Online Supplementary Document

Bayesian Bandwidth Estimation in Nonparametric Time–Varying Coefficient Models [1]

Tingting Cheng† Jiti Gao* [2] Xibin Zhang*

†School of Finance, Nankai University, Tianjin, P. R. China

*Department of Econometrics and Business Statistics, Monash University, Australia

Appendix C: Proofs of Theorems 1 and 2

Observe that

$$\begin{aligned}\widehat{v}_t = \widehat{u}_t - \alpha \widehat{u}_{t-1} &= y_t - x_t^\top \widehat{\beta}(\tau_t; h) - \alpha \left(y_{t-1} - x_{t-1}^\top \widehat{\beta}(\tau_{t-1}; h) \right) \\ &= u_t + x_t^\top (\beta(\tau_t) - \widehat{\beta}(\tau_t; h)) - \alpha \left(u_{t-1} + x_{t-1}^\top (\beta(\tau_{t-1}) - \widehat{\beta}(\tau_{t-1}; h)) \right) \\ &= u_t - \alpha u_{t-1} + \Delta(\tau_t; h) - \alpha \Delta(\tau_{t-1}; h) \\ &= v_t + \Delta(\tau_t; h) - \alpha \Delta(\tau_{t-1}; h) = v_t + \Gamma(\tau_t; \alpha, h),\end{aligned}$$

where $\Delta(\tau_t; h) = x_t^\top (\beta(\tau_t) - \widehat{\beta}(\tau_t; h))$ and $\Gamma(\tau_t; \alpha, h) = \Delta(\tau_t; h) - \alpha \Delta(\tau_{t-1}; h)$.

Therefore, we can further show

$$l_n(\lambda, \theta) = \log L_n(\lambda, \theta) = \sum_{t=1}^{n} \log f(\widehat{v}_t; \theta)$$

[1] We extend our sincere thanks to the Co-Editor, an Associate Editor and three referees for their insightful comments that helped improve the paper substantially. We thank participants to several seminars and conferences, at which earlier versions of this paper were presented, for their constructive comments and suggestions. Thanks also go to Maxwell King for enlightening discussion. This research is supported under the Australian Research Council's *Discovery Projects Scheme* (under Grant Numbers: DP1095838, DP1096374 and DP130104229).

[2] *Correspondence*: Jiti Gao, Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia. Telephone: +61 3 99031675. Fax: +61 3 99032007. Email: jiti.gao@monash.edu.

$$
\begin{aligned}
&= \sum_{t=1}^{n} \log f(v_t + \Gamma(\tau_t;\alpha,h);\theta) \\
&= \sum_{t=1}^{n} \log\left[\left(f(v_t;\theta) + f^{(1)}(v_t;\theta)\Gamma(\tau_t;\alpha,h)\right)(1+o_P(1))\right] \\
&= \sum_{t=1}^{n} \log\left[f(v_t;\theta)\left(1 + \frac{f^{(1)}(v_t;\theta)}{f(v_t;\theta)}\Gamma(\tau_t;\alpha,h)\right)(1+o_P(1))\right] \\
&= \sum_{t=1}^{n} \log f(v_t;\theta) + \sum_{t=1}^{n} \log\left(1+\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\right) + o_P(1) \\
&= \sum_{t=1}^{n} \log f(v_t;\theta) + \sum_{t=1}^{n} \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)(1+o_P(1)), \qquad (1)
\end{aligned}
$$

where $\gamma(v_t;\theta) = \frac{f^{(1)}(v_t;\theta)}{f(v_t;\theta)}$.

Thus, we have

$$
\begin{aligned}
L_n(\lambda,\theta) &= e^{\sum_{t=1}^{n} \log f(v_t;\theta)} e^{\sum_{t=1}^{n} \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)(1+o_P(1))} \qquad (2) \\
&= (1+o_P(1)) \prod_{t=1}^{n} f(v_t;\theta)\left(\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)+1\right) \\
&= (1+o_P(1))\ G_n(\theta)\left(\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)+1\right),
\end{aligned}
$$

where $G_n(\theta) = \prod_{t=1}^{n} f(v_t;\theta)$.

It is easy to see that

$$
L_n(\lambda,\theta) - G_n(\theta) = G_n(\theta)\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h).
$$

As $\Gamma(\tau_t;\alpha,h) = \Delta(\tau_t;h) - \alpha\Delta(\tau_{t-1};h)$, we have

$$
\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h) = \sum_{t=1}^{n}\gamma(v_t;\theta)\Delta(\tau_t;h) - \alpha\sum_{t=1}^{n}\gamma(v_t;\theta)\Delta(\tau_{t-1};h).
$$

We investigate $\Delta(\tau_t;h)$ as follows:

$$
\begin{aligned}
&\Delta(\tau_t;h) = x_t^\top(\beta(\tau_t) - \widehat{\beta}(\tau_t;h)) = x_t^\top\left(\beta(\tau_t) - [p_n(\tau_t;h)]^{-1}q_n(\tau_t;h)\right) \\
&= x_t^\top\left\{[p_n(\tau_t;h)]^{-1}[p_n(\tau_t;h)]\beta(\tau_t) - [p_n(\tau_t;h)]^{-1}\frac{1}{nh}\sum_{s=1}^{n} x_s(x_s^\top\beta(\tau_s)+u_s)K\left(\frac{\tau_s-\tau_t}{h}\right)\right\} \\
&= -x_t^\top[p_n(\tau_t;h)]^{-1}\frac{1}{nh}\sum_{s=1}^{n} x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right) \\
&+ x_t^\top[p_n(\tau_t;h)]^{-1}\frac{1}{nh}\sum_{s=1}^{n} x_s x_s^\top(\beta(\tau_t)-\beta(\tau_s))K\left(\frac{\tau_s-\tau_t}{h}\right) = -\Delta_1(\tau_t;h) + \Delta_2(\tau_t;h),
\end{aligned}
$$

in which $p_n(\tau_t;h) = \frac{1}{nh}\sum_{s=1}^n x_s x_s^\top K\left(\frac{\tau_s-\tau_t}{h}\right)$, $q_n(\tau_t;h) = \frac{1}{nh}\sum_{s=1}^n x_s y_s K\left(\frac{\tau_s-\tau_t}{h}\right)$,

$\Delta_1(\tau_t;h) = x_t^\top [p_n(\tau_t;h)]^{-1} \frac{1}{nh}\sum_{s=1}^n x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right)$, and

$\Delta_2(\tau_t;h) = x_t^\top [p_n(\tau_t;h)]^{-1} \frac{1}{nh}\sum_{s=1}^n x_s x_s^\top (\beta(\tau_t)-\beta(\tau_s)) K\left(\frac{\tau_s-\tau_t}{h}\right)$.

Therefore, we have

$$\sum_{t=1}^n \gamma(v_t;\theta)\Delta(\tau_t;h) = -\sum_{t=1}^n \gamma(v_t;\theta)\Delta_1(\tau_t;h) + \sum_{t=1}^n \gamma(v_t;\theta)\Delta_2(\tau_t;h). \tag{3}$$

Note that $p_n(\tau;h)$ can be expressed as

$$\begin{aligned} p_n(\tau;h) &= \frac{1}{nh}\sum_{s=1}^n x_s x_s^\top K\left(\frac{\tau_s-\tau}{h}\right) = (1+o(1))\mathbb{E}[x_1 x_1^\top]\frac{1}{nh}\sum_{s=1}^n K\left(\frac{\tau_s-\tau}{h}\right) \\ &= (1+o(1))\mathbb{E}[x_1 x_1^\top]\int \frac{1}{h}K\left(\frac{u-\tau}{h}\right)du = (1+o(1))\mathbb{E}[x_1 x_1^\top]\int_{-\frac{\tau}{h}}^{\frac{1-\tau}{h}} k(v)dv. \end{aligned}$$

As we assume that $h = a_n\lambda$ with $a_n \to 0$ as $n\to\infty$ and λ being a continuous random variable, so $h\to 0$ as $n\to\infty$. Therefore, $\int_{-\frac{\tau}{h}}^{\frac{1-\tau}{h}} k(v)dv = 1+o(1)$ and

$$p_n(\tau;h) = (1+o(1))\mathbb{E}[x_1 x_1^\top] = \Sigma_x(1+o(1)).$$

Therefore, it follows that

$$\begin{aligned} \sum_{t=1}^n \gamma(v_t;\theta)\Delta_1(\tau_t;h) &= \sum_{t=1}^n \gamma(v_t;\theta)x_t^\top [p_n(\tau_t;h)]^{-1}\frac{1}{nh}\sum_{s=1}^n x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right) \\ &= \sum_{t=1}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}\frac{1}{nh}\sum_{s=1}^n x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right) \\ &= \frac{1}{nh}\sum_{t=1}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}\sum_{s=1}^n x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right) \\ &= \frac{K(0)}{nh}\sum_{t=1}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}x_t u_t + \frac{1}{nh}\sum_{s=1}^n\sum_{t=1,t\neq s}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right) \\ &= I_{nt}(h) + J_{nt}(h), \end{aligned}$$

where $I_{nt}(h) = \frac{K(0)}{nh}\sum_{t=1}^n \gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_t u_t$ and $J_{nt}(h) = \frac{1}{nh}\sum_{s=1}^n\sum_{t=1,t\neq s}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}x_s u_s K\left(\frac{\tau_s-\tau_t}{h}\right)$.

Since the error term u_t follows an AR(1) process, we can further express $I_{nt}(h)$ and $J_{nt}(h)$ as follows:

$$I_{nt}(h) = \frac{K(0)}{nh}\sum_{t=1}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}x_t u_t = \frac{K(0)}{nh}\sum_{t=1}^n \gamma(v_t;\theta)x_t^\top \Sigma_x^{-1}x_t(\alpha u_{t-1}+v_t)$$

$$= \alpha \frac{K(0)}{nh} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t u_{t-1} + \frac{K(0)}{nh} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t v_t,$$

$$J_{nt}(h) = \frac{1}{nh} \sum_{s=1}^{n} \sum_{t=1, t\neq s}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s u_s K\left(\frac{\tau_s - \tau_t}{h}\right)$$

$$= \frac{1}{nh} \sum_{s=1}^{n} \sum_{t=1, t\neq s}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s \left(\alpha u_{s-1} + v_s\right) K\left(\frac{\tau_s - \tau_t}{h}\right)$$

$$= \alpha \frac{1}{nh} \sum_{s=1}^{n} \sum_{t=1, t\neq s}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s u_{s-1} K\left(\frac{\tau_s - \tau_t}{h}\right)$$

$$+ \frac{1}{nh} \sum_{s=1}^{n} \sum_{t=1, t\neq s}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s v_s K\left(\frac{\tau_s - \tau_t}{h}\right).$$

Similarly, we can get

$$\alpha \sum_{t=1}^{n} \gamma(v_t;\theta) \Delta(\tau_{t-1};h) = -\alpha \sum_{t=1}^{n} \gamma(v_t;\theta) \Delta_1(\tau_{t-1};h) + \alpha \sum_{t=1}^{n} \gamma(v_t;\theta) \Delta_2(\tau_{t-1};h), \tag{4}$$

where $\sum_{t=1}^{n} \gamma(v_t;\theta)\Delta_1(\tau_{t-1};h) = I_{n(t-1)}(h) + J_{n(t-1)}(h)$, $I_{n(t-1)}(h) = \frac{K(0)}{nh} \sum_{t=1}^{n} \gamma(v_t;\theta) x_{t-1}^\top \Sigma_x^{-1} x_{t-1} u_{t-1}$ and $J_{n(t-1)}(h) = \frac{1}{nh} \sum_{s=1}^{n} \sum_{t-1\neq s}^{n} \gamma(v_t;\theta) x_{t-1}^\top \Sigma_x^{-1} x_s u_s K\left(\frac{\tau_s - \tau_{t-1}}{h}\right)$.

Observe that

$$\sum_{t=1}^{n} \gamma(v_t;\theta) \Delta_1(\tau_t;h) - \alpha \sum_{t=1}^{n} \gamma(v_t;\theta) \Delta_1(\tau_{t-1};h)$$
$$= \frac{K(0)}{nh} \sum_{t=1}^{n} v_t \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t + \frac{1}{nh} \sum_{s=1}^{n} \sum_{t=1, t\neq s}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s v_s K\left(\frac{\tau_s - \tau_t}{h}\right) + o_P(1)$$
$$= Q_{1n} + Q_{2n} + o_P(1),$$

where $Q_{1n} = \frac{K(0)}{nh} \sum_{t=1}^{n} v_t \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t$ and $Q_{2n} = \frac{1}{nh} \sum_{s=1}^{n} \sum_{t=1, t\neq s}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s v_s K\left(\frac{\tau_s - \tau_t}{h}\right)$.

