Research Article

The Profile Likelihood Estimation for Single-Index ARCH(p)-M Model

Qiang Xiong,¹ Yuan Li,^{1,2} and XingFa Zhang^{1,2}

¹ School of Economics and Statistics, Guangzhou University, Guangzhou, Guangdong 510006, China

² Center for Statistical Science of Lingnan, Guangzhou University, Guangzhou, Guangdong, 510006, China

Correspondence should be addressed to Yuan Li; mathly@gzhu.edu.cn

Received 8 July 2013; Revised 22 November 2013; Accepted 30 December 2013; Published 20 February 2014

Academic Editor: Xiaojun Wang

Copyright © 2014 Qiang Xiong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a class of single-index ARCH(p)-M models and investigate estimators of the parametric and nonparametric components. We first estimate the nonparametric component using local linear smoothing technique and then construct an estimator of parametric component by using profile quasimaximum likelihood method. Under regularity conditions, the asymptotic properties of our estimators are established.

1. Introduction

As an extension of autoregressive conditional heteroskedastic (ARCH) model, Engle et al. [1] introduced conditional variance into the conditional mean return equation and proposed the generalized ARCH-in mean (GARCH-M) model. In the past decades, (G)ARCH-M model has been widely studied by many researchers. It is useful to describe the relationship between the risk and return in finance. Consider a special conditional mean of ARCH-M model with the form $y_t = \delta h_t + \varepsilon_t$, where y_t denotes a return series, ε_t denotes an error series, and h_t denotes a conditional return volatility series. The above equality gives straightforward linear relationship between risk and return: high risk causes high return. The volatility coefficient δ can be addressed as relative risk aversion parameter in Das and Sarkar [2] and price of volatility in Chou et al. [3]. Many empirical studies have been done based on the above conditional mean. However, some researchers found that δ is time varying rather than constant. To capture the variation of the volatility coefficient, Chou et al. [3] proposed a time-varying parameter GARCH-M model. Zhang [4] introduced a functional coefficient GARCH-M model in which the volatility coefficient was treated as an unknown univariate function. Zhang's [4] model enables us to study the relationship between risk aversion and certain variable. However, Chou et al. [3] suggested

that the risk aversion could be simultaneously influenced by some macroeconomic variables such as inflation rate, interest rate, and consumption price index. Consequently, it is reasonable to regard volatility coefficient as a function of several economic variables. In this paper, we will consider the following single-index ARCH(p)-M model:

$$\begin{aligned} y_t &= m(\beta^\tau x_t) h_t + \varepsilon_t, \quad \varepsilon_t = e_t \sqrt{h_t}, \\ h_t &= \omega_0 + \omega_1 y_{t-1}^2 + \omega_2 y_{t-2}^2 + \cdots + \omega_p y_{t-p}^2, \end{aligned} \quad (1)$$

where $\omega = (\omega_0, \omega_1, \dots, \omega_p)^\tau$ and $\beta = (\beta_1, \beta_2, \dots, \beta_q)^\tau$ are $p \times 1$ and $q \times 1$ vectors of unknown parameters, respectively. The risk aversion $m(\cdot)$ is an unknown smooth function of some macroeconomic variables $x_t = (x_{t1}, x_{t2}, \dots, x_{tq})^\tau$. $\{e_t\}_{t=1}^\infty$ is an independent and identically distributed (i.i.d.) $(0, 1)$ sequence. $\{y_s\}$ is independent of $\{e_t\}$ for any $t > s$ and $\{e_t\}$ is independent of $\{h_t, x_t\}$. $E(\varepsilon_t^2 \mid F_{t-1}) = h_t$ is the conditional variance, F_{t-1} denotes the sigma field generated by the information available up to time $t-1$. Let $\theta = (\omega^\tau, \beta^\tau)^\tau$ and the superscript τ denotes the transpose of a matrix or a vector.

In (1), motivated by Xia and Li [5] and Xue and Pang [6], we assume a single-index form for the volatility coefficient to avoid the so-called curse of dimensionality problem. Namely,

we treat the volatility coefficient as an unknown smooth function $m(\cdot)$ which is a function of linear combination of certain explanatory variables. Following Ling [7] and Christensen et al. [8], we make a modification for the conditional variance h_t . Namely, we use the squared time-lagged observable returns series in the conditional variance equation rather than the squared unobservable errors series. Such a modification is helpful to model estimation in that h_t is deterministic once the parameter ω is known. For simplicity and estimability of model, we consider a ARCH(p) process for the conditional variance h_t . When h_t is observable, model (1) is reduced to the single-index coefficient regression models introduced by Xia and Li [5]. Model (1) is extended from Zhang's [4] model by generalizing the univariate function coefficient to multivariate function coefficient.

The paper is arranged as follows. In Section 2, we describe the estimation methods of nonparametric and parametric components and study their asymptotic properties. The concluding remark is given in Section 3. All technicalities are put in the Appendix.

2. Methodology and Results

In this paper, we will consider the model (1). For the sake of model identifiability, the commonly used assumptions are that $\omega_i > 0$ ($i = 0, 1, 2, \dots, p$), $\|\beta\| = 1$, and $\beta_1 > 0$.

When we analyze the asymptotic properties of the estimator of β , we need to calculate the derivative of $m(\beta^\tau x_t)$ at the point β . However, $\|\beta\| = 1$ means that the true value of β is on the boundary of the unit sphere and it causes that $m(\beta^\tau x_t)$ does not have a derivative at the point β . Hence β needs to be reparameterized. We apply the "remove-one-component" method, introduced by Yu and Ruppert [9], on the parameter β . By removing β_1 , let $\beta^{(1)} = (\beta_2, \beta_3, \dots, \beta_q)^\tau$ be a $(q - 1)$ dimensional parameter vector and we have $\beta =: \beta(\beta^{(1)}) = (\beta_1, \beta^{(1)\tau})^\tau$, $\beta_1 = (1 - \|\beta^{(1)}\|^2)^{1/2}$. The new parameter $\beta^{(1)}$ satisfies the constraint $\|\beta^{(1)}\| < 1$. Then, $m(\beta^\tau x_t)$ is differentiable with respect to β_1 . The Jacobian matrix of β with respect to β_1 is obtained by

$$J_{\beta^{(1)}} = \frac{\partial \beta}{\partial \beta^{(1)}} = \begin{pmatrix} -(1 - \|\beta^{(1)}\|^2)^{-1/2} \beta^{(1)\tau} \\ I_{q-1} \end{pmatrix}, \quad (2)$$

where I_{q-1} is a $(q - 1)$ dimensional unit matrix.

In this way, the parameter space $\Theta = \mathfrak{W} \times \mathfrak{B}^{(1)}$, where

$$\mathfrak{W} = \{\omega = (\omega_0, \omega_1, \dots, \omega_p)^\tau : \omega_i > 0, i = 0, 1, 2, \dots, p\},$$

$$\mathfrak{B}^{(1)} = \{\beta^{(1)} = (\beta_2, \beta_3, \dots, \beta_q)^\tau :$$

$$\beta = \left((1 - \|\beta^{(1)}\|^2)^{1/2}, \beta^{(1)\tau} \right)^\tau, \|\beta^{(1)}\| < 1 \}. \quad (3)$$

Let $\theta^0 = (\omega^{0\tau}, \beta_0^{(1)\tau})^\tau$ and an inner point of Θ , denotes the true parameter point, where

$$\begin{aligned} \omega^0 &= (\omega_{00}, \omega_{10}, \dots, \omega_{p0})^\tau, \\ \beta^0 &= (\beta_{10}, \beta_0^{(1)\tau})^\tau, \quad \beta_0^{(1)} = (\beta_{20}, \dots, \beta_{q0})^\tau, \\ \beta_{10} &= (1 - \|\beta_0^{(1)}\|^2)^{1/2}. \end{aligned} \quad (4)$$

2.1. Estimation Method. In this section, we give details of model estimation. We first estimate the nonparametric component by local linear smoothing technique and then construct an estimate of parametric component by the profile quasimaximum likelihood method.

For any fixed $z \in S_\beta = \{\beta^\tau x : x \in S\}$, where S is a compact support of the density function of random variable x , we have

$$\begin{aligned} \sigma(z; \beta) &=: E(y_t | \beta^\tau x_t = z) \\ &= E(m(\beta^\tau x_t) h_t + e_t \sqrt{h_t} | \beta^\tau x_t = z) \\ &= m(z) E(h_t | \beta^\tau x_t = z) + E(e_t \sqrt{h_t} | \beta^\tau x_t = z) \\ &= m(z) E\left(\omega_0 + \sum_{k=1}^p \omega_k y_{t-k}^2 | \beta^\tau x_t = z\right) \\ &= m(z) \left[\omega_0 + \sum_{k=1}^p \omega_k E(y_{t-k}^2 | \beta^\tau x_t = z) \right] \\ &=: m(z) \left[\omega_0 + \sum_{k=1}^p \omega_k \sigma_k(z; \beta) \right]. \end{aligned} \quad (5)$$

Let

$$m(z; \theta) = \frac{\sigma(z; \beta)}{\omega_0 + \sum_{k=1}^p \omega_k \sigma_k(z; \beta)}. \quad (6)$$

When the true values for parameter θ are known, say θ^0 , we know $m(\beta^{0\tau} x_t; \theta^0) = m(\beta^{0\tau} x_t)$. Using local linear smoothing, we obtain the estimates of $\sigma(z; \beta)$, $\sigma_k(z; \beta)$, say $\hat{\sigma}(z; \beta)$, $\hat{\sigma}_k(z; \beta)$. Then we define the estimate of the nonparametric component $m(z; \theta)$ in (6) given by

$$\hat{m}(z; \theta) = \frac{\hat{\sigma}(z; \beta)}{\omega_0 + \sum_{k=1}^p \omega_k \hat{\sigma}_k(z; \beta)}, \quad (7)$$

where

$$\begin{aligned} \hat{\sigma}(z; \beta) &= \frac{\sum_{t=1}^n d_t(z; \beta) y_t}{\sum_{t=1}^n d_t(z; \beta)}, \\ \hat{\sigma}_k(z; \beta) &= \frac{\sum_{t=1}^n d_t(z; \beta) y_{t-k}^2}{\sum_{t=1}^n d_t(z; \beta)}, \end{aligned}$$

$$\begin{aligned}
d_t(z; \beta) &= K_h(\beta^T x_t - z) \left\{ S_{n,2}(z; \beta) - \left(\frac{\beta^T x_t - z}{h} \right) S_{n,1}(z; \beta) \right\}, \\
S_{n,j}(z; \beta) &= \frac{1}{n} \sum_{t=1}^n K_h(\beta^T x_t - z) \left(\frac{\beta^T x_t - z}{h} \right)^j, \quad j = 0, 1, 2, \dots
\end{aligned} \quad (8)$$

Let $K(\cdot)$ denote a kernel density function and h denote the bandwidth, satisfying $K_h(\cdot) = (1/h)K(\cdot/h)$.

Based on (7), we employ the quasimaximum likelihood approach to estimate parameter θ . For each $\theta \in \Theta$, let

$$L(\theta) = E \left\{ \left[\log h_t(\theta) + \frac{\varepsilon_t^2(\theta)}{h_t(\theta)} \right] \pi(x_t) \right\}, \quad (9)$$

$$L_n(\theta) = \frac{1}{n} \sum_{t=1}^n \left\{ \left[\log h_t(\theta) + \frac{\varepsilon_t^2(\theta)}{h_t(\theta)} \right] \pi(x_t) \right\}, \quad (10)$$

$$\hat{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n \left\{ \left[\log h_t(\theta) + \frac{\hat{\varepsilon}_t^2(\theta)}{h_t(\theta)} \right] \pi(x_t) \right\}, \quad (11)$$

where

$$\begin{aligned}
h_t(\theta) &= \omega_0 + \omega_1 y_{t-1}^2 + \omega_2 y_{t-2}^2 + \dots + \omega_p y_{t-p}^2, \quad h_t = h_t(\theta^0), \\
\varepsilon_t(\theta) &= y_t - m(\beta^T x_t; \theta) h_t(\theta), \quad \varepsilon_t = \varepsilon_t(\theta^0) = e_t \sqrt{h_t}, \\
\hat{\varepsilon}_t(\theta) &= y_t - \hat{m}(\beta^T x_t; \theta) h_t(\theta),
\end{aligned} \quad (12)$$

and $\pi(x_t)$ is a nonnegative weight function whose compact support is contained in S . We can get the estimate of the parametric component θ ; that is,

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \hat{L}_n(\theta). \quad (13)$$

Remark 1. (a) The weight function $\pi(\beta^T x_t)$ is substituted by $\pi(x_t)$, in that

$$\|\beta^T x_t - \beta^T x\| \leq \|\beta\| \|x_t - x\| = \|x_t - x\|. \quad (14)$$

(b) In terms of the expression of $h_t(\theta)$, we can get that

$$\begin{aligned}
\frac{\partial h_t(\theta)}{\partial \theta} &= (1, y_{t-1}^2, y_{t-2}^2, \dots, y_{t-p}^2, O_{1 \times (q-1)})^T, \\
\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta^T} &= O_{(p+q) \times (p+q)},
\end{aligned} \quad (15)$$

where $O_{s \times r}$ denotes a $s \times r$ zero matrix.

