Statistica Sinica **22** (2012), 123-148
doi:http://dx.doi.org/10.5705/ss.2009.229

TESTS FOR VARIANCE COMPONENTS IN VARYING COEFFICIENT MIXED MODELS

Zaixing Li, Yuedong Wang, Ping Wu, Wangli Xu and Lixing Zhu

China University of Mining & Technology, Beijing,
University of California, Santa Barbara, East China Normal University,
RenMin University of China and Hong Kong Baptist University

Abstract: We consider a general class of varying coefficient mixed models where random effects are introduced to account for between-subject variation. To address the question of whether a varying coefficient mixed model can be reduced to a simpler varying coefficient model, we develop one-sided tests for the null hypothesis that all the variance components are zero. In addition to the purely null-based standard quasi-score test (SQT), we propose an extended quasi-score test (EQT) by constructing estimators that are consistent under both the null and alternative hypotheses. No assumptions are required for the distributions of random effects and random errors. Both SQT and EQT are consistent for global alternatives and local alternatives distinct at certain rates from the null. Furthermore, the asymptotic null distributions are simple and easy to use in practice. For comparison, we also adapt the one-sided score test (SST) in Silvapulle and Silvapulle (1995) and the likelihood ratio test (LRT) in Fan, Zhang, and Zhang (2001). Extensive simulations indicate that all proposed tests perform well and the EQT is more powerful than SQT, SST, and LRT. A data example is analyzed for illustration.

Key words and phrases: Extended quasi-likelihood, likelihood ratio test, longitudinal data, random effects, score test, smoothing spline, variance components, varying coefficient models.

1. Introduction

Varying coefficient models (VCM) are widely used to analyze longitudinal data because of their flexibility and relative simplicity (Hoover et al. (1998); Wu and Chiang (2000)). To deal with between-subject variation and within-subject correlation, various forms of varying coefficient mixed models (VCMM) have been proposed recently (Wu and Liang (2004); Zhang (2004); and Wu and Zhang (2006)). One important question is whether a VCMM can be reduced to a VCM since inference for a VCM is much simpler and potentially more efficient. In many applications testing homogeneity is of primary interest (Jacqmin-Gadda and Commenges (1995); Commenges and Jacqmin-Gadda (1997)). The problem can be cast in a framework of hypothesis testing on variance components of random effects. The goal of this article is to develop robust, powerful and easy-to-use tests for the null hypothesis that all variance components are zero.

Consider the VCMM

$$y_{ij} = \sum_{l=1}^{p} x_{ijl} f_l(t_{ij}) + \mathbf{Z}_{ij}^{\mathrm{T}} \mathbf{b}_i + \epsilon_{ij}, \quad i = 1, \ldots, m, \ \ j = 1, \ldots, n_i, \qquad (1.1)$$

where y_{ij} is the observation from subject i at time point t_{ij}, x_{ijl}'s are covariates for the fixed effects with time varying coefficient functions $f_l(\cdot)$'s, $\mathbf{Z}_{ij} = (z_{ij1}, \ldots, z_{ijq})^{\mathrm{T}}$ is a q-dimensional covariate vector for the random effects, $\mathbf{b}_i$'s are i.i.d. q-dimensional random effects with mean zero and covariance $\sigma^2\mathbf{D}_1$, and ϵ_{ij}'s are i.i.d. random errors with mean zero, variance σ^2, and finite fourth moment $\kappa = \mathrm{E}\epsilon_{ij}^4$. We assume that ϵ_{ij}'s are independent of $\mathbf{b}_i$'s, and $f_l(\cdot)$'s are twice-differentiable functions on a finite interval. Without loss of generality, we take the interval to be $[0, 1]$.

Model (1.1) is a natural extension of the VCM where the random effects are introduced to model between-subject variation. The coefficient functions $f_l(\cdot)$'s are modeled nonparametrically using smoothing splines. When the $\mathbf{b}_i$'s are normally distributed and the y_{ij}'s belong to an exponential family, (1.1) is a special case of the generalized linear mixed models with varying coefficients proposed in Zhang (2004). Since the fixed-effects component is a time-varying coefficients model and the random-effects component is a parametric model, (1.1) may also be regarded as a time-varying coefficients semiparametric mixed effects model as defined in Chapter 9 of Wu and Zhang (2006). Note that normality assumptions were made in Wu and Zhang (2006), and they did not consider inference about variance components. In this paper we propose estimation and inference methods for VCMM.

We concentrate on testing the hypothesis that all variance components are zero. This has not been studied for the VCMM. Many hypothesis testing methods have been developed for various parametric and nonparametric mixed models. In particular, Lin (1997) and Zhu and Fung (2004) developed score tests for variance components in a generalized linear mixed model (GLMM) and a semiparametric mixed model, respectively. Both tests considered two-sided alternative hypotheses on variance components. Often, variance components are known to be non-negative which leads to a one-sided alternative hypothesis (Verbeke and Molenberghs (2003)). Some existing one-sided tests can be found in Silvapulle and Silvapulle (1995) and Silvapulle and Sen (2005) When distributions of the random effects and random errors are known, a common approach is to construct the LRT. For example, Stram and Lee (1994) applied the theory in Self and Liang (1987) to linear mixed models with normality assumptions on both the random effects and random errors.

In this article we develop quasi-score tests for one-sided alternative hypothesis. We first propose an SQT that uses estimators of parameters that are consistent under the null and inconsistent under the alternative. We then propose an EQT by constructing estimators that are consistent under both the null and alternative hypotheses. We show that both SQT and EQT are consistent for global alternatives and sensitive to local alternatives converging at certain rates to the null. Furthermore, the quasi-score test statistics asymptotically follow distributions of maximums of several standard normal random variables. Consequently, these tests are easy to use in practice. Other than moment conditions on random effects and random errors, no assumptions are made on their distributions. Construction of consistent estimators, in particular for the fourth moment of random errors κ, is technically challenging. The proposed consistent estimators for the moments of random errors are of interest in themselves. For comparison, we also adapt the one-sided score test (SST) in Silvapulle and Silvapulle (1995) and the LRT in Fan, Zhang, and Zhang (2001). Since nonparametric components appear in both the null and the alternative hypotheses in our setup, extension of the test in Stram and Lee (1994) is difficult and beyond the scope of this paper. Simulations indicate that the EQT is more powerful than SQT, SST, and LRT.

The paper is organized as follows. In Section 2, we derive estimates for the parameters and nonparametric functions in a VCMM. In Section 3, we construct quasi-score tests for one-sided hypothesis and present their asymptotic properties. We also present the SST and LRT in this section. Simulations results are reported in Section 4. The application to a data set is presented in Section 5. Section 6 gives some conclusions. Some lengthy notations, regularity conditions, sketch of proofs, and brief derivations are deferred to the Appendix.

2. Estimation

We rewrite model (1.1) in matrix form. Let $n = \sum_{i=1}^m n_i$, $\mathbf{Y}_i = (y_{i1}, \ldots, y_{in_i})^{\mathrm{T}}$ and $\mathbf{Y} = (\mathbf{Y}_1^{\mathrm{T}}, \ldots, \mathbf{Y}_m^{\mathrm{T}})^{\mathrm{T}}$. Notations $\boldsymbol{\epsilon}_i$, $\boldsymbol{\epsilon}$, and $\mathbf{b}$ are defined similarly. Let $\mathbf{X}_l = \mathrm{diag}(x_{11l}, \ldots, x_{1n_1l}, \ \ldots, x_{m1l}, \cdots x_{mn_ml})$ be an $n \times n$ matrix and $\mathbf{Z}_i = (\mathbf{Z}_{i1}, \ldots, \mathbf{Z}_{in_i})^{\mathrm{T}}$ be a $n_i \times q$ matrix. Let $\mathbf{t}^0 = (t_1^0, \ldots, t_r^0)^{\mathrm{T}}$ be the vector of ordered distinct values of the collection of all time points $\{t_{ij} : i = 1, \ldots, m; \ j = 1, \ldots, n_i\}$. Let $\mathbf{f}_l = (f_l(t_1^0), \ldots, f_l(t_r^0))^{\mathrm{T}}$ and $\mathbf{f} = (\mathbf{f}_1^{\mathrm{T}}, \ldots, \mathbf{f}_p^{\mathrm{T}})^{\mathrm{T}}$. Let $\mathbf{N}_l$ be the incidence matrix mapping $\{t_{ij}\}$ to $\mathbf{t}^0$ such that $(f_l(t_{11}), \ldots, f_l(t_{1n_1}), \ldots, f_l(t_{m1}), \ldots, f_l(t_{mn_m}))^{\mathrm{T}} = \mathbf{N}_l\mathbf{f}_l$. Let $\mathbf{N} = \mathrm{diag}(\mathbf{N}_1, \ldots, \mathbf{N}_p)$, $\mathbf{Z} = \mathrm{diag}(\mathbf{Z}_1, \ldots, \mathbf{Z}_m)$, $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_p)$ and $\tilde{\mathbf{X}} = \mathbf{X}\mathbf{N}$. Then (1.1) can be rewritten as

$$\mathbf{Y} = \tilde{\mathbf{X}}\mathbf{f} + \mathbf{Z}\mathbf{b} + \boldsymbol{\epsilon}. \tag{2.1}$$

We model the $f_l(\cdot)$'s using natural cubic smoothing splines (Wahba (1990); Green and Silverman (1994)). Specifically, we use the value-second derivative representation in Green and Silverman (1994) to represent coefficient functions $f_1, \ldots, f_p$.

Note that (2.1) is a nonparametric model where the vector $\mathbf{f}$ is part of the value-second derivative representation rather than a parameter. We assume that the number of knots tends to infinity as the sample size tends to infinity. This assumption is implied by condition $C1$ in Appendix B.

The covariance matrix of $\mathbf{b}$ is $\sigma^2\mathbf{D}$, where $\mathbf{D} = \text{diag}(\mathbf{D}_1, \ldots, \mathbf{D}_1)$. Here we develop estimates for $f_1(\cdot), \ldots, f_p(\cdot)$, σ^2, and $\mathbf{D}$. We note that our estimation methods are robust in the sense that they do not require specifying distributions for random effects and random errors. For fixed $\mathbf{D}$, as Lin and Zhang (1999), Gu and Ma (2005), and Wu and Zhang (2006), we estimate $f_1(\cdot), \ldots, f_p(\cdot)$ and $\mathbf{b}$ by minimizing the penalized least squares

$$(\mathbf{Y} - \tilde{\mathbf{X}}\mathbf{f} - \mathbf{Z}\mathbf{b})^{\mathrm{T}}(\mathbf{Y} - \tilde{\mathbf{X}}\mathbf{f} - \mathbf{Z}\mathbf{b}) + \mathbf{b}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{b} + \sum_{l=1}^{p}\lambda_l \int_0^1 \{f_l''(t)\}^2 dt. \quad (2.2)$$

Similar to O'Sullivan, Yandell, and Raynor (1986), it can be shown that the solutions of $f_l(\cdot)$'s belong to finite dimensional spaces. Thus, (2.2) reduces to

$$(\mathbf{Y} - \tilde{\mathbf{X}}\mathbf{f} - \mathbf{Z}\mathbf{b})^{\mathrm{T}}(\mathbf{Y} - \tilde{\mathbf{X}}\mathbf{f} - \mathbf{Z}\mathbf{b}) + \mathbf{b}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{b} + \mathbf{f}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{K}\mathbf{f}, \quad (2.3)$$

where $\boldsymbol{\Lambda} = \text{diag}(\lambda_1\mathbf{I}_r, \ldots, \lambda_p\mathbf{I}_r)$, $\mathbf{K} = \text{diag}(\mathbf{K}_1, \ldots, \mathbf{K}_p)$, λ_l's are smoothing parameters, and $\mathbf{K}_l$ are the non-negative definite smoothing matrices defined in (2.3) of Green and Silverman (1994). The solutions to (2.3) are

$$\begin{aligned}\hat{\mathbf{f}} &= (\tilde{\mathbf{X}}^{\mathrm{T}}\mathbf{V}^{-1}\tilde{\mathbf{X}} + \boldsymbol{\Lambda}\mathbf{K})^{-1}\tilde{\mathbf{X}}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{Y},\\ \hat{\mathbf{b}} &= (\mathbf{Z}^{\mathrm{T}}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}^{\mathrm{T}}(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}}),\end{aligned} \quad (2.4)$$

where $\mathbf{V} = \mathbf{I}_n + \mathbf{Z}\mathbf{D}\mathbf{Z}^{\mathrm{T}}$. We construct moment-based estimators of $\mathbf{D}_1$ and σ^2 as

$$\hat{\mathbf{D}}_1 = (\tilde{\boldsymbol{\Sigma}}_z)^{-1/2}\hat{\mathbf{B}}(\tilde{\Sigma}_z)^{-1/2}, \quad (2.5)$$

$$\hat{\sigma}^2 = \frac{(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}})^{\mathrm{T}}(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}})}{n - tr\tilde{\mathbf{H}}_\lambda + tr(\mathbf{Z}\hat{\mathbf{D}}\mathbf{Z}^{\mathrm{T}})}, \quad (2.6)$$

where $\tilde{\boldsymbol{\Sigma}}_z$ and $\hat{\mathbf{B}}$ are defined in (A.1) and (A.2) in Appendix A, $\hat{\mathbf{D}} = \text{diag}(\hat{\mathbf{D}}_1, \ldots, \hat{\mathbf{D}}_1)$, $\tilde{\mathbf{H}}_\lambda = \tilde{\mathbf{X}}(\tilde{\mathbf{X}}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\tilde{\mathbf{X}} + \boldsymbol{\Lambda}\mathbf{K})^{-1}\tilde{\mathbf{X}}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}$, and $\hat{\mathbf{V}} = \mathbf{I}_n + \mathbf{Z}\hat{\mathbf{D}}\mathbf{Z}^{\mathrm{T}}$. Appendix D provides a brief derivation of $\hat{\mathbf{D}}_1$. $\hat{\mathbf{D}}_1$ and $\hat{\sigma}^2$ are consistent under both the null and the alternative (Lemma 2 in Appendix C).

