Bayesian Local Influence of Semiparametric Structural Equation Models

Ming Ouyang*

Abstract
This research develops a Bayesian local influence method for semiparametric structural equation models. The effects of minor perturbations to individual observations, sampling distributions, and prior distributions on the statistical inference are assessed with the use of various perturbation schemes. We construct a Bayesian perturbation manifold to characterize such perturbation schemes. The first- and second-order influence measures are proposed to quantify the degree of minor perturbations to different aspects of a statistical model on the basis of a variety of objective functions such as Bayes factor. We conduct simulation studies to evaluate the empirical performance of the Bayesian local influence procedure.

Key Words: Bayesian local influence, perturbation schemes, latent variables, semiparametric modeling, MCMC method

1. Introduction

In the behavioral, social, psychological, and medical sciences, latent variables that are not directly assessable via a single observed variable and instead should be measured by multiple observed indicators are commonly encountered. In the analysis of the interrelationships among observed and latent variables, structural equation model (SEM) [20] is one of the most popular and efficient methods. In the conventional SEM, outcome latent variables are usually regressed on explanatory latent variables in a linear form. However, this linear and parametric SEM may be too restrictive to correctly reflect the reality. To relax the assumption of the conventional SEM, we consider a semiparametric SEM that incorporates the nonlinear functions of latent variables and covariates as well as their interactions. Abundant nonparametric modeling and smoothing techniques such as smooth splines [10], kernel methods with local polynomials [7, 8], and penalized splines [23], have been developed in a frequentist framework over the past decades. Recently, due to the nice features of the Bayesian approach and the rapid development of the sampling-based computing tools, Bayesian nonparametric methods have received increasing attention in substantive research [see, for example, 25, 26]. In this study, we consider a semiparametric SEM, which comprises a confirmatory factor analysis model and a nonparametric structural equation, wherein the univariate and bivariate nonparametric functions are modeled via the Bayesian P-splines approach [6, 16].

In substantive research, erroneous results of the statistical inference could be caused by several improper model inputs such as data and model specification. To identify potential outliers or influential observations and assess the stability of estimation outputs with respect to model and data inputs, Cook [4] proposed local influence analysis in the context of linear regression models. Since the pioneering work of Cook [4], local influence has received considerable attention as an important statistical inference beyond estimation, and has already been applied to a large number of statistical models including SEMs. For example, Zhu and Lee [31] extended Cook's approach to incomplete data models. Lee and Tang [18] applied the local influence approach to nonlinear SEMs. Lee *et al.* [19] assessed local influence for nonlinear SEMs with ignorable missing data. Song and Lee [24] conducted

*Department of Statistics, The Chinese University of Hong Kong, Hong Kong China

a local influence analysis for mixture of SEMs. However, the abovementioned researches were developed through the maximum likelihood (ML)-based estimation methods. Compared to the rapid development of local influence in the frequentist framework, Bayesian local influence analysis results are quite limited. Among a very first work was conducted by McCulloch [21], which extended the local influence approach to assess the effect of perturbation to prior on the Bayesian analysis. Recently, Zhu *et al.* [31] considered a geometric approach to simultaneously assess outliers and/or influential observations, perform model comparison, and conduct sensitivity analysis by introducing various perturbations to the data, sampling distributions, and the prior distributions of model parameters, respectively. However, the previous studies mainly focused on parametric models and thus are not directly applicable to the present model framework. In this article, we develop Bayesian local influence procedure in the context of semiparametric SEM. To our knowledge, no study has focused on the Bayesian local influence of such semiparametric latent variable model.

The remainder of the article is organized as follows. Section 2 defines the semiparametric SEM. The Bayesian P-splines approach for estimating univariate and bivariate nonlinear functions is briefly described as well. Section 3 introduces the Bayesian perturbation model and manifold, the first- and second-order local influence measures based on the objective function of Bayes factor, and the associated posterior computation. Section 4 presents simulation studies to assess the empirical performance of the proposed methodology. Section 5 reports an application to the study of bone mineral density (BMD). Section 5 concludes the paper.

2. Semiparametric Structural Equation Model

2.1 Model description

Let $\boldsymbol{y}_i$ be a $p \times 1$ random vector of observed variables. A measurement model for characterizing latent variables on the basis of multiple observed indicators is defined as follows:

$$\boldsymbol{y}_i = \boldsymbol{A}\boldsymbol{x}_i + \boldsymbol{\Lambda}\boldsymbol{\varpi}_i + \boldsymbol{\epsilon}_i, \tag{1}$$

where $\boldsymbol{x}_i$ is a vector of fixed covariates, $\boldsymbol{A}$ is a matrix of coefficients, $\boldsymbol{\varpi}_i = (\varpi_{i1}, \cdots, \varpi_{iq})^T$ is a $q \times 1$ random vector of latent variables with $q < p$, $\boldsymbol{\Lambda}$ is a $p \times q$ factor loading matrix, $\boldsymbol{\epsilon}_i$ is a $p \times 1$ residual random vector independent of $\boldsymbol{\varpi}_i$ and distributed as $\mathcal{N}[0, \boldsymbol{\Psi}]$ with a diagonal covariance matrix $\boldsymbol{\Psi}$. To examine the interrelationships among latent variables, we partition $\boldsymbol{\varpi}_i$ into $(\boldsymbol{\eta}_i^T, \boldsymbol{\xi}_i^T)^T$, where $\boldsymbol{\eta}_i(q_1 \times 1)$ and $\boldsymbol{\xi}_i(q_2 \times 1)$ denote the outcome and explanatory latent vectors, respectively, and $\boldsymbol{\xi}_i$ is assumed distributed as $\mathcal{N}[0, \boldsymbol{\Phi}^*]$ with a covariance matrix $\boldsymbol{\Phi}^*$. A nonparametric structural equation for assessing the functional effects of the explanatory observed and latent variables on the outcome latent variables is defined as follows. For an arbitrary element η_{ih} in $\boldsymbol{\eta}_i$, $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, q_1$,

$$\eta_{ij} = g_{j1}(z_{i1}) + \cdots + g_{jD}(z_{iD}) + f_{j1}(\xi_{i1}) + \cdots + f_{jq_2}(\xi_{iq_2}) + \sum_{u<v} h_{juv}(\xi_{iu}, \xi_{iv}) + \delta_{ij}, \tag{2}$$

where $z_{i1}, \cdots, z_{iD}$ are fixed covariates, $g_{j1}, \cdots, g_{jD}, f_{j1}, \cdots, f_{jq2}$ and h_{juv} are the unspecified univariate and bivariate smooth functions, δ_{ij} is the residual error distributed as $\mathcal{N}[0, \psi_{\delta_j}]$ and independent of $\boldsymbol{\xi}_i$ and δ_{ih} for $j \neq h$. For notational simplify, in what follows, we suppress the subscript j in (2) by assuming $q_1 = 1$. Then, (2) can be simplified to

$$\eta_i = g_1(z_{i1}) + \cdots + g_D(z_{iD}) + f_1(\xi_{i1}) + \cdots + f_{q_2}(\xi_{iq_2}) + \sum_{u<v} h_{uv}(\xi_{iu}, \xi_{iv}) + \delta_i. \tag{3}$$

An extension to the case with $q_1 > 1$ is straightforward. The semiparametric SEM defined by (1) and (2) is an impotent extension of the ordinary linear SEM. The proposed model not only releases the linear assumption on the relations between explanatory and outcome variables, but also accommodates the pairwise functional interactions of latent variables.

The proposed model is not identifiable without imposing appropriate identifiability constraints. To identify the measurement equation, a simple and common method is to fix some elements of $\mathbf{\Lambda}$ at preassigned values (see Section 4). To identify the structural equation, we need to identify all the univariate and bivariate unknown functions. Using the idea of Song *et al.* [26], we impose the identifiability constraints for $g_d(\cdot)$, $f_k(\cdot)$, and $h_{uv}(\cdot,\cdot)$ as follows:

$$\bar{g}_d = \int_{min(z_{1d},\cdots,z_{nd})}^{max(z_{1d},\cdots,z_{nd})} g_d(z)dz/range(z_{1d},\cdots,z_{nd}) = 0, \tag{4}$$

$$\bar{f}_k = \int_{-\infty}^{\infty} f_k(\xi)p(\xi)d\xi = 0, \tag{5}$$

$$\bar{h}_{uv}(\xi_v) = \int_{-\infty}^{\infty} h_{uv}(\xi_u,\xi_v)p(\xi_u)d\xi_u = 0, for\ all\ values\ of\ \xi_v, \tag{6}$$

$$\bar{h}_{uv}(\xi_u) = \int_{-\infty}^{\infty} h_{uv}(\xi_u,\xi_v)p(\xi_v)d\xi_v = 0, for\ all\ values\ of\ \xi_u, \tag{7}$$

where $p(\xi_i)$ is the prior distribution of ξ_i. According to Song *et al.* [26], the constraints (4)–(7) enforce $g_d(\cdot)$, $f_k(\cdot)$, and $h_{uv}(\cdot,\cdot)$ to be identified and orthogonal to each other.