2.2. Asymptotic Properties. In order to obtain the asymptotic properties of the estimators of parametric and nonparametric component, we first introduce some regularity conditions as follows.

(A1) The probability density function $f(\beta^T x)$ of the random variable $\beta^T x$ is bounded and satisfies $\inf_{\beta^T x \in S_\beta} f(\beta^T x) > 0$, where $S_\beta = \{\beta^T x : x \in S\}$ and S is a compact support set of the probability density function $p(x)$ of x . $f(\cdot)$ and $p(\cdot)$ have two bounded and continuous derivatives.

(A2) The parameter space Θ is a bounded metric space, satisfying $\theta_L \leq \theta \leq \theta_U$.

(A3) For any $z \in S_\beta$, $\theta \in \Theta$, $k = 1, 2, \dots, p$, $\sigma(z; \beta)$ and $\sigma_k(z; \beta)$ are uniformly bounded and have continuous second-order derivatives, satisfying Lipschitz condition of order 1.

(A4) For any $z \in S_\beta$, $\theta \in \Theta$, $m(z; \theta)$, $\|\partial m(z; \theta)/\partial \theta\|$ and $\|\partial^2 m(z; \theta)/\partial \theta \partial \theta^T\|$ are uniformly bounded.

(A5) For each $\theta \in \Theta$, $x \in S$, $j = 1, 2$, $E(h_t^j(\theta) | x_t = x)$ is bounded away from infinity.

(A6) The process $\{x_t, y_t\}$ is stationary and ergodic. The process $\{e_t, x_t, y_t\}$ is α -mixing with the mixing coefficients satisfying the condition $\alpha(k) \leq ak^{-b}$ for some positive constants $a, b > r/(r-2)$ and $r > 2$. e_t satisfies $E(|e_t|^{4+\delta}) < \infty$ for some $\delta > 0$.

(A7) The function $L(\theta)$ defined in (9) is continuous and has unique minimum θ^0 .

(A8) The kernel $K(\cdot)$ is a bounded and symmetric probability density function with a bounded support and derivative and satisfies

$$\begin{aligned}
\int_{-\infty}^{\infty} u^2 K(u) du &\neq 0, \\
\int_{-\infty}^{\infty} |u|^l K(u) du &< \infty, \quad l = 1, 2, \dots
\end{aligned} \quad (16)$$

Remark 2. Condition (A1) means that the probability density function of the random variable $\beta^T x_t$ is positive, which is used to ensure that the denominators of nonparametric estimators are bounded away from 0 with high probability. Conditions (A2) and (A7) have been analogously adopted by Yang [10]. Conditions (A3)–(A6) are essential conditions for the asymptotic properties of the estimators. Condition (A6) is a standard assumption in time series literature. Condition (A8) is commonly assumed smoothness condition for the consistency of kernel based estimators.

In Section 2.1, we have described two-step estimation method and obtain the final estimate of the nonparametric component, that is, $\hat{m}(z; \hat{\theta}_n)$. We now establish the asymptotic results for the estimates of θ and $m(\cdot)$. And the related rates of convergence are given.

Theorem 3. If conditions (A1)–(A8) are satisfied, then as $n \rightarrow \infty$ and $h \rightarrow 0$ such that $nh \rightarrow 0$, it holds that

$$\sup_{\theta \in \Theta} |\hat{\sigma}(z; \beta) - \sigma(z; \beta)| = O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\},$$

$$\sup_{\theta \in \Theta} |\hat{\sigma}_k(z; \beta) - \sigma_k(z; \beta)| = O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\},$$

$$k = 1, 2, \dots, p. \quad (17)$$

Theorem 4. Under the conditions of Theorem 3, it holds that

$$\sup_{\theta \in \Theta} |\hat{m}(z; \theta) - m(z; \theta)| = O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (18)$$

Theorem 5. Under the conditions of Theorem 3, consistency for our estimates holds: $\hat{\theta}_n \xrightarrow{P} \theta^0$, $\hat{m}(z; \hat{\theta}_n) \xrightarrow{P} m(z; \theta^0)$.

Theorem 6. Under the conditions of Theorem 3, if $nh^2 \rightarrow \infty$ and $nh^4 \rightarrow 0$, $\sqrt{n}(\hat{\theta}_n - \theta^0)$ has an asymptotic normal distribution with mean zero and covariance matrix $\Sigma_I^{-1} \Sigma \Sigma_I^{-1}$, where Σ_I and Σ are given in (A.101) and (A.105) in Appendix, respectively.

3. Concluding Remark

In this paper, we consider a single-index ARCH(p)-M model. We applied profile likelihood approach to investigate the estimators for the parametric and nonparametric components. We first estimate the nonparametric component using local linear smoothing technique then construct an estimator of parametric component by using profile quasimaximum likelihood method. Under some regularity conditions, the asymptotic properties, namely, consistency and asymptotic normality with related rates of convergence, are derived.

Appendix

Throughout Appendix, M denotes a generic constant which may take different values at different places. Let $t_i \in (0, 1)$, $i = 1, 2, 3, 4$. Proofs of our theorems are based on the following lemmas. First of all, we introduce the following notations:

$$s_{n,j}(z; \beta) = \frac{1}{n} \sum_{t=1}^n K_h(\beta^\tau x_t - z) \left(\frac{\beta^\tau x_t - z}{h} \right)^j y_t,$$

$$j = 0, 1, 2, \dots,$$

$$s_{n,j,k}(z; \beta) = \frac{1}{n} \sum_{t=1}^n K_h(\beta^\tau x_t - z) \left(\frac{\beta^\tau x_t - z}{h} \right)^j y_{t-k}^2,$$

$$k = 1, 2, \dots, p, \quad j = 0, 1, 2, \dots,$$

$$g(z; \beta) = \frac{1}{n} \sum_{t=1}^n d_t(z; \beta) y_t, \quad r(z; \beta) = \frac{1}{n} \sum_{t=1}^n d_t(z; \beta),$$

$$g_k(z; \beta) = \frac{1}{n} \sum_{t=1}^n d_t(z; \beta) y_{t-k}^2, \quad k = 1, 2, \dots, p,$$

$$\mu_j = \int v^j K(v) dv. \quad (A.1)$$

Lemma A.1. Under the conditions of Theorem 3, one has

$$\sup_{\theta \in \Theta} |S_{n,j}(\beta^\tau x; \beta) - \mu_j f(\beta^\tau x)| = O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (A.2)$$

Proof. Note that

$$\begin{aligned} & E[S_{n,j}(\beta^\tau x; \beta)] \\ &= E \left[K_h(\beta^\tau x_t - \beta^\tau x) \left(\frac{\beta^\tau x_t - \beta^\tau x}{h} \right)^j \right] \\ &= \int \frac{1}{h} K \left(\frac{z_t - \beta^\tau x}{h} \right) \left(\frac{z_t - \beta^\tau x}{h} \right)^j f(z_t) dz_t \\ &= \int K(v) v^j f(vh + \beta^\tau x) dv \\ &= \mu_j f(\beta^\tau x) + R_1(\beta^\tau x; \beta), \end{aligned} \quad (A.3)$$

where $R_1(\beta^\tau x; \beta) = \int K(v) v^j [f(vh + \beta^\tau x) - f(\beta^\tau x)] dv = h \int K(v) v^{j+1} f'(t_1 vh + \beta^\tau x) dv$.

By conditions (A1) and (A8), as $h \rightarrow 0$, we have

$$\sup_{\theta \in \Theta} |R_1(\beta^\tau x; \beta)| = O(h). \quad (A.4)$$

Next, we show that

$$\sup_{\theta \in \Theta} |S_{n,j}(\beta^\tau x; \beta) - E[S_{n,j}(\beta^\tau x; \beta)]| = O_p \left\{ \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (A.5)$$

By condition (A6) and the law of large numbers, we have $S_{n,j}(\beta^\tau x; \beta) \xrightarrow{P} E[S_{n,j}(\beta^\tau x; \beta)]$.

Note

$$\begin{aligned} & |S_{n,j}(\beta_1^\tau x; \beta_1) - S_{n,j}(\beta_2^\tau x; \beta_2)| \\ & \leq \frac{1}{n} \sum_{t=1}^n K_h(\beta_1^\tau x_t - \beta_1^\tau x) \\ & \quad \times \left| \left[\frac{\beta_1^\tau x_t - \beta_1^\tau x}{h} \right]^j - \left[\frac{\beta_2^\tau x_t - \beta_2^\tau x}{h} \right]^j \right| \\ & \quad + \frac{1}{n} \sum_{t=1}^n \left| \frac{\beta_2^\tau x_t - \beta_2^\tau x}{h} \right|^j \\ & \quad \times |K_h(\beta_1^\tau x_t - \beta_1^\tau x) - K_h(\beta_2^\tau x_t - \beta_2^\tau x)| \\ & = I_1 + I_2. \end{aligned} \quad (A.6)$$

For I_1 ,

$$\begin{aligned} & \left| \left[\frac{\beta_1^\tau x_t - \beta_1^\tau x}{h} \right]^j - \left[\frac{\beta_2^\tau x_t - \beta_2^\tau x}{h} \right]^j \right| \\ & \leq \sum_{i=1}^j \left| \frac{\beta_1^\tau x_t - \beta_1^\tau x}{h} \right|^{j-i} \left| \frac{\beta_2^\tau x_t - \beta_2^\tau x}{h} \right|^{i-1} \left| \frac{(\beta_1 - \beta_2)^\tau (x_t - x)}{h} \right| \\ & \leq j \left\| \frac{x_t - x}{h} \right\|^j \|\beta_1 - \beta_2\|. \end{aligned} \quad (\text{A.7})$$

We have

$$I_1 \leq \frac{j}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} K_h(\beta^\tau x_t - \beta^\tau x) \right] \left\| \frac{x_t - x}{h} \right\|^j \|\beta_1 - \beta_2\|. \quad (\text{A.8})$$

For I_2 ,

$$\begin{aligned} & |K_h(\beta_1^\tau x_t - \beta_1^\tau x) - K_h(\beta_2^\tau x_t - \beta_2^\tau x)| \\ & \leq \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} x_t - \beta^{*\tau} x)| \right] \left\| \frac{x_t - x}{h} \right\| \|\beta_1 - \beta_2\|, \end{aligned} \quad (\text{A.9})$$

where β^* lies between β_1 and β_2 . Thus,

$$I_2 \leq \frac{1}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} x_t - \beta^{*\tau} x)| \right] \left\| \frac{x_t - x}{h} \right\|^{j+1} \|\beta_1 - \beta_2\|. \quad (\text{A.10})$$

Combining (A.8) with (A.10), we have $|S_{n,j}(z; \beta_1) - S_{n,j}(z; \beta_2)| \leq B_n \|\beta_1 - \beta_2\|$, where

$$\begin{aligned} B_n &= \frac{j}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} K_h(\beta^\tau x_t - \beta^\tau x) \right] \left\| \frac{x_t - x}{h} \right\|^j \\ &+ \frac{1}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} x_t - \beta^{*\tau} x)| \right] \left\| \frac{x_t - x}{h} \right\|^{j+1} \quad (\text{A.11}) \\ &= B_{n1} + B_{n2}. \end{aligned}$$

Note that

$$\begin{aligned} E(B_{n1}) &= \frac{j}{h} \int \left[\sup_{\theta \in \Theta} K_h(\beta^\tau x_t - \beta^\tau x) \right] \left\| \frac{x_t - x}{h} \right\|^j p(x_t) dx_t \\ &= j \int \left[\sup_{\theta \in \Theta} K(\beta^\tau v) \right] \|v\|^j p(vh + x) dv \\ &= jp(x) \int \left[\sup_{\theta \in \Theta} K(\beta^\tau v) \right] \|v\|^j dv \\ &+ jh \int \left[\sup_{\theta \in \Theta} K(\beta^\tau v) \right] \|v\|^j \frac{\partial p(t_2 vh + x)}{\partial x^\tau} v dv, \end{aligned}$$

$$\begin{aligned} E(B_{n2}) &= \frac{1}{h} \int \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} x_t - \beta^{*\tau} x)| \right] \\ &\quad \times \left\| \frac{x_t - x}{h} \right\|^{j+1} p(x_t) dx_t \\ &= \int \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} v)| \right] \|v\|^{j+1} p(vh + x) dv \\ &= p(x) \int \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} v)| \right] \|v\|^{j+1} dv \\ &+ h \int \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} v)| \right] \|v\|^{j+1} \frac{\partial p(t_2 vh + x)}{\partial x^\tau} v dv. \end{aligned} \quad (\text{A.12})$$

On the compact support set S , it is easy to obtain $E(B_n) < \infty$, implying $B_n = O_p(1)$.