When $\mathbf{D} = \mathbf{0}$, which corresponds to a VCM, we have

$$\begin{aligned}\hat{\mathbf{f}}_0 &= (\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}} + \boldsymbol{\Lambda}\mathbf{K})^{-1}\tilde{\mathbf{X}}^{\mathrm{T}}\mathbf{Y},\\ \hat{\sigma}_0^2 &= \frac{(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}}_0)^{\mathrm{T}}(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}}_0)}{n - tr\mathbf{H}_\lambda},\end{aligned} \quad (2.7)$$

where $\mathbf{H}_\lambda = \tilde{\mathbf{X}}(\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}} + \mathbf{\Lambda}\mathbf{K})\tilde{\mathbf{X}}^{\mathrm{T}}$. The estimator $\hat{\sigma}_0^2$ is consistent under the null and inconsistent under the alternative (Lemma 2 in Appendix C).

Smoothing parameters $\lambda_1, \ldots, \lambda_p$ are crucial to the performance of the spline estimators. We apply generalized cross-validation (GCV) to select these smoothing parameters in simulations and data analysis. Specifically, the GCV estimates of smoothing parameters are minimizers of

$$GCV(\lambda_1, \ldots, \lambda_p) = \frac{n\mathbf{Y}^{\mathrm{T}}(\mathbf{I}_n - \mathbf{H}_\lambda)^2\mathbf{Y}}{(n - tr(\mathbf{H}_\lambda))^2}.$$

As in Fan and Huang (2005), the same smoothing parameters are used for constructing estimators under both the null and the alternative hypotheses.

3. Tests

When making inference for variance components, it is often the case that these variances are known to be non-negative. One-sided test should be used in these situations (Verbeke and Molenberghs (2003)). In this section we assume that the diagonal of $\mathbf{D}_1$ is a linear function of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$ with $\theta_i \geq 0$, and $\mathbf{D}_1 = \mathbf{0}$ if $\boldsymbol{\theta} = \mathbf{0}$. Thus, we consider the one-sided hypothesis

$$H_0: \ \boldsymbol{\theta} = \mathbf{0} \quad \text{versus} \quad H_A: \ \boldsymbol{\theta} \geq \mathbf{0} \text{ and } \theta_i > 0 \text{ for some } 1 \leq i \leq d. \qquad (3.1)$$

When all diagonal elements of $\mathbf{D}_1$ are free parameters, one may set $\boldsymbol{\theta}$ to be these diagonal elements. The above formulation is more general and allows certain relationships among diagonal elements of $\mathbf{D}_1$. For example, $\boldsymbol{\theta}$ may represent distinct diagonal elements when some diagonal elements are equal. The nonnegativity constraint on $\boldsymbol{\theta}$ is used in the construction of one-sided tests, while the information that $\mathbf{D}_1$ is a non-negative definite matrix is ignored.

Our construction of test statistics is motivated by arguments based on extended quasi-likelihood. Conditional on $\mathbf{b}$, the extended quasi-likelihood is (Nelder and Pregibon (1987))

$$q_y(\mathbf{b}) = -\frac{\|\mathbf{Y} - \tilde{\mathbf{X}}\mathbf{f} - \mathbf{Z}\mathbf{b}\|^2}{2\sigma^2} - \frac{n}{2}\log\sigma^2.$$

The marginal extended quasi-likelihood is

$$\begin{aligned} & l(\mathbf{Y}; \mathbf{f}, \sigma^2, \boldsymbol{\theta}) \\ & = \log \int \exp\{q_y(\mathbf{b})\} dF(\mathbf{b}, \boldsymbol{\theta}) \\ & = \log \int \exp\{q_y(0)\} \left[1 + \dot{q}_y(0)\mathbf{b} + \frac{1}{2}\mathbf{b}^{\mathrm{T}} \left\{\dot{q}_y(0)\dot{q}_y^{\mathrm{T}}(0) + \ddot{q}_y(0)\right\} \mathbf{b} + \Omega\right] dF(\mathbf{b}, \boldsymbol{\theta}) \end{aligned}$$

$$= \log[\exp\{q_y(0)\}] + \log\left(1 + \frac{\sigma^2}{2} tr\left[\{\dot{q}_y(0)\dot{q}_y^{\rm T}(0) + \ddot{q}_y(0)\}\mathbf{D}\right] + R\right)$$

$$\approx q_y(0) + \frac{\sigma^2}{2} tr\left[\{\dot{q}_y(0)\dot{q}_y^{\rm T}(0) + \ddot{q}_y(0)\}\mathbf{D}\right] + R, \tag{3.2}$$

where F is the distribution of $\mathbf{b}$, $\dot{q}_y(0) = \{\partial q_y(\mathbf{b})/\partial\mathbf{b}\}\mid_{\mathbf{b}=0}$, $\ddot{q}_y(0) = \{\partial^2 q_y(0)/\partial\mathbf{b}\partial\mathbf{b}^{\rm T}\}\mid_{\mathbf{b}=0}$, and the residual Ω contains the third and higher order terms of $\mathbf{b}$. We derive test statistics based on approximating the extended quasi-likelihood using the first two terms in (3.2). This approximation is appropriate for the normal case since R is of $o(\|\boldsymbol{\theta}\|)$. In the general situation, R may not be of $o(\|\boldsymbol{\theta}\|)$. We emphasize that both extended quasi-likelihood and its approximation are used for motivating test statistics only, neither of them is required in the theoretical development.

Let $\tilde{\boldsymbol{\epsilon}} = \mathbf{Y} - \tilde{\mathbf{X}}\mathbf{f}$. It is easy to check that $q_y(0) = -\|\tilde{\boldsymbol{\epsilon}}\|^2/(2\sigma^2) - n\log\sigma^2/2$, $\dot{q}_y(0) = \mathbf{Z}^{\rm T}\tilde{\boldsymbol{\epsilon}}/\sigma^2$ and $\ddot{q}_y(0) = -\mathbf{Z}^{\rm T}\mathbf{Z}/\sigma^2$. Then we have the approximate marginal extended quasi-likelihood

$$l_a(\mathbf{Y};\mathbf{f},\sigma^2,\boldsymbol{\theta}) = -\frac{\|\tilde{\boldsymbol{\epsilon}}\|^2}{2\sigma^2} + \frac{\tilde{\boldsymbol{\epsilon}}^{\rm T}\mathbf{Z}\mathbf{D}\mathbf{Z}^{\rm T}\tilde{\boldsymbol{\epsilon}}}{2\sigma^2} - \frac{1}{2}tr(\mathbf{Z}\mathbf{D}\mathbf{Z}^{\rm T}) - \frac{1}{2}n\log\sigma^2. \tag{3.3}$$

Let $\mathbf{U}_\theta(\mathbf{f},\sigma^2,\boldsymbol{\theta}) = \partial l_a/\partial\boldsymbol{\theta}$ be the quasi-score. It is easy to verify that the $i^{\rm th}$ element of $\mathbf{U}_\theta(\mathbf{f},\sigma^2,\boldsymbol{\theta})$ is $\tilde{\boldsymbol{\epsilon}}^{\rm T}\mathbf{Q}_i\tilde{\boldsymbol{\epsilon}}/(2\sigma^2) - tr\mathbf{Q}_i/2$, where $\mathbf{Q}_i = \mathbf{Z}\dot{\mathbf{D}}_i\mathbf{Z}^{\rm T}$ and $\dot{\mathbf{D}}_i = \partial\mathbf{D}/\partial\theta_i$.

Lemma 1. *Under H_0 and conditions* C.1, C.2, C.3(i), C.4, *and* C.5 *in Appendix B,*

$$n^{-1/2}\boldsymbol{U}_\theta(\hat{\boldsymbol{f}}_0,\hat{\sigma}_0^2,0) \xrightarrow{d} N(0,\boldsymbol{M}) \quad and \quad n^{-1/2}\boldsymbol{U}_\theta(\hat{\boldsymbol{f}},\hat{\sigma}^2,0) \xrightarrow{d} N(0,\boldsymbol{M}),$$

where $\boldsymbol{M} = \{2\sigma^4\boldsymbol{M}_0 + (\kappa - 3\sigma^4)\boldsymbol{M}_z\}/(4\sigma^4)$, $\lim_{n\to\infty} n^{-1}\boldsymbol{M}_{n0} = \boldsymbol{M}_0$, $\lim_{n\to\infty} n^{-1}\boldsymbol{M}_{nz} = \boldsymbol{M}_z$,

$$\boldsymbol{M}_{n0} = \{(\boldsymbol{M}_{n0})_{i,j}\}_{d\times d}, \quad (\boldsymbol{M}_{n0})_{i,j} = tr(\boldsymbol{Q}_i\boldsymbol{Q}_j) - \frac{(tr\,\boldsymbol{Q}_i)(tr\,\boldsymbol{Q}_j)}{n}, \tag{3.4}$$

$$\boldsymbol{M}_{nz} = \{(\boldsymbol{M}_{nz})_{i,j}\}_{d\times d}, \quad (\boldsymbol{M}_{nz})_{i,j} = tr\{diag(\boldsymbol{Q}_i)diag(\boldsymbol{Q}_j)\} - \frac{(tr\,\boldsymbol{Q}_i)(tr\,\boldsymbol{Q}_j)}{n}.$$

Remark 1. (i) Up to a scalar constant, $\mathbf{M}_{n0}$ in (3.4) is the same as the efficient information matrix $\mathbf{I}_\theta$ of $\boldsymbol{\theta}$ in Zhu and Fung (2004). (ii) The form of quasi-score $\mathbf{U}_\theta(\mathbf{f},\sigma^2,0)$ is first computed under the assumption that the off-diagonal elements of $\mathbf{D}_1$ are known. This quasi-score is used to construct the test statistics for the general case where the off-diagonal elements may be unknown. We note that the assumption about the off-diagonal entries was used for test statistics construction rather than for theoretical development. Specifically, conclusions in Lemma 1 and the theorems hold when unknown off-diagonal entries are replaced by their estimates.

We can construct a quasi-score test based on Lemma 1. Note that $\mathbf{M}_n \equiv \{2\sigma^4\mathbf{M}_{n0}+(\kappa-3\sigma^4)\mathbf{M}_{nz}\}/(4\sigma^4)$. The dependence of $\mathbf{M}_n$ on σ^2 and κ is expressed explicitly as $\mathbf{M}_n(\sigma^2,\kappa)$. Estimators of σ^2 are given in Section 2. We construct two estimators for κ, $\widehat{\kappa}_0$, and $\widehat{\kappa}$, where the definitions of $\widehat{\kappa}_0$ and $\widehat{\kappa}$ are given in Appendix A; $\widehat{\kappa}_0$ is consistent under the null and $\widehat{\kappa}$ is consistent under both the null and alternative (Lemma 3 in Appendix C).

We now consider a one-sided test as in Verbeke and Molenberghs (2003) and Bolfarine and Valenca (2005). From Lemma 1, extending one-sided test based on univariate normal distribution to the multivariate case, we consider the one-sided test statistic

$$T_{n0} = \max_{1\le i\le d}\left\{\mathbf{l}_i^{\mathrm{T}}\mathbf{M}_n^{-1/2}(\hat{\sigma}_0^2,\hat{\kappa}_0)\mathbf{U}_\theta(\hat{\mathbf{f}}_0,\hat{\sigma}_0^2,0)\right\}, \tag{3.5}$$

where $\mathbf{l}_i$ is the d-dimensional vector with the i^{th} element being 1 and all other elements being 0. Since T_{n0} in (3.5) is constructed using estimators that are consistent under the null, the test based on T_{n0} is referred to as the standard quasi-score test (SQT). Note that $\hat{\sigma}_0^2$ is inflated when the alternative is true (Lemma 2). Thus using such an estimator may deteriorate the power of the test; see relevant discussions in Chen, Härdle, and Li (2003) and Koul and Song (2008). The following test statistic is then derived by replacing $\mathbf{f}$, σ^2, and κ with estimators that are consistent under both the null and the alternative,

$$T_n = \max_{1\le i\le d}\left\{\mathbf{l}_i^{\mathrm{T}}\mathbf{M}_n^{-1/2}(\hat{\sigma}^2,\hat{\kappa})\mathbf{U}_\theta(\hat{\mathbf{f}},\hat{\sigma}^2,0)\right\}. \tag{3.6}$$

The test statistic T_n is referred to as the extended quasi-score test (EQT).

Theorem 1. *Suppose that conditions* C.1, C.2, C.3(i)*, and* C.4–C.8 *in Appendix B hold. Then under H_0, T_{n0} and T_n have an asymptotic distribution $\Phi^d(\cdot)$ as $n\to\infty$, where $\Phi(\cdot)$ is the standard normal distribution function.*

It is easy to see that the α upper quantile C_α of the asymptotic distribution is the $1-(1-\alpha)^{1/d}$ upper quantile of the standard normal distribution.