2.2 Modeling of nonparametric functions

The unknown smooth functions $g_d(\cdot)$, $f_r(\cdot)$ and $h_{uv}(\cdot)s$ in Equation (2) are unspecified and require estimation based on the observed data. The Bayesian P-splines [6, 16] approach can be used to approximate the unknown functions of both covariates and latent variables. For covariates z_{ij}, we can approximate $g_d(\cdot)$ using traditional B-splines approximation $\sum_{k=1}^{K_d^z} \gamma_{dk}B_{dk}^z(\cdot)$, where K_d^z is the number of splines determined by the number of knots, γ_{dk} are unknown coefficients, and $B_{dk}^z(\cdot)$ are B-splines basis functions of appropriate orders. For the unknown functions of latent variables, however, the traditional B-splines approximation cannot be used because latent variables are unobservable and their observations need to be updated in MCMC iterations. Consequently, predefining the finite ranges of latent variables and determining the positions of the knots beforehand is impossible. We propose the use of "probit transformation" method introduced in [26] to address the problem. Let $\Phi(\cdot)$ be the cumulative distribution of $\mathcal{N}(0,1)$. The unknown smooth function of latent variable $f_r(\cdot)$ is modeled by $\sum_{k=1}^{K_j} \beta_{jk}B_{jk}(\Phi(\cdot))$, where $B_{jk}(\Phi(\cdot))$ are the composite functions of B-splines basis $B_{jk}(\cdot)$ and probit transformation $\Phi(\cdot)$. For bivariate nonparametric functions of latent variables, instead of using tradition tensor product B-splines approach, we consider the modified tensor product B-splines approximation $\sum_{s=1}^{S}\sum_{t=1}^{T} b_{uvst}U_{us}(\Phi(\xi_{iu}))V_{vt}(\Phi(\xi_{iv}))$, where $U_{us}(\Phi(\cdot))$ and $V_{vt}(\Phi(\cdot))$ are likewise the composite functions of B-splines basis and probit transformation. With the B-splines approximation, equation (3) can be rewritten as

$$\begin{aligned}\eta_i = &\sum_{d=1}^{D}\sum_{k=1}^{K_d^z} \gamma_{dk}B_{dk}^z(z_{id}) + \sum_{j=1}^{q_2}\sum_{k=1}^{K_j} \beta_{jk}B_{jk}(\Phi(\xi_{ij})) \\ &+ \sum_{u=1}^{q_2}\sum_{v=1}^{q_2}\sum_{s=1}^{S}\sum_{t=1}^{T} b_{uvst}U_{us}(\Phi(\xi_{iu}))V_{vt}(\Phi(\xi_{iv})) + \delta_i.\end{aligned} \tag{8}$$

In practices, a common choice of the B-splines basis function is cubic B-splines. To control the smoothness of the approximations and prevent overfitting, we introduce various penalties to the coefficients of B-splines basis functions in (8) (see Section 3.2).

3. Bayesian Local Influence Analysis

3.1 Bayesian perturbation model

In this section, we first develop the Bayesian perturbations model. Let $\boldsymbol{\omega} = (\boldsymbol{\omega}_y^T, \boldsymbol{\omega}_s^T, \boldsymbol{\omega}_p^T)^T$ be the perturbation mapping from the infinite set $\boldsymbol{\Omega}$, where $\boldsymbol{\omega}_y, \boldsymbol{\omega}_s$, and $\boldsymbol{\omega}_p$ correspond to the perturbation to the data, the sampling distribution, and the prior distributions of parameters, respectively. Let $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta})$ be the joint probability density in Bayesian model. We introduce perturbation $\boldsymbol{\omega}$ into the density function denoted by $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$. Thus, for $\boldsymbol{\omega}$ varies in $\boldsymbol{\Omega}$, the Bayesian perturbation model $\mathfrak{M}$ can be defined as a family of probability density $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega} \in \boldsymbol{\Omega})$. We assume that $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$ have a common dominating measure for all $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ and there is a central point $\boldsymbol{\omega}^0$ of $\boldsymbol{\Omega}$ representing no perturbation, such that $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}^0) = p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta})$.

To measure each perturbations $\boldsymbol{\omega}$, we propose the Bayesian perturbation manifold based on the Bayesian perturbation model $\mathfrak{M}$. For a given $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}) \in \mathfrak{M}$, $\boldsymbol{\omega}$ may be different from $\boldsymbol{\omega}^0$, we consider a smooth curve $C(t) = p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))$ through the space of $\mathfrak{M}$ with open interval domains containing $\mathbf{0}$, by which we have $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(0)) = p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$. We likewise assume that $C(t)$ is smooth enough such that the tangent vector $\dot{\ell}(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) = d \log p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))/dt$ exists with $E_{\boldsymbol{y},\boldsymbol{\varpi},\boldsymbol{\theta}}(\dot{\ell}(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))^2) < \infty$. For all possible smooth $C(t)$, we formed the tangent space $T_\omega\mathfrak{M}$ based on the tangent vectors $\dot{\ell}(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(0)) = d \log C(t)/dt|_{t=0}$, which satisfies $\int \dot{\ell}(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(0))C(0)dzd\theta = 0$. For any two tangent vectors, $v1(\boldsymbol{\omega})$, $v2(\boldsymbol{\omega})$, we have the inner product

$$< v1(\boldsymbol{\omega}), v2(\boldsymbol{\omega}) >= \int v1(\boldsymbol{\omega}) \cdot v2(\boldsymbol{\omega}) p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}) d\boldsymbol{y} d\boldsymbol{\varpi} d\boldsymbol{\theta}. \tag{9}$$

Consider a perturbation vector $\boldsymbol{\omega} = (\omega_j)$, let $\partial_{\omega_j} = \partial/\partial_{\omega_j}$ be the partial derivatives, the inner product of tangent vectors $\partial_{\omega_i}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$ and $\partial_{\omega_j}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$ is

$$\begin{aligned} g_{ij}(\boldsymbol{\omega}) &= E_{\boldsymbol{y},\boldsymbol{\varpi},\boldsymbol{\theta}} \left[\partial_{\omega_i}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})\partial_{\omega_j}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})\right] \\ &= -E_{\boldsymbol{y},\boldsymbol{\varpi},\boldsymbol{\theta}} \left[\partial^2_{\omega_j\omega_k}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})\right], \end{aligned} \tag{10}$$

where $E_{\boldsymbol{y},\boldsymbol{\varpi},\boldsymbol{\theta}}[\cdot]$ represents the expectation with respect to the distribution $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$. The metric matrix $G(\boldsymbol{\omega}) = (g_{ij}(\boldsymbol{\omega}))$ is associated with the expected Fisher information matrix. Thus, the elements of $G(\boldsymbol{\omega})$ can measure the amount of perturbation to the model. An appropriate perturbation to the semiparametric SEM requires that $G(\boldsymbol{\omega})$ is a diagonal matrix [32], we then interpret the diagonal element $g_{jj}(\boldsymbol{\omega})$ as the amount of perturbations generated by ω_j.

Next, we consider an appropriate objective function $f(\boldsymbol{\omega}, \boldsymbol{\omega}^0)$, which assesses the sensitivity of the interested inference on the perturbation under consideration. Notably, a large value of the objective function can be caused in two ways. The first way is the discrepancy between the observed data and the fitted model, which is what we need to detect; whereas the second way is the deviance between the perturbed distribution and the baseline distribution, which is independent of what we observe. Thus, the local influence method developed for quantifying the effect of perturbation in the Bayesian analysis should be rescaled by dividing the minimal geodesic distance between $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$ and $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}^0)$, denoted by $d(\boldsymbol{\omega}, \boldsymbol{\omega}^0)$. As a result, the influence measure for comparing $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$ and $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}^0)$ is defined as

$$IM_{f(\boldsymbol{\omega},\boldsymbol{\omega}^0)} = \frac{f(\boldsymbol{\omega}, \boldsymbol{\omega}^0)^2}{d(\boldsymbol{\omega}, \boldsymbol{\omega}^0)^2}. \tag{11}$$

3.2 Local influence measures

As pointed out by [33], all possible smooth curves $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))$ pass through $\boldsymbol{\omega}^0 = \boldsymbol{\omega}(0)$, the local behaviour of $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}^0) = f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$ can then be measured by $lim_{t\to 0} IM_{f(\boldsymbol{\omega}(t),\boldsymbol{\omega}(0))}$. Since the Taylor's expansion of $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$ at 0 is equal to $f(\boldsymbol{\omega}(0), \boldsymbol{\omega}(0)) + \dot{f}(\boldsymbol{\omega}(0)) + \frac{1}{2}\ddot{f}(\boldsymbol{\omega}(0))t^2 + o(t^2)$, where $\dot{f}(\boldsymbol{\omega}(0))$ and $\ddot{f}(\boldsymbol{\omega}(0))$ denote the first- and second-order derivatives of $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$ with respect to t evaluated at 0. The first-order local influence measure is defined as

$$FI_f(\boldsymbol{\omega}(0)) = \lim_{t\to 0} IM_{f(\boldsymbol{\omega}(t),\boldsymbol{\omega}(0))} = \lim_{t\to 0} \frac{f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))^2}{d(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))^2} = \frac{[\boldsymbol{v}^T \partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))]^2}{\boldsymbol{v}^T G(\boldsymbol{\omega}(0))\boldsymbol{v}}, \quad (12)$$

where $\boldsymbol{v} = d\boldsymbol{\omega}(t)/dt|_{t=0}$, $\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$ denotes the first-order derivatives of $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$ with respect to $\boldsymbol{\omega}(t)$ evaluated at $t = 0$ and $G(\boldsymbol{\omega}(0))$ is Fisher information matrix formed by g_{ij} in (10). Based on the Cauchy-Schwarta inequality and Chen et al. [3] we can prove that

$$arg \max_{\boldsymbol{v}} FI_f(\boldsymbol{\omega}(0)) = [G(\boldsymbol{\omega}(0))]^{-1}[\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))]^2, \quad (13)$$

where $\boldsymbol{v} = [G(\boldsymbol{\omega}(0))]^{-1}\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$, and $arg \max_{\boldsymbol{v}} FI_f(\boldsymbol{\omega}(0))$ is a vector of the maximum first-order local inference measures of all components.