According to Lemma 1 and Theorem 1 of Andrews [11], together with Lemma 1 of Xue and Pang [6], we can obtain (A.5).

Combining (A.4) with (A.5), we complete the proof of this lemma. $\square$

Lemma A.2. Under the conditions of Theorem 3, one has

$$\sup_{\theta \in \Theta} |r(z; \beta) - \mu_2 f^2(z)| = O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (\text{A.13})$$

Proof. Note that

$$\begin{aligned} r(z; \beta) &= S_{n,0}(z; \beta) S_{n,2}(z; \beta) - S_{n,1}^2(z; \beta) \\ &= [S_{n,0}(z; \beta) - f(z) + f(z)] \\ &\quad \times [S_{n,2}(z; \beta) - \mu_2 f(z) + \mu_2 f(z)] \\ &\quad - [S_{n,1}(z; \beta) - 0]^2 \\ &= [S_{n,0}(z; \beta) - f(z)] [S_{n,2}(z; \beta) - \mu_2 f(z)] \\ &\quad + \mu_2 f(z) [S_{n,0}(z; \beta) - f(z)] \\ &\quad + f(z) [S_{n,2}(z; \beta) - \mu_2 f(z)] \\ &\quad + \mu_2 f^2(z) - [S_{n,1}(z; \beta) - 0]^2. \end{aligned} \quad (\text{A.14})$$

Then, we have

$$\begin{aligned} & \sup_{\theta \in \Theta} |r(z; \beta) - \mu_2 f^2(z)| \\ & \leq \sup_{\theta \in \Theta} |S_{n,0}(z; \beta) - f(z)| \sup_{\theta \in \Theta} |S_{n,2}(z; \beta) - \mu_2 f(z)| \\ & \quad + \mu_2 f(z) \sup_{\theta \in \Theta} |S_{n,0}(z; \beta) - f(z)| \\ & \quad + f(z) \sup_{\theta \in \Theta} |S_{n,2}(z; \beta) - \mu_2 f(z)| \\ & \quad + \left[\sup_{\theta \in \Theta} |S_{n,1}(z; \beta) - 0| \right]^2. \end{aligned} \quad (\text{A.15})$$

According to Lemma A.1, it is easy to obtain the result of this lemma. $\square$

Lemma A.3. *Under the assumption of Theorem 3, one has*

$$\begin{aligned} & \sup_{\theta \in \Theta} |s_{n,j}(\beta^\tau x; \beta) - \mu_j f(\beta^\tau x) \sigma(\beta^\tau x; \beta)| \\ &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \\ & \sup_{\theta \in \Theta} |s_{n,j,k}(\beta^\tau x; \beta) - \mu_j f(\beta^\tau x) \sigma_k(\beta^\tau x; \beta)| \\ &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.16})$$

Proof. Note that

$$\begin{aligned} & E[s_{n,j}(\beta^\tau x; \beta)] \\ &= E \left\{ E \left[K_h(\beta^\tau x_t - \beta^\tau x) \left(\frac{\beta^\tau x_t - \beta^\tau x}{h} \right)^j y_t \mid \beta^\tau x_t \right] \right\} \\ &= E \left[K_h(\beta^\tau x_t - \beta^\tau x) \left(\frac{\beta^\tau x_t - \beta^\tau x}{h} \right)^j \sigma(\beta^\tau x; \beta) \right] \\ &= \int \frac{1}{h} K \left(\frac{z_t - \beta^\tau x}{h} \right) \left(\frac{z_t - \beta^\tau x}{h} \right)^j \sigma(z_t; \beta) f(z_t) dz_t \\ &= \int K(v) v^j \sigma(vh + \beta^\tau x; \beta) f(vh + \beta^\tau x) dv \\ &= \mu_j f(\beta^\tau x) \sigma(\beta^\tau x; \beta) + R_2(\beta^\tau x; \beta), \end{aligned} \quad (\text{A.17})$$

where

$$\begin{aligned} R_2(z; \beta) &= f(\beta^\tau x) \int K(v) v^j [\sigma(vh + \beta^\tau x; \beta) - \sigma(\beta^\tau x; \beta)] dv \\ &+ \sigma(\beta^\tau x; \beta) \int K(v) v^j [f(vh + \beta^\tau x) - f(\beta^\tau x)] dv \\ &+ \int K(v) v^j [\sigma(vh + \beta^\tau x; \beta) - \sigma(\beta^\tau x; \beta)] \\ &\quad \times [f(vh + \beta^\tau x) - f(\beta^\tau x)] dv \\ &= hf(\beta^\tau x) \int K(v) v^{j+1} \sigma'(t_3 vh + \beta^\tau x; \beta) dv \\ &+ h\sigma(\beta^\tau x; \beta) \int K(v) v^{j+1} f'(t_4 vh + \beta^\tau x) dv \\ &+ h^2 \int K(v) v^{j+2} \sigma'(t_3 vh + \beta^\tau x; \beta) f'(t_4 vh + \beta^\tau x) dv. \end{aligned} \quad (\text{A.18})$$

By conditions (A1) and (A3), as $h \rightarrow 0$, we have

$$\sup_{\theta \in \Theta} |R_2(\beta^\tau x; \beta)| = O(h). \quad (\text{A.19})$$

Next, we show that

$$\sup_{\theta \in \Theta} |s_{n,j}(\beta^\tau x; \beta) - E[s_{n,j}(\beta^\tau x; \beta)]| = O_p \left\{ \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (\text{A.20})$$

By condition (A6) and the law of large numbers, we have

$$s_{n,j}(\beta^\tau x; \beta) \xrightarrow{P} E[s_{n,j}(\beta^\tau x; \beta)].$$

Note

$$\begin{aligned} & |s_{n,j}(\beta_1^\tau x; \beta_1) - s_{n,j}(\beta_1^\tau x; \beta_2)| \\ &\leq \frac{1}{n} \sum_{t=1}^n K_h(\beta_1^\tau x_t - \beta_1^\tau x) \left| \left[\frac{\beta_1^\tau x_t - \beta_1^\tau x}{h} \right]^j - \left[\frac{\beta_2^\tau x_t - \beta_2^\tau x}{h} \right]^j \right| |y_t| \\ &\quad + \frac{1}{n} \sum_{t=1}^n \left[\frac{\beta_2^\tau x_t - \beta_2^\tau x}{h} \right]^j |K_h(\beta_1^\tau x_t - \beta_1^\tau x) - K_h(\beta_2^\tau x_t - \beta_2^\tau x)| |y_t| \\ &\leq \frac{Mj}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} K_h(\beta^\tau x_t - \beta^\tau x) \right] [e_t^2 + h_t(\theta_U)] \\ &\quad \times \left\| \frac{x_t - x}{h} \right\|^j \|\beta_1 - \beta_2\| \\ &\quad + \frac{M}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} x_t - \beta^{*\tau} x)| \right] \\ &\quad \times [e_t^2 + h_t(\theta_U)] \left\| \frac{x_t - x}{h} \right\|^{j+1} \|\beta_1 - \beta_2\| \\ &= C_n \|\beta_1 - \beta_2\|, \end{aligned} \quad (\text{A.21})$$

where

$$\begin{aligned} C_n &= \frac{Mj}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} K_h(\beta^\tau x_t - \beta^\tau x) \right] [e_t^2 + h_t(\theta_U)] \\ &\quad \times \left\| \frac{x_t - x}{h} \right\|^j \\ &\quad + \frac{M}{n} \sum_{t=1}^n \left[\sup_{\theta \in \Theta} |K'_h(\beta^{*\tau} x_t - \beta^{*\tau} x)| \right] \\ &\quad \times [e_t^2 + h_t(\theta_U)] \left\| \frac{x_t - x}{h} \right\|^{j+1}. \end{aligned} \quad (\text{A.22})$$

As discussing B_n in the proof of Lemma A.1, we can obtain $E(C_n) < \infty$, implying $C_n = O_p(1)$.

According to Lemma 1 and Theorem 1 of Andrews [11], together with Lemma 1 of Xue and Pang [6], we can obtain (A.20).

Combining (A.19) with (A.20), we prove the first result of this lemma. By mimicking the above proof, we can show that the other result holds. $\square$

Lemma A.4. Under the assumption of Theorem 3, one has

$$\begin{aligned} \sup_{\theta \in \Theta} |g(z; \beta) - \mu_2 f^2(z) \sigma(z; \beta)| &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \\ \sup_{\theta \in \Theta} |g_k(z; \beta) - \mu_2 f^2(z) \sigma_k(z; \beta)| &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.23})$$

Proof. The proof of Lemma A.4 is the same as that of Lemma A.2, so we omit it. $\square$

Lemma A.5. Under the conditions of Theorem 3, one has

$$\sup_{\theta \in \Theta} \left| \frac{1}{r(z; \beta)} - \frac{1}{\mu_2 f^2(z)} \right| = O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (\text{A.24})$$

Proof. By conditions (A1) and (A8), we know $f(z) \neq 0, \mu_2 \neq 0$.

According to Lemma A.2, for all $\theta \in \Theta$, it is uniformly followed that $r(z; \beta) = \mu_2 f^2(z) + o_p(1)$. Thus, for all $\theta \in \Theta$, it is uniformly followed that $1/r(z; \beta) = 1/\mu_2 f^2(z) + o_p(1)$. $\square$

Proof of Theorem 3. Note that

$$\begin{aligned} &|\hat{\sigma}(z; \beta) - \sigma(z; \beta)| \\ &= \left| \frac{g(z; \beta)}{r(z; \beta)} - \frac{\mu_2 f^2(z) \sigma(z; \beta)}{\mu_2 f^2(z)} \right| \\ &\leq \left| \frac{1}{r(z; \beta)} \right| |g(z; \beta) - \mu_2 f^2(z) \sigma(z; \beta)| \\ &\quad + \mu_2 f^2(z) |\sigma(z; \beta)| \left| \frac{1}{r(z; \beta)} - \frac{1}{\mu_2 f^2(z)} \right|. \end{aligned} \quad (\text{A.25})$$

Then, we have

$$\begin{aligned} &\sup_{\theta \in \Theta} |\hat{\sigma}(z; \beta) - \sigma(z; \beta)| \\ &\leq \left[\sup_{\theta \in \Theta} \left| \frac{1}{r(z; \beta)} - \frac{1}{\mu_2 f^2(z)} \right| + \frac{1}{\mu_2 f^2(z)} \right] \\ &\quad \times \left[\sup_{\theta \in \Theta} |g(z; \beta) - \mu_2 f^2(z) \sigma(z; \beta)| \right] \\ &\quad + \mu_2 f^2(z) \sup_{\theta \in \Theta} |\sigma(z; \beta)| \sup_{\theta \in \Theta} \left| \frac{1}{r(z; \beta)} - \frac{1}{\mu_2 f^2(z)} \right| \\ &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.26})$$

With conditions (A1), (A3), and (A8), we can obtain the first result of this theorem by Lemmas A.4 and A.5. The other result can be similarly proved. $\square$

