Psychometrika (2025), **90**, 251–279
doi:10.1017/psy.2024.13

PSYCHOMETRIC SOCIETY

THEORY AND METHODS

Mediation Analysis in Bayesian Extended Redundancy Analysis with Mixed Outcome Variables

Ji Yeh Choi[1], Minjung Kyung[2] and Ju-Hyun Park[3]

[1]Department of Psychology, York University, 4700 Keele St., Toronto, ON, Canada; [2]Department of Statistics, Duksung Women's University, 33 Samyang-ro, 144-gil, Dobong-gu, Seoul, Republic of Korea; [3]Department of Statistics, Dongguk University, 30 Phildong-ro 1-gil, Jung-gu, Seoul, Republic of Korea
Corresponding author: Ju-Hyun Park; Email: juhyunp@dongguk.edu.

(Received 11 October 2024; accepted 17 October 2024; published online 3 January 2025)

Abstract
Extended redundancy analysis (ERA) is a statistical approach to component-based multivariate regression modeling that explores interrelationships among multiple sets of while incorporating regression with a data-reduction technique. The extant models that utilize ERA have assumed the outcome variables with the same data type. Also, ERA models focused on estimating direct pathways only without explicitly addressing mediation effects. In this paper, ERA is extended to handle multiple mediators and mixed types of outcome variables by adopting a Bayesian framework, taking into account correlation structure among all of the outcome variables. The proposed method develops an algorithm that derives the joint posterior distribution of parameters using a Markov chain Monte Carlo algorithm. Simulations and an empirical dataset are provided to illustrate the usefulness of the proposed method.

Keywords: Bayesian statistics; Extended redundancy analysis; mediation analysis; multivariate regression with mixed types of variables

1. Introduction

Regression models, encompassing multiple predictors and outcome variables, are pervasive in the social sciences, where research endeavors seek to comprehend the relationships among various variables (e.g., Aiken et al., 1991; Cohen et al., 2003). The inclusion of numerous variables, particularly as predictors, often introduces a level of dependence among them, potentially resulting in the well-known collinearity issue. Several methodologies have been proposed to tackle collinearity, one of which involves integrating a regression model with a data-reduction technique. Extended redundancy analysis (ERA) is one such technique (e.g., DeSarbo et al., 2015; Hwang et al., 2012; Lee et al., 2016, 2018; Lovaglio & Vacca, 2016; Lovaglio & Vittadini, 2014; Takane & Hwang, 2005; Tan et al., 2015).

In ERA, a component is derived from a set of predictors as a weighted composite, maximizing the explained variance in the outcome variables. From a technical standpoint, ERA can be seen as a special case of structural equation models (SEM). This classification arises from ERA's specification of a formative relation between the predictor set and each component and its exploration of the relationship between the constructed components and outcome variables.

To date, the majority of regression models employing data-reduction techniques, including ERA, have typically assumed that the outcome variables share the same data type, such as being either all continuous or all binary. However, in practical scenarios, encountering multivariate outcome data with diverse types is common—examples include mixed measurements involving both continuous and categorical outcomes. In such instances, we would no longer assume a simplified correlation structure such as compound symmetry or independence, which were conventionally used for most applications with multivariate data in behavioral and social sciences studies (e.g., Henningsson et al., 2001; Lovaglio & Vittadini, 2014).

Moreover, the application of mediation analysis has seen a growing prevalence across diverse fields, including psychology, sociology, economics, and other social sciences (e.g., Bullock et al., 2010; MacKinnon & Fairchild, 2009; Selig & Preacher, 2021; VanderWeele, 2016). According to Pieters (2017), mediation plays a crucial role in theory development, serving as an indispensable tool for assessing potential intermediate effects through intervening variables, referred to as mediators. It is conceptualized as a third-variable effect that illuminates the relationship between a predictor variable and an outcome variable (e.g., Baron & Kenny, 1986; Preacher & Hayes, 2008). Researchers frequently utilize latent variables with SEM framework to investigate and validate theoretical models that encompass both direct and indirect pathways among variables. This involves examining how the influence of a predictor on an outcome variable operates through intermediary mediators. Despite these advancements, extant ERA models have given limited attention to explicating and analyzing such mediation effects.

This paper presents a Bayesian extension to ERA, expanding its application to facilitate not only the estimation of indirect effects involving multiple mediators but also the accommodation of a diverse array of outcome variable types. Implementing ERA with ordinal variables is challenging in the Frequentist approach due to its reliance on the alternating least squares algorithm. In contrast, the Bayesian approach offers greater flexibility in modeling complex structure, such as ordinal variables, by incorporating prior information and probabilistic reasoning. By leveraging a Markov Chain Monte Carlo (MCMC) algorithm, we derive the joint posterior distribution of key parameters, including indirect effects arising from components influencing outcome variables through mediators. To estimate indirect effects as a distinct set of parameters, we build on the fundamental ERA model specification, constructing components as weighted composites of observed predictors, and formulate a unified objective function encompassing both direct and mediation effects.

Furthermore, to enhance the efficiency of the MCMC algorithm while avoiding constraints on the covariance structure of outcome variables with mixed types, we adopt the assumption that a set of continuous latent variables, which underlies ordinal outcome variables, conforms to a multivariate t distribution (Park et al., 2021). This strategy enables the incorporation of correlations with continuous outcome variables, which are presumed to adhere to a multivariate t distribution, thereby facilitating a more flexible and realistic modeling approach for mixed-type outcome data.

The proposed method is designed to serve several key objectives. Firstly, it aims to preserve the conceptual associations between predictors and their components, thereby handling multicollinearity in cases where a set of predictors exhibits high levels of correlation. Secondly, it seeks to incorporate the advantageous features of the Bayesian framework into ERA, allowing for the joint modeling of mixed outcome data that might otherwise be overlooked in social science research. Lastly, it provides a viable alternative for exploring potential intermediate effects through mediators within the ERA framework.

The remaining part of this paper is organized as follows. Section 2 first explains ERA with mixed types of outcome variables and then the full version of the proposed method that also includes multiple mediators in detail with parameter estimation. Section 3 provides a simulation study. Section 4 provides an application to illustrate the empirical usefulness of the proposed method. The final section summarizes the implications and possible extensions of the proposed method.

2. Method

2.1. Extended redundancy analysis

Let y_{it} denote the ith value of the tth outcome variable $(i = 1, \cdots N; t = 1, \cdots, T)$ and x_{ilk} the ith value of the lth predictor in the kth set $(l = 1, \cdots, p_k$ and $k = 1, \cdots, K)$, where p_k refers to the number of predictors in the kth set. Let $P = \sum_{k=1}^{K} p_k$ be the total number of predictors in K sets. Let w_{lk} denote a component

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

weight assigned to x_{ilk}. Let f_{ik} denote the ith component score for the kth component defined as a linear combination or weighted composite of the predictors in the kth set, i.e., $f_{ik} = \left[\sum_{l=1}^{p_k} x_{ilk} w_{lk}\right]$. Let a_{kq} denote the kth regression coefficient connecting the kth component to the outcome variable y_{it}, and e_{it} denote the ith residual value for y_{it}. ERA model can be written as follows:

$$\begin{aligned} y_{iq} &= \sum_{k=1}^{K} \left[\sum_{l=1}^{p_k} x_{ilk} w_{lk}\right] a_{kq} + e_{iq} \\ &= \sum_{k=1}^{K} f_{ik} a_{kq} + e_{iq}. \end{aligned} \tag{1}$$

In matrix notation, (1) is re-expressed as

$$\begin{aligned} \mathbf{Y} &= \mathbf{XWA} + \mathbf{E} \\ &= \mathbf{FA} + \mathbf{E}, \end{aligned}$$

where $\mathbf{Y}$ is an N by Q matrix of outcome variables, $\mathbf{X}$ is an N by P matrix of predictors, $\mathbf{W}$ is a P by K matrix of weights, $\mathbf{A}$ is a K by Q matrix of regression coefficients, and $\mathbf{E}$ is an N by Q matrix of residuals. For identifiability of $\mathbf{F}$, a standardization constraint is imposed on $\mathbf{F}$ such that $diag\left(\mathbf{F}'\mathbf{F}\right) = N\mathbf{I}$. More details on the Frequentist ERA can be found in Takane and Hwang (2005) or Choi et al. (2020).

There have been several methodologies described to handle the collinearity problem, among which incorporating regression model with a data-reduction technique is an option. Principal component regression (PCR; Jolliffe, 1982), partial least squares (PLS; Wold, 1975; Wold et al., 1984), and extended redundancy analysis (ERA; Takane & Hwang, 2005) share similarity such that they employ data reduction techniques to transform the predictors into a new set of uncorrelated underlying or latent constructs, more specifically called components. However, PCR extracts components by maximizing the explained variance of the predictors only without considering their associations with an outcome variable. Subsequently, the components are used as the predictors (to fit a regression model by least squares) in a regression model to predict the outcome variable (e.g., Abdi, 2010; Wehrens & Mevik, 2007). In PCR, because components are extracted independently of a regression model, they may not be optimal in explaining the outcome variable the best.

Different from PCR, PLS and ERA do take into account the association between the predictors and an outcome variable when extracting the components. A major distinct feature that differentiates ERA from PLS is on whether or not a single or unified objective function is derived for parameter estimation. While PLS sets up two different objective functions for extracting components and employing a regression model, respectively, ERA estimates the unknown parameters using a single (global) objective function. Also, ERA allows to handle multiple sets of predictors simultaneously whereas PLS involves only one set of predictors. In this paper, we will focus on the most inclusive form of regression with components, i.e., ERA, employing a unified single objective function.

2.2. ERA with mixed types of outcome variables

We consider multiple sets of Q outcome variables which are combinations of a set of T continuous responses and a set of $Q - T$ ordinal responses of C categories. In many cases, outcome variables are correlated, and we need to consider the interdependency among outcome variables in the model regardless of the response structure to avoid biased statistical inference.

2.2.1. Univariate model structure with original regression mean

For continuous outcome variables, we consider a robust regression model to outliers, in the sense that a single out of bounds data point can strongly affect the inference for all the parameters in the model. We are able to reduce the influence of the influential points with considering a longer-tailed family of distributions compared to the normal population model, which allows for the possibility of extreme

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

observations. One of the longer-tailed distributions which are considered frequently is the family of t distributions. In our work, we consider the distribution of errors with t distribution in the place of the normal. Thus, within the multiple regression context, we consider that for the qth response:

$$\mathbf{Y}_{[i,q]} \sim t_\nu\left(\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}, \sigma_{iq}^2\right) \text{ if } \mathbf{Y}_{[i,q]} \text{ is a continuous response,}$$

$$\mathbf{Y}_{[i,q]} \sim multinomial\left(1, \left(p_1^q, p_2^q, \cdots, p_C^q\right)\right) \text{ if } \mathbf{Y}_{[i,q]} \text{ is an ordinal response.}$$

Here the t distribution $t_\nu\left(\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}, \sigma_{iq}^2\right)$ is characterized by three parameters, center $\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}$, scale σ_{iq}, and a degrees of freedom parameters ν that determines the shape of the distribution, and $p_j^q = \Pr\left(\mathbf{Y}_{[i,q]} = j\right)$ is the probability that the ith response for the qth outcome variable is the jth category with $j = 1, 2, \ldots, C$ for $i = 1, \ldots, N$, $q = 1, \ldots, Q$ and $\sum_{j=1}^{C} p_j^q = 1$.

For the ordinal response, we consider latent variables with a continuous distributional assumption and cutpoints of the continuous scale for the latent variable. As in Albert and Chib (1993) with the underlying latent observation $\mathbf{Z}_{[i,q]}$ such that

$$\begin{aligned}\Pr\left(\mathbf{Y}_{[i,q]} = j\right) &= \Pr\left(\gamma_{j-1}^q \le \mathbf{Z}_{[i,q]} \le \gamma_j^q\right)\\ &= \Psi\left(\frac{\gamma_j^q - \mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}}{\sigma_{iq}}\right) - \Psi\left(\frac{\gamma_{j-1}^q - \mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}}{\sigma_{iq}}\right),\end{aligned}$$

where $\mathbf{Z}_{[i,q]} \sim L\left(\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}, \sigma_{iq}\right)$, $\Psi(\cdot)$ is the cumulative distribution function of the standard logistic distribution, center $\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}$, scale σ_{iq}, and cutpoints $-\infty = \gamma_0^q < \gamma_1^q < \cdots < \gamma_C^q = \infty$.

Gelman et al. (2014) argued that the t distribution can be considered alternative to logistic and probit regression. Logistic and probit regressions can be nonrobust in the sense that for large absolute values of the linear predictors $\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}$, the inverse logit or probit transformations give probabilities close to 0 or 1. Such models could be made more robust by allowing the occasional misprediction for large values of $\mathbf{X}_{[i,]}\boldsymbol{\beta}_{[,q]}$. This form of robustness is defined not in terms of the data $\mathbf{Y}_{[i,q]}$ but with respect to the predictors $\mathbf{X}_{[i,]}$. A robust model can be implemented using the latent-variable formulation of discrete-data regression models which is described above, replacing the logistic or normal distribution of the latent continuous data $\mathbf{Z}_{[i,q]}$ with the model, $\boldsymbol{Z}_{[i,q]} \sim t_\nu\left(\boldsymbol{X}_{[i,]}\boldsymbol{\beta}_{[,q]}, \sigma_{iq}\right)$. Gelman et al. (2014) argued that in realistic settings it is impractical to estimate ν from the data, since the latent data are not directly observed, it is essentially impossible to form inference about the shape of their continuos underlying distribution, so it is set as a low value to ensure robustness. Setting $\nu = 4$ yields a distribution that is close to the logistic, and as $\nu \to \infty$, the model approaches the probit.

In the Bayesian inference, Gibbs sampler computations can often be simplified or convergence accelerated by adding auxiliary variables, and it is called data augmentation. One of the simple but important example of auxiliary variables is the t distribution which can be expressed as a mixture of normal distributions. The $t_\nu\left(\mu, \sigma^2\right)$ likelihood is equivalent to the model

$$\begin{aligned}Y &\sim N\left(\mu, \sigma^2 V\right)\\ V^{-1} &\sim Gamma\left(\frac{\nu}{2}, \frac{\nu}{2}\right)\end{aligned}$$

where the V is auxiliary variable that cannot be directly observed.

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

2.2.2. Multivariate model structure with ERA mean

For the mean structure, we consider ERA with predefined components but unknown weights on predictors for each of components. ERA considers multiple sets (or blocks) of predictors and reduces each set into a component. Such formation is based on some substantive theories or domain knowledge about how certain predictors can be grouped into the same block and aggregated into a component. Each observed variable is hypothesized to be linked to only one component. Accordingly, some elements in $\mathbf{W}$ will be constrained to be zero: more specifically, each row of $\mathbf{W}$ has only one non-zero element (whose component weight will be freely estimated), and the remaining elements in each row will be constrained to zero. Following ERA model specification, the ith component $\mathbf{F}_{[i,]}$ is $\mathbf{F}_{[i,]} = \mathbf{X}_{[i,]}\mathbf{W}$ for $i = 1,\cdots,N$, and hence regression coefficients for the qth outcome variable $\mathbf{A}_{1[,q]}$ for $q = 1,\cdots,Q$ are unidentifiable, so $diag\left(\mathbf{F}'\mathbf{F}\right) = diag\left(\mathbf{W}'\mathbf{X}'\mathbf{X}\mathbf{W}\right) = N\mathbf{I}$ with the $N \times K$ component matrix $\mathbf{F}' = \left(\mathbf{F}'_{[1,]},\cdots,\mathbf{F}'_{[N,]}\right)$ is applied as a standardization contraint for identifiablity (Takane and Hwang, 2005).

For the multivariate logistic distribution, O'Brien and Dunson (2004) proposed a Bayesian models with consideing a multivariate t distributed latent variables. The proposed multivariate logistic dstirbution results in the close approximation of the two densities, a multivariate logistic distribution of v degrees of freedom with center $\boldsymbol{\mu}$ and scale matrix $\mathbf{R}$ and a multivariate t distribution of v degrees of freedom with center $\boldsymbol{\mu}$ and scale matrix $\sigma^2\mathbf{R}$ when v and σ^2 are chosen appropriately. To make the approximation almost exact, O'Brien and Dunson (2004) set $\sigma^2 = \pi^2\left(v-2\right)/3v$ (a value chosen to make the variances of the univariate t and logistic distributions equail) and set $v = 7.3$ (a value chosen to minimize the integrated squared dsitnace between the univariate t and univariate logistic densities).

