Unfolding the Stroke - Ambulatory Blood Pressure Association: A Functional Framework

Fei Jiang
Department of Statistics & Actuarial Science
University of Hong Kong
Jie Xu
Beijing Tiantan Hospital, Capital Medical University
Hui Zhi
Biostatistics and Clinical Methodology Research Center
University of Hong Kong
Yongjun Wang
Beijing Tiantan Hospital, Capital Medical University
Haipeng Shen
Innovation and Information Management
Faculty of Business and Economics, University of Hong Kong

Abstract

Stroke is one of the leading causes of death across the world, and creates serious healthcare and economic burdens. Hypertension is the most important risk factor for stroke; as such it is of great clinical interests to have effective ways of monitoring blood pressure. We use the BOSS study to investigate the controversial relationship between stroke-related clinical outcomes and post stroke blood pressure, in particular ambulatory blood pressure trajectories. We first show that the conventional dipping pattern analysis of ambulatory blood pressure trajectories fail to yield clinically meaningful patient clusters. We then propose a functional analysis pipeline to analyze the ambulatory blood pressure trajectories more effectively. We apply a functional clustering approach to obtain patient clusters that differ in the risk of the clinical outcomes. Each cluster exhibits distinct blood pressure trajectories that offer clinical insights to explain the risk difference. We then adopt a functional logistic regression model to study the time-varying effect of the blood pressure trajectory on the risk of the clinical outcomes. A Markov imputation method is applied to impute missing blood pressure measurements. Our analysis pipeline not only establishes the significant association between the ambulatory blood pressure trajectories and the stroke risk, but also characterizes the time dependent structure of the blood pressure effect. The results provide empirical support to advocate the use of the entire ambulatory blood pressure trajectory as a tool to monitor the risk for stroke-related clinical outcomes.

Keywords: Ambulatory Blood Pressure; BOSS; B-spline; Functional Clustering; Functional Logistic Regression; Markov Imputation.

1 Introduction

Stroke is a "chronic disease with acute events" that causes long-term healthcare issues and enormous economic cost. It is the second-leading cause of death in the world and the leading cause of death in China (Yang et al., 2013; Vos et al., 2015; Feigin et al., 2015; Zhou et al., 2016). According to the Internet Stroke Center, every year, 15 million people suffer from stroke; 5 million of these people die, another 5 million people suffer from long-term disability.

Hypertension is the primary risk factor for stroke (O'Donnell et al., 2010), and accounts for about 54% of strokes worldwide. As such, healthcare professional associations pay special attention to the criterion for monitoring high blood pressure, and recently updated it from 140/90 to 130/80 (Whelton et al., 2017). Besides traditional office (one-time) measurements, American, European, and Japanese guidelines (James et al., 2014; Mancia et al., 2013; Ogihara et al., 2009) have confirmed the importance of ambulatory blood pressure monitoring, which measures the blood pressure trajectory continuously across time, for example, over a period of 24 hours or multiple days. Portable or wearable blood pressure monitors have made such measurements feasible and convenient.

So, a question of clinical interest is – besides the 130/80 criterion for hypertension, *does it also make sense to look at the entire ambulatory blood pressure trajectory*?

We think the answer is positive, as we shall show below via analysis of the BOSS study, which produced the first large blood pressure database for transient ischemic attack (TIA) or ischemic stroke (IS) patients in China. As a nation wide hospital-based, prospective cohort study, its goal is to investigate association between blood pressure (both one-time and ambulatory) and clinical outcomes in stroke patients (Xu et al., 2016). Besides blood pressure measurements, the study also collects patient baseline information and three clinical outcomes, including stroke recurrence, combined vascular events, and disability. See Xu et al. (2016) and Section 2 for more details of the study and the data collected.

Ambulatory blood pressure trajectories have been analyzed in the medical literature via "the dipping analysis" (Metoki et al., 2006). The idea behind the analysis is rather simple: it calculates the nocturnal decline in blood pressure using the difference between *average* daytime pressure and *average* nighttime pressure. Intuitively, the dipping pattern is an important feature to differentiate blood pressure curves. Studies have been conducted

to evaluate the association between the dipping patterns and stroke clinical outcomes, but the results remain controversial (Kario et al., 2001; Sargento-Freitas et al., 2015). In the BOSS study, the patients were classified to four groups - "Dipper", "Extreme Dipper", "Non Dipper", "Reverse Dipper," as suggested by the American Heart Association. Unfortunately our analysis in Section 2 shows that none of the clinical outcomes are significantly associated with the dipping patterns.

Although simple and intuitive, the dipping method only uses the average systolic blood pressure information during the sleeping period and awake period. As we will show in Section 2, patients in the same dipping pattern group can have dramatically different ambulatory blood pressure trajectories, which means that the dipping classification loses substantial information in the trajectories. This motivates us to use a more sophisticated functional clustering method to cluster the patients using their entire trajectories.

We view the ambulatory trajectories as functional data. Traditional clustering tools, such as k-means (Ball and Hall, 1967) and hierarchical clustering (Hartigan and Wong, 1979), are not directly applicable. To handle functional data, one often projects the data onto a subspace, and then uses multivariate clustering methods such as Abraham et al. (2003); Garcia-Escudero and Gordaliza (2005); Serban and Wasserman (2005). These methods focus on grouping the random curves based on their first moment properties.

We propose to use the k-centres functional clustering method of Chiou and Li (2007) to cluster the trajectories. The method accounts for both the means and the modes of variation differentials among clusters. It greatly improves the clustering accuracy by concurrently updating the mean/covariance structures, and the cluster membership. Furthermore, it can directly incorporate irregular spaced functional data, which is desirable for clustering the BOSS blood pressure curves. The clustering results (Section 3.2) are clinically meaningful, and the cluster memberships are highly associated with the clinical outcomes.

After the clustering, we perform a functional logistic regression analysis to study the time-varying effect of ambulatory blood pressure on the risk of the clinical outcomes. One major obstacle is that the blood pressure curves are not fully observed, while the missing measurements are distributed randomly along the trajectories. To amend this issue, we first apply the method discussed in Delaigle and Hall (2016) to impute the missing measurements. Then, we fit a functional logistic regression model, incorporating baseline covariates.

To make inference, we construct a point-wise confidence band for the time-varying blood pressure effect function. The results (Section 4.2) identify time periods during which the ambulatory blood pressure has a significant effect on the clinical outcomes.

The rest of the article is organized as follows. In Section 2, we describe the motivating BOSS study and perform the traditional dipping pattern clustering analysis. Section 3, we introduce the functional clustering procedure and report the more meaningful clustering results. In Section 4, we provide technical details of the functional logistic regression analysis and the associated results. We conclude the article in Section 5.

2 The BOSS Study and the Dipping Pattern Analysis

2.1 The BOSS Study

The relationship between post stroke blood pressure and clinical outcomes in ischemic stroke remains controversial. Xu et al. (2016) report the BOSS study, a nation wide hospital-based, prospective cohort study on the blood pressure and clinical outcome in transient ischemic attack (TIA) or ischemic stroke (IS). The study aims at assessing associations between blood pressure and clinical outcomes among such patients. Clinical outcomes include stroke recurrence, combined vascular events, and disability.

A total of 2,608 IS/TIA patients were consecutively enrolled in 61 hospitals from October 2012 to February 2014. The patients were followed up for clinical outcomes at 3 months through face-to-face interview and at 12 months by telephone. In this paper, our neurologist collaborators are interested in studying the association between a patient's ambulatory systolic blood pressure (SBP) trajectory and her/his clinical outcomes. If such association exists, they can potentially identify high risk patients timely through continuously monitoring the blood pressure. During a patient's hospitalization, the ambulatory blood pressure measurements were taken every 15 minutes during the day and every 30 minutes at night. Daytime episodes are 6:00 am to 9:59 pm and night episodes are from 10 pm to 5:59 am.

The times of the measurements are irregularly distributed across the patients. Some measurements are missing; hence the patients' blood pressure trajectories are not fully observed. After excluding patients with severe missing, we include the remaining 1,996 patients in our analysis.