Proof of Theorem 4. Note that

$$\begin{aligned} &\hat{m}(z; \theta) - m(z; \theta) \\ &= \frac{\hat{\sigma}(z; \beta)}{\omega_0 + \sum_{k=1}^p \omega_k \hat{\sigma}_k(z; \beta)} - \frac{\sigma(z; \beta)}{\omega_0 + \sum_{k=1}^p \omega_k \sigma_k(z; \beta)} \end{aligned}$$

$$\begin{aligned} &= \frac{\hat{\sigma}(z; \beta) - \sigma(z; \beta)}{\omega_0 + \sum_{k=1}^p \omega_k \hat{\sigma}_k(z; \beta)} \\ &\quad - \sigma(z; \beta) \left[\frac{1}{\omega_0 + \sum_{k=1}^p \omega_k \sigma_k(z; \beta)} - \frac{1}{\omega_0 + \sum_{k=1}^p \omega_k \hat{\sigma}_k(z; \beta)} \right] \\ &= \frac{\hat{\sigma}(z; \beta) - \sigma(z; \beta)}{\omega_0 + \sum_{k=1}^p \omega_k \hat{\sigma}_k(z; \beta)} \\ &\quad - \frac{\sigma(z; \beta) \sum_{k=1}^p \omega_k [\hat{\sigma}_k(z; \beta) - \sigma_k(z; \beta)]}{\left[\omega_0 + \sum_{k=1}^p \omega_k \hat{\sigma}_k(z; \beta) \right] \left[\omega_0 + \sum_{k=1}^p \omega_k \sigma_k(z; \beta) \right]}. \end{aligned} \quad (\text{A.27})$$

Then, we have

$$\begin{aligned} &\sup_{\theta \in \Theta} |\hat{m}(z; \theta) - m(z; \theta)| \\ &\leq \frac{1}{\omega_0} \sup_{\theta \in \Theta} |\hat{\sigma}(z; \beta) - \sigma(z; \beta)| \\ &\quad + \frac{\sup_{\theta \in \Theta} |\sigma(z; \beta)|}{\omega_0^2} \sum_{k=1}^p \omega_k \sup_{\theta \in \Theta} |\hat{\sigma}_k(z; \beta) - \sigma_k(z; \beta)|. \end{aligned} \quad (\text{A.28})$$

By condition (A3), applying Theorem 3, we can complete the proof of Theorem 4. $\square$

Proof of Theorem 5. By the law of large numbers, we have

$$L_n(\theta) \xrightarrow{p} L(\theta).$$

It is easy to get that

$$\begin{aligned} &L_n(\theta_1) - L_n(\theta_2) \\ &= \frac{1}{n} \sum_{t=1}^n \left\{ [\log h_t(\theta_1) - \log h_t(\theta_2)] \right. \\ &\quad \left. + \left[\frac{\varepsilon_t^2(\theta_1)}{h_t(\theta_1)} - \frac{\varepsilon_t^2(\theta_2)}{h_t(\theta_2)} \right] \right\} \pi(x_t) \\ &= \frac{1}{n} \sum_{t=1}^n \left\{ [\log h_t(\theta_1) - \log h_t(\theta_2)] \right. \\ &\quad \left. + \varepsilon_t^2(\theta_1) \left[\frac{1}{h_t(\theta_1)} - \frac{1}{h_t(\theta_2)} \right] \right. \\ &\quad \left. + \frac{1}{h_t(\theta_2)} [\varepsilon_t^2(\theta_1) - \varepsilon_t^2(\theta_2)] \right\} \pi(x_t) \\ &= \frac{1}{n} \sum_{t=1}^n \left[\frac{\partial \log h_t(\theta_1^*)}{\partial \theta^r} + \varepsilon_t^2(\theta_1) \frac{\partial}{\partial \theta^r} \left(\frac{1}{h_t(\theta_2^*)} \right) \right] \\ &\quad \times \pi(x_t) (\theta_1 - \theta_2) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n} \sum_{t=1}^n \frac{\pi(x_t)}{h_t(\theta_2)} [2y_t - m(\beta^\tau x_t; \theta_1) h_t(\theta_1) \\
& \quad - m(\beta^\tau x_t; \theta_2) h_t(\theta_2)] \\
& \quad \times \{m(\beta^\tau x_t; \theta_1) [h_t(\theta_1) - h_t(\theta_2)] \\
& \quad + h_t(\theta_2) [m(\beta^\tau x_t; \theta_1) - m(\beta^\tau x_t; \theta_2)]\} \\
& = \frac{1}{n} \sum_{t=1}^n \left[\frac{1}{h_t(\theta_1^*)} \frac{\partial h_t(\theta_1^*)}{\partial \theta^\tau} - \frac{\varepsilon_t^2(\theta_1)}{h_t^2(\theta_2^*)} \frac{\partial h_t(\theta_2^*)}{\partial \theta^\tau} \right] \\
& \quad \times \pi(x_t) (\theta_1 - \theta_2) \\
& + \frac{1}{n} \sum_{t=1}^n [2y_t - m(\beta^\tau x_t; \theta_1) h_t(\theta_1) - m(\beta^\tau x_t; \theta_2) h_t(\theta_2)] \\
& \quad \times \left[\frac{m(\beta^\tau x_t; \theta_1)}{h_t(\theta_2)} \frac{\partial h_t(\theta_3^*)}{\partial \theta^\tau} + \frac{\partial m(\beta^\tau x_t; \theta_4^*)}{\partial \theta^\tau} \right] \\
& \quad \times \pi(x_t) (\theta_1 - \theta_2), \tag{A.29}
\end{aligned}$$

where θ_i^* ($i = 1, 2, 3, 4$) lies between θ_1 and θ_2 .

Note that

$$\begin{aligned}
\frac{1}{h_t(\theta)} & < \frac{1}{\omega_0}, \quad \frac{y_{t-i}^2}{h_t(\theta)} < \frac{1}{\omega_i}, \quad i = 1, 2, \dots, p. \\
\frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \theta^\tau} & = \left(\frac{1}{h_t(\theta)}, \frac{y_{t-1}^2}{h_t(\theta)}, \dots, \frac{y_{t-p}^2}{h_t(\theta)}, \mathbf{O}_{1 \times (q-1)} \right). \tag{A.30}
\end{aligned}$$

Then, we have $\|(1/h_t(\theta))(\partial h_t(\theta)/\partial \theta^\tau)\| < M$.

We can derive that

$$\begin{aligned}
|\varepsilon_t(\theta)| & = |y_t - m(\beta^\tau x_t; \theta) h_t(\theta)| \\
& = |m(\beta^\tau x_t; \theta_0) h_t + e_t \sqrt{h_t} - m(\beta^\tau x_t; \theta) h_t(\theta)| \\
& \leq M h_t(\theta_U) + |e_t \sqrt{h_t}| \\
& \leq M h_t(\theta_U) + \frac{1}{2} (e_t^2 + h_t) \\
& \leq M [e_t^2 + h_t(\theta_U)]. \tag{A.31}
\end{aligned}$$

Similarly we have

$$|\varepsilon_t^2(\theta)| \leq M [e_t^2 h_t(\theta_U) + h_t^2(\theta_U)]. \tag{A.32}$$

Following (A.29), we have

$$|L_n(\theta_1) - L_n(\theta_2)| \leq D_n \|\theta_1 - \theta_2\|, \tag{A.33}$$

where

$$D_n = \frac{M}{n} \sum_{t=1}^n \{1 + e_t^2 + e_t^2 h_t(\theta_U) + h_t^2(\theta_U) + h_t(\theta_U)\} \pi(x_t). \tag{A.34}$$

By condition (A5), we can obtain $E(D_n) < \infty$, implying $D_n = O_p(1)$.

According to Lemma 1 and Theorem 1 of Andrews [11], we can obtain

$$\sup_{\theta \in \Theta} |L_n(\theta) - L(\theta)| = o_p(1). \tag{A.35}$$

Note that $\hat{\varepsilon}_t(\theta) = \varepsilon_t(\theta) + [m(\beta^\tau x_t; \theta) - \hat{m}(\beta^\tau x_t; \theta)] h_t(\theta)$; we have

$$\begin{aligned}
& \hat{L}_n(\theta) - L_n(\theta) \\
& = \frac{1}{n} \sum_{t=1}^n \frac{\pi(x_t)}{h_t(\theta)} [\hat{\varepsilon}_t^2(\theta) - \varepsilon_t^2(\theta)] \\
& = \frac{1}{n} \sum_{t=1}^n [\hat{m}(\beta^\tau x_t; \theta) - m(\beta^\tau x_t; \theta)]^2 h_t(\theta) \pi(x_t) \tag{A.36} \\
& \quad + \frac{2}{n} \sum_{t=1}^n [m(\beta^\tau x_t; \theta) - \hat{m}(\beta^\tau x_t; \theta)] \\
& \quad \times [y_t - m(\beta^\tau x_t; \theta) h_t(\theta)] \pi(x_t).
\end{aligned}$$

Then, we can get that

$$\begin{aligned}
& \sup_{\theta \in \Theta} |\hat{L}_n(\theta) - L_n(\theta)| \\
& \leq \frac{1}{n} \sum_{t=1}^n h_t(\theta) \pi(x_t) \left[\sup_{\theta \in \Theta} |\hat{m}(\beta^\tau x_t; \theta) - m(\beta^\tau x_t; \theta)| \right]^2 \\
& \quad + \frac{2}{n} \sum_{t=1}^n [e_t^2 + h_t(\theta_U)] \pi(x_t) \\
& \quad \times \sup_{\theta \in \Theta} |\hat{m}(\beta^\tau x_t; \theta) - m(\beta^\tau x_t; \theta)|. \tag{A.37}
\end{aligned}$$

Note that $E[h_t(\theta)\pi(x_t)] = E\{\pi(x_t)E[h_t(\theta) | x_t]\}$ and $E[e_t^2\pi(x_t)] = E[\pi(x_t)]$. Invoking Theorem 3, condition (A5), and the law of large numbers for $(1/n) \sum_{t=1}^n h_t(\theta)\pi(x_t)$ and $(1/n) \sum_{t=1}^n e_t^2\pi(x_t)$, we can show that

$$\sup_{\theta \in \Theta} |\hat{L}_n(\theta) - L_n(\theta)| = o_p(1). \tag{A.38}$$

With (A.35) and (A.38), we have

$$\sup_{\theta \in \Theta} |\hat{L}_n(\theta) - L(\theta)| = o_p(1). \tag{A.39}$$

According to Theorem 2.12 and Lemma 14.3 of Kosorok [12], we can obtain that $\hat{\theta}$ converges in probability to the true parameter θ^0 .

Next we continue to show the consistency of nonparametric estimate. Indeed, we have

$$\begin{aligned}
& \widehat{m}(z; \widehat{\theta}_n) - m(z; \theta^0) \\
&= \widehat{m}(z; \widehat{\theta}_n) - m(z; \widehat{\theta}_n) + m(z; \widehat{\theta}_n) - m(z; \theta^0) \\
&= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\} + \frac{\partial m(z; \theta)}{\partial \theta^\tau} \Big|_{\theta=\theta^0} \cdot (\widehat{\theta}_n - \theta^0) [1 + o_p(1)] \\
&= o_p(1). \quad \square
\end{aligned} \tag{A.40}$$

To prove Theorem 6, we make a Taylor expansion about $\partial \widehat{L}_n(\theta)/\partial \theta$ at θ^0 ,

$$0 = \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}_n} = \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} \Big|_{\theta=\theta^0} + \frac{\partial^2 \widehat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \Big|_{\theta=\bar{\theta}} \cdot (\widehat{\theta}_n - \theta^0), \tag{A.41}$$

where $\bar{\theta}$ lies between $\widehat{\theta}_n$ and θ^0 .

Thus one has

$$\sqrt{n}(\widehat{\theta}_n - \theta^0) = \left\{ -\frac{\partial^2 \widehat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \Big|_{\theta=\bar{\theta}} \right\}^{-1} \left\{ \sqrt{n} \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} \Big|_{\theta=\theta^0} \right\}. \tag{A.42}$$

We need the following lemmas to deal with $\sqrt{n} \partial \widehat{L}_n(\theta)/\partial \theta|_{\theta=\theta^0}$ and $-\partial^2 \widehat{L}_n(\theta)/\partial \theta \partial \theta^\tau|_{\theta=\bar{\theta}}$, respectively.