To investigate the power of the proposed SQT and EQT, we consider a sequence of alternative hypotheses indexed by n as $H_{An}:\ \boldsymbol{\theta} = c_n\boldsymbol{\theta}_0$ for some fixed nonzero and non-negative vector $\boldsymbol{\theta}_0$. As $n\to\infty$, H_{An} is a global alternative when c_n is bounded away from zero or a local alternative when c_n converges to zero.

Theorem 2. *Under H_{An} and conditions* C.1$\sim$ C.8 *in Appendix B, if* $\lim_{n\to\infty} n^{\alpha_0}c_n = k_0$ *where α_0 is defined in condition C.3 in Appendix B and k_0 is a fixed* *constant, then* T_n *and* T_{n0} *have an asymptotic distribution* $\Pi_{i=1}^{d}\Phi(x-m_i)$ *where* m_i *is the* i^{th} *component of* $k_0\boldsymbol{M}^{-1/2}\boldsymbol{\omega}/2$ *and* $\boldsymbol{\omega}=(\omega_1,\ldots,\omega_d)^{\text{T}}$*, with* ω_i *defined in* C.3 *in Appendix B.*

Remark 2. Condition C.3 in Appendix B implies that the $m_i's$ are nonnegative and that there exists at least one $m_i>0$. Consequently the mean of the asymptotic distribution of T_n is positive under H_{An} and large values of T_n support H_{An}. The asymptotic null distributions of T_{n0} and T_n do not depend on convergence rates of the alternatives to the null. In addition, based on Conditions C.3 and C.6(i), we have $\alpha_0=1/2$ when $q=1$. In fact, $\alpha_0=1/2$ is the key for C.3. Thus the tests T_{n0} and T_n can both detect the local alternatives approaching the null at rates up to root n.

For comparison, we now adapt the score-test in Silvapulle and Silvapulle (1995) and LRT test in Fan, Zhang, and Zhang (2001). Following the argument in Silvapulle and Silvapulle (1995) we have the SST statistic

$$\begin{aligned}SST=&\mathbf{U}_{\tilde\theta}^{\text{T}}(\hat{\mathbf{f}}_0,\hat\sigma_0^2,0)\tilde{\mathbf{M}}_n^{-1}(\hat\sigma_0^2,\hat\kappa_0)\mathbf{U}_{\tilde\theta}(\hat{\mathbf{f}}_0,\hat\sigma_0^2,0)\\&-\inf\{(\mathbf{U}_{\tilde\theta}(\hat{\mathbf{f}}_0,\hat\sigma_0^2,0)-\tilde{\boldsymbol{\theta}})^{\text{T}}\tilde{\mathbf{M}}_n^{-1}(\hat\sigma_0^2,\hat\kappa_0)(\mathbf{U}_{\tilde\theta}(\hat{\mathbf{f}}_0,\hat\sigma_0^2,0)-\tilde{\boldsymbol{\theta}}):\tilde{\boldsymbol{\theta}}\in\mho\},\end{aligned}$$

where $\mathbf{U}_{\tilde\theta}(\mathbf{f},\sigma^2,\tilde{\boldsymbol{\theta}})=\partial l_a/\partial\tilde{\boldsymbol{\theta}}$ with l_a defined in (3.3), $\tilde{\boldsymbol{\theta}}$ includes all free parameters in $\mathbf{D}_1$, and $\mho$ is the parametric space of $\tilde{\boldsymbol{\theta}}$, $\tilde{\mathbf{M}}_n(\hat\sigma_0^2,\hat\kappa_0)=\{2\hat\sigma_0^4\tilde{\mathbf{M}}_{n0}+(\hat\kappa_0-3\hat\sigma_0^4)\tilde{\mathbf{M}}_{nz})/(4\hat\sigma_0^4\}$ with $\tilde{\mathbf{M}}_{n0}$ and $\tilde{\mathbf{M}}_{nz}$ matrices with the $(i,j)^{\text{th}}$ elements $tr(\mathbf{Q}_i\mathbf{Q}_j)-n^{-1}(tr\mathbf{Q}_i)(tr\mathbf{Q}_j)$ and $tr\{\text{diag}(\mathbf{Q}_i)\text{diag}(\mathbf{Q}_j)\}-n^{-1}(tr\mathbf{Q}_i)(tr\mathbf{Q}_j)$, respectively. Note that the first part of SST is the standard two-sided score test statistic.

Following arguments in the proofs of Lemma 1 and Theorem 3, under H_0 and $H_{\tilde{A}n}:\ \tilde{\boldsymbol{\theta}}=n^{-1/2}\tilde{\boldsymbol{\theta}}_0$ for some fixed nonzero vector $\tilde{\boldsymbol{\theta}}_0$, we have

$$n^{-1/2}\mathbf{U}_{\tilde\theta}^{\text{T}}(\hat{\mathbf{f}}_0,\hat\sigma_0^2,0)\xrightarrow{d}\text{N}(0,\tilde{\mathbf{M}}),\quad\text{and}\quad n^{-1/2}\mathbf{U}_{\tilde\theta}^{\text{T}}(\hat{\mathbf{f}}_0,\hat\sigma_0^2,0)\xrightarrow{d}\text{N}(\frac{\tilde{\boldsymbol{\omega}}}{2},\tilde{\mathbf{M}}),$$

where $\tilde{\mathbf{M}}=\lim_{n\to\infty}n^{-1}\tilde{\mathbf{M}}_n$ and the i^{th} element of $\tilde{\boldsymbol{\omega}}$ is $\lim_{n\to\infty}n^{-1}\{tr(\tilde{\mathbf{Q}}_0\mathbf{Q}_i)-n^{-1}tr\tilde{\mathbf{Q}}_0tr\mathbf{Q}_i\}$ for $i=1,\ldots,d$, d is the dimension of $\tilde{\boldsymbol{\theta}}$, and $\tilde{\mathbf{Q}}_0=\mathbf{Z}\tilde{\mathbf{D}}_0\mathbf{Z}^{\text{T}}$ with $\tilde{\mathbf{D}}_0$ the covariance matrix corresponding to $\tilde{\boldsymbol{\theta}}_0$. Applying the results in Silvapulle and Silvapulle (1995, p.180) the asymptotic null distribution of the SST is a mixture of χ^2-distributions $\sum_{l=0}^{d}\omega_i(d,\tilde M,\mathcal{A})\chi_l^2$, where $\mathcal{A}$ is the approximating cone of $\mho$ defined in Section 4.7 in Silvapulle and Sen (2005) As pointed out by Silvapulle and Sen (2005) the exact computation of this distribution is difficult. The critical value is usually computed by Monte Carlo (Silvapulle and Sen (2005, p.78)).

With Gaussian assumptions for both the random effects and random errors, we have $\mathbf{Y}\sim\text{N}(\tilde{\mathbf{X}}\mathbf{f},\sigma^2\mathbf{I})$ under H_0 and $\mathbf{Y}\sim\text{N}(\tilde{\mathbf{X}}\mathbf{f},\sigma^2\mathbf{V})$ under H_A. Let $l(H_0)$

and $l(H_A)$ be log-likelihoods under H_0 and H_A, respectively. As in Fan, Zhang, and Zhang (2001), we define the MLE of $f_1(\cdot), \ldots, f_p(\cdot)$, and σ^2 under H_0 as the minimizers of $-l(H_0)$, and the MLE of $f_1(\cdot), \ldots, f_p(\cdot)$, σ^2 and $\mathbf{D}_1$ under H_A as the minimizers of $-l(H_A)$, both subject to the constraints $\int_0^1 \{f_l''(t)\}^2 dt \le C_l$, $l = 1, \ldots, p$. By introducing Lagrange multipliers, the estimates of σ^2 are $\hat{\sigma}_0^2 = (\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}}_0)^T(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}}_0)/n$ under H_0 and $\hat{\sigma}^2 = (\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}})^T\mathbf{V}^{-1}(\mathbf{Y} - \tilde{\mathbf{X}}\hat{\mathbf{f}})/n$ under H_A. Similar to Fan, Zhang, and Zhang (2001), we define the LRT statistic as

$$LRT = 2\{l(H_A) - l(H_0)\} = n \log \hat{\sigma}_0^2 - n \log \hat{\sigma}^2 - \log |\mathbf{I} + \mathbf{Z}\hat{\mathbf{D}}\mathbf{Z}^T|,$$

where $\hat{\mathbf{D}}$ is the MLE of $\mathbf{D}$. We note that the LRT in Fan, Zhang, and Zhang (2001) was not initially designed for testing for variance components in VCMM. Unlike theirs, we estimate nonparametric functions using smoothing splines and test hypothesis about variance components.

Theoretical properties and asymptotic null distributions of the LRT are beyond the scope of the current article. The Bootstrap method can be used to determine the critical value of the LRT. An alternative approach to constructing a LRT is to use the connection between smoothing splines and linear mixed-effects models (Wang (1998); Liu and Wang (2004); Crainiceanu et al. (2005)). This alternative approach merits future research.

4. Simulations

We conducted simulations to evaluate the finite sample performance of the SQT and EQT, and to compare them with the SST and LRT. Data were generated from the model

$$y_{ij} = x_{ij1} f_1(t_{ij}) + x_{ij2} f_2(t_{ij}) + Z_{ij}^{\mathrm{T}} b_i + \epsilon_{ij}, \quad i = 1, \ldots, m;\ j = 1, \ldots, 10, \quad (4.1)$$

where $t_{ij} = trun[\{i + (m/5 - 1)\}/(m/5)]/50 + 0.10(j - 1)$, $trun$ denotes the truncation operator, the x_{ij1} are independently Uniform$[t_{ij}/10, i + t_{ij}/10]$, the x_{ij2} are independently $\mathrm{N}(10t_{ij}, 0.60^2)$, $f_1(t) = t^2 + 2t$, and $f_2(t) = \cos(\pi t)$. We considered two cases of random errors: $\epsilon_{ij} \overset{i.i.d.}{\sim} N(0, 1)$ and $\epsilon_{ij} \overset{i.i.d.}{\sim} (\Gamma(2, 1) - 2)/\sqrt{2}$, where $\Gamma(2, 1)$ is the Gamma distribution with shape parameter 2 and scale parameter 1, standardized to have mean zero and variance 1; $Z_{ij} = 1$ and the random effects $\mathbf{b}_i$ are i.i.d. samples from

- Normal $\mathrm{N}(0, \theta)$ with $\theta = 0, 0.04, 0.07$ and 0.10.
- Normal mixture $0.25\mathrm{N}(-0.75\varphi, \phi^2) + 0.75\mathrm{N}(0.25\varphi, \phi^2)$, with four combinations of φ and ϕ such that the variance $\theta = \phi^2 + 0.1875\varphi^2$ takes on the same four values 0, 0.04, 0.07, and 0.10 as above; the four values for φ are $0, 0.05, 0.20$, and 0.70, respectively.

The null hypothesis for this setting is $H_0 : \theta = 0$. We considered three sample sizes $m = 50, 75, 100$ and repeated the simulation 500 times. Cubic B-splines were used to estimate the $f_l(\cdot)'s$; the smoothing parameters were selected by the extended GCV method; the significance level was set to be 0.05.

Critical values of SQT and EQT are calculated from their asymptotic distributions, the critical values of SST are calculated by the recommended Monte Carlo method (with 10,000 replication) in Silvapulle and Sen (2005) according to the asymptotic distribution, and the critical values of LRT are calculated by the bootstrap method (with 2,000 the bootstrap sample size). The estimates for the LRT are computed using the EM algorithm in Laird and Ware (1982) and Laird, Lange, and Stram (1987). Powers of all four tests are listed in Tables 1 and 2. We conclude that all four tests performed well: type I errors were close to the nominal level and powers approached 1 quickly. Power increased with the increase of variance components and/or the increase of sample size m. As expected, the EQT was more powerful than SQT, SST, and LRT. No test was uniformly better between SQT and SST. The LRT did not perform better than others in the normal case for the following possible reasons: the LRT was not initially designed for testing for variance components in VCMM; the smoothing parameters were not selected to optimize the performance of LRT as in Fan, Zhang, and Zhang (2001). Further research on LRT is necessary. It is interesting to note that, even though derived under normality assumptions, the LRT performed well for non-Gaussian case; this agrees with the comments made in Fan, Zhang, and Zhang (2001).

To further investigate the performance of EQT and SQT, we conducted another simulation with bivariate random effects $\mathbf{b}_i = (b_{i1}, b_{i2})^{\mathrm{T}}$ and $\mathbf{Z}_{ij} = (1, t_{ij})^{\mathrm{T}}$. Let $\mathrm{Var}(b_{i1}) = \theta_1$, $\mathrm{Var}(b_{i2}) = \theta_2$, and $\mathrm{Cor}(b_{i1}, b_{i2}) = \rho$. The random effects $\mathbf{b}_i$ are generated from a bivariate normal mixture, $0.25\mathrm{N}(-0.75\varphi, D) + 0.75\mathrm{N}(0.25\varphi, D)$, with six combinations of φ and D such that $(\theta_1, \theta_2, \rho)$ equals $(0, 0, 0)$, $(0.02, 0.03, 0)$, $(0.02, 0.03, 0.30)$, $(0.03, 0.04, -0.50)$, $(0.03, 0.04, 0)$, and $(0.03, 0.04, 0.70)$. The choices of φ corresponding to these six combinations are $(0, 0)^{\mathrm{T}}$, $(0, 0)^{\mathrm{T}}$, $(0, 0)^{\mathrm{T}}$, $(0.10, 0.20)^{\mathrm{T}}$, $(0.10, 0.20)^{\mathrm{T}}$, and $(0, 0)^{\mathrm{T}}$. Random errors are generated from the standard normal distribution. The null hypothesis for this two-dimensional setting is $H_0 : \theta_1 = \theta_2 = 0$. Smoothing parameters are selected by the extended GCV method. Since more parameters are involved in the two-dimensional case, three sample sizes $m = 50, 100, 200$ are considered. The simulation was repeated 1,000 times and the results are listed in Table 3. Type I errors got closer to the nominal level 0.05 with the increase of sample size. The EQT was more powerful than SQT.