2.2 The Dipping Pattern Analysis

The BOSS study used the dipping pattern idea (Metoki et al., 2006) to cluster the ambulatory blood pressure trajectories into four different groups: "Dipper", "Extreme Dipper", "Non Dipper", "Reverse Dipper." The clustering algorithm works as follows. First, the percentage decline in nocturnal SBP is calculated:

$$\text{nocturnal decline in SBP} = \frac{\text{average daytime SBP} - \text{average nighttime SBP}}{\text{average daytime SBP}} \times 100.$$

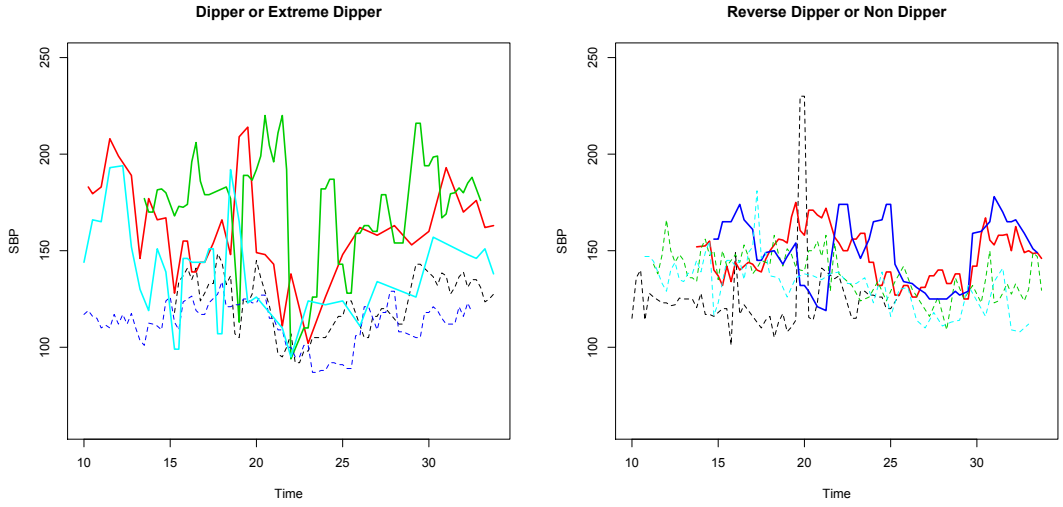
Then, the dipping groups are defined as "Extreme Dipper" if nocturnal decline in SBP $\geq 20\%$; "Dipper" is the decline is within 10% to 19%; "Non Dipper" if the decline is within 0% to 9%; and "Reverse Dipper" if the decline is negative. The cutoff values are based on earlier medical studies.

As defined above, the "Dipper" and "Extreme Dipper" groups are expected to experience normal blood pressure decrease at night and increase in the morning; on the other hand, the "Non Dipper" and "Reverse Dipper" groups are expected to have persistently high or early rising blood pressure. As such, the "Dipper" and "Extreme Dipper" groups should have lower risk for the clinical outcomes, while the "Non Dipper" and "Reverse Dipper" groups will have higher risk.

To examine the performance of the dipper pattern clustering, we randomly show a few blood pressure curves in Figure 1. It can be seen that the "Extreme Dipper" or "Dipper" groups actually contain patients (highlighted in bold) whose blood pressures rise sharply in the morning (which have been shown to associate with high risk for stroke), while the "Non Dipper" or "Reverse Dipper" groups indeed contain patients with normal blood pressure patterns (again highlighted in bold). This pictorially suggests that the dipper pattern is not precise to cluster the patients into different risk groups.

Furthermore, in the BOSS study, the primary endpoints are the binary indicators for stroke recurrence, combined vascular events, and disability. We perform the Chi-square test to examine the association between the three endpoints and the dipping pattern clusters, separately. In contrast to the clinical experience, the dipping patterns are not significantly associated with any of the endpoints, with the p-values to be 0.42, 0.165, and 0.087. This motivates us to use a more sophisticated functional clustering method to group the blood pressure trajectories below in Section 3.

Figure 1: Sample Trajectories from the Four Dipper Groups. The solid lines in the left panel represent trajectories with sharp increasing in the morning. The solid lines in the right panel represent trajectories with normal patterns.



3 Clustering of the Blood Pressure Curves

3.1 Functional Clustering

For the BOSS study, we focus on the 24-hour ambulatory SBP trajectory that starts from 10 am to 10 am of the next day. To align the SBP measurements across different patients, we construct a 15-minute spaced time grid $t_1, \ldots, t_T$, where the 15-minute gap is chosen because the SBP measurements are taken roughly every 15 minutes during the day (Xu et al., 2016). For time t_j, $j = 1, \ldots, T$, we use the average of the SBP measurements within the jth grid cell, denoted as $X_i(t_j)$, to represent the blood pressure level at t_j.

Because of the irregular measurement times, some patients may have missing measurements in some intervals. To indicate the missing measurement, we define $R_{ij} = 1$ if patient i has blood pressure measurement at t_j; otherwise, $R_{ij} = 0$. Note that such missing is not disease related, so that the missing at random assumption is appropriate in our study.

Suppose there are M clusters, and let C_i denote the cluster index of patient i. We model the blood pressure trajectory through the following functional model:

$$X_i(t) = \sum_{c=1}^{M} I(C_i = c)\mu_c(t) + f_i(t) + \epsilon_i, \tag{1}$$

where $\mu_c(t) = E\{X_i(t)|C_i = c\}$ is the mean function of the c-th cluster, $c = 1, \ldots, M$, and

ϵ_i is mean 0 random error, and f_i is an unknown function that describes the within cluster fluctuation for the i-th patient. Furthermore, we approximate $f_i(t)$ via basis expansion, in the form of $\mathbf{P}^{\mathrm{T}}(t)\boldsymbol{\gamma}_i$, where $\mathbf{P}(t)$ represents a set of d-dimensional basis function.

Let $\mathbf{C} = (C_1, \ldots, C_n)^{\mathrm{T}}$ denote the cluster memberships. If they are known, we can estimate the model parameters $\boldsymbol{\gamma}_i$ and $\mu_c(t)$ by minimizing the following sum of squared error (SSE):

$$\mathrm{SSE}(\mathbf{C}, \mu_c, \boldsymbol{\gamma}) = \sum_{i=1}^{n} \sum_{j=1}^{T} R_{ij} I(C_i = c)\{X_i(t_j) - \mu_c(t_j) - \mathbf{P}^{\mathrm{T}}(t_j)\boldsymbol{\gamma}_i\}^2,$$

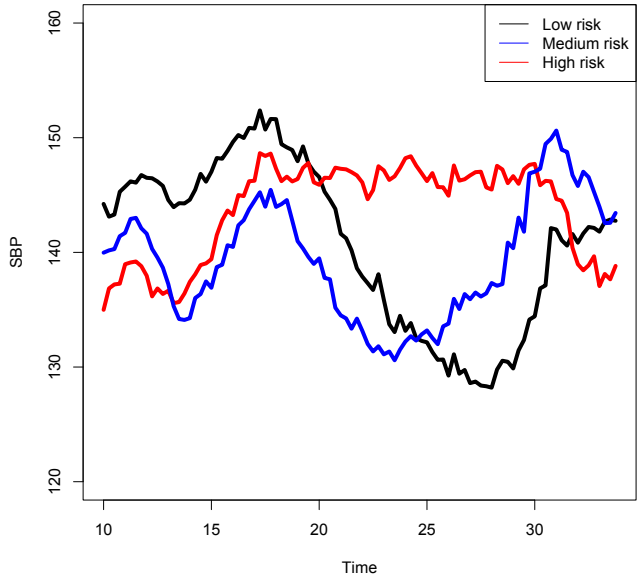
where $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^{\mathrm{T}}, \ldots, \boldsymbol{\gamma}_n^{\mathrm{T}})^{\mathrm{T}}$. When the cluster membership is unknown, we implement the iterative functional clustering method proposed by Chiou and Li (2007). Essentially, we iteratively minimize $\mathrm{SSE}(\mathbf{C}, \mu_c, \boldsymbol{\gamma})$, alternating with respect to the cluster membership $\mathbf{C}$, and the model parameters $\mu_c(t)$ and $\boldsymbol{\gamma}$. It is worth mentioning that this functional clustering method does not require regularly distributed measurement times as it automatically handles cases with missing measurements. A R implementation of the clustering algorithm is available at *funcy* package (Yassouridis, 2017).

3.2 Ambulatory Blood Pressure Clustering Analysis

We implement the functional clustering method discussed in Section 3.1 to cluster the BOSS blood pressure trajectories. Following the neurologists suggestion, we consider clustering the SBP curves into $M = 3$ clusters. We choose five Fourier bases for the basis expansion. Figure 2 plots the mean SBP trajectory for the three clusters. According to the neurologists, the three curves nicely distinguish the High, Medium, and Low risk groups, and matches well with their clinical experience, which we shall discuss below.

