



Bayesian Elastic-Net and Fused Lasso for Semiparametric Structural Equation Models: An Application in Understanding the Relationship Between Alcohol Morbidity and Other Substance Abuse Factors Among American Youth

Zhenyu Wang¹ · Sounak Chakraborty² · Phillip Wood³

Accepted: 1 May 2024

© The Indian Society for Probability and Statistics (ISPS) 2024

Abstract

In contemporary times, high-dimensional datasets have become increasingly prevalent, owing to the expansion and complexity of data collection facilitated by advancements in computer science, biology, and related fields. Analyzing such high-dimensional data poses distinct challenges compared to traditional data analysis, particularly in the realm of variable selection. Structural Equation Modeling (SEM) serves as a pivotal tool for scrutinizing the relationships between observable (manifest) variables and underlying (latent) variables. Traditionally, SEM primarily focuses on elucidating these relationships among latent variables. This paper proposes an extension of semiparametric structural equation modeling, which employs natural cubic splines to approximate nonlinear functional relationships. Moreover, we introduce priors based on Fused Lasso and Elastic Net to address correlations within both covariates and spline expansions. Through comprehensive simulation studies and real-world data analyses, we validate the efficacy of our approach. Our semiparametric structural equation models, enhanced with Bayesian fused Lasso and Bayesian elastic-net priors, consistently outperform conventional Bayesian Lasso models in both simulated and real-world datasets.

Keywords Bayesian variable selection · Elastic net · Fused Lasso · Markov chain Monte Carlo · Natural cubic spline · Structural equation model

1 Introduction

Structural Equation Models (SEMs) represent a sophisticated statistical tool particularly suited for datasets featuring latent variables, which are not directly observable but are inferred from observed variables. Comprising two integral components, SEMs

Extended author information available on the last page of the article

include a measurement equation and a structural equation. The measurement equation explores the connections between unobserved latent variables and observable manifest variables, while the structural equation delves into the interplay among endogenous latent variables, exogenous latent variables, and covariates. Typically, the primary focus of research lies within the structural equation. SEMs find wide application across disciplines such as Psychology, Biology, and others, where latent variables are prevalent. For further illustration, refer to Martens (2005), Lee and Zhu (2000), and Liu et al. (2008).

Traditionally, SEMs assume linear relationships among latent variables in the structural equation. Kenny and Judd (1984) introduced a nonlinear SEM (NSEM) that extended this methodology to include relationships such as interaction and quadratic terms. Lee (2007) generalized NSEM to include a broader set of nonlinear relationships. However, misspecification of the parametric form at the latent level, whether the model is linear or nonlinear, can result in very poor estimation. Recently, some semiparametric approaches have been developed. For example, Bauer (2005), Fahrmeir and Raach (2007), and Guo et al. (2012) used basis expansions to approximate the nonlinear structural relationships using semiparametric SEM (SSEM). To achieve simultaneous estimation and model selection (Guo et al. 2012) applied the Bayesian Lasso method to the SSEM. The Bayesian Lasso performs well in SSEM, however, it ignores correlation of the features which leads to inefficient parameter estimation and model selection.

This is highly concerning when cubic splines are used, because they tend to be highly correlated since each column is a transformed version of the same variables (Keele 2008). This paper addresses this correlation by putting fused Lasso and elastic net prior on the cubic spline coefficient parameters. The fused lasso is shown to be a good method for multiple linear regression when the features have a natural order, specifically when there is side by side correlation (Tibshirani et al. 2005). On the other hand (Zou and Hastie 2005) proved that elastic net can often outperform a regular Lasso in both real world data set and simulation studies with similar sparse representation. In addition, the elastic net encourages a grouping effect, where strongly correlated predictors tend to be in or out of the model together.

The rest of this paper is organized as follows. In Sect. 2, we introduce our Bayesian SSEM framework with its associated basis representation with Fused Lasso prior and the Elastic Net prior. In this section we propose the Bayesian Fused Lasso and Bayesian Elastic Net based methods to achieve simultaneous estimation and model selection. In Sect. 4 we describe our MCMC algorithm to fit our models. To illustrate our proposed methods we introduce two simulation studies in Sect. 5. Subsequently in Sect. 6 we apply our Fused Lasso and Elastic Net based Bayesian SEMs to analyze Monitoring the Future: A Continuing Study of American Youth (12th-Grade Survey) data. Finally in Sect. 7 we discuss some related issues and possible extensions for future work.

2 Bayesian Semiparametric Structural Equation Models

2.1 Semiparametric Structural Equation Models

Semiparametric structural equation models consist of two parts, a measurement equation and a structural equation part. For a random sample of n independent subjects, the measurement equation defines the relationship between the observed $p \times 1$ vector of manifest variables y_i and the unobserved $q \times 1$ vector of latent variables w_i as follows:

$$y_i = A c_i + \Lambda w_i + \epsilon_i, \quad i = 1, 2, \dots, n, \tag{1}$$

where c_i is an $r \times 1$ vector of known functions of the $s \times 1$ vector of fixed covariates x_i , A and Λ are unknown parameter matrices, ϵ_i is a $p \times 1$ vector of measurement errors.

The latent variable w_i is written in two parts, a $q_1 \times 1$ vector of endogenous latent variables η_i and a $q_2 \times 1$ vector of exogenous latent variables ξ_i , i.e. $w_i = (\eta_i^T, \xi_i^T)^T$. We have the following general model which defines the relationship between the exogenous and endogenous latent variables,

$$\eta_i = \Pi \eta_i + F(x_i, \xi_i) + \zeta_i, \quad i = 1, 2, \dots, n, \tag{2}$$

where ζ_i is a vector of residuals and $F(x_i, \xi_i)$ is a vector of unknown functions of the covariates x_i and exogenous latent variables ξ_i .

For the model introduced in Eqs. 1 and 2 model, we require the following assumptions:

- ϵ_i are independently distributed as $N(\mathbf{0}, \Psi_\epsilon)$ with $\Psi_\epsilon = \text{diag}(\psi_{\epsilon 1}, \psi_{\epsilon 2}, \dots, \psi_{\epsilon p})$.
- w_i and ϵ_i are independent, and w_i are independently distributed.
- ζ_i follows $N(\mathbf{0}, \Psi_\zeta)$ with $\Psi_\zeta = \text{diag}(\psi_{\zeta 1}, \psi_{\zeta 2}, \dots, \psi_{\zeta q_1})$.
- ξ_i and ζ_i are independently distributed, and ξ_i follows $N(\mathbf{0}, \Phi)$
- $\Pi_0 = I - \Pi$ is nonsingular and $|\Pi_0|$ is independent of the elements of Π .

Theoretically, $F(x_i, \xi_i)$ can be any linear or nonlinear function of x_i and ξ_i with or without interaction terms like $\xi_{i1} \xi_{i2}$. In this paper, we consider a nonparametric structural equation similar to Guo et al. (2012) and we approximate the nonparametric function $F(x_i, \xi_i)$ using basis expansions. Using basis functions the structural equation 2, in general case, can be represented as

$$\eta_i = \Pi \eta_i + B H(x_i, \xi_i) + \zeta_i, \tag{3}$$

where $H(x_i, \xi_i)$ is an $N_H \times 1$ vector of basis functions, and $B_{q_1 \times N_H}$ is the coefficient parameter matrix associated with $H(x_i, \xi_i)$.

To illustrate the structural equation with basis functions, consider a simple example with $\Pi = 0$, one covariate, one endogenous and two exogenous latent variables. Any function $F(x_i, \xi_i)$ can be decomposed into two parts, functions with only one

variable as f_1, f_2 and f_3 , which could be constant, and functions with interactions as f_{12}, f_{13} and f_{23} , which must be functions of both two parameters, i.e.,

$$\begin{aligned} \eta_i &= F(x_i, \xi_{i1}, \xi_{i2}) + \zeta_i \\ &= f_1(x_i) + f_2(\xi_{i1}) + f_3(\xi_{i2}) + f_{12}(x_i, \xi_{i1}) + f_{13}(x_i, \xi_{i2}) \\ &\quad + f_{23}(\xi_{i1}, \xi_{i2}) + f_{123}(x_i, \xi_{i1}, \xi_{i2}) + \zeta_i, \end{aligned}$$

The above formulation indicates that for modeling f_1, f_2 and f_3 , a linear basis expansion can be used, such as piece-wise polynomials, natural cubic splines, etc. In such cases,

$$f_j(\cdot) = \sum_{m_j=1}^{M_j} \beta_{jm_j} h_{jm_j}(\cdot), \quad j = 1, 2, 3$$

where $\{h_{jm_j}(\cdot), m_j = 1, \dots, M_j\}$ are basis functions. For modeling f_{12}, f_{13} and f_{23} , tensor product basis expansion can be used as follows:

$$f_{kl}(\cdot, \cdot) = \sum_{m_k=1}^{M_k} \sum_{m_l=1}^{M_l} \beta_{m_k m_l}^{(kl)} h_{km_k}(\cdot) h_{lm_l}(\cdot), \quad k, l = 1, 2, 3.$$

2.2 Bayesian Fused Lasso Semiparametric SEM (BFLSEM)

The unknown parameters in the measurement equation 1 are $\Lambda_y = (\mathbf{A}, \Lambda)$ and Ψ_e . On the other hand in structural equation 3, the unknown parameters are $\Lambda_w = (\mathbf{\Pi}, \mathbf{B}), \Psi_\zeta$ and Φ . Some elements of Λ_y must be fixed for identifiability purposes.

For the measurement equation 1, an index matrix $\mathbf{M} = (m_{kj})_{p \times (r+q)}$ is created as follows (Lee and Zhu 2000),

$$m_{kj} = \begin{cases} 1 & \text{if } \lambda_{y_{kj}} \text{ is unknown} \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda_{y_{kj}}$ is the kj -th element of Λ_y . If there is an unknown parameter in k -th row of Λ_y for $k = 1, \dots, p$, this means that $r_{yk} = \sum_{j=1}^{r+q} m_{kj} > 0$. We denote Λ_{yk}^* as the $r_{yk} \times 1$ vector of unknown parameters and specified a conjugate prior for $\{\Lambda_{yk}^*, \psi_{ek}\}$,

$$\Lambda_{yk}^* | \psi_{ek} \sim N_{r_{yk}}(\mu_{0yk}^*, \psi_{ek} \mathbf{H}_{0yk}^*) \tag{4}$$

$$\psi_{ek}^{-1} \sim \text{Gamma}(\alpha_{0ek}, \beta_{0ek}) \tag{5}$$

where $\mu_{0yk}^*, \mathbf{H}_{0yk}^*, \alpha_{0ek}$ and β_{0ek} are hyperparameters.

