Bayesian Structural Equations Modeling for Ordinal Response Data with Missing Responses and Missing Covariates

SUNGDUK KIM[1], SONALI DAS[2], MING-HUI CHEN[3], AND NICHOLAS WARREN[4]

[1]Division of Epidemiology, Statistics and Prevention Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, NIH, Rockville, Maryland, USA
[2]Logistics and Quantitative Methods, CSIR BE, PO Box 395, Pretoria, South Africa
[3]Department of Statistics, University of Connecticut, Storrs, Connecticut, USA
[4]University of Connecticut Health Center, Farmington Avenue, Connecticut, USA

Structural equations models (SEMs) have been extensively used to model survey data arising in the fields of sociology, psychology, health and economics with increasing applications where self assessment questionnaires are the means to collect the data. We propose the SEM for multilevel ordinal response data from a large multilevel survey conducted by the US Veterans Health Administration (VHA). The proposed model involves a set of latent variables to capture dependence between different responses, a set of facility level random effects to capture facility heterogeneity and dependence between individuals within the same facility, and a set of covariates to account for individual heterogeneity. An effective and practically useful modeling strategy is developed to deal with missing responses and to model missing covariates in the structural equations framework. A Markov chain Monte Carlo sampling algorithm is developed for sampling from the posterior distribution. The deviance information criterion measure is used to compare several variations of the proposed model. The proposed methodology is motivated and illustrated by using the VHA All Employee Survey data.

Keywords DIC; Latent variable; Markov chain Monte Carlo; Missing at random; Ordinal response data; Random effects; VHA all employee survey data.

Mathematics Subject Classification 62F15, 91C99, 65C05.

Address correspondence to Ming-Hui Chen, Department of Statistics, University of Connecticut, 215 Glenbrook Road, U-4120, Storrs, Connecticut, USA;
E-mail: mhchen@stat.uconn.edu

1 Introduction

The use of survey questionnaires with ordinal response variables has been used extensively throughout times. Such response variables usually have a number of categories, often on a Likert-like scale. Ordinal response data arising from self-reported survey questionnaires are also common in assessment studies (Eaton and Bohrnstedt, 1989; Meredith and Mislap, 1992; Hagenaars, 1990). An example of such a data is from the survey conducted by the Veterans Health Administration (VHA) in 2001 via self reported questionnaires to target areas for intervention, with the objective of improving employee work environment. The VHA All Employee Survey (AES) had variables categorized into mostly 5 response levels such as 'Strongly Disagree', 'Disagree', 'Neither Agree Nor Disagree', 'Agree' and 'Strongly Agree'. Such survey data often include covariates that can help explain the variations in the responses.

Generalized linear models with appropriate link are commonly used to model the relationship between ordinal response and covariates, that may be either continuous, or nominal (McCullagh and Nelder, 1989). Studies involving latent variables have been carried out in the past (Chen, 1981; Bollen, 1989; Skrondal and Rabe-Hesketh, 2005; Branden-Roche et al., 1997). Albert (1992) used latent data to estimate the polychoric correlation between two ordinal variables. There has been an extensive theoretical development of linear relationships between manifest variables and latent variables. However, non-linear relationships like quadratic and interaction terms among variables are logical but non-trivial in SEM (Li et al., 1998). More statistically involved Bayesian estimation has been developed for non-linear relationships in SEM (Arminger and Muthén, 1998; Zhu and Lee, 1999). All these methods have assumed that the data at hand are multivariate normal. However, in most assessment surveys, as also in the investigation in this paper, variables are ordinal or binary. Assuming normality for such variables may lead to erroneous conclusions (Olsson, 1979; Lee, Poon and Bentler, 1992). In the Bayesian framework, Chen and Dey (1998, 2000a) and Chib and Greenberg (1998) have explored multivariate probit models for correlated binary variables. Albert and Chib (1993) introduced a Bayesian method to analyze data in the generalized linear model framework in which they introduced latent variables to facilitate the Gibbs sampler. Following from the modeling schemes in the lines proposed by Nandram and Chen (1996) and Chen and Dey (2000b), we accommodate ordinal outcomes by including threshold parameters to link each ordinal outcome to an underlying latent variable. In practice, the ordinal categorical data are considered to be imprecise measurements on some corresponding latent continuous and normally distributed variable.

The rest of the paper is organized as follows. Section 2 provides a detailed description of the the VHA AES 2001 ordinal response data. The development of the structural equations model for such survey data is given in Section 3. The likelihood, the priors, and the posterior based on the proposed model are discussed in Section 4. Model assessment via Bayesian deviance information criterion is considered and appropriate deviance function is derived in Section 5. A comprehensive data analysis of the 2001 VHA AES data is given in Section 6. We conclude the paper with a brief discussion of the various issues encountered in this development in Section 7. Details of the computational algorithm to sample from the posterior distributions are given in the Appendix.

2 Data

We consider the data from the survey conducted in 2001 by the US Veterans Health Administration (VHA) via self reported questionnaires to target areas for intervention, with the objective to improve employee work environment. The target participants were all VHA employees. The VHA all employee survey had variables categorized into mostly 5 response levels such as 'Strongly Disagree', 'Disagree', 'Neither Agree Nor Disagree', 'Agree' and 'Strongly Agree'. The VHA AES data also include covariates that can help explain the variations in the response. The part of the VHA AES 2001 data has 70,458 respondents (*all cases - AC*) belonging to one of 132 facilities. Only 32% were *complete cases (CC)*. We consider 25 response variables from the AES, 24 of which are ordinal, while one is dichotomized to a binary response. The ordinal responses are on a Likert-like scale of $\{1, 2, 3, 4, 5\}$ with '1' corresponding to 'Strongly Disagree', '2' corresponding to 'Disagree', '3' corresponding to 'Neither Agree Nor Disagree', '4' corresponding to 'Agree' and '5' corresponding to 'Strongly Disagree'; or analogously on a 5 point scale from 'Not At All Satisfied' to 'Very Satisfied' format. For the binary response variable '0' corresponds to 'Likely to Leave' and '1' corresponds to 'Likely to Stay'. Of the 25 response variables, 4 are the outcome variables are of interest, viz., Customer Satisfaction, Employee Satisfaction, Quality, and Retention — with Retention being the binary response variable. The other 21 responses are manifest variables for the 3 latent variables in the model, viz., Leadership, Support and Resource.

We also consider 3 covariates viz., 'gender', 'age' and 'years in VHA'. These three covariates are dichotomized as follows: gender (female: 0, male: 1), age ($\leq$ 49 year: 0, $>$ 49 years: 1) and years in VHA ($\leq$ 5 years: 0, $>$ 5 years: 1). About 60% of the respondents are females, 87% are 49 years or younger, while about 73% served in the VHA for over 5 years. Further, the AES 2001 has missing data, both in the 25 response variables as well as in the 3 covariates. The missing percentages for the variables considered in the AES 2001 data are given in Table 1. The largest missing percentage is for the question relating to Planning-evaluation (18.7%), while the least missing is for Retention (0.8%), suggesting perhaps that Retention is an important aspect for the employees to report for possible impact on the outcome of the analysis of the survey. Among the covariates, about 39% did not report their age, while almost every respondent reported gender and years of service with the VHA. For the rest of the variables, the missing percentages are all below 10%.

Table 1
Missing percentage of the variables from the VHA AES 2001

Variable	Missing	Variable	Missing
Reward fair	4.16	Employee resources	2.11
Rewards for service	7.25	Safety	2.60
Zero tolerance	5.22	Work/family balance	4.38
Differences valued	4.02	Teamwork	1.69
Customer needs	3.43	Planning-evaluation	18.69
Customer informed	6.01	Different background	8.86
Pay satisfaction	1.25	Supervisor support	7.27
Employee development	1.86	Customer satisfaction	6.83
Innovation	2.94	Employee satisfaction	1.43
Manager goals	5.28	Quality	1.38
Respect	1.48	Retention	0.83
Conflict resolution	9.88	Age	39.34
Employee involvement	4.03	Gender	3.07
Employee needs	3.97	Years in VA	2.79

3 Model

To fit the AES 2001 data, we consider the SEM with the probit link for the ordinal response variables. Let y denote an ordinal response with L levels. Using the latent variable approach of Albert and Chib (1993), we introduce a continuous latent variable y^* such that

$$y = l \text{ iff } \lambda_{l-1} \leq y^* < \lambda_l,$$

where the cut-points, $-\infty = \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_{L-1} < \lambda_L = \infty$, divide the real line into L intervals. The probit model for the ordinal response y is then obtained by assuming $y^* \sim \mathcal{N}(\mu, \sigma^2)$. Therefore we have the following probability $P(y = l)$ by integrating out y^*,

$$P(y = l) = \Phi\Big(\frac{\lambda_l - \mu}{\sigma}\Big) - \Phi\Big(\frac{\lambda_{l-1} - \mu}{\sigma}\Big) \tag{3.1}$$

for $l = 1, 2, \ldots, L$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function (cdf). From (3.1), we clearly see that μ and σ are confounded with the cut-points λ_l's. To remedy this non-identifiability problem, we need to fix two parameters. Following

Nandram and Chen (1996), we fix $\lambda_1 = 0$ and $\lambda_{L-1} = 1$ so that

$$-\infty = \lambda_0 < \lambda_1 = 0 < \lambda_2 < \cdots < \lambda_{L-1} = 1 < \lambda_L = \infty. \tag{3.2}$$

With the constraint (3.2), all parameters are now identifiable and both μ and σ^2 are free. In addition, in (3.2), all unknown cut-points are bounded.