Comparing the mean trajectories of the three clusters, Figure 2 shows that the SBP curves behave dramatically different in the period from Hour 22 (10pm) to Hour 30 (6am the next day). For the Low Risk cluster, the mean trajectory exhibits nicely the normal diurnal blood pressure pattern. For the High Risk cluster, the mean trajectory stays consistently high overnight. For the Medium Risk cluster, the mean trajectory also drops at night, but rises earlier and higher prior to 6am, in comparison with the mean of the Low Risk cluster. This pattern associates nicely with the morning blood pressure surge phenomenon, which has brought attentions in the stroke literature (Bursztyn, 2003; Pierdomenico et al., 2014).

Figure 2: The Cluster Mean SBP Trajectories for Low, Medium, and High Risk Groups.



We compute the event rate for the three clinical outcomes for each trajectory cluster in Table 1. The p-values for the pairwise comparison show that the event rate in the High Risk group is significantly larger than that in the Low Risk group for all the three clinical outcomes. The difference between the Medium Risk group and the High Risk group is only significant for the disability endpoint.

Table 1: Comparison of Clinical Outcome Rate among the Three Risk Groups.

	Stroke Recurrence	Combined Vascular Events	Disability	Cluster Size
High	0.0661	0.0828	0.0944	604
Medium	0.0524	0.0720	0.0638	611
Low	0.0371	0.0512	0.0538	781
p-value High - Low	**0.0171**	**0.0216**	**0.0049**	
p-value Mid - Low	0.1770	0.1129	0.4315	
p-value High - Mid	0.3073	0.4830	**0.0488**	

Furthermore, we compare the functional clustering results with the dipping pattern clustering reported in Section 2 . We tabulate the number of patients in each cluster based on the two clustering methods in Table 2. We observe some association among the cluster memberships, but clearly the two methods give different clusters. To certain extent, this can explain why the dipping groups are not associated with clinical outcomes while the

trajectory clustering groups are.

Table 2: The dipping patterns for the three clusters on the SBP.

	Low	Medium	High
Dipper	277	84	5
Extreme Dipper	28	2	1
Non-Dipper	441	401	186
Reverse Dipper	25	112	410

4 Functional Logistic Regression

The functional clustering analysis in Section 3 suggests that the ambulatory SBP trajectory does contain information that is associated with risk level of clinical outcomes of interest. For example, the mean ambulatory trajectories of the three risk groups (Figure 2) show that the blood pressure impacts on the outcome risk are time dependent: higher from 10pm to 6am of the next day than the other time periods. In this section, we present a functional logistic regression model to further examine how the ambulatory blood pressure trajectory affects the risk of a clinical outcome. In addition, we incorporate additional baseline covariates to eliminate potential confounding effects.

4.1 Model and Estimation

4.1.1 Model

For the ith patient, let Y_i denote the binary indicator of the clinical outcome of interest, and $\mathbf{W}_i$ be the d-dimensional baseline covariates. We study the effect of the ambulatory blood pressure trajectory $X_i(t)$ on the clinical outcome through the following functional logistic regression model:

$$E\{Y_i|X_i(t), \mathbf{W}_i; \boldsymbol{\beta}, \alpha\} = h\left\{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{W}_i + \int \alpha(t)X_i(t)dt\right\}, \tag{2}$$

where h is the logit link function in the form of $h(x) = \exp(x)/\{1+\exp(x)\}$, $\boldsymbol{\beta}$ is the baseline coefficient vector, and $\alpha(t)$ is the blood pressure effect function. Here, the integral is over

the compact support of the measurement times.

4.1.2 Trajectory Imputation

Note that the trajectory $X_i(t)$ may contain missing measurements; hence before model estimation, we first use the imputation method of Delaigle and Hall (2016) to impute the missing values. The imputation incorporates the clustering memberships obtained in Section 3. We start with assuming $X_i(t)$ as a Markov process so that the distribution of $X_i(t)$ depends only on the closest observed values in its blood pressure measurement sequence.

To begin with, we divide the blood pressure measurements to m states, denoted by $z_s, s = 1, \ldots, m$, at their equally spaced quantiles. Furthermore, let $Z(t_j) = z_k$ if $(z_{k-1} + z_k)/2 < X(t_j) \leq (z_k + z_{k+1})/2$, $j = 1, \ldots, T$, which are the transformed observations used in the imputation procedure.

For the cluster c, we denote the transition probability as $\text{pr}_c\{Z(t_s) = z_s | Z(t_v) = z_v\}$ and let E_c denote the corresponding expectation. Furthermore, let $(\tau_{i1}, \ldots, \tau_{in_i}) \subset (t_1, \ldots, t_T)$ denote the observed time points for the ith patient, $n_i = \sum_{j=1}^{T} R_{ij}$, whose cluster membership is C_i.

There can be three types of missing measurements: missing before τ_{i1}, after τ_{in_i}, and between τ_{il} and $\tau_{i(l+1)}$, $l = 1, \ldots, n_i - 1$. We define the imputed trajectory $\nu_i(t)$ as follows:

$$\nu_i(t) = \begin{cases} Z_i(t), & t \in \{\tau_{ik}, k = 1, \ldots, n_i\}; \\ E_{C_i}\{Z_i(t) | Z_i(\tau_{i1})\}, & t_1 \leq t < \tau_{i1}; \\ E_{C_i}\{Z_i(t) | Z_i(\tau_{in_i})\}, & \tau_{in_i} \leq t < t_l; \\ E_{C_i}\{Z_i(t) | Z_i(\tau_{il}), Z_i(\tau_{i(l+1)})\}, & \tau_{il} \leq t < \tau_{i(l+1)}. \end{cases} \tag{3}$$

For the first case where $t_j < \tau_{i1}$, assuming $Z(\tau_{i1}) = z_s$, then

$$E_{C_i}\{Z_i(t_j) | Z_i(\tau_{i1})\} = \sum_{v=1}^{m} z_v \text{pr}_{C_i}\{Z_i(t_j) = z_v | Z_i(\tau_{i1}) = z_s\}. \tag{4}$$

Here because $t_j < \tau_{i1}$, $\text{pr}_c\{Z_i(t_j) = z_v | Z_i(\tau_{i1}) = z_s\}$ is a backward transition probability. Similarly, for the second case where $t_j > \tau_{in_i}$, assuming $Z_i(\tau_{in_i}) = z_s$, then

$$E_{C_i}\{Z_i(t) | Z_i(\tau_{in_i})\} = \sum_{v=1}^{m} z_v \text{pr}_{C_i}\{Z_i(t_j) = z_v | Z_i(\tau_{in_i}) = z_s\}. \tag{5}$$

Because $t_j > \tau_{in_i}$, $\mathrm{pr}_{C_i}\{Z_i(t_j) = z_v | Z_i(\tau_{in_i}) = z_s\}$ is a forward transition probability. Finally, if $\tau_{il} < t < \tau_{i(l+1)}$, assuming $Z_i(\tau_{il}) = z_{s1}$ and $Z_i(\tau_{il}) = z_{s2}$, then

$$\begin{aligned} & E_{C_i}\{Z_i(t)|Z_i(\tau_{il}), Z_i(\tau_{i(l+1)})\} \\ = & \frac{\sum_{v=1}^m z_v \mathrm{pr}_{C_i}\{Z_i(t_j) = z_v | Z_i(\tau_{il}) = z_{s_1}\} \mathrm{pr}_{C_i}\{Z_i(\tau_{i(l+1)}) = z_{s_2} | Z_i(t_j) = z_v\}}{\sum_{v=1}^m \mathrm{pr}_{C_i}\{Z_i(t_j) = z_v | Z_i(\tau_{il}) = z_{s_1}\} \mathrm{pr}_{C_i}\{Z_i(\tau_{i(l+1)}) = z_{s_2} | Z_i(t_j) = z_v\}}. \end{aligned} \quad (6)$$

The forward and backward transition probabilities can be estimated by the corresponding entries in the transition matrix, which are the products of one-stage forward transition probability $\mathrm{pr}_c\{Z(t_{j+1}) = z_v | Z(t_j) = z_s\}$ and one-stage backward transition probability $\mathrm{pr}_c\{Z(t_j) = z_v | Z(t_{j+1}) = z_s\}$. Furthermore, the transition probabilities can be estimated, respectively, through

$$\widehat{\mathrm{pr}}_c\{Z(t_{j+1}) = z_v | Z(t_j) = z_s\} = \frac{\sum_{i=1}^n I(C_i = c, Z_i(t_{j+1}) = z_v, Z_i(t_j) = z_s)}{\sum_{v=1}^m \sum_{i=1}^n I(C_i = c, Z_i(t_{j+1}) = z_v, Z_i(t_j) = z_s)},$$

and

$$\widehat{\mathrm{pr}}_c\{Z(t_j) = z_v | Z(t_{j+1}) = z_s\} = \frac{\sum_{i=1}^n I(C_i = c, Z_i(t_j) = z_v, Z_i(t_{j+1}) = z_s)}{\sum_{v=1}^m \sum_{i=1}^n I(C_i = c, Z_i(t_j) = z_v, Z_i(t_{j+1}) = z_s)}.$$

These estimators can then be plugged into (3)–(6) to obtain the imputed trajectory $\widehat{\nu}_i(t_j)$.