As $n \to \infty$, the law of large numbers implies

$$h Q_{1n} = K(0) \frac{1}{n} \sum_{t=1}^{n} v_t \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t \to_P K(0) \cdot \mathbb{E}_\theta[\gamma(v_1;\theta) v_1] \cdot \mathbb{E}[w_1] \neq 0,$$

where $w_t = x_t^\top \Sigma_x^{-1} x_t$. So $Q_{1n} = O_P(1/h)$.

We now show that Q_{2n} is a higher–order term than Q_{1n}. Because $\mathbb{E}_\theta[v_1] = \mathbb{E}_\theta[\gamma(v_1;\theta)] = 0$, we have

$$\mathbb{E}_{\theta,\lambda}\left[Q_{2n}^2\right] = \frac{1}{n^2 h^2} \mathbb{E}_{\theta,\lambda}\left[\sum_{t=1}^{n} \left(\sum_{s=1,\neq t}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_s v_s K\left(\frac{\tau_s - \tau_t}{h}\right)\right)\right]^2$$
$$= \frac{1}{n^2 h^2} \sum_{t=1}^{n} \left(\sum_{s=1,\neq t}^{n} \mathbb{E}_\theta\left[\gamma^2(v_t;\theta)\right] \mathbb{E}_\theta[v_s^2] \cdot \mathbb{E}[w_s w_t] K^2\left(\frac{\tau_s - \tau_t}{h}\right)\right)$$

$$
\begin{aligned}
&\leq \frac{C_{vw}}{n^2h^2}\sum_{t=1}^{n}\sum_{s=1,t\neq s}^{n} K^2\left(\frac{\tau_s-\tau_t}{h}\right) = \frac{C_{vw}}{nh^2}\sum_{t=1}^{n}\frac{1}{n}\sum_{s=1,t\neq s}^{n} K^2\left(\frac{\tau_s-\tau_t}{h}\right)\\
&= \frac{C_{vw}(1+o(1))}{nh^2}\sum_{t=1}^{n}\int_0^1 K^2\left(\frac{u-\tau_t}{h}\right)du = \frac{C_{vw}}{nh}\sum_{t=1}^{n}\int_{-\tau_t/h}^{1-\tau_t/h} K^2(v)dv\\
&= \frac{C_{vw}}{h}\frac{1}{n}\sum_{t=1}^{n} m(\tau_t;h) = \frac{C_{vw}(1+o(1))}{h}\int K^2(u)du,
\end{aligned}
$$

where $C_{vw}>0$ is some constant and $m(\tau;h)=\int_{-\tau/h}^{1-\tau/h}K^2(u)du=(1+o(1))\int K^2(u)du$. So we have $Q_{2n}=O_P(1/\sqrt{h})$. This deduces that $Q_{1n}+Q_{2n}=O_P(1/h)\left(1+O_P(\sqrt{h})\right)$, which implies that Q_{1n} is the leading term.

Meanwhile, we have

$$
\begin{aligned}
&\sum_{t=1}^{n}\gamma(v_t;\theta)\Delta_2(\tau_t;h) = \sum_{t=1}^{n}\gamma(v_t;\theta)x_t^\top[p_n(\tau_t;h)]^{-1}\frac{1}{nh}\sum_{s=1}^{n}x_sx_s^\top(\beta(\tau_t)-\beta(\tau_s))K\left(\frac{\tau_s-\tau_t}{h}\right)\\
&= \frac{1+o(1)}{nh}\sum_{t=1}^{n}\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}\sum_{s=1}^{n}x_sx_s^\top K\left(\frac{\tau_s-\tau_t}{h}\right)(\beta(\tau_t)-\beta(\tau_s))\\
&= \frac{1+o(1)}{nh}\sum_{t=1}^{n}\sum_{s=1,s\neq t}^{n}\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top K\left(\frac{\tau_s-\tau_t}{h}\right)\left(\beta^{(1)}(\tau_t)(\tau_t-\tau_s)+1/2\beta^{(2)}(\tau_t)(\tau_t-\tau_s)^2\right)\\
&= \frac{1+o(1)}{n}\sum_{t=1}^{n}\sum_{s=1,s\neq t}^{n}\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top K_1\left(\frac{\tau_s-\tau_t}{h}\right)\beta^{(1)}(\tau_t)\\
&+ \frac{(1+o(1))h}{2n}\sum_{t=1}^{n}\sum_{s=1,s\neq t}^{n}\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top K_2\left(\frac{\tau_s-\tau_t}{h}\right)\beta^{(2)}(\tau_t)\\
&\equiv (1+o(1))\,Q_{3n}+(1+o(1))\,Q_{4n}.
\end{aligned} \tag{5}
$$

We then show that $\mathbb{E}_{\theta,\lambda}[Q_{jn}]\to 0$ and $\mathbb{E}_{\theta,\lambda}\left(Q_{jn}-\mathbb{E}_{\theta,\lambda}[Q_{jn}]\right)^2\to 0$ for $j=3,4$. Obviously, we have

$$
\begin{aligned}
\mathbb{E}_{\theta,\lambda}[Q_{3n}] &= \frac{1}{n}\sum_{t=1}^{n}\sum_{s=1,s\neq t}^{n}\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]K_1\left(\frac{\tau_s-\tau_t}{h}\right)\beta^{(1)}(\tau_t),\\
\mathbb{E}_{\theta,\lambda}[Q_{4n}] &= \frac{h}{2n}\sum_{t=1}^{n}\sum_{s=1,s\neq t}^{n}\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]K_2\left(\frac{\tau_s-\tau_t}{h}\right)\beta^{(2)}(\tau_t),
\end{aligned}
$$

where $K_1(u)=uK(u)$ and $K_2(u)=u^2K(u)$.

Under Assumption 3(ii), by Lemma A.1 of Gao (2007), we obtain[3]

$$\begin{aligned}
&\left\|\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]-\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}\right]\mathbb{E}\left[x_sx_s^\top\right]\right\|\\
&=\left\|\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]-\mathbb{E}_\theta\left[\gamma(v_t;\theta)\right]\mathbb{E}\left[x_t^\top\Sigma_x^{-1}\right]\mathbb{E}\left[x_sx_s^\top\right]\right\|\\
&=\left\|\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]\right\|\\
&\le\left(\mathbb{E}_\theta\left[\|\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}\|^p\right]\right)^{1/p}\left(\mathbb{E}\left[\|x_sx_s^\top\|^q\right]\right)^{1/q}\rho^{1-Q}(|s-t|),
\end{aligned}$$

where $Q=\frac{1}{p}+\frac{1}{q}<1$. Without loss of generosity, we choose $p=q=4$. Then under Assumptions 1(i) and 3(ii), we have

$$\begin{aligned}
&\left\|\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]\right\|\le\left(\mathbb{E}_\theta\left[\|\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}\|^4\right]\right)^{1/4}\left(\mathbb{E}\left[\|x_sx_s^\top\|^4\right]\right)^{1/4}\rho^{1/2}(|s-t|)\\
&=\left(\mathbb{E}_\theta\left[\gamma^4(v_t;\theta)\right]\right)^{1/4}\left(\mathbb{E}\left[\|x_t^\top\Sigma_x^{-1}\|^4\right]\right)^{1/4}\left(\mathbb{E}\left[\|x_sx_s^\top\|^4\right]\right)^{1/4}\rho^{1/2}(|s-t|)\equiv A_0\,\rho^{1/2}(|s-t|),
\end{aligned}$$

where $A_0=\left(\mathbb{E}_\theta\left[\gamma^4(v_t;\theta)\right]\right)^{1/4}\left(\mathbb{E}\left[\|x_t^\top\Sigma_x^{-1}\|^4\right]\right)^{1/4}\left(\mathbb{E}\left[\|x_sx_s^\top\|^4\right]\right)^{1/4}$.

Therefore, we obtain

$$\begin{aligned}
\left|\sum_{t=1}^n\gamma(v_t;\theta)\Delta_2(\tau_t;h)\right|&\le\frac{M}{n}\sum_{t=1}^n\sum_{s=1,s\neq t}^n\left\|\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]\right\|\left|K_1\left(\frac{\tau_s-\tau_t}{h}\right)\right|\left\|\beta^{(1)}(\tau_t)\right\|\\
&+\frac{Mh}{2n}\sum_{t=1}^n\sum_{s=1,s\neq t}^n\left\|\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]\right\|K_2\left(\frac{\tau_s-\tau_t}{h}\right)\left\|\beta^{(2)}(\tau_t)\right\|\\
&\le\frac{M}{n}\sum_{t=1}^n\sum_{s=1,s\neq t}^n\rho^{1/2}(|s-t|)\left|K_1\left(\frac{\tau_s-\tau_t}{h}\right)\right|\left\|\beta^{(1)}(\tau_t)\right\|\\
&+\frac{Mh}{2n}\sum_{t=1}^n\sum_{s=1,s\neq t}^n\rho^{1/2}(|s-t|)K_2\left(\frac{\tau_s-\tau_t}{h}\right)\left\|\beta^{(2)}(\tau_t)\right\|\\
&=\frac{M}{n}\sum_{t=1}^n\sum_{s=1,s\neq t}^n\rho^{1/2}(n|\tau_s-\tau_t|)\left|K_1\left(\frac{\tau_s-\tau_t}{h}\right)\right|\left\|\beta^{(1)}(\tau_t)\right\|\\
&+\frac{Mh}{2n}\sum_{t=1}^n\sum_{s=1,s\neq t}^n\rho^{1/2}(n|\tau_s-\tau_t|)K_2\left(\frac{\tau_s-\tau_t}{h}\right)\left\|\beta^{(2)}(\tau_t)\right\|\\
&=Mn\,(1+o(1))\int_0^1\int_0^1\rho^{1/2}(n|v-u|)\left|K_1\left(\frac{v-u}{h}\right)\right|\left\|\beta^{(1)}(u)\right\|dudv
\end{aligned}$$

[3]As discussed in the remark about Assumption 1 in Appendix A1, we need only to require $\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t\right]=\mathbb{E}_\theta\left[\gamma(v_t;\theta)\right]\,\mathbb{E}[x_t]=0$ in an application of Lemma A.1 of Gao (2007).

Meanwhile, we will have $\mathbb{E}_\theta\left[\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]=\mathbb{E}_\theta\left[\gamma(v_t;\theta)\right]\mathbb{E}\left[x_t^\top\Sigma_x^{-1}x_sx_s^\top\right]=0$ if we assume that $\{v_t\}$ and $\{x_s\}$ are independent for all (s,t). In this case, there is no need to apply Lemma A.1 of Gao (2007) with the choice of (p,q). As a matter of fact, $\mathbb{E}_{\theta,\lambda}[Q_{jn}]=0$ and the derivation of $\mathbb{E}_{\theta,\lambda}\left[Q_{jn}\right]^2\to 0$ for $j=3,4$ follows similarly from that of $\mathbb{E}_{\theta,\lambda}\left[Q_{2n}\right]^2\to 0$.

$$+\frac{Mnh}{2}(1+o(1))\int_0^1\int_0^1 \rho^{1/2}(n|v-u|)K_2\left(\frac{v-u}{h}\right)\left\|\beta^{(2)}(u)\right\| du dv$$
$$= nhM\cdot A_{1n}\int_0^1 \left\|\beta^{(1)}(u)\right\| du + \frac{Mnh^2}{2}\cdot A_{2n}\int_0^1 \left\|\beta^{(2)}(u)\right\| du,$$

where $A_{1n} = \int_{-\infty}^{+\infty} \rho^{1/2}(nh|w|)\,|K_1(w)|\,dw$, $A_{2n} = \int_{-\infty}^{+\infty} \rho^{1/2}(nh|w|)K_2(w)dw$ and M is a positive constant.