If the first-order local influence $FI_f(\boldsymbol{\omega}(0))$ equals 0, following Zhu et al. [33], we can use $\ddot{f}(\boldsymbol{\omega}(0))$ to assess the second-order local influence measure of perturbation $\boldsymbol{\omega}$ to the Baeysian model. We consider a specific smooth curve $C(t) = Exp_{\boldsymbol{\omega}(0)}(tv)$, where the $Exp_{\boldsymbol{\omega}(0)}(tv)$ stands for the geodesic $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))$, passing through $Exp_{\boldsymbol{\omega}(0)}(tv)|_{t=0} = \boldsymbol{\omega}(0)$. Thus, we have the Taylor's series expansion of object function as follows:

$$f(Exp_{\boldsymbol{\omega}(0)}(tv), \boldsymbol{\omega}(0)) = f(\boldsymbol{\omega}(0), \boldsymbol{\omega}(0)) + t\dot{f}(\boldsymbol{\omega}(0)) + \frac{1}{2}t^2\ddot{f}(Exp_{\boldsymbol{\omega}(0)}(tv))|_{t=0} + o(t^2), \quad (14)$$

where $\ddot{f}(Exp_{\boldsymbol{\omega}(0)}(tv)) = d^2 f(Exp_{\boldsymbol{\omega}(0)}(tv), \boldsymbol{\omega}(0))/dt^2$ called the Riemannian Hessian [32]. Then, we can introduce the second-order influence measure for perturbation vector $\boldsymbol{\omega}$ as follows:

$$SI_f(\boldsymbol{\omega}(0)) = \lim_{t\to 0} \frac{f(Exp_{\boldsymbol{\omega}(0)}(tv), \boldsymbol{\omega}(0))^2}{d(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))^2} = \frac{\boldsymbol{v}^T \partial^2_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))\boldsymbol{v}}{\boldsymbol{v}^T G(\boldsymbol{\omega}(0))\boldsymbol{v}}, \quad (15)$$

where $\partial^2_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$ denotes the second-order derivatives of $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$ with respect to $\boldsymbol{\omega}(t)$ evaluate at $\boldsymbol{\omega}(0)$. Similar to that of first-order influence measure, we have the maximum second-order local influence measures of all components as follows:

$$arg \max_{\boldsymbol{v}} SI_f(\boldsymbol{\omega}(0)) = diag\{G(\boldsymbol{\omega}(0))]^{-1}[\partial^2_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))\}, \quad (16)$$

where $\boldsymbol{v}$ is the eigenvector of $[G(\boldsymbol{\omega}(0))]^{-1}\partial^2_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$ corresponding to the largest eigenvalue.

Regarding the choice of the objective function, Ibrahim et al. [12] have introduced several candidates such as the Bayes factor, $\phi-$divergence, and the posterior mean distance. In the present study, we focus on the Bayes factor because $arg \max_v FI_f(\boldsymbol{\omega}(0)) \neq 0$ when the objective function is taken as the Bayes factor. Consequently, the computation of the second-order influence measure is unnecessary. We consider the logarithm of the Bayes factor as follows:

$$\begin{aligned} B(\boldsymbol{\omega}(0)) &= \log p(\boldsymbol{y}|\boldsymbol{\omega}(t)) - \log p(\boldsymbol{y}|\boldsymbol{\omega}(0)) \\ &= \log \int p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))\, d\boldsymbol{\varpi}\, d\boldsymbol{\theta} - \log \int p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(0))\, d\boldsymbol{\varpi}\, d\boldsymbol{\theta}. \end{aligned} \quad (17)$$

Based on the (13), $\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$ needs to be calculated. By assuming the legitimacy of interchange of integration with differentiation, we have

$$\begin{aligned}
\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0)) &= \partial_{\boldsymbol{\omega}(t)} B(\boldsymbol{\omega}(t))|_{t=0} \\
&= \frac{\int \partial_{\boldsymbol{\omega}(t)} p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) d\boldsymbol{\varpi}\, d\boldsymbol{\theta}}{\int p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) d\boldsymbol{\varpi}\, d\boldsymbol{\theta}}|_{t=0} \\
&= \frac{\int \partial_{\boldsymbol{\omega}(t)} \ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) d\boldsymbol{\varpi}\, d\boldsymbol{\theta}}{\int p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) d\boldsymbol{\varpi}\, d\boldsymbol{\theta}}|_{t=0} \\
&= \int \partial_{\boldsymbol{\omega}} \ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) \frac{p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))}{\int p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) d\boldsymbol{\varpi}\, d\boldsymbol{\theta}} d\boldsymbol{\varpi}\, d\boldsymbol{\theta}|_{t=0} \\
&= \int \partial_{\boldsymbol{\omega}} \ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) p(\boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{\omega}(t)) d\boldsymbol{\varpi}\, d\boldsymbol{\theta}|_{t=0} \\
&= E_{\boldsymbol{\varpi}, \boldsymbol{\theta}} \left[\partial_{\boldsymbol{\omega}} \ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) \right] |_{t=0},
\end{aligned} \tag{18}$$

where $E_{\boldsymbol{\varpi},\boldsymbol{\theta}}[\cdot]$ denotes the expectation with respect to the posterior distribution $p(\boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{\omega}(t))$.

3.3 Prior distribution

An important issue in implementing a full Bayesian analysis is to specify the prior distribution of the unknown parameters. Firstly, we discuss the regression coefficients in equation (8). For the univariate functions of covariates and latent variables, random walk priors coped with the constrained Gaussian density are assigned to $\boldsymbol{\gamma}_d = (\gamma_{d1}, \cdots, \gamma_{dK_d^z})$ and $\boldsymbol{\beta}_j = (\beta_{j1}, \cdots, \beta_{jK_j})$ as follows:

$$p(\boldsymbol{\gamma}_d|\tau_{\boldsymbol{\gamma}_d}) = \left(\frac{\tau_{\boldsymbol{\gamma}_d}}{2\pi}\right)^{K_d^*/2} \exp\left\{-\frac{\tau_{\boldsymbol{\gamma}_d}}{2} \boldsymbol{\gamma}_d^\top \mathcal{M}_{\boldsymbol{\gamma}_d} \boldsymbol{\gamma}_d\right\}, \tag{19}$$

$$p(\boldsymbol{\beta}_j|\tau_{\boldsymbol{\beta}_j}) = \left(\frac{\tau_{\boldsymbol{\beta}_j}}{2\pi}\right)^{K_j^*/2} \exp\left\{-\frac{\tau_{\boldsymbol{\beta}_j}}{2} \boldsymbol{\beta}_j^\top \mathcal{M}_{\boldsymbol{\beta}_j} \boldsymbol{\beta}_j\right\}, \tag{20}$$

where $K_d^* = rank(\boldsymbol{M}_{\boldsymbol{\gamma}_d})$ and $K_j^* = rank(\boldsymbol{M}_{\boldsymbol{\beta}_j})$, $\boldsymbol{M}_{\boldsymbol{\gamma}_d}$ and $\boldsymbol{M}_{\boldsymbol{\beta}_j}$ are the penalty matrices derived according to the random walk priors [25], $\tau_{\boldsymbol{\gamma}_d}$ and $\tau_{\boldsymbol{\beta}_j}$ are the inverse smoothing parameters that control the amount of penalty. For the regression coefficients associated with the interactions terms, to avoid over-smoothing of the bivariate functions of latent variables, we set the following prior distributions [26]:

$$\begin{aligned}
p(\boldsymbol{b}_{uv}|\tau_1, \tau_2) = \;& \prod_{s=1}^{S} \prod_{t=2}^{T} [\sqrt{\tfrac{\vartheta_{1st}\tau_1}{2\pi}} \exp\{-\tfrac{\vartheta_{1st}\tau_1 (b_{uvst} - b_{uvs(t-1)})^2}{2}\}] \times \\
& \prod_{t=1}^{T} \prod_{s=2}^{S} [\sqrt{\tfrac{\vartheta_{2st}\tau_2}{2\pi}} \exp\{-\tfrac{\vartheta_{2st}\tau_2 (b_{uvst} - b_{uvs(t-1)})^2}{2}\}],
\end{aligned} \tag{21}$$

where τ_1 and τ_2 follow the gamma distribution $Gamma(\alpha_{\boldsymbol{b}_{uv}}, \beta_{\boldsymbol{b}_{uv}})$, and ϑ_{lkr}s follow the uniform distribution $U(0, 1)$.

Secondly, we specify the prior distributions for the structural parameters such as $\boldsymbol{A}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}$, ψ_δ, and $\boldsymbol{\Phi}^*$. The following conjugate prior distributions [20] are considered. For $j = 1, \cdots, p$,

$$\begin{aligned}
&\boldsymbol{A}_j \sim \mathcal{N}[\boldsymbol{A}_{j0}, \boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}}], \quad \boldsymbol{\Lambda}_j \sim \mathcal{N}[\boldsymbol{\Lambda}_{j0}, \psi_j \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}}], \\
&\psi_j^{-1} \sim Gamma[\alpha_{j0}, \beta_{j0}], \quad \psi_\delta^{-1} \sim Gamma[\alpha_{\delta 0}, \beta_{\delta 0}], \\
&\boldsymbol{\Phi}^{*^{-1}} \sim Wishart[\boldsymbol{R}_0, \rho_0],
\end{aligned}$$

where $\boldsymbol{A}_j^T$ and $\boldsymbol{\Lambda}_j^T$ denote the jth row of $\boldsymbol{A}_j$ and $\boldsymbol{\Lambda}$, respectively; ψ_j is the jth diagonal element of $\boldsymbol{\Psi}$; $\boldsymbol{A}_{j0}$, $\boldsymbol{\Lambda}_{j0}$, α_{j0}, β_{j0}, $\alpha_{\delta 0}$, $\beta_{\delta 0}$, ρ_0, as well as positive definite matrices $\boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}}$ and $\boldsymbol{R}_0$, are hyperparameters with preassigned values.