4.1.3 Parameter Estimation

We further approximate $\alpha(t)$ using the B-spline expansion in the form of $\mathbf{B}^{\mathrm{T}}(t)\boldsymbol{\eta}$ with order r and N interior knots. The resulting working model with $\widehat{\nu}_i(t)$ in place of $X_i(t)$ is

$$E\{Y_i | \widehat{\nu}_i, \mathbf{W}_i, \boldsymbol{\beta}, \boldsymbol{\eta}\} = h\left\{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}^{\mathrm{T}}(t)\widehat{\nu}_i(t)dt\ \boldsymbol{\eta}\right\}.$$

Then the resulting score function is

$$\mathbf{S}_i(\boldsymbol{\beta}, \boldsymbol{\eta}) = \left[\mathbf{W}_i^{\mathrm{T}}, \left\{\int \mathbf{B}^{\mathrm{T}}(t)\widehat{\nu}_i(t)dt\right\}^{\mathrm{T}}\right] \left[Y_i - \boldsymbol{\beta}^{\mathrm{T}}\mathbf{W}_i - h\left\{\int \mathbf{B}^{\mathrm{T}}(t)\widehat{\nu}_i(t)dt\ \boldsymbol{\eta}\right\}\right].$$

Treating $\int \mathbf{B}^{\mathrm{T}}(t)\widehat{\nu}_i(t)dt$ as the new covariate, the problem reduces to the usual parameter estimation in a logistic regression model. We obtain the estimators for $\boldsymbol{\eta}$, denoted by $\widehat{\boldsymbol{\eta}}$, by solving $\sum_{i=1}^n \mathbf{S}_i(\boldsymbol{\beta}, \boldsymbol{\eta}) = \mathbf{0}$. The asymptotic convergence and the estimation variances are provided below in Theorem 1, where for simplicity, we consider Z_i as the realization. Although the transformation from X_i to Z_i may induce additional estimation error, the error can be controlled by choosing smaller grid size in the state space.

Theorem 1. *Assume Conditions (A1) – (A3) hold. Let* $\boldsymbol{\beta}_0$ *and* α_0 *be the true values for* $\boldsymbol{\beta}$ *and* α*, and* $h_b = O(N^{-1})$ *be the distance between knots. Then*

$$\sqrt{n}[(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^{\mathrm{T}}, h_b^{1/2}\{\mathbf{B}^{\mathrm{T}}(t)\widehat{\boldsymbol{\eta}} - \alpha_0(t)\}]^{\mathrm{T}}$$

converges to a normal random vector with mean $\mathbf{0}$ *and variance* $\boldsymbol{\Sigma}$ *defined in (7) of Appendix.*

Practically, to achieve the estimation accuracy, we require (1) the grid in the state space is dense enough so that $\sup_t |Z_i(t) - X_i(t)|$ converges to 0: we chose the grid in the state space based on the empirical quantiles of the blood pressure levels so that there are roughly 5% observations in each grid cell; (2) the grid in the time space is dense enough so that the summations in (4)–(6) approximate expectations accurately: we chose 15 minutes spaced time grid, which is the minimal gap between the measurement times. (3) the estimators $\widehat{\mathrm{pr}}_c\{Z(t_{j+1}) = z_s | Z(t_j) = z_v\}$ and $\widehat{\mathrm{pr}}_c\{Z(t_j) = z_s | Z(t_{j+1}) = z_v\}$ converge to the true values asymptotically, which readily hold by the law of large numbers.

4.2 Functional Regression Results

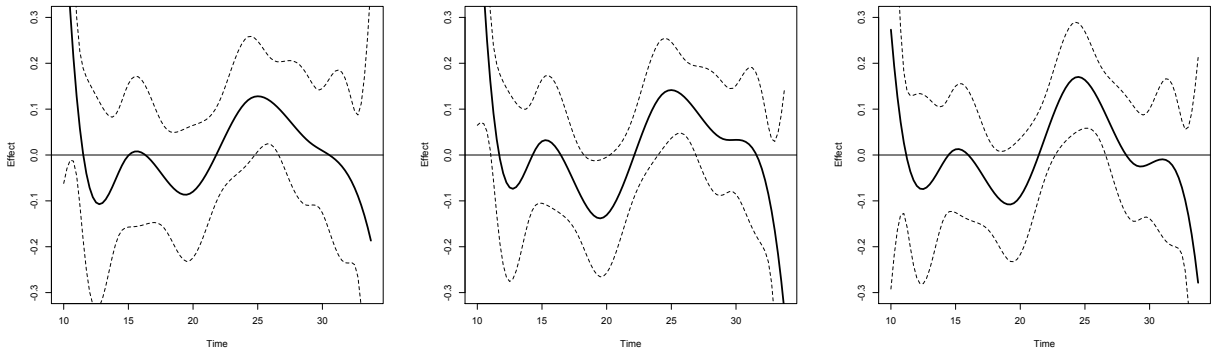
We apply Model (2) to the BOSS data to further examine the time varying effect of ambulatory blood pressure on the three clinical outcomes: stroke recurrence, combined vascular event, and disability. To approximate $\alpha(t)$, we consider cubic splines with uniform knots on the time support, and perform leave-one-out cross-validation to select eight basis functions.

To adjust for potential confounding factors, we include the following covariates: age, gender, the histories of smoking (SMK), drinking (DRINK), diabetes (DIAB) , hypertension (HYPT), dyslipidemia (LIPID), stroke (STROKE), the National Institution of Health stroke scale (NIHSS), and the histories of taking anti-platelet drug (DAP), anti-hypertensive drugs (DHT), and anti-dyslipidemia (DLIPID).

We present the estimated coefficient function $\hat{\alpha}(t)$ and the associated 95% confidence band for the three clinical outcomes, respectively, in Figure 3. The time varying effects of the ambulatory SBP reach the peak and show significance around 1am for all three clinical outcomes. This finding is consistent with the message exhibited in Figure 2. The results also show that the blood pressure patterns are strongly linked with daily activities. For example, the risk is low around noon and 8pm, corresponding to usual lunch and dinner

times, respectively. Furthermore, there is a small bump around 3pm, which may attribute to the Chinese habit of taking nap after lunch.

Figure 3: Estimated Time Varying Effects of the Ambulatory SBP on three Clinical Outcomes: Stroke Recurrence (left), Combined Vascular Events (middle), Disability (right)



In addition, Tables 3-5 show the covariate effects for the clinical outcomes, respectively, with the covariates signifiant at the 5% level highlighted. It can be seen that the history of stroke significantly increases the risk of all three outcomes; Age and the history of taking anti-platelet drug are significant for stroke recurrence; moreover, for combined vascular events, the age effect is significant, while for disability, the history of dyslipidemia and the National Institution of Health stroke scale are significant risk factors.

Table 3: The Covariate Effects for Stroke Recurrence. STD stands for estimated standard deviation.

Covariate	Estimate	STD	p-value	Covariate	Estimate	STD	p-value
AGE	0.0205	0.0103	**0.0458**	GENDER	-0.2111	0.2650	0.4258
SMK	-0.1580	0.2784	0.5703	DRINK	-0.0544	0.1723	0.7521
DIAB	0.2063	0.2389	0.3879	HYPT	-0.3967	0.2439	0.1038
LIPID	0.4182	0.3062	0.1720	STROKE	0.6859	0.2227	**0.0021**
NIHSS	0.0154	0.2104	0.9416	DAP	-0.7402	0.3720	**0.0466**
DHT	-0.0513	0.2532	0.8395	DLIPID	0.2034	0.3158	0.5195

Table 4: The Covariate Effects for Combined Vascular Events. STD stands for estimated standard deviation.