Under Assumption 1(i), we have

$$|A_{1n}| = \left|\int_{-\infty}^{+\infty} \rho^{1/2}(nh|w|)\,|K_1(w)|\,dw\right| \le \int_{-\infty}^{+\infty} \rho^{1/2}(nh|w|)|w|K(w)dw$$
$$\le c(nh)^{-c_0/2}\int_{-\infty}^{+\infty} |w|^{-c_0/2}|w|K(w)dw \le c(nh)^{-c_0/2}\int_{-\infty}^{+\infty} |w|^{1-c_0/2}K(w)dw.$$

Thus, under Assumptions 1(i)(iii) and 2(i), as $n \to \infty$, we have

$$nh|A_{1n}| \le c(nh)^{1-c_0/2}\int_{-\infty}^{+\infty} |w|^{1-c_0/2}K(w)dw = O\left((nh)^{1-c_0/2}\right) \to 0.$$

Similarly, we have

$$|A_{2n}| = \left|\int_{-\infty}^{+\infty} \rho^{1/2}(nh|w|)K_2(w)dw\right| \le \int_{-\infty}^{+\infty} \rho^{1/2}(nh|w|)|w|^2K(w)dw$$
$$\le c(nh)^{-c_0/2}\int_{-\infty}^{+\infty} |w|^{2-c_0/2}K(w)dw.$$

Hence, under Assumptions 1(i)(iii) and 2(i), as $n \to \infty$, we have

$$nh^2|A_{2n}| \le cn^{1-c_0/2}h^{2-c_0/2}\int_{-\infty}^{+\infty} |w|^{2-c_0/2}K(w)dw = O\left(n^{1-c_0/2}h^{2-c_0/2}\right) \to 0.$$

Thus, under Assumptions 1(i)(iii), 2(i) and 3(i)(ii), it follows that

$$\left|\sum_{t=1}^n \gamma(v_t;\theta)\Delta_2(\tau_t;h)\right| \le nhM\int_0^1 \left\|\beta^{(1)}(u)\right\| du\, A_{1n} + \frac{nh^2}{2}M\int_0^1 \left\|\beta^{(2)}(u)\right\| du\, A_{2n} = o_P(1).$$

Similarly, we can show that $\sum_{t=1}^n \gamma(v_t;\theta)\Delta_2(\tau_{t-1};h) = o_P(1)$. Under Assumptions 1(ii) and 1(iii), we therefore have

$$\int \left|\sum_{t=1}^n \gamma(v_t;\theta)\Delta_2(\tau_t;h)\right| \pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$$
$$\le \int \tilde{c}_1(nh)^{1-c_0/2}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta + \int \tilde{c}_2 n^{1-c_0/2}h^{2-c_0/2}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$$
$$= \tilde{c}_1 n^{1-c_0/2}a_n^{1-c_0/2}\int \lambda^{1-c_0/2}\pi_\lambda(\lambda)d\lambda + \tilde{c}_2 n^{1-c_0/2}a_n^{2-c_0/2}\int \lambda^{2-c_0/2}\pi_\lambda(\lambda)d\lambda \to 0,$$

where $|\tilde{c}_1| < \infty$ and $|\tilde{c}_2| < \infty$.

In order to show that $\mathbb{E}_{\theta,\lambda}\left[\left(Q_{jn} - \mathbb{E}_{\theta,\lambda}[Q_{jn}^2]\right)^2\right] \to 0$ for $j = 3,4$, we need only to deal with a leading term of the form

$$\frac{1}{n^2}\mathbb{E}_{\theta,\lambda}\left[\left(\sum_{t=1}^{n}\sum_{s=1,s\neq t}^{n}\left(\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}\right)\left(x_s x_s^\top - \mathbb{E}[x_s x_s^\top]\right)K_1\left(\frac{\tau_s-\tau_t}{h}\right)\beta^{(1)}(\tau_t)\right)^2\right] \to 0,$$

which follows similarly from the derivation of $\mathbb{E}[Q_{2n}^2] \to 0$.

Let $w_t = x_t^\top \Sigma_x^{-1} x_t$. To sum up, we have

$$\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h) = \sum_{t=1}^{n}\gamma(v_t;\theta)\Delta(\tau_t;h) - \alpha\sum_{t=1}^{n}\gamma(v_t;\theta)\Delta(\tau_{t-1};h)$$
$$= -\frac{K(0)}{nh}\sum_{t=1}^{n}v_t\gamma(v_t;\theta)x_t^\top\Sigma_x^{-1}x_t\,(1+o_P(1)) = -\frac{K(0)\cdot(1+o_P(1))}{nh}\sum_{t=1}^{n}v_t\gamma(v_t;\theta)w_t.$$

We are now ready to complete the proofs of Theorems 1 and 2.

Proof of Theorem 1.

Observe that

$$\mathbb{E}_\star[\lambda|X_n,Y_n] = \frac{\iint \lambda e^{l_n(\lambda,\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\iint e^{l_n(\lambda,\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta} = \frac{\iint \lambda L_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\iint L_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta} = \frac{q_1(X_n,Y_n)}{p_1(X_n,Y_n)}, \tag{6}$$

where $q_1(X_n,Y_n) = \iint \lambda L_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$ and $p_1(X_n,Y_n) = \iint L_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$.

$$\mathbb{E}[\lambda] = \int \lambda\pi_\lambda(\lambda)d\lambda = \frac{\int\lambda\pi_\lambda(\lambda)d\lambda}{\int\pi_\lambda(\lambda)d\lambda}\frac{\int G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} = \frac{\iint \lambda G_n(\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\iint G_n(\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta} = \frac{q(X_n,Y_n)}{p(X_n,Y_n)}, \tag{7}$$

where $q(X_n,Y_n) = \iint \lambda G_n(\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$ and $p(X_n,Y_n) = \iint G_n(\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$.

Based on (6) and (7), we have that

$$q_1(X_n,Y_n) - q(X_n,Y_n) = \iint \lambda\left(L_n(\lambda,\theta) - G_n(\theta)\right)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta,$$
$$p_1(X_n,Y_n) - p(X_n,Y_n) = \iint \left(L_n(\lambda,\theta) - G_n(\theta)\right)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta.$$

Therefore, by equation (2) we have

$$\mathbb{E}_\star[\lambda|X_n,Y_n] - \mathbb{E}[\lambda] = \frac{q_1(X_n,Y_n)}{p_1(X_n,Y_n)} - \frac{q(X_n,Y_n)}{p(X_n,Y_n)}$$
$$= \frac{1}{p_1(X_n,Y_n)p(X_n,Y_n)}\left[q_1(X_n,Y_n)p(X_n,Y_n) - p_1(X_n,Y_n)q(X_n,Y_n)\right]$$

$$
\begin{aligned}
&= \frac{1}{p_1(X_n,Y_n)p(X_n,Y_n)}\left[(q_1(X_n,Y_n)-q(X_n,Y_n))p(X_n,Y_n)-(p_1(X_n,Y_n)-p(X_n,Y_n))q(X_n,Y_n)\right]\\
&= \frac{1}{p_1(X_n,Y_n)}\left[(q_1(X_n,Y_n)-q(X_n,Y_n))-(p_1(X_n,Y_n)-p(X_n,Y_n))\mathbb{E}[\lambda]\right]\\
&= \frac{1}{p_1(X_n,Y_n)}\iint(\lambda-\mathbb{E}[\lambda])\left(L_n(\lambda,\theta)-G_n(\theta)\right)\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta\\
&= \frac{1+o_P(1)}{p_1(X_n,Y_n)}\iint(\lambda-\mathbb{E}[\lambda])\,G_n(\theta)\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta\\
&= (1+o_P(1))\int\frac{G_n(\theta)}{p_1(X_n,Y_n)}\left(\int(\lambda-\mathbb{E}[\lambda])\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda\right)\pi_\theta(\theta)\,d\theta,
\end{aligned}
$$

where

$$
\begin{aligned}
p_1(X_n,Y_n) &= \iint L_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta\\
&= \iint G_n(\theta)\left(\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)+1\right)\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta\\
&= \int G_n(\theta)\left(1+\int\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda\right)\pi_\theta(\theta)\,d\theta.
\end{aligned}
$$

Therefore, it follows that

$$
\begin{aligned}
\mathbb{E}_\star[\lambda|X_n,Y_n]-\mathbb{E}[\lambda] &= \frac{\int G_n(\theta)\left(\int(\lambda-\mathbb{E}[\lambda])\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda\right)\pi_\theta(\theta)\,d\theta}{\int G_n(\theta)\left(1+\int\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda\right)\pi_\theta(\theta)\,d\theta}\\
&= \frac{\int G_n(\theta)R_{2n}(\lambda,\theta)\pi_\theta(\theta)\,d\theta}{\int G_n(\theta)R_{1n}(\lambda,\theta)\pi_\theta(\theta)\,d\theta},
\end{aligned}
$$

where $R_{1n}(\lambda,\theta)=1+\int\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda$ and

$$
R_{2n}(\lambda,\theta)=\int(\lambda-\mathbb{E}[\lambda])\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda.
$$

Equivalently, we have

$$
\int G_n(\theta)R_{1n}(\lambda,\theta)\pi_\theta(\theta)\,d\theta\cdot(\mathbb{E}_\star[\lambda|X_n,Y_n]-\mathbb{E}[\lambda])=\int G_n(\theta)R_{2n}(\lambda,\theta)\pi_\theta(\theta)\,d\theta. \tag{8}
$$

Divide (8) by $\int G_n(\theta)\pi_\theta(\theta)\,d\theta$. Then we have

$$
\frac{\int G_n(\theta)R_{1n}(\lambda,\theta)\pi_\theta(\theta)\,d\theta}{\int G_n(\theta)\pi_\theta(\theta)\,d\theta}\cdot(\mathbb{E}_\star[\lambda|X_n,Y_n]-\mathbb{E}[\lambda])=\frac{\int G_n(\theta)R_{2n}(\lambda,\theta)\pi_\theta(\theta)\,d\theta}{\int G_n(\theta)\pi_\theta(\theta)\,d\theta}. \tag{9}
$$

We now express $R_{2n}(\lambda,\theta)$ as follows:

$$
R_{2n}(\lambda,\theta)=\int(\lambda-\mathbb{E}[\lambda])\sum_{t=1}^{n}\gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\pi_\lambda(\lambda)\,d\lambda
$$

$$
\begin{aligned}
&= (1+o_P(1)) \int (\lambda - \mathbb{E}[\lambda]) \left(-\frac{K(0)}{nh} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t v_t \right) \pi_\lambda(\lambda) d\lambda \\
&= -(1+o_P(1)) K(0) \frac{1}{n a_n} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t v_t \int (\lambda - \mathbb{E}[\lambda]) \frac{1}{\lambda} \pi_\lambda(\lambda) d\lambda \\
&= -(1+o_P(1)) K(0) \frac{1}{n a_n} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t v_t \left(1 - \mathbb{E}[\lambda]\mathbb{E}[\lambda^{-1}]\right) + o_P(1) \\
&= -(1+o_P(1)) K(0) \left(1 - \mathbb{E}[\lambda]\mathbb{E}[\lambda^{-1}]\right) \frac{1}{n a_n} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t v_t + o_P(1) \\
&= \frac{d_1}{n a_n} \sum_{t=1}^{n} v_t \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t = \frac{d_1(1+o_P(1))}{n a_n} \sum_{t=1}^{n} v_t \gamma(v_t;\theta) w_t,
\end{aligned}
$$

where $d_1 = -K(0)\left(1 - \mathbb{E}[\lambda]\mathbb{E}[\lambda^{-1}]\right)$.

By ignoring the high order $o_P(1)$ in the following derivations, we therefore have

$$
\begin{aligned}
\int G_n(\theta) R_{2n}(\lambda,\theta) \pi_\theta(\theta) d\theta &= \int G_n(\theta) \left[\frac{d_1}{n a_n} \sum_{t=1}^{n} \gamma(v_t;\theta) w_t v_t \right] \pi_\theta(\theta) d\theta \\
&= \frac{d_1}{n a_n} \sum_{t=1}^{n} w_t v_t \int \gamma(v_t;\theta) G_n(\theta) \pi_\theta(\theta) d\theta. \qquad (10)
\end{aligned}
$$

Divide (10) by $\int G_n(\theta)\pi_\theta(\theta) d\theta$. It then follows that

$$
\frac{\int G_n(\theta) R_{2n}(\lambda,\theta) \pi_\theta(\theta) d\theta}{\int G_n(\theta)\pi_\theta(\theta) d\theta} = \frac{d_1}{n a_n} \sum_{t=1}^{n} w_t v_t \frac{\int \gamma(v_t;\theta) G_n(\theta) \pi_\theta(\theta) d\theta}{\int G_n(\theta)\pi_\theta(\theta) d\theta}.
$$

We know that

$$
G_n(\theta) = \prod_{t=1}^{n} f(v_t;\theta) = e^{\sum_{t=1}^{n} \log f(v_t;\theta)} = e^{n \mathbb{E}_\theta[\log f(v_1;\theta)]}(1+o_P(1)), \qquad (11)
$$

where $\mathbb{E}_\theta[\log f(v_1;\theta)] = \int \log f(v_1;\theta) f(v;\theta) dv$.