Park et al. (2021) introduced a Bayesian methodology for a component-based model that accounts for unstructured residual covariances, while regressing multivariate ordinal outcomes on pre-defined sets of predictors. The proposed Bayesian multivariate ordinal logistic model re-expresses ordinal outcomes of interest with a set of latent continuous variables based on an approximate multivariate t distribution based on the method of O'Brien and Dunson (2004). This contributes not only to developing an efficient Gibbs sampler, a Markov Chain Monte Carlo algorithm, but also to facilitating the interpretation of regression coefficients as log-transformed odds ratio.

In this work, we consider multivariate t distribution for the residual error on continuous outcomes and multivariate t approximation of the latent continuous variables for the ordinal outcomes. We also need to consider the interdependency among outcome variables in the model regardless of the response structure to avoid biased statistical inference. Thus, with an ERA model of K components, for $i = 1,\ldots,N$ and $j = 1,\ldots,C$

$$\mathbf{Y}_{[i,q]} = \mathbf{Z}_{[i,q]} \text{ for a continuous response } (q = 1,\cdots,T)$$

$$\mathbf{Y}_{[i,q]} = j \quad \text{if } \gamma^q_{j-1} \le \mathbf{Z}_{[i,q]} \le \gamma^q_j \text{ for an ordianal response } (q = T+1,\cdots,Q),$$

we express our model as following:

$$\begin{aligned}
&\mathbf{Z}'_{[i,]} \mid \mathbf{X}_{[i,]},\mathbf{B}_1,\mathbf{H},\mathbf{W},\mathbf{A}_1,\boldsymbol{\Sigma}_1,\mathbf{a}_1,\phi_{1i},\xi_{1i} \\
&\qquad \sim N_Q\left(\mathbf{a}_1 + \mathbf{B}'_1\mathbf{H}'_{[i,]} + \mathbf{A}'_1\mathbf{W}'\mathbf{X}'_{[i,]}, \mathbf{D}^{1/2}_{1i}\boldsymbol{\Sigma}_1\mathbf{D}^{1/2}_{1i}\right) \\
&\xi_{1i} \mid \mathbf{X}_{[i,]},\mathbf{B}_1,\mathbf{H},\mathbf{W},\mathbf{A}_1,\boldsymbol{\Sigma}_1,\mathbf{a}_1,\phi_{1i},\mathbf{Z}_{[i,]} \sim Gamma\left(\frac{v^*_1}{2},\frac{v^*_1}{2}\right) \\
&\phi_{1i} \mid \mathbf{X}_{[i,]},\mathbf{B}_1,\mathbf{H},\mathbf{W},\mathbf{A}_1,\boldsymbol{\Sigma}_1,\mathbf{a}_1,\xi_{1i},\mathbf{Z}_{[i,]} \sim Gamma\left(\frac{\tilde{v}}{2},\frac{\tilde{v}}{2}\right),
\end{aligned} \tag{2}$$

where $\mathbf{a}_1$ is a length Q vector of intercepts, $\mathbf{H}_{[i,]}$ is a $1 \times D$ vector of regression predictors, $\mathbf{B}_1$ is a $D \times Q$ matrix of regression parameters, $\mathbf{X}_{[i,]}$ is a $1 \times Q$ vector of ERA, $\mathbf{W}$ is a P by K matrix of weights, $\mathbf{A}_1$ is a $K \times Q$ matrix of component coefficients, $\boldsymbol{D}_{1i} = diag\left(\xi^{-1}_{1i}\mathbf{1}_T, \tilde{\sigma}^2\phi^{-1}_{1i}\mathbf{1}_{Q-T}\right)$ with length T vector of 1's $\mathbf{1}_T\mathbf{1}_T$ and length $Q-T$ vector of 1's $\mathbf{1}_{Q-T}$, and $\boldsymbol{\Sigma}_1$ is a $Q \times Q$ unstructured variance-covariance matrix. Here ξ^{-1}_{1i} and ϕ^{-1}_{1i} are precision parameters to form multivariate t distribution for the T continuous

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

responses and $Q-T$ ordinal responses, respectively. As discussed in O'Brien and Dunson (2004), we set $\tilde{v}=7.3$ and $\tilde{\sigma}^2=\pi^2\left(\tilde{v}-1\right)/3\tilde{v}$ to make the multivariate t distribution approximate the multivariate logistic regression. Because a degrees of freedom parameters v_1^* determines the shape of the distribution for continuous outcomes, we treat v_1^* as an unknown parameter which need to be estimated.

2.3. The proposed method with mediating effects

A conventional single-level mediation model is expressed as follows:

$$M=\beta_1+\alpha X+e_1$$
$$Y=\beta_2+\beta M+\gamma X+e_2,$$

where M, Y, and X denote a mediator, an outcome variable, and a predictor, respectively, with M being unobservable while Y and X are observable. Miočević et al. (2017) included an interaction term (XM) in their model predicting Y

$$Y=\beta_3+\beta M+\gamma X+\theta XM+e_3,$$

to test if the moderation effect is statistically significant based on Bayesian inference. They considered diffuse (non-informative) priors for the coefficients in their empirical example. In the Markov Chain Monte Carlo Estimation section, they used trace plots to evaluate whether the chains were mixing well, which is indicative of convergence, and provided indexes for diagnosing convergence. Even with the interaction term, there was no issue with the identifiability of model coefficient estimates.

For the indirect or mediation effect of intervening impacts on outcomes, we consider S contemporaneous mediators. For S mediators, we consider a multivariate t distribution with correlation structure on error term and a component-based ERA structure for the mean trend. As we described above, we express a multivariate t distribution with the degrees of freedom v_2^* as the hierarchical form of a normal mixture such that

$$\begin{aligned}&\mathbf{M}'_{[i,]}\mid\mathbf{X}_{[i,]},\mathbf{B}_2,\mathbf{H},\mathbf{W},\mathbf{A}_2,\mathbf{\Sigma}_2,\mathbf{a}_2,\xi_{2i}\sim N_S\left(\mathbf{a}_2+\mathbf{B}_2'\mathbf{H}'_{[i,]}+\mathbf{A}_2'\mathbf{W}'\mathbf{X}'_{[i,]},\quad \xi_{2i}^{-1}\mathbf{\Sigma}_2\right)\\&\xi_{2i}\mid\mathbf{X}_{[i,]},\mathbf{B}_2,\mathbf{H},\mathbf{W},\mathbf{A}_2,\mathbf{M}_{[i,]},\mathbf{\Sigma}_2,\mathbf{a}_2\sim Gamma\left(\frac{v_2^*}{2},\quad\frac{v_2^*}{2}\right)\end{aligned}\tag{3}$$

where $\mathbf{a}_2$ is a length S vector of intercepts, $\mathbf{B}_2$ is a D by S matrix of regression parameters, $\mathbf{A}_2$ is a K by S matrix of component coefficients, and $\mathbf{\Sigma}_2$ is a S by S unstructured variance-covariance matrix. We treat the degrees of freedom v_2^* as an unknown parameter for the shape of the distribution.

Similar to the response model in (2), with the T continuous responses and $Q-T$ ordinal responses, the multivariate mixed outcomes model with mediators can be expressed as

$$\begin{aligned}&\mathbf{Z}'_{[i,]}\mid\mathbf{X}_{[i,]},\mathbf{B}_3,\mathbf{H},\mathbf{W},\mathbf{A}_3,\mathbf{M},\mathbf{A}_4,\mathbf{\Sigma}_3,\mathbf{a}_3,\phi_{3i},\xi_{3i}\\&\sim N_Q\left(\mathbf{a}_3+\mathbf{B}_3'\mathbf{H}'_{[i,]}+\mathbf{A}_3'\mathbf{W}'\mathbf{X}'_{[i,]}+\mathbf{A}_4'\mathbf{M}'_{[i,]},\quad \mathbf{D}_{3i}^{1/2}\mathbf{\Sigma}_3\mathbf{D}_{3i}^{1/2}\right)\\&\xi_{3i}\mid\mathbf{X}_{[i,]},\mathbf{B}_3,\mathbf{H},\mathbf{W},\mathbf{A}_3,\mathbf{M},\mathbf{A}_4,\mathbf{\Sigma}_3,\mathbf{a}_3,\phi_{3i},\mathbf{Z}_{[i,]}\sim Gamma\left(\frac{v_3^*}{2},\quad\frac{v_3^*}{2}\right)\\&\phi_{3i}\mid\mathbf{X}_{[i,]},\mathbf{B}_3,\mathbf{H},\mathbf{W},\mathbf{A}_3,\mathbf{M},\mathbf{A}_4,\mathbf{\Sigma}_3,\mathbf{a}_3,\xi_{3i},\mathbf{Z}_{[i,]}\sim Gamma\left(\frac{\tilde{v}}{2},\quad\frac{\tilde{v}}{2}\right),\end{aligned}\tag{4}$$

for $i=1,\ldots,N$ and $j=1,\ldots,C$. Here $\mathbf{a}_3$ is a length Q vector of intercepts, $\mathbf{B}_3$ is a D by Q matrix of regression parameters, $\mathbf{A}_3$ is a K by Q matrix of component coefficients, $\mathbf{A}_4$ is a S by Q matrix of coefficients relating the mediators to the dependent variables adjusted for the independent variables, $\boldsymbol{D}_{3i}=diag\left(\xi_{3i}^{-1}\mathbf{1}_T,\quad \tilde{\sigma}^2\phi_{3i}^{-1}\mathbf{1}_{Q-T}\right)$ with length T vector of 1's $\mathbf{1}_T$ and length $Q-T$ vector of 1's $\mathbf{1}_{Q-T}$, $\mathbf{\Sigma}_3$ is a Q by Q unstructured variance-covariance matrix and ξ_{3i}^{-1} and ϕ_{3i}^{-1} are precision parameters to form multivariate t distribution. As we discussed above, $\tilde{v}=7.3$ and $\tilde{\sigma}^2=\pi^2\left(\tilde{v}-1\right)/3\tilde{v}$ for the approximation

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

to multivariate logistic regression and we treat v_3^* as an unknown parameter for the shape of the distribution for continuous outcomes.

In our model (2)–(4), the pre-determined components $\mathbf{F} = \mathbf{XW}$ are considered to explain the relationship between predictors and multivariate outcomes (2), to extent to which components changes the mediators (3), and to explain which components are related to the multivariate outcomes adjusted for the mediation (4). Thus the mediated effects can be calculated in two ways as either $\mathbf{A}_1 - \mathbf{A}_3$ or $\mathbf{A}_2\mathbf{A}_4$. The value of the indirect effect can be estimated by taking the difference in the coefficients $\mathbf{A}_1 - \mathbf{A}_3$ from (2) and (3) corresponds to the reduction in the independent variables effect on the dependence variables when adjusted for the mediator. The product of coefficients method is based on the estimation of (3) and (4) to form the mediated or indirect effect. It can be interpreted as that mediation depends on the extent to which the components change the mediators $\mathbf{A}_2$ and the extent to which the mediator affects the outcome variables adjusted for the predictors $\mathbf{A}_4$. For the Bayesian inference with rich explanation, we use (3) and (4) models in our work.

2.4. Parameter estimation

2.4.1. Prior distributions

In the proposed method, with the likelihood of multivariate t distribution formation of the mediating effects in (3) and of the response with mediators in (4), we consider the conjugate distributions for specifying priors on the parameters of interest. For the prior on the weight $\mathbf{W}$ and the component coefficients $\mathbf{A}$, we consider the large valued hyper-parameters of the covariance. For a more flexible Bayesian model, we might be able to consider a prior distribution on the hyper-parameters such as an inverse gamma prios distribution for the conjugacy. However, in our proposed method, to mitigate the effects fo priors, we consider flat priors with large valued hyper-parameters and because of the conjugacy, the convergence of chains could be gauranteed. Note that specifying a prior distribution for the unconstrained covariance matrix $\mathbf{\Sigma}$ makes our approach different from O'Brien and Dunson (2004), in which unique off-diagonal elements of a correlation matrix are assumed to follow a prior distribution of no specific form. This simple modification leads us more efficient but easier posterior computation, as detailed in Park et al. (2021) for the multivariate ordinal responses. Details about the prior distributions can be found in Appendix A1.

For the prior on v_2^* and v_3^*, degrees of freedom parameters of the distribution for mediators and continuous outcomes, we try a uniform density on $1/\nu^*$ for the range $[0,1]$. Gelman et al. (2014) explained that the parameterization in terms of $1/\nu^*$ rather than ν^* has the advantage of including the normal distribution at $1/\nu^* = 0$ and encompassing the entire range from normal to Cauchy distribution in the finite interval $[0,1]$. This prior distribution favors longer-tailed models. We cannot parameterize the t_ν distribution in terms of their variance, because the variance is infinite for $\nu \leq 2$. Instead the interquartile range would be a more reasonable parameter than the curvature for setting up a prior distribution. The interquartile range varies mildly as a function of ν, and we consider the convenient parameterization in terms of mean and variance, and set $g\left(\nu^*\right) \propto 1$.

2.4.2. Markov chain Monte Carlo

When employing conjugate priors for model parameters, the majority of the full conditional posterior distributions can be derived in closed form. However, closed-form solutions are not available for the posterior distributions of precision parameters and degrees of freedom parameters. Therefore, we update these parameters using a Metropolis step within the MCMC iterations. For the other parameters, Gibbs sampling is a straightforward and easily implementable method. A code programmed in software R (version 4.3.1, R Core Team, 2023) is available on GitHub.

To update parameters from the posterior distributions, we choose initial values of parameters from the prior distributions and update parameters based on the full conditional distributions. First, we update the common weight parameter $\mathbf{W}$ in mediator and response mean terms and standardize it for

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

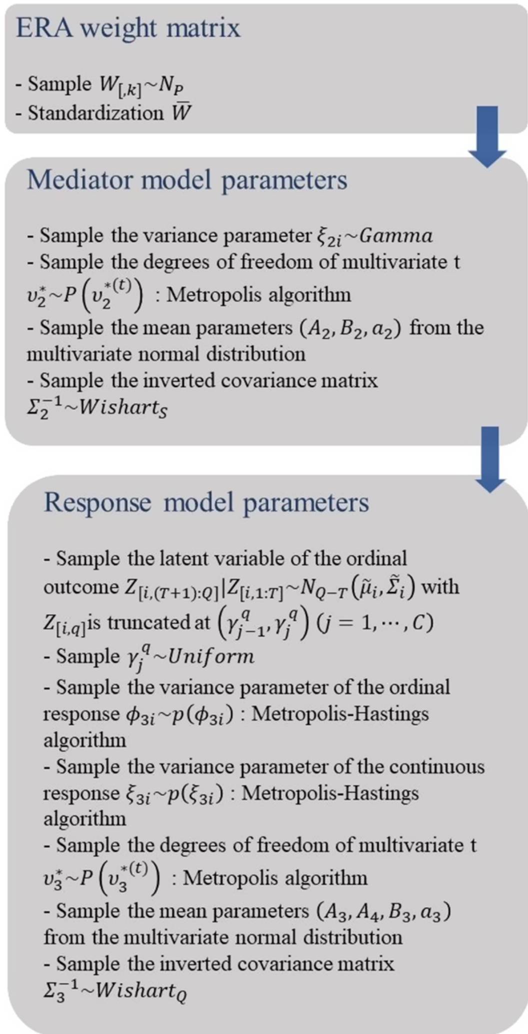

Figure 1. MCMC algorithm to update parameters from the posterior distribution.

the identifiability. Next, we update the parameters of the mediator model $\left(\boldsymbol{\xi}_2, \nu_2^*, \mathbf{A}_2, \mathbf{B}_2, \mathbf{a}_2, \boldsymbol{\Sigma}_2^{-1}\right)$ simultaneously and update the latent variable of the ordinal outcomes from the truncated conditional normal distribution. Finally, we update the other parameters in response model $\left(\boldsymbol{\gamma}, \boldsymbol{\phi}_3, \boldsymbol{\xi}_3, \nu_3^*, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_3, \mathbf{a}_3, \boldsymbol{\Sigma}_3^{-1}\right)$ from the conditional posterior distribution of each. We repeat the described steps and the process of the sampling schemes is in Figure 1. Also, the full conditional posterior distributions of parameters and the detailed sampling schemes are in Appendix A2.