Covariate	Estimate	STD	p-value	Covariate	Estimate	STD	p-value
AGE	0.0273	0.0091	**0.0027**	GENDER	-0.4133	0.2321	0.0749
SMK	-0.1911	0.2424	0.4307	DRINK	-0.1345	0.1534	0.3807
DIAB	0.1180	0.2152	0.5836	HYPT	-0.2328	0.2181	0.2858
LIPID	0.2205	0.2858	0.4404	STROKE	0.6566	0.1964	**0.0008**
NIHSS	0.0868	0.1848	0.6388	DAP	-0.5102	0.3419	0.1356
DHT	-0.0609	0.2221	0.7837	DLIPID	0.0089	0.2668	0.9733

Table 5: The Covariate Effects for Disability. STD stands for estimated standard deviation.

Covariate	Estimate	STD	p-value	Covariate	Estimate	STD	p-value
AGE	0.0088	0.0088	0.3149	GENDER	0.0062	0.2286	0.9784
SMK	-0.2557	0.2450	0.2967	DRINK	0.0182	0.1484	0.9023
DIAB	0.1489	0.2089	0.4762	HYPT	-0.0809	0.2160	0.7079
LIPID	0.5860	0.2571	**0.0226**	STROKE	0.4864	0.1966	**0.0134**
NIHSS	-0.3634	0.1848	**0.0493**	DAP	-0.5605	0.3292	0.0886
DHT	-0.1968	0.2125	0.3544	DLIPID	0.0505	0.2586	0.8451

5 Conclusion

Our research is motivated by China's BOSS study that investigates association between blood pressure and risk of stroke related clinical outcomes. We add to the controversial views about whether ambulatory blood pressure trajectories can effectively predict clinical outcomes of interests. Utilizing China's first large blood pressure database, we showed that the conventional dipping analysis didn't produce clustering that is significantly associated with the outcomes. We then developed an analysis pipeline for clustering the ambulatory blood pressure trajectories and making inference on the time-varying effect of ambulatory blood pressure on clinical outcomes. To identify different blood pressure groups, we introduced a more powerful clustering analysis that makes use of the entire trajectories, instead of just the nocturnal decline in blood pressure as in the dipping analysis. The clustering analysis detected clinically meaningful blood pressure trajectory patterns, which associated significantly with the outcomes. To study the time-varying effect of blood pressure, we first

imputed the missing blood pressure measurements under the Markov assumption, and then fitted a functional logistic regression model, while incorporating baseline covariates. The analysis identified time periods that significantly affect the clinical outcome risk, along with significant baseline covariates. Our research adds a new way of incorporating ambulatory blood pressure in clinical practice.

References

Abraham, C., Cornillon, P.-A., Matzner-Løber, E., and Molinari, N. (2003), "Unsupervised curve clustering using B-splines," *Scandinavian Journal of Statistics*, 30, 581–595.

Ball, G. H. and Hall, D. J. (1967), "A clustering technique for summarizing multivariate data," *Systems Research and Behavioral Science*, 12, 153–155.

Bursztyn, M. (2003), "Morning blood pressure surge and the risk of stroke," *Circulation*, 108, e110–e111.

Chiou, J.-M. and Li, P.-L. (2007), "Functional clustering and identifying substructures of longitudinal data," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69, 679–699.

De Boor, C. (1978), *A practical guide to splines*, vol. 27, New York: Springer-Verlag.

Delaigle, A. and Hall, P. (2016), "Approximating fragmented functional data by segments of Markov chains," *Biometrika*, 103, 779–799.

Feigin, V. L., Krishnamurthi, R. V., Parmar, P., Norrving, B., Mensah, G. A., Bennett, D. A., Barker-Collo, S., Moran, A. E., Sacco, R. L., Truelsen, T., et al. (2015), "Update on the global burden of ischemic and hemorrhagic stroke in 1990-2013: the GBD 2013 study," *Neuroepidemiology*, 45, 161–176.

Garcia-Escudero, L. A. and Gordaliza, A. (2005), "A proposal for robust curve clustering," *Journal of Classification*, 22, 185–201.

Hartigan, J. A. and Wong, M. A. (1979), "Algorithm AS 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28, 100–108.

Huang, J. Z. (1998), "Projection estimation in multiple regression with application to functional ANOVA models," *The Annals of Statistics*, 26, 242–272.

James, P. A., Oparil, S., Carter, B. L., Cushman, W. C., Dennison-Himmelfarb, C., Handler, J., Lackland, D. T., LeFevre, M. L., MacKenzie, T. D., Ogedegbe, O., et al. (2014), "2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8)," *Journal of the American Medical Association*, 311, 507–520.

Jiang, F., Baek, S., Cao, J., and Ma, Y. (2017), "A Functional Single Index Model," *Statistica Sinica*, under review.

Jiang, F. and Ma, Y. (2017), "A spline-assisted semiparametric approach to nonparametric measurement error models," *Annals of Statistics*, under review.

Kario, K., Pickering, T. G., Matsuo, T., Hoshide, S., Schwartz, J. E., and Shimada, K. (2001), "Stroke prognosis and abnormal nocturnal blood pressure falls in older hypertensives," *Hypertension*, 38, 852–857.

Mancia, G., Fagard, R., Narkiewicz, K., Redán, J., Zanchetti, A., Böhm, M., Christiaens, T., Cifkova, R., De Backer, G., Dominiczak, A., et al. (2013), "2013 Practice guidelines for the management of arterial hypertension of the European Society of Hypertension (ESH) and the European Society of Cardiology (ESC): ESH/ESC Task Force for the Management of Arterial Hypertension," *Journal of Hypertension*, 31, 1925–1938.

Metoki, H., Ohkubo, T., Kikuya, M., Asayama, K., Obara, T., Hashimoto, J., Totsune, K., Hoshi, H., Satoh, H., and Imai, Y. (2006), "Prognostic significance for stroke of a morning pressor surge and a nocturnal blood pressure decline," *Hypertension*, 47, 149–154.

O'Donnell, M. J., Xavier, D., Liu, L., Zhang, H., Chin, S. L., Rao-Melacini, P., Rangarajan, S., Islam, S., Pais, P., McQueen, M. J., et al. (2010), "Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study," *The Lancet*, 376, 112–123.

Ogihara, T., Kikuchi, K., Matsuoka, H., Fujita, T., Higaki, J., Horiuchi, M., Imai, Y.,

Imaizumi, T., Ito, S., Iwao, H., et al. (2009), "The Japanese Society of Hypertension guidelines for the management of hypertension (JSH 2009)," *Hypertens Res*, 32, 3–107.

Pierdomenico, S. D., Pierdomenico, A. M., and Cuccurullo, F. (2014), "Morning blood pressure surge, dipping, and risk of ischemic stroke in elderly patients treated for hypertension," *American Journal of Hypertension*, 27, 564–570.

Sargento-Freitas, J., Laranjinha, I., Galego, O., Rebelo-Ferreira, A., Moura, B., Correia, M., Silva, F., Machado, C., Cordeiro, G., and Cunha, L. (2015), "Nocturnal blood pressure dipping in acute ischemic stroke," *Acta Neurologica Scandinavica*, 132, 323–328.

Serban, N. and Wasserman, L. (2005), "CATS: clustering after transformation and smoothing," *Journal of the American Statistical Association*, 100, 990–999.

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., Charlson, F., Davis, A., Degenhardt, L., Dicker, D., et al. (2015), "Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013," *The Lancet*, 386, 743.

Whelton, P. K., Carey, R. M., and AronowLancet, W. S. (2017), "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults," *Journal of the American College of Cardiology*, to appear.

Xu, J., Liu, Y., Tao, Y., Xie, X., Gu, H., Pan, Y., Zhao, X., Wang, Y., Yan, A., and Wang, Y. (2016), "The design, rationale, and baseline characteristics of a nationwide cohort registry in china: blood pressure and clinical outcome in TIA or ischemic stroke," *Patient Preference and Adherence*, 10, 2419.

Yang, G., Wang, Y., Zeng, Y., Gao, G. F., Liang, X., Zhou, M., Wan, X., Yu, S., Jiang, Y., Naghavi, M., et al. (2013), "Rapid health transition in China, 1990–2010: findings from the Global Burden of Disease Study 2010," *The Lancet*, 381, 1987–2015.