Let $A_n(\theta) = \frac{e^{n\mathbb{E}_\theta[\log f(v_1;\theta)]}}{\int e^{n\mathbb{E}_\theta[\log f(v_1;\theta)]}\pi_\theta(\theta) d\theta}$ and $\phi_n(v_t) = \int \gamma(v_t;\theta) A_n(\theta) \pi_\theta(\theta) d\theta$. Then

$$
\frac{\int G_n(\theta) R_{2n}(\lambda,\theta) \pi_\theta(\theta) d\theta}{\int G_n(\theta)\pi_\theta(\theta) d\theta} = \frac{d_1}{n a_n} \sum_{t=1}^{n} w_t v_t \phi_n(v_t). \qquad (12)
$$

Meanwhile, we have

$$
\begin{aligned}
&\int G_n(\theta) R_{1n}(\lambda,\theta) \pi_\theta(\theta) d\theta = \int G_n(\theta) \left(1 + \int \sum_{t=1}^{n} \gamma(v_t;\theta) \Gamma(\tau_t;\alpha,h) \pi_\lambda(\lambda) d\lambda \right) \pi_\theta(\theta) d\theta \\
&= \int G_n(\theta)\pi_\theta(\theta) d\theta + \int G_n(\theta) \left(\int \sum_{t=1}^{n} \gamma(v_t;\theta) \Gamma(\tau_t;\alpha,h) \pi_\lambda(\lambda) d\lambda \right) \pi_\theta(\theta) d\theta \\
&= \int G_n(\theta)\pi_\theta(\theta) d\theta + \int G_n(\theta) \left(\int \left(-\frac{K(0)}{nh} \sum_{t=1}^{n} \gamma(v_t;\theta) x_t^\top \Sigma_x^{-1} x_t v_t \right) \pi_\lambda(\lambda) d\lambda \right) \pi_\theta(\theta) d\theta
\end{aligned}
$$

$$= \int G_n(\theta)\pi_\theta(\theta)d\theta - K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_t v_t \int \gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta. \tag{13}$$

Divide (13) by $\int G_n(\theta)\pi_\theta(\theta)d\theta$. Then we have

$$\begin{aligned}\frac{\int G_n(\theta)R_{1n}(\lambda,\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} &= 1 - K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_t v_t \frac{\int \gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\ &= 1 - K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_t v_t \phi_n(v_t). \end{aligned}\tag{14}$$

Based on (12) and (14), we have

$$\left(1 - K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_t v_t \phi_n(v_t)\right)\cdot (\mathbb{E}_\star[\lambda|X_n,Y_n] - \mathbb{E}[\lambda]) = \frac{d_1}{na_n}\sum_{t=1}^n w_t v_t \phi_n(v_t). \tag{15}$$

Denote $Z_n = \frac{1}{\sqrt{n}}\sum_{t=1}^n w_t(v_t\phi_n(v_t) - \mathbb{E}[v_1\phi_n(v_1)])$. Since w_t and v_t are independent, we have $\mathbb{E}[Z_n] = 0$ and

$$\begin{aligned}\mathbb{E}[Z_n^2] &= \frac{1}{n}\sum_{t=1}^n \mathbb{E}[w_t^2]\mathbb{E}\left[\left(v_t\phi_n(v_t) - \mathbb{E}[v_1\phi_n(v_1)]\right)^2\right] \\ &= \mathbb{E}[w_1^2]\mathbb{E}\left[\left(v_t\phi_n(v_t) - \mathbb{E}[v_1\phi_n(v_1)]\right)^2\right] \equiv \delta_n^2. \end{aligned}$$

By a standard central limit theorem for the i.i.d. random variable case, we have

$$\delta_n^{-1} Z_n \to_D \mathcal{N}(0,1). \tag{16}$$

We then have $\left(1 - K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_t v_t\phi_n(v_t)\right)(\mathbb{E}_\star[\lambda|X_n,Y_n] - \mathbb{E}[\lambda]) = \frac{d_1}{na_n}\sum_{t=1}^n w_t v_t\phi_n(v_t) = \frac{d_1}{\sqrt{n}a_n}\left(Z_n + \sqrt{n}\mathbb{E}[w_1]\mathbb{E}[v_1\phi_n(v_1)]\right)$.

Using $\frac{1}{n}\sum_{t=1}^n\left(w_t v_t\phi_n(v_t) - \mathbb{E}[w_1]\mathbb{E}[v_1\phi_n(v_1)]\right) \to 0$ as $n\to\infty$, and recalling the definitions of $b_{1n} = a_n^{-1}\mathbb{E}[w_1]\mathbb{E}[v_1\phi_n(v_1)]$ and $b_{0n} = 1 - K(0)\mathbb{E}[\lambda^{-1}]b_{1n}$ listed in Theorem 1, we obtain as $n\to\infty$

$$\sqrt{n}a_n\delta_n^{-1}\left(b_{0n}\ (\mathbb{E}_\star[\lambda|X_n,Y_n] - \mathbb{E}[\lambda]) - d_1 b_{1n}\right) = \frac{d_1}{\delta_n}Z_n \to_D \mathcal{N}(0, d_1^2) \tag{17}$$

by Assumption 4 and equation (16).

Therefore, under Assumptions 1–4, when $m\to\infty$ and $n\to\infty$, we have

$$\begin{aligned}&\sqrt{n}\delta_n^{-1}a_n\left(b_{0n}\left(\widehat{\lambda}_{mn} - \mathbb{E}[\lambda]\right) - d_1 b_{1n}\right) \\ &= \sqrt{n}\delta_n^{-1}a_n\left(\frac{b_{0n}}{m}\sum_{j=1}^m\left(\lambda_{jn} - \mathbb{E}_\star[\lambda|X_n,Y_n]\right) + b_{0n}(\mathbb{E}_\star[\lambda|X_n,Y_n] - \mathbb{E}[\lambda]) - d_1 b_{1n}\right)\end{aligned}$$

$$= \frac{\sqrt{n}\delta_n^{-1} a_n b_{0n}}{\sqrt{m}} \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \left(\lambda_{jn} - \mathbb{E}_{\star}[\lambda|X_n, Y_n]\right) + \sqrt{n}\delta_n^{-1} a_n \left(b_{0n}\left(\mathbb{E}_{\star}[\lambda|X_n, Y_n] - \mathbb{E}[\lambda]\right) - d_1 b_{1n}\right)$$

$$= o_P(1) + \sqrt{n}\delta_n^{-1} a_n \left(b_{0n}\left(\mathbb{E}_{\star}[\lambda|X_n, Y_n] - \mathbb{E}[\lambda]\right) - d_1 b_{1n}\right) \longrightarrow_D \mathcal{N}(0, d_1^2). \tag{18}$$

Therefore, we have proved Theorem 1. In order to prove Theorem 2, we introduce a useful lemma in Lemma C below.

Lemma C. Under Assumption 5 listed in Appendix A1, we have as $n \to \infty$,

$$\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n, Y_n] - \mathbb{E}[\theta]\right) \to_D \mathcal{N}(0, \Sigma_0),$$

where Σ_0 is same as that in Assumption 5 (iv) and it can be estimated by a traditional method, such as the maximum likelihood estimation method.

Proof of Lemma C. By definition, we have that

$$\mathbb{E}[\theta|X_n, Y_n] = \int \theta f_n(\theta|X_n, Y_n) d\theta = \frac{\int \theta e^{g_n(\theta)} \pi_\theta(\theta) d\theta}{\int e^{g_n(\theta)} \pi_\theta(\theta) d\theta} = \frac{q_n(X_n, Y_n)}{p_n(X_n, Y_n)},$$

where $q_n(X_n, Y_n) = \int \theta e^{g_n(\theta)} \pi_\theta(\theta) d\theta$ and $p_n(X_n, Y_n) = \int e^{g_n(\theta)} \pi_\theta(\theta) d\theta$.

Taking the first order Taylor expansion of $g_n(\theta)$ in the neighborhood of $\widehat{\theta}_n$, we have that

$$g_n(\theta) = g_n(\widehat{\theta}_n) + g_n^{(1)}(\widehat{\theta}_n)(\theta - \widehat{\theta}_n) + \frac{1}{2} g_n^{(2)}(\widehat{\theta}_n)(\theta - \widehat{\theta}_n)^2 = g_n(\widehat{\theta}_n) - \frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta - \widehat{\theta}_n)^2,$$

where $g_n^{(1)}(\widehat{\theta}_n)$ and $g_n^{(2)}(\widehat{\theta}_n)$ are the first order and second order derivatives of $g_n(\theta)$ evaluated at the point $\widehat{\theta}_n$, respectively.

So $q_n(X_n, Y_n)$ can be written as

$$q_n(X_n, Y_n) = \int_{C_n} \theta e^{g_n(\widehat{\theta}_n)} e^{-\frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta - \widehat{\theta}_n)^2} \pi_\theta(\theta) d\theta + \int_{D_n} \theta e^{g_n(\theta)} \pi_\theta(\theta) d\theta = q_{1n}(X_n, Y_n) + q_{2n}(X_n, Y_n),$$

where $q_{1n}(X_n, Y_n) = \int_{C_n} \theta e^{g_n(\widehat{\theta}_n)} e^{-\frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta - \widehat{\theta}_n)^2} \pi_\theta(\theta) d\theta$ and $q_{2n}(X_n, Y_n) = \int_{D_n} \theta e^{g_n(\theta)} \pi_\theta(\theta) d\theta$.

Similarly, $p_n(X_n, Y_n)$ can be written as

$$p_n(X_n, Y_n) = \int_{C_n} e^{g_n(\widehat{\theta}_n)} e^{-\frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta - \widehat{\theta}_n)^2} \pi_\theta(\theta) d\theta + \int_{D_n} e^{g_n(\theta)} \pi_\theta(\theta) d\theta = p_{1n}(X_n, Y_n) + p_{2n}(X_n, Y_n),$$

where $p_{1n}(X_n, Y_n) = \int_{C_n} e^{g_n(\widehat{\theta}_n)} e^{-\frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta - \widehat{\theta}_n)^2} \pi_\theta(\theta) d\theta$ and $p_{2n}(X_n, Y_n) = \int_{D_n} e^{g_n(\theta)} \pi_\theta(\theta) d\theta$.

We first express $q_{1n}(X_n, Y_n)$ as follows:

$$
\begin{aligned}
&q_{1n}(X_n, Y_n) = \int_{C_n} \theta e^{g_n(\widehat{\theta}_n)} e^{-\frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta-\widehat{\theta}_n)^2} \pi_\theta(\theta) d\theta \\
&= \int_{|y|\le c} \left(\widehat{\theta}_n + \Delta_n^{-1/2}(\widehat{\theta}_n) y\right) \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n + \Delta_n^{-1/2}(\widehat{\theta}_n) y) e^{-\frac{1}{2}y^2} dy \\
&= \widehat{\theta}_n \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \int_{|y|\le c} \pi(\widehat{\theta}_n + \Delta_n^{-1/2}(\widehat{\theta}_n) y) e^{-\frac{1}{2}y^2} dy \\
&+ \Delta_n^{-1/2}(\widehat{\theta}_n) \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \int_{|y|\le c} \pi(\widehat{\theta}_n + \Delta_n^{-1/2}(\widehat{\theta}_n) y) y e^{-\frac{1}{2}y^2} dy \\
&= \widehat{\theta}_n \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n) \int_{|y|\le c} e^{-\frac{1}{2}y^2} dy + \widehat{\theta}_n \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \Delta_n^{-1/2}(\widehat{\theta}_n) \pi^{(1)}(\widehat{\theta}_n) \int_{|y|\le c} y e^{-\frac{1}{2}y^2} dy \\
&+ \Delta_n^{-1/2}(\widehat{\theta}_n) \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n) \int_{|y|\le c} y e^{-\frac{1}{2}y^2} dy + o_P(\Delta_n^{-1}(\widehat{\theta}_n)) \\
&= \widehat{\theta}_n \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n) \int_{|y|\le c} e^{-\frac{1}{2}y^2} dy + o_P(\Delta_n^{-1}(\widehat{\theta}_n)).
\end{aligned}
\tag{19}
$$