There are a few comments to make regarding the proposed algorithm. First, to ensure the identifiability of $\mathbf{F} = \mathbf{XW}$, a standardization constraint is applied to $\mathbf{F}$ such that $diag\left(\mathbf{F}'\mathbf{F}\right) = diag\left(\mathbf{W}'\mathbf{X}'\mathbf{XW}\right) = N\mathbf{I}$, i.e., a procedure inherent in any ERA model. Second, sampling $\boldsymbol{\Sigma}^{-1}$ from a Wishart distribution can be easily done in most statistical softwares, for example, with function *rWishart* from the stats library in R. Furthermore, noting that scaled coefficients with $diag(\boldsymbol{\Sigma})$ being the diagonal variance matrix corresponding to $\boldsymbol{\Sigma}$ are identifiable and correlations from $\boldsymbol{\Sigma}$ are constrained to be $[-1, 1]$, our proposed algorithm employs the parameter expansion of Gelman et al. (2004), which proves improved convergence of Gibbs sampler in generalized linear models. In contrast, O'Brien and Dunson (2004)

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

draw unique correlations corresponding to $\boldsymbol{\Sigma}$ instead using a Metropolis algorithm (Metropolis & Ulam, 1949), for which it is difficult to find a good value of a tuning parameter to control an acceptance probability, especially in a case of sampling multiple variates, often resulting in slow convergence of the algorithm. Third, there is still an additional identifiability problem in intercept $\mathbf{a}_{3[(T+1):Q]}$, cutpoints $\boldsymbol{\Gamma}$, and $diag\left(\tilde{\boldsymbol{\Sigma}}_3\right)$. As Hirk et al. (2019) pointed out, $\frac{\gamma_j^q - a_{3,q}}{\sigma_{qq}}$ with $a_{3,q}$ being the qth element of $\mathbf{a}_3$ and σ_{qq} being the qth diagonal element of $\tilde{\boldsymbol{\Sigma}}_3$, $j = 1,\ldots,C-1$ and $q = T+1,\ldots,Q$, is only identifiable, and we set for all $q = T+1,\ldots,Q$ to secure the identifiability of scaled intercepts $\frac{a_{3,q}}{\sigma_{qq}}$ and scaled cutpoints $\frac{\gamma_j^q}{\sigma_{qq}}$. Such a constraint of $\gamma_1^q = 0$ is a common practice in Bayesian logistic regression models for both univariate and multivariate binary outcomes, which is a special case of the proposed multivariate ordinal logistic regression model when $C = 2$. The fixing of the cutpoint at zero in the process of MCMC is to prevent the identifiability problem of the parameter estimation.

3. Simulation study

To validate the performance of the proposed method, we conducted simulation studies varying sample sizes of N = 100, 300, and 500 for a model with two continuous ($T = 2$) and two ordinal outcome variables ($Q - T = 2$) using (3) and (4). The hypothesized model had two mediators, and there were no covariate or explanatory variables affecting mediators and outcome variables, as shown in Figure 2. For data generation, we used

$$\begin{aligned} \mathbf{M}'_{[i,]} &= \mathbf{a}_{02} + \mathbf{A}'_2\mathbf{W}'\mathbf{X}'_{[i,]} + \mathbf{E}'_{2[i,]} \\ \mathbf{Z}'_{[i,]} &= \mathbf{a}_{03} + \mathbf{A}'_3\mathbf{W}'\mathbf{X}'_{[i,]} + \mathbf{A}'_4\mathbf{M}'_{[i,]} + \mathbf{E}'_{3[i,]} \end{aligned} \tag{5}$$

where $\mathbf{E}'_{2[i,]} \sim N_S\left(\mathbf{0}, \xi_{2i}^{-1}\boldsymbol{\Sigma}_2\right)$ with $\xi_{2i} \sim Gamma(1.5, 1.5)$ and $\mathbf{E}'_{3[i,]} \sim N_Q\left(\mathbf{0}, \mathbf{D}_{3i}^{1/2}\boldsymbol{\Sigma}_3\mathbf{D}_{3i}^{1/2}\right)$ with $\mathbf{D}_{3i} = diag\left(\xi_{3i}^{-1}\mathbf{1}_T, \tilde{\sigma}^2\phi_{3i}^{-1}\mathbf{1}_{Q-T}\right)$, $\xi_{3i} \sim Gamma(2.5, 2.5)$, $\tilde{\sigma}^2 = \pi^2(7.3-2)/(3\times 7.3)$, and $\phi_{3i} \sim$

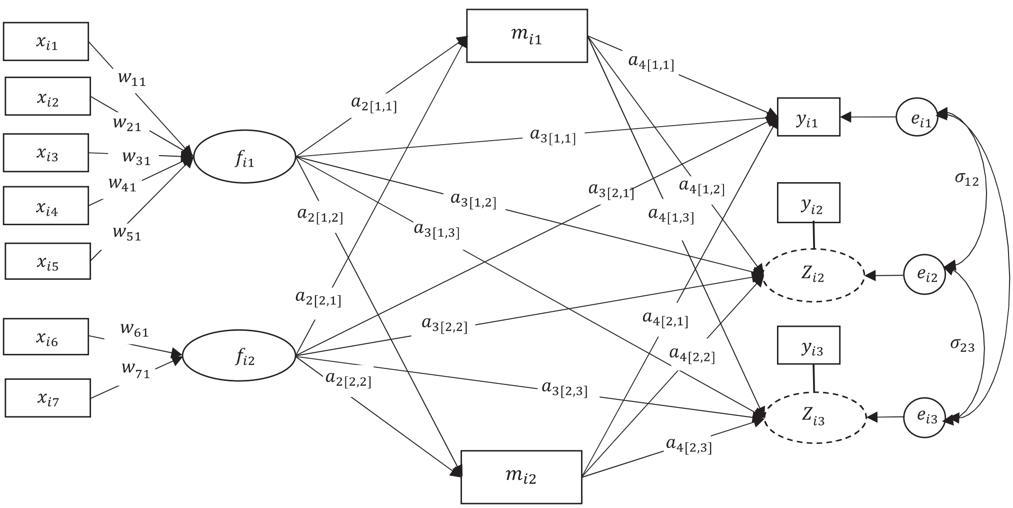

Figure 2. A hypothesized model for the first simulation study (with one continuous and two ordinal outcome variables).

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

Gamma$(7.3/2, 7.3/2)$. In (5), the true parameter values were set at $\mathbf{a}_{02} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \\ w_{41} & w_{42} \\ w_{51} & w_{52} \\ w_{61} & w_{62} \\ w_{71} & w_{72} \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.5 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix}$, $\mathbf{A}_2 = \begin{bmatrix} a_{2[1,1]} & a_{2[1,2]} \\ a_{2[2,1]} & a_{2[2,1]} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $\mathbf{a}_{03} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{A}_3 = \begin{bmatrix} a_{3[1,1]} & a_{3[1,2]} & a_{3[1,3]} & a_{3[1,4]} \\ a_{3[2,1]} & a_{3[2,2]} & a_{3[2,3]} & a_{3[2,4]} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1.5 & 1.5 & 0 \end{bmatrix}$, and $\mathbf{A}_4 = \begin{bmatrix} a_{4[1,1]} & a_{4[1,2]} & a_{4[1,3]} & a_{4[1,4]} \\ a_{4[2,1]} & a_{4[2,2]} & a_{4[2,3]} & a_{4[2,4]} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0.5 & 0 \\ 2 & 1 & 1.5 & 2 \end{bmatrix}$ with the covariance matrices $\mathbf{\Sigma}_2 = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$ and $\mathbf{\Sigma}_3 = \begin{bmatrix} 0.8 & 0.3 & 0.3 & 0.3 \\ 0.3 & 0.8 & 0.3 & 0.3 \\ 0.3 & 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 0.3 & 1 \end{bmatrix}$. In addition, $\mathbf{X}_{[i,]}$ was sampled from a multivariate normal distribution $N_p(\mathbf{0}, \mathbf{\Omega})$ with $\mathbf{\Omega} = \begin{bmatrix} 1 & 0.3 & 0.1 & 0.1 \\ 0.3 & 1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 1 & 0.6 \\ 0.1 & 0.1 & 0.6 & 1 \end{bmatrix}$. Here $P = 7$, $K = 2$, $S = 2$, $T = 2$, and $Q = 4$.

Note that the first two element of $\mathbf{E}'_{3[i,]}$ was sampled from a multivariate t distribution with mean 0 and scale matrix of $\mathbf{\Sigma}_{3[1:2,1:2]}$ after integrating over ξ_{3i}, and the next two elements were sampled from a multivariate t distribution with mean 0 and scale matrix of $\tilde{\sigma}^2 \mathbf{\Sigma}_{3[3:4,3:4]}$ after integrating over ϕ_{3i}, while these three elements were correlated. A multivariate logistic random variables underlying the two ordinal outcomes were then generated by transforming each of the multivariate t distributed random variables of $\mathbf{E}_{3[i,3:4]}$ using the following formula proposed by O'Brien and Dunsion (2004),

$$\mathbf{Z}^*_{[i,q]} = \mu_q + \log\left[\frac{F_v\left(\mathbf{E}_{3[i,j]}/\tilde{\sigma}\right)}{1 - F_v\left(\mathbf{E}_{3[i,j]}/\tilde{\sigma}\right)}\right] \quad (\text{for } q = 3,4)$$

where $\mu_q = \mathbf{a}_{03[q]} + \mathbf{A}'_{3[,q]}\mathbf{W}'\mathbf{X}'_{[i,]} + \mathbf{A}'_{4[,q]}\mathbf{M}'_{[i,]}$ and $F_v(\cdot)$ is the cumulative Student's t distribution with v degrees of freedom. The transformation from multivariate t variables to multivariate logistic variables is derived in detail in the supplementary information of Kyung et al. (2021). Once $\mathbf{Z}^*_{[i,q]}$ for $q = 3,4$, was generated, two sets of cutoffs were used to obtain the ordinal variables; $\mathbf{\Gamma}_{[,3]} = (0,3,6)'$and $\mathbf{\Gamma}_{[,4]} = (0,2,4)'$.

For the proposed method, the hyperparameters for diffuse prior distributions were set with all the prior mean parameters such as $\mathbf{W}^0_{[,k]}$ being zero, $\tau^2\mathbf{\Sigma}^0_{\mathbf{W}_k} = 10\mathbf{I}_7$, $\eta_2^2\mathbf{\Sigma}^0_{\mathbf{A}_{2k}} = 100\mathbf{I}_2$, $\eta_3^2\mathbf{\Sigma}^0_{\mathbf{A}_{3k}} = \eta_4^2\mathbf{\Sigma}^0_{\mathbf{A}_{4s}} = 100\mathbf{I}_4$, $\sigma_2^2\mathbf{\Sigma}^0_{\mathbf{a}_{02}} = 100\mathbf{I}_2$, $\sigma_3^2\mathbf{\Sigma}^0_{\mathbf{a}_{03}} = 100\mathbf{I}_4$, $v_2 = v_3 = 0.01$, $\mathbf{\Sigma}^{-1}_{20} = 100\mathbf{I}_2$, $\mathbf{\Sigma}^{-1}_{30} = 100\mathbf{I}_4$, and the total number of iterations were set at 30,000. Note that these hyperparameters' setting and the total number of iterations remained the same for all of the analyses conducted in this manuscript (also same for the empirical dataset presented in the next section). Fast convergence and good mixing of the MCMC chain were observed for all parameters of interest across all the simulation scenarios considered. The first 2,000 iterations were discarded as a burn-in period and every fifth posterior sample was used by applying a thinning approach to calculate the posterior means and credible intervals (CI) for the parameters of interest. Relevant simulation code written in R (R Core Team, 2023) is available on GitHub. Fast convergence and

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

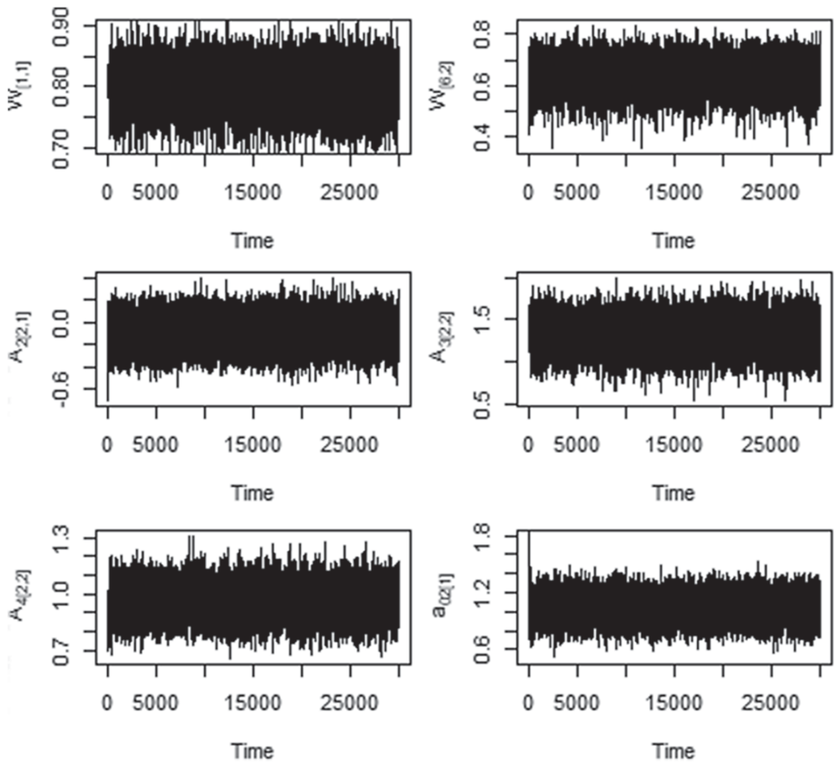

Figure 3. Trace plots of selected parameters for the simulation study with $N = 100$.

good mixing of the MCMC chain were observed for all parameters of interest across all the simulation scenarios considered. As an example, trace plots are presented in Figure 3.

Table 1 presents the results of analyzing the simulated data with two continuous and two ordinal outcome variables varying sample sizes. With increased sample size, posterior mean estimates obtained from the proposed method became on average closer to the true parameter values that were well embedded by narrower corresponding CIs. The residual variances and covariances for the two mediators and four outcome variables ($\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_3$) exhibited similar patterns: with increased sample sizes, their estimates became closer to the prescribed true values with more precision (resulting in narrower CIs). This indicated that the proposed method is capable of taking into account the dependency among the outcome variables, recovering the true parameter values well.

Table 2 displays the results for indirect effects, whose true values were calculated based on the product term of the prescribed true values for $\mathbf{A}_2$ and $\mathbf{A}_4$, i.e. $\mathbf{A}_2\mathbf{A}_4$. For instance, the true value of an indirect effect from F_2 on Y_1 via M_1 was calculated by multiplying the true values of the effect of F_2 on M_1 and that of M_1 on Y_1 (i.e., $a_{2[2,1]} \times a_{4[1,1]}$). At each iteration of MCMC, the obtained posterior samples on $\mathbf{A}_2$ and $\mathbf{A}_4$, were used to get the posterior samples for the indirect effects. Based on those posterior samples, we computed posterior means and CIs for the indirect effects. As shown in Table 2, the posterior means of all indirect effect estimates were close to the prescribed true values, and this pattern became salient with narrower CIs as the sample size increased.

