sum_{i=1}^{n}\sum_{j=q+1}^{m_i} e_{i,j}\left\{\delta_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k}\delta_{i,j-k}\right\} + o_p(1).$$

Since $e_{i,j}$ and $\delta_{i,j}$ are not correlated, it follows that $J_2 \to_D u^\top Z_1$ with $Z_1 \sim N(\mathbf{0}, \Delta_1)$, where

$$\frac{1}{n}\sigma_e^2\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\delta^*_{i,j}\delta^{*\top}_{i,j} \to_p \Delta_1$$

For J_3, we have

$$J_3 = -2n^{-\frac{1}{2}}(v^\top,w^\top)\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}(e_{i,j}\zeta_{i,j}) + o_p(1).$$

Therefore, $J_3 \to_D (v^\top, w^\top)Z_2$, where $Z_2 \sim N(0,\Lambda)$. For J_4, we have

$$\begin{aligned}
J_4 &= 2n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\left\{\delta_{i,j} - \sum_{k=1}^{q} d^0_{i,j,k}\delta_{i,j-k})\right\}\sum_{k=1}^{q}(v_k + w_k D_{i,j,k})\varepsilon_{i,j-k} + o_p(1) \\
&= 2u^\top \cdot O_p(N^{-1/2}) = o_p(1).
\end{aligned}$$

Similarly, we have $J_5 = o_p(1)$ and $J_6 = o_p(1)$. For J_7, it holds that

$$J_7 = -2n^{-\frac{1}{2}}u^\top\sum_{i=1}^{n}\sum_{j=1}^{q}(\tilde{g}(t_{i,j})+\varepsilon_{i,j}-\tilde{\varepsilon}_{i,j})(\widehat{\eta}(t_{i,j})+\widehat{\delta}_{i,j}) = -2n^{-\frac{1}{2}}u^\top\sum_{i=1}^{n}\sum_{j=1}^{q}\varepsilon_{i,j}\delta_{i,j} + o_p(1).$$

By the Lindeberg conditions, we have $J_7 \to_D u^\top Z_1$ with $Z_1 \sim N(0, \Delta_2)$, where

$$\frac{1}{n}\sum_{i=1}^{n}(\delta_{i,1}, \ldots, \delta_{i,q})\mathrm{Cov}\{(\varepsilon_{i,1}, \ldots, \varepsilon_{i,q})^\top\}(\delta_{i,1}, \ldots, \delta_{i,q})^\top \to_p \Delta_2,$$

which in combination with the the covariance matrix in J_2 leads to Δ. For J_8, we have

$$\begin{aligned} J_8 &= n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=1}^{q}\widehat{\eta}(t_{i,j})\widehat{\eta}(t_{i,j})^\top u + n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=1}^{q}\widehat{\delta}_{i,j}\widehat{\delta}_{i,j}^\top u + 2n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=1}^{q}\widehat{\eta}(t_{i,j})\widehat{\delta}_{i,j}^\top u \\ &= O_p(h_N^4) + O_p(h_N^2/\sqrt{Nh_N}) + O_p(h_N^2/\sqrt{N}) + n^{-1}u^\top \sum_{i=1}^{n}\sum_{j=q+1}^{m_i}\widehat{\delta}_{i,j-k}\widehat{\delta}_{i,j-k}^\top u \\ &\to_p u^\top \sum_{j=1}^{q} E(\delta_{i,j}\delta_{i,j}^\top)u \ , \end{aligned}$$

which combing the term $J_{1,1}$ leads to the term D.

Thus,

$$V_n(\theta) \equiv Q(\theta_0 + n^{-1/2}\theta) - Q(\theta_0) \to_p \theta^\top \boldsymbol{\Sigma}\theta - 2\theta^\top Z = V(\theta), \text{ say}$$

where $Z \sim N(0, \Pi)$ with $\Pi = \begin{pmatrix} D^{-1}\Delta D^{-1} & \mathbf{0} \\ \mathbf{0} & \Delta^{-1} \end{pmatrix}$. Therefore, by the argmax continuous mapping theorem of Kim and Pollard (1990) (applied to the negative of the criterion here), to prove $\mathrm{argmin}V_n(\theta) \to \mathrm{argmin}V(\theta)$, and therefore parts (i) and (ii) of the theorem, it suffices to show that $\mathrm{argmin}V_n(\theta) = O_p(1)$. This can be shown by a standard argument and is omitted here for brevity. Finally, part (iii) of the theorem follows from the facts that Π is block diagonal and that uncorrelated random vectors with a joint multivariate normal distribution are independent. This completes the proof of Theorem 1.