3.4 Posterior computation

To calculate the first- and second-order local influence measures, we need to calculate the expectation $E_{\boldsymbol{\varpi},\boldsymbol{\theta}}(\cdot)$ in (10) and (18). However, this expectation is intractable because it involves high-dimensional integration. Thus, we employ MCMC methods such as the Gibbs sampler [9] and the Metropolis-Hastings algorithm [11, 22] to conduct the Bayesian estimation and local influence analysis. The Gibbs sampler algorithm can be implemented as follows: at the t iteration with current values $\{\boldsymbol{\varpi}^{(t)}, \boldsymbol{\theta}^{(t)}\}$: (a) draw $\boldsymbol{\varpi}$ from $p(\boldsymbol{\varpi}|\boldsymbol{Y}, \boldsymbol{A}^{(t)}, \boldsymbol{\Lambda}^{(t)}, \boldsymbol{\Psi}^{(t)}, \psi_\delta^{(t)}, \boldsymbol{\Phi}^{*(t)}, \boldsymbol{\gamma}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{b}^{(t)})$; (b) draw $\boldsymbol{\theta}^{(t)}$ from $p(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{\varpi}^{(t+1)})$. Due to its complexity, step (b) is further decomposed into: (b1) draw $\boldsymbol{A}^{(t+1)}$ from $p(\boldsymbol{A}|\boldsymbol{Y}, \boldsymbol{\varpi}^{(t+1)}, \boldsymbol{\Lambda}^{(t)}, \boldsymbol{\Psi}^{(t)})$; (b2) draw $(\boldsymbol{\Lambda}^{(t+1)}, \boldsymbol{\Psi}^{(t+1)})$ from $p(\boldsymbol{\Lambda}, \boldsymbol{\Psi}|\boldsymbol{Y}, \boldsymbol{\varpi}^{(t+1)}, \boldsymbol{A}^{(t+1)})$; (b3) draw $\psi_\delta^{(t+1)}$ from $p(\psi_\delta|\boldsymbol{\varpi}^{(t+1)}, \boldsymbol{\gamma}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{b}^{(t)})$; (b4) draw $\boldsymbol{\Phi}^{*(t+1)}$ from $p(\boldsymbol{\Phi}^*|\boldsymbol{\varpi}^{(t+1)})$; (b5) draw $\boldsymbol{\gamma}^{(t+1)}$ from $p(\boldsymbol{\gamma}|\boldsymbol{\varpi}^{(t+1)}, \psi_\delta^{(t+1)}, \boldsymbol{\tau_\gamma}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{b}^{(t)})$; (b6) draw $\boldsymbol{\beta}^{(t+1)}$ from $p(\boldsymbol{\gamma}|\boldsymbol{\varpi}^{(t+1)}, \psi_\delta^{(t+1)}, \boldsymbol{\tau_\beta}^{(t)}, \boldsymbol{\gamma}^{(t+1)}, \boldsymbol{b}^{(t)})$; (b7) draw $\boldsymbol{b}^{(t+1)}$ from $p(\boldsymbol{b}|\boldsymbol{\varpi}^{(t+1)}, \psi_\delta^{(t+1)}, \boldsymbol{\tau}_1^{(t)}, \boldsymbol{\tau}_2^{(t)}, \boldsymbol{\gamma}^{(t+1)}, \boldsymbol{\vartheta}^{(t)}, \boldsymbol{\beta}^{(t+1)})$; (b8) draw $\boldsymbol{\tau_\gamma}^{(t+1)}$ from $p(\boldsymbol{\tau_\gamma}|\boldsymbol{\gamma}^{(t+1)})$ and $\boldsymbol{\tau_\beta}^{(t+1)}$ from $p(\boldsymbol{\tau_\beta}|\boldsymbol{\beta}^{(t+1)})$; (b9) draw $\boldsymbol{\tau}_1^{(t+1)}$ from $p(\boldsymbol{\tau}_1|\boldsymbol{b}^{(t+1)}, \boldsymbol{\vartheta}^{(t)})$ and $\boldsymbol{\tau}_2^{(t+1)}$ from $p(\boldsymbol{\tau}_2|\boldsymbol{b}^{(t+1)}, \boldsymbol{\vartheta}^{(t)})$; and (b10) draw $\boldsymbol{\vartheta}^{(t+1)}$ from $p(\boldsymbol{\vartheta}|\boldsymbol{\tau}_1^{(t+1)}, \boldsymbol{\tau}_2^{(t+1)}, \boldsymbol{b}^{(t+1)})$.

The full conditional distributions involved in (b1)–(b9) are the normal, inverse gamma, and the inverse Wishart distributions, respectively. Sampling from them is efficient and straightforward. However, those involved in (a) and (b10) are complex. In particular, the conditional distribution in (b10) is the gamma distribution truncated in $[0, 1]$. Sampling from these non-standard distributions requires additional sampling technique such as the Metropolis-Hastings algorithm.

3.5 Perturbation schemes

In this section, we discuss the computation of the first-order local influence. The main task is to compute $G(\boldsymbol{\omega}(0))$ involved in the Fisher information matrix (10) and $\partial_{\boldsymbol{\omega}}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))|_{t=0}$ involved in the Bayes factor (18). Considering that $\boldsymbol{\omega}_y(t)$, $\boldsymbol{\omega}_s(t)$, and $\boldsymbol{\omega}_p(t)$ denote the perturbation to the data, sampling distributions, and the prior distributions of parameters, respectively, they are independent of each other. Thus, we have $\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t)) = \ell(\boldsymbol{y}|\boldsymbol{\varpi}, \boldsymbol{\theta}, \boldsymbol{\omega}_y(t)) + \ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\omega}_s(t)) + \ell(\boldsymbol{\xi}|\boldsymbol{\theta}) + \ell(\boldsymbol{\theta}|\boldsymbol{\omega}_p(t))$, and we can separately discuss $G(\boldsymbol{\omega}(0))$ and $\partial_{\boldsymbol{\omega}}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(t))|_{t=0}$ that are related to the data, sampling distributions, or the prior distributions of parameters only.

First, we discuss the perturbation schemes corresponding to the observed-data log-likelihood. Let $\boldsymbol{\omega}_y(t) = (\omega_{y_1}(t), \cdots, \omega_{y_n}(t))^\top$ and $\boldsymbol{\omega}(0) = (1, \cdots, 1)^\top$. The perturbation scheme for the observed data is given by

$$\ell(\boldsymbol{y}|\boldsymbol{\varpi}, \boldsymbol{\theta}, \boldsymbol{\omega}_y(t)) = -\tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left\{\omega_{y_i}(t)(y_{ij} - \boldsymbol{A_j}^T\boldsymbol{x_i} - \boldsymbol{\Lambda_j}^T\boldsymbol{\varpi_i})^2/\psi_j + \log(\tfrac{2\pi\psi_j}{\omega_{y_i}})\right\},$$

$$\partial_{\omega_{y_i}(tp)}\ell(\boldsymbol{y}_i|\boldsymbol{\varpi}_i, \boldsymbol{\theta}, \boldsymbol{\omega}_y(t))|_{t=0} = -\tfrac{1}{2}\sum_{j=1}^{p}\left\{(y_{ij} - \boldsymbol{A_j}^T\boldsymbol{x_i} - \boldsymbol{\Lambda_j}^T\boldsymbol{\varpi_i})^2/\psi_j - 1\right\},$$

$$G(\boldsymbol{\omega}_y(0)) = \tfrac{p}{2}\boldsymbol{I}.$$

By directly perturbing the observed data log-likelihood, we can diagnose the overall impact of outliers or influential observations on the model.

Second, to diagnose whether the nonparametric structural equation (8) is adequate to illustrate the nonlinear relationships among the observable and latent explanatory variables on the outcome latent variables, we discuss the perturbation $\boldsymbol{\omega}_s(t)$ corresponding to the sampling distribution. For the purpose of illustration, we take the following simple structural equation as an example:

$$\eta_i = g(z_i) + f_1(\xi_{i1}) + f_2(\xi_{i2}) + h_{12}(\xi_{i1}, \xi_{i2}) + \delta_i. \quad (22)$$

Perturbation of covariates z_i : Let $\omega_{s_1}(t)$ denote a perturbation and $\omega_{s_1}(0) = 0$, then

$$\begin{aligned}
\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_1}(t)) &= -\tfrac{1}{2}\sum_{i=1}^{n}\{(\eta_i - \omega_{s_1}(t)g(z_i) - f_1(\xi_{i1}) - f_2(\xi_{i2}) - h_{12}(\xi_{i1}, \xi_{i2}))^2/\psi_\delta \\
&\quad + \log(2\pi\psi_\delta)\}, \\
\partial_{\omega_{s_1}(t)}\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_1}(t))|_{t=0} &= \tfrac{1}{\psi_\delta}\sum_{i=1}^{n} g(z_i)(\eta_i - f_1(\xi_{i1}) - f_2(\xi_{i2}) - h_{12}(\xi_{i1}, \xi_{i2})), \\
G_{z_i p}(\boldsymbol{\omega}_{s_1}(0)) &= \sum_{i=1}^{n} E_{\boldsymbol{\varpi},\boldsymbol{\theta}}[\tfrac{1}{\psi_\delta}g(z_i)^2].
\end{aligned}$$