For comparison, using the same generated datasets from this simulation study, we fitted the same model but excluding the two mediators, that is corresponding to the model in Equation (2) without **H**. Table 3 displays the results without the mediators. Although there were no true values assigned to the direct effects between predictors and outcome variables because the simulation setting was generated with multiple mediators, each of the direct effects could be indirectly inferred from the total effect (i.e., addition of indirect and direct effects) obtained from a model with the mediators. Their direct effects,

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

Table 1. Results of the simulation study varying sample sizes (N = 100, 300, 500)

Parameters		Truth	N = 100			N = 300			N =500		
			Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up
W	w_{11}	0.9	0.92	0.85	0.99	0.95	0.90	1.00	0.90	0.87	0.94
	w_{21}	0.5	0.48	0.37	0.59	0.51	0.44	0.58	0.54	0.50	0.59
	w_{31}	0	−0.01	−0.13	0.09	−0.07	−0.16	0.01	−0.01	−0.07	0.04
	w_{41}	0	0.00	−0.12	0.11	0.03	−0.05	0.11	0.02	−0.03	0.07
	w_{51}	0	0.05	−0.07	0.17	0.01	−0.07	0.09	−0.03	−0.09	0.03
	w_{62}	0.5	0.52	0.43	0.60	0.49	0.43	0.55	0.46	0.42	0.51
	w_{72}	0.5	0.46	0.37	0.55	0.49	0.43	0.55	0.53	0.49	0.57
$\mathbf{A}_2$	$a_{2[1,1]}$	2	2.06	1.85	2.27	1.95	1.84	2.07	1.94	1.86	2.03
	$a_{2[2,1]}$	0	−0.11	−0.41	0.19	0.03	−0.12	0.19	0.00	−0.12	0.12
	$a_{2[1,2]}$	0	0.09	−0.16	0.34	−0.06	−0.17	0.06	−0.08	−0.17	0.00
	$a_{2[2,2]}$	2	2.06	1.73	2.38	1.97	1.81	2.13	2.01	1.87	2.13
$\mathbf{A}_3$	$a_{3[1,1]}$	1	0.95	0.53	1.35	0.92	0.72	1.10	0.94	0.78	1.09
	$a_{3[2,1]}$	0	0.00	−0.39	0.40	−0.10	−0.32	0.11	−0.06	−0.23	0.11
	$a_{3[1,2]}$	2	2.02	1.59	2.43	1.92	1.72	2.11	2.07	1.92	2.22
	$a_{3[2,2]}$	1.5	1.63	1.19	2.05	1.42	1.19	1.65	1.43	1.26	1.60
	$a_{3[1,3]}$	1	0.84	−0.17	1.95	1.08	0.63	1.54	0.72	0.34	1.09
	$a_{3[2,3]}$	1.5	1.63	0.46	2.93	1.04	0.53	1.56	1.28	0.87	1.69
	$a_{3[1,4]}$	2	2.93	1.81	4.28	1.59	1.07	2.16	2.03	1.60	2.50
	$a_{3[2,4]}$	0	−0.28	−1.40	0.83	0.26	−0.35	0.86	0.14	−0.27	0.55
$\mathbf{A}_4$	$a_{4[1,1]}$	0	−0.02	−0.20	0.16	0.01	−0.07	0.09	0.00	−0.07	0.07
	$a_{4[2,1]}$	2	2.02	1.88	2.16	2.01	1.92	2.10	2.00	1.93	2.06
	$a_{4[1,2]}$	1	0.93	0.75	1.12	0.98	0.90	1.06	0.96	0.89	1.03
	$a_{4[2,2]}$	1	0.96	0.81	1.11	1.04	0.95	1.13	1.04	0.98	1.11

(Continued)

Table 1. (Continued)

Parameters		Truth	*N* = 100			*N* = 300			*N* =500		
			Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up
	$a_{4[1,3]}$	0.5	0.83	0.36	1.36	0.37	0.17	0.59	0.50	0.33	0.67
	$a_{4[2,3]}$	1.5	1.96	1.20	2.82	1.58	1.27	1.94	1.37	1.14	1.62
	$a_{4[1,4]}$	0	−0.56	−0.99	−0.16	0.13	−0.07	0.34	−0.05	−0.22	0.11
	$a_{4[2,4]}$	2	2.00	1.31	2.82	1.97	1.59	2.39	1.94	1.65	2.25
$\mathbf{a}_{02}$	$\mathrm{a}_{02[1]}$	1	0.84	0.60	1.08	1.01	0.88	1.14	0.99	0.90	1.09
	$\mathrm{a}_{02[2]}$	1	0.92	0.64	1.20	1.16	1.04	1.29	1.01	0.91	1.12
$\mathbf{a}_{03}$	$\mathrm{a}_{03[1]}$	1	1.04	0.80	1.28	0.90	0.73	1.06	0.95	0.82	1.08
	$\mathrm{a}_{03[2]}$	0	0.14	−0.11	0.39	0.04	−0.12	0.21	−0.06	−0.17	0.06
	$\mathrm{a}_{03[3]}$	2	2.89	1.77	4.18	1.45	0.99	1.94	1.80	1.40	2.19
	$\mathrm{a}_{03[4]}$	−1	−0.34	−1.28	0.57	−1.04	−1.61	−0.49	−0.57	−0.99	−0.16
$\mathbf{\Sigma}_2$	σ_{11}	1	1.14	0.73	1.68	0.97	0.76	1.22	0.98	0.81	1.16
	σ_{12}	0.3	0.35	0.05	0.72	0.31	0.17	0.46	0.25	0.14	0.36
	σ_{22}	1	1.52	0.99	2.24	0.95	0.74	1.20	0.99	0.82	1.17
$\mathbf{\Sigma}_3$	σ_{11}	0.8	0.91	0.58	1.34	1.04	0.83	1.29	0.98	0.84	1.15
	σ_{12}	0.3	0.27	0.04	0.55	0.39	0.25	0.56	0.39	0.29	0.50
	σ_{13}	0.3	−0.01	−0.33	0.30	0.32	0.18	0.48	0.21	0.08	0.33
	σ_{22}	0.3	0.22	−0.06	0.50	0.09	−0.09	0.26	0.29	0.16	0.41
	σ_{23}	0.8	1.00	0.61	1.52	1.03	0.83	1.27	0.92	0.78	1.08
	σ_{33}	0.3	0.15	−0.18	0.48	0.33	0.17	0.49	0.10	−0.02	0.22
	σ_{34}	0.3	0.35	0.01	0.68	0.24	0.06	0.42	0.21	0.08	0.34
	σ_{44}	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Note: Post.mean = posterior mean; 95CI.low = lower bound of the 95% credible interval; and 95CI.up = upper bound of the 95% credible interval. These abbreviated terms remained the same hereinafter.

Table 2. Results of indirect effect estimates obtained from the simulation study varying across sample sizes

Parameters		Truth	N = 100			N = 300			N =500		
			Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up
Mediator $\mathbf{M_1}$											
$\mathbf{A_2A_4}$	$a_{2[1,1]} \times a_{4[1,1]}$	0	−0.053	−0.422	0.347	0.010	−0.148	0.177	−0.008	−0.143	0.131
	$a_{2[2,1]} \times a_{4[1,1]}$	0	0.003	−0.033	0.047	0.000	−0.008	0.008	0.000	−0.005	0.005
	$a_{2[1,1]} \times a_{4[1,2]}$	2	1.917	1.524	2.359	1.909	1.721	2.110	1.864	1.713	2.024
	$a_{2[2,1]} \times a_{4[1,2]}$	0	−0.112	−0.403	0.170	0.031	−0.125	0.188	0.002	−0.115	0.122
	$a_{2[1,1]} \times a_{4[1,3]}$	1	1.732	0.708	2.869	0.780	0.370	1.209	1.013	0.664	1.377
	$a_{2[2,1]} \times a_{4[1,3]}$	0	−0.102	−0.413	0.153	0.013	−0.052	0.083	0.001	−0.064	0.067
	$a_{2[1,1]} \times a_{4[1,4]}$	0	−1.188	−2.142	-0.326	0.239	−0.165	0.674	−0.110	−0.439	0.243
	$a_{2[2,1]} \times a_{4[1,4]}$	0	0.070	−0.103	0.286	0.004	−0.021	0.036	0.000	−0.016	0.014
Mediator $\mathbf{M_2}$											
$\mathbf{A_2A_4}$	$a_{2[1,2]} \times a_{4[2,1]}$	0	0.182	−0.316	0.685	−0.118	−0.354	0.124	−0.167	−0.348	0.005
	$a_{2[2,2]} \times a_{4[2,1]}$	4	4.166	3.458	4.894	3.969	3.613	4.336	4.015	3.739	4.302
	$a_{2[1,2]} \times a_{4[2,2]}$	0	0.086	−0.150	0.327	−0.061	−0.182	0.063	−0.087	−0.180	0.003
	$a_{2[2,2]} \times a_{4[2,2]}$	2	1.977	1.551	2.437	2.049	1.802	2.301	2.092	1.912	2.281
	$a_{2[1,2]} \times a_{4[2,3]}$	0	0.181	−0.318	0.730	−0.098	−0.300	0.103	−0.120	−0.256	0.004
	$a_{2[2,2]} \times a_{4[2,3]}$	3	4.150	2.501	6.062	3.300	2.589	4.048	2.887	2.392	3.437
	$a_{2[1,1]} \times a_{4[1,4]}$	0	0.186	−0.324	0.729	−0.119	−0.365	0.129	−0.169	−0.356	0.005
	$a_{2[2,1]} \times a_{4[1,4]}$	4	4.247	2.712	6.047	4.033	3.170	4.943	4.043	3.409	4.730

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

Table 3. Results of total effect estimates without the mediators from the simulation study varying across sample sizes

Parameters		Truth*	*N* = 100			*N* = 300			*N* =500		
			Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up	Post.mean	95CI.low	95CI.up
W	w_{11}	0.9	0.93	0.87	0.99	0.94	0.91	0.98	0.91	0.88	0.94
	w_{21}	0.5	0.46	0.37	0.55	0.50	0.44	0.55	0.54	0.50	0.57
	w_{31}	0	0.00	−0.09	0.09	−0.02	−0.09	0.04	−0.03	−0.08	0.01
	w_{41}	0	0.08	−0.01	0.17	0.01	−0.05	0.07	0.03	−0.01	0.08
	w_{51}	0	−0.06	−0.15	0.03	0.03	−0.04	0.09	−0.02	−0.06	0.03
	w_{62}	0.5	0.54	0.44	0.64	0.50	0.43	0.56	0.46	0.41	0.51
	w_{72}	0.5	0.44	0.32	0.54	0.49	0.42	0.55	0.53	0.48	0.57
$\mathbf{A_2}$	$a_{2[1,1]}$	2	–	–	–	–	–	–	–	–	–
	$a_{2[2,1]}$	0	–	–	–	–	–	–	–	–	–
	$a_{2[1,2]}$	0	–	–	–	–	–	–	–	–	–
	$a_{2[2,2]}$	2	–	–	–	–	–	–	–	–	–
$\mathbf{A_3}$	$a_{3[1,1]}$	1	1.24	0.71	1.78	0.79	0.50	1.07	0.81	0.60	1.01
	$a_{3[2,1]}$	0	4.08	3.34	4.80	3.73	3.33	4.10	3.83	3.52	4.14
	$a_{3[1,2]}$	2	4.08	3.68	4.48	3.76	3.52	3.99	3.86	3.68	4.03
	$a_{3[2,2]}$	1.5	3.40	2.83	3.97	3.45	3.13	3.77	3.45	3.20	3.70
	$a_{3[1,3]}$	1	1.39	0.99	1.82	0.95	0.74	1.16	1.03	0.86	1.21
	$a_{3[2,3]}$	1.5	2.71	1.95	3.52	2.27	1.90	2.65	2.49	2.19	2.81
	$a_{3[1,4]}$	2	1.24	0.85	1.69	0.90	0.69	1.12	1.02	0.85	1.20
	$a_{3[2,4]}$	0	2.18	1.51	2.90	2.14	1.77	2.52	2.12	1.84	2.41
$\mathbf{A_4}$	$a_{4[1,1]}$	0	–	–	–	–	–	–	–	–	–
	$a_{4[2,1]}$	2	–	–	–	–	–	–	–	–	–
	$a_{4[1,2]}$	1	–	–	–	–	–	–	–	–	–

(Continued)

Table 3. (Continued)

Parameters		Truth*	N = 100 Post.mean	N = 100 95CI.low	N = 100 95CI.up	N = 300 Post.mean	N = 300 95CI.low	N = 300 95CI.up	N = 500 Post.mean	N = 500 95CI.low	N = 500 95CI.up
	$a_{4[2,2]}$	1	–	–	–	–	–	–	–	–	–
	$a_{4[1,3]}$	0.5	–	–	–	–	–	–	–	–	–
	$a_{4[2,3]}$	1.5	–	–	–	–	–	–	–	–	–
	$a_{4[1,4]}$	0	–	–	–	–	–	–	–	–	–
	$a_{4[2,4]}$	2	–	–	–	–	–	–	–	–	–
$\mathbf{a}_{02}$	$\mathrm{a}_{02[1]}$	1	–	–	–	–	–	–	–	–	–
	$\mathrm{a}_{02[2]}$	1	–	–	–	–	–	–	–	–	–
$\mathbf{a}_{03}$	$\mathrm{a}_{03[1]}$	1	2.68	2.05	3.29	3.07	2.75	3.38	2.89	2.66	3.13
	$\mathrm{a}_{03[2]}$	0	1.65	1.17	2.11	2.12	1.84	2.39	1.90	1.70	2.09
	$\mathrm{a}_{03[3]}$	2	2.55	1.91	3.28	1.94	1.63	2.26	2.26	1.99	2.53
	$\mathrm{a}_{03[4]}$	−1	0.47	0.00	0.95	0.57	0.30	0.84	0.69	0.49	0.89
$\mathbf{\Sigma}_2$	σ_{11}	1	–	–	–	–	–	–	–	–	–
	σ_{12}	0.3	–	–	–	–	–	–	–	–	–
	σ_{22}	1	–	–	–	–	–	–	–	–	–
$\mathbf{\Sigma}_3$	σ_{11}	0.8	8.74	5.98	12.40	7.12	5.73	8.72	6.80	5.79	7.97
	σ_{12}	0.3	5.07	3.28	7.51	4.31	3.32	5.44	4.03	3.33	4.80
	σ_{13}	0.3	2.37	1.73	3.05	2.18	1.86	2.50	1.96	1.73	2.21
	σ_{22}	0.3	2.51	1.93	3.15	2.13	1.80	2.46	2.23	2.01	2.47
	σ_{23}	0.8	4.96	3.37	7.08	5.01	4.00	6.16	4.57	3.90	5.32
	σ_{33}	0.3	1.85	1.41	2.35	1.68	1.40	1.96	1.48	1.28	1.69
	σ_{34}	0.3	1.36	0.86	1.89	1.57	1.29	1.86	1.38	1.18	1.59
	σ_{44}	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

* True values are equivalent to Table 1 as the simulation datasets were generated based on a model with the two mediators.

however, were just rough estimates after controlling for other predictor(s) and mediator(s) (an exact calculation would become only feasible in a setting with single predictor and mediator but without any covariates). When a model was mis-specified by omitting mediators, we found out that the estimates of intercepts $\mathbf{a}$ and residual variance-covariance matrix $\mathbf{\Sigma}_3$ would be more likely to be biased than the weight estimates $\mathbf{W}$. The weight estimates remained relatively consistent with the earlier results shown in Table 1. However, unlike the weight estimates, elements of $\mathbf{a}$ and $\mathbf{\Sigma}_3$ exhibited more bias and susceptibility to the absence of mediators. This indicated that the direct effects alone were insufficient in accounting for all the omitted information associated with the mediators and rather remained as unexplained variation in outcome variables. Consequently, this contributed to biases in the estimates of their intercepts and the residual variance-covariance matrix.

4. An illustrative example

We used a subset of the National Survey on Drug Use and Health (NSDUH) data (Substance Abuse and Mental Health Services Administration (SAMHSA), United States Department of Health and Human Services, 2015), from which 2,347 observations (N = 2,347 whose age was 12 or older) responded to a number of questionnaire items concerning substance use and health in 2012. As shown in Figure 4, we constructed two components, i.e., *socioeconomic status* (SES) and *drug history* (DRH). The first component SES was defined as a linear combination of four observed variables: education, insurance, family income, and employment status. The second component DRH was constructed as a linear combination of use of cigarette and alcohol, asking the age of first use. The extent to which participants were dependent on nicotine (measured by nicotine dependence syndrome scale) (nicotine) was used one of the outcome variables. Another continuous outcome variable was measuring past month psychological distress (Pdistress). The remaining two outcome variables were binary variables asking whether or not they have been dependent on alcohol (alcohol) and pain reliever (pain reliver) in the past year. In this example, we examined the effect of SES and drug history on alcohol dependence, psychological distress, and other drug addictions while exploring perceived mental and physical health conditions as potential mediators. Specifically, there were three mediators named distress (measuring

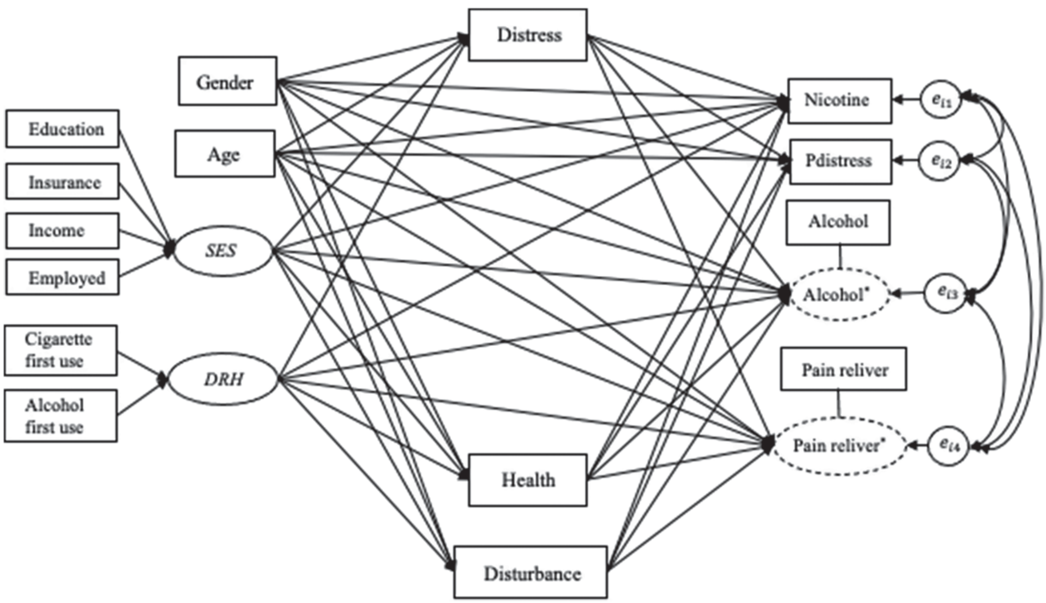

Figure 4. A hypothesized model for the empirical data.