Yassouridis, C. (2017), *funcy: Functional Clustering Algorithms*, r package version 0.8.6.

Zhou, M., Wang, H., Zhu, J., Chen, W., Wang, L., Liu, S., Li, Y., Wang, L., Liu, Y., Yin, P., et al. (2016), "Cause-specific mortality for 240 causes in China during 1990–2013: a systematic subnational analysis for the Global Burden of Disease Study 2013," *The Lancet*, 387, 251–272.

A Definition and Regularity Conditions

Let $\mathcal{H}$ be the separable Hilbert space of square integrable functions on $[0, T_{\max}]$, the support of $X_i(s)$. Further let $< \phi_1, \phi_2 > \equiv \int \phi_1(x)\phi_2(x)dx$ denote the inner product of the functions ϕ_1, ϕ_2 in $\mathcal{H}$. Then the L_2 norm $\|\cdot\|_2$ is the norm induced by the inner product. We denote the functional sup norm as $\|\cdot\|_\infty$, and the L_p norm as $\|\cdot\|_p$.

In addition, for any random function $g(x)$ with the ith observation $g_i(x)$, we define $\Gamma(g)$ to be a second moment based linear operator such that

$$\begin{aligned}\Gamma(g)\phi(x) &\equiv \int E\{g_i(x)g_i(s)\}\phi(s)ds \\ &= E\{< g_i, \phi > g_i(x)\},\end{aligned}$$

while its empirical version can be written as

$$\Gamma_n(g)\phi(x) \equiv n^{-1}\sum_{i=1}^{n} < g_i, \phi > g_i(x).$$

Using these definitions, we can write

$$\begin{aligned}< \Gamma(g)B_k, B_l > &= E\left(\int B_k(x)g_i(x)dx \int B_l(x)g_i(x)dx\right), \\ < \Gamma_n(g)B_k, B_l > &= n^{-1}\sum_{i=1}^{n}\int B_k(x)g_i(x)dx \int B_l(x)g_i(x)dx.\end{aligned}$$

Define $\mathbf{C}(g)$ as a $d_{\boldsymbol{\theta}} \times d_{\boldsymbol{\theta}}$ matrix with its (k,l) element $< \Gamma(g)B_k, B_l >$, and $\widehat{\mathbf{C}}(g)$ as a $d_{\boldsymbol{\theta}} \times d_{\boldsymbol{\theta}}$ matrix matrix with its (k,l) element $< \Gamma_n(g)B_k, B_l >$.

We define

$$\begin{aligned}\boldsymbol{\Sigma} &= \{\mathbf{I}_p, \mathbf{B}(t)\}^{\mathrm{T}}\left(E\left[\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}^{\mathrm{T}}(t)Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}^{\mathrm{T}}(t)Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right)^{-1} \\ &\quad \times E\left[\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}^{\mathrm{T}}(t)Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0^{\mathrm{T}}(t)Z_i(t)dt\right\}\right]^{\otimes 2}\right] \\ &\quad \times\left(E\left[\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}^{\mathrm{T}}(t)Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}^{\mathrm{T}}(t)Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right)^{-1}\{\mathbf{I}_p, \mathbf{B}(t)\},\end{aligned} \tag{7}$$

where h' is the derivative of the link function h, $\mathbf{I}_p$ is the $p \times p$ identity matrix, and $\mathbf{a}^{\otimes 2} = \mathbf{a}\mathbf{a}^{\mathrm{T}}$ for a vector $\mathbf{a}$.

We first present the necessary conditions for Theorem 1 to hold.

(A1) Assume $\alpha_0 \in \{\alpha \in L^p([0, T_{\max}])\}$, and is q-times continuous differentiable.

(A2) $\boldsymbol{\gamma}_0$ is a spline coefficient such that $\sup_{t\in[0,T_{\max}]} |\mathbf{B}(t)^{\mathrm{T}}\boldsymbol{\eta}_0 - \alpha_0(t)| = O_p(h_b^q)$. The existence of such $\boldsymbol{\eta}_0$ has been shown in Huang (1998).

(A3) $T \to \infty$, $T/n \to 0$, $N \to \infty$ as $n \to \infty$, and $N \to \infty$ as $n \to \infty$, and $N^{-1}n(\log n)^{-1} \to \infty$ and $Nn^{-1/(2q+1)} = O(1)$.

(A4) $\mathbf{W}_i$ and $\sup_s \|Z(s)\|$ are bounded above in probability.

B Proof of Theorem 1

Lemma 1. *Assume that $C_i(\cdot)$ is a continuous random function of $t \in [0, T_{\max}]$. At each t, $|E\{C_i(t)\}| < \infty$. $\|C_i(\cdot)\|_2 < \infty$ a.s.. Then*

$$|n^{-1}\sum_{i=1}^{n}\int B_k(t)C_i(t)dt| = O_p(h_b) \qquad \text{if} \quad E\{C_i(t)\} \neq 0,$$

and

$$|n^{-1}\sum_{i=1}^{n}\int B_k(t)C_i(t)dt| = O_p\{\sqrt{h_b n^{-1} log(n)}\} \qquad \text{if} \quad E\{C_i(t)\} = 0.$$

The lemma is a direct consequence of Lemma 3 in Jiang et al. (2017).

Lemma 2. *Let $\mathbf{V}_i = \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt$. For a bounded positive random variable $D(\mathbf{W}_i, \mathbf{V}_i)$, a p-dimensional vector $\mathbf{Q}_1 = O_p(1)$, and a random function $\|Q_2(\cdot)\| = O_p(1)$, we have*

$$E[\{(\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\mathbf{V}_i^{\mathrm{T}})^{\mathrm{T}}\}^{\otimes 2} D(\mathbf{W}_i, \mathbf{V}_i)]^{-1} \left\{ \begin{array}{c} E(\mathbf{Q}_{i1}) \\ \int h_b^{-1/2}\mathbf{B}(s)E\{Q_{i2}(s)\}ds \end{array} \right\} = (\mathbf{R}_1^{\mathrm{T}}, \mathbf{R}_2^{\mathrm{T}})^{\mathrm{T}},$$

where $\mathbf{R}_1$ is a d-dimensional vector with $\|\mathbf{R}_1\| = O_p\{\|E(\mathbf{Q}_{i1})\| + \|E(Q_{i2}(\cdot)\|\}$ and $\mathbf{R}_2$ is a $N + r$-dimensional vector with $\|\mathbf{R}_2\| = O\{\|E(Q_{i2}(\cdot)\|\}$. Furthermore, we have

$$E[\{(\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\mathbf{V}_i^{\mathrm{T}})^{\mathrm{T}}\}^{\otimes 2} D(\mathbf{W}_i, \mathbf{V}_i)]^{-1} \left\{ \begin{array}{c} n^{-1}\sum_{i=1}^{n}\mathbf{Q}_{i1} \\ \int h_b^{-1/2}\mathbf{B}(s)n^{-1}\sum_{i=1}^{n} Q_{i2}(s)ds \end{array} \right\} = (\mathbf{T}_1^{\mathrm{T}}, \mathbf{T}_2^{\mathrm{T}})^{\mathrm{T}},$$

where $\mathbf{T}_1$ is a d-dimensional vector with $\|\mathbf{T}_1\| = O_p\{\|n^{-1}\sum_{i=1}^{n}\mathbf{Q}_{i1}\| + \|n^{-1}\sum_{i=1}^{n} Q_{i2}(s)\|\}$ and $\mathbf{T}_2$ is a $N + r$-dimensional vector with $\|\mathbf{T}_2\| = O\{\|n^{-1}\sum_{i=1}^{n} Q_{i2}(\cdot)\|\}$.