To capture the association structure of a set of latent variables and a set of responses of interest, we need to identify a set of response variables that can be considered as reasonable manifestations of the latent variables. We recall that latent variables represent the constructs we want to study, and are forced to do so via a set of observable variables we can study. Let $\boldsymbol{y}_{ij} = (y_{ij1}, ..., y_{ijK})'$ denote the $K \times 1$ vector of responses by the j^{th} individual belonging to the i^{th} facility for $i = 1, 2, \ldots, I$, and $j = 1, 2, \ldots, n_i$, where I denotes the total number of the facilities, n_i is the number of individuals within the i^{th} facility, and K is the total number of responses considered. Let L_k be the number of levels of the k^{th} ordinal response. We propose the ordinal response model incorporating the measurement part of the SEM as follows

$$y_{ijk} = l \text{ iff } \lambda_{k,l-1} \leq y^*_{ijk} < \lambda_{kl},$$

where

$$y^*_{ijk} = \mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} + \boldsymbol{\phi}'_k \mathbf{z}_{ij} + \epsilon_{ijk}. \tag{3.3}$$

As discussed above, in (3.3), the cut-points are subject to the constraints:

$$\lambda_{k0} = -\infty < \lambda_{k1} = 0 < \lambda_{k2} < \cdots \lambda_{k,L_k-2} < \lambda_{k,L_k-1} = 1 < \lambda_{k,L_k} = \infty, \tag{3.4}$$

for $k = 1, 2, \ldots, K$. Note that since $y_{ij,25}$ is a binary response, there are no unknown cut-points. Thus, for the AES 2001 data, $L_k = 5$ for $k = 1, 2, \ldots, 24$ and $L_{25} = 2$. Terms τ_i and τ_{ik} introduced in (3.3) are the facility level random effects and the facility-response interactions random effects, respectively, and $\boldsymbol{\beta}_k$ is a $p_k \times 1$ vector of loading coefficients between the k^{th} variable y^*_{ijk} and the r-dimensional latent vector $\boldsymbol{\eta}_{ij}$, where the loading's existence is set via $\boldsymbol{\omega}_k$, a $p_k \times r$ matrix whose elements are either of 1's or 0's. In (3.3), $\boldsymbol{\phi}_k$ is a q-dimensional vector of the regression coefficients corresponding to a q-dimensional vector of covariates, $\boldsymbol{z}_{ij}$, the random error terms $\epsilon_{ijk} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_k)$, $k = 1, 2, \ldots, K$. We assume that $\tau_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_\tau)$, $\tau_{ik} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_{\tau^*})$, and ϵ_{ijk}, τ_i, and τ_{ik} are mutually independent for $i = 1, \ldots, I$, $j = 1, 2, \ldots, n_i$, and $k = 1, \ldots, K$. Let $\boldsymbol{\lambda}_k = (\lambda_{k2}, \ldots, \lambda_{k,L-2})'$, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}'_1, \boldsymbol{\lambda}'_2, \ldots, \boldsymbol{\lambda}'_K)'$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)'$, $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \ \ldots, \ \boldsymbol{\beta}'_K)'$, $\boldsymbol{\phi} = (\boldsymbol{\phi}'_1, \boldsymbol{\phi}'_2, \ldots, \boldsymbol{\phi}'_K)'$, and $\boldsymbol{\sigma}^2 = (\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_K)'$. Then, $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\sigma}^2_\tau, \boldsymbol{\sigma}^2_{\tau^*}, \boldsymbol{\lambda})'$, is the collection of all the parameters of interest in the model.

From the ordinal response SEM in (3.3), the mean and variance of latent variable y^*_{ijk} conditional on $(\mu_k, \boldsymbol{\beta}_k, \boldsymbol{\phi}_k, \mathbf{z}_{ij})$ are

$$\mu_{ijk} = E(y^*_{ijk} | \mu_k, \boldsymbol{\beta}_k, \boldsymbol{\phi}_k, \mathbf{z}_{ij}) = \mu_k + \boldsymbol{\phi}'_k \mathbf{z}_{ij} \tag{3.5}$$

and

$$\sigma_{ijk}^2 = \text{Var}(y_{ijk}^*|\mu_k, \boldsymbol{\beta}_k, \boldsymbol{\phi}_k, \mathbf{z}_{ij}) = \boldsymbol{\beta}_k'\boldsymbol{\omega}_k \text{Var}(\boldsymbol{\eta}_{ij})\boldsymbol{\omega}_k'\boldsymbol{\beta}_k + \sigma_\tau^2 + \sigma_{\tau^*}^2 + \sigma_k^2. \quad (3.6)$$

For a particular individual indexed by i and j, the covariance between answering different questions can be quantified as

$$\text{Cov}(y_{ijk}^*, y_{ijk'}^*) = \boldsymbol{\beta}_k'\boldsymbol{\omega}_k \text{Var}(\boldsymbol{\eta}_{ij})\boldsymbol{\omega}_{k'}'\boldsymbol{\beta}_{k'} + \sigma_\tau^2$$

for $k \neq k'$; for different individuals in the same facility i answering different questions, this covariance equals $\text{Cov}(y_{ijk}^*, y_{ij'k'}^*) = \sigma_\tau^2$ for $j \neq j'$ and $k \neq k'$; while the covariance between two individuals in the same facility i answering the same question is $\text{Cov}(y_{ijk}^*, y_{ij'k}^*) = \sigma_\tau^2 + \sigma_{\tau^*}^2$ for $j \neq j'$. Observe that the covariance structure is in tandem with the natural response pattern of different individuals belonging to the same facility who answer different questions, as well as different individuals within the same facility who answer the same question. The variability in response in the former is solely due to the random effect due to facility effect. However, in the latter, the variability is accounted for by the facility randomness as well as an additional component from the variability due to individual effect as they respond to the same question. While in the former the covariation between responding to different questions by the same individual in a particular facility is accounted for by the facility effect variability, in the latter the variability is accounted for by the structural dependency as well as the facility effect.

The structural part of the ordinal response SEM is given by

$$\boldsymbol{\eta}_{ij} = \Gamma\boldsymbol{\eta}_{ij} + \boldsymbol{\xi}_{ij}, \quad (3.7)$$

where $\boldsymbol{\xi}_{ij} \sim \mathcal{N}(0, diag(\sigma_{\eta_1}^2, \ldots, \sigma_{\eta_r}^2))$, $\boldsymbol{\eta}_{ij}$ is the vector of latent variables corresponding to individual j belonging to facility i, and the Γ matrix is the loading matrix, with the elements parameterized such that the variance of each of the latent variables equals 1. Let $\text{Var}(\boldsymbol{\eta}_{ij}) = V_\eta$ and $D = \text{diag}(\sigma_{\eta_1}^2, \ldots, \sigma_{\eta_r}^2))$. Then, the variance of the latent variables $\boldsymbol{\eta}_{ij}$ is given by

$$V\boldsymbol{\eta} = \text{Var}(\boldsymbol{\eta}_{ij}) = D^{\frac{1}{2}}(I - \Gamma^*)^{-1}\left[(I - \Gamma^*)'\right]^{-1}D^{\frac{1}{2}}. \quad (3.8)$$

We define D to ensure that the diagonal elements of (3.8) are exactly all equal to unity.

For the AES 2001 data, the loading of the 25 manifest variables and 4 outcome variables in the measurement model as well as the structural model of the SEM is illustrated in Figure 1. We note that (3.7) is a conventional representation of the structural part of the SEM, which is used in SAS PROC CALIS and also discussed in detail in Hatcher (2000).

Similarly to Das et al. (2008), we propose the structural model of the SEM for the

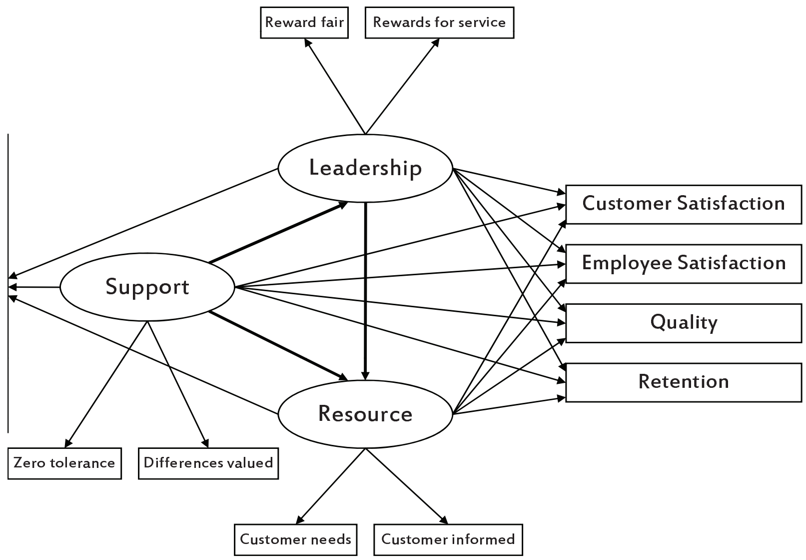

Figure 1. Path diagram for analyzing the VHA AES 2001 data.

ordinal response model as follows

$$\eta_{ij1} = \frac{\gamma_1}{\sqrt{1+\gamma_1^2}}\eta_{ij2} + \xi_{ij1},$$
$$\eta_{ij2} = \xi_{ij2},$$
$$\eta_{ij3} = \frac{\gamma_2}{\sqrt{1+\gamma_3^2+(\gamma_2+\gamma_1\gamma_3)^2}}\ \eta_{ij2} + \frac{\sqrt{(1+\gamma_1^2)}\gamma_3}{\sqrt{1+\gamma_3^2+(\gamma_2+\gamma_1\gamma_3)^2}}\ \eta_{ij_1} + \xi_{ij3},$$

where $\xi_{ij1} \sim \mathcal{N}\left(0, \frac{1}{1+\gamma_1^2}\right)$, $\xi_{ij2} \sim \mathcal{N}(0,1)$, $\xi_{ij3} \sim \mathcal{N}\left(0, \frac{1}{1+\gamma_3^2+(\gamma_2+\gamma_1\gamma_3)^2}\right)$, and ξ_{ij1}, ξ_{ij2} and ξ_{ij3} are independent with each other and they are also independent of ϵ_{ijk}, τ_i, and τ_{ik}. The above representation of $\boldsymbol{\eta}_{ij}$ ensures that the diagonal elements of V_η are exactly all equal to unity.

As discussed in Section 2, there were missing values in both response variables y_{ijk}'s and covariates $\boldsymbol{z}_{ijq}$'s. With the protection of the labor/management confidentiality agreement, it is unlikely that missing responses are due to the sensitive nature of questions as the identification of individual respondents was not recorded. In addition, it does not appear that there were any apparent systematic patterns in the missing values of y_{ijk}'s or $\boldsymbol{z}_{ijq}$'s in the AES 2001 data. Thus, it is reasonable to assume that any missingness in y_{ijk} and covariates $\boldsymbol{z}_{ijq}$ is missing at random (MAR) (Rubin, 1976; Little and Rubin, 2002). As discussed in Ibrahim et al. (1999, 2005), we do not need to model the missing data mechanism for MAR missing responses and covariates. However, we need to model missing covariates. To this end, we assume a series of one-dimensional distributions proposed by Lipsitz and Ibrahim (1996) for the missing covariate variables

$$\begin{aligned}&f(z_{ij1}, z_{ij2}, \dots, z_{ij,q}|\boldsymbol{\alpha})\\ =&f(z_{ijq}|z_{ij1}, \dots, z_{ij,q-1}, \boldsymbol{\alpha}_q)f(z_{ij,q-1}|z_{ij,1}, \dots, z_{ij,q-2}, \boldsymbol{\alpha}_{q-1}) \dots f(z_{ij1}|\boldsymbol{\alpha}_1),\end{aligned} \tag{3.9}$$

where $\boldsymbol{\alpha}_h$ is a vector of parameters for the h^{th} conditional distribution, the $\boldsymbol{\alpha}'_h$s are distinct, and moreover, $\boldsymbol{\alpha} = (\boldsymbol{\alpha}'_1, \dots, \boldsymbol{\alpha}'_q)'$, and extend $\boldsymbol{\theta}$ to include $\boldsymbol{\alpha}$. Of course, for CC data, we do not need to model the covariates.