For the structural equation 3, let Λ_{wh} be the h -th row of Λ_w where $h = 1, \dots, q_1$. As mentioned earlier, we assigned Bayesian fused Lasso priors for each Λ_{wh} and assigned the inverse-Wishart prior for Φ .

$$\begin{aligned} \Lambda_{wh} | \psi_{\zeta h}, \tau_{\Lambda_{wh}}, \mathbf{v}_{\Lambda_{wh}} &\sim N(0, \psi_{\zeta h} \mathbf{D}_{\Lambda_{wh}}), \\ \psi_{\zeta h}^{-1} &\sim \text{Gamma}(\alpha_{0\zeta h}, \beta_{0\zeta h}), \\ \pi(\tau_{\Lambda_{wh}}^2) &\propto \prod_{j=1}^{q_1} \frac{\lambda_{\Pi_h}^2}{2} e^{-\lambda_{\Pi_h}^2 \tau_{\Pi_h j}^2 / 2} \prod_{j=1}^{N_X} \frac{\lambda_{B_{1h}}^2}{2} e^{-\lambda_{B_{1h}}^2 \tau_{B_{1h} j}^2 / 2} \prod_{j=1}^{N_T} \frac{\lambda_{B_{2h}}^2}{2} e^{-\lambda_{B_{2h}}^2 \tau_{B_{2h} j}^2 / 2}, \\ \pi(\mathbf{v}_{\Lambda_{wh}}^2) &\propto \prod_{j=1}^{N_T} \frac{\mu_{B_{2h}}^2}{2} e^{-\mu_{B_{2h}}^2 v_{B_{2h} j}^2 / 2}, \\ \Phi &\sim IW(\mathbf{R}_0, \rho_0), \end{aligned}$$

where N_h is the number of non-constant spline basis functions, and $N_h = N_x + N_T$, where N_x is the number of basis functions related x 's, and N_T is the number of basis functions related to exogenous latent variables. $\mathbf{B}_h = (\mathbf{B}_{1h}^T, \mathbf{B}_{2h}^T)^T$, where \mathbf{B}_{1h} are the coefficients corresponding to the x 's and \mathbf{B}_{2h} are the coefficients corresponding to the exogenous latent variables. $\tau_{\Lambda_{wh}}$ and $\mathbf{v}_{\Lambda_{wh}}$ are mutually independent, and the covariance matrix $\mathbf{D}_{\Lambda_{wh}}^{-1}$ is a diagonal tridiagonal mixed matrix.

$\mathbf{D}_{\Lambda_{wh}}^{-1} = \text{diag}(\mathbf{D}_{q_1 \times q_1}^{11}, \mathbf{D}_{N_x \times N_x}^{22}, \mathbf{D}_{N_T \times N_T}^{33})$, where $\mathbf{D}_{q_1 \times q_1}^{11}$ is a diagonal matrix with

$$\text{main diagonal} = \left\{ \frac{1}{\tau_{\Pi_h j}^2}, j = 1, \dots, q_1 \right\}$$

$\mathbf{D}_{N_x \times N_x}^{22}$ is also a diagonal matrix with

$$\text{main diagonal} = \left\{ \frac{1}{\tau_{B_{1h} j}^2}, j = 1, \dots, N_x \right\}$$

$\mathbf{D}_{N_T \times N_T}^{33}$ is a tridiagonal matrix with

$$\begin{aligned} \text{main diagonal} &= \left\{ \frac{1}{\tau_{B_{2h} j}^2} + \frac{1}{v_{B_{2h} j-1}^2} + \frac{1}{v_{B_{2h} j}^2}, j = 1, \dots, N_T \right\} \\ \text{off diagonals} &= \left\{ -\frac{1}{v_{B_{2h} j}^2}, j = 1, \dots, N_T - 1 \right\} \end{aligned}$$

All the λ 's are tuning parameters with gamma priors.

The extended Bayesian Fused Lasso prior has additional parameters, however, with the priors specified as above, it is straightforward to derive the full conditional distribution (Kyung et al. 2010). As a result we can use MCMC methods to generate samples from the joint posterior distribution of parameters.

The model can be easily extended to the case where X 's has side by side correlation. We only need to change $\mathbf{D}_{N_x \times N_x}^{22}$ to tridiagonal matrix with

$$\begin{aligned} \text{main diagonal} &= \left\{ \frac{1}{\tau_{B_{1h}j}^2} + \frac{1}{v_{B_{1h}j-1}^2} + \frac{1}{v_{B_{1h}j}^2}, j = 1, \dots, N_X \right\} \\ \text{off diagonals} &= \left\{ -\frac{1}{v_{B_{1h}j}^2}, j = 1, \dots, N_X - 1 \right\} \\ \pi(\mathbf{v}_{\Lambda_{wh}}) &\propto \prod_{j=1}^{N_X} \frac{\mu_{B_{1h}}^2}{2} e^{-\mu_{B_{1h}}^2 v_{B_{1h}j}^2 / 2} \prod_{j=1}^{N_T} \frac{\mu_{B_{2h}}^2}{2} e^{-\mu_{B_{2h}}^2 v_{B_{2h}j}^2 / 2}, \end{aligned}$$

It is easy to derive the full conditional distribution and use MCMC methods to generate samples from the joint posterior distribution of parameters for our Bayesian Fused Lasso Semiparametric SEM (BFLSEM).

2.3 Bayesian Elastic Net Semiparametric SEM (BENSEM)

The measurement equation model with prior is exactly the same as in Sect. 2.2, however, for the structural equation part we assign priors based on Elastic Net as follows,

$$\begin{aligned} \Lambda_{wh} | \psi_{\zeta_h}, \tau_{\Lambda_{wh}}, \mathbf{v}_{\Lambda_{wh}} &\sim N(0, \psi_{\zeta_h} \mathbf{D}_{\Lambda_{wh}}), \\ \psi_{\zeta_h}^{-1} &\sim \text{Gamma}(\alpha_{0\zeta_h}, \beta_{0\zeta_h}), \\ \pi(\tau_{\Lambda_{wh}}) &\propto \prod_{j=1}^{q_1} \frac{\lambda_{\Pi_h}^2}{2} e^{-\lambda_{\Pi_h}^2 \tau_{\Pi_h j}^2 / 2} \prod_{k=1}^{N_G} \prod_{j=1}^{N_k} \frac{\lambda_{1B_h k}^2}{2} e^{-\lambda_{1B_h k}^2 \tau_{B_h k j}^2 / 2} \\ \Phi &\sim IW(\mathbf{R}_0, \rho_0), \end{aligned}$$

where \mathbf{X} is reordered. Strongly correlated covariates are grouped together, so we have N_G blocks of \mathbf{X} 's, including one block for independent \mathbf{X} 's if any exists. And $k = 1, \dots, N_G$. For block, k , N_k is the total number of members in the block. $\mathbf{D}_{\Lambda_{wh}}$ is a diagonal matrix with diagonal elements. If \mathbf{X} 's in the corresponding block k are correlated, the diagonal elements are $(\tau_{B_h k j}^{-2} + \lambda_{2B_h k})^{-1}$; if \mathbf{X} 's in the corresponding block k are independent, the diagonal elements are $\tau_{2B_h k j}^2$, in other words $\lambda_{2B_h k} = 0$. And similar to the Bayesian fused lasso, all the λ 's have gamma priors. It is still straightforward to derive the full conditional distribution (Li and Lin 2010), and use MCMC methods to generate samples from the joint posterior distribution of parameters from our Bayesian Elastic Net Semiparametric SEM (BENSEM).

3 Posterior Distributions in Our Bayesian Semiparametric SEM

3.1 Posterior Distribution in the Measurement Equation

Using the conjugate prior for Λ_{yk}^* and ψ_{ek} from 4 and 5, we can easily get the posterior distributions as:

$$\Lambda_{yk}^* | rest \sim N_{r_{yk}}(\mathbf{H}_{yk}(\mathbf{H}_{0yk}^{*-1} \boldsymbol{\mu}_{0yk}^* + \mathbf{G}_{yk} \mathbf{y}_k^*), \boldsymbol{\psi}_{ek}(\mathbf{H}_{0yk}^{*-1} + \mathbf{G}_{yk} \mathbf{G}_{yk}^T)^{-1}) \tag{6}$$

$$\begin{aligned} \boldsymbol{\psi}_{ek}^{-1} | rest &\sim \text{Gamma}(\alpha_{0ek} + n/2, \beta_{0ek} \\ &+ \frac{1}{2}(\mathbf{y}_k^{*T} \mathbf{y}_k^* + \boldsymbol{\mu}_{0yk}^{*T} \mathbf{H}_{0yk}^{*-1} \boldsymbol{\mu}_{0yk}^* - \boldsymbol{\mu}_{yk}^T \mathbf{H}_{yk}^{-1} \mathbf{b} \boldsymbol{\mu}_{yk})) \end{aligned} \tag{7}$$

where $\mathbf{G}_y = (\mathbf{C}^T, \boldsymbol{\Omega}^T)^T$, $\mathbf{C} = \{c_1, \dots, c_n\}$ and $\boldsymbol{\Omega} = \{\omega_1, \dots, \omega_n\}$.

3.2 Posterior Distribution in the Bayesian Structure Equation of Fused Lasso (BFLSEM)

Let $\mathbf{G}_\omega = (\mathbf{g}_{\omega 1}, \dots, \mathbf{g}_{\omega n})$, where $\mathbf{g}_{\omega i} = (\boldsymbol{\eta}_i^T, \mathbf{H}(x_i, \boldsymbol{\xi}_i)^T)^T$. Full conditionals in the structure equation for the h -th row of Λ_ω is:

$$\begin{aligned} \Lambda_{\omega h} | \boldsymbol{\Omega}, \boldsymbol{\psi}_{\zeta h}, \boldsymbol{\tau}_{\Lambda_{\omega h}}, \mathbf{v}_{\Lambda_{\omega h}} &\sim N_{q_1 + N_H}((\mathbf{G}_\omega^T \mathbf{G}_\omega + \mathbf{D}_{\Lambda_{\omega h}}^{-1})^{-1} \mathbf{G}_\omega^T (\boldsymbol{\eta}_h - \beta_{0h} \mathbf{1}_n), \\ \boldsymbol{\psi}_{\zeta h} (\mathbf{G}_\omega^T \mathbf{G}_\omega + \mathbf{D}_{\Lambda_{\omega h}}^{-1})^{-1}), \end{aligned} \tag{8}$$

where $\Lambda_{\omega h} = (\boldsymbol{\Pi}_h^T, \mathbf{B}_h^T)^T$. N_h is the number of non-constant spline basis functions, and $N_h = N_x + N_T$, where N_x is the number of basis functions related x 's, and N_T is the number of basis functions related to exogenous latent variables.