Perturbation of latent variable ξ_{i1} : Let $\omega_{s_2}(t)$ denote a perturbation and $\omega_{s_2}(0) = 0$, then

$$\begin{aligned}
\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_2}(t)) &= -\tfrac{1}{2}\sum_{i=1}^{n}\{(\eta_i - g(z_i) - \omega_{s_2}(t)[f_1(\xi_{i1}) + h_{12}(\xi_{i1}, \xi_{i2})] - f_2(\xi_{i2}))^2/\psi_\delta \\
&\quad + \log(2\pi\psi_\delta)\}, \\
\partial_{\omega_{s_2}(t)}\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_2}(t))|_{t=0} &= \tfrac{1}{\psi_\delta}\sum_{i=1}^{n}(f_1(\xi_{i1}) + h_{12}(\xi_{i1}, \xi_{i2}))(\eta_i - g(z_i) - f_2(\xi_{i2})), \\
G_{\xi_{i1}}(\boldsymbol{\omega}_{s_2}(0)) &= \sum_{i=1}^{n} E_{\boldsymbol{\varpi},\boldsymbol{\theta}}[\tfrac{1}{\psi_\delta}(f_1(\xi_{i1}) + h_{12}(\xi_{i1}, \xi_{i2}))^2].
\end{aligned}$$

Perturbation of latent variable ξ_{i2} : Let $\omega_{s_3}(t)$ denote a perturbation and $\omega_{s_3}(0) = 0$, then

$$\begin{aligned}
\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_3}(t)) &= -\tfrac{1}{2}\sum_{i=1}^{n}\{(\eta_i - g(z_i) - f_1(\xi_{i1}) - \omega_{s_3}(t)[f_2(\xi_{i2}) + h_{12}(\xi_{i1}, \xi_{i2})])^2/\psi_\delta \\
&\quad + \log(2\pi\psi_\delta)\}, \\
\partial_{\omega_{s_3}(t)}\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_3}(t))|_{t=0} &= \tfrac{1}{\psi_\delta}\sum_{i=1}^{n}(f_2(\xi_{i2}) + h_{12}(\xi_{i1}, \xi_{i2}))(\eta_i - g(z_i) - f_1(\xi_{i1})), \\
G_{\xi_{i2}}(\boldsymbol{\omega}_{s_3}(0)) &= \sum_{i=1}^{n} E_{\boldsymbol{\varpi},\boldsymbol{\theta}}[\tfrac{1}{\psi_\delta}(f_2(\xi_{i2}) + h_{12}(\xi_{i1}, \xi_{i2}))^2].
\end{aligned}$$

Perturbation of interaction $h_{12}(\xi_{i1}, \xi_{i2})$ Let $\omega_{s_4}(t)$ denote a perturbation and $\omega_{s_4}(0) = 0$, then

$$\begin{aligned}
\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_4}(t)) &= -\tfrac{1}{2}\sum_{i=1}^{n}\{(\eta_i - g(z_i) - f_1(\xi_{i1}) - f_2(\xi_{i2}) - \omega_{s_4}(t)h_{12}(\xi_{i1}, \xi_{i2}))^2/\psi_\delta \\
&\quad + \log(2\pi\psi_\delta)\}, \\
\partial_{\omega_{s_4}(t)}\ell(\boldsymbol{\eta}|\boldsymbol{\xi}, \boldsymbol{\theta}, \omega_{s_4}(t))|_{t=0} &= \tfrac{1}{\psi_\delta}\sum_{i=1}^{n} h_{12}(\xi_{i1}, \xi_{i2})(\eta_i - g(z_i) - f_1(\xi_{i1}) - f_2(\xi_{i2})), \\
G_{(\xi_{i1},\xi_{i2})}(\boldsymbol{\omega}_{s_3}(0)) &= \sum_{i=1}^{n} E_{\boldsymbol{\varpi},\boldsymbol{\theta}}[\tfrac{1}{\psi_\delta}h_{12}(\xi_{i1}, \xi_{i2})^2].
\end{aligned}$$

Finally, to assess the sensitivity of Bayesian results to prior inputs, we discuss the perturbation schemes corresponding to the prior distributions of model parameters. We consider the structural parameters of primary interest, including regression-type parameters $\boldsymbol{A}$ and $\boldsymbol{\Lambda}$, as well as variance/covariance parameters $\boldsymbol{\Psi}$ and ψ_δ. The perturbation schemes are given below:

Perturbation of $\boldsymbol{A}_j$: Let $\omega_A(t)$ denote a perturbation and $\omega_A(0) = 1$, then

$$\begin{aligned}
\ell(\boldsymbol{A}_j|\boldsymbol{A}_{j0}, \boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}}, \omega_A(t)) &= -\tfrac{1}{2}\Big\{p_{\boldsymbol{A}_j}\log(\tfrac{2\pi}{\omega_A(t)}) + \log|\boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}}| \\
&\quad + \omega_A(t)(\boldsymbol{A}_j - \boldsymbol{A}_{j0})^\top(\boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}})^{-1}(\boldsymbol{A}_j - \boldsymbol{A}_{j0})\Big\}, \\
\partial_{\omega_A(t)}\ell(\boldsymbol{A}|\boldsymbol{A}_{j0}, \boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}}, \omega_A(t))|_{t=0} &= \tfrac{1}{2}\Big\{p_{\boldsymbol{A}} - \sum_{j=1}^{p}(\boldsymbol{A}_j - \boldsymbol{A}_{j0})^\top(\boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}})^{-1}(\boldsymbol{A}_j - \boldsymbol{A}_{j0})\Big\}, \\
G_{\boldsymbol{A}}(\omega_A(0)) &= p_{\boldsymbol{A}}/2,
\end{aligned}$$

where $p_{\boldsymbol{A}_j}$ denotes the dimension of $\boldsymbol{A}_j$ and $p_{\boldsymbol{A}} = \sum_{j=1}^{p} p_{\boldsymbol{A}_j}$.
Perturbation of $\boldsymbol{\Lambda}_j$: Let $\omega_\Lambda(t)$ denote a perturbation and $\omega_\Lambda(0) = 1$, then

$$\begin{aligned}
&\ell(\boldsymbol{\Lambda}_j|\boldsymbol{\Lambda}_{j0}, \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}}, \omega_\Lambda(t)) = -\tfrac{1}{2}\Big\{p_{\boldsymbol{\Lambda}_j}\log(\tfrac{2\pi}{\omega_\Lambda(t)}) + \log|\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}}| \\
&\qquad\qquad + \omega_\Lambda(t)(\boldsymbol{\Lambda}_j - \boldsymbol{\Lambda}_{j0})^\top(\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}})^{-1}(\boldsymbol{\Lambda}_j - \boldsymbol{\Lambda}_{j0})\Big\}, \\
&\partial_{\omega_\Lambda(t)}\ell(\boldsymbol{\Lambda}|\boldsymbol{\Lambda}_{j0}, \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}}, \omega_\Lambda(t))|_{t=0} = \tfrac{1}{2}\Big\{p_{\boldsymbol{\Lambda}} - \sum_{j=1}^{p}(\boldsymbol{\Lambda}_j - \boldsymbol{\Lambda}_{j0})^\top(\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}})^{-1}(\boldsymbol{\Lambda}_j - \boldsymbol{\Lambda}_{j0})\Big\}, \\
&G_{\boldsymbol{\Lambda}}(\omega_\Lambda(0)) = p_{\boldsymbol{\Lambda}}/2,
\end{aligned}$$

where $p_{\boldsymbol{\Lambda}_j}$ denotes the number of parameters of $\boldsymbol{\Lambda}_j$ excluding the fixed elements and $p_{\boldsymbol{\Lambda}} = \sum_{j=1}^{p} p_{\boldsymbol{\Lambda}_j}$.

Perturbation of $\boldsymbol{\Psi}$: Let $\omega_{\boldsymbol{\Psi}}(t)$ denote a perturbation and $\omega_{\boldsymbol{\Psi}}(t) = 1$, then

$$\begin{aligned}
&\ell(\psi_j^{-1}|\alpha_{j0}, \beta_{j0}, \omega_{\boldsymbol{\Psi}}(t)) = \alpha_{j0}\log(\omega_{\boldsymbol{\Psi}}(t)\beta_{j0}) - \log(\Gamma(\alpha_{j0})) + (\alpha_{j0} - 1)\log(\psi_j^{-1}) \\
&\qquad\qquad - \omega_{\boldsymbol{\Psi}}(t)\beta_{j0}\psi_j^{-1}, \\
&\partial_{\omega_{\boldsymbol{\Psi}}(t)}\ell(\boldsymbol{\Psi}^{-1}|\alpha_{j0}, \beta_{j0}, \omega_{\boldsymbol{\Psi}}(t))|_{t=0} = \sum_{j=1}^{p}[\alpha_{j0} - \beta_{j0}\psi_j^{-1}], \\
&G_{\omega_{\boldsymbol{\Psi}}(0)}(\omega_\Lambda(0)) = \sum_{j=1}^{p}\alpha_{j0},
\end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function.
Perturbation of ψ_δ: Let $\omega_{\psi_\delta}(t)$ denote a perturbation and $\omega_{\psi_\delta}(t) = 1$, then

$$\begin{aligned}
&\ell(\psi_\delta^{-1}|\alpha_{\delta 0}, \beta_{\delta 0}, \omega_{\psi_\delta}(t)) = \alpha_{\delta 0} log(\omega_{\psi_\delta}(t)\beta_{\delta 0}) - \log(\Gamma(\alpha_{\delta 0})) + (\alpha_{\delta 0} - 1)\log(\psi_\delta^{-1}) \\
&\qquad\qquad - \omega_{\psi_\delta}(t)\beta_{\delta 0}\psi_\delta^{-1}, \\
&\partial_{\omega_{\psi_\delta}(t)}\ell(\psi_\delta^{-1}|\alpha_{\delta 0}, \beta_{\delta 0}, \omega_{\psi_\delta}(t))|_{t=0} = \alpha_{\delta 0} - \beta_{\delta 0}\psi_\delta^{-1}, \\
&G_{\omega_{\psi_\delta}(0)}(\omega_\Lambda(0)) = \alpha_{\delta 0},
\end{aligned}$$

For a specific perturbation scheme and an objective function, we use the following steps to implement Bayesian influence analysis:

Step 1. Construct a Bayesian perturbation model $p(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega})$ and choose the objective function $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$.