Proof of Theorem 2. Let $m_i^* = m_i - d$. By Theorem 1, we have

$$\begin{aligned} \widehat{\sigma}_{e,N}^2 &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i^*}\sum_{j=q+1}^{m_i}\left\{\varepsilon_{i,j} - \sum_{k=1}^{q}(\widehat{a}_{k,N} + \widehat{b}_{k,N}d_{i,j,k})\varepsilon_{i,j-k}\right\}^2 + o_p\left(\frac{1}{\sqrt{N}}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i^*}\sum_{j=q+1}^{m_i}\left\{\varepsilon_{i,j} - \sum_{k=1}^{q}(a_{k,N} + b_{k,N}d_{i,j,k})\varepsilon_{i,j-k}\right\}^2 + o_p\left(\frac{1}{\sqrt{N}}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i^*}\sum_{j=q+1}^{m_i} e_{i,j}^2 + o_p\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

Since $e_{i,j}^2$ are i.i.d random variables with mean σ_e^2 and variance $Ee_{i,j}^4 - \sigma_e^4$, the first claim of Theorem 2 follows from the central limit theorem and Slutsky's lemma. The remaining claims of Theorem 2 follow from Theorem 1, the law of large numbers and Slutsky's lemma.

Proof of Theorem 3. Denote $M_t^* = D_t^{*\top} W_t^* D_t^*, h_{i,j}(t) = (1, t_{i,j} - t)^\top$ and $K_{i,j}^*(t) = K((t_{i,j} - t)/h_N^*)/h_N^*$. According to the definition of $(\widehat{g}_N^{TS}(t), \widehat{g}_N^{'TS}(t))$ given in (4.2), it can be shown that

$$(\widehat{g}_N^{TS}(t), \widehat{g}_N^{'TS}(t))^\top - (g(t), g^{'}(t))^\top = J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7,$$

where

$$J_1 = M_t^{*-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) X_{i,j}^\top (\beta - \widehat{\beta}_N)$$

$$J_2 = M_t^{*-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) g(t_{i,j}) - (g(t), g'(t))^\top$$

$$J_3 = M_t^{*-1} \sum_{i=1}^{n} \sum_{j=1}^{q} h_{i,j}(t) K_{i,j}^*(t) \varepsilon_{i,j}$$

$$J_4 = M_t^{*-1} \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) e_{i,j}$$

$$J_5 = -M_t^{*-1} \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) \sum_{k=1}^{q} d_{i,j,k}(\widehat{a}_{k,N}, \widehat{b}_{k,N}) X_{i,j-k}^\top (\beta - \widehat{\beta}_N)$$

$$J_6 = M_t^{*-1} \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) \sum_{k=1}^{q} d_{i,j,k}(\widehat{a}_{k,N}, \widehat{b}_{k,N}) \{\widehat{g}_N(t_{i,j-k}) - g_(t_{i,j-k})\}$$

$$J_7 = -M_t^{*-1} \sum_{i=1}^{n} \sum_{j=q+1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) \sum_{k=1}^{q} \Big\{ ((\widehat{a}_{k,N} - a_k) + (\widehat{b}_{k,N} - b_k) d_{i,j,k}) \varepsilon_{i,j-k} \Big\}.$$

First note that each element of $M_t^* = D_t^{*\top} W_t^* D_t^*$ has the form of a kernel regression, that is,

$$M_t^* = \begin{pmatrix} \sum_{i=1}^{n} \sum_{j=1}^{m_i} K_{h_N^*}(t_{i,j} - t) & \sum_{i=1}^{n} \sum_{j=1}^{m_i} (t_{i,j} - t) K_{h_N^*}(t_{i,j} - t) \\ \sum_{i=1}^{n} \sum_{j=1}^{m_i} (t_{i,j} - t) K_{h_N^*}(t_{i,j} - t) & \sum_{i=1}^{n} \sum_{j=1}^{m_i} (t_{i,j} - t)^2 K_{h_N^*}(t_{i,j} - t) \end{pmatrix}.$$