Let $\mathbf{B}_h = (\mathbf{B}_{1h}^T, \mathbf{B}_{2h}^T)^T$, where \mathbf{B}_{1h} are the coefficients corresponding to the x 's and \mathbf{B}_{2h} are the coefficients corresponding to the exogenous latent variables. Note that $\boldsymbol{\tau}_{\Lambda_{\omega h}} = (\tau_{\Pi_{h1}}^2, \dots, \tau_{\Pi_{hq_1}}^2, \tau_{\Pi_{B_{h1}}}^2, \dots, \tau_{\Pi_{B_{hN_H}}}^2)^T$, and the full conditional distribution for $\boldsymbol{\tau}_{\Lambda_{\omega h}}$ are:

$$\begin{aligned} 1/\tau_{\Pi_{hj}}^2 | \boldsymbol{\Pi}_h, \boldsymbol{\psi}_{\zeta h} &\sim IN \left(\sqrt{\frac{\lambda_{\Pi_h}^2 \boldsymbol{\psi}_{\zeta h}}{\Pi_{hj}^2}}, \lambda_{\Pi_h}^2 \right) \\ 1/\tau_{B_{1hj}}^2 | \mathbf{B}_{1h}, \boldsymbol{\psi}_{\zeta h} &\sim IN \left(\sqrt{\frac{\lambda_{B_{1h}}^2 \boldsymbol{\psi}_{\zeta h}}{(B_{1hj})^2}}, \lambda_{B_{1h}}^2 \right) \\ 1/\tau_{B_{2hj}}^2 | \mathbf{B}_{2h}, \boldsymbol{\psi}_{\zeta h} &\sim IN \left(\sqrt{\frac{\lambda_{B_{2h}}^2 \boldsymbol{\psi}_{\zeta h}}{(B_{2hj})^2}}, \lambda_{B_{2h}}^2 \right) \\ 1/v_{B_{2hj}}^2 | \mathbf{B}_{2h}, \boldsymbol{\psi}_{\zeta h} &\sim IN \left(\sqrt{\frac{\lambda_4^2 \boldsymbol{\psi}_{\zeta h}}{(B_{2h(j+1)} - B_{2h(j)})^2}}, \lambda_4^2 \right) \end{aligned}$$

for $j = 1, \dots, NT - 1$.

The full conditional of $\psi_{\zeta h}$ is:

$$\psi_{\zeta h} | \Lambda_{wh}, \mathbf{G}_\omega \sim IG \left(\alpha_{0\zeta h} + \frac{n + q_1 + N_H + 1}{2}, \beta_{1\zeta h} \right)$$

where $\beta_{1\zeta h} = \beta_{0\zeta h} + \frac{1}{2} [(\boldsymbol{\eta}_h - \beta_{0h} \mathbf{1}_n - \mathbf{G}_\omega^T \Lambda_{\omega h})^T (\boldsymbol{\eta}_h - \beta_{0h} \mathbf{1}_n - \mathbf{G}_\omega^T \Lambda_{\omega h}) + \Lambda_{\omega h}^T \mathbf{D}_{\omega h}^{-1} \Lambda_{\omega h}]$

Let the prior of λ 's to be Gamma distribution and the full conditional distributions of them is:

$$\begin{aligned} \lambda_{\Pi_h}^2 | \boldsymbol{\tau}_{\Pi_h} &\sim \text{Gamma} \left(q_1 + r_{0\omega}, \sum_{j=1}^{q_1} \tau_{\Pi_{hj}}^2 / 2 + \delta_{0\Pi} \right) \\ \lambda_{B_{1h}}^2 | \boldsymbol{\tau}_{B_{1h}} &\sim \text{Gamma} \left(N_X + r_{0B_1}, \sum_{j=1}^{N_X} \tau_{B_{1hj}}^2 / 2 + \delta_{0B_1} \right) \\ \lambda_{B_{2h}}^2 | \boldsymbol{\tau}_{B_{2h}} &\sim \text{Gamma} \left(N_T + r_{0B_2}, \sum_{j=1}^{N_T} \tau_{B_{2hj}}^2 / 2 + \delta_{0B_2} \right) \\ \lambda_4^2 | \mathbf{v}_{B_{2h}} &\sim \text{Gamma} \left(N_T + r_{0B_{22}} - 1, \sum_{j=1}^{N_T-1} v_{B_{2hj}}^2 / 2 + \delta_{0B_{22}} \right) \end{aligned}$$

3.3 Posterior Distribution in the Bayesian Structure Equation of Elastic Net (BENSEM)

Full conditionals in the structure equation for the h -th row of Λ_ω is:

$$\begin{aligned} \Lambda_{wh} | \boldsymbol{\Omega}, \boldsymbol{\psi}_{\zeta h}, \boldsymbol{\tau}_{\Lambda_{wh}} &\sim N_{q_1 + N_H} ((\mathbf{G}_\omega^T \mathbf{G}_\omega + \mathbf{D}_{\Lambda_{wh}}^{-1})^{-1} \mathbf{G}_\omega^T (\boldsymbol{\eta}_h - \beta_{0h} \mathbf{1}_n), \\ &\boldsymbol{\psi}_{\zeta h} (\mathbf{G}^T \mathbf{G}_\omega + \mathbf{D}_{\Lambda_{wh}}^{-1})^{-1}), \\ 1 / \tau_{B_{hkj}} | \Lambda_{\omega h}, \boldsymbol{\psi}_{\zeta h} &\sim IG \left(\sqrt{\frac{\lambda_{1\Lambda_{hk}}^2 \boldsymbol{\psi}_{\zeta h}}{\Lambda_{\omega hk}^2}}, \lambda_{1\Lambda_{hk}}^2 \right) \end{aligned}$$

for $j = 1, \dots, N_k$, where $\lambda_{1\Lambda_{hk}} = \lambda_{1\Pi_{hk}}$, when Λ_{whk} are the coefficients of the endogenous latent variables; and $\lambda_{1\Lambda_{hk}} = \lambda_{1B_{hk}}$, when Λ_{whk} are the coefficients of the exogenous latent variables.

The full conditional of $\psi_{\zeta h}$ is:

$$\psi_{\zeta h} | \Lambda_{wh}, \mathbf{G}_\omega \sim IG \left(\alpha_{0\zeta h} + \frac{n + q_1 + N_H + 1}{2}, \beta_{1\zeta h} \right)$$

where $\beta_{1\zeta h} = \beta_{0\zeta h} + \frac{1}{2}[(\boldsymbol{\eta}_h - \beta_{0h}\mathbf{1}_n - \mathbf{G}_\omega^T \Lambda_{\omega h})^T (\boldsymbol{\eta}_h - \beta_{0h}\mathbf{1}_n - \mathbf{G}_\omega^T \Lambda_{\omega h}) + \Lambda_{\omega h}^T \mathbf{D}_{\omega h}^{-1} \Lambda_{\omega h}]$

Let the prior of λ 's to be Gamma distribution and the full conditional distributions of them is:

$$\begin{aligned} \lambda_{\Pi_h}^2 | \boldsymbol{\tau}_{\Pi_h} &\sim \text{Gamma}\left(q_1 + r_{0\Pi}, \sum_{j=1}^{q_1} \tau_{\Pi_h j}^2 / 2 + \delta_{0\Pi}\right) \\ \lambda_{1B_{hk}}^2 | \boldsymbol{\tau}_{\Lambda_{hk}} &\sim \text{Gamma}\left(N_k + r_{1B_{hk}}, \sum_{j=1}^{N_k} \tau_{1B_{hk}j}^2 / 2 + \delta_{1B_{hk}}\right) \\ \lambda_{2B_{hk}}^2 | \mathbf{B} &\sim \text{Gamma}\left(N_k + r_{2B_{hk}}, \frac{1}{2\psi_{\zeta h}} \sum_{j=1}^{N_k} \Lambda_{\omega hkj}^2 + \delta_{2B_{hk}}\right) \end{aligned}$$

where $\Lambda_{\omega hk}$ represent the Λ 's belong to the group k .

4 MCMC Algorithm to Fit Our Bayesian Semiparametric SEM

The parameters from the measurement equation are denoted as $\theta_1^T = \{\Lambda_y, \Psi_\epsilon\}$, while the parameters from the structure equation are denoted as $\theta_2^T = \{\Lambda_\omega, \Psi_\xi, \Phi\}$. Let the parameter of interest be $\theta = (\theta_1^T, \theta_2^T)^T$.

Here are the variables we use in MCMC Algorithm:

- $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$, and \mathbf{y}_i is $p \times 1$ vector of manifest variables.
- $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, and \mathbf{x}_i is $s \times 1$ vector of fixed covariates.
- $\mathbf{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$, and \mathbf{c}_i is $r \times 1$ vector of known function of \mathbf{x}_i .
- $\boldsymbol{\Omega} = \{\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_n\}$, and $\boldsymbol{\omega}_i$ is $q \times 1$ vector of latent variables.

where $i = 1, \dots, n$

$\boldsymbol{\Omega}$ are unobservable latent variables, we can generate it from the full conditional distribution $p(\boldsymbol{\Omega} | \mathbf{Y}, \mathbf{X}, \mathbf{C}, \boldsymbol{\theta})$. Because the latent variables are independent among the subjects, we can write the full conditional distribution as $p(\boldsymbol{\Omega} | \mathbf{Y}, \mathbf{X}, \mathbf{C}, \boldsymbol{\theta}) = \prod_{i=1}^n p(\boldsymbol{\omega}_i | \mathbf{y}_i, \mathbf{x}_i, \mathbf{c}_i, \boldsymbol{\theta})$. Let $g_{yi} = (\mathbf{c}_i^T, \boldsymbol{\omega}_i^T)^T$. The full conditional distribution of $\boldsymbol{\omega}_i$ is:

$$\begin{aligned}
 p(\boldsymbol{\omega}_i | \mathbf{y}_i, \mathbf{x}_i, \mathbf{c}_i, \boldsymbol{\theta}) &\propto p(\mathbf{y}_i | \mathbf{c}_i, \boldsymbol{\omega}_i, \theta_1) p(\boldsymbol{\eta}_i | \mathbf{x}_i, \boldsymbol{\xi}_i, \theta_2) p(\boldsymbol{\xi}_i | \theta_2) \\
 &\propto \exp\left\{-\frac{1}{2}(\mathbf{y}_i - \Lambda_y \mathbf{g}_{y_i})^T \boldsymbol{\Psi}_\epsilon^{-1} (\mathbf{y}_i - \Lambda_y \mathbf{g}_{y_i}) - \frac{1}{2} \boldsymbol{\xi}_i^T \boldsymbol{\Phi}_i^{-1} \boldsymbol{\Phi}_i \right. \\
 &\quad \left. - \frac{1}{2}(\boldsymbol{\eta}_i - \beta_0 - \Lambda_\omega \mathbf{g}_{\omega_i})^T \boldsymbol{\Psi}_\zeta^{-1} (\boldsymbol{\eta}_i - \beta_0 - \Lambda_\omega \mathbf{g}_{\omega_i})\right\}
 \end{aligned} \tag{9}$$