Step 2. Calculate $\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$, $\partial^2_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0))$, and $G(\boldsymbol{\omega}(0)) = (\partial^2_{\omega_j\omega_k}\ell(\boldsymbol{y}, \boldsymbol{\varpi}, \boldsymbol{\theta}|\boldsymbol{\omega}(0)))$.

Step 3. If $\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0)) \neq 0$, we calculate $arg\max_{\boldsymbol{v}} FI_f(\boldsymbol{\omega}(\boldsymbol{0}))$, the first-order local influence measure. If $\partial_{\boldsymbol{\omega}} f(\boldsymbol{\omega}(0)) = 0$, we calculate $arg\max_{\boldsymbol{v}} SI_f(\boldsymbol{\omega}(\boldsymbol{0}))$, the second-order local influence measure.

Step 4. Let $M(0)_j$ be the jth element of $arg\max_{\boldsymbol{v}} FI_f(\boldsymbol{\omega}(\boldsymbol{0}))$ or $arg\max_{\boldsymbol{v}} SI_f(\boldsymbol{\omega}(\boldsymbol{0}))$, and $\bar{C} = \bar{M}(0) + 3SM(0)$, where $\bar{M}(0)$ and $SM(0)$ are the mean and the standard deviation of $\{M(0)_j : j = 1, \cdots, m\}$. For a selected objective function $f(\boldsymbol{\omega}(t), \boldsymbol{\omega}(0))$, if $M(0)_j \geq \bar{C}$, then the jth observation is detected as influential.

4. Simulation Study

4.1 Simulation 1

In the first simulation study, we focus on the diagnosis of outliers and/or influential observations. We consider the perturbation corresponding to the data, $\boldsymbol{\omega} = \boldsymbol{\omega}_y$, and a semiparametric SEM with $p = 9$, $q = 3$, $q_1 = 1$, and $q_2 = 2$. For $i = 1, \cdots, 1500$ and $j = 1, \cdots, 9$,

$$y_{ij} = \boldsymbol{A_j}^T \boldsymbol{x_i} + \boldsymbol{\Lambda_j}^T \boldsymbol{\varpi_i} + \epsilon_{ij}, \quad (23)$$

$$\eta_i = g_1(z_{i1}) + g_2(z_{i2}) + f_1(\xi_{i1}) + f_2(\xi_{i2}) + h_{12}(\xi_{i1}, \xi_{i2}) + \delta_i, \quad (24)$$

where $g_1(z_{i1}) = 2z_{i1}^3$, $g_2(z_{i2}) = 0.2\exp(3z_{i2}) - 2/3$, $f_1(\xi_{i1}) = 3\exp(3\xi_{i1})/[1+3\exp(3\xi_{i1})]$ -1.5, $f_2(\xi_{i2}) = \sin(1.5\xi_{i2} - \xi_{i2})$, $h_{12}(\xi_{i1}, \xi_{i2}) = 5(\Phi(\xi_{i1}) - 0.5)\cos(2\pi\Phi(\xi_{i2}))$, $\Phi(\cdot)$ is the cumulative distribution function of $N(0, 1)$; x_i, z_{i1}, and z_{i2} are generated from uniform distribution $U(-1, 1)$,

$$\boldsymbol{\Lambda}_j^T = \begin{bmatrix} 1 & \lambda_{21} & \lambda_{31} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda_{52} & \lambda_{62} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{83} & \lambda_{93} \end{bmatrix}$$

in which 1s and 0s are fixed to obtain an identified model, and λs are unknown factor loadings; $\boldsymbol{\Phi}^*$ is the covariance matrix of $\boldsymbol{\xi}_i = (\xi_{i1}, \xi_{i2})$. The true population values of parameters are $a_j = 0.6$, $\psi_j = \psi_\delta = 0.3$, $j = 1, \cdots, 9$; $\lambda_{21} = \lambda_{31} = \lambda_{52} = \lambda_{62} = \lambda_{83} = \lambda_{93} = 1$, and $\{\phi_{11}, \phi_{12}, \phi_{22}\}$ in $\boldsymbol{\Phi}^*$ are $\{1.0, 0.3, 1.0\}$. To obtain the influential cases, we replace the 500th observed variable as follows: for $j = 1, \cdots, 9$, $y_{500,j} = y_{500,j} + 1.5sign(y_{500,j})$.

The prior distributions discussed in Section 3.3 are used with the following prespecified hyperparameters: the elements of $\boldsymbol{A}_{j0}$ and $\boldsymbol{\Lambda}_{j0}$ are taken as 0 and 0.8, respectively; $\boldsymbol{\Sigma}_{\boldsymbol{A}_{j0}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{j0}}$ are taken as the identity matrices of proper dimensions; $\alpha_{j0} = 6$, $\alpha_{\delta 0} = 9$, $\beta_{j0} = 1.8$, $\beta_{\delta 0} = 4$; $\rho_0 = 7$, $\boldsymbol{R}_0 = 3\boldsymbol{I}_2$, where $\boldsymbol{I}_2$ is the two-dimensional identity matrix; $\alpha_{\tau_{\boldsymbol{\gamma}_d}} = \alpha_{\tau_{\boldsymbol{\beta}_d}} = \alpha_{\boldsymbol{b}_{uv}} = 1$ and $\beta_{\tau_{\boldsymbol{\gamma}_d}} = \beta_{\tau_{\boldsymbol{\beta}_d}} = \beta_{\boldsymbol{b}_{uv}} = 0.005$. We utilize a total of 16 equidistant knots to construct $B_k^z(\cdot)$s, $B_k(\Phi(\cdot))$s, $U_{us}(\Phi(\cdot))$s and $V_{vt}(\Phi(\cdot))$s in the approximation of $g_1(z_{i1})$, $g_2(z_{i2})$, $f_1(\xi_{i1})$, $f_2(\xi_{i2})$, and $h_{12}(\xi_{i1}, \xi_{i2})$. The second-order and first-order random walk priors are used to prevent the overfitting in the estimation of the unknown univariate and bivariate smooth functions, respectively.

We carried out several pilot runs to check the convergence of the MCMC algorithm and found that it converged within 10,000 observations. Thus, we collected additional 10,000 observations after discarding the 10,000 burn-in iterations to obtain the Bayesian estimates of model parameters and to calculate the first-order local influence measures for various perturbation schemes. The obtained results are reported in Figure 1.

In the first graph, the data and the prior distributions of the four kinds of parameters discussed in Section 3.5 are simultaneously perturbed. Thus, we need to calculate a total of 1504 first-order local influence measures, wherein the 1st to 1500th elements correspond to the data, and the 1501st to 1504th elements correspond to the prior distributions. Based on the first graph, the 500th observation is detected as influential. However, the 1501st to 1504th measures are not significantly different from others, which implies that the Bayesian results are not sensitive to the given prior input.

In the second to sixth graph, the 1501th first-local influence measure correspond to $g_1(z_{i1})$, $g_2(z_{i2})$, $f_1(\xi_{i1})$, $f_2(\xi_{i2})$, and $h_{12}(\xi_{i1}, \xi_{i2})$, respectively. That is, we simultaneously perturb the data and the sampling distribution by introducing perturbation to one of the nonlinear terms of (24) in turn. The obtained results successfully detect the outlier and also show that the main effects of the covariates and the latent variables as well as the interaction effect between the two latent variables are all important and cannot be removed from the model.

4.2 Simulation 2

In this section, we conduct a simulation to examine whether or not the Bayesian local influence analysis erroneously diagnoses unimportant effects as influential. We consider a SEM, where the measurement equation is the same as that of Simulation 1, but the structural equation has the following simpler form:

$$\eta_i = g_1(z_{i1}) + f_1(\xi_{i1}) + \delta_i. \tag{25}$$

We generate the observed data using this simpler model and then perform the Bayesian local influence procedure based on larger model (24). The main purpose is to check whether the perturbation of the sampling distribution can correctly detect important effects such as $g_1(z_{i1})$ and $f_1(\xi_{i1})$, and discard unimportant ones such as $g_2(z_{i2})$, $f_2(\xi_{i2})$, and $h_{12}(\xi_{i1}, \xi_{i2})$. Here, except the data are generated based on equation (25), all the other settings are similar to those of Simulation 1. Likewise, in Figure 2 from the first to the last graph, the 1st to 1500th first-order local influence measures correspond to the data, and the 1501st corresponds to $g_1(z_{i1})$, $g_2(z_{i2})$, $f_1(\xi_{i1})$, $f_2(\xi_{i2})$, and $h_{12}(\xi_{i1}, \xi_{i2})$, respectively. As expected, the first graph simultaneously detects the outlier and the first important factor $g_1(z_{i1})$, the third graph simultaneously detects the outlier and the second important factor $f_1(\xi_{i1})$, whereas the other three graphs only show the outlier but do not display any sampling distribution-related influential point, implying that $g_2(z_{i2})$, $f_2(\xi_{i2})$, and $h_{12}(\xi_{i1}, \xi_{i2})$ that are perturbed respectively in these three graphs are nonsignificant and should be removed from the model. Thus, by introducing perturbation to the sampling distribution, the proposed Bayesian local influence procedure is capable of identifying the model structure, thereby playing a role that is similar to model/variable selection.