The last equality is true, because

$$
\begin{aligned}
&\int_{|y|\le c} y e^{-\frac{1}{2}y^2} dy = \int_0^c y e^{-\frac{1}{2}y^2} dy + \int_{-c}^0 y e^{-\frac{1}{2}y^2} dy \\
&= \int_0^c y e^{-\frac{1}{2}y^2} dy + \int_c^0 (-y) e^{-\frac{1}{2}y^2} d(-y) = \int_0^c y e^{-\frac{1}{2}y^2} dy - \int_0^c y e^{-\frac{1}{2}y^2} dy = 0.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&p_{1n}(X_n, Y_n) = \int_{C_n} e^{g_n(\widehat{\theta}_n)} e^{-\frac{1}{2}\Delta_n(\widehat{\theta}_n)(\theta-\widehat{\theta}_n)^2} \pi_\theta(\theta) d\theta = \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \int_{|y|\le c} \pi(\widehat{\theta}_n + \Delta_n^{-1/2}(\widehat{\theta}_n) y) e^{-\frac{1}{2}y^2} dy \\
&= \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n) \int_{|y|\le c} e^{-\frac{1}{2}y^2} dy + \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \Delta_n^{-1/2}(\widehat{\theta}_n) \pi^{(1)}(\widehat{\theta}_n) \int_{|y|\le c} y e^{-\frac{1}{2}y^2} dy \\
&+ o_P(\Delta_n^{-1}(\widehat{\theta}_n)) = \Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n) \int_{|y|\le c} e^{-\frac{1}{2}y^2} dy + o_P(\Delta_n^{-1}(\widehat{\theta}_n)).
\end{aligned}
\tag{20}
$$

Then, based on (19) and (20), we obtain

$$
\frac{q_{1n}(X_n, Y_n)}{p_{1n}(X_n, Y_n)} = \widehat{\theta}_n + \frac{o_P(\Delta_n^{-1}(\widehat{\theta}_n))}{\Delta_n^{-1/2}(\widehat{\theta}_n) e^{g_n(\widehat{\theta}_n)} \pi(\widehat{\theta}_n) \int_{|y|\le c} e^{-\frac{1}{2}y^2} dy + o_P(\Delta_n^{-1}(\widehat{\theta}_n))} = \widehat{\theta}_n + o_P(\Delta_n^{-1/2}(\widehat{\theta}_n)). \tag{21}
$$

Note that under Assumption 5 (iii), we have

$$
\frac{q_{2n}(X_n, Y_n)}{p_n(X_n, Y_n)} = o_P(\Delta_n^{-1/2}(\widehat{\theta}_n)) \text{ and } \frac{p_{2n}(X_n, Y_n)}{p_n(X_n, Y_n)} = o_P(\Delta_n^{-1/2}(\widehat{\theta}_n)). \tag{22}
$$

Therefore, in view of (21) and (22), with Assumption 5 (iii), it is shown that

$$
\left|\mathbb{E}[\theta|X_n, Y_n] - \widehat{\theta}_n\right| = \left|\frac{q_n(X_n, Y_n)}{p_n(X_n, Y_n)} - \widehat{\theta}_n\right| = \left|\frac{q_{1n}(X_n, Y_n)}{p_n(X_n, Y_n)} + \frac{q_{2n}(X_n, Y_n)}{p_n(X_n, Y_n)} - \widehat{\theta}_n\right|
$$

$$
\begin{aligned}
&\le \left|\frac{q_{1n}(X_n,Y_n)}{p_n(X_n,Y_n)} - \widehat{\theta}_n\right| + \left|\frac{q_{2n}(X_n,Y_n)}{p_n(X_n,Y_n)}\right| \\
&= \left|\frac{q_{1n}(X_n,Y_n)}{p_n(X_n,Y_n)} - \frac{q_{1n}(X_n,Y_n)}{p_{1n}(X_n,Y_n)} + \frac{q_{1n}(X_n,Y_n)}{p_{1n}(X_n,Y_n)} - \widehat{\theta}_n\right| + \left|\frac{q_{2n}(X_n,Y_n)}{p_n(X_n,Y_n)}\right| \\
&\le \left|\frac{q_{1n}(X_n,Y_n)}{p_n(X_n,Y_n)} - \frac{q_{1n}(X_n,Y_n)}{p_{1n}(X_n,Y_n)}\right| + \left|\frac{q_{1n}(X_n,Y_n)}{p_{1n}(X_n,Y_n)} - \widehat{\theta}_n\right| + \left|\frac{q_{2n}(X_n,Y_n)}{p_n(X_n,Y_n)}\right| \\
&= \left|q_{1n}(X_n,Y_n)\right|\left|\frac{1}{p_n(X_n,Y_n)} - \frac{1}{p_{1n}(X_n,Y_n)}\right| + \left|\frac{q_{1n}(X_n,Y_n)}{p_{1n}(X_n,Y_n)} - \widehat{\theta}_n\right| + \left|\frac{q_{2n}(X_n,Y_n)}{p_n(X_n,Y_n)}\right| \\
&= \left|q_{1n}(X_n,Y_n)\right|\left|\frac{p_{2n}(X_n,Y_n)}{p_n(X_n,Y_n)p_{1n}(X_n,Y_n)}\right| + o_P(\Delta_n^{-1/2}(\widehat{\theta}_n)) \\
&= \left|\frac{q_{1n}(X_n,Y_n)}{p_{1n}(X_n,Y_n)}\right|\left|\frac{p_{2n}(X_n,Y_n)}{p_n(X_n,Y_n)}\right| + o_P(\Delta_n^{-1/2}(\widehat{\theta}_n)) = o_P(\Delta_n^{-1/2}(\widehat{\theta}_n)).
\end{aligned}
$$

Thus, $\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n,Y_n] - \widehat{\theta}_n\right) \to_P 0$.

Therefore, with Assumption 5(iv), we obtain

$$
\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n,Y_n] - \mathbb{E}[\theta]\right) = \Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n,Y_n] - \widehat{\theta}_n\right) + \Delta_n^{1/2}(\widehat{\theta}_n)\left(\widehat{\theta}_n - \mathbb{E}[\theta]\right) \to_D \mathcal{N}(0,\Sigma_0), \quad (23)
$$

which completes the proof of Lemma C.

We next give the proof of Theorem 2 by Lemma C.

Proof of Theorem 2(i).

From equation (1) in the proof of Theorem 1, we have obtained that

$$
l_n(\lambda,\theta) = \log L_n(\lambda,\theta) = \sum_{t=1}^n \log f(\widehat{v}_t;\theta) = \sum_{t=1}^n \log f(v_t;\theta) + \sum_{t=1}^n \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\left(1+o_P(1)\right),
$$

where $\gamma(v_t;\theta) = \frac{f^{(1)}(v_t;\theta)}{f(v_t;\theta)}$.

Therefore, we have

$$
e^{l_n(\lambda,\theta)} = e^{\sum_{t=1}^n \log f(v_t;\theta)} e^{\sum_{t=1}^n \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)(1+o_P(1))} = e^{\sum_{t=1}^n \log f(v_t;\theta)} + R_n(\lambda,\theta),
$$

where $R_n(\lambda,\theta) = \left(e^{\sum_{t=1}^n \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)+o_P(1)} - 1\right) e^{\sum_{t=1}^n \log f(v_t;\theta)} = \left(\sum_{t=1}^n \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\right) G_n(\theta)$.

By ignoring the high order $o_P(1)$ in the following derivations, we then have the following two equations:

$$
\begin{aligned}
\iint e^{l_n(\lambda,\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta &= \iint e^{\sum_{t=1}^n \log f(v_t;\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta + \iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta \\
&= \int G_n(\theta)\pi_\theta(\theta)\,d\theta + \iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)\,d\lambda d\theta, \quad (24)
\end{aligned}
$$

$$\iint \theta e^{l_n(\lambda,\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta = \iint \theta e^{\sum_{t=1}^n \log f(v_t;\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta + \iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta$$
$$= \int \theta G_n(\theta)\pi_\theta(\theta)d\theta + \iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta. \tag{25}$$

It follows from (24) and (25) that

$$\begin{aligned}
\mathbb{E}_\star[\theta|X_n,Y_n] &= \frac{\iint \theta e^{l_n(\lambda,\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\iint e^{l_n(\lambda,\theta)}\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta} \\
&= \frac{\int \theta G_n(\theta)\pi_\theta(\theta)d\theta + \iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta + \iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}.
\end{aligned}$$

As $\mathbb{E}[\theta|X_n,Y_n] = \frac{\int \theta G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}$, we have

$$\begin{aligned}
&\mathbb{E}_\star[\theta|X_n,Y_n] - \mathbb{E}[\theta|X_n,Y_n] \\
&= \frac{\int \theta G_n(\theta)\pi_\theta(\theta)d\theta + \iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta + \iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta} - \frac{\int \theta G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= \frac{\iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} - \mathbb{E}[\theta|X_n,Y_n]\frac{\iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}.
\end{aligned} \tag{26}$$

From the proof of Theorem 1, we have obtained that

$$\sum_{t=1}^n \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h) = -\frac{K(0)}{nh}\sum_{t=1}^n v_t\gamma(v_t;\theta)w_t\,(1+o_P(1)).$$

Recall that $R_n(\lambda,\theta) = \left(\sum_{t=1}^n \gamma(v_t;\theta)\Gamma(\tau_t;\alpha,h)\right)G_n(\theta)$. Therefore, we obtain

$$\begin{aligned}
&\frac{\iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} = -\frac{\iint \left(\frac{K(0)}{nh}\sum_{t=1}^n \gamma(v_t;\theta)w_tv_t\right)G_n(\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{\int \left(\frac{1}{na_n}\sum_{t=1}^n \gamma(v_t;\theta)w_tv_t\right)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_tv_t\frac{\int \gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}.
\end{aligned}$$

Similarly, we can have

$$\begin{aligned}
&\frac{\iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} = -\frac{\iint \theta\left(\frac{K(0)}{nh}\sum_{t=1}^n \gamma(v_t;\theta)w_tv_t\right)G_n(\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{\int \left(\frac{1}{na_n}\sum_{t=1}^n \gamma(v_t;\theta)w_tv_t\right)\theta G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{na_n}\sum_{t=1}^n w_tv_t\frac{\int \theta\gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}.
\end{aligned}$$

Recall that

$$G_n(\theta) = \prod_{t=1}^{n} f(v_t;\theta) = e^{\sum_{t=1}^{n}\log f(v_t;\theta)} = e^{n\mathbb{E}_\theta[\log f(v_1;\theta)]}(1+o_P(1)), \tag{27}$$

where $\mathbb{E}_\theta[\log f(v_1;\theta)] = \int \log f(v_1;\theta) f(v;\theta)\,dv$. Denote $A_n(\theta) = \frac{e^{n\mathbb{E}_\theta[\log f(v_1;\theta)]}}{\int e^{n\mathbb{E}_\theta[\log f(v_1;\theta)]}\pi_\theta(\theta)d\theta}$.

Simple decompositions give

$$\begin{aligned}
&\mathbb{E}_\star[\theta|X_n,Y_n] - \mathbb{E}[\theta|X_n,Y_n] \\
&= \frac{\iint \theta R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} - \mathbb{E}[\theta|X_n,Y_n]\frac{\iint R_n(\lambda,\theta)\pi_\lambda(\lambda)\pi_\theta(\theta)d\lambda d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^{n} w_t v_t \frac{\int \theta\gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&+ \mathbb{E}[\theta|X_n,Y_n]K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^{n} w_t v_t \frac{\int \gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^{n} w_t v_t \left(\frac{\int \theta\gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} - \mathbb{E}[\theta|X_n,Y_n]\frac{\int \gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}\right) \\
&= -K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^{n} w_t v_t \frac{\int (\theta - \mathbb{E}[\theta|X_n,Y_n])\,\gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta} \qquad (28)\\
&= K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^{n} w_t v_t \left(\int (\mathbb{E}[\theta]-\theta)\,\gamma(v_t;\theta)A_n(\theta)\pi_\theta(\theta)d\theta\right) \\
&+ K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^{n} w_t v_t \left(\int \gamma(v_t;\theta)A_n(\theta)\pi_\theta(\theta)d\theta\right)\cdot(\mathbb{E}[\theta|X_n,Y_n]-\mathbb{E}[\theta]).
\end{aligned}$$

By Lemma C, we have $\mathbb{E}[\theta|X_n,Y_n] - \mathbb{E}[\theta] = o_P(\Delta_n^{-1/2}(\widehat{\theta}_n))$. Therefore, it follows that

$$\begin{aligned}
\mathbb{E}_\star[\theta|X_n,Y_n] - \mathbb{E}[\theta|X_n,Y_n] &= K(0)\mathbb{E}[\lambda^{-1}]\ (1+o_P(1)) \\
&\times \frac{1}{na_n}\sum_{t=1}^{n} w_t v_t \left(\int (\mathbb{E}[\theta]-\theta)\,\gamma(v_t;\theta)A_n(\theta)\pi_\theta(\theta)d\theta\right).
\end{aligned}$$

Let $\gamma_n(v_t) = \int (\mathbb{E}[\theta]-\theta)\gamma(v_t;\theta)A_n(\theta)\pi_\theta(\theta)d\theta$. Denote $S_n = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_t(v_t\gamma_n(v_t) - \mathbb{E}[v_1\gamma_n(v_1)])$.