$\boldsymbol{\omega}_i$ can be sampled using Metropolis Hastings (MH) algorithm with a proposal distribution $q(\boldsymbol{\omega}_i^* | \sigma_\omega^2) \sim N(\boldsymbol{\omega}_i^{(j)}, \sigma_\omega^2 \Sigma_\omega)$, where $\boldsymbol{\omega}_i^*$ is the proposed new value and $\boldsymbol{\omega}_i^{(j)}$ is the value from previous step (j th step). From Guo et al. (2012),

$$\Sigma_\omega^{-1} = \Lambda^T \boldsymbol{\Psi}^{-1} \Lambda + \begin{pmatrix} \Pi_0^T \boldsymbol{\Psi}_\zeta^{-1} \Pi_0 & -\Pi_0^T \boldsymbol{\Psi}_\zeta^{-1} \mathbf{B} \Delta_H \\ -\Delta_H^T \mathbf{B}^T \boldsymbol{\Psi}_\zeta^{-1} \Pi_0 & \boldsymbol{\Phi}^{-1} + \Delta_H^T \mathbf{B}^T \boldsymbol{\Psi}_\zeta^{-1} \mathbf{B} \Delta_H \end{pmatrix} \tag{10}$$

where $\Delta_H = \partial \mathbf{H}(\mathbf{x}_i, \boldsymbol{\xi}_i) / \partial \boldsymbol{\xi}_i^T |_{\boldsymbol{\xi}_i = \mathbf{0}}$. The proposed $\boldsymbol{\omega}_i^*$ can be accepted with the probability $\min\{1, \frac{p(\boldsymbol{\omega}_i^* | \mathbf{y}_i, \mathbf{x}_i, \mathbf{c}_i, \boldsymbol{\theta})}{p(\boldsymbol{\omega}_i^{(j)} | \mathbf{y}_i, \mathbf{x}_i, \mathbf{c}_i, \boldsymbol{\theta})}\}$. $\boldsymbol{\Omega}$ can be sampled using Gibbs sampler.

For θ_1 , sample $\Lambda_{yk}^* | rest$ and $\psi_{ek} | rest$ from 6 and 7 respectively.

For θ_2 , the posterior distribution of the parameters are different between Bayesian fused Lasso and Bayesian Elastic Net. We can sample the unknown parameters from the posterior distribution we get on Sects. 3.2 and 3.3.

5 Simulation Study

To illustrate the use of our Fused Lasso and Elastic Net prior based SEMs we have considered the case where the covariates have correlations. Under this framework it is of interest to compare among our BFLSEM (based on Fused Lasso prior), BENSEM (based on Bayesian Elastic Net prior) with Guo et al. (2012) (based on Bayesian standard Lasso prior or BSLSEM).

5.1 Simulation 1

We follow the simulation setup on Guo et al. (2012), setting $n = 500$, $p = 9$, $q_1 = 1$, $q_2 = 2$ and $\mathbf{A} = \text{diag}(0^*, 0^*, 0^*, \mu_4, \dots, \mu_9)$, $\mathbf{c}_i = (1, \dots, 1)^T$,

$$\Lambda^T = \begin{bmatrix} 1.0^* & \lambda_{21} & \lambda_{31} & 0^* & 0^* & 0^* & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 1.0^* & \lambda_{52} & \lambda_{62} & 0^* & 0^* & 0^* \\ 0^* & 0^* & 0^* & 0^* & 0^* & 0^* & 1.0^* & \lambda_{83} & \lambda_{93} \end{bmatrix},$$

where $\mu_4 = \dots = \mu_9 = \lambda_{21} = \dots = \lambda_{93} = \zeta = .36$ and $\{\phi_{11}, \phi_{12}, \phi_{22}\} = \{1, .25, 1\}$. The function, $f(\xi_{i1}, \xi_{i2}) = f_1(\xi_{i1}) + f_2(\xi_{i2}) + f_{12}(\xi_{i1}, \xi_{i2})$, where $f_1(\xi_{i1}) = \sin(\xi_{i1}) - \xi_{i1}$, $f_2(\xi_{i2}) = \exp(\xi_{i2})/2.5 - 3.0$ and $f_{12}(\xi_{i1}, \xi_{i2}) = 0$, has been used to define the underlying relationship between the endogenous and exogenous latent variables. Also, this function is considered unknown and will be approximated using natural cubic splines, i.e.,

$$f_j(\xi_{ij}) \approx \beta_{j2}\xi_{ij} \sum_{m=1}^{K-2} \beta_{j,m+2} (d_m(\xi_{ij}) - d_{K-1}(\xi_{ij}))$$

$$f_{12}(\xi_{i1}, \xi_{i2}) \approx \beta_{12}^{(12)} \xi_{i1}\xi_{i2} + \sum_{m_1=1}^{K-2} \xi_{i2} (d_{m_1}(\xi_{i1}) - d_{K-1}(\xi_{i1}))$$

$$+ \sum_{m_2=1}^{K-2} \xi_{i1} (d_{m_2}(\xi_{i2}) - d_{K-1}(\xi_{i2}))$$

$$+ \sum_{m_1=1}^{K-2} \sum_{m_2=1}^{K-2} (d_{m_1}(\xi_{i1}) - d_{K-1}(\xi_{i1})) (d_{m_2}(\xi_{i2}) - d_{K-1}(\xi_{i2})),$$

with $d_k(\xi_{ij}) = \left[(\xi_{ij} - \kappa_k)_+ - (\xi_{ij} - \kappa_k)_+ \right] / (\kappa_K - \kappa_k)$ where K is the number of knots and $(\kappa_k, k = 1, \dots, K)$ are the location of the knots. The knot locations are selected using a truncated power series basis developed in Hastie et al. (2009). In general cubic splines will be correlated, thus the use of the fused Lasso is appropriate.

We consider $s = 35$ with true parameter values

$$b_l = \begin{cases} 0.5 & \text{if } l \in \{1, 2, 3\} \\ -0.7 & \text{if } l \in \{4, 5\} \\ 0.85 & \text{if } l \in \{6, \dots, 15\} \\ 0.7 & \text{if } l = 32 \\ 0.5 & \text{if } l = 33 \\ 0 & \text{otherwise} \end{cases} .$$

To induce correlation of the covariates $x_1, \dots, x_{31}, x_{34}, x_{35}$ are simulated from a multivariate standard normal distribution where $\text{corr}(x_i, x_j) = .5^{|i-j|}$, $i \neq j \in (6, \dots, 15)$, $\text{corr}(x_i, x_j) = .7$, $i - j = 1, i \in (1, 2, 3)$, $\text{corr}(x_i, x_j) = .9$, $i \neq j \in (4, 5)$ and all other correlations equal to 0. The covariate of $x_{32} \sim 2$ Binomial(1, .5) and $x_{33} \sim N(-0.5, 1)$.

Table 1 summarizes the parameter estimates from the 50 simulations using the BFLSEM (based on fused Lasso prior), BENSEM (based on elastic net prior) and BSLSEM (based on standard Lasso prior). The b_i parameters which relate the covariates to the endogenous latent variable are slightly closer to the true value when BFLSEM is used, however for most of the parameters it is only a slight improvement. The covariates with $\text{corr}(x_i, x_j) = .7$, $i \neq j \in (1, 2, 3)$ have the most marked improvement when BFLSEM is used instead of the BSLSEM or BENSEM. All models are efficient at shrinking the insignificant parameters to 0. As several

Table 1 Simulation 1 result for fused Lasso, elastic net and standard Lasso

Para	True			Fused			Elastic Net			Standard		
	Est	STD	Bias	Est	STD	Bias	Est	STD	Bias	Est	STD	Bias
b_1	0.5000	0.4604	0.0885	-0.0825	0.4417	0.1241	-0.1238	0.4524	0.1767	-0.1000		
b_2	0.5000	0.5996	0.1230	0.1125	0.5768	0.1739	0.1426	0.5685	0.2484	0.1282		
b_3	0.5000	0.4512	0.0909	-0.1026	0.4374	0.1280	-0.1336	0.4469	0.1798	-0.1122		
b_4	-0.7000	-0.6817	0.0808	0.0265	-0.6800	0.0894	0.0290	-0.6793	0.0949	0.0300		
b_5	-0.7000	-0.7120	0.0809	-0.0170	-0.7099	0.0908	-0.0140	-0.7143	0.0955	-0.0202		
b_6	0.8500	0.8430	0.0429	-0.0083	0.8460	0.0485	-0.0047	0.8460	0.0481	-0.0047		
b_7	0.8500	0.8466	0.0537	-0.0040	0.8480	0.0612	-0.0024	0.8481	0.0608	-0.0022		
b_8	0.8500	0.8465	0.0519	-0.0041	0.8405	0.0598	-0.0112	0.8409	0.0586	-0.0108		
b_9	0.8500	0.8408	0.0542	-0.0109	0.8448	0.0634	-0.0061	0.8455	0.0643	-0.0053		
b_{10}	0.8500	0.8579	0.0555	0.0093	0.8543	0.0623	0.0050	0.8523	0.0638	0.0027		
b_{11}	0.8500	0.8470	0.0513	-0.0035	0.8440	0.0574	-0.0071	0.8439	0.0562	-0.0072		
b_{12}	0.8500	0.8454	0.0516	-0.0054	0.8433	0.0537	-0.0079	0.8428	0.0535	-0.0085		
b_{13}	0.8500	0.8430	0.0510	-0.0083	0.8499	0.0549	-0.0001	0.8495	0.0547	-0.0006		
b_{14}	0.8500	0.8499	0.0502	-0.0001	0.8404	0.0568	-0.0114	0.8415	0.0566	-0.0101		
b_{15}	0.8500	0.8513	0.0488	0.0015	0.8582	0.0507	0.0096	0.8581	0.0490	0.0095		
b_{16}	0.0000	0.0043	0.0382	2.0000	0.0014	0.0387	2.0000	0.0026	0.0425	2.0000		
b_{17}	0.0000	0.0049	0.0380	2.0000	0.0051	0.0347	2.0000	0.0063	0.0388	2.0000		
b_{18}	0.0000	0.0050	0.0432	2.0000	0.0036	0.0418	2.0000	0.0036	0.0454	2.0000		
b_{19}	0.0000	-0.0003	0.0382	-2.0000	-0.0003	0.0362	-2.0000	-0.0001	0.0389	-2.0000		
b_{20}	0.0000	0.0048	0.0397	2.0000	0.0004	0.0400	2.0000	0.0006	0.0440	2.0000		
b_{21}	0.0000	-0.0006	0.0393	-2.0000	-0.0034	0.0367	-2.0000	-0.0036	0.0400	-2.0000		
b_{22}	0.0000	-0.0030	0.0377	-2.0000	-0.0017	0.0399	-2.0000	-0.0025	0.0427	-2.0000		
b_{23}	0.0000	-0.0024	0.0429	-2.0000	-0.0021	0.0417	-2.0000	-0.0014	0.0452	-2.0000		
b_{24}	0.0000	-0.0028	0.0413	-2.0000	0.0022	0.0353	2.0000	0.0030	0.0380	2.0000		