Letting $\boldsymbol{y}_{ij} = (y_{ij1}, y_{ij2}, \dots, y_{ijK})'$, we thus partition the response vector $\boldsymbol{y}'_{ij}$ into $(\boldsymbol{y}'_{ij,obs},\ \boldsymbol{y}'_{ij,mis})$. Similarly, we partition the vector of covariates $\boldsymbol{z}'_{ij}$ into $(\boldsymbol{z}'_{ij,obs}, \boldsymbol{z}'_{ij,mis})$. Let $\boldsymbol{\eta} = (\boldsymbol{\eta}'_{ij},\ j = 1, 2, \dots, n_i,\ i = 1, 2, \dots, I)'$, $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_I)'$, and $\boldsymbol{\tau}^* = (\tau_{ik},\ i = 1, 2, \dots, I, k = 1, 2, \dots, K)'$. Also let $D_{obs} = (\boldsymbol{y}_{obs}, \boldsymbol{z}_{obs})$ denote the observed data, where $\boldsymbol{y}_{obs} = (\boldsymbol{y}'_{ij,obs},\ j = 1, 2, \dots, n_i,\ i = 1, 2, \dots, I)'$ and $\boldsymbol{z}_{obs} = (\boldsymbol{z}_{ij,obs},\ j = 1, 2, \dots, n_i,\ i = 1, 2, \dots, I)'$. In addition, we introduce a missing indicator for the response in the data

$$\delta_{ijk} = \begin{cases} 1 \text{ if } y_{ijk} \text{ is observed} \\ 0 \text{ if } y_{ijk} \text{ is missing} \end{cases}.$$

Thus, using the result for MAR missing responses given in Chen et al. (2008), the likelihood function given D_{obs} can be written as

$$\begin{aligned}&L(\boldsymbol{\theta}, \boldsymbol{y}^*, \boldsymbol{\tau}, \boldsymbol{\tau}^*|D_{obs})\\ \propto&\int \prod_{i=1}^{I} \Big\{ \prod_{j=1}^{n_i} \Big[\prod_{k=1}^{K} \Big(f(y^*_{ijk}|\mu_k, \tau_i, \tau_{ik}, \boldsymbol{\beta}_k, \boldsymbol{\eta}_{ij}, \boldsymbol{\phi}_k, \sigma^2_k, \boldsymbol{z}_{ij}) 1\{\lambda_{y_{ijk}-1} \le y^*_{ijk} < \lambda_{y_{ijk}}\} \Big)^{\delta_{ijk}} \Big]\\ &\times f(\boldsymbol{\eta}_{ij}|\boldsymbol{\gamma}) f(\boldsymbol{z}_{ij,obs}, \boldsymbol{z}_{ij,mis}|\boldsymbol{\alpha}) \Big\} \Big[\prod_{k=1}^{K} f(\tau_{ik}|\sigma^2_{\tau^*}) \Big] f(\tau_i|\sigma^2_\tau) d\boldsymbol{z}_{mis},\end{aligned} \tag{3.10}$$

where $\boldsymbol{y}^* = (y^*_{ijk},\ j = 1, 2, \dots, n_i, i = 1, 2, \dots, I, k = 1, 2, \dots, K)'$, $1\{\lambda_{y_{ijk}-1} \le y^*_{ijk} <$

$\lambda_{y_{ijk}}\}$ is the indicator function, $\boldsymbol{z}_{mis} = (\boldsymbol{z}_{ij,mis},\ j = 1, 2, \ldots, n_i,\ i = 1, 2, \ldots, I)'$,

$$\begin{aligned} &f(y^*_{ijk}|\mu_k, \tau_i, \tau_{ik}, \boldsymbol{\beta}_k, \boldsymbol{\eta}_{ij}, \boldsymbol{\phi}_k, \sigma^2_k, \boldsymbol{z}_{ij}) \\ =&\frac{1}{\sqrt{2\pi}\sigma_k} \exp\Big\{-\frac{1}{2\sigma^2_k}[y^*_{ijk} - (\mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k\boldsymbol{\omega}_k\boldsymbol{\eta}_{ij} + \boldsymbol{\phi}'_k\boldsymbol{z}_{ij})]^2\Big\}, \end{aligned}$$

$f(\tau_i|\sigma^2_\tau) = \frac{1}{\sqrt{2\pi}\sigma_\tau} \exp\big\{-\frac{\tau^2_i}{2\sigma^2_\tau}\big\}$, $f(\tau_{ik}|\sigma^2_{\tau^*}) = \frac{1}{\sqrt{2\pi}\sigma_{\tau^*}} \exp\big\{-\frac{\tau^2_{ik}}{2\sigma^2_{\tau^*}}\big\}$, and $f(\boldsymbol{\eta}_{ij}|\boldsymbol{\gamma}) = \frac{1}{(2\pi)^{\frac{r}{2}}|V_\eta|^{\frac{1}{2}}}$ $\times \exp\big\{-\frac{1}{2}\boldsymbol{\eta}'_{ij}V^{-1}_\eta\boldsymbol{\eta}_{ij}\big\}$.

4 Prior and Posterior Distributions

We take the joint prior for $\boldsymbol{\theta}$ as

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\mu})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\phi})\pi(\boldsymbol{\gamma})\pi(\boldsymbol{\sigma}^2)\pi(\sigma^2_\tau)\pi(\sigma^2_{\tau^*})\pi(\boldsymbol{\lambda})\pi(\boldsymbol{\alpha}). \tag{4.1}$$

The detailed specification for each prior on the right hand side of (4.1) is given as follows: for location parameters, we take $\mu_k \overset{iid}{\sim} \mathcal{N}(\mu_0, \sigma^2_0),\ k = 1, 2, \ldots, K$, for $\pi(\boldsymbol{\mu})$, $\boldsymbol{\beta}_k \overset{iid}{\sim} \mathcal{N}(\beta_0, \Sigma_{\beta_0}),\ k = 1, 2, \ldots, K$, for $\pi(\boldsymbol{\beta})$, $\boldsymbol{\phi}_k \overset{iid}{\sim} \mathcal{N}(\phi_0, \Sigma_{\phi_0}),\ k = 1, 2, \ldots, K$, for $\pi(\boldsymbol{\phi})$, and $\gamma_l \overset{iid}{\sim} \mathcal{N}(\gamma_0, \sigma^2_{0\gamma}),\ l = 1, 2, \ldots, r$, for $\pi(\boldsymbol{\gamma})$. For the scale parameters, we assume inverse gamma priors as follows: $\sigma^2_k \overset{iid}{\sim} \mathcal{IG}(a_0, b_0)$, $k = 1, 2, \ldots, K$, for $\pi(\boldsymbol{\sigma}^2)$, $\sigma^2_\tau \sim \mathcal{IG}(a_1, b_1)$ for $\pi(\sigma^2_\tau)$, and $\sigma^2_{\tau^*} \sim \mathcal{IG}(a_2, b_2)$ for $\pi(\sigma^2_{\tau^*})$. For $\boldsymbol{\lambda}$, we take $\pi(\boldsymbol{\lambda}) = \prod_{k=1}^K \pi(\boldsymbol{\lambda}_k)$, where $\pi(\boldsymbol{\lambda}_k) \propto 1$, which is a proper uniform over the constrained space defined by (3.4), for $k = 1, 2, \ldots, K$. For $\pi(\boldsymbol{\alpha})$, we assume $\pi(\boldsymbol{\alpha}) = \prod_{l=1}^q \pi(\boldsymbol{\alpha}_1)$, where the prior specification for each $\pi(\boldsymbol{\alpha}_1)$ depends on the form of the one-dimensional conditional distribution for z_{ijl}.

Using (3.10) and (4.1), the joint posterior is thus given by

$$\pi(\boldsymbol{\theta}, \boldsymbol{y}^*, \boldsymbol{\tau}, \boldsymbol{\tau}^*|D_{obs}) \propto L(\boldsymbol{\theta}, \boldsymbol{y}^*, \boldsymbol{\tau}, \boldsymbol{\tau}^*|D_{obs})\pi(\boldsymbol{\theta}). \tag{4.2}$$

Due to the complexity of the ordinal structural equation model and the presence of missing responses and covariates, the analytical evaluation of the posterior distribution in (4.2) does not appear to be possible. However, the posterior distribution is computationally attractive partially due to the use of the probit link, as an efficient Markov chain Monte Carlo (MCMC) sampling algorithm can be developed. The detailed steps of the MCMC sampling algorithm to sample from $\pi(\boldsymbol{\theta}, \boldsymbol{y}^*, \boldsymbol{\tau}, \boldsymbol{\tau}^*|D_{obs})$ is given in the Appendix.

5 Model assessment

For the ordinal response SEM developed in Section 3, we account for both facility effects as well as covariate effects. We refer to this as the full model and denote it by $\mathcal{M}_1$. We are interested in investigating how the exclusion of the two random effects terms introduced and the exclusion of covariates from $\mathcal{M}_1$ contribute to the fit of the data. Further, we will investigate the exclusion from $\mathcal{M}_1$ in turns of the facility effects, then of the covariate effects, and then the effect due to the exclusion of both the effects. The forms of these models are given as follows:

$\mathcal{M}_1$ (Facility and covariates effects): $y^*_{ijk} = \mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} + \boldsymbol{\phi}'_k \boldsymbol{z}_{ij} + \epsilon_{ijk}$;

$\mathcal{M}_2$ (Facility effect, no covariates): $y^*_{ijk} = \mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} + \epsilon_{ijk}$;

$\mathcal{M}_3$ (No facility effect, but covariates effect): $y^*_{ijk} = \mu_k + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} + \boldsymbol{\phi}'_k \boldsymbol{z}_{ij} + \epsilon_{ijk}$; and

$\mathcal{M}_4$ (Neither facility nor covariates effects): $y^*_{ijk} = \mu_k + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} + \epsilon_{ijk}$.