Since w_t and v_t are independent, we have $\mathbb{E}[S_n]=0$ and

$$\begin{aligned}
\mathbb{E}[S_n S_n^\top] &= \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[w_t^2]\mathbb{E}\left[\left(v_t\gamma_n(v_t)-\mathbb{E}[v_1\gamma_n(v_1)]\right)\left(v_t\gamma_n(v_t)-\mathbb{E}[v_1\gamma_n(v_1)]\right)^\top\right] \\
&= \mathbb{E}[w_t^2]\mathbb{E}\left[\left(v_t\gamma_n(v_t)-\mathbb{E}[v_1\gamma_n(v_1)]\right)\left(v_t\gamma_n(v_t)-\mathbb{E}[v_1\gamma_n(v_1)]\right)^\top\right] \equiv \sigma_n^2.
\end{aligned}$$

A standard central limit theorem implies that as $n \to \infty$

$$\sigma_n^{-1}S_n \to_D \mathcal{N}(0, I_p), \tag{29}$$

where I_p denotes the $p \times p$ identity matrix and p denotes the number of parameters in the vector θ.

Let $T_n = \mathbb{E}_\star[\theta|X_n, Y_n] - \mathbb{E}[\theta|X_n, Y_n] = K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{\sqrt{n}a_n}\left(S_n + \sqrt{n}\mathbb{E}[w_1]\mathbb{E}[v_1\gamma_n(v_1)]\right)$. Based on (29), under Assumption 3 (iv), we obtain

$$\sqrt{n}\sigma_n^{-1}a_n\left(\frac{1}{K(0)\mathbb{E}[\lambda^{-1}]}T_n - a_n^{-1}\mathbb{E}[w_1]\mathbb{E}[v_1\gamma_n(v_1)]\right) = \sigma_n^{-1}S_n \to_D \mathcal{N}(0, I_p). \tag{30}$$

Equivalently, we have

$$\begin{aligned}&\sqrt{n}\sigma_n^{-1}a_n\left(\mathbb{E}_\star[\theta|X_n, Y_n] - \mathbb{E}[\theta|X_n, Y_n] - a_n^{-1}K(0)\mathbb{E}[\lambda^{-1}]\mathbb{E}[w_1]\mathbb{E}[v_1\gamma_n(v_1)]\right)\\ &\to_D \mathcal{N}(0, K^2(0)\mathbb{E}^2[\lambda^{-1}]I_p).\end{aligned} \tag{31}$$

By a conditional central limit theorem for the i.i.d. case, we have as $m \to \infty$

$$\frac{1}{\sqrt{m}}\sum_{j=1}^m\left(\theta_{jn} - \mathbb{E}_\star[\theta|X_n, Y_n]\right) = O_P(1).$$

Let $b_{2n} = a_n^{-1}K(0)\mathbb{E}[\lambda^{-1}]\mathbb{E}[w_1]\mathbb{E}[v_1\gamma_n(v_1)]$.

Then, under Assumptions 1–3 and 5, we obtain as $m \to \infty$ and $n \to \infty$

$$\begin{aligned}&\sqrt{n}\sigma_n^{-1}a_n\left(\widehat{\theta}_{mn} - \mathbb{E}[\theta] - b_{2n}\right)\\ &= \sqrt{n}\sigma_n^{-1}a_n\left(\frac{1}{m}\sum_{j=1}^m\theta_{jn} - \mathbb{E}_\star[\theta|X_n, Y_n] + \mathbb{E}_\star[\theta|X_n, Y_n] - \mathbb{E}[\theta|X_n, Y_n] - b_{2n} + \mathbb{E}[\theta|X_n, Y_n] - \mathbb{E}[\theta]\right)\\ &= \sqrt{n}\sigma_n^{-1}a_n\left(\frac{1}{m}\sum_{j=1}^m\theta_{jn} - \mathbb{E}_\star[\theta|X_n, Y_n]\right) + \sqrt{n}\sigma_n^{-1}a_n\left(\mathbb{E}_\star[\theta|X_n, Y_n] - \mathbb{E}[\theta|X_n, Y_n] - b_{2n}\right)\\ &+ \sqrt{n}\sigma_n^{-1}a_n\Delta_n^{-1/2}(\widehat{\theta}_n)\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n, Y_n] - \mathbb{E}[\theta]\right)\\ &= \frac{\sqrt{n}\sigma_n^{-1}a_n}{\sqrt{m}}\frac{1}{\sqrt{m}}\sum_{j=1}^m\left(\theta_{jn} - \mathbb{E}_\star[\theta|X_n, Y_n]\right) + \sqrt{n}\sigma_n^{-1}a_n\left(\mathbb{E}_\star[\theta|X_n, Y_n] - \mathbb{E}[\theta|X_n, Y_n] - b_{2n}\right) + o_P(1)\\ &\to_D \mathcal{N}(0, K^2(0)\mathbb{E}^2[\lambda^{-1}]I_p).\end{aligned}$$

To sum up, under Assumptions 1–3 and 5, as $m \to \infty$ and $n \to \infty$, we have

$$\sqrt{n}\sigma_n^{-1}a_n\left(\widehat{\theta}_{mn} - \mathbb{E}[\theta] - b_{2n}\right) \to_D \mathcal{N}(0, K^2(0)\mathbb{E}^2[\lambda^{-1}]I_p),$$

which completes the proof of Theorem 2(i).

Proof of Theorem 2(ii).

Equation (28) implies

$$\mathbb{E}_\star[\theta|X_n,Y_n]-\mathbb{E}[\theta|X_n,Y_n]=-K(0)\mathbb{E}[\lambda^{-1}]\frac{1}{a_n n}\sum_{t=1}^n w_t v_t\frac{\int(\theta-\mathbb{E}[\theta|X_n,Y_n])\gamma(v_t;\theta)G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}$$
$$=-\frac{K(0)\mathbb{E}[\lambda^{-1}]}{a_n n}\sum_{t=1}^n v_t\gamma(v_t)w_t\left(\frac{\int(\theta-\mathbb{E}[\theta|X_n,Y_n])\,G_n(\theta)\pi_\theta(\theta)d\theta}{\int G_n(\theta)\pi_\theta(\theta)d\theta}\right)$$
$$=-\frac{K(0)\mathbb{E}[\lambda^{-1}]}{a_n n}\sum_{t=1}^n v_t\gamma(v_t)w_t\cdot(\mathbb{E}[\theta|X_n,Y_n]-\mathbb{E}[\theta|X_n,Y_n])=0$$

when $\gamma(v;\theta)\equiv\gamma(v)$.

Therefore, we have as $m\to\infty$ and $n\to\infty$

$$\Delta_n^{1/2}(\widehat{\theta}_n)\left(\widehat{\theta}_{mn}-\mathbb{E}[\theta]\right)$$
$$=\Delta_n^{1/2}(\widehat{\theta}_n)\left(\frac{1}{m}\sum_{j=1}^m\theta_{jn}-\mathbb{E}_\star[\theta|X_n,Y_n]+\mathbb{E}_\star[\theta|X_n,Y_n]-\mathbb{E}[\theta|X_n,Y_n]+\mathbb{E}[\theta|X_n,Y_n]-\mathbb{E}[\theta]\right)$$
$$=\Delta_n^{1/2}(\widehat{\theta}_n)\left(\frac{1}{m}\sum_{j=1}^m\theta_{jn}-\mathbb{E}_\star[\theta|X_n,Y_n]\right)+\Sigma_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}_\star[\theta|X_n,Y_n]-\mathbb{E}[\theta|X_n,Y_n]\right)$$
$$+\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n,Y_n]-\mathbb{E}[\theta]\right)$$
$$=\frac{\Delta_n^{1/2}(\widehat{\theta}_n)}{\sqrt{m}}\frac{1}{\sqrt{m}}\sum_{j=1}^m\left(\theta_{jn}-\mathbb{E}_\star[\theta|X_n,Y_n]\right)+\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n,Y_n]-\mathbb{E}[\theta]\right)$$
$$=\Delta_n^{1/2}(\widehat{\theta}_n)\left(\mathbb{E}[\theta|X_n,Y_n]-\mathbb{E}[\theta]\right)+o_P(1)\to_D\mathcal{N}(0,\Sigma_0),$$

which completes the proof of Theorem 2(ii).

Appendix D: Plots for $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ in simulation

Figure 1: Median of $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ based on 1000 replications under DGP1 in Case 1 with sample size $n = 200, 600$ and 1200 (Graphs of $\widehat{\beta}_1(\tau)$ are displayed in the first row and graphs of $\widehat{\beta}_2(\tau)$ are displayed in the second row).

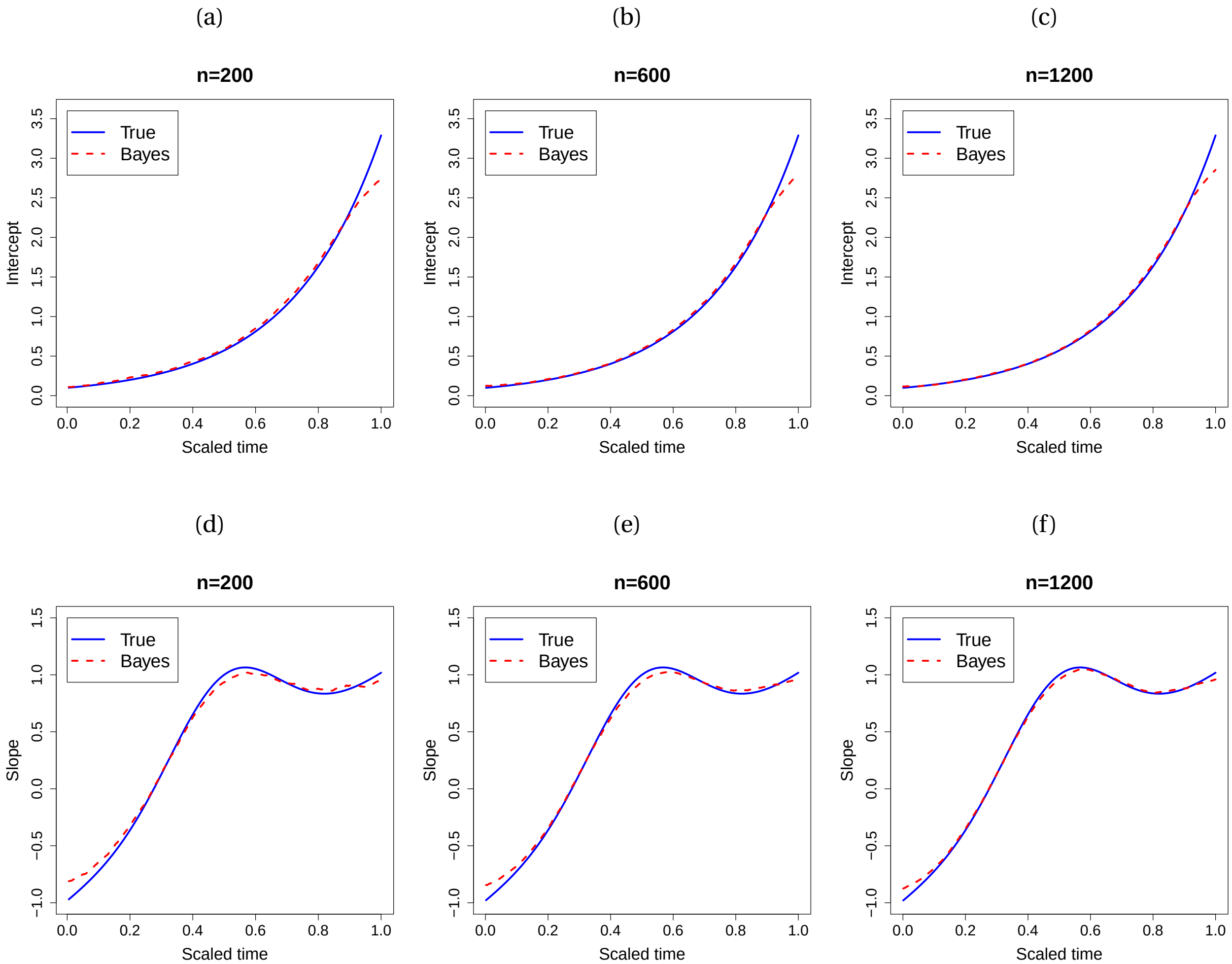


Figure 2: Median of $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ based on 1000 replications under DGP2 in Case 1 with sample size $n = 200, 600$ and 1200 (Graphs of $\widehat{\beta}_1(\tau)$ are displayed in the first row and graphs of $\widehat{\beta}_2(\tau)$ are displayed in the second row).