Lemma A.6. Under the conditions of Theorem 3, it holds that

$$\sup_{\theta \in \Theta} |D_{n,j}(\beta^\tau x; \beta)| = O_p \left\{ \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \quad j = 0, 1, \dots, \tag{A.43}$$

where

$$\begin{aligned}
& D_{n,j}(z; \beta) \\
&= \frac{1}{n} \sum_{t=1}^n K_h(\beta^\tau x_t - z) \left(\frac{\beta^\tau x_t - z}{h} \right)^j [y_t - \sigma(\beta^\tau x_t; \beta)].
\end{aligned} \tag{A.44}$$

Proof. Note that

$$E[D_{n,j}(\beta^\tau x; \beta)] = E\{E[D_{n,j}(\beta^\tau x; \beta) | \beta^\tau x_t]\} = 0. \tag{A.45}$$

By the law of large numbers and condition that the process $\{y_t, x_t\}$ is stationary and ergodic, we can obtain that $D_{n,j}(\beta^\tau x; \beta)$ converges in probability to 0. Then according to Lemma 1 and Theorem 1 of Andrews [11], together with Lemma 1 of Xue and Pang [6], we can prove that the result of this lemma holds as the proof of Lemma A.1. $\square$

Lemma A.7. If conditions (A1)–(A8) are satisfied, then as $n \rightarrow \infty$ and $h \rightarrow 0$ such that $nh^3 \rightarrow 0$, it holds that

$$\sup_{\theta \in \Theta} |\widehat{\sigma}'(z; \beta) - \sigma'(z; \beta)| = O_p \left\{ h + \left(\frac{\log n}{nh^3} \right)^{1/2} \right\},$$

$$\sup_{\theta \in \Theta} |\widehat{\sigma}'_k(z; \beta) - \sigma'_k(z; \beta)| = O_p \left\{ h + \left(\frac{\log n}{nh^3} \right)^{1/2} \right\}, \tag{A.46}$$

where

$$\begin{aligned}
h\widehat{\sigma}'(z; \beta) &= \frac{\sum_{t=1}^n w_t(z; \beta) y_t}{\sum_{t=1}^n d_t(z; \beta)}, \\
h\widehat{\sigma}'_k(z; \beta) &= \frac{\sum_{t=1}^n w_t(z; \beta) y_{t-k}^2}{\sum_{t=1}^n d_t(z; \beta)},
\end{aligned}$$

$$w_t(z; \beta)$$

$$= K_h(\beta^\tau x_t - z) \left\{ \left(\frac{\beta^\tau x_t - z}{h} \right) S_{n,0}(z; \beta) - S_{n,1}(z; \beta) \right\}. \tag{A.47}$$

Proof. Applying the fact that $(1/n) \sum_{t=1}^n w_t(z; \beta) = 0$, we can get that

$$\begin{aligned}
& h\widehat{\sigma}'(z; \beta) - h\sigma'(z; \beta) \\
&= \left(\frac{1}{n} \sum_{t=1}^n w_t(z; \beta) [y_t - \sigma(\beta^\tau x_t; \beta) + \sigma(\beta^\tau x_t; \beta) \right. \\
&\quad \left. - \sigma(z; \beta) - (\beta^\tau x_t - z) \sigma'(z; \beta)] \right) \\
&\quad \times \left(\frac{1}{n} \sum_{t=1}^n d_t(z; \beta) \right)^{-1} \\
&= \frac{(1/n) \sum_{t=1}^n w_t(z; \beta) [y_t - \sigma(\beta^\tau x_t; \beta)]}{(1/n) \sum_{t=1}^n d_t(z; \beta)} \\
&\quad + \left(\frac{1}{n} \sum_{t=1}^n w_t(z; \beta) \left(\frac{\beta^\tau x_t - z}{h} \right)^2 \frac{h^2}{2} \sigma''(z; \beta) \right. \\
&\quad \left. \times [1 + o(1)] \right) \left(\frac{1}{n} \sum_{t=1}^n d_t(z; \beta) \right)^{-1}.
\end{aligned} \tag{A.48}$$

Note that

$$\begin{aligned}
& \frac{1}{n} \sum_{t=1}^n d_t(z; \beta) = S_{n,0}(z; \beta) S_{n,2}(z; \beta) - S_{n,1}^2(z; \beta), \\
& \frac{1}{n} \sum_{t=1}^n w_t(z; \beta) \left(\frac{\beta^\tau x_t - z}{h} \right)^2 \\
&= S_{n,0}(z; \beta) S_{n,3}(z; \beta) - S_{n,1}(z; \beta) S_{n,2}(z; \beta), \\
& \frac{1}{n} \sum_{t=1}^n w_t(z; \beta) [y_t - \sigma(\beta^\tau x_t; \beta)] \\
&= S_{n,0}(z; \beta) D_{n,1}(z; \beta) - S_{n,1}(z; \beta) D_{n,0}(z; \beta).
\end{aligned} \tag{A.49}$$

By Lemmas A.1 and A.6, we have

$$\begin{aligned} \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n d_t(z; \beta) - \mu_2 f^2(z) \right| &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \\ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n w_t(z; \beta) \left(\frac{\beta^\tau x_t - z}{h} \right)^2 \right| &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \\ \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n w_t(z; \beta) [y_t - \sigma(\beta^\tau x_t; \beta)] \right| &= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.50})$$

Then, we can obtain that

$$\begin{aligned} h \sup_{\theta \in \Theta} |\hat{\sigma}'(z; \beta) - \sigma'(z; \beta)| &\leq \left[\sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n w_t(z; \beta) [y_t - \sigma(\beta^\tau x_t; \beta)] \right| \right] \\ &\quad \times \left[\sup_{\theta \in \Theta} \left| \frac{1}{n^{-1} \sum_{t=1}^n d_t(z; \beta)} - \frac{1}{\mu_2 f^2(z)} \right| + \frac{1}{\mu_2 f^2(z)} \right] \\ &\quad + \left[\sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n w_t(z; \beta) \left(\frac{\beta^\tau x_t - z}{h} \right)^2 \right| \right] \\ &\quad \times \left\{ \frac{h^2}{2} [1 + o(1)] \sup_{\theta \in \Theta} |\sigma''(z; \beta)| \right\} \\ &\quad \times \left[\sup_{\theta \in \Theta} \left| \frac{1}{n^{-1} \sum_{t=1}^n d_t(z; \beta)} - \frac{1}{\mu_2 f^2(z)} \right| + \frac{1}{\mu_2 f^2(z)} \right] \\ &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.51})$$

Thus, it can be shown that

$$\sup_{\theta \in \Theta} |\hat{\sigma}'(z; \beta) - \sigma'(z; \beta)| = O_p \left\{ h + \left(\frac{\log n}{nh^3} \right)^{1/2} \right\}. \quad (\text{A.52})$$

By mimicking the above proof, we can show that the other result of this lemma holds. $\square$

Lemma A.8. Under the conditions of Theorem 3, it holds that

$$\begin{aligned} \sup_{\theta \in \Theta} \left| \frac{\partial \hat{\sigma}(\beta^\tau x; \beta)}{\partial \beta^{(1)}} - \sigma'(\beta^\tau x; \beta) J^\tau [x - g(\beta^\tau x)] \right| &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \\ \sup_{\theta \in \Theta} \left| \frac{\partial \hat{\sigma}_k(\beta^\tau x; \beta)}{\partial \beta^{(1)}} - \sigma'_k(\beta^\tau x; \beta) J^\tau [x - g(\beta^\tau x)] \right| &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \quad k = 1, 2, \dots, p, \end{aligned} \quad (\text{A.53})$$

where $g(\beta^\tau x) = E(x_t | \beta^\tau x_t = \beta^\tau x)$.

Proof. Here we only prove the first result of this lemma. In the same token, we can obtain the other result.

For any given point z , we use local approximation as $\sigma(z; \beta) + \sigma'(z; \beta)(\beta^\tau x_t - z)$ to approximate $\sigma(\beta^\tau x_t; \beta)$. The estimators $\hat{\sigma}(z; \beta)$ and $\hat{\sigma}'(z; \beta)$ are obtained by solving the kernel estimating functions with respect to $\sigma(z; \beta)$, $\sigma'(z; \beta)$:

$$\begin{aligned} \sum_{t=1}^n [y_t - \hat{\sigma}(z; \beta) - \hat{\sigma}'(z; \beta)(\beta^\tau x_t - z)] k_h(\beta^\tau x_t - z) &= 0, \\ \sum_{t=1}^n [y_t - \hat{\sigma}(z; \beta) - \hat{\sigma}'(z; \beta)(\beta^\tau x_t - z)] &\quad \times (\beta^\tau x_t - z) k_h(\beta^\tau x_t - z) = 0. \end{aligned} \quad (\text{A.54})$$

The first equation of (A.54) is

$$\begin{aligned} \sum_{t=1}^n [y_t - \hat{\sigma}(\beta^\tau x; \beta) - \hat{\sigma}'(\beta^\tau x; \beta)(\beta^\tau x_t - \beta^\tau x)] &\quad \times k_h(\beta^\tau x_t - \beta^\tau x) = 0. \end{aligned} \quad (\text{A.55})$$

Taking derivatives with respect to $\beta^{(1)}$ on both sides, direct observations lead to

$$\begin{aligned} \frac{\partial \hat{\sigma}(\beta^\tau x; \beta)}{\partial \beta^{(1)}} &= [A(\beta^\tau x)]^{-1} [B_1(\beta^\tau x) + B_2(\beta^\tau x) + B_3(\beta^\tau x)], \end{aligned} \quad (\text{A.56})$$

where

$$A(\beta^\tau x) = \frac{1}{n} \sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x),$$

$$B_1(\beta^\tau x)$$

$$= -\frac{1}{n} \sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x) J^\tau(x_t - x) \hat{\sigma}'(\beta^\tau x; \beta),$$

$$B_2(\beta^\tau x) = -\frac{1}{n} \sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x)$$

$$\times (\beta^\tau x_t - \beta^\tau x) \frac{\partial \hat{\sigma}'(\beta^\tau x; \beta)}{\partial \beta^{(1)}},$$

$$B_3(\beta^\tau x)$$

$$= \frac{1}{nh} \sum_{t=1}^n k'_h(\beta^\tau x_t - \beta^\tau x) J^\tau(x_t - x)$$

$$\times [y_t - \hat{\sigma}(\beta^\tau x; \beta) - \hat{\sigma}'(\beta^\tau x; \beta)(\beta^\tau x_t - \beta^\tau x)], \quad (\text{A.57})$$

and $k'_h(\cdot) = (1/h)k'_h(\cdot/h)$. Then, we have

$$\frac{\partial \hat{\sigma}(\beta^\tau x; \beta)}{\partial \beta^{(1)}}$$

$$\begin{aligned}
&= [A(\beta^\tau x)]^{-1} B_1(\beta^\tau x) \\
&\quad + [A(\beta^\tau x)]^{-1} B_2(\beta^\tau x) + [A(\beta^\tau x)]^{-1} B_3(\beta^\tau x).
\end{aligned} \tag{A.58}$$

The proof of Lemma A.8 is summarized in the following three steps.

Step 1. It is the analysis of term $[A(\beta^\tau x)]^{-1} B_1(\beta^\tau x)$.

First we analyze $[A(\beta^\tau x)]^{-1} B_1(\beta^\tau x)$, which can be decomposed as follows:

$$\begin{aligned}
&[A(\beta^\tau x)]^{-1} B_1(\beta^\tau x) \\
&= \hat{\sigma}'(\beta^\tau x; \beta) J^\tau [x - g(\beta^\tau x)] - \hat{\sigma}'(\beta^\tau x; \beta) J^\tau \\
&\quad \cdot \left\{ \left[\sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x) \right]^{-1} \right. \\
&\quad \left. \times \sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x) x_t - g(\beta^\tau x) \right\}.
\end{aligned} \tag{A.59}$$

Note that $\sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x) x_t$ is the Nadaraya-Watson estimate for $g(\beta^\tau x)$, so it can be shown that

$$\begin{aligned}
&\sup_{\theta \in \Theta} \left| \left[\sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x) \right]^{-1} \right. \\
&\quad \left. \times \sum_{t=1}^n k_h(\beta^\tau x_t - \beta^\tau x) x_t - g(\beta^\tau x) \right| \\
&= O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}.
\end{aligned} \tag{A.60}$$

Combining (A.59) and (A.60) with Lemma A.7, we have

$$\begin{aligned}
&\sup_{\theta \in \Theta} \left| [A(\beta^\tau x)]^{-1} B_1(\beta^\tau x) - \sigma'(\beta^\tau x; \beta) J^\tau [x - g(\beta^\tau x)] \right| \\
&= O_p \left\{ h + \left(\frac{\log n}{nh^3} \right)^{1/2} \right\}.
\end{aligned} \tag{A.61}$$

Step 2. The analysis of term $[A(\beta^\tau x)]^{-1} B_2(\beta^\tau x)$.