To further evaluate the accuracy of asymptotic distributions for the EQT and SQT tests, we computed their distributions with unknown parameters calculated from the simulated data. Figure 1 shows the theoretical and asymptotic

Table 1. Powers of SQT, EQT, SST, and LRT with standard normal random errors.

distribution	θ	tests	m 50	75	100
Normal	0.00	EQT(SQT)	0.054(0.038)	0.062(0.048)	0.056(0.044)
		LRT(SST)	0.036(0.044)	0.040(0.056)	0.044(0.046)
	0.04	EQT(SQT)	0.518(0.420)	0.656(0.562)	0.820(0.790)
		LRT(SST)	0.282(0.424)	0.346(0.566)	0.610(0.800)
	0.07	EQT(SQT)	0.812(0.746)	0.940(0.904)	0.970(0.940)
		LRT(SST)	0.608(0.740)	0.792(0.904)	0.810(0.960)
	0.10	EQT(SQT)	0.950(0.932)	0.986(0.980)	1.000(0.990)
		LRT(SST)	0.822(0.930)	0.912(0.982)	0.940(0.990)
Normal Mixture	0.00	EQT(SQT)	0.064(0.046)	0.060(0.048)	0.054(0.048)
		LRT(SST)	0.046(0.046)	0.042(0.044)	0.044(0.048)
	0.04	EQT(SQT)	0.538(0.448)	0.654(0.574)	0.736(0.668)
		LRT(SST)	0.324(0.442)	0.410(0.508)	0.610(0.670)
	0.07	EQT(SQT)	0.814(0.724)	0.910(0.888)	0.974(0.942)
		LRT(SST)	0.576(0.700)	0.774(0.860)	0.840(0.920)
	0.10	EQT(SQT)	0.946(0.932)	0.988(0.982)	1.000(1.000)
		LRT(SST)	0.760(0.932)	0.948(0.982)	0.990(1.000)

Table 2. Powers of SQT, EQT, SST, and LRT with Gamma random errors.

distribution	θ	tests	m 50	75	100
Normal	0.00	EQT(SQT)	0.066(0.048)	0.045(0.040)	0.056(0.046)
		LRT(SST)	0.034(0.060)	0.040(0.042)	0.038(0.056)
	0.04	EQT(SQT)	0.568(0.480)	0.654(0.578)	0.774(0.712)
		LRT(SST)	0.330(0.484)	0.450(0.554)	0.552(0.746)
	0.07	EQT(SQT)	0.842(0.780)	0.922(0.900)	0.972(0.950)
		LRT(SST)	0.600(0.718)	0.712(0.880)	0.836(0.966)
	0.10	EQT(SQT)	0.934(0.908)	0.996(0.980)	0.996(0.996)
		LRT(SST)	0.748(0.910)	0.904(0.984)	0.952(0.996)
Normal Mixture	0.00	EQT(SQT)	0.058(0.048)	0.045(0.040)	0.056(0.046)
		LRT(SST)	0.034(0.060)	0.040(0.042)	0.038(0.056)
	0.04	EQT(SQT)	0.542(0.466)	0.640(0.554)	0.736(0.678)
		LRT(SST)	0.324(0.468)	0.414(0.562)	0.558(0.678)
	0.07	EQT(SQT)	0.850(0.772)	0.928(0.900)	0.970(0.954)
		LRT(SST)	0.576(0.774)	0.766(0.900)	0.840(0.960)
	0.10	EQT(SQT)	0.950(0.926)	0.988(0.980)	0.996(0.996)
		LRT(SST)	0.768(0.926)	0.932(0.980)	0.946(0.996)

null distributions for normal errors. As expected, as m increases, the asymptotic distributions get closer to the theoretical distributions. Furthermore, as suggested by Theorem 1, the densities for the one-dimensional case are symmetric

Table 3. Powers of EQT and SQT in the two-dimensional case with normal random errors.

$(\theta_1, \theta_2, \rho)$	tests	m		
		50	100	200
(0,0,0)	EQT(SQT)	0.069(0.059)	0.058(0.045)	0.053(0.047)
(0.02,0.03,0)	EQT(SQT)	0.347(0.218)	0.541(0.361)	0.739(0.551)
(0.02,0.03,0.30)	EQT(SQT)	0.451(0.283)	0.658(0.446)	0.893(0.734)
(0.03,0.04,-0.50)	EQT(SQT)	0.372(0.221)	0.575(0.382)	0.775(0.566)
(0.03,0.04,0)	EQT(SQT)	0.582(0.393)	0.827(0.644)	0.976(0.861)
(0.03,0.04,0.70)	EQT(SQT)	0.778(0.559)	0.967(0.831)	1.000(0.973)

and the densities for the two-dimensional case are skewed.

5. Application

We analyzed a subset of the Multi-Center AIDS Cohort Study that includes repeated measurements from 283 homosexual men who were infected with HIV during the period between 1984 and 1991. Not all subjects were observed at a common set of time points due to missed visits and random infection time. The number of repeated measurements per subject ranged from 1 to 14, with a median of 6 and a mean of 6.4205. The number of distinct measurement time points was 59. Further details can be found in Kaslow et al. (1987)

As did Wu and Chiang (2000) and Huang, Wu, and Zhou (2004), we investigated the effects of cigarette smoking, pre-HIV infection CD4 percentage and age at HIV infection on the mean CD4 percentage after the infection. Let t_{ij} be the time in years of the j^{th} measurements from subject i after HIV infection, y_{ij} be the i^{th} subject's CD4 percentage at time t_{ij}, x_{ij1} be smoking status (1 if the i^{th} subject always smokes cigarettes and 0 if he never smokes cigarettes), x_{ij2} be the i^{th} individual's centered age at HIV infection computed by subtracting the sample average age at infection from the i^{th} subject's age at infection, and x_{ij3} be the i^{th} subject's centered pre-infection CD4 percentage computed by subtracting the average pre-infection CD4 percentage of the sample from the i^{th} subject's actual pre-infection CD4 percentage. We fit the VCMM

$$y_{ij} = f_0(t_{ij}) + x_{ij1}f_1(t_{ij}) + x_{ij2}f_2(t_{ij}) + x_{ij3}f_3(t_{ij}) + \mathbf{Z}_{ij}^{\mathrm{T}}\mathbf{b}_i + \epsilon_{ij}, \qquad (5.1)$$

where ϵ_{ij}'s are i.i.d. random error, $\mathbf{Z}_{ij} = (1, t_{ij})^{\mathrm{T}}$, and the $\mathbf{b}_i$'s are i.i.d. random effects with mean zero and an unstructured covariance matrix.

Cubic splines were used to estimate the varying coefficients and the smoothing parameters were selected by the extended GCV method. Figure 2 shows the estimators of coefficient functions and their bootstrap confidence intervals based on 200 bootstrap replications. The conclusions about the fixed effects are

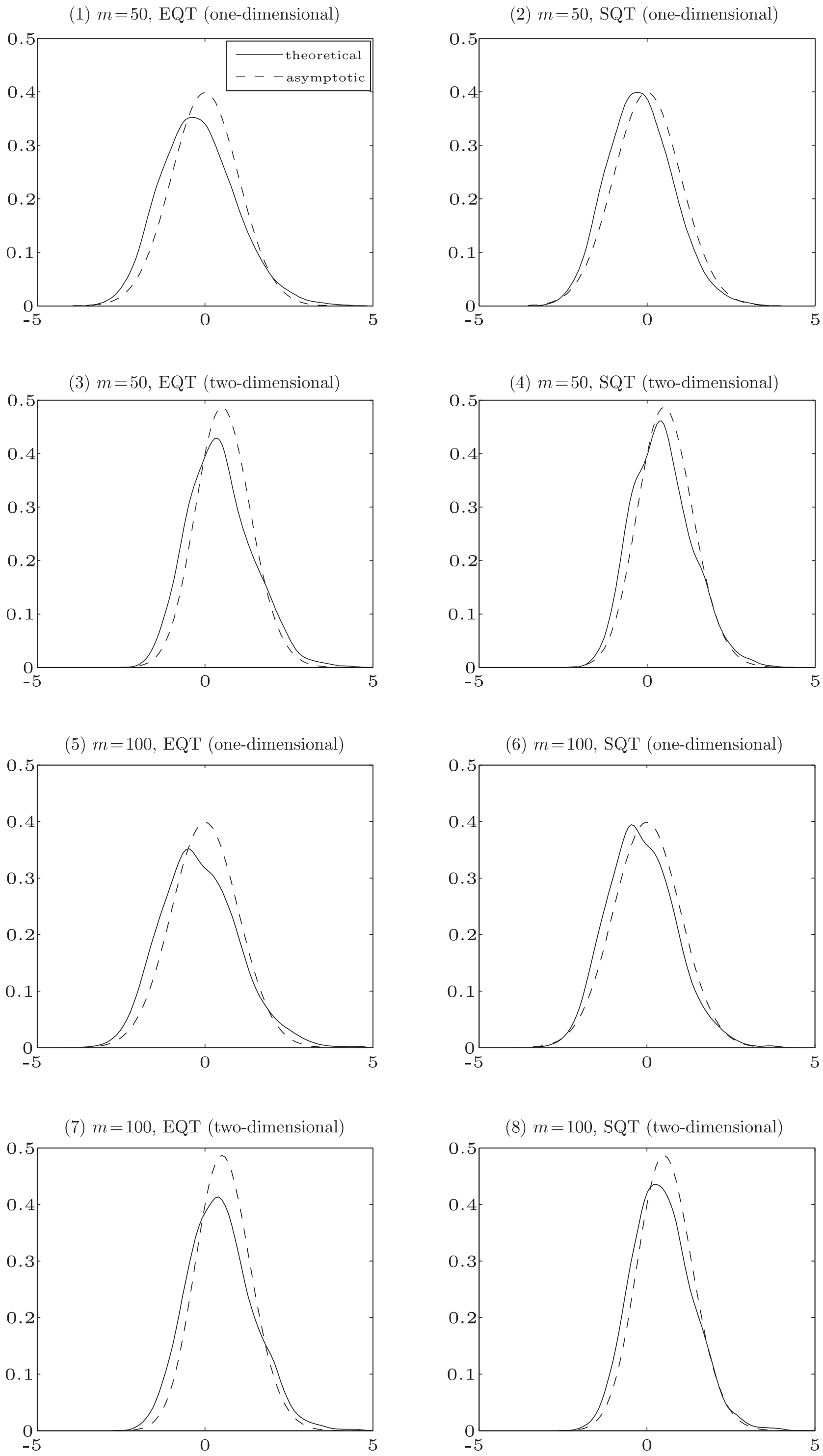

Figure 1. Plots of simulated null distributions in solid lines and asymptotic null distributions in dashed lines.

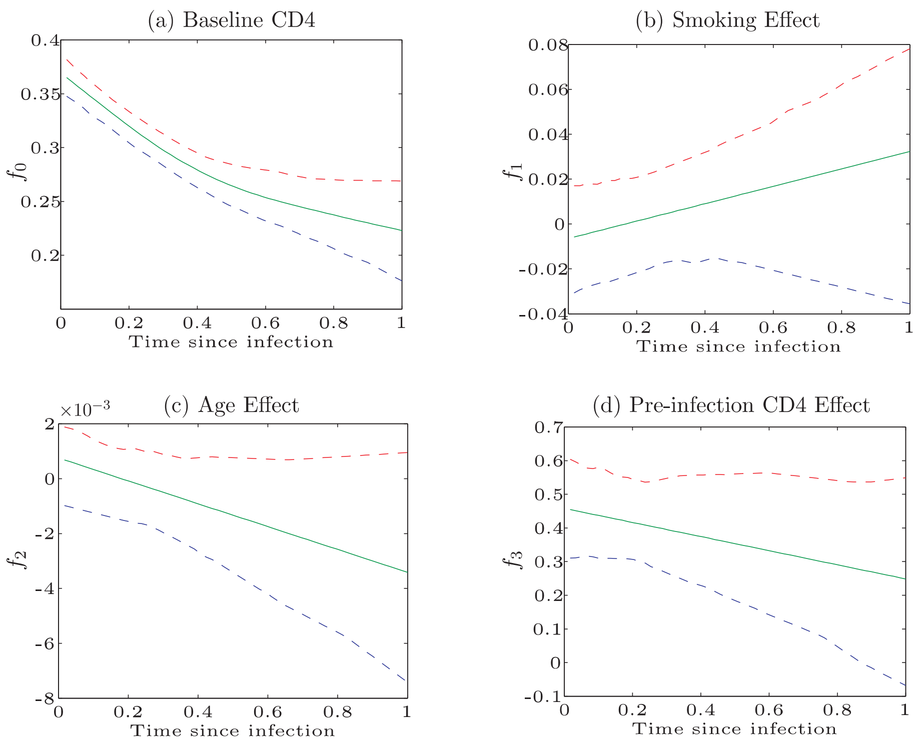

Figure 2. Estimated coefficient functions and their 0.95 bootstrap pointwise confidence intervals.

the same as those in Wu and Chiang (2000) and Huang, Wu, and Zhou (2004): effects of smoking and age of HIV infection are not statistically significant, the baseline CD4 percentage of the population depletes over time with a rate that slows down gradually, and pre-infection CD4 percentage is positively associated with high post-infection percentage, which effect does not change over time.