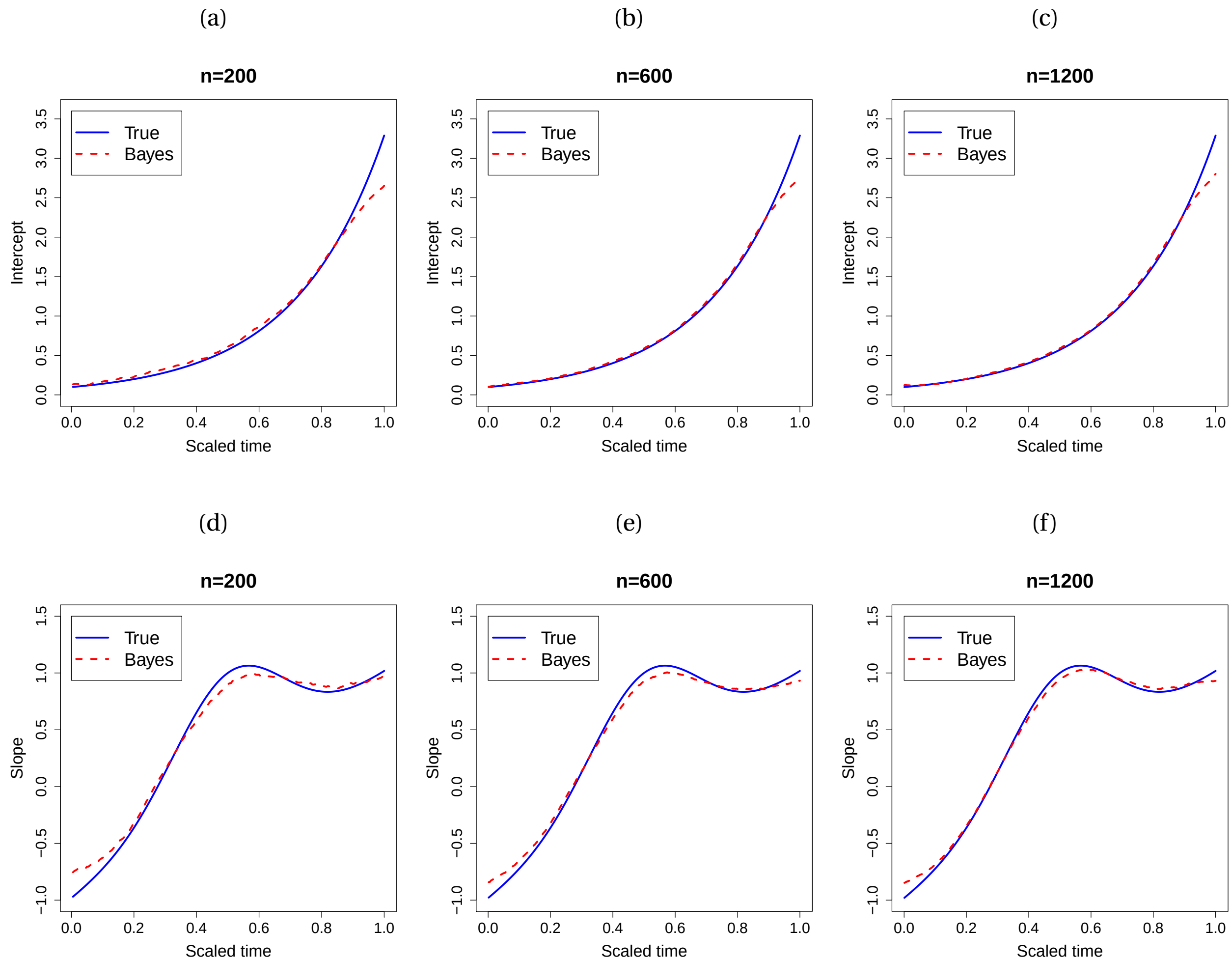


Figure 3: Median of $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ based on 1000 replications under DGP3 in Case 1 with sample size $n = 200, 600$ and 1200 (Graphs of $\widehat{\beta}_1(\tau)$ are displayed in the first row and graphs of $\widehat{\beta}_2(\tau)$ are displayed in the second row).

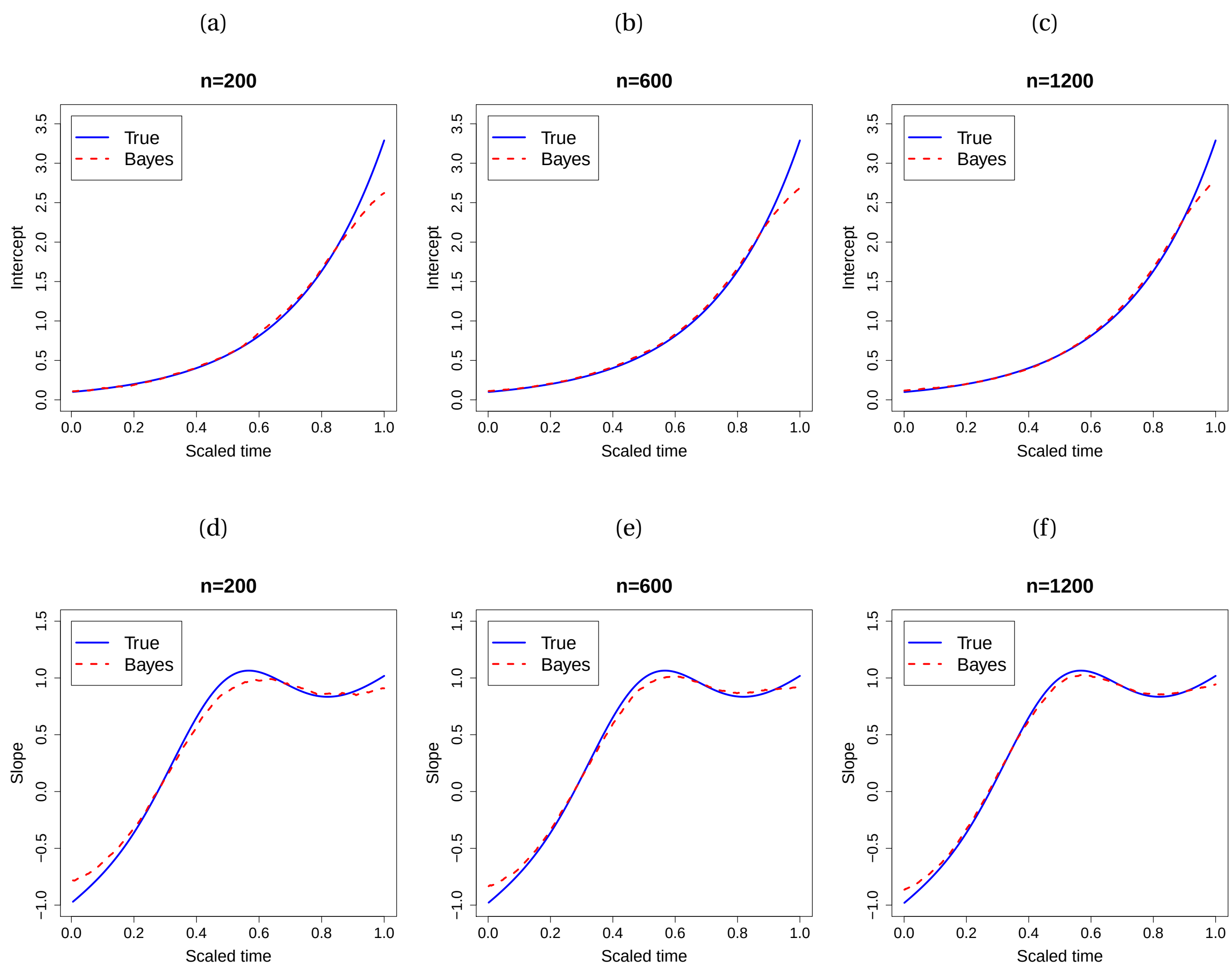


Figure 4: Median of $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ based on 1000 replications under DGP1 in Case 2 with sample size $n = 200, 600$ and 1200 (Graphs of $\widehat{\beta}_1(\tau)$ are displayed in the first row and graphs of $\widehat{\beta}_2(\tau)$ are displayed in the second row).

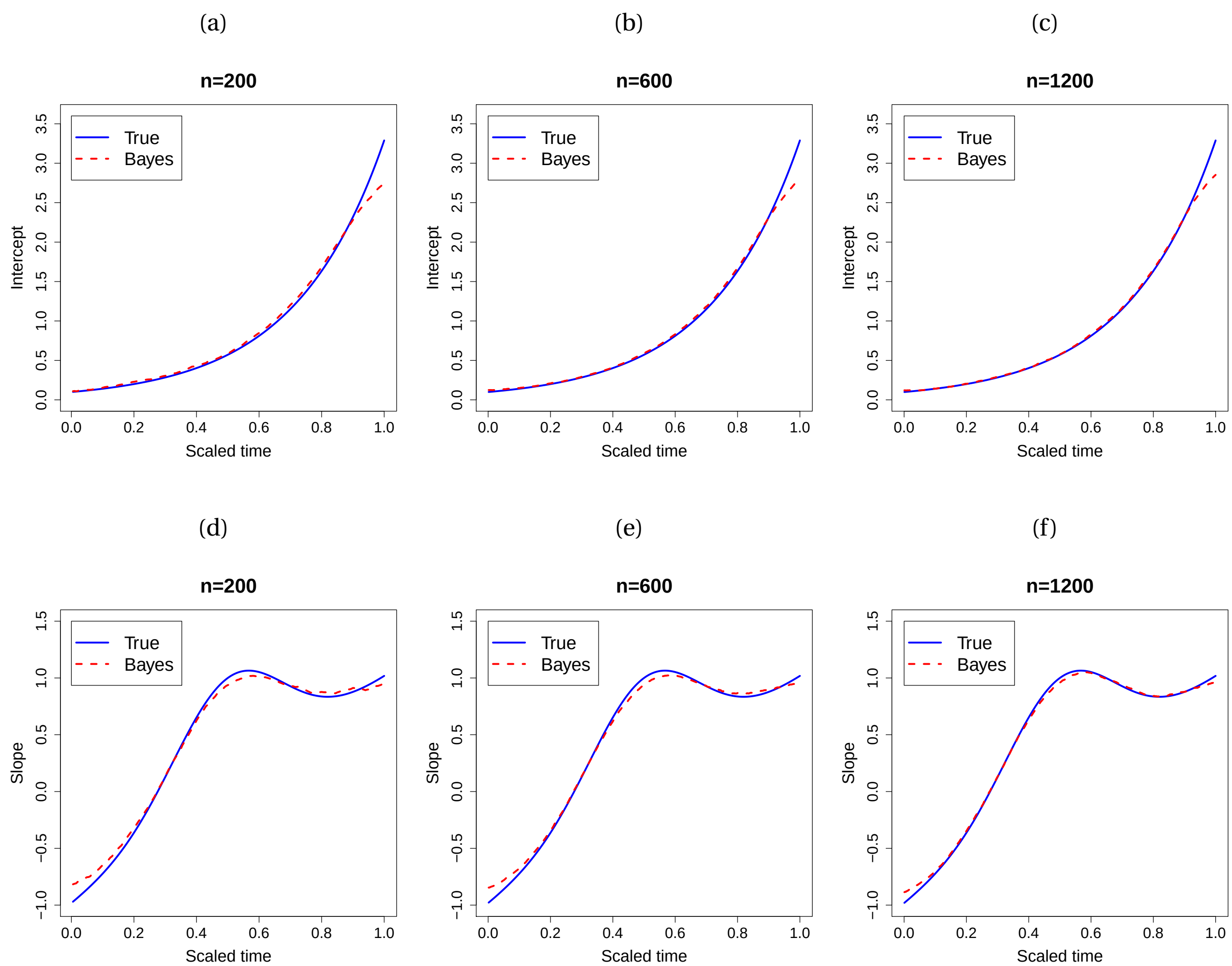
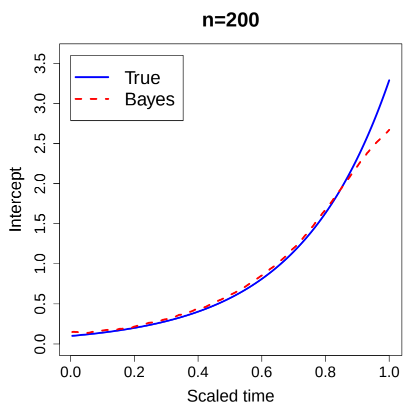
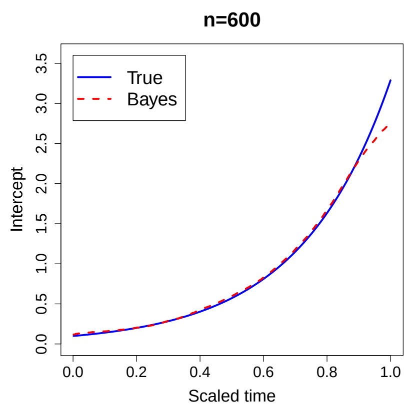


Figure 5: Median of $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ based on 1000 replications under DGP2 in Case 2 with sample size $n = 200, 600$ and 1200 (Graphs of $\widehat{\beta}_1(\tau)$ are displayed in the first row and graphs of $\widehat{\beta}_2(\tau)$ are displayed in the second row).