We assess these four ordinal response models via the Bayesian Deviance Information Criteria (DIC) proposed by Spiegelhalter et al. (2002). For model $\mathcal{M}_1$, we treat all facility effects, τ_i and τ_{ik}, and missing covariates, $\boldsymbol{z}_{ij,mis}$, as parameters. Thus, we define $\boldsymbol{\theta}^* = (\boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{z}_{mis})$. The DIC measure proposed by Spiegelhalter et al. (2002) is given by

$$\text{DIC} = D(\overline{\boldsymbol{\theta}^*}) + 2p_D, \tag{5.1}$$

where $D(\boldsymbol{\theta}^*)$ is a deviance function, $\overline{\boldsymbol{\theta}^*} = E[\boldsymbol{\theta}^* | D_{obs}]$ is the posterior mean of $\boldsymbol{\theta}^*$, and p_D is the penalty due to model dimension and is evaluated as $p_D = \overline{D(\boldsymbol{\theta}^*)} - D(\overline{\boldsymbol{\theta}^*})$. For the ordinal response SEM, the deviance function is defined as follows:

$$D(\boldsymbol{\theta}^*) = -2 \log \prod_{i=1}^{I} \prod_{j=1}^{n_i} \prod_{k=1}^{K} \left\{ \Phi\left(\frac{\lambda_{k,y_{ijk}} - \mu^*_{ijk}}{\sigma_k} \right) - \Phi\left(\frac{\lambda_{k,y_{ijk}-1} - \mu^*_{ijk}}{\sigma_k} \right) \right\}^{\delta_{ijk}}, \tag{5.2}$$

where for the full model $\mathcal{M}_1$ with both facility random effects and covariate effects, $\mu^*_{ijk} = \mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ik} + \boldsymbol{\phi}'_k \boldsymbol{z}_{ij}$. Using the extension to DIC as proposed by Huang et al. (2005) in the presence of missing covariates, we compute $\overline{\mu^*}_{ijk} = E(\mu^*_{ijk} | D_{obs})$, $\overline{\lambda}_{kl} = E(\lambda_{kl} | D_{obs})$, $\overline{\sigma}_k = E(\sigma_k | D_{obs})$, $\overline{D(\boldsymbol{\theta}^*)} = E[D(\boldsymbol{\theta}^*) | D_{obs}]$, and

$$D(\overline{\boldsymbol{\theta}^*}) = -2 \log \prod_{i=1}^{I} \prod_{j=1}^{n_i} \prod_{k=1}^{K} \left\{ \Phi\left(\frac{\overline{\lambda}_{k,y_{ijk}} - \overline{\mu^*}_{ijk}}{\overline{\sigma}_k} \right) - \Phi\left(\frac{\overline{\lambda}_{k,y_{ijk}-1} - \overline{\mu^*}_{ijk}}{\overline{\sigma}_k} \right) \right\}^{\delta_{ijk}}.$$

The DIC measures are computed in a similar fashion for the other three competing models. The detail is omitted for brevity.

6 Analysis of the AES 2001 Ordinal Response Data

The VHA AES 2001 was an all employees survey conducted via self reported questionnaire to ascertain organizational climate and locate intervention points, and was a follow up to a previous VHA AES 1997. The main focus was on 4 outcome variable — Customer Satisfaction, Employee satisfaction, Quality and Retention, investigated via 3 work place traits — Leadership, Support and Resource, which correspond to η_{ij1}, η_{ij2}, and η_{ij3}, respectively. These traits being latent constructs, observed responses from the questionnaire were identified that were considered as manifestations of the latent constructs. The dimension of all the response (outcome + manifest) variables is $K = 4 + 21 = 25$, the number of latent traits is $r = 3$, while the number of facilities is $I = 132$. The loading pattern of the 21 manifest variables was motivated by feedback from the field knowledge of VHA, and further confirmed by factor analytic exploratory data analysis.

6.1 Specification of the model and priors

The dimensions of the vectors involved are as follows: for the single loading manifest variables $p_k = 1$, $k = 1, \ldots, 6$. For the 15 manifest variables and 3 outcome variables $p_k = 3$, $k = 7, \ldots, 25$. The corresponding loading matrix are: $\omega_1 = \omega_2 = (1, 0, 0)'$; $\omega_3 = \omega_4 = (0, 1, 0)'$; $\omega_5 = \omega_6 = (0, 0, 1)'$ and $\omega_k = I_3$, $k = 7, \ldots, 25$; in the structural part of the model, the dimension of the latent variable vector $r = 3$. The other prior distributions have been described in Section 4. We specify the hyperparameters as follows. A $\mathcal{N}(0, 1000)$ or $\mathcal{N}(0, 1000I)$ prior is used for all location parameters including $\mu_k, \boldsymbol{\beta}_k, \boldsymbol{\phi}_k, \gamma_1, \gamma_2$, and γ_3. For the scale parameters, the hyperparameters are $a_0 = a_1 = a_2 = 1$ and $b_0 = b_1 = b_2 = 0.001$. Since the outcome variable Retention ($y_{ij,25}$) is dichotomous, the cut-point is known.

For the covariates, since the $\boldsymbol{z}_{ij}$'s are all dichotomized, the distributions for these covariates are specified as

$$
\begin{aligned}
f(z_{ij1}|\alpha_{11}) &= \alpha_{11}^{z_{ij1}}(1-\alpha_{11})^{1-z_{ij1}}, \\
f(z_{ij2}|z_{ij1}, \alpha_{21}, \alpha_{22}) &= \frac{\exp\{z_{ij2}(\alpha_{21} + \alpha_{22} z_{ij1})\}}{1 + \exp(\alpha_{21} + \alpha_{22} z_{ij1})}, \\
f(z_{ij3}|z_{ij1}, z_{ij2}, \alpha_{31}, \alpha_{32}, \alpha_{33}) &= \frac{\exp\{z_{ij3}(\alpha_{31} + \alpha_{32} z_{ij1} + \alpha_{33} z_{ij2})\}}{1 + \exp(\alpha_{31} + \alpha_{32} z_{ij1} + \alpha_{33} z_{ij2})}.
\end{aligned}
$$

The prior distribution for α_{11} is Beta$(0.001, 0.001)$, and the priors distributions of α_{21}, α_{22}, α_{31}, α_{32}, and α_{33} are all $\mathcal{N}(0, 1000)$ independently. Observe that all priors are non-informative.

6.2 Posterior Computation

In all the computations, we used 10,000 MCMC iterations, after a burn-in of 1000 iterations for each model, to compute all the posterior estimates, including posterior means (Estimates), posterior standard deviations (SDs), 95% highest posterior density intervals (HPDs) and DICs. The computer codes were written in FORTRAN 95 using IMSL subroutines with double-precision accuracy. The convergence of the Gibbs sampler for all parameters has passed the recommendations of Cowles and Carlin (1996). The trace plots and auto-correlation plots all show good convergence and excellent mixing of the MCMC sampling algorithm.

6.3 Model Assessment

We calculate DICs as defined in Section 5. The DIC values for the 4 ordinal response SEMs for the all cases data are given in Table 2.

Table 2

DIC values for different ordinal response sEMs (all cases)

Model	$D(\overline{\boldsymbol{\theta}^*})$	p_D	DIC
$\mathcal{M}_1$	3542898.6	162016.2	3866931.0
$\mathcal{M}_2$	3544164.5	161447.2	3867058.8
$\mathcal{M}_3$	3552819.6	160080.3	3872980.2
$\mathcal{M}_4$	3553943.8	159558.0	3873059.9

From Table 2, we see that for the ordinal response SEMs, the full model $\mathcal{M}_1$ has the smallest DIC among the 4 models considered, followed by the model that has only facility effect terms and no covariate terms, then by the model with covariate terms included but no facility terms, and lastly by the model with neither of the effects. We recall that the AES 2001 data is at the individual level, and we do not have any covariate information about the facilities. The only characteristic of the facility that we have is the number of respondents from each of the I facilities, and we use the random effect terms of τ_i and τ_{ik} in the model with the aim to capture variability in the data due to individuals belonging together in the same facility. We also computed the simulation errors of the Monte Carlo estimates of the DICs using the batch mean method discussed in Chen, Shao, and Ibrahim (2000). For example, the simulation errors were 77.7 and 84.4 for $\mathcal{M}_1$ and $\mathcal{M}_4$, respectively, based on 20 batches of size 2000. We recall that about 39% of the respondents had not specified their age. We also subjected the best model $\mathcal{M}_1$ to the *complete-cases* (CC) data, and $D(\overline{\boldsymbol{\theta}^*})$, p_D, and DIC for CC data are 1186402.6, 54825.9, and 1296054.4, respectively.

6.4 Posterior Estimates

The direct effect and total effect are commonly used to describe the association between two variables in the SEM. The *direct effect* is the direct path from one variable to the other, which is simply the coefficient β_{kj}. The *total effect* takes into account the direction and all possible path coefficients between the two variables in the system (Bollen, 1987). If neither is exogenous, then total effect misses to capture the upstream effect due to other connected relationships in the model. To overcome the limitation of the total effect, Das et. al (2008) introduced the *superbeta* measure to capture the overall association between the cause variable (either endogenous or exogenous) and the effect variable (endogenous). If the cause variable is exogenous, then the *superbeta* equals the corresponding total effect. For the ordinal response SEM, we use these measures to describe the association between latent variables y^*_{ijk} and $\boldsymbol{\eta}_{ij}$. Following from Das et. al (2008), the *total effect* between y^*_{ijk} and $\boldsymbol{\eta}_{ij}$ can be calculated as follows

$$\boldsymbol{\beta}^T_{y^*_{ijk},\boldsymbol{\eta}_{ij}} = \begin{pmatrix} \beta_{k1} + \beta_{k3}\frac{a\gamma_3}{b} \\ \beta_{k1}\frac{\gamma_1}{a} + \beta_{k2} + \beta_{k3}\frac{\gamma_2+\gamma_1\gamma_3}{b} \\ \beta_{k3} \end{pmatrix}$$

where $a = \sqrt{1+\gamma_1^2}, b = \sqrt{1+\gamma_3^2+(\gamma_2+\gamma_1\gamma_3)^2}$, and the superbeta between y^*_{ijk} and η_{ijr^*}, $r^* = 1,2,3$, is given by

$$\beta^{sb}_{(y^*_{ijk},\eta_{ij1})} = \text{Cov}(y^*_{ijk},\eta_{ij1})/\text{Var}(\eta_{ij1}) = \beta_{k1} + \beta_{k2}\frac{\gamma_1}{a} + \beta_{k3}\left[\frac{\gamma_1\gamma_2}{ab} + \frac{\gamma_3\ a}{b}\right],$$

$$\beta^{sb}_{(y^*_{ijk},\eta_{ij2})} = \text{Cov}(y^*_{ijk},\eta_{ij2})/\text{Var}(\eta_{ij2}) = \beta_{k1}\frac{\gamma_1}{a} + \beta_{k2} + \beta_{k3}\left[\frac{\gamma_2}{b} + \frac{\gamma_1\gamma_3}{b}\right],$$

and

$$\beta^{sb}_{(y^*_{ijk},\eta_{ij3})} = \text{Cov}(y^*_{ijk},\eta_3)/\text{Var}(\eta_{ij3}) = \beta_{k1}\left[\frac{\gamma_1\gamma_2}{ab} + \frac{\gamma_3 a}{b}\right] + \beta_{k2}\left[\frac{\gamma_2}{b} + \frac{\gamma_1\gamma_3}{b}\right] + \beta_{k3}.$$

In Table 3 (AC) and Table 4 (CC), we present the posterior estimates of *direct effect*, *total effect*, and *superbeta* corresponding to latent outcome variables y^*_{ijk}, $k = 22, 23, 24, 25$, $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and σ^2_τ and $\sigma^2_{\tau^*}$ under the best model $\mathcal{M}_1$, where symbols used are as follow — $\boldsymbol{\beta}_k$: direct effect; $\boldsymbol{\beta}^T_k$: total effect; and $\boldsymbol{\beta}^{sb}_k$: superbeta. For the AC analysis (Table 3), based on the direct effect, we observe that Customer satisfaction is most strongly directly associated with Resources, Quality is directly associated with Leadership and Resources, and Employee satisfaction and Retention are most directly associated with Leadership. Based on either the total effect or the superbeta, all four outcomes are associated with Leadership, Support and Resource. However, based on the superbeta, Leadership is most strongly associated with Employee satisfaction and Retention while Customer satisfaction is most strongly associated with Resources. These results suggest that based on perception of service received, Resource in a facility is most important in improving quality of service provided, while from the service provider perception, Leadership was the most important intervention area.