We checked whether the VCMM (5.1) could be reduced to a VCM by testing the hypothesis $H_0 : \theta_1 = \theta_2 = 0$. The SQT, EQT, SST and LRT statistics were SQT = 40.95, EQT = 42.15, SST = 2479.26, and LRT = 914.18, with P-values approximately zero for all tests. With the null hypothesis of homogeneity rejected, we find the VCMM useful for predicting subject-specific trajectories.

6. Conclusions

We have proposed and compared four one-sided tests for the null hypothesis that all variance components are zero in a VCMM. All tests performed well in simulations, with the EQT more powerful than SQT, SST and LRT. The SQT, EQT and SST are robust in the sense that they do not require specifying distributions for random effects and random errors. The LRT test performed well even when the distributions for random errors and random effects were non-Gaussian. The asymptotic null distributions of SQT and EQT are simple and easy to use. The proposed estimation method is new and is itself of interest. In particular, the development of consistent estimators for the fourth moment is technically difficult.

When the null hypothesis that all variance components are zero is rejected, it is of interest to test whether some of the variance components are zero. We will investigate this problem in our future research. Another interesting future research topic is to extend the proposed quasi-score tests to the setting of generalize linear models.

Acknowledgement

The authors thank Colin O. Wu for allowing us to use the “MACS Public Use Data Set Release PO4 (1984-1991)”. We also thank the Editor, An Associate editor, and two referees for constructive comments that substantially improved an earlier draft. The research was supported by a grant (HKBU2030/07P) from the Research Grants Council of Hong Kong, a grant from the National Natural Science Foundation of China (No.11001267, 11071253) and the Fundamental Research Funds for the Central Universities in China(NO.2009QS02); Wang’s research was supported by a grant from the National Science Foundation (DMS-0706886).

Appendix A: Notations

The Vec function stacks column vectors of a matrix; $\otimes$ denotes the Kronecker operation of two matrices or vectors. $\tilde{\mathbf{X}}_i$ is the submatrix of $\tilde{\mathbf{X}}$ with rows from $(n_1 + \cdots + n_{i-1} + 1)$ to $(n_1 + \cdots + n_i)$; $\mathbf{C}_{(ij)}$ denotes the $(i, j)^{\text{th}}$ element of a matrix $\mathbf{C}$. Let

$$\tilde{\boldsymbol{\Sigma}}_z = n^{-1} \sum_{i=1}^{m} \mathbf{Z}_i^{\mathrm{T}} \mathbf{Z}_i, \quad \tilde{\mathbf{Z}}_i = \mathbf{Z}_i \tilde{\boldsymbol{\Sigma}}_z^{-1/2}, \quad \tilde{\boldsymbol{\Sigma}}_{zz} = n^{-1} \sum_{i=1}^{m} (\mathbf{Z}_i^{\mathrm{T}} \mathbf{Z}_i) \otimes (\mathbf{Z}_i^{\mathrm{T}} \mathbf{Z}_i); \qquad \text{(A.1)}$$

$$\tilde{c}_{11} = n^{-1} \sum_{i=1}^{m} (\tilde{\mathbf{Z}}_i^{\mathrm{T}} \tilde{\mathbf{Z}}_i)_{11}^2, \; \tilde{c}_{12} = n^{-1} \sum_{i=1}^{m} (\tilde{\mathbf{Z}}_i^{\mathrm{T}} \tilde{\mathbf{Z}}_i)_{11} (\tilde{\mathbf{Z}}_i^{\mathrm{T}} \tilde{\mathbf{Z}}_i)_{22}, \; \tilde{c}_{13} = n^{-1} \sum_{i=1}^{m} (\tilde{\mathbf{Z}}_i^{\mathrm{T}} \tilde{\mathbf{Z}}_i)_{12}^2;$$

$$\begin{aligned} \hat{\mathbf{A}} &= n^{-1} \sum_{i=1}^{m} \tilde{\mathbf{Z}}_i^{\mathrm{T}} (\mathbf{Y}_i - \tilde{\mathbf{X}}_i \hat{\mathbf{f}}_0)(\mathbf{Y}_i - \tilde{\mathbf{X}}_i \hat{\mathbf{f}}_0)^{\mathrm{T}} \tilde{\mathbf{Z}}_i \\ &\quad - n^{-1} \sum_{i=1}^{m} (\mathbf{Y}_i - \tilde{\mathbf{X}}_i \hat{\mathbf{f}}_0)^{\mathrm{T}} (\mathbf{Y}_i - \tilde{\mathbf{X}}_i \hat{\mathbf{f}}_0) \mathbf{I}_q; \\ \hat{\mathbf{B}} &= \frac{\hat{\mathbf{A}}}{\tilde{c}_{12} + \tilde{c}_{13}} + \frac{\{2\tilde{c}_{13} - (\tilde{c}_{11} - \tilde{c}_{12})\} \mathrm{diag}(\hat{\mathbf{A}})}{(\tilde{c}_{11} - \tilde{c}_{13})(\tilde{c}_{12} + \tilde{c}_{13})} \end{aligned}$$

$$-\frac{(\tilde{c}_{13}-1)tr(\hat{\mathbf{A}})\mathbf{I}_q}{(\tilde{c}_{11}-\tilde{c}_{13})\{\tilde{c}_{11}-\tilde{c}_{13}+q(\tilde{c}_{13}-1)\}}; \tag{A.2}$$

$$\hat{\mathbf{I}}_0 = n^{-1}\sum_{i=1}^m \{(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}})^{\mathrm{T}}(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}})\}^2;$$

$$\hat{\mathbf{I}}_1 = \hat{\sigma}^4\{tr(\tilde{\mathbf{\Sigma}}_z\hat{\mathbf{D}}_1)+(n^{-1}\sum_{i=1}^m n_i^2-1)+2n^{-1}\sum_{i=1}^m n_i tr(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i\hat{\mathbf{D}}_1)\};$$

$$\hat{\mathbf{J}} = n^{-1}\sum_{i=1}^m Vec[Vec\{\mathbf{Z}_i^{\mathrm{T}}(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}})(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}})^{\mathrm{T}}\mathbf{Z}_i\}$$
$$\times Vec^{\mathrm{T}}\{\mathbf{Z}_i^{\mathrm{T}}(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}})(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}})^{\mathrm{T}}\mathbf{Z}_i\}];$$

$$\mathbf{J}_0 = n^{-1}\sum_{i=1}^m \{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}\otimes\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\};$$

$$\mathbf{J}_2 = n^{-1}\sum_{i=1}^m\sum_{j=1}^{n_i} Vec\{(\mathbf{Z}_{ij}\mathbf{Z}_{ij}^{\mathrm{T}})\otimes(\mathbf{Z}_{ij}\mathbf{Z}_{ij}^{\mathrm{T}})\};$$

$$\mathbf{R}_n^0 = Vec^{\mathrm{T}}(\tilde{\mathbf{\Sigma}}_{zz})Vec\left[E\left\{Vec(\mathbf{b}_1\mathbf{b}_1^{\mathrm{T}})Vec^{\mathrm{T}}(\mathbf{b}_1\mathbf{b}_1^{\mathrm{T}})\right\}\right]$$
$$+\sigma^4(6-2n^{-1}\sum_{i=1}^m n_i^2)tr(\tilde{\mathbf{\Sigma}}_z\mathbf{D}_1)$$
$$+\sigma^4[2n^{-1}\sum_{i=1}^m n_i tr(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i\mathbf{D}_1)-(n^{-1}\sum_{i=1}^m n_i^2-1)\{tr(\tilde{\mathbf{\Sigma}}_z\mathbf{D}_1)\}^2]. \tag{A.3}$$

Let $\mathbf{E}_{n_i}(l_1,l_2)$ be an $n_i\times n_i$ matrix with the $(l_1,l_2)^{\mathrm{th}}$ element 1 and others 0. Take

$$\mathbf{E}_{n_i^2} = \begin{pmatrix} \mathbf{I}_{n_i}-\mathbf{E}_{n_i}(1,1) & \mathbf{E}_{n_i}(1,2)+\mathbf{E}_{n_i}(2,1) & \cdots & \mathbf{E}_{n_i}(1,n_i)+\mathbf{E}_{n_i}(n_i,1) \\ \mathbf{E}_{n_i}(2,1)+\mathbf{E}_{n_i}(1,2) & \mathbf{I}_{n_i}-\mathbf{E}_{n_i}(2,2) & \cdots & \mathbf{E}_{n_i}(2,n_i)+\mathbf{E}_{n_i}(n_i,2) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{E}_{n_i}(n_i,1)+\mathbf{E}_{n_i}(1,n_i) & \mathbf{E}_{n_i}(n_i,2)+\mathbf{E}_{n_i}(2,n_i) & \cdots & \mathbf{I}_{n_i}-\mathbf{E}_{n_i}(n_i,n_i) \end{pmatrix},$$

where $\mathbf{I}_{n_i}$ is the $n_i\times n_i$ identity matrix. Let

$$\hat{\mathbf{J}}_1 = \hat{\sigma}^4 n^{-1}\sum_{i=1}^m(\mathbf{Z}_i^{\mathrm{T}}\otimes\mathbf{Z}_i^{\mathrm{T}})\otimes\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}Vec\{Vec(\hat{\mathbf{D}}_1)Vec^{\mathrm{T}}(\mathbf{I}_{n_i})\}$$
$$+\hat{\sigma}^4 n^{-1}\sum_{i=1}^m\{(\mathbf{Z}_i^{\mathrm{T}}\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}\otimes\{\mathbf{Z}_i^{\mathrm{T}}\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}Vec(\mathbf{I}_{n_i}\otimes\hat{\mathbf{D}}_1)$$
$$+\hat{\sigma}^4 n^{-1}\sum_{i=1}^m\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes\mathbf{Z}_i^{\mathrm{T}}\}\otimes\{\mathbf{Z}_i^{\mathrm{T}}\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}$$

$$\times Vec\{(\mathbf{I}_{n_i}\otimes\hat{\mathbf{d}}_1,\mathbf{I}_{n_i}\otimes\hat{\mathbf{d}}_2,\ldots,\mathbf{I}_{n_i}\otimes\hat{\mathbf{d}}_q)\}$$

$$+\hat{\sigma}^4 n^{-1}\sum_{i=1}^{m}\{\mathbf{Z}_i^{\mathrm{T}}\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}\otimes\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes\mathbf{Z}_i^{\mathrm{T}}\}$$

$$\times Vec\{(\mathbf{I}_{n_i}\otimes\hat{\mathbf{d}}_1,\mathbf{I}_{n_i}\otimes\hat{\mathbf{d}}_2,\ldots,\mathbf{I}_{n_i}\otimes\hat{\mathbf{d}}_q)^{\mathrm{T}}\}$$

$$+\hat{\sigma}^4 n^{-1}\sum_{i=1}^{m}\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes\mathbf{Z}_i^{\mathrm{T}}\}\otimes\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes\mathbf{Z}_i^{\mathrm{T}}\}Vec(\hat{\mathbf{D}}_1\otimes\mathbf{I}_{n_i})$$

$$+\hat{\sigma}^4 n^{-1}\sum_{i=1}^{m}\{(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\otimes(\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)\}\otimes(\mathbf{Z}_i^{\mathrm{T}}\otimes\mathbf{Z}_i^{\mathrm{T}})Vec\{Vec(\mathbf{I}_{n_i})Vec^{\mathrm{T}}(\hat{\mathbf{D}}_1)\}$$

$$+\hat{\sigma}^4 n^{-1}\sum_{i=1}^{m}(\mathbf{Z}_i^{\mathrm{T}}\otimes\mathbf{Z}_i^{\mathrm{T}})\otimes(\mathbf{Z}_i^{\mathrm{T}}\otimes\mathbf{Z}_i^{\mathrm{T}})Vec(\mathbf{E}_{n_i^2}),$$

where $\hat{\mathbf{d}}_l$ is the l^{th} column vector of $\hat{\mathbf{D}}_1$, $l=1,\ldots,q$. Define

$$\widehat{\kappa}_0 \triangleq n^{-1}\sum_{i=1}^{m}\Big\{(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}}_0)^{\mathrm{T}}(\mathbf{Y}_i-\tilde{\mathbf{X}}_i\hat{\mathbf{f}}_0)\Big\}^2-n^{-1}\sum_{i=1}^{m}n_i(n_i-1)\hat{\sigma}_0^4,$$

$$\widehat{\kappa}\triangleq\{1-Vec^{\mathrm{T}}(\tilde{\boldsymbol{\Sigma}}_{zz})\mathbf{J}_0^{-1}\mathbf{J}_2\}^{-1}\{\hat{\mathbf{I}}_0-\hat{\mathbf{I}}_1-Vec^{\mathrm{T}}(\tilde{\boldsymbol{\Sigma}}_{zz})\mathbf{J}_0^{-1}(\hat{\mathbf{J}}-\hat{\mathbf{J}}_1)\}.$$

Appendix B: Regularity Conditions

C.1 $\{n_i\}$ is a bounded sequence of positives integers, and there exists $1\le \upsilon<\infty$ such that $\upsilon\mathbf{I}_r-\mathbf{N}_l^{\mathrm{T}}\mathbf{N}_l$ is positive definite, $l=1,\ldots,p$.