Table 1 (continued)

Para	True		Fused		Elastic Net			Standard			
	Est	Bias	Est	STD	Bias	Est	STD	Bias	Est	STD	Bias
b_{25}	0.0000	0.0388	0.0027	0.0388	2.0000	-0.0042	0.0374	-2.0000	-0.0048	0.0399	-2.0000
b_{26}	0.0000	0.0367	-0.0030	0.0367	-2.0000	-0.0031	0.0368	-2.0000	-0.0017	0.0422	-2.0000
b_{27}	0.0000	0.0388	0.0039	0.0388	2.0000	0.0026	0.0337	2.0000	0.0022	0.0376	2.0000
b_{28}	0.0000	0.0385	-0.0015	0.0385	-2.0000	0.0024	0.0368	2.0000	0.0018	0.0403	2.0000
b_{29}	0.0000	0.0408	0.0061	0.0408	2.0000	0.0052	0.0377	2.0000	0.0052	0.0402	2.0000
b_{30}	0.0000	0.0360	0.0008	0.0360	2.0000	-0.0029	0.0367	-2.0000	-0.0033	0.0401	-2.0000
b_{31}	0.0000	0.0361	0.0036	0.0361	2.0000	-0.0009	0.0337	-2.0000	-0.0016	0.0364	-2.0000
b_{32}	0.7000	0.0452	0.6908	0.0452	-0.0132	0.6870	0.0442	-0.0187	0.6948	0.0443	-0.0075
b_{33}	-0.5000	0.0407	-0.4932	0.0407	0.0137	-0.4932	0.0412	0.0137	-0.5001	0.0409	-0.0002
b_{34}	0.0000	0.0368	-0.0064	0.0368	-2.0000	-0.0028	0.0363	-2.0000	-0.0029	0.0406	-2.0000
b_{35}	0.0000	0.0384	0.0055	0.0384	2.0000	0.0047	0.0384	2.0000	0.0052	0.0421	2.0000
β_0	NA	0.0917	-2.1529	0.0917	NA	-2.1641	0.0940	NA	-2.2231	0.1641	NA
β_{12}	NA	0.0772	-0.1947	0.0772	NA	-0.2165	0.0788	NA	-0.2744	0.1721	NA
β_{13}	NA	0.0233	-0.0607	0.0233	NA	0.0117	0.0403	NA	0.3379	0.4499	NA
β_{14}	NA	0.0283	-0.0380	0.0283	NA	-0.0308	0.0267	NA	-0.0199	0.0465	NA
β_{15}	NA	0.0322	-0.0213	0.0322	NA	-0.0793	0.0456	NA	-0.3637	0.3492	NA
β_{22}	NA	0.0404	0.0805	0.0404	NA	0.0781	0.0611	NA	0.1753	0.1352	NA
β_{23}	NA	0.0324	0.1528	0.0324	NA	0.1163	0.0354	NA	-0.0002	0.1777	NA
β_{24}	NA	0.0556	0.1985	0.0556	NA	0.1766	0.0478	NA	0.1118	0.0704	NA
β_{25}	NA	0.0650	0.1866	0.0650	NA	0.2782	0.1173	NA	0.3291	0.2340	NA
$\beta_{22}^{(12)}$	0.0000	0.0632	0.0411	0.0632	2.0000	0.0208	0.0551	2.0000	0.1011	0.1753	2.0000
$\beta_{23}^{(12)}$	0.0000	0.0634	-0.0320	0.0634	-2.0000	-0.0256	0.0553	-2.0000	-0.1101	0.1630	-2.0000
$\beta_{24}^{(12)}$	0.0000	0.0808	-0.0446	0.0808	-2.0000	-0.0198	0.0618	-2.0000	-0.0912	0.1752	-2.0000

Table 1 (continued)

Para	True	Fused			Elastic Net			Standard		
		Est	STD	Bias	Est	STD	Bias	Est	STD	Bias
$\beta_{25}^{(12)}$	0.0000	-0.0369	0.0732	-2.0000	-0.0148	0.0688	-2.0000	-0.0600	0.2716	-2.0000
$\beta_{32}^{(12)}$	0.0000	-0.0207	0.0331	-2.0000	-0.0093	0.0234	-2.0000	-0.0823	0.1394	-2.0000
$\beta_{33}^{(12)}$	0.0000	0.0197	0.0190	2.0000	0.0190	0.0273	2.0000	0.0406	0.2214	2.0000
$\beta_{34}^{(12)}$	0.0000	0.0415	0.0304	2.0000	0.0279	0.0344	2.0000	0.1178	0.1546	2.0000
$\beta_{35}^{(12)}$	0.0000	0.0444	0.0329	2.0000	0.0378	0.0431	2.0000	0.2527	0.2123	2.0000
$\beta_{42}^{(12)}$	0.0000	0.0009	0.0318	2.0000	-0.0110	0.0263	-2.0000	-0.0676	0.1108	-2.0000
$\beta_{43}^{(12)}$	0.0000	0.0205	0.0329	2.0000	0.0155	0.0319	2.0000	-0.0275	0.1377	-2.0000
$\beta_{44}^{(12)}$	0.0000	0.0362	0.0439	2.0000	0.0246	0.0390	2.0000	0.0510	0.1011	2.0000
$\beta_{45}^{(12)}$	0.0000	0.0382	0.0472	2.0000	0.0352	0.0480	2.0000	0.1807	0.2341	2.0000
$\beta_{52}^{(12)}$	0.0000	-0.0029	0.0461	-2.0000	-0.0118	0.0310	-2.0000	-0.0572	0.1702	-2.0000
$\beta_{53}^{(12)}$	0.0000	0.0190	0.0569	2.0000	0.0116	0.0381	2.0000	-0.1153	0.1802	-2.0000
$\beta_{54}^{(12)}$	0.0000	0.0469	0.0848	2.0000	0.0215	0.0451	2.0000	-0.0199	0.1947	-2.0000
$\beta_{55}^{(12)}$	0.0000	0.0760	0.1133	2.0000	0.0323	0.0538	2.0000	0.1275	0.4199	2.0000
RMSE (\hat{f})	0	0.6676	0.3864	NA	0.6493	0.3778	NA	1.5127	1.5713	NA
RMSE ($\hat{\beta}$)	0	0.2323	0.0299	NA	0.2350	0.0291	NA	0.2436	0.0278	NA
RMSE	0	0.6966	0.3738	NA	0.6764	0.3693	NA	1.5311	1.5627	NA

parameter true value are set to be zero we cannot calculate the relative bias, however in Table 1 we include the $RelativeChange = \frac{(\hat{\beta} - \beta)}{(|\hat{\beta}| + |\beta|)/2}$.

There is a fairly significant difference in the spline estimates between the BSLSEM and our proposed two models (BFLSEM and BENSEM). For the spline parameters that are not equal to zero it is not possible to determine which of the models is better in terms of estimation. However, in many of these cases the standard deviations of BSLSEM are significantly higher; while BFLSEM and BENSEM are similar to each other. For the spline parameters that are equal to zero both BFLSEM and BENSEM shrink the estimates nearer to zero than BSLSEM and many have significantly lower standard deviations. Moreover, BENSEM is slightly better than BFLSEM.

To measure the models efficiency at predicting the endogenous latent variable using the covariates and exogenous latent variables, we consider three measures of RMSE.

- $RMSE(\hat{f}) = \sqrt{\sum_{i=1}^n (\hat{f}(\xi_{i1}, \xi_{i2}) - f(\xi_{i1}, \xi_{i2}))^2 / n}$ is a measure of the models ability to approximate the nonlinear relationship between the endogenous and exogenous latent variables,
- $RMSE(\hat{B}) = \sqrt{\sum_{i=1}^n (X\hat{B} - XB)^2 / n}$ is a measure of the models ability to relate the covariates to the endogenous latent variables and
- $RMSE = \sqrt{\sum_{i=1}^n ((X\hat{B} + \hat{f}(\xi_{i1}, \xi_{i2})) - (XB + f(\xi_{i1}, \xi_{i2})))^2 / n}$ is a measure of the models overall ability to predict the endogenous latent variable.

The most significant improvement in the BFLSEM and BENSEM appears to be in the $RMSE(\hat{f})$ which suggests that it is much better at defining the relationship