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

the level of psychological distress in the past year excluding the latest past month used for measuring Pdistress), health (measuring the overall perceived health condition), and disturb (measuring the level of psychological impairment and disturbances in social adjustment and behavior). It was hypothesized that the perception of mental and physical health conditions would both mediate the relationship between SES and drug history to alcohol and drug addictions.

We fitted this dataset with the proposed method, fast convergence was observed across all parameters of interest as shown in Figure 5. The results for the parameters are presented in Table 4. Looking at the results of direct effects from components on outcome variables given in $\mathbf{A}_3$, the two components SES and DRH showed significant and negative direct effects on nicotine addition and pain reliever but not on pain reliver. It suggested that those who have higher socioeconomic status and exposed to cigarette later were less likely to show nicotine dependence ($a_{3[3,1]} = -0.090$, 95% CI = [−0.117, −0.063]; $a_{3[4,1]} = -0.120$, 95% CI = [−0.143, −0.098]) and pain reliver dependence ($a_{3[3,4]} = 0.383$, 95% CI = [0.209, 0.701]; $a_{3[4,4]} = 0.381$, 95% CI = [0.223, 0.646]). The odds of pain reliver dependence was reduced by approximately 62% as both SES and DRH increased by a one-unit. Similarly, the directionality between the two components and past month psychological distress (Pdistress) were all negative. The statistical significance, however, seemed to be present for SES only among the two components. Among the covariates, age showed consistently significant impacts on nicotine, past month distress and pain reliever dependencies ($a_{3[2,1]} = 0.137$, 95% CI = [0.107, 0.167]; $a_{3[2,2]} = -0.041$, 95% CI = [−0.067, −0.015]; $a_{3[2,4]} = 0.433$, 95% CI = [0.205, 0.925]): older people tended to show more dependency on nicotine but less on past month distress and pain reliever. Each additional increase of one year in age was associated with a 56.7% decrease in the odds of being more dependent on pain reliever. Females showed much higher dependency rate on alcohol (1.44 times larger in the odds for alcohol dependency: $a_{3[1,2]} = 1.439$, 95% CI = [1.171, 1.776]) compared to males, but no significant direct impact on nicotine dependence, past month distress and pain reliever dependence.

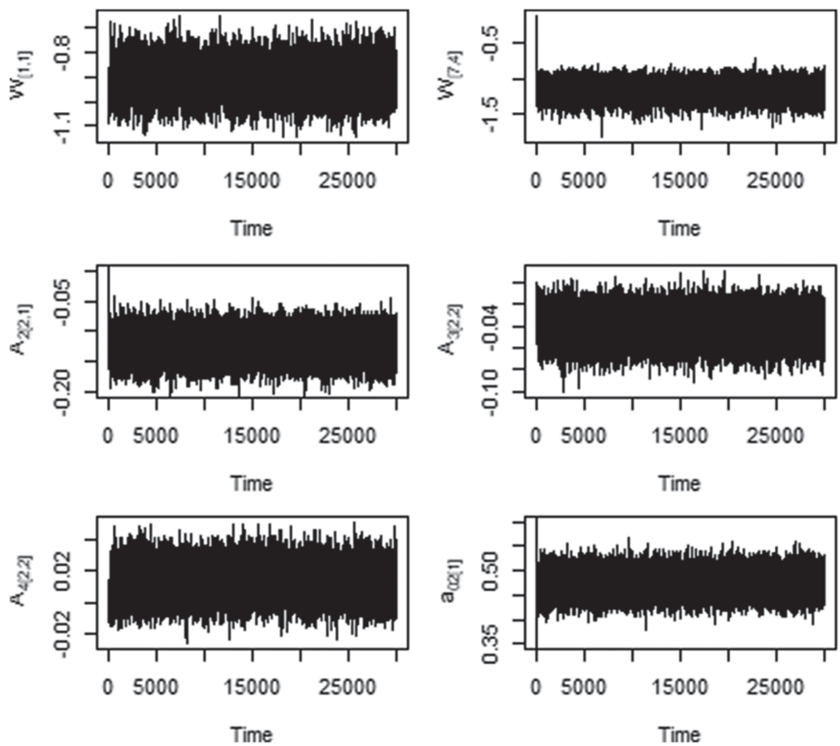

Figure 5. Trace plots of selected parameters for the empirical data.

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

Table 4. Results of fitting the National Survey on Drug Use and Health (NSDUH) data

Parameters		Estimate*		
		Post.mean	95CI.low	95CI.up
W	w_{11}	0.905	0.783	1.025
	w_{21}	−0.052	−0.143	0.032
	w_{31}	0.264	0.158	0.378
	w_{41}	0.384	0.304	0.466
	w_{52}	1.175	0.958	1.407
	w_{62}	0.248	0.006	0.469
$\mathbf{A}_2$	$a_{2[1,1]}$	−0.045	−0.067	−0.024
	$a_{2[2,1]}$	−0.124	−0.165	−0.082
	$a_{2[3,1]}$	−0.100	−0.133	−0.068
	$a_{2[4,1]}$	−0.064	−0.095	−0.034
	$a_{2[1,2]}$	−0.078	−0.099	−0.057
	$a_{2[2,2]}$	0.034	−0.007	0.075
	$a_{2[3,2]}$	−0.113	−0.146	−0.080
	$a_{2[4,2]}$	−0.047	−0.077	−0.017
	$a_{2[1,3]}$	−0.016	−0.035	0.003
	$a_{2[2,3]}$	0.175	0.138	0.212
	$a_{2[3,3]}$	−0.204	−0.236	−0.174
	$a_{2[4,3]}$	−0.053	−0.080	−0.026
$\mathbf{A}_3$	$a_{3[1,1]}$	0.006	−0.009	0.022
	$a_{3[2,1]}$	0.137	0.107	0.167
	$a_{3[3,1]}$	−0.090	−0.117	−0.063
	$a_{3[4,1]}$	−0.120	−0.143	−0.098
	$a_{3[1,2]}$	0.006	−0.008	0.019
	$a_{3[2,2]}$	−0.041	−0.067	−0.015
	$a_{3[3,2]}$	−0.088	−0.109	−0.068
	$a_{3[4,2]}$	−0.014	−0.032	0.005
	$a_{3[1,3]}$	1.439	1.171	1.776
	$a_{3[2,3]}$	0.681	0.454	1.030
	$a_{3[3,3]}$	1.121	0.814	1.536
	$a_{3[4,3]}$	0.806	0.597	1.083
	$a_{3[1,3]}$	1.408	0.949	2.096
	$a_{3[2,4]}$	0.433	0.205	0.925
	$a_{3[3,4]}$	0.383	0.209	0.701
	$a_{3[4,4]}$	0.381	0.223	0.646
$\mathbf{A}_4$	$a_{4[1,1]}$	0.085	0.049	0.122
	$a_{4[2,1]}$	0.029	−0.006	0.064

(Continued)

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

Table 4. (Continued)

Parameters		Estimate*		
		Post.mean	95CI.low	95CI.up
	$a_{4[3,1]}$	0.089	0.056	0.122
	$a_{4[1,2]}$	0.474	0.441	0.506
	$a_{4[2,2]}$	0.052	0.021	0.084
	$a_{4[3,2]}$	0.116	0.089	0.145
	$a_{4[1,3]}$	3.736	2.306	6.061
	$a_{4[2,3]}$	2.301	1.484	3.639
	$a_{4[3,3]}$	0.876	0.573	1.351
	$a_{4[1,4]}$	2.357	0.970	5.973
	$a_{4[2,4]}$	4.413	1.892	10.445
	$a_{4[3,4]}$	0.362	0.165	0.834
$\mathbf{a}_{02}$	$a_{02[1]}$	0.758	0.712	0.803
	$a_{02[2]}$	0.555	0.511	0.601
	$a_{02[3]}$	0.551	0.511	0.592
$\mathbf{a}_{03}$	$a_{03[1]}$	0.475	0.432	0.519
	$a_{03[2]}$	0.143	0.107	0.178
	$a_{03[3]}$	−2.330	−2.912	−1.743
	$a_{03[4]}$	−1.855	−2.845	−0.892
$\mathbf{\Sigma}_2$	σ_{11}	0.061	0.058	0.065
	σ_{12}	0.037	0.034	0.041
	σ_{13}	0.009	0.007	0.012
	σ_{22}	0.065	0.061	0.069
	σ_{23}	0.008	0.006	0.011
	σ_{33}	0.049	0.047	0.052
$\mathbf{\Sigma}_3$	σ_{11}	0.032	0.030	0.034
	σ_{12}	0.002	0.001	0.003
	σ_{13}	−0.006	−0.016	0.005
	σ_{14}	0.026	0.009	0.043
	σ_{22}	0.023	0.022	0.025
	σ_{23}	0.027	0.018	0.036
	σ_{24}	0.025	0.010	0.038
	σ_{33}	1.000	1.000	1.000
	σ_{34}	0.023	−0.117	0.159
	σ_{44}	1.000	1.000	1.000

* The parameters loaded on binary outcome variables were reported by exponentiating the original estimate values. Those reported in odds ratio were displayed with grey shading behind the text. This applied the same to the following table.

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

The estimates for mediation effects of distress, health, and disturb are presented in Table 5. The indirect effect of distress as a mediator was significant on the two components as well as the two covariates. It was observed that SES and DRH were significant direct predictors of the lower level of nicotine dependence (as shown in A_3 from Table 3) and also had significant indirect effects through distress ($a_{2[3,1]} \times a_{4[1,1]} = -0.008$, 95% CI = [−0.013, −0.004] for the indirect effect of SES on nicotine; $a_{2[4,1]} \times a_{4[1,1]} = -0.005$, 95% CI = [−0.009, −0.002] for the indirect effect of DRH on nicotine). SES and DRH led to lower levels of distress, which in turn yielded a positive impact on nicotine dependence. That is, those who have higher socioeconomic status and were exposed to cigarettes later showed lower level of distress ($a_{2[3,1]} = -0.100$, 95% CI = [−0.133, −0.068]; $a_{2[4,1]} = -0.064$, 95% CI = [−0.095, −0.034]), and this also led to less nicotine addiction ($a_{4[1,1]} = 0.085$, 95% CI = [0.049, 0.122]). This same relationship was also present for the mediation effect of distress on past month distress ($a_{2[3,1]} \times a_{4[1,2]} = -0.047$, 95%

Table 5. Results of indirect effect estimates obtained from the empirical NSDUH data

Parameters		post.mean	95CI.low	95CI.high
Mediator **Distress**				
A$_2$**A**$_4$	$a_{2[1,1]} \times a_{4[1,1]}$	−0.004	−0.007	−0.002
	$a_{2[2,1]} \times a_{4[1,1]}$	−0.011	−0.017	−0.005
	$a_{2[3,1]} \times a_{4[1,1]}$	−0.008	−0.013	−0.004
	$a_{2[4,1]} \times a_{4[1,1]}$	−0.005	−0.009	−0.002
	$a_{2[1,1]} \times a_{4[1,2]}$	−0.022	−0.032	−0.012
	$a_{2[2,1]} \times a_{4[1,2]}$	−0.059	−0.078	−0.039
	$a_{2[3,1]} \times a_{4[1,2]}$	−0.047	−0.063	−0.032
	$a_{2[4,1]} \times a_{4[1,2]}$	−0.030	−0.045	−0.016
	$a_{2[1,1]} \times a_{4[1,3]}$	0.942	0.905	0.972
	$a_{2[2,1]} \times a_{4[1,3]}$	0.850	0.778	0.915
	$a_{2[3,1]} \times a_{4[1,3]}$	0.876	0.817	0.930
	$a_{2[4,1]} \times a_{4[1,3]}$	0.919	0.870	0.962
	$a_{2[1,1]} \times a_{4[1,4]}$	0.962	0.913	1.001
	$a_{2[2,1]} \times a_{4[1,4]}$	0.900	0.792	1.004
	$a_{2[3,1]} \times a_{4[1,4]}$	0.918	0.827	1.003
	$a_{2[4,1]} \times a_{4[1,4]}$	0.947	0.880	1.002
Mediator **Health**				
A$_2$**A**$_4$	$a_{2[1,2]} \times a_{4[2,1]}$	−0.001	−0.003	0.000
	$a_{2[2,2]} \times a_{4[2,1]}$	0.016	0.009	0.023
	$a_{2[3,2]} \times a_{4[2,1]}$	−0.018	−0.026	−0.011
	$a_{2[4,2]} \times a_{4[2,1]}$	−0.005	−0.008	−0.002
	$a_{2[1,2]} \times a_{4[2,2]}$	−0.002	−0.004	0.000
	$a_{2[2,2]} \times a_{4[2,2]}$	0.020	0.014	0.028
	$a_{2[3,2]} \times a_{4[2,2]}$	−0.024	−0.031	−0.017
	$a_{2[4,2]} \times a_{4[2,2]}$	−0.006	−0.010	−0.003

(Cotinued)

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

Table 5. (Continued)

Parameters		post.mean	95CI.low	95CI.high
	$a_{2[1,2]} \times a_{4[2,3]}$	1.002	0.995	1.012
	$a_{2[2,2]} \times a_{4[2,3]}$	0.977	0.904	1.055
	$a_{2[3,2]} \times a_{4[2,3]}$	1.027	0.940	1.122
	$a_{2[4,2]} \times a_{4[2,3]}$	1.007	0.983	1.033
	$a_{2[1,2]} \times a_{4[2,4]}$	1.016	0.997	1.047
	$a_{2[2,2]} \times a_{4[2,4]}$	0.837	0.723	0.969
	$a_{2[3,2]} \times a_{4[2,4]}$	1.231	1.038	1.457
	$a_{2[4,2]} \times a_{4[2,4]}$	1.056	1.009	1.120
Mediator **Disturbance**				
$\mathbf{A_2A_4}$	$a_{2[1,3]} \times a_{4[3,1]}$	−0.002	−0.005	0.000
	$a_{2[2,3]} \times a_{4[3,1]}$	0.001	0.000	0.003
	$a_{2[3,3]} \times a_{4[3,1]}$	−0.003	−0.008	0.001
	$a_{2[4,3]} \times a_{4[3,1]}$	−0.001	−0.004	0.000
	$a_{2[1,3]} \times a_{4[3,2]}$	−0.004	−0.007	−0.002
	$a_{2[2,3]} \times a_{4[3,2]}$	0.002	0.000	0.005
	$a_{2[3,3]} \times a_{4[3,2]}$	−0.006	−0.010	−0.002
	$a_{2[4,3]} \times a_{4[3,2]}$	−0.002	−0.005	−0.001
	$a_{2[1,3]} \times a_{4[3,3]}$	0.937	0.898	0.971
	$a_{2[2,3]} \times a_{4[3,3]}$	1.029	0.995	1.077
	$a_{2[3,3]} \times a_{4[3,3]}$	0.910	0.855	0.959
	$a_{2[4,3]} \times a_{4[3,3]}$	0.962	0.926	0.989
	$a_{2[1,3]} \times a_{4[3,4]}$	0.891	0.823	0.954
	$a_{2[2,3]} \times a_{4[3,4]}$	1.052	0.990	1.141
	$a_{2[3,3]} \times a_{4[3,4]}$	0.846	0.753	0.934
	$a_{2[4,3]} \times a_{4[3,4]}$	0.933	0.867	0.981

CI = [−0.063, −0.032] for the indirect effect of SES on distress; $a_{2[4,1]} \times a_{4[1,2]}$ = −0.030 95%CI = [−0.045, −0.016] for the indirect effect of DRH on distress). For alcohol dependence, the odds were reduced by about 10% through the mediation effect via distress ($a_{2[3,1]} \times a_{4[1,3]}$ = 0.876, 95% CI = [0.817, 0.930] for the indirect effect SES on alcohol; $a_{2[4,1]} \times a_{4[1,3]}$ = 0.919, 95% CI = [0.870, 0.962] for the indirect effect of DRH on alcohol). While there was a notable and significant indirect effect observed for nicotine, past month distress, and alcohol dependence through a lower level of distress in the past year, the indirect effects on pain reliever were not found to be significant.