By condition (A1) and Lemma A.1, it can be shown that

$$\sup_{\theta \in \Theta} \left| [A(\beta^\tau x)]^{-1} - f^{-1}(\beta^\tau x) \right| = O_p \left\{ h + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \tag{A.62}$$

According to the symmetric assumption for the kernel function, we have

$$\begin{aligned}
&\sup_{\theta \in \Theta} |B_2(\beta^\tau x)| \\
&\leq h \sup_{\theta \in \Theta} \left| \frac{\partial \hat{\sigma}'(\beta^\tau x; \beta)}{\partial \beta^{(1)}} \right| \left[\sup_{\theta \in \Theta} |S_{n,1}(\beta^\tau x; \beta) - \mu_1 f(\beta^\tau x)| \right] \\
&= O_p \left\{ h^2 + \left(\frac{h \log n}{n} \right)^{1/2} \right\}.
\end{aligned} \tag{A.63}$$

It is not difficult to get that

$$\begin{aligned}
&\sup_{\theta \in \Theta} \left| [A(\beta^\tau x)]^{-1} B_2(\beta^\tau x) \right| \\
&\leq \left[\sup_{\theta \in \Theta} \left| [A(\beta^\tau x)]^{-1} - f^{-1}(\beta^\tau x) \right| \right. \\
&\quad \left. + f^{-1}(\beta^\tau x) \right] \left[\sup_{\theta \in \Theta} |B_2(\beta^\tau x)| \right],
\end{aligned} \tag{A.64}$$

which implies that

$$\sup_{\theta \in \Theta} \left| [A(\beta^\tau x)]^{-1} B_2(\beta^\tau x) \right| = O_p \left\{ h^2 + \left(\frac{h \log n}{n} \right)^{1/2} \right\}. \tag{A.65}$$

Step 3. It is the analysis of term $[A(\beta^\tau x)]^{-1} B_3(\beta^\tau x)$.

We will deal with $B_3(\beta^\tau x)$, which can be rewritten as follows:

$$B_3(\beta^\tau x) = \sum_{i=1}^4 I_i(\beta^\tau x), \tag{A.66}$$

where

$$\begin{aligned}
I_1(\beta^\tau x) &= \frac{1}{nh} J^\tau \sum_{t=1}^n k'_h(\beta^\tau x_t - \beta^\tau x) (x_t - x) \\
&\quad \times [y_t - \sigma(\beta^\tau x_t; \beta)], \\
I_2(\beta^\tau x) &= \frac{1}{nh} J^\tau \sum_{t=1}^n k'_h(\beta^\tau x_t - \beta^\tau x) (x_t - x) \\
&\quad \times (\beta^\tau x_t - \beta^\tau x)^2 \left[\frac{1}{2} \sigma''(\beta^\tau x) + o(1) \right], \\
I_3(\beta^\tau x) &= \frac{1}{nh} J^\tau \sum_{t=1}^n k'_h(\beta^\tau x_t - \beta^\tau x) (x_t - x) \\
&\quad \times [\sigma(\beta^\tau x; \beta) - \hat{\sigma}(\beta^\tau x; \beta)], \\
I_4(\beta^\tau x) &= \frac{1}{nh} J^\tau \sum_{t=1}^n k'_h(\beta^\tau x_t - \beta^\tau x) (x_t - x) \\
&\quad \times (\beta^\tau x_t - \beta^\tau x) \\
&\quad \times [\sigma'(\beta^\tau x; \beta) - \hat{\sigma}'(\beta^\tau x; \beta)].
\end{aligned} \tag{A.67}$$

In the following part, we will prove that every term $I_i(\beta^\tau x)$ is of order $o_p(1)$ uniformly.

By the law of large numbers, we have

$$\frac{1}{n} \sum_{t=1}^n [y_t - \sigma(\beta^\tau x_t; \beta)] \xrightarrow{p} 0. \quad (\text{A.68})$$

By condition (A3), we can get that

$$\begin{aligned} & |y_t - \sigma(\beta^{1\tau} x_t; \beta^1) - [y_t - \sigma(\beta^{2\tau} x_t; \beta^2)]| \\ &= |\sigma(\beta^{2\tau} x_t; \beta^2) - \sigma(\beta^{1\tau} x_t; \beta^1)| \\ &\leq M \|x_t\| \|\beta^1 - \beta^2\|. \end{aligned} \quad (\text{A.69})$$

According to Lemma 1 and Theorem 1 of Andrews [11], we can show that

$$\sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n [y_t - \sigma(\beta^\tau x_t; \beta)] \right| = o_p(1). \quad (\text{A.70})$$

Denote

$$F_{nl} = \frac{1}{nh} \|J\| \sum_{t=1}^n \sup_{\theta \in \Theta} |k'_h(\beta^\tau x_t - \beta^\tau x)| \|x_t - x\|^{l+1}, \quad (\text{A.71})$$

$$l = 0, 1, 2, \dots$$

Furthermore, we can get that

$$\begin{aligned} E(F_{nl}) &= \frac{1}{h} \|J\| \int \sup_{\theta \in \Theta} |k'_h(\beta^\tau x_t - \beta^\tau x)| \\ &\quad \times \left\| \frac{x_t - x}{h} \right\|^{l+1} h^{l+1} p(x_t) dx_t \\ &= h^l \|J\| p(x) \int \sup_{\theta \in \Theta} |k'_h(\beta^\tau v)| \|v\|^{l+1} dv \\ &\quad + h^{l+1} \|J\| \int \sup_{\theta \in \Theta} |k'_h(\beta^\tau v)| \|v\|^{l+1} v p'(tvh + x) dv \\ &= O(h^l), \end{aligned} \quad (\text{A.72})$$

which implies $F_{nl} = O_p(h^l)$.

For $I_1(\beta^\tau x)$, the Abel inequality implies

$$\begin{aligned} & \sup_{\theta \in \Theta} |I_1(\beta^\tau x)| \\ &\leq \frac{M}{h} \sup_t |J^\tau k'_h(\beta^\tau x_t - \beta^\tau x)(x_t - x)| \\ &\quad \times \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n [y_t - \sigma(\beta^\tau x_t; \beta)] \right| = o_p(1). \end{aligned} \quad (\text{A.73})$$

For $I_2(\beta^\tau x)$, we have

$$\begin{aligned} & \sup_{\theta \in \Theta} |I_2(\beta^\tau x)| \\ &\leq F_{n2} \left[\frac{1}{2} \sup_{\theta \in \Theta} |\sigma''(\beta^\tau x; \beta)| + o(1) \right] = O_p(h^2). \end{aligned} \quad (\text{A.74})$$

For $I_3(\beta^\tau x)$, we have

$$\begin{aligned} & \sup_{\theta \in \Theta} |I_3(\beta^\tau x)| \\ &\leq F_{n0} \left(\sup_{\theta \in \Theta} |\sigma(\beta^\tau x; \beta) - \hat{\sigma}(\beta^\tau x; \beta)| \right) \\ &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.75})$$

For $I_4(\beta^\tau x)$, we have

$$\begin{aligned} & \sup_{\theta \in \Theta} |I_4(\beta^\tau x)| \\ &\leq F_{n1} \left(\sup_{\theta \in \Theta} |\sigma'(\beta^\tau x; \beta) - \hat{\sigma}'(\beta^\tau x; \beta)| \right) \\ &= O_p(h) O_p \left\{ h + \left(\frac{\log n}{nh^3} \right)^{1/2} \right\} \\ &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.76})$$

Consequently, we have shown that

$$\sup_{\theta \in \Theta} |B_3(\beta^\tau x)| = O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (\text{A.77})$$

Together with (A.62), we can obtain that

$$\sup_{\theta \in \Theta} | [A(\beta^\tau x)]^{-1} B_3(\beta^\tau x) | = O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \quad (\text{A.78})$$

For (A.58), we can get that

$$\begin{aligned} & \sup_{\theta \in \Theta} \left| \frac{\partial \hat{\sigma}(\beta^\tau x; \beta)}{\partial \beta^{(1)}} - \sigma'(\beta^\tau x; \beta) J^\tau [x - g(\beta^\tau x)] \right| \\ &\leq \sup_{\theta \in \Theta} | [A(\beta^\tau x)]^{-1} B_1(\beta^\tau x) \\ &\quad - \sigma'(\beta^\tau x; \beta) J^\tau [x - g(\beta^\tau x)] | \\ &\quad + \sup_{\theta \in \Theta} | [A(\beta^\tau x)]^{-1} B_2(\beta^\tau x) | \\ &\quad + \sup_{\theta \in \Theta} | [A(\beta^\tau x)]^{-1} B_3(\beta^\tau x) |. \end{aligned} \quad (\text{A.79})$$

Combination of (A.61), (A.65), and (A.78) can prove the first result. $\square$

Lemma A.9. Under the conditions of Theorem 3, it holds that

$$\begin{aligned} & \sup_{\theta \in \Theta} \left| \frac{\partial m(\beta^\tau x; \theta)}{\partial \theta} - \frac{\partial \widehat{m}(\beta^\tau x; \theta)}{\partial \theta} - H(\beta^\tau x; \theta) \right| \\ &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \end{aligned} \quad (\text{A.80})$$

where

$$H(\beta^\tau x; \theta) = (O_{1 \times (p+1)}, m'(\beta^\tau x; \theta) g^\tau(\beta^\tau x) J)^\tau. \quad (\text{A.81})$$

Proof. Let $\sigma_0(z; \beta) = 1$, $\widehat{\sigma}_0(z; \beta) = 1$. Note that

$$\begin{aligned} & \frac{\partial m(z; \theta)}{\partial \theta} - \frac{\partial \widehat{m}(z; \theta)}{\partial \theta} \\ &= \left(\frac{\partial m(z; \theta)}{\partial \omega^\tau} - \frac{\partial \widehat{m}(z; \theta)}{\partial \omega^\tau}, \frac{\partial m(z; \theta)}{\partial \beta^{(1)\tau}} - \frac{\partial \widehat{m}(z; \theta)}{\partial \beta^{(1)\tau}} \right)^\tau, \end{aligned} \quad (\text{A.82})$$

where

$$\begin{aligned} & \frac{\partial m(z; \theta)}{\partial \omega_i} - \frac{\partial \widehat{m}(z; \theta)}{\partial \omega_i} \\ &= \sigma(z; \beta) \sigma_i(z; \beta) \\ & \times \left\{ \frac{1}{\left[\sum_{k=0}^p \omega_k \widehat{\sigma}_k(z; \beta) \right]^2} - \frac{1}{\left[\sum_{k=0}^p \omega_k \sigma_k(z; \beta) \right]^2} \right\} \\ & + \left\{ \frac{1}{\left[\sum_{k=0}^p \omega_k \widehat{\sigma}_k(z; \beta) \right]^2} - \frac{1}{\left[\sum_{k=0}^p \omega_k \sigma_k(z; \beta) \right]^2} \right\} \\ & \times \{ \widehat{\sigma}(z; \beta) \widehat{\sigma}_i(z; \beta) - \sigma(z; \beta) \sigma_i(z; \beta) \} \\ & + \frac{1}{\left[\sum_{k=0}^p \omega_k \sigma_k(z; \beta) \right]^2} \\ & \times \{ \widehat{\sigma}(z; \beta) \widehat{\sigma}_i(z; \beta) - \sigma(z; \beta) \sigma_i(z; \beta) \}, \\ & \frac{\partial m(z; \theta)}{\partial \beta^{(1)}} - \frac{\partial \widehat{m}(z; \theta)}{\partial \beta^{(1)}} \\ &= \left\{ \frac{\partial \sigma(z; \beta)}{\partial \beta^{(1)}} \frac{1}{\sum_{k=0}^p \omega_k \sigma_k(z; \beta)} \right. \\ & \quad \left. - \frac{\partial \widehat{\sigma}(z; \beta)}{\partial \beta^{(1)}} \frac{1}{\sum_{k=0}^p \omega_k \widehat{\sigma}_k(z; \beta)} \right\} \\ & + \left\{ \frac{\widehat{\sigma}(z; \beta) \left[\sum_{k=0}^p \omega_k (\partial \widehat{\sigma}_k(z; \beta) / \partial \beta^{(1)}) \right]}{\left[\sum_{k=0}^p \omega_k \widehat{\sigma}_k(z; \beta) \right]^2} \right. \\ & \quad \left. - \frac{\sigma(z; \beta) \left[\sum_{k=0}^p \omega_k (\partial \sigma_k(z; \beta) / \partial \beta^{(1)}) \right]}{\left[\sum_{k=0}^p \omega_k \sigma_k(z; \beta) \right]^2} \right\}. \end{aligned} \quad (\text{A.83})$$