Fig. 1 Simulation 2: true surface for $\eta = F(x, \xi)$

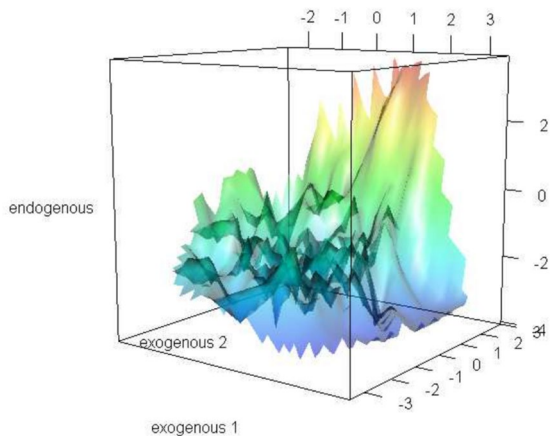


Fig. 2 Simulation 2: true surface for simulated data

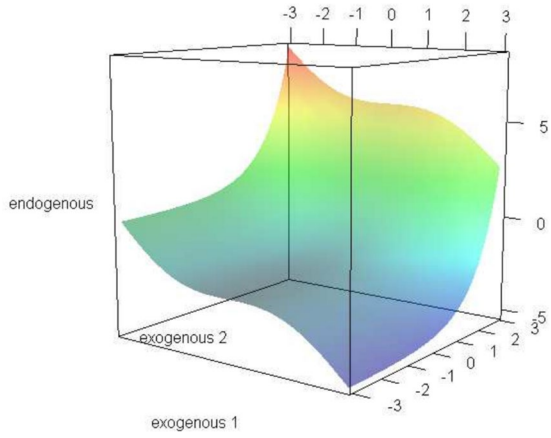


Fig. 3 Simulation 2: estimated surface via BLSSEM

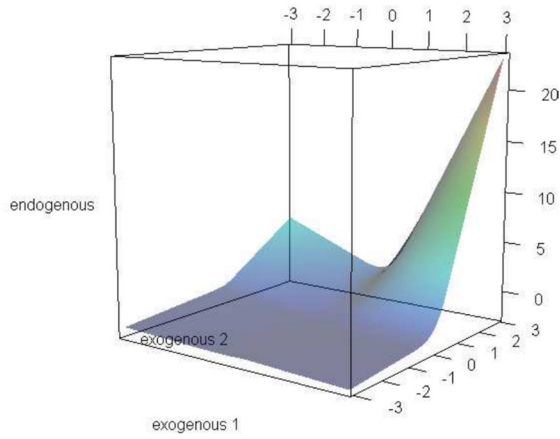


Fig. 4 Simulation 2: estimated surface via BFLSEM

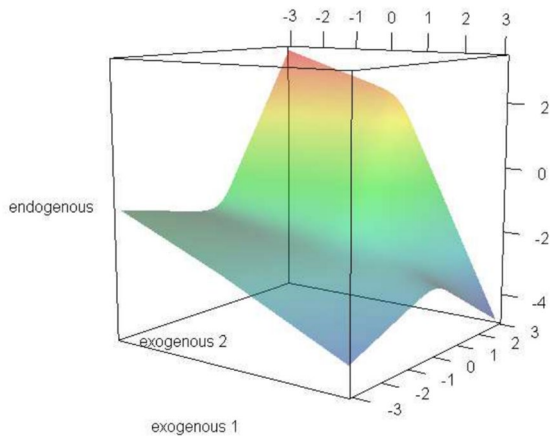
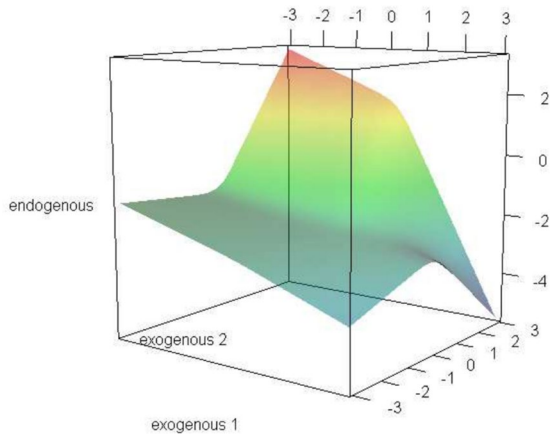


Fig. 5 Simulation 2: estimated surface via BENSEM



between the endogenous and exogenous latent variables. And $RMSE(\hat{f})$ of BENSEM is slightly lower than BFLSEM's. A possible reason there was little impact from on the covariate parameters is that it is very difficult to simulate complex correlation structures. If more covariance structures are examined we believe the difference could be significant.

5.2 Simulation 2

In order to compare the difference in defining the relationship between the endogenous and exogenous latent variables among the three competing model, we randomly choose one of the simulation study and let the coefficient of the covariate to be zeros and plot the surface of $f(\xi_{i1}, \xi_{i2})$. Figure 1 shows the true relationship between exogenous latent variables and endogenous latent variable based on function $\eta = F(x, \xi)$; Fig. 2 shows the relationship between them based on the simulation data, and some of the surface does not have data. Figures 3, 4, and 5 show the estimated surface via original Lasso (BSLSEM), Fused Lasso (BFLSEM), and Elastic Net (BENSEM). In Fig. 3, BSLSEM perform badly when η_1 and η_2 both greater than 0. From Fig. 2, there are no data when both η_1 and η_2 are greater than 2.5. BFLSEM and BENSEM perform similarly. In this simulation, BFLSEM performed marginally better, when both η_1 and η_2 are less than 0.

6 Application in Monitoring the Future: A Continuing Study of American Youth

We apply our BFLSEM and BENSEML to analyze Monitoring the Future: A Continuing Study of American Youth (12th-Grade Survey). There are three exogenous latent variables of interests, cigarette morbidity, marijuana morbidity and

Table 2 Non-spline parameter estimation using posterior means and posterior standard deviations

Parameters	BFLSEM		BENSEM		BSLSEM	
	Mean	SD	Mean	SD	Mean	SD
$\lambda_{2,1}$	0.8477	0.0356	0.8358	0.0348	0.836	0.0417
$\lambda_{3,1}$	0.5202	0.0405	0.5067	0.0397	0.5069	0.0396
$\lambda_{4,1}$	0.4098	0.0271	0.3982	0.0245	0.3983	0.0301
$\lambda_{6,2}$	1.0505	0.0343	0.9976	0.0329	1.0470	0.0399
$\lambda_{8,3}$	1.2825	0.0417	1.2866	0.0442	1.2965	0.0408
$\lambda_{9,3}$	1.2088	0.0314	1.2127	0.0288	1.2219	0.0298
$\lambda_{11,4}$	0.7485	0.0284	0.5973	0.0329	0.6634	0.0342
$\lambda_{12,4}$	0.3364	0.0119	0.4089	0.0238	0.3993	0.0266
$\lambda_{13,4}$	0.1659	0.0162	0.1558	0.0191	0.1583	0.0186
$\lambda_{14,4}$	0.109	0.0274	0.1978	0.0251	0.1796	0.0247
μ_5	3.133	0.3004	3.1667	0.3139	3.1099	0.2764
μ_6	1.9161	0.3270	1.9526	0.3420	1.8925	0.3109
μ_7	5.3888	0.4233	5.4149	0.3969	5.4285	0.4080
μ_8	4.4709	0.4033	4.5048	0.3981	4.5217	0.4699
μ_9	2.8753	0.2186	2.9067	0.2392	2.9222	0.2115
μ_{10}	2.0841	0.1459	2.1565	0.1407	2.1203	0.1531
μ_{11}	1.7787	0.3488	1.8275	0.2960	1.8065	0.3250
μ_{12}	3.9871	0.2796	4.0141	0.2934	3.999	0.3009
μ_{13}	3.1242	0.3645	3.1356	0.3511	3.1303	0.3731
μ_{14}	3.5832	0.2609	3.5943	0.3645	3.5879	0.3829

behavior risk index; one endogenous latent variable, alcohol morbidity. We want to analyze how cigarette morbidity, marijuana morbidity and behavior risk index affect alcohol morbidity. We used the subset from the Monitoring the Future data: 1878 students who had drinking experience. More details about the data and all descriptions can be obtained from <https://monitoringthefuture.org/>.

The endogenous latent variable, alcohol morbidity, is measured by the following items:

- The occasions that students had alcoholic beverages to drink, more than just a few sips in their lifetime.
- The occasions that students had alcoholic beverages to drink, more than just a few sips last year.
- The occasions that students had alcoholic beverages to drink, more than just a few sips last month.
- The number of times that the students had five or more drinks in a row in the last two weeks.

The first exogenous latent variable, cigarette morbidity, is measure by the following items:

- The occasions that students smoked cigarettes in their lifetime.

Table 3 Spline parameter estimation using BFLSEM

Para	Est	Para	Est	Para	Est	Para	Est
b_1	0.1216 ^a	$\beta_1^{(12)}$	- 0.0553	$\beta_1^{(13)}$	- 0.0156	$\beta_1^{(23)}$	- 0.1149 ^a
b_2	0.0243	$\beta_2^{(12)}$	- 0.008	$\beta_2^{(13)}$	0.0073	$\beta_2^{(23)}$	- 0.0056
b_3	0.0835	$\beta_3^{(12)}$	0.0081	$\beta_3^{(13)}$	0.0179	$\beta_3^{(23)}$	0.0246
b_4	0.0845	$\beta_4^{(12)}$	0.0152	$\beta_4^{(13)}$	0.0248	$\beta_4^{(23)}$	0.0327
b_5	- 0.0322	$\beta_5^{(12)}$	0.0258	$\beta_5^{(13)}$	0.033	$\beta_5^{(23)}$	0.0286
β_0	5.8534 ^a	$\beta_6^{(12)}$	0.0063	$\beta_6^{(13)}$	0.007	$\beta_6^{(23)}$	0.0165
β_{12}	0.2729 ^a	$\beta_7^{(12)}$	- 0.0045	$\beta_7^{(13)}$	- 0.0041	$\beta_7^{(23)}$	0.009
β_{13}	0.0303	$\beta_8^{(12)}$	- 0.0044	$\beta_8^{(13)}$	- 0.0029	$\beta_8^{(23)}$	0.0068
β_{14}	- 0.0037	$\beta_9^{(12)}$	0.0166	$\beta_9^{(13)}$	0.0163	$\beta_9^{(23)}$	0.0113
β_{15}	0.0067	$\beta_{10}^{(12)}$	0.0007	$\beta_{10}^{(13)}$	- 0.006	$\beta_{10}^{(23)}$	0.0002
β_{22}	0.1142 ^a	$\beta_{11}^{(12)}$	- 0.0086	$\beta_{11}^{(13)}$	- 0.0139	$\beta_{11}^{(23)}$	- 0.0047
β_{23}	0.0654 ^a	$\beta_{12}^{(12)}$	- 0.0065	$\beta_{12}^{(13)}$	- 0.012	$\beta_{12}^{(23)}$	- 0.0066
β_{24}	0.0382	$\beta_{13}^{(12)}$	0.0176	$\beta_{13}^{(13)}$	0.0048	$\beta_{13}^{(23)}$	- 0.002
β_{25}	0.0441	$\beta_{14}^{(12)}$	- 0.0025	$\beta_{14}^{(13)}$	- 0.0273	$\beta_{14}^{(23)}$	- 0.0168
β_{32}	0.1745 ^a	$\beta_{15}^{(12)}$	- 0.0151	$\beta_{15}^{(13)}$	- 0.0529	$\beta_{15}^{(23)}$	- 0.0252
β_{33}	0.0986 ^a	$\beta_{16}^{(12)}$	- 0.0217	$\beta_{16}^{(13)}$	- 0.0824	$\beta_{16}^{(23)}$	- 0.0314
β_{34}	0.0417						
β_{35}	- 0.0039						

Para. = parameters; Est. = posterior estimates; ^aMarked values indicates 90% of the distribution is greater than 0 or less than 0

- The occasions have students smoked cigarettes during the past 30 days.

The second exogenous latent variable, marijuana morbidity, is measure by the following items:

- The occasions that students smoked marijuana in their lifetime.
- The occasions that students smoked marijuana last year.
- The occasions that students smoked marijuana last month.

The third exogenous latent variable, behavior risk index, is measure by the following items:

- During the last four weeks, the number of whole days of school students have missed because they skipped.
- During the last four weeks, the number of whole days of school students have missed because other reasons.
- During a typical week, the number of evenings students go out for fun and recreation.