In addition to significant direct effects from SES and DRH on nicotine dependence ($a_{3[3,1]}$ = −0.090, 95% CI = [−0.117, −0.063] for SES; $a_{3[4,1]}$ = −0.120, 95% CI = [−0.143, −0.098] for DRH), SES and DRH had a significant indirect effect on nicotine dependence ($a_{2[3,2]} \times a_{4[2,1]}$ = −0.018, 95% CI = [−0.026, −0.011] for the indirect effect of SES on nicotine via health; $a_{2[4,2]} \times a_{4[2,1]}$ = −0.005, 95% CI = [−0.008, −0.002] for the indirect effect of DRH on nicotine via health) and past month distress ($a_{2[3,2]} \times a_{4[2,2]}$ = −0.024, 95% CI = [−0.031, −0.017] for the indirect effect of SES on past month distress via health; $a_{2[4,2]} \times a_{4[2,2]}$ = −0.006, 95% CI = [−0.010, −0.003] for the indirect effect of DRH on distress via health) through lower risk for health problems, respectively. The odds of pain reliever dependence increased

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

by 1.231 ($a_{2[3,2]} \times a_{4[2,4]}$ = 1.231, 95% CI = [1.038, 1.457]) and 1.056 ($a_{2[4,2]} \times a_{4[2,4]}$ = 1.056, 95% CI = [1.009, 1.120]) times, respectively, through the mediation effect via health. Such a pattern was not shown in alcohol dependence. Lastly, one's perceived level of disturbances in social adjustment and behavior significantly mediated the relationships between the two components and three outcome variables (past month distress, alcohol, and pain reliever dependence). The estimated indirect effects from SES and DRH on distress ($a_{2[3,3]} \times a_{4[3,2]}$ = −0.006 (95% CI = [−0.010, −0.002]); $a_{2[4,3]} \times a_{4[3,2]}$ = −0.002 (95% CI = [−0.005, −0.001]), alcohol ($a_{2[3,3]} \times a_{4[3,3]}$ = 0.910 (95% CI = [0.855, 0.959]); $a_{2[4,3]} \times a_{4[3,3]}$ = 0.962 (95% CI = [0.926, 0.989]), and pain reliever dependence ($a_{2[3,3]} \times a_{4[3,4]}$ = 0.846 (95% CI = [0.753, 0.934]); $a_{2[4,3]} \times a_{4[3,4]}$ = 0.933 (95% CI = [0.867, 0.981]) via disturbances, respectively. SES and DRH showed a negative and significant effect on disturbance, which in turn led to a positive and significant impact on all three outcome variables. Individuals with a lower perceived level of disturbance, influenced by higher SES status and later onset of cigarette consumption, exhibited reduced addictive consumption of alcohol and pain reliever (though not for nicotine dependence). Specifically, when considering odds ratios obtained by exponentiating the corresponding posterior estimates, a one-unit increase in SES and DRH was associated with a 9.00% and 3.80% decrease in the odds of being more dependent on alcohol, respectively, mediated by the level of psychological impairment and disturbance. This statistical significance and interpretation remained consistent for the mediated effect of SES and DRH on pain reliever dependency via disturbance, resulting in a 15.40% and 6.70% decrease in the odds for pain reliever dependency, respectively.

5. Concluding remarks

We introduced a multivariate component-based regression model designed to handle mixed types of outcomes and estimate multiple pathways of indirect effects with multiple mediators within a Bayesian framework. The efficacy of our proposed approach was validated through simulated and real data instances. The simulation design intentionally aimed to depict a simplified scenario, characterized by significant discrepancies in actual values, with a particular focus on parameter estimation. In a real-world application, we analyzed a subset of NSDUH data to elucidate how the underlying mechanism of perceived mental and physical health conditions influences the relationship between components (SES and DRH) and drug dependence (nicotine, alcohol, and pain reliever).

In addition to its technical and empirical implications, the proposed method holds the potential for further enhancement in flexibility. An intriguing extension could involve the incorporation of variable selection techniques. While the current conceptualization of the method relies on a predetermined model, there are scenarios where it may be preferable to select an optimal subset of mediators, especially when dealing with a large number of potential mediators. The inclusion of techniques such as lasso (Tibshirani, 1996) or elastic net (Zou & Hastie, 2005) for variable selection could help eliminate irrelevant mediators, thereby improving the interpretability of potential indirect effects within the model.

Moreover, the current proposed method assumes that subjects in the dataset are randomly sampled from the same population. Consequently, it is not possible with the current version to assess differences in effect estimates for different subject clusters using the mediation model with mixed types of outcome data. Considering a recent study proposing a Bayesian approach to ERA with mixture modeling (Kyung et al., 2021), future investigations are warranted to explore the technical and empirical feasibility of incorporating such settings.

Acknowledgements. This research was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1F1A1049834; NRF-2021R1F1A1049834).

Software availability. To ensure applicability, we have shared an R function along with the simulation code specific to this paper on GitHub (https://github.com/jhppack/BERA_mediation).

References

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(1), 97–106. 10.1002/wics.51.

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Albert, J.. & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, *88*(422), 669–679. 10.2307/2290350.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173.

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550–558. 10.1037/a0018933.

Choi, J. Y., Kyung, M., Hwang, H., & Park, J-H. (2020). Bayesian extended redundancy analysis: A Bayesian approach to component-based regression with dimension reduction. *Multivariate Behavioral Research*, *55*(1), 30–48. 10.1080/00273171.2019.1598837.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Rouotledge.

DeSarbo, W. S., Hwang, H., Blank, A. S., & Kappe, E. (2015). Constrained stochastic extended redundancy analysis. *Psychometrika*, *80*(2), 516–534. 10.1007/s11336-013-9385-6.

Gelman, A., Huang, Z., van Dyk, D. A., & Boscardin, W. J. (2004). Transformed and parameter-expanded Gibbs samplers for hierarchical linear and generalized linear models. *Technical report*. Department of Statistics, Columbia University.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). CRC Press.

Henningsson, M., Sundbom, E., Armelius, B., & Erdberg, P. (2001). PLS model building: A multivariate approach to personality test data. *Scandinavian Journal of Psychology*, *42*(5), 399–409. 10.1111/1467-9450.00252.

Hirk, R., Hornik, K. & Vana, L. (2019). Multivariate ordinal regression models: An analysis of corporate credit ratings. *Statistical Methods and Applications*, *28*, 507–539. 10.1007/s10260-018-00437-7.

Hwang, H., Suk, H. W., Lee, J. H., Moskowitz, D. S., & Lim, J. (2012). Functional extended redundancy analysis. *Psychometrika*, *77*(3), 524–542. 10.1007/s11336-012-9268-2.

Jolliffe, I. T. (1982). A note on the use of principal components in regression. *Applied Statistics*, 300–303. 10.2307/2348005.

Kyung, M., Park, J.-H., Choi, J. Y. (2021). Bayesian mixture extended redundancy analysis. *Psychometrika*. 10.1007/s11336-021-09809-7.

Lee, S., Choi, S., Kim, Y. J., Kim, B.-J., Consortium, T. D.-G., Hwang, H., & Park, T. (2016). Pathway-based approach using hierarchical components of collapsed rare variants. *Bioinformatics*, *32*(17), i586–i594. 10.1093/bioinformatics/btw425.

Lee, S., Kim, Y., Choi, S., Hwang, H., & Park, T. (2018). Pathway-based approach using hierarchical components of rare variants to analyze multiple phenotypes. *BMC Bioinformatics*, *19*(4), 79, 85–97. 10.1186/s12859-018-2066-9.

Lovaglio, P. G., & Vacca, G. (2016). %ERA: A SAS Macro for Extended Redundancy Analysis. *Journal of Statistical Software*, 74(Code Snippet 1). https://doi.org/10.18637/jss.v074.c01.

Lovaglio, P. G., & Vittadini, G. (2014). Structural equation models in a redundancy analysis framework with covariates, *Multivariate Behavioral Research*, *49*(5), 486–501. 10.1080/00273171.2014.931798.

MacKinnon, D. P., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science*, *18*(1), 16–20. 10.1111/j.1467-8721.2009.01598.x.

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341. 10.2307/2280232.

Miočević, M., Gonzalez, O., Valente, M. J. & MacKinnon, D. P. (2017). A tutorial in Bayesian potential outcomes mediation analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(1), 121–136. 10.1080/10705511.2017.1342541.

O'Brien, S. M., & Dunson, D. B. (2004). Bayesian multivariate logistic regression. *Biometrics*, *60*(3), 739–746. 10.1111/j.0006-341X.2004.00224.x.

Park, J.-H., Choi, J. Y., Lee, J., & Kyung, M. (2021). Bayesian approach to multivariate component-based logistic regression: Analyzing correlated multivariate ordinal data. *Multivariate Behavioral Research*. 10.1080/00273171.2021.1874260.

Pieters, R. (2017). Meaningful mediation analysis: Plausible causal inference and informative communication. *Journal of Consumer Research*, *44*(3), 692–716. 10.1093/jcr/ucx081.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. 10.3758/BRM.40.3.879.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Selig, J. P., & Preacher, K. J. (2021). Mediation analysis with multiple mediators and complex models. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 229–249. 10.1080/10705511.2020.1805074.

Takane, Y., & Hwang, H. (2005). An extended redundancy analysis and its applications to two practical examples. *Computational Statistics & Data Analysis*, *49*(3), 785–808. 10.1016/j.csda.2004.06.004.

Tan, T., Choi, J. Y., & Hwang, H. (2015). Fuzzy clusterwise functional extended redundancy analysis. *Behaviormetrika*, *42*(1), 37–62. 10.2333/bhmk.42.37.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288. 10.1111/j.2517-6161.1996.tb02080.x.

VanderWeele, T. J. (2016). Mediation analysis: A practitioner's guide. *Annual Review of Public Health*, *37*, 17–32. 10.1146/annurev-publhealth-032315-021402.

Wehrens, R., & Mevik, B. H. (2007). The pls package: Principal component and partial least squares regression in R. *Journal of Statistical Software*, *018* (i02). 10.18637/jss.v018.i02.

Wold, H. (1975). Soft modeling by latent variables: The nonlinear iterative partial least squares approach. In *Perspectives in Probability and Statistics, Papers in Honour of M. S. Bartlett* (pp. 117–144). 10.1017/S0021900200047604.

Wold, S., Ruhe, A., Wold, H., & Dunn, W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, *5*(3), 735–743. 10.1137/0905052.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B (Methodological)*, *67*(2), 301–320. 10.1111/j.1467-9868.2005.00503.x.

1. Appendix

A1. Prior distribution

In the proposed method, we consider the following conjugate distributions for specifying priors on the parameters of interest:

- $\mathbf{W}_{[,k]} \mid \tau^2 \sim N_{p_k}\left(\mathbf{W}^0_{[,k]},\ \tau^2\boldsymbol{\Sigma}^0_{\mathbf{W}_k}\right)$ for $k = 1,\ldots,K$
- $\mathbf{A}'_{2[k,]} \mid \eta_2^2 \sim N_S\left(\mathbf{A}^{0\prime}_{2[k,]},\ \eta_2^2\boldsymbol{\Sigma}^0_{\mathbf{A}_{2k}}\right)$ for $k = 1,\ldots,K$
- $\mathbf{B}'_{2[d,]} \mid \delta_2^2 \sim N_S\left(\mathbf{B}^{0\prime}_{2[d,]},\ \delta_2^2\boldsymbol{\Sigma}^0_{B_{2d}}\right)$ for $d = 1,\ldots,D$
- $\mathbf{a}_2 \mid \sigma_2^2 \sim N_S\left(\boldsymbol{\alpha}_2^0,\ \sigma_2^2\boldsymbol{\Sigma}^0_{\mathbf{a}_{20}}\right)$
- $\boldsymbol{\Sigma}_2^{-1} \mid v_2 \sim Wishart_S\left(v_2,\ \boldsymbol{\Sigma}_{20}^{-1}\right)$ where $v_2 \geq S$
- $\mathbf{A}'_{3[k,]} \mid \eta_3^2 \sim N_Q\left(\mathbf{A}^{0\prime}_{3[k,]},\ \eta_3^2\boldsymbol{\Sigma}^0_{\mathbf{A}_{3k}}\right)$ for $k = 1,\ldots,K$
- $\mathbf{A}'_{4[s,]} \mid \eta_4^2 \sim N_Q\left(\mathbf{A}^{0\prime}_{4[s,]},\ \eta_4^2\boldsymbol{\Sigma}^0_{\mathbf{A}_{4s}}\right)$ for $s = 1,\ldots,S$
- $\mathbf{B}'_{3[d,]} \mid \delta_3^2 \sim N_Q\left(\mathbf{B}^{0\prime}_{3[d,]},\ \delta_3^2\boldsymbol{\Sigma}^0_{B_{3d}}\right)$ for $d = 1,\ldots,D$
- $\mathbf{a}_3 \mid \sigma_3^2 \sim N_Q\left(\boldsymbol{\alpha}_3^0,\ \sigma_3^2\boldsymbol{\Sigma}^0_{\mathbf{a}_{30}}\right)$
- $\boldsymbol{\Sigma}_3^{-1} \mid v_3 \sim Wishart_Q\left(v_3,\ \boldsymbol{\Sigma}_{30}^{-1}\right)$ where $v_3 \geq Q$
- $1/v_2^* \sim U[0,1]$
- $1/v_3^* \sim U[0,1]$
- $\boldsymbol{\Gamma}_{[j,q]} \equiv \gamma_j^q \sim \pi\left(\gamma_j^q\right) \propto constant$ for $j = 1,\ldots,C-1$ and $q = T+1,\ldots,Q$.

Here, $\mathbf{W}_{[,k]}$ refers to a p_k by 1 column vector containing the weight estimates for the kth component, $\mathbf{A}'_{2[k,]}$ is a S by 1 column vector containing the component coefficients for the kth component affecting the mediators, $\mathbf{A}'_{3[k,]}$ is a Q by 1 column vector containing the component coefficients for the kth component affecting outcome variables when adjusted for the mediator and $\mathbf{A}'_{4[s,]}$ is a Q by 1 column component coefficients vector of the sth mediation variable adjusted for the predictors. Also, $\mathbf{B}'_{2[d,]}$ is a S by 1 column vector containing the regression coefficients for the dth predictor affecting the mediators and $\mathbf{B}'_{3[d,]}$ is a Q by 1 column vector containing the regression coefficients for the dth predictor affecting the outcomes. $\boldsymbol{\Gamma}$ is a $C-1$ by $Q-T$ matrix of $C-1$ cutpoints for $Q-T$ ordinal responses.

A2. Sampling scheme

An overview of the general sampling scheme is as follows.

ERA structure parameter

- $\mathbf{W}_{[,k]} \mid \mathbf{W}_{[,-k]}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{B}_3, \mathbf{a}_2, \mathbf{a}_3, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Gamma}, \phi_3, \boldsymbol{\xi}_3, v_3^*, \boldsymbol{\xi}_2, v_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$

  $\sim N_{p_k}\left(\mathbf{W}^*_{[,k]}, \boldsymbol{\Sigma}_{\mathbf{W}_k}\right)$

  for $k = 1,\cdots,K$ where

  $$\boldsymbol{\Sigma}_{\mathbf{W}_k} = \left(\sum_{i=1}^{N} \xi_{2i}\mathbf{X}'_{[i,]}\mathbf{A}_{2[k,]}\boldsymbol{\Sigma}_2^{-1}\mathbf{A}'_{2[k,]}\mathbf{X}_{[i,]} + \sum_{i=1}^{N} \mathbf{X}'_{[i,]}\mathbf{A}_{3[k,]}\mathbf{D}_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{-1}\mathbf{D}_{3i}^{-1/2}\mathbf{A}'_{3[k,]}\mathbf{X}_{[i,]} + \frac{1}{\tau^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{W}_k}\right)^{-1}$$

$$\mathbf{W}^*_{[,k]} = \boldsymbol{\Sigma}_{\mathbf{W}_k}\left(\sum_{i=1}^{N}\xi_{2i}\mathbf{X}'_{[i,]}\mathbf{A}_{2[k,]}\boldsymbol{\Sigma}_2^{-1}\mathbf{M}^{-k\prime}_{[i,]} + \sum_{i=1}^{N}\mathbf{X}'_{[i,]}\mathbf{A}_{3[k,]}\mathbf{D}_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{-1}\mathbf{D}_{3i}^{-1/2}\mathbf{Z}^{-k\prime}_{[i,]} + \tfrac{1}{\tau^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{W}_k}\mathbf{W}^0_{[,k]}\right),$$
$\mathbf{M}^{-k}_{[i,]} = \mathbf{M}_{[i,]} - \mathbf{a}'_2 - \mathbf{H}_{[i,]}\mathbf{B}_2 - \mathbf{X}_{[i,]}\mathbf{W}_{[,-k]}\mathbf{A}_{2[-k,]}$, and $\mathbf{Z}^{-k}_{[i,]} = \mathbf{Z}_{[i,]} - \mathbf{a}'_3 - \mathbf{H}_{[i,]}\mathbf{B}_3 - \mathbf{X}_{[i,]}\mathbf{W}_{[,-k]}\mathbf{A}_{3[-k,]} - \mathbf{M}_{[i,]}\mathbf{A}_4$.