By Lemma 1,

$$\frac{1}{N} M_t^* = f(t) \otimes \begin{pmatrix} 1 & \mu_1 \\ \mu_1 & \mu_2 \end{pmatrix} \cdot O_p\Big(1 + \{\frac{\log N}{N h_N^*}\}^{1/2}\Big)$$

with probability approaching to 1. Based on the fact $\beta - \widehat{\beta}_N = O_p(N^{-\frac{1}{2}})$, we have $H^* J_1 = O_p(N^{-\frac{1}{2}}) = o_p(h_N^{*2} + 1/\sqrt{N h_N^*})$, where recall that $H^* = \mathrm{diag}(1, h_N^*)$. Since

$$g(t_{i,j}) = g(t) + h_N^* g'(t)\Big(\frac{t_{i,j} - t}{h_N^*}\Big) + \frac{h_N^{*2} g^{''}(t)}{2}\Big(\frac{t_{i,j} - t}{h_N^*}\Big)^2 + o(h_N^{*2}),$$

J_2 can be rewritten as

$$M_t^{*-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} h_{i,j}(t) K_{i,j}^*(t) \Big\{ \frac{h^{*2} g^{''}(t)}{2} \Big(\frac{t_{i,j} - t}{h_N^*}\Big)^2 + o(h_N^{*2}) \Big\}.$$

Therefore,

$$\sqrt{Nh_N^*}\Big[H^*\Big\{J_2 - \begin{pmatrix} g(t) \\ g^{'}(t) \end{pmatrix}\Big\} - \frac{h_N^{*2}}{2}\begin{pmatrix} \kappa_1 g^{''}(t) \\ \kappa_2 g^{''}(t) \end{pmatrix} + o(h_N^{*2})\Big] = o_p(1).$$

We now show that

$$\sqrt{Nh_N^*}H^*J_3 \to_D N(0, \Gamma_1^{TS}), \tag{S2.1}$$

where

$$\Gamma_1^{TS} = \Big\{\lim_{n\to\infty} \frac{\sum_{i=1}^n \sum_{j=1}^q \mathrm{Var}(\varepsilon_{i,j})}{N}\Big\} \frac{1}{f(t)(\mu_2 - \mu_1^2)^2} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}.$$

For any constants d_1 and d_2, let $Z_t = N^{-1}\sum_{i=1}^n \xi_i$, where

$$\xi_i = \sqrt{h_N^*}\sum_{j=1}^q \left\{d_1 + d_2\left(\frac{t_{i,j} - t}{h_N^*}\right)\right\} K_{i,j}^*(t)\varepsilon_{i,j},$$

We have $E(Z_t) = 0$ and some calculation shows

$$\mathrm{Var}(\sqrt{N}Z_t) = d_1^2\mathrm{Var}(\varepsilon_{i,j})f(u)\nu_0 + d_2^2\mathrm{Var}(\varepsilon_{i,j})f(u)\nu_2 + 2d_1d_2\mathrm{Var}(\varepsilon_{i,j})f(u)\nu_1 + o(1),$$

and

$$\sum_{i=1}^n E|\xi_i|^3 \le O(1)\cdot\sum_{i=1}^n h_N^{*3/2}E\left\{|d_1| + |d_2|\cdot\left|\frac{t_{i,j} - t}{h_N^*}\right|\right\}^3 K_{h_N^*}^3(t_{i,j} - t) = O(Nh_N^{*-1/2}).$$

There the Lyapunov condition for the central limit theorem is satisfied. Hence, (S2.1) holds. By the same argument, we can show that

$$\sqrt{Nh_N^*}H^*J_4 \to_D N(0, \Gamma_2^{TS}), \tag{S2.2}$$

where

$$\Gamma_2^{TS} = \Big\{\lim_{n\to\infty} \frac{\sigma_e^2\sum_{i=1}^n (m_i - d)}{N}\Big\} \frac{1}{f(t)(\mu_2 - \mu_1^2)^2} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}.$$

In addition, since $\{e_{i,q+1}, \ldots, e_{i,m_i}\}$ and $\{\varepsilon_{i,1}, \ldots, \varepsilon_{i,q}\}$ are uncorrelated for all $i = 1, \ldots, n$, we have

$$\sqrt{Nh_N^*}H^*(J_3 + J_4) \xrightarrow{D} N(0, \Gamma^{TS}).$$

By the same argument as for J_1, we can show that $H^*J_5 = O_p(N^{-\frac{1}{2}}) = o_p(h_N^{*2} + 1/\sqrt{Nh_N^*})$ and $H^*J_7 = O_p(N^{-\frac{1}{2}}) = o_p(h_N^{*2} + 1/\sqrt{Nh_N^*})$.