Table 4 Spline parameter estimation using BENSEM

Para	Est	Para	Est	Para	Est	Para	Est
b_1	0.0169	$\beta_1^{(12)}$	- 0.0026	$\beta_1^{(13)}$	- 0.1193	$\beta_1^{(23)}$	0.0023
b_2	0.0018	$\beta_2^{(12)}$	- 0.0054	$\beta_2^{(13)}$	- 0.4759	$\beta_2^{(23)}$	0.0031
b_3	0.0035	$\beta_3^{(12)}$	- 0.005	$\beta_3^{(13)}$	0.462	$\beta_3^{(23)}$	0.0039
b_4	0.004	$\beta_4^{(12)}$	- 0.0045	$\beta_4^{(13)}$	2.5223 ^a	$\beta_4^{(23)}$	0.0046
b_5	- 0.0042	$\beta_5^{(12)}$	0.0022	$\beta_5^{(13)}$	- 0.3639	$\beta_5^{(23)}$	- 0.0008
β_0	5.9121	$\beta_5^{(12)}$	0.0012	$\beta_6^{(13)}$	- 0.4635	$\beta_6^{(23)}$	0.0015
β_{12}	- 0.0291	$\beta_7^{(12)}$	0.0013	$\beta_7^{(13)}$	- 0.6698	$\beta_7^{(23)}$	0.0018
β_{13}	- 0.0016	$\beta_8^{(12)}$	0.0011	$\beta_8^{(13)}$	- 1.2979 ^a	$\beta_8^{(23)}$	0.0027
β_{14}	0.0043	$\beta_9^{(12)}$	0.003	$\beta_9^{(13)}$	- 0.1049	$\beta_9^{(23)}$	- 0.0011
β_{15}	0.0156	$\beta_{10}^{(12)}$	0.0021	$\beta_{10}^{(13)}$	- 0.1743	$\beta_{10}^{(23)}$	0.0021
β_{22}	- 0.0002	$\beta_{11}^{(12)}$	0.0022	$\beta_{11}^{(13)}$	- 0.5128	$\beta_{11}^{(23)}$	0.0021
β_{23}	- 0.0026	$\beta_{12}^{(12)}$	0.0022	$\beta_{12}^{(13)}$	- 1.5011 ^a	$\beta_{12}^{(23)}$	0.003
β_{24}	- 0.0025	$\beta_{13}^{(12)}$	0.0044	$\beta_{13}^{(13)}$	1.0263 ^a	$\beta_{13}^{(23)}$	- 0.0011
β_{25}	- 0.0026	$\beta_{14}^{(12)}$	0.0045	$\beta_{14}^{(13)}$	1.3305 ^a	$\beta_{14}^{(23)}$	0.0017
β_{32}	2.4961 ^a	$\beta_{15}^{(12)}$	0.0042	$\beta_{15}^{(13)}$	0.4733	$\beta_{15}^{(23)}$	0.0024
β_{33}	18.3222 ^a	$\beta_{16}^{(12)}$	0.0041	$\beta_{16}^{(13)}$	- 0.53	$\beta_{16}^{(23)}$	0.0031
β_{34}	- 35.6403 ^a						
β_{35}	16.0324 ^a						

Para. = parameters; Est. = posterior estimates; ^aIndicates the estimates that did not converge

- On the average, how often students go out with a date.
- During an average week, how much students usually drive.

As a result, there are totally 14 manifest variables. The Λ in the measurement equation is given by:

$$\Lambda^T = \begin{pmatrix} 1 & \lambda_{21} & \lambda_{31} & \lambda_{41} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \lambda_{62} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{83} & \lambda_{93} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{11,4} & \lambda_{12,4} & \lambda_{13,4} & \lambda_{14,4} \end{pmatrix} \quad (11)$$

Let $\mathbf{A} = \text{diag}(0, \dots, 0, \mu_5, \dots, \mu_{14})$ and $\mathbf{c}_i = (1, \dots, 1)^T$. In addition, we have five covariates, which are gender, geographic area, living with siblings, father education level and mother education level. Let $\mathbf{x}_i = (x_{1i}, \dots, x_{5i})$. To study the interaction between the exogenous latent variables and endogenous latent variable, we proposed following structure equation model:

$$\eta_i = \mathbf{x}_i \mathbf{b}^T + f_1(\xi_{1i}) + f_2(\xi_{2i}) + f_3(\xi_{3i}) + f_{12}(\xi_{1i}, \xi_{2i}) + f_{13}(\xi_{1i}, \xi_{3i}) + f_{23}(\xi_{2i}, \xi_{3i}) \quad (12)$$

Table 5 Spline parameter estimation using BSLSEM

Para	Est	Para	Est	Para	Est	Para	Est
b_1	0.0159	$\beta_1^{(12)}$	- 2.7382 ^a	$\beta_1^{(13)}$	1.0451 ^a	$\beta_1^{(23)}$	0.2214 ^a
b_2	0.0014	$\beta_2^{(12)}$	8.284 ^a	$\beta_2^{(13)}$	22.0392 ^a	$\beta_2^{(23)}$	0.3895 ^a
b_3	0.0008	$\beta_3^{(12)}$	- 3.644 ^a	$\beta_3^{(13)}$	- 3.7627 ^a	$\beta_3^{(23)}$	4.5003 ^a
b_4	0.0074	$\beta_4^{(12)}$	- 1.4991 ^a	$\beta_4^{(13)}$	- 15.6498 ^a	$\beta_4^{(23)}$	- 2.7686 ^a
b_5	- 0.0038	$\beta_5^{(12)}$	9.7528 ^a	$\beta_5^{(13)}$	- 7.5459 ^a	$\beta_5^{(23)}$	1.2014 ^a
β_0	5.9608	$\beta_6^{(12)}$	- 5.0116 ^a	$\beta_6^{(13)}$	- 11.514 ^a	$\beta_6^{(23)}$	- 7.5515 ^a
β_{12}	- 1.3933 ^a	$\beta_7^{(12)}$	- 3.7479 ^a	$\beta_7^{(13)}$	- 10.617 ^a	$\beta_7^{(23)}$	- 18.5468 ^a
β_{13}	- 9.2749 ^a	$\beta_8^{(12)}$	- 1.6634 ^a	$\beta_8^{(13)}$	4.8088 ^a	$\beta_8^{(23)}$	- 6.5395 ^a
β_{14}	8.9574 ^a	$\beta_9^{(12)}$	- 7.53 ^a	$\beta_9^{(13)}$	2.7655 ^a	$\beta_9^{(23)}$	- 4.1878 ^a
β_{15}	0.6507 ^a	$\beta_{10}^{(12)}$	- 0.5455 ^a	$\beta_{10}^{(13)}$	0.9769 ^a	$\beta_{10}^{(23)}$	3.9873 ^a
β_{22}	- 1.7371 ^a	$\beta_{11}^{(12)}$	3.0895 ^a	$\beta_{11}^{(13)}$	0.6814 ^a	$\beta_{11}^{(23)}$	4.1692 ^a
β_{23}	3.1955 ^a	$\beta_{12}^{(12)}$	5.5655 ^a	$\beta_{12}^{(13)}$	18.2107 ^a	$\beta_{12}^{(23)}$	34.5997 ^a
β_{24}	- 1.5914 ^a	$\beta_{13}^{(12)}$	0.4208 ^a	$\beta_{13}^{(13)}$	3.3887 ^a	$\beta_{13}^{(23)}$	2.5759 ^a
β_{25}	- 0.5296 ^a	$\beta_{14}^{(12)}$	- 0.4141 ^a	$\beta_{14}^{(13)}$	- 1.0071 ^a	$\beta_{14}^{(23)}$	8.6044 ^a
β_{32}	4.3805 ^a	$\beta_{15}^{(12)}$	1.5452 ^a	$\beta_{15}^{(13)}$	- 2.2612 ^a	$\beta_{15}^{(23)}$	- 0.4574 ^a
β_{33}	3.9027 ^a	$\beta_{16}^{(12)}$	- 1.8298 ^a	$\beta_{16}^{(13)}$	- 2.0676 ^a	$\beta_{16}^{(23)}$	- 21.5666 ^a
β_{34}	- 5.7997 ^a						
β_{35}	1.9083 ^a						

Para. = parameters; Est. = posterior estimates; ^aIndicates the estimates that did not converge

where $\mathbf{b} = (b_1, \dots, b_5)$. Similar to simulation study, natural cubic splines are used in function $f(\cdot)$ with 5 knots. MCMC chains of 100,000 iterations are generated and the burnin is 30,000. We use both BFLSEM and BENSEM in this case, and compare the result with the BSLSEM. Table 2 shows the estimates from measurement equation. The estimates are very similar among all three methods.

The structure equation results for BFLSEM, BENSEM, and BSLSEM are presented in Tables 3, 4, and 5 respectively. We notice that some of the β 's in BENSEM and BSLSEM did not converge completely. Comparing parameter estimates from Tables 3, 4, and 5 we observe that BFLSEM performed best in this application, with all the β 's properly converged. The result from BFLSEM Table 3 shows that there is an interaction between marijuana morbidity and behavior risk index. The main effect of cigarette morbidity is also significant. The graphs of the two-way interaction of these three exogenous latent variables shows their relation with endogenous latent variable. Figure 6 shows there a weak interaction between cigarette morbidity and marijuana morbidity, but both main effects are highly significant. When cigarette morbidity or/and marijuana morbidity increase, alcohol morbidity increases. Figure 7 shows similar pattern with cigarette morbidity and behavior risk index. Figure 8 shows the interaction between marijuana morbidity and behavior risk index.

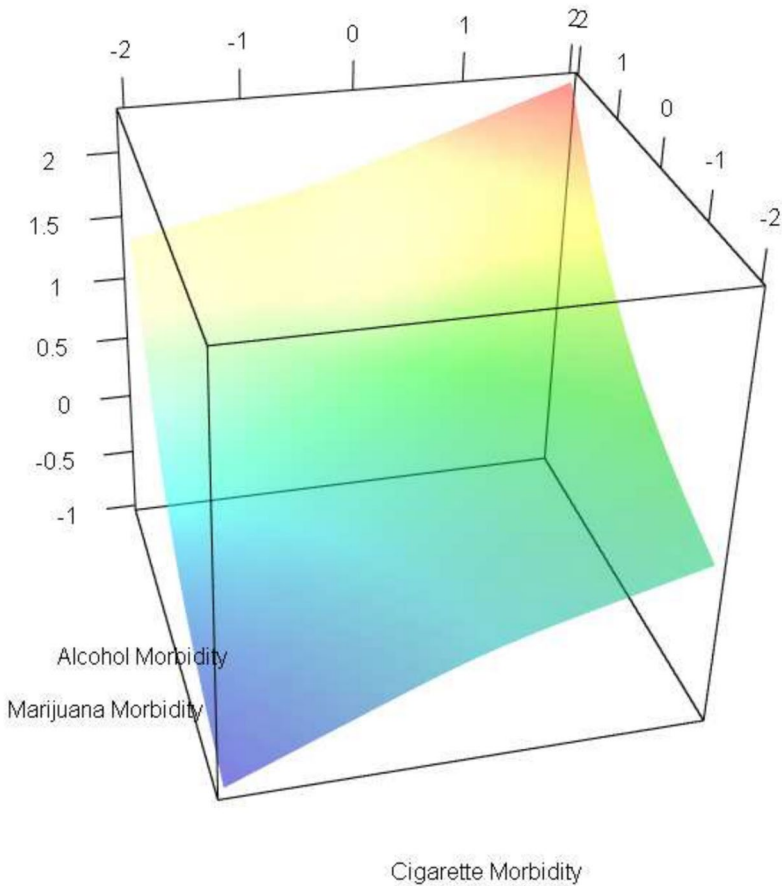


Fig. 6 Estimated surface for cigarette morbidity and marijuana morbidity

When behavior risk index is in the higher level, as marijuana morbidity increases, alcohol morbidity increases faster.