C.2 The second derivatives of function $f_l(\cdot)$ are bounded, and the eigenvalues of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ are bounded.

C.3 (i) The eigenvalues of $\mathbf{Q}_i$ are bounded, and the $\mathbf{M}$ of Lemma 1 is positive definite.

(ii) The eigenvalues of $\mathbf{Q}_0$ are bounded, where $\mathbf{Q}_0=\mathbf{Z}\mathbf{D}_0\mathbf{Z}^{\mathrm{T}}$, $\mathbf{D}_0$ the covariance matrix of random effects $\mathbf{b}_i$ corresponding to $\boldsymbol{\theta}_0$ under the global alternative. In addition, there exists a positive constant α_0 such that

$$\lim_{n\to\infty}\frac{tr(\mathbf{Q}_0\mathbf{Q}_i)-n^{-1}tr\mathbf{Q}_0 tr\mathbf{Q}_i}{n^{\alpha_0+1/2}}=\omega_i,\ \frac{tr(\mathbf{Q}_0\mathbf{Q}_i)}{n^{\alpha_0+1/2}}=O(1),\ i=1,\ldots,d,$$

where ω_i are constants and not all ω_i are zero.

C.4 There exists a continuous strictly positive density function $w(\cdot)$ on [0,1] such that $\int_0^{t_i^0}w(t)dt=(2i-1)/2r$, $i=1,\ldots,r$.

C.5 For $l=1,\ldots,p$, $n\to\infty$ and $\lambda_l\to 0$ in such a way that $n\lambda_l\to\infty$; $n^{-1/2}\|(\mathbf{I}_n-\mathbf{H}_\lambda)\boldsymbol{\mu}\|^2\to 0$ as $n\to\infty$, where $\boldsymbol{\mu}=\tilde{\mathbf{X}}\mathbf{f}$ and $\mathbf{H}_\lambda=\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}}+\boldsymbol{\Lambda}\mathbf{K})^{-1}\tilde{\mathbf{X}}^{\mathrm{T}}$.

C.6 (i) There exists a $q \times q$ positive definite matrix $\mathbf{\Sigma}_z$ such that $\lim_{n\to\infty} \tilde{\mathbf{\Sigma}}_z = \mathbf{\Sigma}_z$.

(ii) There exist a $q^2 \times q^2$ non-negative definite matrix $\mathbf{\Sigma}_{zz}$, a $q^4 \times q^4$ positive definite matrix $\mathbf{J}_{00}$ and a q^4-dimensional vector $\mathbf{J}_{20}$ such that $\lim_{n\to\infty} \tilde{\mathbf{\Sigma}}_{zz} = \mathbf{\Sigma}_{zz}$, $\lim_{n\to\infty} \mathbf{J}_0 = \mathbf{J}_{00}$, $\lim_{n\to\infty} \mathbf{J}_2 = \mathbf{J}_{20}$, and $Vec^{\mathrm{T}}(\mathbf{\Sigma}_{zz})\mathbf{J}_{00}^{-1}\mathbf{J}_{20} < 1$.

C.7 There exist positive constants c_{11}, c_{12}, and c_{13} such that

$$\lim_{n\to\infty} n^{-1}\sum_{i=1}^{m}(\tilde{\mathbf{Z}}_i^{\mathrm{T}}\tilde{\mathbf{Z}}_i)_{l_1,l_1'}(\tilde{\mathbf{Z}}_i^{\mathrm{T}}\tilde{\mathbf{Z}}_i)_{l_2,l_2'}$$
$$= \begin{cases} c_{11}, 1 \le l_1 = l_1' = l_2 = l_2' \le q, \\ c_{12}, l_1 = l_1' \ne l_2 = l_2', 1 \le l_1, l_2 \le q, \\ c_{13}, l_1 = l_2 \ne l_1' = l_2', \quad or \quad l_1 = l_2' \ne l_1' = l_2, 1 \le l_1, l_1', l_2, l_2' \le q, \\ 0, \quad \text{otherwise.} \end{cases}$$

C.8 There exist a $\delta > 0$ such that the $(4+\delta)^{\mathrm{th}}$ moments of both $\mathbf{b}_i$ and ϵ_{ij} are finite.

The condition on sample sizes in condition C.1 is reasonable since it is common in longitudinal studies that the number of subjects is large and the number of observations for each subject is limited. In addition, condition C.1 ensures that, as $m \to \infty$, the number of distinct time points $r \to \infty$ and $r = O(n)$. Thus, under the condition C.1, $n \to \infty$ indicates $m \to \infty$ and the n_i are bounded. Conditions C.2 and C.4 are standard. As far as $\mathbf{M}_n$ and its limit $\mathbf{M}$ in C.3(i) are concerned, it is sufficient that $n^{-1}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\{Vec(\mathbf{Z}_{ij}\mathbf{Z}_{ij}^{\mathrm{T}})Vec^{\mathrm{T}}(\mathbf{Z}_{ij}\mathbf{Z}_{ij}^{\mathrm{T}})\}$ converges to some non-negative definite matrix. This condition can be seen as a natural extension of the second part of Condition 3 in Zhu and Fung (2004) when the normal distribution assumption is removed. The assumptions in Condition C.3(ii) are similar to those of Crainiceanu et al. (2005) for LMM with only one unknown variance component. Condition C.5 was also assumed by Chiang, Rice, and Wu (2001), and is similar to a condition in Eubank and Thomas (1993) and Zhu and Fung (2004). Condition C.6(i) is a commonly used condition for fixed design points. When $q = 1$, Condition C.6(ii) reduces to $\{n^{-1}\sum_{i=1}^{m}(\sum_{j=1}^{n_i} Z_{ij}^2)^2\}^2/\{n^{-1}\sum_{i=1}^{m}(\sum_{j=1}^{n_i} Z_{ij}^2)^4\} < 1$, which is always true as long as $m < n$. Condition C.7 is mild. As a special case, when $\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i$ is the identity matrix, $c_{11} = c_{12} = \lim_{n\to\infty} n/m$ and $c_{13} = 0$.

Appendix C: Theoretical Results and Sketch of Proofs

For simplicity, c is a generic constant which may be different in different places.

Lemma 2. *As* $n \to \infty$,

(i) *under conditions* C.1, C.4–C.6, *and* C.8, $\hat{\sigma}_0^2 \xrightarrow{p} \sigma^2 + \sigma^2 tr(\boldsymbol{\Sigma}_z \boldsymbol{D}_1)$, *where* $\boldsymbol{\Sigma}_z = \lim_{n\to\infty} \tilde{\boldsymbol{\Sigma}}_z$;

(ii) *under conditions* C.1, C.4, C.5, C.6(i), C.7, *and* C.8, $\hat{\boldsymbol{D}}_1 = \boldsymbol{D}_1 + O_p(n^{-1/2})$;

(iii) *under conditions* C.1 *and* C.5–C.8, $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$.

Remark 3. Note that even though the varying coefficient functions are modeled nonparametrically using cubic splines, averages of the nonparametric estimators evaluated at the design points lead to the root n convergence rate for $\hat{\mathbf{D}}_1$. This phenomenon has previously been shown for kernel estimation (Stute and Zhu (2005); Zhu and Fang (1996) and splines estimation (Zhu and Yu (2007)).

Lemma 3. *As* $n \to \infty$,

(i) *under* H_0 *and conditions* C.1, C.4, C.5, *and* C.8, $\widehat{\kappa}_0 \xrightarrow{p} \kappa$; *under* H_A *and conditions* C.1, C.2, *and* C.5–C.8, $\widehat{\kappa}_0 \xrightarrow{p} \kappa + \lim_{n\to\infty} R_n^0$, *where* R_n^0 *is defined in* (A.3) *in Appendix A.*

(ii) *under conditions* C.1, C.2, *and* C.5–C.8, $\widehat{\kappa} \xrightarrow{p} \kappa$.

Lemma 4. *Under Conditions* C.1, C.2, C.4, *and* C.5, *we have*

(i) *all eigenvalues of* $\boldsymbol{I}_n - \boldsymbol{H}_\lambda$ *lie between 0 and 1, and* $n^{-1/2} tr(\boldsymbol{H}_\lambda^j) \to 0$ *as* $n \to \infty$ *for* $j = 1, \ldots$.

(ii) *all eigenvalues of* $\boldsymbol{I}_n - \tilde{\boldsymbol{H}}_\lambda$ *lie between 0 and 1,* $n^{-1/2}\|(\boldsymbol{I}_n - \tilde{\boldsymbol{H}}_\lambda)\boldsymbol{\mu}\|^2 \to 0$, $n^{-1/2} tr(\tilde{\boldsymbol{H}}_\lambda^j) \to 0$, *and* $n^{-1/2} tr(\tilde{\boldsymbol{H}}_\lambda^{\mathrm{T}} \tilde{\boldsymbol{H}}_\lambda)^j \to 0$ *for* $j = 1, \ldots$.

Lemma 5. *Let* $\boldsymbol{A}_n$ *be a symmetric matrix,* $\delta_{(n)}$ *be the maximum eigenvalue of* $\boldsymbol{A}_n^2$, *and* $\boldsymbol{\epsilon}$ *be as in model* (2.1). *Then under the null hypothesis and condition* C.8, $\delta_{(n)}/tr(\boldsymbol{A}_n^2) \to 0$ *implies that*

$$\frac{\boldsymbol{\epsilon}^{\mathrm{T}} \boldsymbol{A}_n \boldsymbol{\epsilon} - \sigma^2 tr \boldsymbol{A}_n}{\{(\kappa - 3\sigma^4)\boldsymbol{a}_n^{\mathrm{T}} \boldsymbol{a}_n + 2\sigma^4 tr(\boldsymbol{A}_n^2)\}^{1/2}} \xrightarrow{d} N(0,1),$$

where $\boldsymbol{a}_n$ *is the column vector composed of the diagonal components of* $\boldsymbol{A}_n$.

We provide a sketch of proof for Lemma 1 only; proofs of the other lemmas are given in the supplement to the paper.

Proof of Lemma 1. We only need verify that, for any d-dimensional constant vector $\mathbf{a} = (a_1, a_2, \ldots, a_d)^{\mathrm{T}}$, $n^{-1/2}\mathbf{a}^{\mathrm{T}}\mathbf{U}_\theta(\hat{\mathbf{f}}, \hat{\sigma}^2, 0)$ converges to $\mathrm{N}(0, \mathbf{a}^{\mathrm{T}}\mathbf{M}\mathbf{a})$ in distribution under the null.

Note that, under the null hypothesis,

$$\begin{aligned}
&n^{-1/2}\mathbf{a}^{\mathrm T}\mathbf{U}_\theta(\hat{\mathbf f},\hat\sigma^2,0)\\
&= \frac{\mathbf{Y}^{\mathrm T}(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)^{\mathrm T}\left[\sum_{i=1}^d a_i\left\{\mathbf{Q}_i-\mathbf{I}_n tr\mathbf{Q}_i/(n-tr\tilde{\mathbf H}_\lambda+tr(\mathbf{Z}\hat{\mathbf D}\mathbf{Z}^{\mathrm T}))\right\}\right](\mathbf{I}_n-\tilde{\mathbf H}_\lambda)\mathbf{Y}}{2\sqrt{n}\hat\sigma^2}\\
&= \frac{\boldsymbol\epsilon^{\mathrm T}\mathbf{A}_n\boldsymbol\epsilon-\sigma^2 tr(\mathbf{A}_n)}{2\sqrt{n}\hat\sigma^2}+\frac{2\boldsymbol\mu^{\mathrm T}\mathbf{A}_n\boldsymbol\epsilon}{2\sqrt{n}\hat\sigma^2}+\frac{\sigma^2 tr\mathbf{A}_n}{2\sqrt{n}\hat\sigma^2}+\frac{\boldsymbol\mu^{\mathrm T}\mathbf{A}_n\boldsymbol\mu}{2\sqrt{n}\hat\sigma^2}, \qquad \text{(C.1)}
\end{aligned}$$

where $\mathbf{A}_n=(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)^{\mathrm T}\left(\sum_{i=1}^d a_i\left[\mathbf{Q}_i-\mathbf{I}_n tr\mathbf{Q}_i/\{n-tr\tilde{\mathbf H}_\lambda+tr(\mathbf{Z}\hat{\mathbf D}\mathbf{Z}^{\mathrm T})\}\right]\right)(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)$.

We show that, as $n\to\infty$,

$$n^{-1}tr(\mathbf{A}_n^2)\to\mathbf{a}^{\mathrm T}\mathbf{M}_0\mathbf{a},\quad \frac{tr\mathbf{A}_n}{\sqrt{tr(\mathbf{A}_n^2)}}\to 0,\quad \frac{\boldsymbol\mu^{\mathrm T}\mathbf{A}_n\boldsymbol\mu}{\sqrt{tr(\mathbf{A}_n^2)}}\to 0,$$

which implies that the third term and the fourth term in (C.1) converge to zero.