5. An illustrative study

In this section, we applied the proposed methodology to a study of osteoporosis prevention and control. A total of 1460 Chinese men aged 65 years and above were recruited using a combination of private solicitation and public advertising from community centers and public housing estates. A primary concern of this study is to investigate the functional relationships between bone mineral density (BMD) and its correlated determinants including 'Estrogen', 'Androgen', 'Precursors', and 'Metabolites'. However, BMD and its associated determinants are latent constructs that should be measured by multiple observed variables. The observed variables include spine BMD, hip BMD, estrone (E1), estrone sulphate (E1-S), estradiol (E2), testosterone (TESTO), 5-Androstenediol (5-DIOL), dihydrotestosterone (DHT), androstenedione (4-DIONE), dehydroepiandrosterone (DHEA), DHEA sulphate (DHEA-S), androsterone (ADT), ADT glucuronide (ADT-G), 3a-diol-3G (3G), and 3d-diol-17G (17G). According to medical knowledge, 'BMD, η', 'Estrogen, ξ_1', 'Androgen, ξ_2', 'Precursors, ξ_3', and 'Metabolites, ξ_4' are measured by $\{$Spine BMD, Hip BMD$\}$, $\{$E1, E1-S, E2$\}$, $\{$TESTO, 5-DIOL, DHT$\}$, $\{$4-DIONE, DHEA, DHEA-S$\}$, and $\{$ADT, ADT-G, 3G, 3G-17G$\}$, respectively. Due to the complex nature of the above mentioned latent variables, we expect that a simple linear SEM is inadequate to provide accurate analysis of the true functional relationships between latent variables. Hence, a semiparametric SEM was considered here.

In the measurement equation, $p = 15$, $q = 5$, $q_1 = 1$, $q_2 = 4$, $\boldsymbol{\varpi} = (\eta, \xi_1, \xi_2, \xi_3, \xi_4)^T$, and $\boldsymbol{A} = \mathbf{0}$. Given that each observed variable is clearly associated with only one latent variable, the factor loading matrix $\mathbf{\Lambda}$ has the following non-overlapping structure:

$$\begin{bmatrix} 1 & \lambda_{2,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \lambda_{4,2} & \lambda_{5,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{7,3} & \lambda_{8,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{10,4} & \lambda_{11,4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{13,5} & \lambda_{14,5} & \lambda_{15,5} \end{bmatrix}^T,$$

where the 1s and 0s are fixed to obtain an identified model and provide a clear interpretation of the latent variables. The λ_{ij}s are unknown factor loadings reflecting the associations between each latent variable and its associated observed variables. To examine the potential determinants of BMD, we considered the following nonparametric structural equation:

$$\begin{aligned} \eta_i =& g_1(z_{i1}) + g_2(z_{i2}) + f_1(\xi_{i1}) + f_2(\xi_{i2}) + f_3(\xi_{i3}) + f_4(\xi_{i4}) \\ &+ \sum_{u<v,\ u=1,\cdots,4,\ v=1,\cdots,4} h_{uv}(\xi_{iu}, \xi_{iv}) + \delta_i, \end{aligned} \tag{26}$$

where z_{i1} and z_{i2} indicate respondents' weight and age. Equation (26) is a full model with all the main effects and all the interaction effects of the latent variables. We then use the local influence procedure to check the significance of all the terms. The aforementioned continuous measurements were standardized in advance of analysis. In the posterior computation, the prior distributions discussed in Section 3.3 were specified for the unknown parameters. The hyperparameters were taken to those that were assigned in the simulation study. Again, a total of 16 knots was used to construct $B_k^z(\cdot)$s, $B_k(\Phi(\cdot))$s, $U_{us}(\Phi(\cdot))$s and $V_{vt}(\Phi(\cdot))$s in the approximation of $g_1(z_{i1})$, $g_2(z_{i2})$, $f_1(\xi_{i1})$, $f_2(\xi_{i2})$, and $h_{12}(\xi_{i1}, \xi_{i2})$. The MCMC algorithm converged within 15,000 iterations. After convergence, we collected additional 10,000 observations for the posterior inference. Because the sample size is 1460, we set that the 1st to 1460th elements correspond to the data, whereas the 1461st to 1463th elements correspond to the prior distributions in Figure 3, which showed that five observations, including the 303th, 390th, 398th, 405th, and 1279th observations, were identified as influential, and the Bayesian results are not sensitive to the given prior input. Furthermore, we checked the significance of each term in (26) one by one. Figure 4 showed that 'androgen, ξ_2' and 'metabolites, ξ_4' have important effects on BMD. However, Figure 5 did not provide any evidence of interaction effects. Thus, the Bayesian local influence procedure detected five influential observations and two important determinants of BMD in this study. Exploration of the reason for each influential observation may be worthy of further investigation.

6. Discussion

In this paper, we developed a Bayesian local influence procedure in the context of semiparametric SEMs. We introduced a Bayesian perturbation model through perturbing $p(\boldsymbol{y}|\boldsymbol{\varpi}, \boldsymbol{\theta})$, $p(\boldsymbol{\theta})$, and $p(\boldsymbol{\varpi}|\boldsymbol{\theta})$ to characterize perturbations to the data, prior distributions, and the sampling distribution. We proposed to use the first- and second-order local influence measures with the Bayes factor as the objective function to quantify the degree of various perturbations to the interested feature of the analysis. The empirical performance of the proposed method was demonstrated by simulation studies. An application to the BMD study revealed new insights into osteoporosis prevention and control.

We can consider several possible extensions. First, this study adopts the Bayes factor as the objective function in computing the first-order local influence measure. To achieve better local influence efficiency, the use of other objective functions, such as $\phi-$divergence, posterior mean distance, and the χ^2-divergence, may be worthy of further investigation. Second, the proposed model only accommodates continuous observed variables. Given the popularity of various types of data in substantive research, extending the current model framework to incorporate mixed data types is of great importance. Finally, we can consider a generalization of the proposed methodology to a longitudinal setting, which enables us to assess the dynamic effects of time-varying covariates and latent variables on the time-varying outcome of interest.

References

[1] Berry, S.M., Carroll, R.J., Ruppert, D. (2002). Bayesian Smoothing and Regression Splines for Measurement Error Problems. *Journal of the American Statistical Association* **97** 160–169.

[2] Bollen, K.A. (1989). *Structural Equations with Latent Variables*. Wiley, New York.

[3] Chen, J., Liu, P.F., Song, X.Y. (2013). Bayesian diagnostics of transformation structural equation models. *Computational Statistics and Data Analysis* **68** 111–128.

[4] Cook, R.D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society, Series B* **48** 133–169.

[5] Cox, D.R., Reid, N. (1987). Parameter orthogonality and approximate conditional influence (with discussion). *Journal of the Royal Statistical Society, Series B* **49** 1–39.

[6] Fahrmeir L., Raach, A. (2007). A Bayesian Semiparametric Latent Variable Model for Mixed Responses. *Psychometrika* **72** 327–346.

[7] Fan, J., Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, Londan.

[8] Fan, J., Zhang, W. (1999). Statstics Estimation in Varying Coefficient Models. *The Annals of Statistics* **27** 1491–1581.

[9] Geman, S., Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6** 721–741.

[10] Green, P. J., Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall, London.

[11] Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Application. *Biometrika* **57** 97–100.

[12] Ibrahim, J.G., Zhu, H.T., Tang, N.S. (2011). Bayesian local influence for survival models. *Lifetime Data Analysis* **17** 43–70

[13] Jöreskog, K. G., Sörbom, D. (1996). *LISREL 8: User's reference guide*. Scientific Software International.

[14] KASS, R. E., Tierney, L., Kadane, J. B. (1989). Approximate Methods for Assessing Influence and Sensitivity in Bayesian Analysis.*Biometrika* **76** 663–674.

[15] Kenny, D. A., Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological bulletin* **96** 201–210.

[16] Lang, S., Brezger, A. (2004). Bayesian P-splines. *Journal of computational and graphical statistics* **13** 183–212.

[17] Lee, S. Y., Zhu, H. T. (2002). Maximum likelihood estimation of nonlinear structural equation models. *Psychometrika* **67** 189–210.

[18] Lee, S. Y., Tang, N. S. (2004). Local influence analysis of nonlinear structural equation models. *Psychometrika* **69** 573–592.

[19] Lee, S. Y., Lu, B., Song, X. Y. (2006). Assessing local influence for nonlinear structural equation models with ignorable missing data. *Computational statistics and data analysis* **50** 1356–1377.

[20] Lee, S.Y. (2007). *Structural Equation Modeling: A Bayesian Approach*. Wiley, Chichester,UK.

[21] McCulloch, R. E. (1989). Local model influence. *Journal of the American Statistical Association* **84** 473–478.

[22] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics* **21** 1087–1092.

[23] Ruppert, D., Wand, M. P., Carrroll, R.J. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge.

[24] Song, X. Y., Lee, S. Y. (2004). Local influence analysis for mixture of structural equation models. *Journal of classification* **21** 111–137.

[25] Song, X. Y., Lu, Z. H. (2010). Semiparametric latent variable models with Bayesian P-splines. *Journal of Computational and Graphical Statistics* **19** 590–608.

[26] Song, X.Y., Lu, Z.H., Feng, X.N. (2014). Latent variable models with nonparametric interaction effects of latent variables. *Statistics in medicine*, **33** 1723–1737.

[27] Schumacker, R. E., Marcoulides, G. A. (1998). *Interaction and nonlinear effects in structural equation modeling*. Lawrence Erlbaum Associates Publishers, New Jersey.