Table 3

Posterior estimates of Model 1 (all cases)

Variable	Estimate	SD	95% HPD	Variable	Estimate	SD	95% HPD
Customer satisfaction				Quality			
$\beta_{22,1}$	0.019	0.0023	(0.014, 0.023)	$\beta_{24,1}$	0.112	0.0030	(0.107, 0.118)
$\beta^{T}_{22,1}$	0.089	0.0021	(0.084, 0.093)	$\beta^{T}_{24,1}$	0.162	0.0026	(0.157, 0.167)
$\beta^{sb}_{22,1}$	0.134	0.0017	(0.131, 0.137)	$\beta^{sb}_{24,1}$	0.210	0.0022	(0.206, 0.215)
$\beta_{22,2}$	0.037	0.0019	(0.033, 0.040)	$\beta_{24,2}$	0.057	0.0024	(0.053, 0.062)
$\beta^{T}_{22,2}$	0.131	0.0016	(0.128, 0.134)	$\beta^{T}_{24,2}$	0.178	0.0021	(0.174, 0.182)
$\beta^{sb}_{22,2}$	0.131	0.0016	(0.128, 0.134)	$\beta^{sb}_{24,2}$	0.178	0.0021	(0.174, 0.182)
$\beta_{22,3}$	0.182	0.0023	(0.178, 0.187)	$\beta_{24,3}$	0.128	0.0031	(0.122, 0.134)
$\beta^{T}_{22,3}$	0.182	0.0023	(0.178, 0.187)	$\beta^{T}_{24,3}$	0.128	0.0031	(0.122, 0.134)
$\beta^{sb}_{22,3}$	0.209	0.0017	(0.206, 0.212)	$\beta^{sb}_{24,3}$	0.213	0.0025	(0.208, 0.218)
Employee satisfaction				Retention			
$\beta_{23,1}$	0.195	0.0027	(0.190, 0.200)	$\beta_{25,1}$	0.298	0.0083	(0.282, 0.314)
$\beta^{T}_{23,1}$	0.232	0.0023	(0.228, 0.237)	$\beta^{T}_{25,1}$	0.324	0.0072	(0.310, 0.339)
$\beta^{sb}_{23,1}$	0.258	0.0018	(0.254, 0.262)	$\beta^{sb}_{25,1}$	0.375	0.0059	(0.364, 0.386)
$\beta_{23,2}$	0.023	0.0021	(0.018, 0.027)	$\beta_{25,2}$	0.076	0.0074	(0.062, 0.090)
$\beta^{T}_{23,2}$	0.174	0.0018	(0.170, 0.177)	$\beta^{T}_{25,2}$	0.271	0.0058	(0.259, 0.282)
$\beta^{sb}_{23,2}$	0.174	0.0018	(0.170, 0.177)	$\beta^{sb}_{25,2}$	0.271	0.0058	(0.259, 0.282)
$\beta_{23,3}$	0.096	0.0028	(0.091, 0.102)	$\beta_{25,3}$	0.068	0.0083	(0.051, 0.084)
$\beta^{T}_{23,3}$	0.096	0.0028	(0.091, 0.102)	$\beta^{T}_{25,3}$	0.068	0.0083	(0.051, 0.084)
$\beta^{sb}_{23,3}$	0.209	0.0022	(0.204, 0.213)	$\beta^{sb}_{25,3}$	0.259	0.0064	(0.246, 0.271)
α_{11}	-2.224	0.0197	(-2.261, -2.184)	σ^2_τ	0.002	0.0002	(0.001, 0.002)
α_{21}	0.506	0.0288	(0.447, 0.561)	$\sigma^2_{\tau^*}$	0.001	0.0001	(0.001, 0.001)
α_{22}	0.385	0.0019	(0.382, 0.389)	γ_1	0.656	0.0067	(0.642, 0.669)
α_{31}	0.940	0.0114	(0.918, 0.963)	γ_2	0.303	0.0076	(0.289, 0.318)
α_{32}	1.431	0.0463	(1.339, 1.521)	γ_3	0.390	0.0072	(0.376, 0.404)
α_{33}	-0.139	0.0182	(-0.175, 0.104)				

For the CC analysis (Table 4), we observe that for almost all the β coefficients, the CC estimates for the full model $\mathcal{M}_1$ are larger than the AC estimates. The $\boldsymbol{\beta}$ coefficients between outcome variables and Resource are smaller in the AC analysis compared to the AC analysis. But for the superbeta measure, all the CC estimates are stronger than in the AC case. On the other hand, for the γ coefficients, the γ_1 and γ_3 coefficients between latent variables 'Leadership' and 'Support' and between 'Resource' and 'Leadership' significantly increase for the complete cases from the all cases.

Table 4

Posterior estimates of Model 1 (complete cases)

Variable	Estimate	SD	95% HPD	Variable	Estimate	SD	95% HPD
Customer satisfaction				Quality			
$\beta_{22,1}$	0.037	0.0039	(0.029, 0.044)	$\beta_{24,1}$	0.148	0.0054	(0.138, 0.159)
$\beta^T_{22,1}$	0.112	0.0036	(0.105, 0.119)	$\beta^T_{24,1}$	0.199	0.0047	(0.190, 0.208)
$\beta^{sb}_{22,1}$	0.155	0.0028	(0.150, 0.161)	$\beta^{sb}_{24,1}$	0.243	0.0038	(0.235, 0.250)
$\beta_{22,2}$	0.033	0.0031	(0.027, 0.039)	$\beta_{24,2}$	0.046	0.0042	(0.038, 0.055)
$\beta^T_{22,2}$	0.140	0.0026	(0.135, 0.145)	$\beta^T_{24,2}$	0.193	0.0036	(0.186, 0.199)
$\beta^{sb}_{22,2}$	0.140	0.0026	(0.135, 0.145)	$\beta^{sb}_{24,2}$	0.193	0.0036	(0.186, 0.199)
$\beta_{22,3}$	0.174	0.0039	(0.166, 0.182)	$\beta_{24,3}$	0.119	0.0054	(0.108, 0.129)
$\beta^T_{22,3}$	0.174	0.0039	(0.166, 0.182)	$\beta^T_{24,3}$	0.119	0.0054	(0.108, 0.129)
$\beta^{sb}_{22,3}$	0.211	0.0028	(0.205, 0.216)	$\beta^{sb}_{24,3}$	0.226	0.0042	(0.217, 0.233)
Employee satisfaction				Retention			
$\beta_{23,1}$	0.222	0.0046	(0.213, 0.231)	$\beta_{25,1}$	0.342	0.0153	(0.312, 0.373)
$\beta^T_{23,1}$	0.259	0.0039	(0.251, 0.266)	$\beta^T_{25,1}$	0.357	0.0133	(0.331, 0.383)
$\beta^{sb}_{23,1}$	0.280	0.0030	(0.275, 0.286)	$\beta^{sb}_{25,1}$	0.411	0.0105	(0.390, 0.431)
$\beta_{23,2}$	0.016	0.0037	(0.009, 0.024)	$\beta_{25,2}$	0.082	0.0127	(0.056, 0.106)
$\beta^T_{23,2}$	0.191	0.0031	(0.184, 0.196)	$\beta^T_{25,2}$	0.303	0.0100	(0.283, 0.322)
$\beta^{sb}_{23,2}$	0.191	0.0031	(0.184, 0.196)	$\beta^{sb}_{25,2}$	0.303	0.0100	(0.283, 0.322)
$\beta_{23,3}$	0.085	0.0047	(0.076, 0.094)	$\beta_{25,3}$	0.035	0.0145	(0.007, 0.064)
$\beta^T_{23,3}$	0.085	0.0047	(0.076, 0.094)	$\beta^T_{25,3}$	0.035	0.0145	(0.007, 0.064)
$\beta^{sb}_{23,3}$	0.219	0.0036	(0.212, 0.226)	$\beta^{sb}_{25,3}$	0.270	0.0110	(0.248, 0.291)
γ_1	0.744	0.0121	(0.721, 0.768)	σ^2_τ	0.002	0.0003	(0.002, 0.003)
γ_2	0.292	0.0135	(0.265, 0.318)	$\sigma^2_{\tau^*}$	0.001	0.0001	(0.001, 0.001)
γ_3	0.430	0.0128	(0.406, 0.456)				

6.5 Normal-Ordinal SEMs

As discussed in Bollen (1989), the Pearson correlation coefficients between categorical measures are generally less than the correlation of the corresponding continuous variables. This phenomenon may more likely occur with few categories (e.g., < 5). However, as the number of categories increases and the marginal distributions become similar, the difference in correlations lessens. For the AES 2001 data, the first 21 manifest variables are ordinal with 5 levels and the last four are the outcome variables with Retention being the binary response variable. It is of practical interest to investigate whether the structural part of the SEM and the loading coefficients of the last four measurement part of the ordinal response SEM are sensitive to the specification of the marginal distributions of y_{ijk} for $k = 1, 2, \ldots, 21$ for the first 21 manifest variables. Specifically, we assume that $y^*_{ijk} = y_{ijk}$ in (3.3) for $k = 1, 2, \ldots, 21$. That is, we assume the normal distribution models directly for the ordinal responses from the first 21 manifest variables. As the last four outcome variables are of primary interest, we still assume the ordinal (binary) models for these four outcome variables. Under these assumptions, the resulting SEM is called the normal-ordinal SEM.

First, we compare the four variations of the normal-ordinal SEM discussed in Section 5 via DIC. Similar to (5.2), the deviance function for the normal-ordinal SEM is given by

$$D(\boldsymbol{\theta}^*) = -2\sum_{i=1}^{I}\sum_{j=1}^{n_i}\Bigg\{\sum_{k=1}^{21}\delta_{ijk}\Big[-\frac{1}{2}\log(2\pi\sigma_k^2) - \frac{1}{2\sigma_k^2}\big(y_{ijk}-\mu^*_{ijk}\big)^2\Big] + \sum_{k=22}^{K}\delta_{ijk}\log\Big[\Phi\Big(\frac{\lambda_{k,y_{ijk}}-\mu^*_{ijk}}{\sigma_k}\Big) - \Phi\Big(\frac{\lambda_{k,y_{ijk}-1}-\mu^*_{ijk}}{\sigma_k}\Big)\Big]\Bigg\},$$

where for the full model $\mathcal{M}_1$ with both facility random effects and covariate effects, $\mu^*_{ijk} = \mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k\boldsymbol{\omega}_k\boldsymbol{\eta}_{ik} + \boldsymbol{\phi}'_k\boldsymbol{z}_{ij}$. For the all cases AES 2001 data, under the normal-ordinal SEMs, the DIC values are 4214927.9, 4214942.2, 4220289.6, and 4220339.3 for $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and $\mathcal{M}_4$, respectively. Thus, the best normal-ordinal SEM is $\mathcal{M}_1$, the second best DIC normal-ordinal SEM is $\mathcal{M}_2$, and the worst DIC normal-ordinal SEM is $\mathcal{M}_4$. Although the DIC values under the normal-ordinal SEMs are not comparable to those under the ordinal SEMs (Table 2), the orders of these four DIC values under these two types of SEMs are the same.