Let $\mathbf{A}_{na}=\sum_{i=1}^d a_i\left[\mathbf{Q}_i-\mathbf{I}_n tr(\mathbf{Q}_i)/\{n-tr\tilde{\mathbf H}_\lambda+tr(\mathbf{Z}\hat{\mathbf D}\mathbf{Z}^{\mathrm T})\}\right]$, and let $\delta_{(n)}$ be the maximum eigenvalue of $\mathbf{A}_n^2$. Note that under the null hypothesis, by Condition $C3(i)$, Lemmas 2 and 4, $n^{-1}tr(\mathbf{A}_{na}^2)=n^{-1}\mathbf{a}^{\mathrm T}\mathbf{M}_{n0}\mathbf{a}+o_p(1)=\mathbf{a}^{\mathrm T}\mathbf{M}_0\mathbf{a}+o_p(1)$. Then, as $n\to\infty$,

$$\begin{aligned}
n^{-1}tr(\mathbf{A}_n^2) &= n^{-1}tr\{(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)^{\mathrm T}\mathbf{A}_{na}(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)^{\mathrm T}\mathbf{A}_{na}(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)\}\\
&= n^{-1}tr(\mathbf{A}_{na}^2)\\
&\quad +n^{-1}O\{tr(\tilde{\mathbf H}_\lambda+\tilde{\mathbf H}_\lambda^{\mathrm T}+\tilde{\mathbf H}_\lambda\tilde{\mathbf H}_\lambda^{\mathrm T}+\tilde{\mathbf H}_\lambda^{\mathrm T}\tilde{\mathbf H}_\lambda\tilde{\mathbf H}_\lambda+\tilde{\mathbf H}_\lambda\tilde{\mathbf H}_\lambda^{\mathrm T}\tilde{\mathbf H}_\lambda\tilde{\mathbf H}_\lambda^{\mathrm T})\}\\
&\to \mathbf{a}^{\mathrm T}\mathbf{M}_0\mathbf{a}; \qquad \text{(C.2)}\\
\delta_{(n)} &\le \max_{\|\boldsymbol\xi\|^2=1}\boldsymbol\xi^{\mathrm T}(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)^{\mathrm T}\mathbf{A}_{na}(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)^{\mathrm T}\mathbf{A}_{na}(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)\boldsymbol\xi\\
&\le \max_{\|\boldsymbol\xi\|^2=1}\boldsymbol\xi^{\mathrm T}\mathbf{A}_{na}^2\boldsymbol\xi\\
&\le cC_0^2. \qquad \text{(C.3)}
\end{aligned}$$

Condition C.5, Lemma 4, and the fact that $tr(\mathbf{A}_{na})\to 0$ imply

$$n^{-1/2}tr(\mathbf{A}_n)=n^{-1/2}[tr(\mathbf{A}_{na})+O\{tr(\tilde{\mathbf H}_\lambda+\tilde{\mathbf H}_\lambda^{\mathrm T}\tilde{\mathbf H}_\lambda)\}]\to 0,\quad n\to\infty.$$

Combining this with (C.2), we obtain $tr\mathbf{A}_n/\sqrt{tr(\mathbf{A}_n^2)}\to 0$. In addition, (C.2), Condition C.3(i), and Lemma 4(ii) lead to $\boldsymbol\mu^{\mathrm T}\mathbf{A}_n\boldsymbol\mu/\sqrt{tr(\mathbf{A}_n^2)}\to 0$.

Note that as $n\to\infty$,

$$E\left(\frac{\boldsymbol\mu^{\mathrm T}\mathbf{A}_n\boldsymbol\epsilon}{\sqrt n}\right)^2=\frac{\sigma^2\boldsymbol\mu^{\mathrm T}\mathbf{A}_n^2\boldsymbol\mu}{n}\le\frac{cC_0^2\|(\mathbf{I}_n-\tilde{\mathbf H}_\lambda)\boldsymbol\mu\|^2}{n}\to 0.$$

Then by Lemma 4(ii), the second term in (C.1) converges to zero in probability.

By (C.2) and (C.3), we have

$$\frac{\delta_{(n)}}{tr(\mathbf{A}_n^2)} \to 0, \quad n \to \infty. \tag{C.4}$$

Furthermore, if $\mathbf{a}_n$ and $\mathbf{a}_{na}$ are the column vectors composed of the principal diagonal elements of $\mathbf{A}_n$ and $\mathbf{A}_{na}$ respectively, then

$$\lim_{n\to\infty} n^{-1}\mathbf{a}_n^{\mathrm{T}}\mathbf{a}_n = \mathbf{a}^{\mathrm{T}}\mathbf{M}_z\mathbf{a}. \tag{C.5}$$

In fact, under the null hypothesis, Condition C.3(i), Lemma 4(ii), and Lemma 2 imply

$$\begin{aligned}
n^{-1}\mathbf{a}_{na}^{\mathrm{T}}\mathbf{a}_{na} &= n^{-1}\sum_{l_1=1}^{d}\sum_{l_2=1}^{d} a_{l_1}a_{l_2} tr\left[diag\left\{\mathbf{Q}_{l_1} - \frac{\mathbf{I}_n tr\mathbf{Q}_{l_1}}{n - tr\tilde{\mathbf{H}}_\lambda + tr(\mathbf{Z}\hat{\mathbf{D}}\mathbf{Z}^{\mathrm{T}})}\right\}\right. \\
&\qquad \left. diag\left\{\mathbf{Q}_{l_2} - \frac{\mathbf{I}_n tr\mathbf{Q}_{l_2}}{n - tr\tilde{\mathbf{H}}_\lambda + tr(\mathbf{Z}\hat{\mathbf{D}}\mathbf{Z}^{\mathrm{T}})}\right\}\right] \\
&= n^{-1}\sum_{l_1=1}^{d}\sum_{l_2=1}^{d} \mathbf{a}_{l_1}\mathbf{a}_{l_2}\left[tr\{diag(\mathbf{Q}_{l_1})diag(\mathbf{Q}_{l_2})\} - \frac{tr\mathbf{Q}_{l_1}tr\mathbf{Q}_{l_2}}{n}\right] \\
&\qquad + \left\{1 - \frac{n}{n - tr\tilde{\mathbf{H}}_\lambda + tr(\mathbf{Z}\hat{\mathbf{D}}\mathbf{Z}^{\mathrm{T}})}\right\}^2 \sum_{l_1=1}^{d}\sum_{l_2=1}^{d} \mathbf{a}_{l_1}\mathbf{a}_{l_2}\frac{tr\mathbf{Q}_{l_1}tr\mathbf{Q}_{l_2}}{n^2} \\
&= \mathbf{a}^{\mathrm{T}}\mathbf{M}_{nz}\mathbf{a} + o_p(1) \\
&\xrightarrow{p} \mathbf{a}^{\mathrm{T}}\mathbf{M}_z\mathbf{a}, \quad n \to \infty.
\end{aligned} \tag{C.6}$$

Besides, from Condition C.3(i), Lemma 4, and the Cauchy-Schwartz inequality, we have

$$\begin{aligned}
n^{-1}\mathbf{a}_n^{\mathrm{T}}\mathbf{a}_n - n^{-1}\mathbf{a}_{na}^{\mathrm{T}}\mathbf{a}_{na} &= n^{-1}O[tr\{\tilde{\mathbf{H}}_\lambda^{\mathrm{T}}\tilde{\mathbf{H}}_\lambda + \tilde{\mathbf{H}}_\lambda\tilde{\mathbf{H}}_\lambda^{\mathrm{T}}\tilde{\mathbf{H}}_\lambda + (\tilde{\mathbf{H}}_\lambda^{\mathrm{T}}\tilde{\mathbf{H}}_\lambda)^2\}] \\
&\to 0, \quad n \to \infty.
\end{aligned} \tag{C.7}$$

Thus, (C.5) is obtained from (C.6) and (C.7). Using (C.2), (C.4), (C.5), and Lemma 5, we have

$$\begin{aligned}
&\frac{\boldsymbol{\epsilon}^{\mathrm{T}}\mathbf{A}_n\boldsymbol{\epsilon} - \sigma^2 tr\mathbf{A}_n}{2\sqrt{n}\hat{\sigma}^2} \\
&\quad = \frac{\boldsymbol{\epsilon}^{\mathrm{T}}\mathbf{A}_n\boldsymbol{\epsilon} - \sigma^2 tr\mathbf{A}_n}{2\hat{\sigma}^2\sqrt{(\kappa - 3\sigma^4)\mathbf{a}_n^{\mathrm{T}}\mathbf{a}_n + 2\sigma^4 tr(\mathbf{A}_n^2)}}\sqrt{(\kappa - 3\sigma^4)\frac{\mathbf{a}_n^{\mathrm{T}}\mathbf{a}_n}{n} + 2\sigma^4\frac{tr(\mathbf{A}_n^2)}{n}}
\end{aligned}$$

$$\xrightarrow{d} \mathrm{N}(0, \mathbf{a}^{\mathrm{T}}\mathbf{M}\mathbf{a}), \qquad n \to \infty. \tag{C.8}$$

Altogether, we have $n^{-1/2}\mathbf{U}_\theta(\hat{\mathbf{f}}, \hat{\sigma}^2, 0) \xrightarrow{d} \mathrm{N}(0, \mathbf{M})$.

The result $n^{-1/2}\mathbf{U}_\theta(\hat{\mathbf{f}}_0, \hat{\sigma}_0^2, 0) \xrightarrow{d} \mathrm{N}(0, \mathbf{M})$ can be proved similarly.

Proof of Theorem 1. Note that $\mathbf{M}_n^{-1/2}(\hat{\sigma}^2, \hat{\kappa})\mathbf{U}_\theta(\hat{\mathbf{f}}, \hat{\sigma}^2, 0) \xrightarrow{d} \mathrm{N}(0, \mathbf{I}_d)$. Then for any $x \in R$,

$$\lim_{n\to\infty} P(T_n \le x) = \Phi^d(x),$$

where $\Phi(\cdot)$ is the standard normal distribution function. Thus the asymptotic property of T_n under the null is verified. The asymptotic null distribution of T_{n0} can be proved similarly.

Proof of Theorem 2: For notational simplicity, let $\mathbf{Y}_{n0}$ be $\mathbf{Y}$ under the null. Then, for any d-dimensional vector $\mathbf{a}$,

$$\begin{aligned} n^{-1/2}\mathbf{a}^{\mathrm{T}}\mathbf{U}_\theta(\hat{\mathbf{f}}, \hat{\sigma}^2, 0) &= \frac{\mathbf{Y}_{n0}^{\mathrm{T}}\mathbf{A}_n\mathbf{Y}_{n0}}{2\sqrt{n}\sigma^2} + \frac{\mathbf{Y}_{n0}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{\sqrt{n}\sigma^2} + \frac{\mathbf{b}^{\mathrm{T}}\mathbf{Z}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{2\sqrt{n}\sigma^2} \\ &\triangleq II_1 + II_2 + II_3, \end{aligned} \tag{C.9}$$

where $\mathbf{A}_n$ is defined in (C.1). According to the proof for Lemma 1, $II_1 \xrightarrow{d} \mathrm{N}(0, \mathbf{a}_0^{\mathrm{T}}\mathbf{M}\mathbf{a}_0)$. Hence, we only need show that II_2 and II_3 converge in probability to zero and $\mathbf{a}^{\mathrm{T}}\boldsymbol{\omega}/2$, respectively.

By condition C.8, C.3, and (C.2) in the proof for Lemma 1, we have

$$\begin{aligned} E\left(\frac{II_3}{n^{\alpha_0}c_n}\right)^2 &= \frac{E\{tr(\mathbf{Z}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}\mathbf{b}^{\mathrm{T}})\}^2}{4\sigma^4 n^{1+2\alpha_0}c_n^2} \\ &= \frac{Vec^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z})E\{Vec(\mathbf{b}\mathbf{b}^{\mathrm{T}})Vec^{\mathrm{T}}(\mathbf{b}\mathbf{b}^{\mathrm{T}})\}Vec(\mathbf{Z}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z})}{4\sigma^4 n^{1+2\alpha_0}c_n^2} \\ &\le \frac{c\,tr(\mathbf{Z}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z})^2}{n^{2\alpha_0}\cdot n} \\ &\le cn^{-(1+2\alpha_0)}tr(\mathbf{A}_n^2) \to 0, \quad n \to \infty. \end{aligned}$$

In addition, by Condition C.3 and Lemma 4, we have

$$\begin{aligned} \frac{EII_3}{n^{\alpha_0}c_n} &= \frac{tr(\mathbf{Q}_0\mathbf{A}_n)}{2n^{\alpha_0+1/2}} \\ &= \frac{1}{2n^{\alpha_0+1/2}}\sum_{i=1}^d a_i\{tr(\mathbf{Q}_0\mathbf{Q}_i) - n^{-1}tr\mathbf{Q}_0 tr\mathbf{Q}_i\} \\ &\quad + \frac{1}{2n^{\alpha_0+1/2}}\sum_{i=1}^d a_i\frac{tr\mathbf{Q}_0 tr\mathbf{Q}_i}{n}\left\{1 - \frac{n}{n - tr\tilde{\mathbf{H}}_\lambda + c_n tr(\mathbf{Q}_0)}\right\} \end{aligned}$$

$$
\begin{aligned}
&+n^{-\alpha_0}O\left\{n^{-1/2}tr(\tilde{\mathbf{H}}_\lambda+\tilde{\mathbf{H}}_\lambda^{\mathrm{T}}\tilde{\mathbf{H}}_\lambda)\right\}\\
&=\frac{\mathbf{a}^{\mathrm{T}}\boldsymbol{\omega}}{2}+\sum_{i=1}^{d}\frac{a_i tr\mathbf{Q}_0 tr\mathbf{Q}_i}{2nn^{\alpha_0+1/2}}\frac{c_n n^{-1}tr\mathbf{Q}_0}{1+c_n n^{-1}tr\mathbf{Q}_0}+o(1)\\
&=\frac{\mathbf{a}^{\mathrm{T}}\boldsymbol{\omega}}{2}+o(1).
\end{aligned}
$$

Therefore,

$$
II_3=\frac{k_0\mathbf{a}^{\mathrm{T}}\boldsymbol{\omega}}{2}+o_p(1). \tag{C.10}
$$

For II_2, it is easy to see that

$$
\frac{\mathbf{Y}_{n_0}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{\sqrt{n}}=\frac{\boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{\sqrt{n}}+\frac{(\mathbf{Y}_{n_0}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{\sqrt{n}}=o_p(1). \tag{C.11}
$$

In fact, $E(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}/\sqrt{n})=0$. Lemma 4 and Condition C.5 imply

$$
E\left(\frac{\boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{\sqrt{n}}\right)^2=\frac{\sigma^2 c_n tr(\boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}_n\mathbf{Q}_0\mathbf{A}_n\boldsymbol{\mu})}{n}\le\frac{cc_n\boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}_n^2\boldsymbol{\mu}}{n}\to 0,\quad as\quad n\to\infty.
$$

Hence, the first term of (C.11) converges to zero in probability.