Combining Theorem 3 and Lemma A.8, we can obtain that

$$\begin{aligned} & \sup_{\theta \in \Theta} \left| \frac{\partial m(\beta^\tau x; \theta)}{\partial \omega} - \frac{\partial \widehat{m}(\beta^\tau x; \theta)}{\partial \omega} \right| = O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}, \\ & \sup_{\theta \in \Theta} \left| \frac{\partial m(\beta^\tau x; \theta)}{\partial \beta^{(1)}} - \frac{\partial \widehat{m}(\beta^\tau x; \theta)}{\partial \beta^{(1)}} - m'(\beta^\tau x; \theta) J^\tau g(\beta^\tau x) \right| \\ &= O_p \left\{ h^2 + \left(\frac{\log n}{nh} \right)^{1/2} \right\}. \end{aligned} \quad (\text{A.84})$$

Consequently, the desired result directly follows from (A.84). $\square$

Lemma A.10. Under the conditions of Theorem 3, if $nh^2 \rightarrow \infty$ and $nh^4 \rightarrow 0$, it holds that

$$\sqrt{n} \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} \Big|_{\theta=\theta^0} = \frac{1}{\sqrt{n}} \sum_{t=1}^n \lambda_t + o_p(1), \quad (\text{A.85})$$

where

$$\begin{aligned} \lambda_t &= -2\pi(x_t) \left[\frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) h_t \right] \\ & \times \frac{f^{-1}(\beta^{0\tau} x_t) [1 + o_p(1)]}{\sum_{k=0}^p \omega_{k0} \sigma_k(\beta^{0\tau} x_t; \beta^0)} \frac{1}{n} \sum_{i=1}^n K_h(\beta^{0\tau} x_i - \beta^{0\tau} x_t) \xi_i \\ & + 2\pi(x_t) \left[\frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) h_t \right] \\ & \times \frac{f^{-1}(\beta^{0\tau} x_t) [1 + o_p(1)]}{\left[\sum_{k=0}^p \omega_{k0} \sigma_k(\beta^{0\tau} x_t; \beta^0) \right]^2} \sigma(\beta^{0\tau} x_t; \beta^0) \\ & \times \frac{1}{n} \sum_{k=0}^p \omega_{k0} \sum_{i=1}^n K_h(\beta^{0\tau} x_i - \beta^{0\tau} x_t) \xi_{i,k} \\ & + \pi(x_t) \left\{ 2\varepsilon_t \left[\frac{1}{h_t} \frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) \right] \right. \\ & \quad \left. + \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} \left(1 - \frac{\varepsilon_t^2}{h_t} \right) \right\}. \end{aligned} \quad (\text{A.86})$$

Proof. Note that

$$\begin{aligned} \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} &= \frac{1}{n} \sum_{t=1}^n \left\{ \frac{1}{h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \left[1 - \frac{\widehat{\varepsilon}_t^2(\theta)}{h_t(\theta)} \right] \right. \\ & \quad \left. + \frac{2\widehat{\varepsilon}_t(\theta)}{h_t(\theta)} \frac{\partial \widehat{\varepsilon}_t(\theta)}{\partial \theta} \right\} \pi(x_t), \end{aligned} \quad (\text{A.87})$$

where

$$\begin{aligned} \widehat{\varepsilon}_t(\theta) &= \varepsilon_t(\theta) + h_t(\theta) [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)], \\ \frac{\partial \widehat{\varepsilon}_t(\theta)}{\partial \theta} &= \frac{\partial \varepsilon_t(\theta)}{\partial \theta} + h_t(\theta) \left[\frac{\partial m(\beta^\tau x_t; \theta)}{\partial \theta} - \frac{\partial \widehat{m}(\beta^\tau x_t; \theta)}{\partial \theta} \right] \\ & + \frac{\partial h_t(\theta)}{\partial \theta} [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)]. \end{aligned} \quad (\text{A.88})$$

It can be easily got that

$$\begin{aligned} & \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} - \frac{\partial L_n(\theta)}{\partial \theta} \\ &= \frac{1}{n} \sum_{t=1}^n \pi(x_t) \left\{ \frac{2}{h_t(\theta)} \left[\widehat{\varepsilon}_t(\theta) \frac{\partial \widehat{\varepsilon}_t(\theta)}{\partial \theta} - \varepsilon_t(\theta) \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \right] \right. \\ & \quad \left. + \frac{1}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} [\varepsilon_t^2(\theta) - \widehat{\varepsilon}_t^2(\theta)] \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{2}{n} \sum_{t=1}^n \pi(x_t) \left\{ \frac{1}{h_t(\theta)} \hat{\varepsilon}_t(\theta) \left[\frac{\partial \hat{\varepsilon}_t(\theta)}{\partial \theta} - \frac{\partial \varepsilon_t(\theta)}{\partial \theta} \right] \right. \\
&\quad \left. - \frac{1}{h_t(\theta)} \frac{\partial \varepsilon_t(\theta)}{\partial \theta} [\hat{\varepsilon}_t(\theta) - \varepsilon_t(\theta)] \right\} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \pi(x_t) \frac{1}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \\
&\quad \times (2\varepsilon_t(\theta) h_t(\theta) [\widehat{m}(\beta^\tau x_t; \theta) - m(\beta^\tau x_t; \theta)] \\
&\quad - h_t^2(\theta) [\widehat{m}(\beta^\tau x_t; \theta) - m(\beta^\tau x_t; \theta)]^2) \\
&= \frac{1}{n} \sum_{t=1}^n \pi(x_t) \left\{ \frac{\partial h_t(\theta)}{\partial \theta} [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)]^2 \right. \\
&\quad \left. + 2 \frac{\partial \varepsilon_t(\theta)}{\partial \theta} [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)] \right\} \\
&\quad + \frac{2}{n} \sum_{t=1}^n \pi(x_t) \\
&\quad \times [\varepsilon_t(\theta) + h_t(\theta) (m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta))] \\
&\quad \times \left[\frac{\partial m(\beta^\tau x_t; \theta)}{\partial \theta} - \frac{\partial \widehat{m}(\beta^\tau x_t; \theta)}{\partial \theta} \right] \\
&= \frac{2}{n} \sum_{t=1}^n \pi(x_t) \{ H(\beta^\tau x_t; \theta) \varepsilon_t(\theta) \\
&\quad + \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta} + H(\beta^\tau x_t; \theta) h_t(\theta) \right] \\
&\quad \times [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)] \} \\
&\quad + \frac{1}{n} \sum_{t=1}^n \pi(x_t) \frac{\partial h_t(\theta)}{\partial \theta} [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)]^2 \\
&\quad + \frac{2}{n} \sum_{t=1}^n \pi(x_t) \varepsilon_t(\theta) \left[\frac{\partial m(\beta^\tau x_t; \theta)}{\partial \theta} \right. \\
&\quad \left. - \frac{\partial \widehat{m}(\beta^\tau x_t; \theta)}{\partial \theta} - H(\beta^\tau x_t; \theta) \right] \\
&\quad + \frac{2}{n} \sum_{t=1}^n \pi(x_t) h_t(\theta) [m(\beta^\tau x_t; \theta) - \widehat{m}(\beta^\tau x_t; \theta)] \\
&\quad \times \left[\frac{\partial m(\beta^\tau x_t; \theta)}{\partial \theta} - \frac{\partial \widehat{m}(\beta^\tau x_t; \theta)}{\partial \theta} - H(\beta^\tau x_t; \theta) \right]. \tag{A.89}
\end{aligned}$$

Invoking Theorem 4, Lemma A.9, and the law of large numbers for $(1/n) \sum_{t=1}^n \pi(x_t) h_t(\theta)$, together with the facts that $n^{-1/2} \sum_{t=1}^n \pi(x_t) \varepsilon_t = O_p(1)$ and $\|(1/h_t(\theta))(\partial h_t(\theta)/\partial \theta)\| < M$, we can prove that the latter three terms of the right-hand side of the last equation are of order $o_p(n^{1/2})$ if $nh^2 \rightarrow \infty$ and $nh^8 \rightarrow 0$ hold.

Thus, we can obtain that

$$\begin{aligned}
&\sqrt{n} \frac{\partial \widehat{L}_n(\theta)}{\partial \theta} \Big|_{\theta=\theta^0} \\
&= \frac{2}{\sqrt{n}} \sum_{t=1}^n \pi(x_t) \left[\frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) h_t \right] \\
&\quad \times [m(\beta^{0\tau} x_t; \theta^0) - \widehat{m}(\beta^{0\tau} x_t; \theta^0)] \\
&\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n \pi(x_t) \left\{ 2\varepsilon_t \left[\frac{1}{h_t} \frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) \right] \right. \\
&\quad \left. + \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} \left(1 - \frac{\varepsilon_t^2}{h_t} \right) \right\} + o_p(1). \tag{A.90}
\end{aligned}$$

To analyze $\partial \widehat{L}_n(\theta)/\partial \theta$, we will show the asymptotic representation for $m(z; \theta) - \widehat{m}(z; \theta)$. By the fact that $(1/n) \sum_{t=1}^n d_t(z; \beta)(\beta^\tau x_t - z) = 0$, applying the similar argument used in Lemma A.7, we can obtain that

$$\begin{aligned}
&\widehat{\sigma}(z; \beta) - \sigma(z; \beta) \\
&= f^{-1}(z) \frac{1}{n} \sum_{i=1}^n K_h(\beta^\tau x_i - z) \xi_i(\beta) [1 + o_p(1)] \\
&\quad + O_p(h^2), \tag{A.91} \\
&\widehat{\sigma}_k(z; \beta) - \sigma_k(z; \beta) \\
&= f^{-1}(z) \frac{1}{n} \sum_{i=1}^n K_h(\beta^\tau x_i - z) \xi_{i,k}(\beta) [1 + o_p(1)] \\
&\quad + O_p(h^2),
\end{aligned}$$

where $\xi_i(\beta) = y_i - \sigma(\beta^\tau x_i; \beta)$, $\xi_{i,k}(\beta) = y_{i-k}^2 - \sigma_k(\beta^\tau x_i; \beta)$, $\xi_i = \xi_i(\beta^0)$ and $\xi_{i,k} = \xi_{i,k}(\beta^0)$.