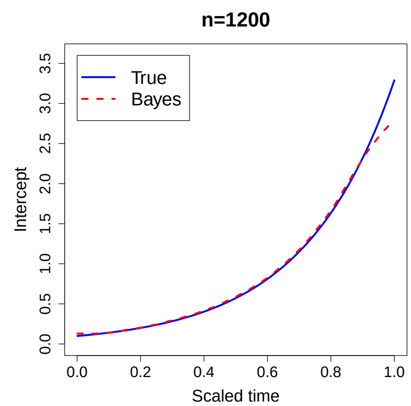
(a)



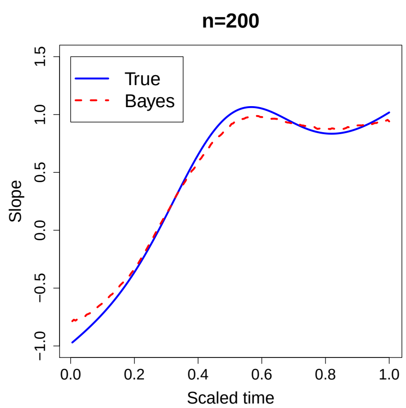
(b)



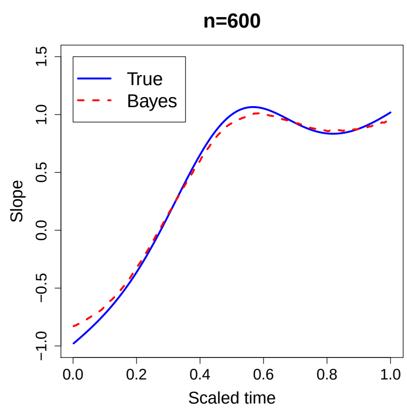
(c)



(d)



(e)



(f)

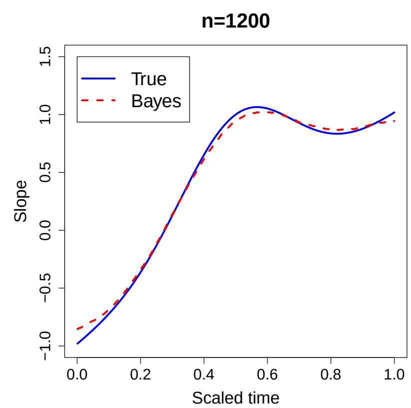
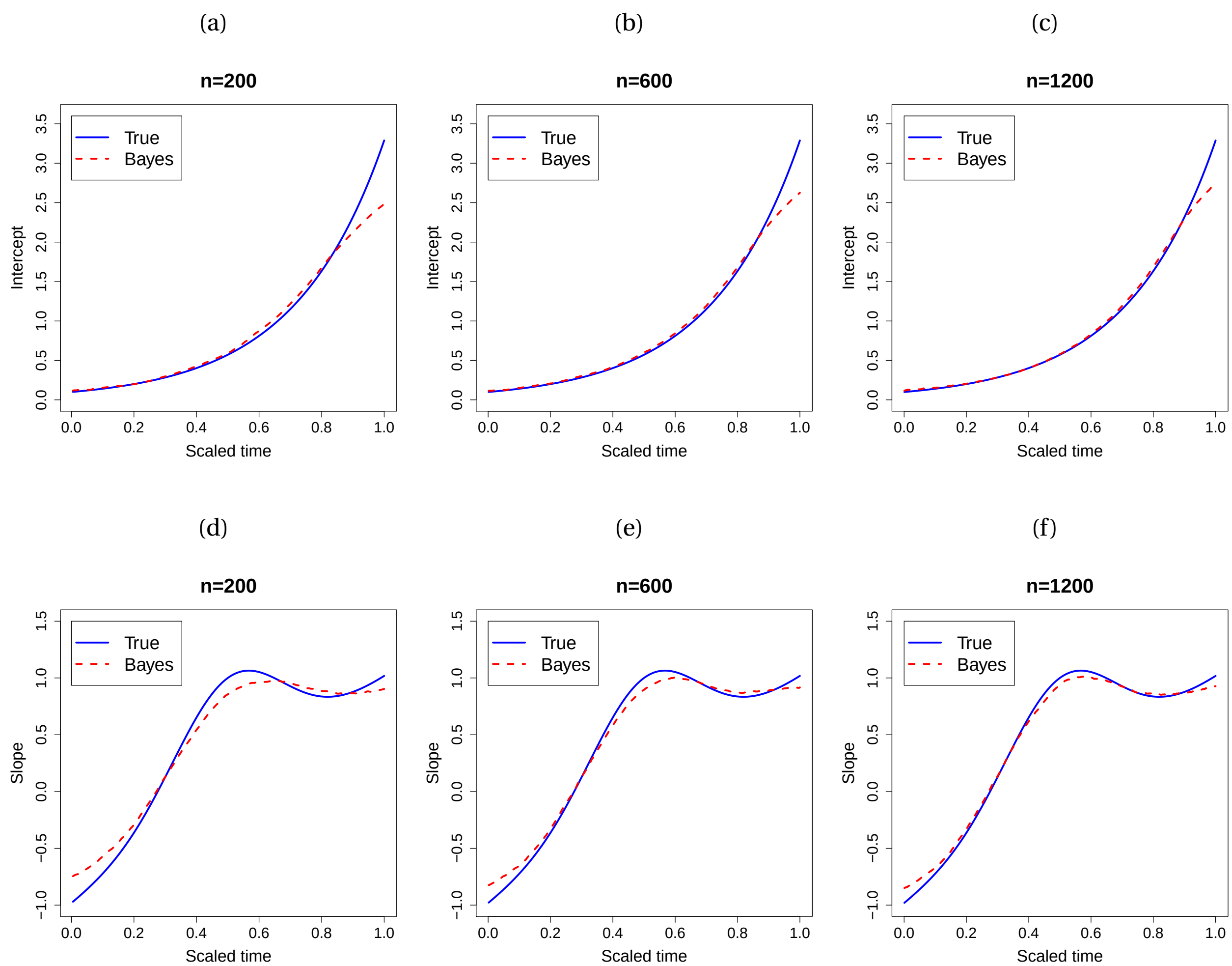


Figure 6: Median of $\widehat{\beta}_1(\tau)$ and $\widehat{\beta}_2(\tau)$ based on 1000 replications under DGP3 in Case 2 with sample size $n = 200, 600$ and 1200 (Graphs of $\widehat{\beta}_1(\tau)$ are displayed in the first row and graphs of $\widehat{\beta}_2(\tau)$ are displayed in the second row).



Appendix E: Plug–in method for bandwidth selection

In this appendix, we aim to obtain a formula for optimal bandwidth by minimizing mean squared error of local constant estimator of $\beta(\tau)$.

The local constant estimator of $\beta(\tau)$ is given by

$$\widehat{\beta}(\tau; h) = \left(\sum_{t=1}^{n} x_t x_t^\top K_h(\tau_t - \tau)\right)^{-1} \sum_{t=1}^{n} x_t y_t K_h(\tau_t - \tau)$$

$$= \left(\sum_{t=1}^{n} x_t x_t^\top K_h(\tau_t - \tau)\right)^{-1} \sum_{t=1}^{n} x_t (x_t^\top \beta(\tau_t) + u_t) K_h(\tau_t - \tau)$$
$$= \left(\sum_{t=1}^{n} x_t x_t^\top K_h(\tau_t - \tau)\right)^{-1} \sum_{t=1}^{n} x_t x_t^\top \beta(\tau_t) K_h(\tau_t - \tau) + \left(\sum_{t=1}^{n} x_t x_t^\top K_h(\tau_t - \tau)\right)^{-1} \sum_{t=1}^{n} x_t u_t K_h(\tau_t - \tau).$$

Let $p_n(\tau; h) = \frac{1}{nh}\sum_{t=1}^{n} x_t x_t^\top K\left(\frac{\tau_t - \tau}{h}\right)$. We then have

$$p_n(\tau; h) = \frac{1}{nh}\sum_{t=1}^{n} x_t x_t^\top K\left(\frac{\tau_t - \tau}{h}\right) = (1 + o(1))\mathbb{E}[x_1 x_1^\top]\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{\tau_t - \tau}{h}\right)$$
$$= (1 + o(1))\mathbb{E}[x_1 x_1^\top]\int \frac{1}{h}K\left(\frac{u - \tau}{h}\right) du = (1 + o(1))\mathbb{E}[x_1 x_1^\top] = \Sigma_x (1 + o(1)),$$

where $\Sigma_x = \mathbb{E}[x_1 x_1^\top]$.

Existing results from Cai (2007) imply

$$\text{bias}(\widehat{\beta}(\tau; h)) = \mathbb{E}[\widehat{\beta}(\tau; h)] - \beta(\tau) = \frac{\mu_2 h^2 \beta^{(2)}(\tau)}{2} + o(h^2), \tag{32}$$

where $\mu_2 = \int w^2 K(w) dw$.

It then follows from Robinson (1989) that the asymptotic variance of $\widehat{\beta}(\tau; h)$ is given by

$$\text{tr}\left(\text{var}(\widehat{\beta}(\tau; h))\right) = \frac{1}{nh}\sigma_u^2 \cdot \int K^2(w) dw \cdot \text{tr}\left(\Sigma_x^{-1}\right) + o\left(\frac{1}{nh}\right), \tag{33}$$

where $\sigma_u^2 = \mathbb{E}[u_1^2]$.

Therefore, based on (32) and (33), we can obtain

$$\text{MSE}(\widehat{\beta}(\tau; h)) = \frac{h^4}{4}\mu_2^2 \|\beta^{(2)}(\tau_t)\|^2 + \text{tr}\left(\text{var}(\widehat{\beta}(\tau_t; h))\right) + o\left(h^4 + (nh)^{-1}\right)$$
$$= \frac{h^4}{4}\mu_2^2 \|\beta^{(2)}(\tau_t)\|^2 + \frac{1}{nh}\sigma_u^2 \int K^2(w) dw \, \text{tr}\left(\Sigma_x^{-1}\right) + o\left(h^4 + (nh)^{-1}\right)$$
$$= C_1 h^4 + \frac{C_2}{nh} + o\left(h^4 + (nh)^{-1}\right),$$

where $C_1 = \frac{1}{4}\mu_2^2 \|\beta^{(2)}(\tau_t)\|^2$ and $C_2 = \sigma_u^2 \int K^2(w) dw \, \text{tr}\left(\Sigma_x^{-1}\right)$.

By minimizing the leading terms of $\text{MSE}(\widehat{\beta}(\tau; h))$, we can obtain the optimal bandwidth as follows:

$$h_{\text{opt}} = \left(\frac{C_2}{4C_1}\right)^{\frac{1}{5}} n^{-\frac{1}{5}}. \tag{34}$$

From (34), we can find that we need to estimate $\beta^{(2)}(\tau)$ to obtain a plug–in estimate for h. In order to do so, one may need to employ a local quadratic kernel method to estimate $\beta(\tau)$. In the simulation and empirical application, in order to avoid estimating $\beta^{(2)}(\tau)$, we use a normal reference rule (NRR), which is equivalent to h_{opt} in terms of the leading order $n^{-\frac{1}{5}}$.

References

Cai, Z. (2007), 'Trending time–varying coefficient time series models with serially correlated errors', *Journal of Econometrics* **136**, 163–188.

Gao, J. (2007), *Nonlinear Time Series: Semi– and Non–Parametric Methods*, Chapman & Hall/CRC, London.

Robinson, P. M. (1989), Nonparametric estimation of time–varying parameters, *in* P. Hackl, ed., 'Statistical Analysis and Forecasting of Economics Structural Change', Springer, Berlin, pp. 253–264.