We now deal with the second term in (C.11). By Condition C.3(i), $\lim_{n\to\infty} n^{\alpha_0}c_n=k_0$, and (C.2) in the proof for Lemma 1,

$$
E\left(\frac{\boldsymbol{\epsilon}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{\sqrt{n}}\right)^2=\frac{c_n\sigma^2 tr(\mathbf{Q}_0\mathbf{A}_n^2)}{n}\to 0,\quad n\to\infty.
$$

Combining (C.9), (C.10), with (C.11) and Lemma 1, we have

$$
\frac{\mathbf{a}^{\mathrm{T}}\mathbf{U}_\theta(\hat{\mathbf{f}},\hat{\sigma}^2,0)}{\sqrt{n}}\xrightarrow{d}\mathrm{N}\left(\frac{k_0\mathbf{a}^{\mathrm{T}}\boldsymbol{\omega}}{2},\mathbf{a}^{\mathrm{T}}\mathbf{M}\mathbf{a}\right).
$$

By the same argument as in the proof of Theorem 1, the power property of T_n is obtained.

The proof for the power property of T_{n0} is similar and is outlined below. From (C.1) in the proof for Lemma 1,

$$
\begin{aligned}
n^{-1/2}\mathbf{a}^{\mathrm{T}}\mathbf{U}_\theta(\hat{\mathbf{f}}_0,\hat{\sigma}_0^2,0)&=\frac{\mathbf{Y}_{n0}^{\mathrm{T}}\mathbf{A}_n\mathbf{Y}_{n0}}{2\sqrt{n}\hat{\sigma}_0^2}+\frac{2\mathbf{Y}_{n0}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{2\sqrt{n}\hat{\sigma}_0^2}+\frac{\mathbf{b}^{\mathrm{T}}\mathbf{Z}^{\mathrm{T}}\mathbf{A}_n\mathbf{Z}\mathbf{b}}{2\sqrt{n}\hat{\sigma}_0^2}\\
&\triangleq II_1^0+II_2^0+II_3^0.
\end{aligned}
$$

By Lemmas 1 and 2, we have

$$
II_1^0\xrightarrow{d}\mathrm{N}\left(0,\frac{\mathbf{a}^{\mathrm{T}}\mathbf{M}\mathbf{a}}{\{1+\lim_{n\to\infty}c_n tr(\boldsymbol{\Sigma}_z\mathbf{D}_1^0)\}^2}\right).
$$

In addition, $II_3^0/(n^{\alpha_0}c_n) \xrightarrow{p} \mathbf{a}^{\mathrm{T}}\boldsymbol{\omega}/\{2(1+\lim_{n\to\infty} c_n tr(\boldsymbol{\Sigma}_z\mathbf{D}_1^0))\} \neq 0$. Therefore, if c_n is bounded away from zero, we have $T_{n0} \xrightarrow{d} \infty$. If $n^{\alpha_0}c_n \to k_0$, then $\hat{\sigma}_0^2 \xrightarrow{p} \sigma^2$ from Lemma 2(i), and $\hat{\kappa}_0 \xrightarrow{p} \kappa$ from Lemma 3(i). Therefore, $\mathbf{M}_n^{-1/2}(\hat{\sigma}_0^2,\hat{\kappa}_0)\mathbf{U}_\theta(\hat{\mathbf{f}}_0,\hat{\sigma}_0^2,0) \xrightarrow{d} \mathrm{N}(k_0\mathbf{M}^{-1/2}\boldsymbol{\omega}/2,\mathbf{I}_d)$.

Appendix D: Sketch of the Derivation of $\hat{\mathbf{D}}_1$

Consider $\tilde{\mathbf{A}} = n^{-1}\sum_{i=1}^m \tilde{\mathbf{Z}}_i^{\mathrm{T}}\tilde{\boldsymbol{\epsilon}}_i\tilde{\boldsymbol{\epsilon}}_i^{\mathrm{T}}\tilde{\mathbf{Z}}_i - n^{-1}\sum_{i=1}^m \tilde{\boldsymbol{\epsilon}}_i^{\mathrm{T}}\tilde{\boldsymbol{\epsilon}}_i I_q$, where $\tilde{\boldsymbol{\epsilon}}_i = \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i$ and $\tilde{\mathbf{Z}}_i$ is defined in (A.1) in Appendix A. Following the arguments of Cui, Ng, and Zhu (2004), we have $\tilde{\mathbf{A}} = \mathbf{A} + o_p(1)$, where

$$\mathbf{A} = \sigma^2\{(c_{12}+c_{13})\mathbf{B} + (c_{13}-1)(tr\mathbf{B})\mathbf{I}_q + (c_{11}-c_{12}-2c_{13})\mathrm{diag}(\mathbf{B})\}, \quad (D.1)$$

the c_{1j}'s are defined in Condition C.7, $\mathbf{B} = (\boldsymbol{\Sigma}_z)^{1/2}\mathbf{D}_1(\boldsymbol{\Sigma}_z)^{1/2}$, diag($\mathbf{B}$) is the diagonal matrix composed of the diagonal elements of $\mathbf{B}$, and $\boldsymbol{\Sigma}_z$ is defined in Condition C.6(i) in Appendix B. Solving equation (D.1) by examining the diagonal and off-diagonal elements separately, we have

$$\mathbf{B} = \frac{\mathbf{A}}{c_{12}+c_{13}} + \frac{2c_{13}-(c_{11}-c_{12})}{(c_{11}-c_{13})(c_{12}+c_{13})}\mathrm{diag}(\mathbf{A}) - \frac{(c_{13}-1)tr\mathbf{A}}{(c_{11}-c_{13})\{c_{11}-c_{13}+q(c_{13}-1)\}}\mathbf{I}_q.$$

Using the fact that $\mathbf{D}_1 = (\boldsymbol{\Sigma}_z)^{-1/2}\mathbf{B}(\boldsymbol{\Sigma}_z)^{-1/2}$ and plugging in estimators of $\boldsymbol{\Sigma}_z$ and $\mathbf{B}$ leads to the estimator (2.5).

References

Bolfarine, H. and Valenca, D. M. (2005). Testing homogeneity in Weibull-regression models. *Biometrical J.* **47**, 1-14.

Chen, S. X., Härdle, W. and Li, M. (2003). An empirical likelihood goodness-of-fit test for time series. *J. Roy. Statist. Soc. Ser. B* **65**, 663-678.

Chiang, C. T., Rice, J. A. and Wu, C. O. (2001). Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables. *J. Amer. Statist. Assoc.* **96**, 605-619.

Commenges, D. and Jacqmin-Gadda, H. (1997). Generalized score test of homogeneity based on correlated random effects models. *J. Roy. Statist. Soc. Ser. B* **59**, 157-171.

Crainiceanu, C. M. Ruppert, D., Claeskens, G. and Wand, M. P. (2005). Exact likelihood ratio tests for penalised splines. *Biometrika* **92**, 91-103.

Cui, H. J., Ng, K. W., and Zhu, L. X. (2004). Estimation in mixed effects model with errors in variables. *J. Mult. Anal.* **91**, 53-73.

Eubank, R. L. and Thomas, W. (1993). Detecting heteroscedasticity in nonparametric regression. *J. Roy. Statist. Soc. Ser. B* **55**, 145-155.

Fan, J. and Huang, T. (2005). Profile likelihood inference on semiparametric varying-coefficient partially linear models. *Bernoulli* **11**, 1031-1057.

Fan, J., Zhang, C. and Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Statist.* **29**, 153-193.

Green, P. and Silverman, B. (1994). *Nonparametric Regression and Generalized Linear Models.* Chapman and Hall, London.

Gu, C. and Ma, P. (2005). Optimal smoothing in nonparametric mixed-effect models. *Ann. Statist.* **33**, 1357-1379.

Hoover, D. R., Rice, J. A., Wu, C. O., and Yang, L. P. (1998). Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika* **85**, 809-822.

Huang, J. H., Wu, C.O., and Zhou, L. (2004). Polynomial spline estimation and inference for varying coefficient models with longitudinal data. *Statist. Sinica* **14**, 763-788.

Jacqmin-Gadda, H. and Commenges, D. (1995). Tests of homogeneity for generalized linear models. *J. Amer. Statist. Assoc.* **90**, 1237-1246.

Kaslow, R. A., Ostrow, D. G., Detels, R., Phair, J. P., Polk, B. F., and Rinaldo, C. R. (1987). The multicenter AIDS cohort study: rationale, organization and selected characteristics of the participants. *Amer. J. Epidemiology* **126**, 310-318.

Koul, H. and Song, W. (2008). Regression model checking with Berkson measurement errors. *J. Stat. Plan. Infer.* **138**, 1615-1628.

Laird, N., Lange, N. and Stram, D. (1987). Maximum likelihood computation with repeated measures: Application of the EM algorithm. *J. Amer. Statis. Assoc.* **82**, 97-105.

Laird, N. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* **38**, 963-974.

Lin, X. H. (1997). Variance component testing in generalized linear models with random effects. *Biometrika* **84**, 309-326.

Lin, X. H. and Zhang, D. W. (1999). Inference in generalized additive mixed models by using smoothing spline. *J. Roy. Statist. Soc. Ser. B* **61**, 381-400.

Liu, A. and Wang, Y. D. (2004). Hypothesis testing in smoothing spline models. *J. Statist. Comput. Simul.* **74**, 581-597.

Nelder, J. A. and Pregibon, D. (1987). An extended quasi-likelihood function. *Biometrika* **74**, 221-232.

O'Sullivan, F., Yandell, B. S., and Raynor, W. J. (1986). Automatic smoothing of regression functions in generalized models. *J. Amer. Statist. Assoc.* **81**, 96-103.

Self, S. G. and Liang, K. Y. (1987). Large sample properties of the maximum likelihood estimator and the likelihood ratio test on the boundary of the parameter space. *J. Amer. Statist. Assoc.* **82**, 605-611.

Silvapulle, M. J. and Silvapulle, P. (1995). A score test against one-sided alternatives . *J. Amer. Statist. Assoc.* **90**, 342-349.

Silvapulle, M. J. and Sen, P. K. (2005). *Constrained Statistical Inference: Inequality, Order and Shape Restrictions.* Wiley, New Jersey.

Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics* **50**, 1171-1177.

Stute, W. and Zhu, L. X.(2005). Nonparametric checks for single-index models. *Ann. Statist.* **33**, 1048-1083.

Verbeke, G. and Molenberghs, G. (2003). The use of score tests for inference on variance components. *Biometrics* **59**, 254-262.

Wahba, G(1990). *Spline methods for observational data.* Philadelphia, PA: Society for Industrial and Applied Mathematics.

Wang, Y. D. (1998). Smoothing spline models with correlated random errors. *J. Amer. Statist. Assoc.* **93**, 341-348.

Wu, C. O. and Chiang, C. T. (2000). Kernel smoothing on varying coefficient models with longitudinal dependent variable. *Statist. Sinica* **10**, 433-456.

Wu, H. and Liang, H. (2004). Backfitting random varying-coefficient models with time-dependent smoothing covariates. *Scand. J. Statist.* **31**, 3-19.

Wu, H. and Zhang, J. T. (2006). *Nonparametric Regression Methods for Longitudinal Data: Mixed-Effects Modeling Approaches.* Wiley, New York.

Zhang, D. W. (2004). Generalized linear mixed models with varying coefficients for longitudinal data. *Biometrics* **60**, 8-15.

Zhu, L. P. and Yu, Z. (2007). On spline approximation of sliced inverse regression, *Sci China Ser A: Mathematics* **50**, 1289-1302.

Zhu, L. X. and Fang, K. T. (1996). Asymptotics for the kernel estimates of sliced inverse regression. *Ann. Statist.* **24**, 1053-1067.

Zhu, Z. Y. and Fung, W. K. (2004). Variance component testing in semiparametric mixed models. *J. Mult. Anal.* **91**, 107-118.

Department of Mathematics, China University of Mining & Technology, Beijing, 100083, China

E-mail: zxlee@amss.ac.cn

Department of Statistics and Applied Probability, College of Letters and Science, University of California - Santa Barbara, Santa Barbara, California 93106-3110, USA.

E-mail: yuedong@pstat.ucsb.edu

College of Life Science, East China Normal University, Shanghai 200062, P.R. China.

E-mail: pwu@stat.ecnu.edu.cn

School of Statistics, Renmin University of China, Bejing, 100072, China.

E-mail: xwlbnu@163.com

Department of Mathematics, The Hong Kong Baptist University, Kowloon Tong, Hong Kong, China.

E-mail: lzhu@hkbu.edu.hk

(Received September 2009; accepted November 2010)