Next, we compute the posterior estimates of *direct effect*, *total effect*, and *superbeta* corresponding to latent outcome variables y^*_{ijk}, $k = 22, 23, 24, 25$, $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and σ^2_τ and $\sigma^2_{\tau^*}$ under the best normal-ordinal model $\mathcal{M}_1$ and the results are shown in Table 5 (AC) and Table 6 (CC). Comparing the results in Table 5 to those in Table 3, we observe that these two sets of the posterior estimates of *direct effect*, *total effect*, *superbeta*, and $\boldsymbol{\gamma}$ are very similar, and the estimates of $\boldsymbol{\alpha}$ are almost identical. However, the posterior estimates of σ^2_τ and $\sigma^2_{\tau^*}$ under the normal-ordinal SEM (Table 5) are much larger than those under the ordinal SEM (Table 3). The similar patterns are observed when we compare the results in Table 6 to those in Table 4 for the CC data.

Table 5
Posterior estimates of Model 1 (all cases) under the normal-ordinal SEM

Variable	Estimate	SD	95% HPD	Variable	Estimate	SD	95% HPD
Customer satisfaction				Quality			
$\beta_{22,1}$	0.017	0.0023	(0.012, 0.021)	$\beta_{24,1}$	0.113	0.0030	(0.107, 0.119)
$\beta^T_{22,1}$	0.088	0.0021	(0.084, 0.092)	$\beta^T_{24,1}$	0.163	0.0026	(0.158, 0.168)
$\beta^{sb}_{22,1}$	0.128	0.0017	(0.125, 0.132)	$\beta^{sb}_{24,1}$	0.203	0.0022	(0.199, 0.207)
$\beta_{22,2}$	0.031	0.0018	(0.028, 0.035)	$\beta_{24,2}$	0.046	0.0022	(0.042, 0.050)
$\beta^T_{22,2}$	0.122	0.0015	(0.119, 0.125)	$\beta^T_{24,2}$	0.163	0.0019	(0.159, 0.167)
$\beta^{sb}_{22,2}$	0.122	0.0015	(0.119, 0.125)	$\beta^{sb}_{24,2}$	0.163	0.0019	(0.159, 0.167)
$\beta_{22,3}$	0.185	0.0023	(0.181, 0.189)	$\beta_{24,3}$	0.127	0.0031	(0.122, 0.134)
$\beta^T_{22,3}$	0.185	0.0023	(0.181, 0.189)	$\beta^T_{24,3}$	0.127	0.0031	(0.122, 0.134)
$\beta^{sb}_{22,3}$	0.207	0.0017	(0.204, 0.211)	$\beta^{sb}_{24,3}$	0.206	0.0024	(0.201, 0.210)
Employee satisfaction				Retention			
$\beta_{23,1}$	0.189	0.0026	(0.184, 0.194)	$\beta_{25,1}$	0.293	0.0083	(0.277, 0.310)
$\beta^T_{23,1}$	0.229	0.0022	(0.224, 0.233)	$\beta^T_{25,1}$	0.325	0.0072	(0.311, 0.339)
$\beta^{sb}_{23,1}$	0.250	0.0018	(0.247, 0.254)	$\beta^{sb}_{25,1}$	0.376	0.0059	(0.365, 0.388)
$\beta_{23,2}$	0.017	0.0020	(0.013, 0.021)	$\beta_{25,2}$	0.075	0.0069	(0.062, 0.089)
$\beta^T_{23,2}$	0.164	0.0017	(0.160, 0.167)	$\beta^T_{25,2}$	0.270	0.0056	(0.259, 0.280)
$\beta^{sb}_{23,2}$	0.164	0.0017	(0.160, 0.167)	$\beta^{sb}_{25,2}$	0.270	0.0056	(0.259, 0.280)
$\beta_{23,3}$	0.102	0.0027	(0.097, 0.107)	$\beta_{25,3}$	0.083	0.0081	(0.067, 0.099)
$\beta^T_{23,3}$	0.102	0.0027	(0.097, 0.107)	$\beta^T_{25,3}$	0.083	0.0081	(0.067, 0.099)
$\beta^{sb}_{23,3}$	0.206	0.0021	(0.202, 0.210)	$\beta^{sb}_{25,3}$	0.266	0.0063	(0.254, 0.278)
α_{11}	-2.226	0.0202	(-2.266, -2.187)	σ^2_τ	0.006	0.0009	(0.005, 0.008)
α_{21}	0.508	0.0293	(0.451, 0.565)	$\sigma^2_{\tau^*}$	0.006	0.0002	(0.005, 0.006)
α_{22}	0.385	0.0018	(0.382, 0.389)	γ_1	0.639	0.0065	(0.627, 0.652)
α_{31}	0.941	0.0113	(0.918, 0.963)	γ_2	0.279	0.0073	(0.265, 0.293)
α_{32}	1.433	0.0457	(1.343, 1.522)	γ_3	0.390	0.0073	(0.375, 0.404)
α_{33}	-0.139	0.0180	(-0.175, -0.104)				

7 Discussion

In this paper, we have carried out an extensive investigation of the ordinal response SEM. We have shown that the use of latent variables linking the unobserved continuous variables y^*_{ijk}'s to the observed ordinal or binary response variables y_{ijk}'s is a useful modeling technique, which facilitates an easy implementation of the posterior computation and provides a great flexibility to incorporate different types of responses such as continuous, ordinal, and binary responses in the survey data. For the AES 2001 data, we have observed that the ordinal SEM with both facility level random effect terms as well as individual covariates fits the data best based on the DIC measure even though it has the largest penalty for dimensionality. Also, including facility random effect terms is more important than including just covariate information, implying that variability within facilities is more pronounced than variability in terms of demographic variables within individuals. This finding confirms that facility characteristics differ not only in size but also in the available service resources, and responses by individuals belonging to different facilities are bound to reflect that. In fact, including facility random effects captures the natural heterogeneity because of the clustering of individuals into specific facilities. We note that the similar results were also obtained by Das et al. (2008) based on the VHA AES 1997 data.

Table 6

Posterior estimates of Model 1 (complete cases) under the normal-ordinal SEM

Variable	Estimate	SD	95% HPD	Variable	Estimate	SD	95% HPD
Customer satisfaction				Quality			
$\beta_{22,1}$	0.032	0.0048	(0.025, 0.040)	$\beta_{24,1}$	0.145	0.0052	(0.134, 0.155)
$\beta^{T}_{22,1}$	0.109	0.0035	(0.101, 0.115)	$\beta^{T}_{24,1}$	0.197	0.0046	(0.188, 0.206)
$\beta^{sb}_{22,1}$	0.147	0.0027	(0.142, 0.153)	$\beta^{sb}_{24,1}$	0.233	0.0037	(0.226, 0.241)
$\beta_{22,2}$	0.027	0.0029	(0.021, 0.033)	$\beta_{24,2}$	0.035	0.0039	(0.028, 0.043)
$\beta^{T}_{22,2}$	0.130	0.0024	(0.125, 0.135)	$\beta^{T}_{24,2}$	0.177	0.0034	(0.171, 0.184)
$\beta^{sb}_{22,2}$	0.130	0.0024	(0.125, 0.135)	$\beta^{sb}_{24,2}$	0.177	0.0034	(0.171, 0.184)
$\beta_{22,3}$	0.179	0.0038	(0.172, 0.187)	$\beta_{24,3}$	0.123	0.0051	(0.114, 0.133)
$\beta^{T}_{22,3}$	0.179	0.0038	(0.172, 0.187)	$\beta^{T}_{24,3}$	0.123	0.0051	(0.113, 0.133)
$\beta^{sb}_{22,3}$	0.210	0.0027	(0.205, 0.215)	$\beta^{sb}_{24,3}$	0.219	0.0040	(0.212, 0.228)
Employee satisfaction				Retention			
$\beta_{23,1}$	0.211	0.0045	(0.202, 0.219)	$\beta_{25,1}$	0.327	0.0149	(0.299, 0.357)
$\beta^{T}_{23,1}$	0.251	0.0039	(0.244, 0.259)	$\beta^{T}_{25,1}$	0.351	0.0127	(0.328, 0.378)
$\beta^{sb}_{23,1}$	0.270	0.0030	(0.264, 0.276)	$\beta^{sb}_{25,1}$	0.409	0.0101	(0.390, 0.429)
$\beta_{23,2}$	0.012	0.0034	(0.005, 0.018)	$\beta_{25,2}$	0.085	0.0119	(0.062, 0.108)
$\beta^{T}_{23,2}$	0.180	0.0030	(0.174, 0.186)	$\beta^{T}_{25,2}$	0.303	0.0095	(0.285, 0.323)
$\beta^{sb}_{23,2}$	0.180	0.0030	(0.174, 0.186)	$\beta^{sb}_{25,2}$	0.303	0.0095	(0.285, 0.323)
$\beta_{23,3}$	0.096	0.0045	(0.087, 0.105)	$\beta_{25,3}$	0.059	0.0143	(0.030, 0.086)
$\beta^{T}_{23,3}$	0.096	0.0045	(0.087, 0.105)	$\beta^{T}_{25,3}$	0.059	0.0143	(0.030, 0.086)
$\beta^{sb}_{23,3}$	0.218	0.0035	(0.211, 0.225)	$\beta^{sb}_{25,3}$	0.279	0.0108	(0.258, 0.301)
γ_1	0.721	0.0118	(0.698, 0.744)	σ^2_τ	0.004	0.0006	(0.002, 0.005)
γ_2	0.270	0.0130	(0.244, 0.295)	$\sigma^2_{\tau*}$	0.004	0.0002	(0.003, 0.004)
γ_3	0.425	0.0123	(0.400, 0.448)				

In Section 6, we have empirically shown that treating the categorical responses as ordinal or continuous for the 21 manifest variables results in very similar posterior estimates of the loading coefficients for the 4 ordinal/binary outcome variables in the measurement part of the SEM and the parameters in the structural part of the SEM. We have also observed that the same best DIC model is obtained under both the ordinal SEMs and the normal-ordinal SEMs. Interestingly, such treatment of the categorical responses for the manifest variables does not at all affect the posterior estimates of parameters in the missing covariates models.