Proof: Without loss of generality, we assume $D(\mathbf{W}_i, \mathbf{V}_i) = 1$, and we have

$$
\begin{aligned}
& E[\{(\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\mathbf{V}_i^{\mathrm{T}})^{\mathrm{T}}\}^{\otimes 2}]^{-1} \left\{ \begin{array}{c} E(\mathbf{Q}_{i1}) \\ \int h_b^{-1/2}\mathbf{B}(s)E\{Q_{i2}(s)\}ds \end{array} \right\} \\
= & \left\{ \begin{array}{cc} \mathbf{I} & \mathbf{0} \\ -h_b^{1/2}E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}}) & \mathbf{I} \end{array} \right\} \\
& \times \left[\begin{array}{cc} \left\{E(\mathbf{W}_i\mathbf{W}_i^{\mathrm{T}}) - E(\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}})E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}})\right\}^{-1} & \mathbf{0} \\ \mathbf{0} & h_b E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1} \end{array} \right] \\
& \times \left\{ \begin{array}{cc} \mathbf{I} & -h_b^{1/2}E(\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}})E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1} \\ \mathbf{0} & \mathbf{I} \end{array} \right\} \left\{ \begin{array}{c} E(\mathbf{Q}_{i1}) \\ \int h_b^{-1/2}\mathbf{B}(s)E\{Q_{i2}(s)\}ds \end{array} \right\} \\
= & \left\{ \begin{array}{cc} \mathbf{I} & \mathbf{0} \\ -h_b^{1/2}E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}}) & \mathbf{I} \end{array} \right\} \\
& \times \left[\begin{array}{cc} \left\{E(\mathbf{W}_i\mathbf{W}_i^{\mathrm{T}}) - E(\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}})E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}})\right\}^{-1} & \mathbf{0} \\ \mathbf{0} & h_b E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1} \end{array} \right] \\
& \times \left[\begin{array}{c} E(\mathbf{Q}_{i1}) - E(\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}})E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}\int \mathbf{B}(s)E\{Q_{i2}(s)\}ds \\ \int h_b^{-1/2}\mathbf{B}(s)E\{Q_{i2}(s)\}ds \end{array} \right] \\
= & \left\{ \begin{array}{cc} \mathbf{I} & \mathbf{0} \\ -E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}}) & \mathbf{I} \end{array} \right\} \\
& \times \left(\begin{array}{c} \left\{E(\mathbf{W}_i\mathbf{W}_i^{\mathrm{T}}) - E(\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}})E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}})\right\}^{-1} \\ \left[E(\mathbf{Q}_{i1}) - E(\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}})E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}\int \mathbf{B}(s)E\{Q_{i2}(s)\}ds\right] \\ h_b^{1/2}E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}\int \mathbf{B}(s)E\{Q_{i2}(s)\}ds \end{array} \right). \qquad (8)
\end{aligned}
$$

First note that

$$
\|\mathbf{W}_i\| = O_p(1), \|\mathbf{W}_i\mathbf{W}_i^{\mathrm{T}}\| = O_p(1), E(\|\mathbf{W}_i\|) = O(1), E(\|\mathbf{W}_i\mathbf{W}_i^{\mathrm{T}}\|) = O(1). \qquad (9)
$$

Furthermore, recall that $\mathbf{V}_i = \int Z_i(s)\mathbf{B}(s)ds$ and $\sup_s \|Z_i(s)\|_\infty < \infty$ by Condition (A4).

Hence

$$\begin{aligned}
\|\mathbf{V}_i\| &= \left[\sum_{l=1}^{d}\left\{\int Z_i(s)B_l(s)ds\right\}^2\right]^{1/2} \\
&= \left[\sum_{l=1}^{d}\left\{Z_i(\eta)\int B_l(s)ds\right\}^2\right]^{1/2} \\
&\leq \left[\sum_{l=1}^{d} D_1(t_l - t_{l-r})^2\right]^{1/2} \\
&= O_p(h_b^{1/2}), \qquad (10)
\end{aligned}$$

where η is a point in on the support of Z_i, D_1 is a positive constant. The second equality holds by the mean value theorem, and the third line holds because $\sup_s \|Z(s)\|$ is bounded by Condition (A4), B_l is positive in the interval (t_{lr}, t_k), and is 0 otherwise (page 88 in De Boor (1978)). Hence

$$\begin{aligned}
&\|\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}}\| = O(h_b^{1/2}), E(\|\mathbf{V}_i\|) = O(h_b^{1/2}), \|E(\mathbf{V}_i)\| = O(h_b^{1/2}), \\
&E(\|\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}}\|) = O(h_b^{1/2}), \|E(\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}})\| = O(h_b^{1/2}). \qquad (11)
\end{aligned}$$

Similarly we have

$$\left\|\int \mathbf{B}(s)E\{Q_{i2}(s)\}ds\right\| = O_p(h_b^{1/2})O_p(\|E\{Q_{i2}(\cdot)\}\|). \qquad (12)$$

In addition,

$$\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}} = \int Z_i(s)\mathbf{B}(s)ds \int Z_i(s)\mathbf{B}(s)^{\mathrm{T}}ds.$$

From Lemma 2 in Jiang and Ma (2017), for any vector $\mathbf{u}$ with $\|\mathbf{u}\| = 1$, there exist positive constants that

$$\begin{aligned}
&D_3 h_b \leq \mathbf{u}^{\mathrm{T}}E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})\mathbf{u} \leq D_4 h_b, \\
&D_4^{-1}h_b^{-1} \leq \mathbf{u}^{\mathrm{T}}E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}\mathbf{u} \leq D_3^{-1}h_b^{-1}, \\
&D_5 h_b \leq \mathbf{u}^{\mathrm{T}}n^{-1}\sum_{i=1}^{n}\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}}\mathbf{u} \leq D_6 h_b, \\
&D_6^{-1}h_b^{-1} \leq \mathbf{u}^{\mathrm{T}}\left(n^{-1}\sum_{i=1}^{n}\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}}\right)^{-1}\mathbf{u} \leq D_5^{-1}h_b^{-1}. \qquad (13)
\end{aligned}$$

This implies $\|E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})\| = O_p(h_b)$, and $\|E(\mathbf{V}_i\mathbf{V}_i^{\mathrm{T}})^{-1}\| = O_p(h_b)$. Furthermore, since $E(\|\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}}\|) = E(\sup_{\mathbf{u},\|u\|=1}\mathbf{u}^{\mathrm{T}}\mathbf{V}_i\mathbf{W}_i^{\mathrm{T}}\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}}\mathbf{u})^{1/2}$.

Using the same argument as those leading to (13), we have

$$E(\|\mathbf{W}_i\mathbf{V}_i^{\mathrm{T}}\|) = O_p(h_b^{1/2}). \tag{14}$$

Now plugging the results (9), (11), (13), (14), (12) into (8), and noting that the first matrix in (8) is of order $O(1)$, we finally have

$$E[\{(\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\mathbf{V}_i^{\mathrm{T}})\}^{\otimes 2}]^{-1}\left\{\begin{array}{c} E(\mathbf{Q}_{i1}) \\ \int h_b^{-1/2}\mathbf{B}(s)E\{Q_{i2}(s)\}ds \end{array}\right\} = (\mathbf{R}_1^{\mathrm{T}}, \mathbf{R}_2^{\mathrm{T}})^{\mathrm{T}}, \tag{15}$$

where $\mathbf{R}_1$ is a d-dimensional vector with $\|\mathbf{R}_1\| = O_p\{\|E(\mathbf{Q}_{i1})\| + \|E(Q_{i2}(\cdot)\|\}$ and $\mathbf{R}_2$ is a $N + r$-dimensional vector with $\|\mathbf{R}_2\| = O\{\|E(Q_{i2}(\cdot))\|\}$.

Using the same argument as those leading to (15), we have

$$E[\{(\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\mathbf{V}_i^{\mathrm{T}})\}^{\otimes 2}]^{-1}\left\{\begin{array}{c} n^{-1}\sum_{i=1}^n \mathbf{Q}_{i1} \\ \int h_b^{-1/2}\mathbf{B}(s)n^{-1}\sum_{i=1}^n Q_{i2}(s)ds \end{array}\right\} = (\mathbf{T}_1^{\mathrm{T}}, \mathbf{T}_2^{\mathrm{T}})^{\mathrm{T}},$$

where $\mathbf{T}_1$ is a d-dimensional vector with $\|\mathbf{T}_1\| = O_p\{\|n^{-1}\sum_{i=1}^n \mathbf{Q}_{i1}\| + \|n^{-1}\sum_{i=1}^n Q_{i2}(\cdot)\|\}$ and $\mathbf{T}_2$ is a $N + r$-dimensional vector with $\|\mathbf{T}_2\| = O_p[\{\|n^{-1}\sum_{i=1}^n Q_{i2}(\cdot)\|\}]$.