- To satisfy the stadardization constraint $diag(\mathbf{F}'\mathbf{F}) = diag(\mathbf{W}'\mathbf{X}'\mathbf{X}\mathbf{W}) = N\mathbf{I}$, we standardize $\mathbf{W}$ such as $\overline{\mathbf{W}} = \sqrt{N}(\mathbf{W}'\mathbf{X}'\mathbf{X}\mathbf{W})^{-1/2}\mathbf{W}$.

Mediator model parameters

- $\xi_{2i} \mid \mathbf{Z}_{[i,]}, \mathbf{X}_{[i,]}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \phi_{3i}, \nu_3^*, \xi_{3i}, \nu_2^*, \mathbf{Y}$
  $\sim Gamma\left(\frac{S+\nu_2^*}{2}, \frac{\nu_2^*}{2} + \frac{1}{2}\mathbf{E}_{2[i,]}\boldsymbol{\Sigma}_2^{-1}\mathbf{E}'_{2[i,]}\right)$
  where $\mathbf{E}_{2[i,]} = \mathbf{M}_{[i,]} - \mathbf{a}'_2 - \mathbf{H}_{[i,]}\mathbf{B}_2 - \mathbf{X}_{[i,]}\mathbf{W}\mathbf{A}_2$ for $i = 1, \cdots, N$.
- $\nu_2^* \mid \mathbf{Z}, \mathbf{X}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \phi_3, \boldsymbol{\xi}_3, \nu_3^*, \boldsymbol{\xi}_2, \mathbf{Y}$
  $\sim \pi(\nu_2^* | \mathbf{Z}, \mathbf{X}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \phi_3, \boldsymbol{\xi}_3, \nu_3^*, \boldsymbol{\xi}_2, \mathbf{Y})$

$$\propto \frac{\left(\frac{\nu_2^*}{2}\right)^{\frac{\nu_2^*}{2}N}}{\Gamma\left(\frac{\nu_2^*}{2}\right)^N}\left(\prod_{i=1}^{N}\xi_{2i}\right)^{\frac{\nu_2^*}{2}-1}\exp\left(-\frac{\nu_2^*}{2}\sum_{i=1}^{N}\xi_{2i}\right)$$

  $\equiv p\left(\nu_2^{*(t)}\right)$ (target distribution at tth iteration)
  Metropolis step

  - ♦ generate $\tilde{\nu}_2^{*-1} \sim N\left(\frac{1}{\nu_2^{*(t-1)}}, 1\right)$ where $N\left(\tilde{\nu}_2^{*-1} \middle| \frac{1}{\nu_2^{*(t-1)}}, 1\right) \equiv q\left(\tilde{\nu}_2^{*-1} | \nu_2^{*(t-1)}\right)$ is a candidate distribution function of ν_2^*.
  - ♦ compute acceptance probability $A\left(\tilde{\nu}_2^{*-1}, \nu_2^{*(t-1)}\right)$ such that

$$A\left(\tilde{\nu}_2^{*-1}, \nu_2^{*(t-1)}\right) = \min\left(1, \frac{p\left(\tilde{\nu}_2^*\right)}{p\left(\nu_2^{*(t-1)}\right)}\right).$$

  - ♦ set $\nu_2^{*(t)} = \tilde{\nu}_2^*$ with probability $A\left(\tilde{\nu}_2^{*-1}, \nu_2^{*(t-1)}\right)$

- $\mathbf{A}'_{2[k,]} \mid \mathbf{A}_{2[-k,]}, \overline{\mathbf{W}}, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{B}_3, \mathbf{a}_2, \mathbf{a}_3, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Gamma}, \phi_3, \boldsymbol{\xi}_3, \nu_3^*, \boldsymbol{\xi}_2, \nu_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$

$$\sim N_S\left(\tilde{\mathbf{A}}'_{2[k,]}, \boldsymbol{\Sigma}_{\mathbf{A}_{2k}}\right)$$

  for $k = 1, \ldots, K$ where $\boldsymbol{\Sigma}_{\mathbf{A}_{2k}} = \left(\sum_{i=1}^{N}\xi_{2i}\mathbf{X}_{[i,]}\overline{\mathbf{W}}_{[,k]}\boldsymbol{\Sigma}_2^{-1}\overline{\mathbf{W}}'_{[,k]}\mathbf{X}'_{[i,]} + \frac{1}{\eta_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{A}_{2k}}\right)^{-1}$ and

$$\tilde{\mathbf{A}}'_{2[k,]} = \boldsymbol{\Sigma}_{\mathbf{A}_{2k}}\left(\sum_{i=1}^{N}\xi_{2i}\mathbf{X}_{[i,]}\overline{\mathbf{W}}_{[,k]}\boldsymbol{\Sigma}_2^{-1}\mathbf{M}^{-k\prime}_{[i,]} + \frac{1}{\eta_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{A}_{2k}}\mathbf{A}^{0\prime}_{2[k,]}\right).$$

- $\mathbf{B}'_{2[d,]} \mid \mathbf{B}_{2[-d,]}, \overline{\mathbf{W}}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_3, \mathbf{a}_2, \mathbf{a}_3, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Gamma}, \phi_3, \boldsymbol{\xi}_3, \nu_3^*, \boldsymbol{\xi}_2, \nu_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$

$$\sim N_S\left(\tilde{\mathbf{B}}'_{2[d,]}, \boldsymbol{\Sigma}_{\mathbf{B}_{2d}}\right)$$

  for $d = 1, \ldots, D$ where $\boldsymbol{\Sigma}_{\mathbf{B}_{2d}} = \left(\sum_{i=1}^{N}\xi_{2i}\mathbf{H}_{[i,d]}\boldsymbol{\Sigma}_2^{-1}\mathbf{H}'_{[i,d]} + \frac{1}{\delta_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{B}_{2d}}\right)^{-1}$ and
  $\tilde{\mathbf{B}}'_{2[d,]} = \boldsymbol{\Sigma}_{\mathbf{B}_{2d}}\left(\sum_{i=1}^{N}\xi_{2i}\mathbf{H}_{[i,d]}\boldsymbol{\Sigma}_2^{-1}\mathbf{M}^{-d\prime}_{[i,]} + \frac{1}{\delta_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{B}_{2d}}\mathbf{B}^{0\prime}_{2[d,]}\right)$ with $\mathbf{M}^{-d}_{[i,]} = \mathbf{M}_{[i,]} - \mathbf{a}'_2 - \mathbf{H}_{[i,-d]}\mathbf{B}_{2[-d,]} - \mathbf{X}_{[i,]}\mathbf{W}\mathbf{A}_2$.
- $\mathbf{a}_2 \mid \overline{\mathbf{W}}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{B}_3, \mathbf{a}_3, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Gamma}, \phi_3, \boldsymbol{\xi}_3, \nu_3^*, \boldsymbol{\xi}_2, \nu_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$

$$\sim N_S\left(\tilde{\mathbf{a}}_2, \left(\sum_{i=1}^{N}\xi_{2i}\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\sigma_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{a}_2}\right)^{-1}\right)$$

  where $\tilde{\mathbf{a}}_2 = \left(\sum_{i=1}^{N}\xi_{2i}\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\sigma_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{a}_2}\right)^{-1}\left(\boldsymbol{\Sigma}_2^{-1}\sum_{i=1}^{N}\xi_{2i}\left(\mathbf{M}_{[i,]} - \mathbf{H}_{[i,]}\mathbf{B}_2 - \mathbf{X}_{[i,]}\overline{\mathbf{W}}\mathbf{A}_2\right)' + \frac{1}{\sigma_2^2}\boldsymbol{\Sigma}^{0-1}_{\mathbf{a}_2}\boldsymbol{\alpha}_2^0\right).$

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

- $\boldsymbol{\Sigma}_2^{-1} \mid \mathbf{W}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{B}_3, \mathbf{a}_2, \mathbf{a}_3, \boldsymbol{\Sigma}_3, \boldsymbol{\Gamma}, \phi_3, \xi_3, \phi_3, \xi_3, v_3^*, \xi_2, v_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$
  $\sim Wishart_S\left(v_2 + N, \left(\boldsymbol{\Sigma}_{20} + \sum_{i=1}^{N} \xi_{2i} \mathbf{E}'_{2[i,]} \mathbf{E}_{2[i,]}\right)^{-1}\right)$ where

$$\mathbf{E}_{2[i,]} = \mathbf{M}_{[i,]} - \mathbf{a}'_2 - \mathbf{H}_{[i,]} \mathbf{B}_2 - \mathbf{X}_{[i,]} \mathbf{W} \mathbf{A}_2.$$

Outcome model parameters

- $\mathbf{Z}'_{[i,]} \mid \mathbf{X}_{[i,]}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \phi_{3i}, \xi_{3i}, v_3^*, \xi_{2i}, v_2^*, \mathbf{Y}$
  $\sim N_Q\left(\mathbf{a}_3 + \mathbf{B}_3{}'\mathbf{H}'_{[i,]} + \mathbf{A}_3{}'\mathbf{W}'\mathbf{X}'_{[i,]} + \mathbf{A}_4{}'\mathbf{M}'_{[i,]}, \mathbf{D}_{3i}^{1/2} \boldsymbol{\Sigma}_3 \mathbf{D}_{3i}^{1/2}\right)$
  The variance matrix of the latent variable of the mixed outcomes will have the form of $\mathbf{D}_{3i}^{1/2} \boldsymbol{\Sigma}_3 \mathbf{D}_{3i}^{1/2} \equiv \begin{pmatrix} \xi_{3i}^{-1} \boldsymbol{\Sigma}_3^{11} & \xi_{3i}^{-1/2} \tilde{\sigma} \phi_{3i}^{-1/2} \boldsymbol{\Sigma}_3^{12} \\ \xi_{3i}^{-1/2} \tilde{\sigma} \phi_{3i}^{-1/2} \boldsymbol{\Sigma}_3^{21} & \tilde{\sigma}^2 \phi_{3i}^{-1} \boldsymbol{\Sigma}_3^{22} \end{pmatrix}$ where $\boldsymbol{\Sigma}_3^{12\prime} = \boldsymbol{\Sigma}_3^{21}$. The continuous outcomes are observed and we set $\mathbf{Y}_{[i,1:T]} = \mathbf{Z}_{[i,1:T]}$, but for the ordinal outcome, we need to update the latent variable while considering the correlation with continuous outcomes. Thus, we consider the conditional distribution of the latent variable of the ordinal outcome given the continuous outcome, $\mathbf{Z}_{[i,(T+1):Q]} \big| \mathbf{Z}_{[i,1:T]}$, and it will have a multivariate normal distribution with mean vector

$$\tilde{\boldsymbol{\mu}}_i = \mathbf{a}_{3[(T+1):Q]} + \mathbf{B}'_{3[,(T+1):Q]} \mathbf{H}'_{[i,]} + \mathbf{A}_{3[,(T+1):Q]}{}' \mathbf{W}' \mathbf{X}'_{[i,]} + \mathbf{A}'_{4[,(T+1):Q]} \mathbf{M}'_{[i,]} + \xi_{3i}^{-1/2} \tilde{\sigma} \phi_{3i}^{-1/2} \boldsymbol{\Sigma}_3^{21} \boldsymbol{\Sigma}_3^{11^{-1}} \mathbf{E}'_{3[i,1:T]}$$

  where $\mathbf{E}_{3[i,1:T]} = \mathbf{Z}_{[i,1:T]} - \mathbf{a}'_{3[1:T]} - \mathbf{H}_{[i,]} \mathbf{B}_{3[,1:T]} - \mathbf{X}_{[i,]} \mathbf{W} \mathbf{A}_{3[,1:T]} - \mathbf{M}_{[i,]} \mathbf{A}_{4[,1:T]}$ and covariance matrix $\tilde{\boldsymbol{\Sigma}}_i = \tilde{\sigma}^2 \phi_{3i}^{-1} \left(\boldsymbol{\Sigma}_3^{22} - \boldsymbol{\Sigma}_3^{21} \boldsymbol{\Sigma}_3^{11^{-1}} \boldsymbol{\Sigma}_3^{12}\right) \equiv \tilde{\sigma}^2 \phi_{3i}^{-1} \tilde{\boldsymbol{\Sigma}}_3$. Thus we update the latent variable of the ordinal outcomes from the conditional distribution

$$\mathbf{Z}_{[i,(T+1):Q]} \big| \mathbf{Z}_{[i,1:T]} \sim N_{Q-T}\left(\tilde{\boldsymbol{\mu}}_i, \tilde{\boldsymbol{\Sigma}}_i\right)$$

  with $\mathbf{Z}_{[i,q]}$ is truncated at the left and right by γ_{j-1}^q and γ_j^q if $\mathbf{Y}_{[i,q]} = j$ for $i = 1, \ldots, N$, $j = 1, \ldots, C$ and $q = T+1, \ldots, Q$.
- $\gamma_j^q \mid \boldsymbol{\Gamma}_{[-j,-q]}, \mathbf{Z}, \mathbf{X}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \phi_3, \xi_3, v_3^*, \xi_2, v_2^*, \mathbf{Y}$
  $\sim U\left(\max\left\{\max\left(\mathbf{Z}_{[i,q]} : \mathbf{Y}_{[i,q]} = j\right), \gamma_{j-1}^q\right\}, \min\left\{\min\left(\mathbf{Z}_{[i,q]} : \mathbf{Y}_{[i,q]} = j+1\right), \gamma_j^q\right\}\right)$
- $\phi_{3i} \mid \mathbf{Z}_{[i,]}, \mathbf{X}_{[i,]}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \xi_{3i}, v_3^*, \xi_{2i}, v_2^*, \mathbf{Y}$
  $\sim \pi\left(\phi_{3i} | \mathbf{Z}_{[i,]}, \mathbf{X}_{[i,]}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \xi_{3i}, v_3^*, \xi_{2i}, v_2^*, \mathbf{Y}\right)$

$$\propto \phi_{3i}^{\frac{Q-T+\tilde{v}}{2}-1} \exp\left[-\phi_{3i}\left\{\frac{\tilde{v}}{2} + \frac{1}{2\tilde{\sigma}^2}\left(\mathbf{Z}'_{[i,(T+1):Q]} - \tilde{\boldsymbol{\mu}}_i\right)' \tilde{\boldsymbol{\Sigma}}_3^{-1} \left(\mathbf{Z}_{[i,(T+1):Q]} - \tilde{\boldsymbol{\mu}}_i\right)\right\}\right]$$

  $\equiv p\left(\phi_{3i}^{(t)}\right)$ (target distribution at tth iteration) for $i = 1, \cdots, N$
  Metropolis–Hastings step

  - ♦ generate $\phi_{3i}^* \sim U\left(\phi_{3i}^{(t-1)} - \delta, \phi_{3i}^{(t-1)} + \delta\right)$ where $U\left(\phi_{3i}^{(t-1)} - \delta, \phi_{3i}^{(t-1)} + \delta\right)$ is a uniform distribution with small number of δ and it is a candidate distribution based on previous iteration sample $\phi_{3i}^{(t-1)}$, thus $U\left(\phi_{3i}^* | \phi_{3i}^{(t-1)} - \delta, \phi_{3i}^{(t-1)} + \delta\right) \equiv q\left(\phi_{3i}^* | \phi_{3i}^{(t-1)}\right)$ is a candidate distribution function of ϕ_{3i}.
  - ♦ compute acceptance probability $\mathrm{A}\left(\phi_{3i}^*, \phi_{3i}^{(t-1)}\right)$ such that

$$\mathrm{A}\left(\phi_{3i}^*, \phi_{3i}^{(t-1)}\right) = \min\left(1, \frac{p\left(\phi_{3i}^*\right) q\left(\phi_{3i}^* | \phi_{3i}^{(t-1)}\right)}{p\left(\phi_{3i}^{(t-1)}\right) q\left(\phi_{3i}^{(t-1)} | \phi_{3i}^*\right)}\right).$$