Therefore, in order to complete the proof, we just need to show that $H^*J_6 = o_p(h_N^{*2} + 1/\sqrt{Nh_N^*})$. Based on Theorem 1, it can be shown that

$$\begin{aligned} H^*J_6 &= H^*M_t^{*-1}\sum_{i=1}^n \sum_{j=q+1}^{m_i} h_{i,j}(t)K_{i,j}^*(t) \\ &\quad \cdot\sum_{k=1}^q (a_k + b_kd_{i,j,k})\{\widehat{g}_N(t_{i,j-k}) - g_(t_{i,j-k})\} + O_p(N^{-\frac{1}{2}}). \end{aligned}$$

By the standard results from nonparametric regression,

$$\begin{aligned}\widehat{g}_N(t_{i,j})-g(t_{i,j}) &= \frac{\mu_2(f(t_{i,j}))^{-1}}{(\mu_2-\mu_1^2)}\frac{1}{N}\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_i}K_{h_N^*}(t_{i_1,j_1}-t_{i,j})\varepsilon_{i_1,j_1}\\ &\quad -\frac{\mu_1\left(f(t_{i,j})\right)^{-1}}{(\mu_2-\mu_1^2)}\frac{1}{N}\sum_{i_1=1}^{n}\sum_{j=1}^{m_i}\left(\frac{t_{i_1,j_1}-t_{i,j}}{h_N^*}\right)K_{h_N^*}(t_{i_1,j_1}-t_{i,j})\varepsilon_{i_1,j_1}\\ &\quad +\frac{h_N^{*2}}{2}\frac{\mu^2-\mu_1\mu_3}{\mu_2-\mu_1^2}g''(t_{i,j})+o_p(h_N^{*2})=\xi_1(t_{i,j})+\xi_2(t_{i,j})+\xi_3(t_{i,j})+o(h_N^{*2}),\text{say}\end{aligned}$$

Let $w(t)=\mu_2(f(t))^{-1}/(\mu_2-\mu_1^2)$. We have

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}K_{h_N^*}(t_{i,j}-t)\sum_{k=1}^{q}d_{i,j,k}(a_k,b_k)X_{i,j-k}^{\tau}\xi_1(t_{i,j-k})\\ =&\ \frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}K_{h_N^*}(t_{i,j}-t)\sum_{k=1}^{q}d_{i,j,k}(a_k,b_k)w(t_{i,j-k})\frac{1}{N}\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_i}K_{h_N^*}(t_{i_1,j_1}-t_{i,j-k})\varepsilon_{i_1,j_1}\\ =&\ \frac{1}{N}\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_i}\varepsilon_{i_1,j_1}\upsilon_{i_1j_1}\end{aligned}$$

with

$$\upsilon_{i_1,j_1}=\frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}K_{h_N^*}(t_{i,j}-t)\sum_{k=1}^{q}d_{i,j,k}(a_k,b_k)w(t_{i,j-k})K_{h_N^*}(t_{i_1,j_1}-t_{i,j-k}).$$

Obviously, ε_{t_1} and J_{t_1i} are independent. We can show that $\upsilon_{i_1j_1}$ is bounded. Therefore,

$$\frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}K_{h_N^*}(t_{i,j}-t)\sum_{k=1}^{q}d_{i,j,k}(a_k,b_k)\xi_1(t_{i,j-k})=O_p(N^{-\frac{1}{2}}).$$

By the same argument, we can show that

$$\frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}K_{h_N^*}(t_{i,j}-t)\sum_{k=1}^{q}d_{i,j,k}(a_k,b_k)\xi_2(t_{i,j-k})=O_p(N^{-\frac{1}{2}}).$$

Moreover, combining Lemma 1 it is easy to see that

$$\frac{1}{N}\sum_{i=1}^{n}\sum_{j=q+1}^{m_i}K_{h_N^*}(t_{i,j}-t)\sum_{k=1}^{q}d_{i,j,k}(a_k,b_k)(\xi_3(t_{i,j-k})+o_p(h_N^2))=O_p(h_N^2)=o_p(h_N^{*2}).$$

This implies that $H^*J_6=o_p(h_N^{*2}+1/\sqrt{Nh_N^*})$. The proof is complete.