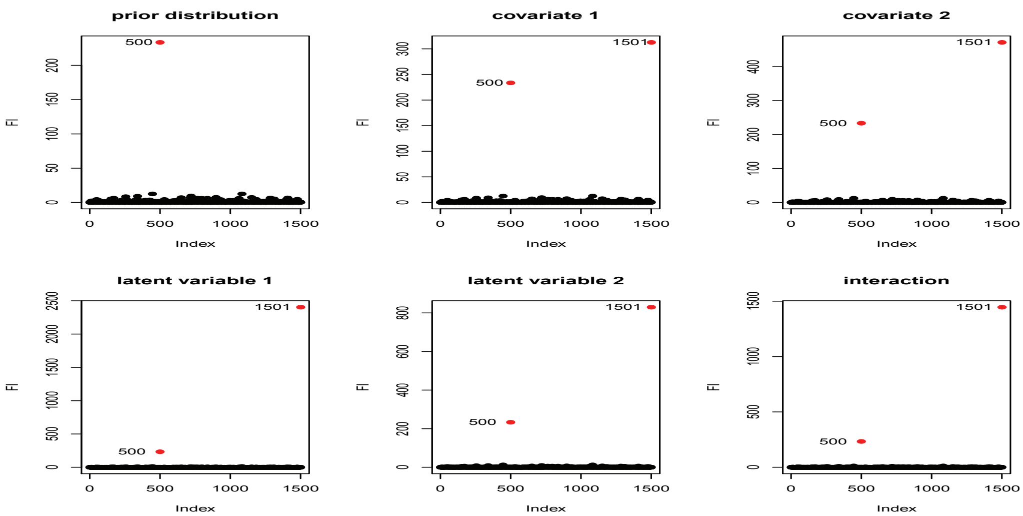

Figure 1: The result of the local influence analysis in Simulation 1.

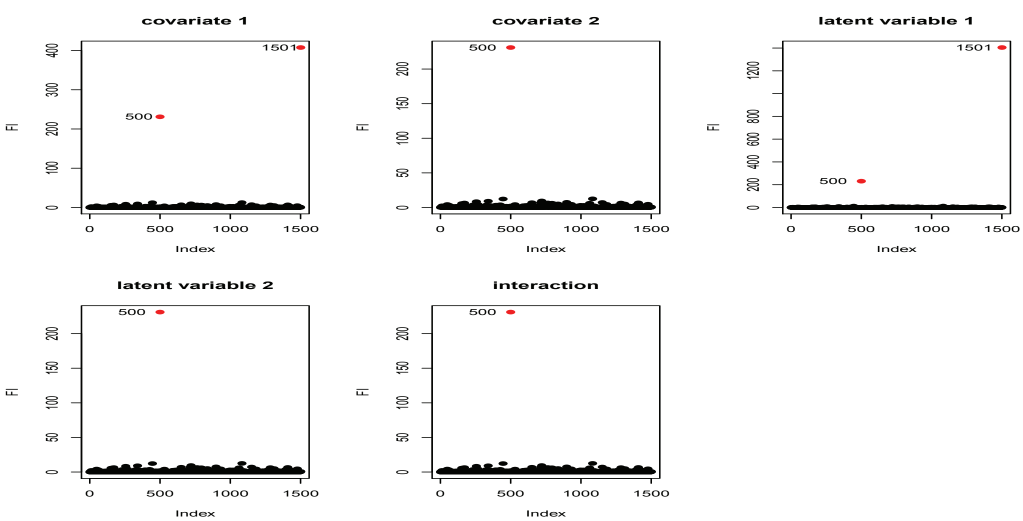

Figure 2: The result of the local influence analysis in Simulation 2.

[28] Tang, N. S., Duan, X. D. (2014). Bayesian influence analysis of generalized partial linear mixed models for longitudinal data. *Journal of Multivariate Analysis* **126** 86–99.

[29] Weiss, R. E., Cook, R. D. (1992). A graphical case statistic for assessing posterior influenceo. *Biometrika* **79** 51–55.

[30] Weiss, R. (1996). An approach to Bayesian sensitivity analysis. *Journal of the Royal Statistical Society. Series B* **58** 739–750.

[31] Zhu, H. T., Lee, S. Y. (2001). Local influence for incomplete-data models. *Journal of the Royal Statistical Society. Series B* **61** 111–126.

[32] Zhu, H. T., Ibrahim, J. G., Lee, S. Y., Zhang, H.P. (2007). Perturbation selection and influence measures in local influence analysis. *The Annals of Statistics* **35** 2565–2588.

[33] Zhu, H. T., Ibrahim, J. G., Tang, N. S. (2011). Bayesian influence analysis: a geometric approach. *Biometrika* **98**, 307–323.

[34] Zhu, H. T., Ibrahim, J. G., Cho, H., Tang, N.S. (2012). Bayesian Case Influence Measures for Statistical Models With Missing Data. *Journal of Computational and Graphical Statistics* **21** 253–271.

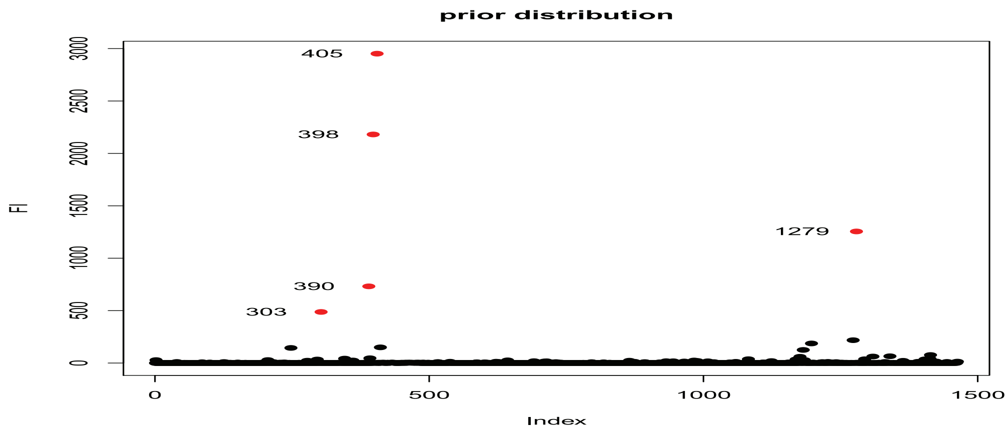

Figure 3: The result of the local influence analysis in the BMD study.

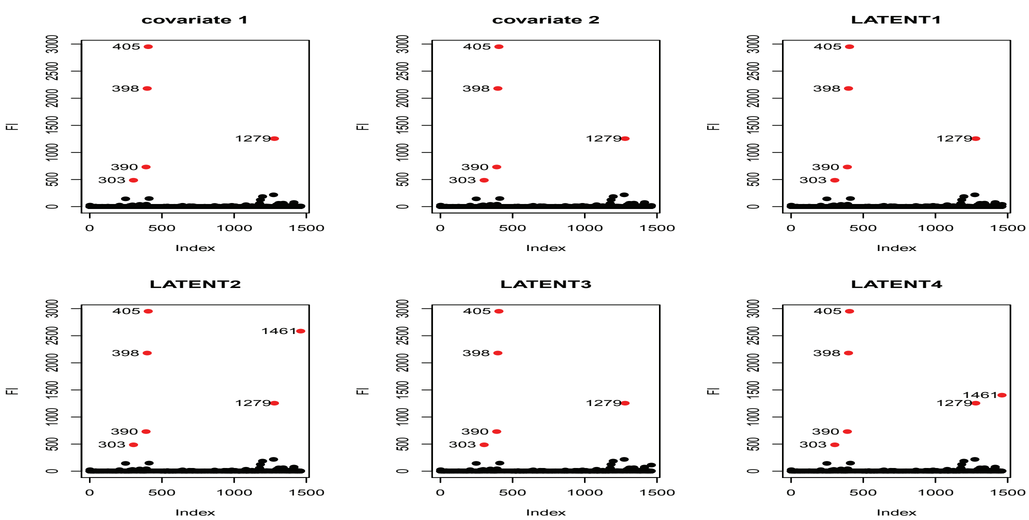

Figure 4: The result of the local influence analysis in the BMD study.

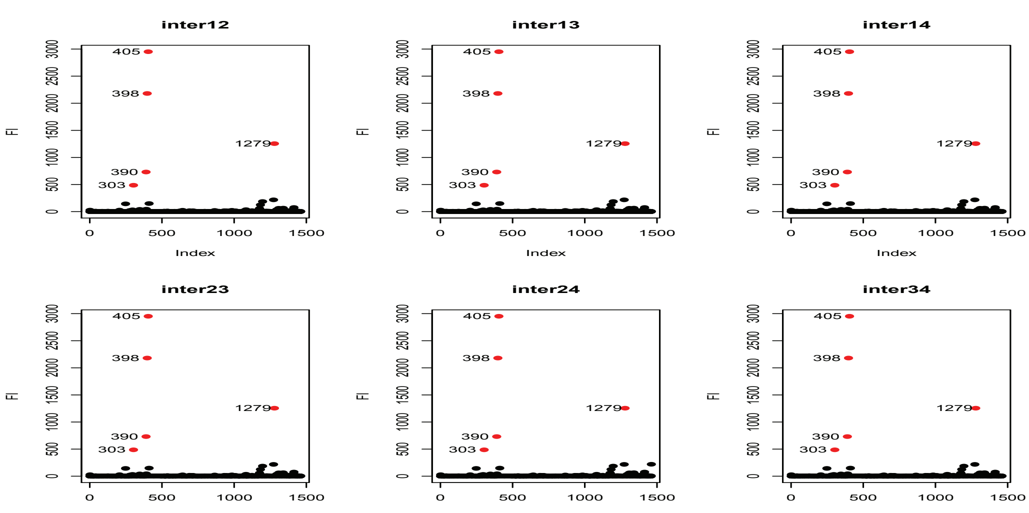

Figure 5: The result of the local influence analysis in the BMD study.