  - ♦ set $\phi_{3i}{}^{(t)} = \phi_{3i}^*$ with probability $\mathrm{A}\left(\phi_{3i}^*, \phi_{3i}^{(t-1)}\right)$

- $\xi_{3i} \mid \mathbf{Z}_{[i,]}, \mathbf{X}_{[i,]}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \phi_{3i}, v_3^*, \xi_{2i}, v_2^*, \mathbf{Y}$
  $\sim \pi\left(\xi_{3i} | \mathbf{Z}_{[i,]}, \mathbf{X}_{[i,]}, \mathbf{W}, \mathbf{A}_3, \mathbf{H}, \mathbf{B}_3, \mathbf{M}, \mathbf{A}_4, \boldsymbol{\Sigma}_3, \mathbf{a}_3, \boldsymbol{\Gamma}, \phi_{3i}, v_3^*, \xi_{2i}, v_2^*, \mathbf{Y}\right)$
  $\propto \xi_{3i}^{\frac{T+v_3^*}{2}-1} \exp\left[-\xi_{3i}\left\{\frac{v_3^*}{2} + \frac{1}{2}\left(\mathbf{Z}'_{[i,1:T]} - \boldsymbol{\mu}_i^*\right)' \boldsymbol{\Sigma}_3^{*-1} \left(\mathbf{Z}_{[i,(T+1):Q]} - \boldsymbol{\mu}_i^*\right)\right\}\right]$
  $\equiv p\left(\xi_{3i}^{(t)}\right)$ (target distribution at tth iteration)

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

where $\mathbf{Z}_{[i,1:T]}\big|\mathbf{Z}_{[i,(T+1):Q]}$, and it will have a multivariate normal distribution with a mean vector

$$\boldsymbol{\mu}_i^* = \mathbf{a}_{3[1:T]} + \mathbf{B}'_{3[,1:T]}\mathbf{H}'_{[i,]} + \mathbf{A}'_{3[,1:T]}\mathbf{W}'\mathbf{X}'_{[i,]} + \mathbf{A}'_{4[,1:T]}\mathbf{M}'_{[i,]} + \xi_{3i}^{-1/2}\tilde{\sigma}\phi_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{21}\boldsymbol{\Sigma}_3^{22-1}\mathbf{E}'_{3[i,(T+1):Q]}$$

with $\mathbf{E}_{3[i,(T+1):Q]} = \mathbf{Z}_{[i,(T+1):Q]} - \mathbf{a}'_{3[(T+1):Q]} - \mathbf{H}_{[i,]}\mathbf{B}_{3[,(T+1):Q]} - \mathbf{X}_{[i,]}\mathbf{W}\mathbf{A}_{3[,(T+1):Q]} - \mathbf{M}_{[i,]}\mathbf{A}_{4[,(T+1):Q]}$ and covariance matrix $\boldsymbol{\Sigma}_3^{**} = \xi_{3i}^{-1}\left(\boldsymbol{\Sigma}_3^{11} - \boldsymbol{\Sigma}_3^{21}\boldsymbol{\Sigma}_3^{22-1}\boldsymbol{\Sigma}_3^{12}\right) \equiv \xi_{3i}^{-1}\boldsymbol{\Sigma}_3^*$ for $i = 1,\cdots,N$.

Metropolis–Hastings step

- generate $\xi_{3i}^* \sim U\left(\xi_{3i}^{(t-1)} - \delta^*, \xi_{3i}^{(t-1)} + \delta^*\right)$ where $U\left(\xi_{3i}^{(t-1)} - \delta^*, \xi_{3i}^{(t-1)} + \delta^*\right)$ is a uniform distribution with small number of δ^* and it is a candidate distribution based on previous iteration sample $\xi_{3i}^{(t-1)}$, thus $U\left(\xi_{3i}^{(t-1)} - \delta^*, \xi_{3i}^{(t-1)} + \delta^*\right) \equiv q\left(\xi_{3i}^*|\xi_{3i}^{(t-1)}\right)$ is a candidate distribution function of ξ_{3i}.
- compute acceptance probability $\mathrm{A}\left(\xi_{3i}^*, \xi_{3i}^{(t-1)}\right)$ such that

$$\mathrm{A}\left(\xi_{3i}^*, \xi_{3i}^{(t-1)}\right) = \min\left(1, \frac{p\left(\xi_{3i}^*\right)q\left(\xi_{3i}^*|\xi_{3i}^{(t-1)}\right)}{p\left(\xi_{3i}^{(t-1)}\right)q\left(\xi_{3i}^{(t-1)}|\xi_{3i}^*\right)}\right).$$

- set $\xi_{3i}{}^{(t)} = \xi_{3i}^*$ with probability$\mathrm{A}\left(\xi_{3i}^*, \xi_{3i}^{(t-1)}\right)$

- $v_3^* \mid \mathbf{Z},\mathbf{X},\mathbf{W},\mathbf{A}_3,\mathbf{H},\mathbf{B}_3,\mathbf{M},\mathbf{A}_4,\boldsymbol{\Sigma}_3,\mathbf{a}_3,\boldsymbol{\Gamma},\phi_3,\boldsymbol{\xi}_3,\boldsymbol{\xi}_2,v_2^*,\mathbf{Y}$
  $\sim \pi\left(v_3^*|\mathbf{Z},\mathbf{X},\mathbf{W},\mathbf{A}_3,\mathbf{H},\mathbf{B}_3,\mathbf{M},\mathbf{A}_4,\boldsymbol{\Sigma}_3,\mathbf{a}_3,\boldsymbol{\Gamma},\xi_{3i},\phi_{3i},\tau^2,\eta^2,\sigma^2,\mathbf{Y}\right)$

$$\propto \frac{\left(\frac{v_3^*}{2}\right)^{\frac{v_3^*}{2}N}}{\Gamma\left(\frac{v_3^*}{2}\right)^N}\left(\prod_{i=1}^N \xi_{3i}\right)^{\frac{v_3^*}{2}-1}\exp\left(-\frac{v_3^*}{2}\sum_{i=1}^N \xi_{3i}\right)$$

$\equiv p\left(v_3^{*\,(t)}\right)$ (target distribution at tth iteration)

Metropolis step

- generate $\tilde{v}_3^{*-1} \sim N\left(\frac{1}{v_3^{*\,(t-1)}}, 1\right)$ where $N\left(\tilde{v}_3^{*-1}\middle|\frac{1}{v_3^{*\,(t-1)}}, 1\right) \equiv q\left(\tilde{v}_3^{*-1}|v_3^{*\,(t-1)}\right)$ is a candidate distribution function of v_3^*.
- compute acceptance probability $A\left(\tilde{v}_3^{*-1}, v_3^{*\,(t-1)}\right)$ such that
  $A\left(\tilde{v}_3^{*-1}, v_3^{*\,(t-1)}\right) = \min\left(1, \frac{p\left(\tilde{v}_3^*\right)}{p\left(v_3^{*\,(t-1)}\right)}\right).$
- set $v_3^{*\,(t)} = \tilde{v}_3^*$ with probability $A\left(\tilde{v}_3^{*-1}, v_3^{*\,(t-1)}\right)$.

- $\mathbf{A}_{3[k,]}{}' \mid \mathbf{A}_{3[-k,]},\overline{\mathbf{W}},\mathbf{A}_2,\mathbf{A}_4,\mathbf{B}_2,\mathbf{B}_3,\mathbf{a}_2,\mathbf{a}_3,\boldsymbol{\Sigma}_2,\boldsymbol{\Sigma}_3,\boldsymbol{\Gamma},\phi_3,\boldsymbol{\xi}_3,v_3^*,\boldsymbol{\xi}_2,v_2^*,\mathbf{Z},\mathbf{X},\mathbf{M},\mathbf{Y}$

$$\sim N_Q\left(\tilde{\mathbf{A}}'_{3[k,]}, \boldsymbol{\Sigma}_{\mathbf{A}_{3k}}\right)$$

for $k = 1,\ldots,K$ where $\boldsymbol{\Sigma}_{\mathbf{A}_{3k}} = \left(\sum_{i=1}^N \mathbf{X}_{[i,]}\overline{\mathbf{W}}_{[,k]}\mathbf{D}_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{-1}\mathbf{D}_{3i}^{-1/2}\overline{\mathbf{W}}'_{[,k]}\mathbf{X}'_{[i,]} + \frac{1}{\eta_3^2}\boldsymbol{\Sigma}_{\mathbf{A}_{3k}}^{0-1}\right)^{-1}$ and $\tilde{\mathbf{A}}'_{3[k,]} = \boldsymbol{\Sigma}_{\mathbf{A}_{3k}}\left(\sum_{i=1}^N \mathbf{X}_{[i,]}\overline{\mathbf{W}}_{[,k]}\mathbf{D}_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{-1}\mathbf{D}_{3i}^{-1/2}\mathbf{Z}_{[i,]}^{-k\prime} + \frac{1}{\eta_3^2}\boldsymbol{\Sigma}_{\mathbf{A}_{3k}}^{0-1}\mathbf{A}_{3[k,]}^{0\prime}\right)$.

- $\mathbf{A}'_{4[s,]} \mid \mathbf{A}_{4[-s,]},\overline{\mathbf{W}},\mathbf{A}_2,\mathbf{A}_3,\mathbf{B}_2,\mathbf{B}_3,\mathbf{a}_2,\mathbf{a}_3,\boldsymbol{\Sigma}_2,\boldsymbol{\Sigma}_3,\boldsymbol{\Gamma},\phi_3,\boldsymbol{\xi}_3,v_3^*,\boldsymbol{\xi}_2,v_2^*,\mathbf{Z},\mathbf{X},\mathbf{M},\mathbf{Y}$

$$\sim N_Q\left(\tilde{\mathbf{A}}'_{4[s,]}, \boldsymbol{\Sigma}_{\mathbf{A}_{4s}}\right)$$

for $s = 1,\ldots,S$ where $\boldsymbol{\Sigma}_{\mathbf{A}_{4s}} = \left(\sum_{i=1}^N \mathbf{M}_{[i,s]}\mathbf{D}_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{-1}\mathbf{D}_{3i}^{-1/2}\mathbf{M}'_{[i,s]} + \frac{1}{\eta_4^2}\boldsymbol{\Sigma}_{\mathbf{A}_{4s}}^{0-1}\right)^{-1}$ and $\tilde{\mathbf{A}}'_{4[s,]} = \boldsymbol{\Sigma}_{\mathbf{A}_{3k}}\left(\sum_{i=1}^N \mathbf{M}_{[i,s]}\mathbf{D}_{3i}^{-1/2}\boldsymbol{\Sigma}_3^{-1}\mathbf{D}_{3i}^{-1/2}\mathbf{Z}_{[i,]}^{-s\prime} + \frac{1}{\eta_4^2}\boldsymbol{\Sigma}_{\mathbf{A}_{4s}}^{0-1}\mathbf{A}_{4[s,]}^{0\prime}\right)$ with $\mathbf{Z}_{[i,]}^{-s} = \mathbf{Z}_{[i,]} - \mathbf{a}'_3 - \mathbf{H}_{[i,]}\mathbf{B}_3 - \mathbf{X}_{[i,]}\mathbf{W}\mathbf{A}_3 - \mathbf{M}_{[i,-s]}\mathbf{A}_{4[-s,]}$

https://doi.org/10.1017/psy.2024.13 Published online by Cambridge University Press

- $\mathbf{B}'_{3[d,]} \mid \mathbf{B}_{3[-d,]}, \overline{\mathbf{W}}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{a}_2, \mathbf{a}_3, \mathbf{\Sigma}_2, \mathbf{\Sigma}_3, \mathbf{\Gamma}, \phi_3, \boldsymbol{\xi}_3, v_3^*, \boldsymbol{\xi}_2, v_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$

$$\sim N_Q\left(\tilde{\mathbf{B}}'_{3[d,]}, \mathbf{\Sigma}_{\mathbf{B}_{3d}}\right)$$

for $d = 1, \ldots, D$ where $\mathbf{\Sigma}_{\mathbf{B}_{2d}} = \left(\sum_{i=1}^{N} \mathbf{H}_{[i,d]} \mathbf{D}_{3i}^{-1/2} \mathbf{\Sigma}_3^{-1} \mathbf{D}_{3i}^{-1/2} \mathbf{H}'_{[i,d]} + \frac{1}{\delta_3^2} \mathbf{\Sigma}_{\mathbf{B}_{3d}}^{0-1}\right)^{-1}$ and

$\tilde{\mathbf{B}}'_{3[d,]} = \mathbf{\Sigma}_{\mathbf{B}_{3d}} \left(\sum_{i=1}^{N} \mathbf{H}_{[i,d]} \mathbf{D}_{3i}^{-1/2} \mathbf{\Sigma}_3^{-1} \mathbf{D}_{3i}^{-1/2} \mathbf{Z}_{[i,]}^{-d\prime} + \frac{1}{\delta_3^2} \mathbf{\Sigma}_{\mathbf{B}_{3d}}^{0-1} \mathbf{B}_{3[d,]}^{0\prime}\right)$ with $\mathbf{Z}_{[i,]}^{-d} = \mathbf{Z}_{[i,]} - \mathbf{a}'_3 - \mathbf{H}_{[i,-d]} \mathbf{B}_{3[-d,]} - \mathbf{X}_{[i,]} \mathbf{W} \mathbf{A}_3 - \mathbf{M}_{[i,]} \mathbf{A}_4$

$\mathbf{a}_3 \mid \overline{\mathbf{W}}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{B}_3, \mathbf{a}_2, \mathbf{\Sigma}_2, \mathbf{\Sigma}_3, \mathbf{\Gamma}, \phi_3, \boldsymbol{\xi}_3, v_3^*, \boldsymbol{\xi}_2, v_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$

$$\sim N_Q\left(\tilde{\boldsymbol{a}}_3, \left(\sum_{i=1}^{N} \boldsymbol{D}_{3i}^{-1/2} \boldsymbol{\Sigma}_3^{-1} \boldsymbol{D}_{3i}^{-1/2} + \frac{1}{\sigma_3^2} \boldsymbol{\Sigma}_{\boldsymbol{a}_3}^{0-1}\right)^{-1}\right)$$

- where $\tilde{\mathbf{a}}_3 = \mathbf{\Sigma}_{\mathbf{a}_3} \left(\sum_{i=1}^{N} \mathbf{D}_{3i}^{-1/2} \mathbf{\Sigma}_3^{-1} \mathbf{D}_{3i}^{-1/2} \left(\mathbf{Z}_{[i,]} - \mathbf{H}_{[i,]} \mathbf{B}_3 - \mathbf{X}_{[i,]} \overline{\mathbf{W}} \mathbf{A}_3 - \mathbf{M}_{[i,]} \mathbf{A}_4\right)' + \frac{1}{\sigma_3^2} \mathbf{\Sigma}_{\mathbf{a}_3}^{0-1} \boldsymbol{\alpha}_3^0\right)$.
- $\mathbf{\Sigma}_3^{-1} \mid \overline{\mathbf{W}}, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{B}_2, \mathbf{B}_3, \mathbf{a}_2, \mathbf{a}_3, \mathbf{\Sigma}_2, \mathbf{\Gamma}, \phi_3, \boldsymbol{\xi}_3, v_3^*, \boldsymbol{\xi}_2, v_2^*, \mathbf{Z}, \mathbf{X}, \mathbf{M}, \mathbf{Y}$
  $\sim Wishart_Q\left(v_3 + N, \left(\mathbf{\Sigma}_{30} + \sum_{i=1}^{N} \mathbf{E}_{3[i,]}^{*\prime} \mathbf{E}_{3[i,]}^*\right)^{-1}\right)$ where
  $\mathbf{E}_{3[i,]}^* = \left(\mathbf{Z}_{[i,]} - \mathbf{a}'_3 - \mathbf{H}_{[i,]} \mathbf{B}_3 - \mathbf{X}_{[i,]} \mathbf{W} \mathbf{A}_3 - \mathbf{M}_{[i,]} \mathbf{A}_4\right) \mathbf{D}_{3i}^{-1/2}$.

Cite this article: Choi, J. Y., Kyung, M. and Park, J-H. (2025). Mediation Analysis in Bayesian Extended Redundancy Analysis with Mixed Outcome Variables. *Psychometrika*, 251–279. https://doi.org/10.1017/psy.2024.13