Recalling the definition of (6) and (7), we have

$$\begin{aligned}
&\widehat{m}(z; \theta) - m(z; \theta) \\
&= \frac{\widehat{\sigma}(z; \beta) - \sigma(z; \beta)}{\sum_{k=0}^p \omega_k \sigma_k(z; \beta)} \\
&\quad - \frac{\sigma(z; \beta) \sum_{k=0}^p \omega_k [\widehat{\sigma}_k(z; \beta) - \sigma_k(z; \beta)]}{\left[\sum_{k=0}^p \omega_k \sigma_k(z; \beta) \right]^2} + o_p(n^{-1/2}) \\
&= \frac{f^{-1}(z) [1 + o_p(1)]}{\sum_{k=0}^p \omega_k \sigma_k(z; \beta)} \frac{1}{n} \sum_{i=1}^n K_h(\beta^\tau x_i - z) \xi_i(\beta)
\end{aligned}$$

$$\begin{aligned}
& - \frac{f^{-1}(z) [1 + o_p(1)]}{\left[\sum_{k=0}^p \omega_k \sigma_k(z; \beta) \right]^2} \\
& \times \frac{\sigma(z; \beta)}{n} \sum_{k=0}^p \omega_k \sum_{i=1}^n K_h(\beta^\tau x_i - z) \xi_{i,k}(\beta) \\
& + O_p(h^2) + o_p(n^{-1/2}).
\end{aligned} \tag{A.92}$$

Finally, we have

$$\begin{aligned}
& \sqrt{n} \frac{\partial \hat{L}_n(\theta)}{\partial \theta} \Big|_{\theta=\theta^0} \\
& = - \frac{2}{\sqrt{n}} \sum_{t=1}^n \pi(x_t) \left[\frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) h_t \right] \\
& \quad \times \frac{f^{-1}(\beta^{0\tau} x_t) [1 + o_p(1)]}{\left[\sum_{k=0}^p \omega_{k0} \sigma_k(\beta^{0\tau} x_t; \beta^0) \right]^2} \\
& \quad \times \frac{1}{n} \sum_{i=1}^n K_h(\beta^{0\tau} x_i - \beta^{0\tau} x_t) \xi_i \\
& \quad + \frac{2}{\sqrt{n}} \sum_{t=1}^n \pi(x_t) \left[\frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) h_t \right] \\
& \quad \times \frac{f^{-1}(\beta^{0\tau} x_t) [1 + o_p(1)]}{\left[\sum_{k=0}^p \omega_{k0} \sigma_k(\beta^{0\tau} x_t; \beta^0) \right]^2} \sigma(\beta^{0\tau} x_t; \beta^0) \\
& \quad \times \frac{1}{n} \sum_{k=0}^p \omega_{k0} \sum_{i=1}^n K_h(\beta^{0\tau} x_i - \beta^{0\tau} x_t) \xi_{i,k} \\
& \quad + \frac{2}{\sqrt{n}} \sum_{t=1}^n \pi(x_t) \left[\frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) h_t \right] \\
& \quad \times [O_p(h^2) + o_p(n^{-1/2})] \\
& \quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n \pi(x_t) \left\{ 2\varepsilon_t \left[\frac{1}{h_t} \frac{\partial \varepsilon_t}{\partial \theta} + H(\beta^{0\tau} x_t; \theta^0) \right] \right. \\
& \quad \left. + \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} \left(1 - \frac{\varepsilon_t^2}{h_t} \right) \right\} + o_p(1) \\
& = \frac{1}{\sqrt{n}} \sum_{t=1}^n \lambda_t + o_p(1),
\end{aligned} \tag{A.93}$$

where λ_t is given in (A.86). The third term above is of order $o_p(1)$ by applying the law of large numbers to $(1/n) \sum_{t=1}^n \pi(x_t) [\partial \varepsilon_t / \partial \theta + H(\beta^{0\tau} x_t; \theta^0) h_t]$ and $nh^4 \rightarrow 0$. $\square$

Lemma A.11. Under the conditions of Theorem 3, it holds that

$$- \frac{\partial^2 \hat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \Big|_{\theta=\bar{\theta}} = \Sigma_I + o_p(1), \tag{A.94}$$

where Σ_I is given in (A.101).

Proof. Note that

$$\begin{aligned}
& \frac{\partial^2 \hat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \\
& = \frac{1}{n} \sum_{t=1}^n \pi(x_t) \\
& \quad \times \left\{ \left[\frac{2\hat{\varepsilon}_t^2(\theta)}{h_t(\theta)} - 1 \right] \frac{1}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta^\tau} \right. \\
& \quad - \frac{2\hat{\varepsilon}_t(\theta)}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial \hat{\varepsilon}_t(\theta)}{\partial \theta^\tau} \\
& \quad + \frac{2}{h_t(\theta)} \frac{\partial \hat{\varepsilon}_t(\theta)}{\partial \theta} \frac{\partial \hat{\varepsilon}_t(\theta)}{\partial \theta^\tau} + \frac{2\hat{\varepsilon}_t(\theta)}{h_t(\theta)} \frac{\partial^2 \hat{\varepsilon}_t(\theta)}{\partial \theta \partial \theta^\tau} \\
& \quad \left. - \frac{2\hat{\varepsilon}_t(\theta)}{h_t^2(\theta)} \frac{\partial \hat{\varepsilon}_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta^\tau} \right\} \\
& = \frac{1}{n} \sum_{t=1}^n \pi(x_t) \\
& \quad \times \left\{ \left[\frac{2\varepsilon_t^2(\theta)}{h_t(\theta)} - 1 \right] \frac{1}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta^\tau} \right. \\
& \quad - \frac{2\varepsilon_t(\theta)}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \\
& \quad \times \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta^\tau} + h_t(\theta) H^\tau(\beta^\tau x_t; \theta) \right] \\
& \quad + \frac{2}{h_t(\theta)} \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta} + h_t(\theta) H(\beta^\tau x_t; \theta) \right] \\
& \quad \times \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta^\tau} + h_t(\theta) H^\tau(\beta^\tau x_t; \theta) \right] \\
& \quad + \frac{2\varepsilon_t(\theta)}{h_t(\theta)} \left[\frac{\partial^2 \varepsilon_t(\theta)}{\partial \theta \partial \theta^\tau} \right. \\
& \quad \left. + h_t(\theta) \left(\frac{\partial^2 m(\beta^\tau x_t; \theta)}{\partial \theta \partial \theta^\tau} \right. \right. \\
& \quad \left. \left. - \frac{\partial^2 \hat{m}(\beta^\tau x_t; \theta)}{\partial \theta \partial \theta^\tau} \right) \right] \\
& \quad - \frac{2\varepsilon_t(\theta)}{h_t^2(\theta)} \\
& \quad \times \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta} + h_t(\theta) H(\beta^\tau x_t; \theta) \right] \frac{\partial h_t(\theta)}{\partial \theta^\tau} \Big\} \\
& \quad + o_p(1) \\
& = \frac{1}{n} \sum_{t=1}^n D_t(\theta) + o_p(1),
\end{aligned} \tag{A.95}$$

where

$$D_t(\theta) = \frac{1}{n} \sum_{i=1}^n \pi(x_t) \times \left\{ \left[\frac{2\varepsilon_t^2(\theta)}{h_t(\theta)} - 1 \right] \frac{1}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta^\tau} + \frac{2}{h_t(\theta)} \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta} + h_t(\theta) H(\beta^\tau x_t; \theta) \right] \times \left[\frac{\partial \varepsilon_t(\theta)}{\partial \theta^\tau} + h_t(\theta) H^\tau(\beta^\tau x_t; \theta) \right] \right\}. \quad (\text{A.96})$$

Thus, we have

$$\left. \frac{\partial^2 \hat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \right|_{\theta=\tilde{\theta}} = \frac{1}{n} \sum_{i=1}^n D_t(\tilde{\theta}) + o_p(1). \quad (\text{A.97})$$

Based on the fact that $\tilde{\theta}$ converges in probability to θ^0 , we can obtain that

$$\frac{1}{n} \sum_{i=1}^n D_t(\tilde{\theta}) = \frac{1}{n} \sum_{i=1}^n D_t(\theta^0) + o_p(1). \quad (\text{A.98})$$

Then, by the law of large numbers, we have

$$\frac{1}{n} \sum_{i=1}^n D_t(\theta^0) = E[D_t(\theta^0)] + o_p(1). \quad (\text{A.99})$$

Therefore, we get

$$-\left. \frac{\partial^2 \hat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \right|_{\theta=\tilde{\theta}} = \Sigma_I + o_p(1), \quad (\text{A.100})$$

where

$$\Sigma_I = -E \left\{ \pi(x_t) \left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta^\tau} + \frac{2}{h_t} \left[\frac{\partial \varepsilon_t}{\partial \theta} + h_t H(\beta^{0\tau} x_t; \theta^0) \right] \times \left[\frac{\partial \varepsilon_t}{\partial \theta^\tau} + h_t H^\tau(\beta^{0\tau} x_t; \theta^0) \right] \right) \right\}. \quad (\text{A.101})$$

□

Proof of Theorem 6. We have shown that

$$\sqrt{n}(\hat{\theta}_n - \theta^0) = \left\{ -\left. \frac{\partial^2 \hat{L}_n(\theta)}{\partial \theta \partial \theta^\tau} \right|_{\theta=\tilde{\theta}} \right\}^{-1} \left\{ \sqrt{n} \left. \frac{\partial \hat{L}_n(\theta)}{\partial \theta} \right|_{\theta=\theta^0} \right\}. \quad (\text{A.102})$$

According to Lemmas A.10 and A.11, we can obtain the following asymptotic expansion:

$$\sqrt{n}(\hat{\theta}_n - \theta^0) = \Sigma_I^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \lambda_i \right) + o_p(1). \quad (\text{A.103})$$

To analyse the asymptotic normality of $\sqrt{n}(\hat{\theta}_n - \theta^0)$, we only need to show the asymptotic normality of $(1/\sqrt{n}) \sum_{i=1}^n \lambda_i$.

Recalling the definition of ξ_i and $\xi_{i,k}$, hence we have $E(\xi_i) = 0$ and $E(\xi_{i,k}) = 0$, together with the assumptions of $\{e_t\}$, which imply $E(\lambda_t) = 0$.

We have to verify that for some $r > 2$, $E|\lambda_t|^r < \infty$, which can be obtained with condition (A.20). Then, by the central limit theorem for strongly mixing sequences (see, e.g., Bosq [13], Theorem 1.7), we can show that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \lambda_i \xrightarrow{d} N(0, \Sigma), \quad (\text{A.104})$$

where

$$\Sigma = \sum_{i=1}^n \text{Cov}(\lambda_i, \lambda_i). \quad (\text{A.105})$$

Finally, by the Slutsky theorem, we have $\sqrt{n}(\hat{\theta}_n - \theta^0) \xrightarrow{d} N(0, \Sigma_I^{-1} \Sigma \Sigma_I^{-1})$. □

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Yuan Li's work is financially supported by National Natural Science Foundation of China (Grant no. 11271095) and the Specialized Research Fund for the Doctoral Program of Higher Education (Grant no. 20124410110002).

References

- [1] R. F. Engle, D. M. Lilien, and R. P. Robins, "Estimating time varying risk premia in the term structure: the ARCH-M model," *Econometrica*, vol. 55, pp. 391–407, 1987.
- [2] S. Das and N. Sarkar, "Is the relative risk aversion parameter constant over time? A multi-country study," *Empirical Economics*, vol. 38, no. 3, pp. 605–617, 2010.
- [3] R. Chou, R. F. Engle, and A. Kane, "Measuring risk aversion from excess returns on a stock index," *Journal of Econometrics*, vol. 52, no. 1-2, pp. 201–224, 1992.
- [4] X. F. Zhang, *Some parametric and semiparametric models for financial time series analysis [Ph.D. thesis]*, The Hong Kong Polytechnic University, 2011.
- [5] Y. Xia and W. K. Li, "On single-index coefficient regression models," *Journal of the American Statistical Association*, vol. 94, no. 448, pp. 1275–1285, 1999.
- [6] L. Xue and Z. Pang, "Statistical inference for a single-index varying-coefficient model," *Statistics and Computing*, vol. 23, no. 5, pp. 589–599, 2013.
- [7] S. Ling, "A double AR(p) model: structure and estimation," *Statistica Sinica*, vol. 17, no. 1, pp. 161–175, 2007.
- [8] B. J. Christensen, C. M. Dahl, and E. M. Iglesias, "Semiparametric inference in a GARCH-in-mean model," *Journal of Econometrics*, vol. 167, no. 2, pp. 458–472, 2012.

- [9] Y. Yu and D. Ruppert, "Penalized spline estimation for partially linear single-index models," *Journal of the American Statistical Association*, vol. 97, no. 460, pp. 1042–1054, 2002.
- [10] L. Yang, "A semiparametric GARCH model for foreign exchange volatility," *Journal of Econometrics*, vol. 130, no. 2, pp. 365–384, 2006.
- [11] D. W. K. Andrews, "Generic uniform convergence," *Econometric Theory*, vol. 8, no. 2, pp. 241–257, 1992.
- [12] M. R. Kosorok, *Introduction to Empirical Processes and Semiparametric Inference*, Springer, New York, NY, USA, 2006.
- [13] D. Bosq, *Nonparametric Statistics for Stochastic Processes*, Springer, New York, NY, USA, 1996.