This proves the results. □

Proof of Theorem 1:

$$\begin{aligned}
\mathbf{0} &= \mathbf{S}_i(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\eta}}) \\
&= \sum_{i=1}^n \left\{\mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}}\widehat{\nu}_i(t)dt\right\}^{\mathrm{T}}\left(\left[Y_i - h\left\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dtI\widehat{\boldsymbol{\eta}}\right\}\right]\right. \\
&\quad + h\left\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\widehat{\boldsymbol{\eta}}\right\} - h\left\{\rho^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}\nu_i(t)dt\widehat{\boldsymbol{\eta}}\right\} \\
&\quad \left. + h\left\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}\nu_i(t)dt\widehat{\boldsymbol{\eta}}\right\} - h\left\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}\widehat{\nu}_i(t)dt\widehat{\boldsymbol{\eta}}\right\}\right) \\
&= \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3,
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{R}_1 &= \sum_{i=1}^{n} \left\{ \mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} \widehat{\nu}_i(t) dt \right\}^{\mathrm{T}} \left[Y_i - h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t) dt I \widehat{\boldsymbol{\eta}} \right\} \right] \\
\mathbf{R}_2 &= \sum_{i=1}^{n} \left\{ \mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} \widehat{\nu}_i(t) dt \right\}^{\mathrm{T}} \\
&\quad \times \left[h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t) dt \widehat{\boldsymbol{\eta}} \right\} - h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \right] \\
\mathbf{R}_3 &= \sum_{i=1}^{n} \left\{ \mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} \widehat{\nu}_i(t) dt \right\}^{\mathrm{T}} \\
&\quad \times \left[h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} - h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \widehat{\nu}_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \right].
\end{aligned}$$

First note that

$$\begin{aligned}
\mathbf{R}_2 &= \sum_{i=1}^{n} \left\{ \mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} \widehat{\nu}_i(t) dt \right\} \\
&\quad \times \left[h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t) dt \widehat{\boldsymbol{\eta}} \right\} - h\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \right] \\
&= \sum_{i=1}^{n} \left\{ \mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} \widehat{\nu}_i(t) dt \right\} h'\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \nu_{ic}(t) dt \widehat{\boldsymbol{\eta}} \right\} \\
&\quad \times \left\{ \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t) dt \widehat{\boldsymbol{\eta}} - \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \\
&= \left(\sum_{i=1}^{n} \left\{ \int \mathbf{B}(t) \nu_i(t) dt \right\} h'\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \right. \\
&\quad \left. \times \left\{ \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t) dt \widehat{\boldsymbol{\eta}} - \int \mathbf{B}(t)^{\mathrm{T}} \nu_{ic}(t) dt \widehat{\boldsymbol{\eta}} \right\} \right) \{1 + o_p(1)\} \\
&= O_p(T^{-1/2}),
\end{aligned}$$

where h' is the first derivative of h. The last equality holds by the fact

$$\begin{aligned}
& E\left(\sum_{i=1}^{n} \left\{ \int \mathbf{B}(t) \nu_i(t) dt \right\} h'\left\{ \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \right. \\
& \left. \times \left\{ \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t) dt \widehat{\boldsymbol{\eta}} - \int \mathbf{B}(t)^{\mathrm{T}} \nu_i(t) dt \widehat{\boldsymbol{\eta}} \right\} \right) = \mathbf{0}.
\end{aligned}$$

Similarly, we have $\mathbf{R}_3 = O_p(T^{-1/2})$. The convergences of $\mathbf{R}_2, \mathbf{R}_3$ depend only on the grid

size specifications. Hence

$$
\begin{aligned}
\mathbf{0} &= n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}\widehat{\nu}_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\widehat{\boldsymbol{\eta}}\right\}\right]\{1+o_p(1)\} \\
&= \left(n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right. \\
&\quad -n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\otimes 2} \\
&\quad \left.\times h'\left\{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}\right\}\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})^{\mathrm{T}}, h_b^{1/2}(\widehat{\boldsymbol{\eta}}-\boldsymbol{\eta}_0)^{\mathrm{T}}\}\right)\{1+o_p(1)\},
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})^{\mathrm{T}}, h_b^{1/2}(\widehat{\boldsymbol{\eta}}-\boldsymbol{\eta}_0)^{\mathrm{T}}\}^{\mathrm{T}} \\
&= \left(n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}\right\}\right)^{-1} \\
&\quad \times\left(n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right)\{1+o_p(1)\} \\
&= \left\{E\left(\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}_0\right\}\right)\right\}^{-1} \\
&\quad \times\left(n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0(t)^{\mathrm{T}}Z_i(t)dt\right\}\right]\right)\{1+o_p(1)\} \\
&\quad + \left\{E\left(\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}_0\right\}\right)\right\}^{-1} \\
&\quad \times n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\mathrm{T}}\left[h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0(t)Z_i(t)dt\right\}\right. \\
&\quad \left.-h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\Bigg\} \\
&\quad \times\{1+o_p(1)\} \\
&= \left\{E\left(\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\boldsymbol{\eta}_0\right\}\right)\right\}^{-1} \\
&\quad \times\left(n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}}Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0(t)^{\mathrm{T}}Z_i(t)dt\right\}\right]\right) \\
&\quad \times\{1+o_p(1)\}. \qquad (16)
\end{aligned}
$$

The last equality holds because by Lemma 1, we have

$$\begin{aligned}
&\left\| n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\right\}^{\mathrm{T}}\left[h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0(t)Z_i(t)dt\right\}\right.\right. \\
&\left.\left. -h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right\|_\infty \\
= \;& O_p\left(h_b^{q+1}\right),
\end{aligned}$$

and

$$\begin{aligned}
&\left\| n^{-1}\sum_{i=1}^{n}\left\{\mathbf{W}_i^{\mathrm{T}}, \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0(t)^{\mathrm{T}} Z_i(t)dt\right\}\right]\right\|_\infty \\
= \;& O_p\left\{\sqrt{h_b n^{-1}\log(n)}\right\}.
\end{aligned}$$

Furthermore, by Condition (A3), the second display dominates the first one. Now by Lemma 2, (16) is of order $O_p(1)$. For each t,

$$\begin{aligned}
&\mathbf{B}(t)^{\mathrm{T}}\widehat{\boldsymbol{\eta}} - \alpha_0(t) \\
= \;& \mathbf{B}(t)^{\mathrm{T}}\widehat{\boldsymbol{\eta}} - \mathbf{B}(t)^{\mathrm{T}}\boldsymbol{\eta}_0 + \mathbf{B}(t)^{\mathrm{T}}\boldsymbol{\eta}_0 - \alpha_0(t) \\
= \;& \mathbf{B}(t)^{\mathrm{T}}(\widehat{\boldsymbol{\eta}} - \widehat{\boldsymbol{\eta}}_0)\{1 + o_p(1)\}.
\end{aligned}$$

The second equality holds because $\mathbf{B}(t)^{\mathrm{T}}\boldsymbol{\eta}_0 - \alpha_0(t) = h_b^{-q}$ by Condition (A2) and $h_b n^{1/(2q+1)} \to 0$ by Condition (A3). Therefore,

$$\sqrt{n}[(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^{\mathrm{T}}, h_b^{1/2}\{\mathbf{B}(t)^{\mathrm{T}}\widehat{\boldsymbol{\eta}} - \alpha_0(t)\}]^{\mathrm{T}} = \sqrt{n}\{\mathbf{I}_p, \mathbf{B}(t)\}^{\mathrm{T}}\{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^{\mathrm{T}}, h_b^{1/2}(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0)^{\mathrm{T}}\}^{\mathrm{T}}\{1 + o_p(1)\}$$

converges to a normal random vector with mean $\mathbf{0}$ and variance

$$\begin{aligned}
\boldsymbol{\Sigma} = \;& \{\mathbf{I}_p, \mathbf{B}(t)\}^{\mathrm{T}}\left(E\left[\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right)^{-1} \\
&\times E\left[\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\right\}^{\mathrm{T}}\left[Y_i - h\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \alpha_0(t)^{\mathrm{T}} Z_i(t)dt\right\}\right]^{\otimes 2}\right] \\
&\times\left(E\left[\left\{\mathbf{W}_i^{\mathrm{T}}, h_b^{-1/2}\int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\right\}^{\otimes 2} h'\left\{\boldsymbol{\beta}_0^{\mathrm{T}}\mathbf{W}_i + \int \mathbf{B}(t)^{\mathrm{T}} Z_i(t)dt\boldsymbol{\eta}_0\right\}\right]\right)^{-1}\{\mathbf{I}_p, \mathbf{B}(t)\}.
\end{aligned}$$

□