Finally, we mention that in the ordinal SEM, the probit links were used for all ordinal or binary responses. As discussed in Chen, Dey, and Shao (1999) and Kim, Chen, and Dey (2008), the choice of links is important in fitting categorical response data. However, the literature on the importance of the choice of links in the structural equations framework is still sparse. In addition, in this paper, we assume that missing responses and/or missing covariates are missing at random. This assumption needs to be further examined. These important issues are currently under investigation.

Appendix: Posterior Computation

To implement the MCMC sampling algorithm, we need to sample from the following conditional posterior distributions in turn:

(i) $[\boldsymbol{y}^*, \boldsymbol{\lambda} | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(ii) $[\boldsymbol{\mu} | \boldsymbol{y}^*, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(iii) $[\boldsymbol{\sigma}^2 | \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(iv) $[\boldsymbol{\beta} | \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(v) $[\boldsymbol{\phi} | \boldsymbol{y}^*, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(vi) $[\sigma^2_{\tau} | \boldsymbol{\tau}, D_{obs}]$;

(vii) $[\sigma^2_{\tau^*} | \boldsymbol{\tau}^*, D_{obs}]$;

(viii) $[\boldsymbol{\eta} | \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{z}_{mis}, D_{obs}]$;

(ix) $[\boldsymbol{\tau} | \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \sigma^2_{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(x) $[\boldsymbol{\tau}^* | \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \sigma^2_{\tau^*}, \boldsymbol{\tau}, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}]$;

(xi) $[\boldsymbol{\gamma} | \boldsymbol{\eta}, D_{obs}]$;

(xii) $[\boldsymbol{\alpha} | \boldsymbol{z}_{mis}, D_{obs}]$; and

(xiii) $[\boldsymbol{z}_{mis} | \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{\alpha}, D_{obs}]$.

We briefly discuss how we sample from each of the above posterior conditional distributions. For (i), we apply the collapsed Gibbs technique of Liu (1994) via the following identity:

$$[\boldsymbol{y}^*, \boldsymbol{\lambda} | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}] = [\boldsymbol{\lambda} | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}] \\ \times [\boldsymbol{y}^* | \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}].$$

That is, we sample $\boldsymbol{\lambda}$ after collapsing out $\boldsymbol{y}^*$. It can be shown that given $\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}$, $\boldsymbol{z}_{mis}$, and D_{obs}, $\boldsymbol{\lambda}_k$'s are conditionally independent. Thus, we can sample each $\boldsymbol{\lambda}_k$ independently. Instead of directly sampling $\boldsymbol{\lambda}_k$, we use the "de-constrained" transformation for the Metropolis-Hastings algorithm proposed by Chen and Dey (2000b) as follows

$$\lambda_{kl} = \frac{\lambda_{k,l-1} + \exp(\zeta_{kl})}{1 + \exp(\zeta_{kl})},$$

where $-\infty < \zeta_{kl} < \infty,\ l = 2, \ldots, L_k - 2$. Let $\pi(\boldsymbol{\lambda}_k | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs})$ denote the conditional distribution of $\boldsymbol{\lambda}_k$. Then, the conditional distribution of $\boldsymbol{\zeta}_k = (\zeta_{kl},\ l = 2, \ldots, L_k - 2)'$ is obtained by

$$\pi(\boldsymbol{\zeta}_k | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}) \\ \propto \pi(\boldsymbol{\lambda}_k | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}) \prod_{l=2}^{L_k-2} \frac{(1 - \lambda_{k,l-1}) \exp(\zeta_{kl})}{[1 + \exp(\zeta_{kl})]^2}.$$

Let $\hat{\boldsymbol{\zeta}}_k = \text{argmax}_{\boldsymbol{\zeta}_k} \log \pi(\boldsymbol{\zeta}_k | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs})$ and $\hat{\Sigma}_{\zeta_k}$ is minus the inverse matrix of the second derivative of $\log \pi(\boldsymbol{\zeta}_k | \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs})$ evaluated at $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}}_k$. We then sample $\boldsymbol{\zeta}_k$ via the standard Metropolis algorithm using the proposal $\mathcal{N}(\hat{\boldsymbol{\zeta}}_k, \hat{\Sigma}_{\zeta_k})$. Since the conditional distribution of y^*_{ijk} is a truncated normal, $\mathcal{N}(\mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} + \boldsymbol{\phi}'_k \boldsymbol{z}_{ij}, \sigma^2_k) 1\{\lambda_{y_{ijk}-1} \leq y^*_{ijk} < \lambda_{y_{ijk}}\}$, sampling y^*_{ijk} is straightforward. Note that when $\delta_{ijk} = 0$, we do not need to sample y^*_{ijk}.

For (ii), given $\boldsymbol{y}^*, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}$, the μ_k are conditionally independent and

$$\mu_k | \boldsymbol{y}^*, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs} \\ \sim \mathcal{N}\left(\frac{\frac{\sum_{i=1}^I \sum_{j=1}^{n_i} \delta_{ijk} (y^*_{ijk} - \tau_i - \tau_{ik} - \boldsymbol{\beta}'_k \boldsymbol{\omega}_k \boldsymbol{\eta}_{ij} - \boldsymbol{\phi}'_k \boldsymbol{z}_{ij})}{\sigma^2_k} + \frac{\mu_0}{\sigma^2_0}}{\frac{\sum_{i=1}^I \sum_{j=1}^{n_i} \delta_{ijk}}{\sigma^2_k} + \frac{1}{\sigma^2_0}},\ \frac{1}{\frac{\sum_{i=1}^I \sum_{j=1}^{n_i} \delta_{ijk}}{\sigma^2_k} + \frac{1}{\sigma^2_0}} \right),$$

for $k = 1, 2, \ldots, K$. For (iii), again the σ_k^2 are conditionally independent and distributed as

$$\sigma_k^2|\boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}$$
$$\sim \mathcal{IG}\left(\frac{\sum_{i=1}^{I}\sum_{j=1}^{n_i}\delta_{ijk}}{2} + a_0,\ \frac{\sum_i\sum_j \delta_{ijk}[y_{ijk}^* - (\mu_k + \tau_i + \tau_{ik} + \boldsymbol{\beta}_k'\boldsymbol{\omega}_k\boldsymbol{\eta}_{ij} + \boldsymbol{\phi}_k'\boldsymbol{z}_{ij})]^2}{2} + b_0 \right),$$

for $k = 1, 2, \ldots, K-1$. For (iv), we have

$$\boldsymbol{\beta}_k \mid \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis} \ \sim\ N_{p_k}(B_{\boldsymbol{\beta}_k}^{-1} A_{\boldsymbol{\beta}_k}, B_{\boldsymbol{\beta}_k}^{-1})$$

for $k = 1, 2, \ldots, K$, where $A_{\boldsymbol{\beta}_k} = \frac{1}{\sigma_k^2}\sum_{i=1}^{I}\sum_{j=1}^{n_i}\boldsymbol{\omega}_k\boldsymbol{\eta}_{ij}\delta_{ijk}(y_{ijk}^* - \mu_k - \tau_i - \tau_{ik} - \boldsymbol{\phi}_k'\boldsymbol{z}_{ij}) + \Sigma_0^{-1}\boldsymbol{\beta}_0$ and $B_{\boldsymbol{\beta}_k} = \frac{1}{\sigma_k^2}\boldsymbol{\omega}_k\Big[\sum_{i=1}^{I}\sum_{j=1}^{n_i}\boldsymbol{\eta}_{ij}\boldsymbol{\eta}_{ij}'\Big]\boldsymbol{\omega}_k' + \Sigma_0^{-1}$. For (v),

$$\boldsymbol{\phi}_k \mid \boldsymbol{y}^*, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs} \ \sim\ \mathcal{N}_q(B_{\boldsymbol{\phi}_k}^{-1} A_{\boldsymbol{\phi}_k}, B_{\boldsymbol{\phi}_k}^{-1}),$$

where

$$A_{\boldsymbol{\phi}_k} = \frac{1}{\sigma_k^2}\sum_{i=1}^{I}\sum_{j=1}^{n_i}(y_{ijk}^* - \mu_k - \tau_i - \tau_{ik} - \boldsymbol{\beta}_k'\boldsymbol{\omega}_k\boldsymbol{\eta}_{ij})\boldsymbol{z}_{ij} + \Sigma_{\phi_0}^{-1}\boldsymbol{\phi}_0$$

and $B_{\boldsymbol{\phi}_k} = \frac{1}{\sigma_k^2}\sum_{i=1}^{I}\sum_{j=1}^{n_i}\boldsymbol{z}_{ij}\boldsymbol{z}_{ij}' + \Sigma_{\phi_0}^{-1}$. For (vi)

$$\sigma_\tau^2 \mid \boldsymbol{\tau}, D_{obs} \sim \mathcal{IG}\left(a_1 + \frac{I}{2},\ \frac{1}{2}\sum_{i=1}^{I}\tau_i^2 + b_1\right),$$

and for (vii)

$$\sigma_{\tau*}^2|\boldsymbol{\tau}^*, D_{obs} \sim \mathcal{IG}\left(\frac{IK}{2} + a_2, \frac{\sum_i\sum_k\tau_{ik}^2}{2} + b_2\right).$$

For the latent variables in (viii),

$$\boldsymbol{\eta}_{ij}|\boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{z}_{mis}, D_{obs} \ \sim\ \mathcal{N}_r(B_{\boldsymbol{\eta}_{ij}}^{-1} A_{\boldsymbol{\eta}_{ij}}, B_{\boldsymbol{\eta}_{ij}}^{-1}),$$

where $A_{\boldsymbol{\eta}_{ij}} = \sum_{k=1}^{K}\Big[\frac{1}{\sigma_k^2}\boldsymbol{\omega}_k'\boldsymbol{\beta}_k\delta_{ijk}(y_{ijk}^* - \mu_k - \tau_i - \tau_{ik} - \boldsymbol{\phi}_k'\boldsymbol{z}_{ij})\Big]$ and $B_{\boldsymbol{\eta}_{ij}} = \sum_{k=1}^{K}\Big[\frac{\delta_{ijk}}{\sigma_k^2}\boldsymbol{\omega}_k'\boldsymbol{\beta}_k\boldsymbol{\beta}_k'\boldsymbol{\omega}_k\Big]$ $+V_\eta^{-1}$. For (ix), the τ_i are conditionally independent and distributed as

$$\tau_i \mid \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \sigma_\tau^2, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}$$
$$\sim \mathcal{N}\left(\frac{\sum_{j=1}^{n_i}\sum_{k=1}^{K}\delta_{ijk}(y_{ijk}^* - \mu_k - \tau_{ik} - \boldsymbol{\beta}_k'\boldsymbol{\omega}_k\boldsymbol{\eta}_{ij} - \boldsymbol{\phi}_k'\boldsymbol{z}_{ij})/\sigma_k^2}{\sum_{j=1}^{n_i}\sum_{k=1}^{K}\frac{\delta_{ijk}}{\sigma_k^2} + \frac{1}{\sigma_\tau^2}},\ \frac{1}{\sum_{j=1}^{n_i}\sum_{k=1}^{K}\frac{\delta_{ijk}}{\sigma_k^2} + \frac{1}{\sigma_\tau^2}} \right),$$

and for (x),

$$\begin{aligned}&\tau_{ik} \mid \boldsymbol{y}^*, \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \sigma^2_{\tau^*}, \boldsymbol{\tau}, \boldsymbol{\eta}, \boldsymbol{z}_{mis}, D_{obs}\\ &\sim \mathcal{N}\left(\frac{\sum_{j=1}^{n_i}\delta_{ijk}(y^*_{ijk}-\mu_k-\tau_i-\boldsymbol{\beta}'_k\boldsymbol{\omega}_k\boldsymbol{\eta}_{ij}-\boldsymbol{\phi}'_k\boldsymbol{z}_{ij})}{\sigma^2_k(\frac{\sum_{j=1}^{n_i}\delta_{ijk}}{\sigma^2_k}+\frac{1}{\sigma^2_{\tau*}})},\ \frac{1}{\frac{\sum_{j=1}^{n_i}\delta_{ijk}}{\sigma^2_k}+\frac{1}{\sigma^2_{\tau*}}}\right).\end{aligned}$$

The conditional distributions for (ii) to (x) are either normal or inverse gamma distributions and therefore, sampling from these distributions is straightforward.