7 Discussion

In this paper we adapted Bayesian fused Lasso prior and Bayesian elastic net prior for using in semiparametric structural equation models. Basis expansions are used to approximate the nonparametric relationships between the endogenous latent variables and the exogenous latent variables and covariates. When cubic splines are used as the basis expansion, it is beneficial to use the fused Lasso or the elastic net based priors (BFLSEM and BENSEM) to estimate the parameters since cubic splines are correlated in general. In the simulation study, both our BFLSEM and BENSEM reduce the standard deviations of the spline parameters

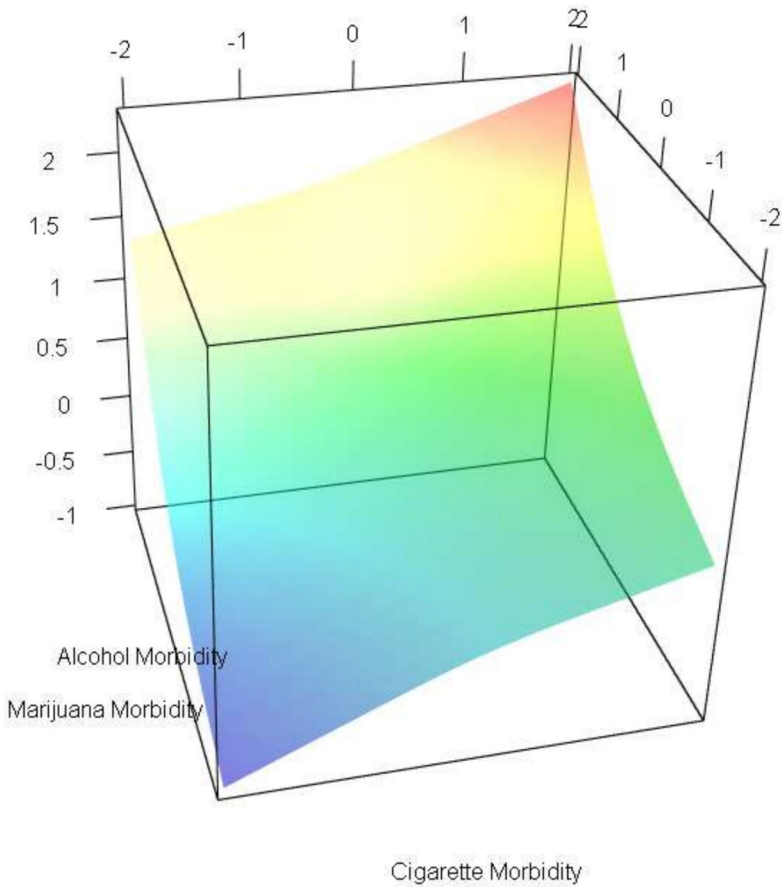


Fig. 7 Estimated surface for cigarette morbidity and marijuana morbidity

and shrink the estimates of the spline parameters closer to zero when the true values of those parameters are equal to zero. More importantly, $RMSE(\hat{f})$ of BFLSEM and BENSEM is about half of $RMSE(\hat{f})$ of the BSLSEM (which is based on the standard Lasso prior, Guo et al. (2012)).

There are clear benefits to use the fused Lasso prior to estimate the coefficients of the covariates, however, it is difficult to generate realistic correlation structures. The usefulness of our methods will depend greatly on the type of the underlying correlation structures. In our simulation study, the fused Lasso prior based SEM (BFLSEM) has a remarkable improvement over the standard Lasso prior based SEM (BSLSEM, Guo et al. (2012)) for the tridiagonal structure with correlation equal to 0.70. However, it is difficult to simulate tridiagonal structures since we often get negative eigenvalues. We believe that if a natural order are present in a real data set our fused Lasso prior based SEM (BFLSEM) would lead to much better results.

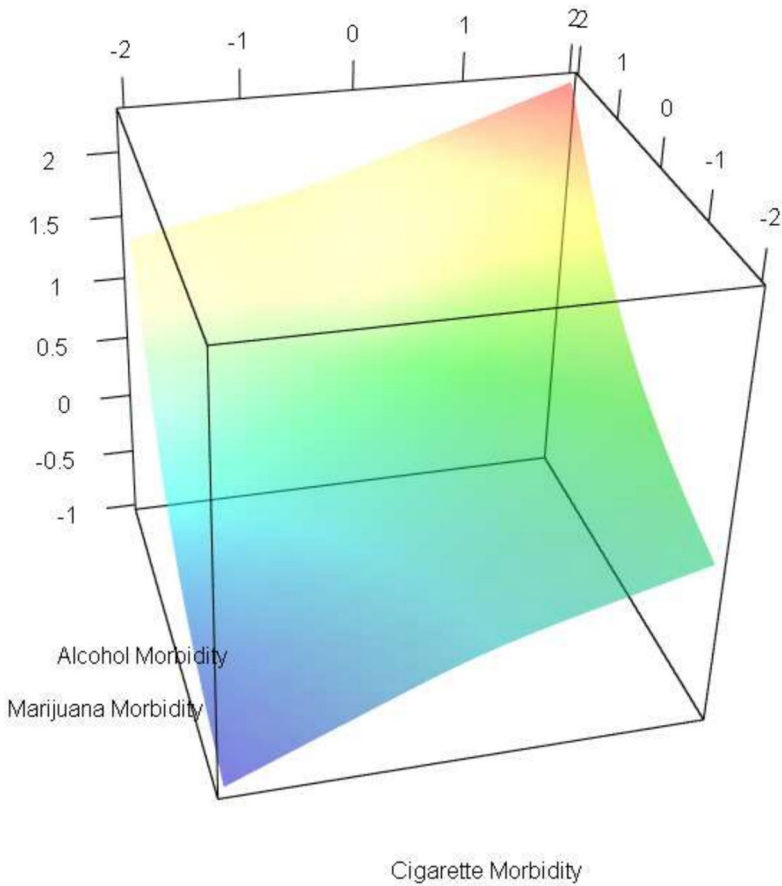


Fig. 8 Estimated surface for cigarette morbidity and marijuana morbidity

In the application all of three methods BFLSEM, BENSEM and BSLSEM have similar estimates for the measurement equations. However in terms of structural equation parameters estimates our BFLSEM based on Bayesian fused Lasso comes out to be the winner and indicates strong interaction between behavior risk index and marijuana morbidity.

However, in real-world data if the Gaussian assumptions on the random components ϵ_i and ζ_i are not met, the model's performance can be compromised, leading to biased parameter estimates, incorrect inference, and poor predictive accuracy. However violations of the Gaussian assumptions can be fixed by adopting a contaminated Gaussian error structure on ϵ_i and ζ_i . Another approach can be using standard transformations like logarithmic, square root, or Box-Cox transformations on the manifest variables and the endogenous latent variables.

In all our two proposed models we include two way interaction of the exogenous latent variables. It is straightforward to extend our model to three way interaction,

when the problem has at least three exogenous latent variables. However, that will increase the number of coefficients need to be estimated significantly, depending on the number of knots. In our study, the options of the psychology survey are mostly ordinal data. In some cases, the options might be dichotomous and that would violate the continuous assumption of the manifest variable. Further research is needed to extend the manifest variable to binary and nominal response. Also it is worthwhile to extend it to other basis expansion methods.

Funding No funding.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no Conflict of interest.

References

- Bauer DJ (2005) A semiparametric approach to modeling nonlinear relations among latent variables. *Struct Equ Model* 12:513–535
- Fahrmeir L, Raach A (2007) A Bayesian semiparametric latent variable model for mixed responses. *Psychometrika* 72:327–346
- Guo R, Zhu H, Chow S, Ibrahim JG (2012) Bayesian lasso for semiparametric structural equation models. *Biometrics* 68:567–577
- Hastie T, Tibshirani R, Friedman J (2009) *Elements of statistical learning: data mining, inference and prediction*, 5th edn. Springer, New York
- Keele L (2008) *Semiparametric regression for the social sciences*. Wiley, Chichester
- Kenny DA, Judd CM (1984) Estimating the nonlinear and interactive effects of latent variables. *Psychol Bull* 96:201–210
- Kyung M, Gill J, Ghosh M, Casella G (2010) Penalized regression, standard errors, and Bayesian lassos. *Bayesian Anal* 5:369–412
- Lee S, Zhu H (2000) Statistical analysis of nonlinear structural equation models with continuous and polytomous data. *Br J Math Stat Psychol* 53:209–232
- Lee SY (2007) *Structural equation modeling: a Bayesian approach*. Wiley, Chichester
- Li Q, Lin N (2010) The Bayesian elastic net. *Bayesian Anal* 5:151–170
- Liu B, Fuente A, Hoeschele I (2008) Gene network inference via structural equation modeling in genetic genomics experiments. *Genetics* 178:1763–1776
- Martens MP (2005) The use of structural equation modeling in counseling psychology research. *Couns Psychol* 33:269–298
- Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K (2005) Sparsity and smoothness via the fused lasso. *J R Stat Soc B* 67:91–108
- Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *J R Stat Soc B* 67:301–320

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Zhenyu Wang¹ · Sounak Chakraborty² · Phillip Wood³

✉ Sounak Chakraborty
chakrabortys@missouri.edu

Zhenyu Wang
zw.stats@gmail.com

Phillip Wood
woodph@missouri.edu

¹ Pfizer Pharmaceuticals, New York, USA

² Department of Statistics, University of Missouri, 209F Middlebush Hall, Columbia, MO 65211, USA

³ Department of Psychology, University of Missouri, 220 McAlester Hall, Columbia, MO 65211, USA