For (xi), we use the localized Metropolis algorithm discussed in Chen, Shao, and Ibrahim (2000, Chapter 2) to sample $\boldsymbol{\gamma}$ from $[\boldsymbol{\gamma}|\boldsymbol{\eta}, D_{obs}]$. Let

$$\pi^*(\boldsymbol{\gamma}|\boldsymbol{\eta}, D_{obs}) = \Big[\prod_{i=1}^{I}\prod_{j=1}^{n_i}|V_\eta|^{-1/2}\exp\Big\{-\frac{1}{2}\boldsymbol{\eta}'_{ij}V_\eta^{-1}\boldsymbol{\eta}_{ij}\Big\}\Big]\pi(\boldsymbol{\gamma}),$$

where $\pi(\boldsymbol{\gamma})$ is the prior for $\boldsymbol{\gamma}$. We compute

$$\widehat{\boldsymbol{\gamma}} = \operatorname*{argmax}_{\gamma_1,\gamma_2,\gamma_3}\log\pi^*(\boldsymbol{\gamma}|\boldsymbol{\eta}, D_{obs}) \ \text{ and }\ \widehat{\Sigma} = \left[-\frac{\partial^2\log\pi^*(\boldsymbol{\gamma}|\boldsymbol{\eta}, D_{obs})}{\partial\gamma_i\partial\gamma_j}\Bigg|_{\boldsymbol{\gamma}=\widehat{\boldsymbol{\gamma}}}\right]^{-1}.$$

We use $\mathcal{N}(\widehat{\boldsymbol{\gamma}}, c^*\widehat{\Sigma})$ as the proposed density for the localized Metropolis algorithm, where c^* is a tuning parameter.

For (xii), $\pi(\boldsymbol{\alpha}|\boldsymbol{z}_{mis}, D_{obs}) \propto \Big[\prod_{i=1}^{I}\prod_{j=1}^{n_i} f(\boldsymbol{z}_{ij,obs}, \boldsymbol{z}_{ij,mis}|\boldsymbol{\alpha})\Big]\pi(\boldsymbol{\alpha})$. For various covariate distributions specified through a series of one dimensional conditional distributions, sampling $\boldsymbol{\alpha}$ is straightforward. For (xiii), given $\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{\alpha}$, and D_{obs}, $\boldsymbol{z}_{ij,mis}$'s are independent across all i and j, and the conditional distribution for $\boldsymbol{z}_{ij,mis}$ is

$$\begin{aligned}&\pi(\boldsymbol{z}_{ij,mis}|\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{\alpha}, D_{obs})\\ &\propto\prod_{k=1}^{K}\Big\{\Phi\Big(\frac{\lambda_{k,y_{ijk}}-\mu^*_{ijk}}{\sigma_k}\Big)-\Phi\Big(\frac{\lambda_{k,y_{ijk}-1}-\mu^*_{ijk}}{\sigma_k}\Big)\Big\}^{\delta_{ijk}} f(\boldsymbol{z}_{ij,obs}, \boldsymbol{z}_{ij,mis}|\boldsymbol{\alpha}),\end{aligned}$$

where $\mu^*_{ijk} = \mu_k+\tau_i+\tau_{ik}+\boldsymbol{\beta}'_k\boldsymbol{\omega}_k\boldsymbol{\eta}_{ik}+\boldsymbol{\phi}'_k\boldsymbol{z}_{ij}$. Thus, sampling $\boldsymbol{z}_{mis}$ depends on the form of $f(\boldsymbol{z}_{ij,obs}, \boldsymbol{z}_{ij,mis}|\boldsymbol{\alpha})$. For the VHA AES 2001 data, $\pi(\boldsymbol{z}_{ij,mis}|\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\tau}^*, \boldsymbol{\eta}, \boldsymbol{\alpha}, D_{obs})$ is simply a multinomial distribution, which is easy to sample from.

Acknowledgments

The authors wish to thank the Guest Editor and a referee for helpful comments which have led to an improvement of this article. The authors also wish to acknowledge and thank the

Veterans Health Administration and the National Center for Organizational Development for access to the All Employee Survey (AES) 2001 data and their substantial efforts to insure the quality and integrity of these data. Dr. Kim's research is supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development; National Institutes of Health. Dr. Chen's research was partially supported by NIH grants #GM 70335 and #CA 74015 and a Faculty Large Research Grant from University of Connecticut. Dr. Warren's research was funded by the VHA National Center for Organizational Development.

References

Albert, J. H. (1992). Bayesian estimation of normal ogive item response curves using Gibbs sampling. *J. Educ. Statist.* 17:251-269.

Albert, J. H., Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* 88:669-679.

Arminger, G., Muthén, B. O. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika* 63:271-300.

Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. *Sociolog. Methodol.* 17:37-69.

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

Branden-Roche, K., Miglioretti, D., Zeger, S., Rathouz, P. (1997). Latent variable regression for multiple discrete outcomes. *J. Amer. Statist. Assoc.* 92:1375-1386.

Chen, C.-F. (1981). The EM approach to the multiple indicators and multiple causes model via the estimation of the latent variable. *J. Amer. Statist. Assoc.* 76:704-708.

Chen, M.-H., Dey, D. K. (1998). Bayesian modeling of correlated binary responses via scale mixture of multivariate normal link functions. *Sankhyā, Ser. A*, 60:322-343.

Chen, M.-H., Dey, D. K. (2000a). Bayesian data analysis of longitudinal binary response models. In Ghosh, J.K., Sen, P.K., Sinha, B.K., eds. *Proceedings of the Third International Triennial Calcutta Symposium on Probability and Statistics.* Oxford: Oxford University Press, pp. 100-123.

Chen, M.-H., Dey, D. K. (2000b). A unified Bayesian analysis for correlated ordinal data models. *Brazilian J. Probab. Statist.* 14:87-111.

Chen, M.-H., Dey, D. K., Shao, Q.-M. (1999). A new skewed link model for dichotomous quantal response data. *J. Amer. Statist. Assoc.* 94:1172-1186.

Chen, M.-H., Shao, Q.-M., Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation.* New York: Springer-Verlag.

Chen, Q., Ibrahim, J.G., Chen, M.-H., and Senchaudhuri, P. (2008). Theory and inference for regression models with missing response and covariates. *J. Multivariate Anal.* 99:1302-1331.

Chib, S., Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika* 85:347-361.

Cowles, C., Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *J. Amer. Statist. Assoc.* 91:883-904.

Das, S., Chen, M.-H., Kim, S., Warren, N. (2008), A Bayesian structural equations model for multilevel data with missing responses and missing covariates. *Bayesian Analysis* 3:197-224.

Eaton, W. W., Bohrnstedt, G. (1989). Introduction to latent variable models for dichotomous outcomes: analysis of data from the epidemiologic catchment area program. *Sociolog. Methods Research* 18:4-18.

Hagenaars, J. A. (1990). *Categorical Longitudinal Data - Log-Linear Analysis of Panel, Trend and Cohort Data.* Newbury Park: Sage.

Hatcher, L. (2000). *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling.* Cary, NC: SAS Institute Inc.

Huang, L., Chen, M.-H., Ibrahim, J. G. (2005). Bayesian analysis for generalized linear models with nonignorably missing covariates. *Biometrics* 61:767-780.

Ibrahim, J. G., Chen, M.-H., Lipsitz, S. R., Herring, A. H. (2005). Missing data methods in regression models. *J. Amer. Statist. Assoc.* 100:332-346.

Ibrahim, J. G., Lipsitz, S. R., Chen, M.-H. (1999). Missing covariates in generalized linear models when the missing data mechanism is nonignorable. *J. Roy. Statist. Soc., Ser. B* 61:173-190.

Kim, S., Chen, M.-H., Dey, D. K. (2008). Flexible generalized t-link models for binary response data. *Biometrika* 95:93-106.

Lee, S.-Y., Poon, W.-Y., Bentler, P. M. (1992). Structural equation models with continuous and polytomous variables. *Psychometrika* 57:89-105.

Li, F., Harmer, P., Duncan, T. E., Duncan, S. C., Acock, A., Boles, S. (1998). Approaches to testing interaction effects using structural equation modeling methodology. *Multivariate Behavioral Research* 33:1-39.

Lipsitz, S. R., Ibrahim, J. G. (1996). A conditional model for incomplete covariates in parametric regression models. *Biometrika* 83:916-922.

Little, R. J. A., Rubin, D. B. (2002). *Statistical Analysis with Missing Data.* Second Edition. New York: John Wiley.

Liu, J. S. (1994). The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *J. Amer. Statist. Assoc.* 89:958-966.

McCullagh, P., Nelder, J. A. (1989). *Generalized Linear Models.* London: Chapman & Hall.

Meredith, W., Millsap, R. E. (1992). On the misuse of manifest variables in the detection of measurement bias. *Psychometrika* 57:289-311

Nandram, B. and Chen, M.-H. (1996). Accelerating Gibbs sampler convergence in the generalized linear models via a reparameterization. *J. Statist. Comput. Simul.* 54:129-144.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation. *Psychometrika* 44:443-460.

Rubin, D. B. (1976). Inference and missing data. *Biometrika* 63:581-592.

Skrondal, A., Rabe-Hesketh, S. (2005). Structural equation modeling: categorical variables. In Everitt, B., Howell, D., eds., *Encyclopedia of Statistics in Behavioral Science*. New York: Wiley, pp. 2352.

Spiegelhalter, D. J., Best, N. G., Carlin., B. P., van der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. Roy. Statist. Soc., Ser. B* 62:583-639.

Zhu, H. T., Lee, S. Y. (1999). Statistical analysis of nonlinear factor analysis models. *British J. Mathemat. Statist. Psychology* 52:225-242.