Appendix 5.2: Conditional Distributions and Implementation of MH Algorithm Related to SEMs with EFDs

It can be shown that the full conditional distribution of $\mathbf{\Omega}$ is given by

$$p(\mathbf{\Omega}|\mathbf{Y},\boldsymbol{\theta}) = \prod_{i=1}^{n} p(\boldsymbol{\omega}_i|\mathbf{y}_i,\boldsymbol{\theta}) \propto \prod_{i=1}^{n} p(\mathbf{y}_i|\boldsymbol{\omega}_i,\boldsymbol{\theta})p(\boldsymbol{\eta}_i|\boldsymbol{\xi}_i,\boldsymbol{\theta})p(\boldsymbol{\xi}_i|\boldsymbol{\theta}),$$

where $p(\boldsymbol{\omega}_i|\mathbf{y}_i,\boldsymbol{\theta})$ is proportional to

$$\exp\Bigg\{ \sum_{k=1}^{p} \Big[y_{ik}\vartheta_{ik} - b(\vartheta_{ik}) \Big] / \psi_{\epsilon k} \tag{5.A10}$$
$$-\frac{1}{2}\Big[(\boldsymbol{\eta}_i - \mathbf{B}\mathbf{d}_i - \mathbf{\Pi}\boldsymbol{\eta}_i - \mathbf{\Gamma}\mathbf{F}(\boldsymbol{\xi}_i))^T \mathbf{\Psi}_\delta^{-1} (\boldsymbol{\eta}_i - \mathbf{B}\mathbf{d}_i - \mathbf{\Pi}\boldsymbol{\eta}_i - \mathbf{\Gamma}\mathbf{F}(\boldsymbol{\xi}_i)) + \boldsymbol{\xi}_i^T \mathbf{\Phi}^{-1}\boldsymbol{\xi}_i \Big] \Bigg\}.$$

Under the conjugate prior distributions given in (5.14), it can be shown that the full conditional distributions of the components of $\boldsymbol{\theta}$ are given by

$$p(\mathbf{A}_k|\mathbf{Y},\mathbf{\Omega},\mathbf{\Lambda}_k,\psi_{\epsilon k}) \propto \exp\Bigg\{ \sum_{i=1}^{n} \frac{y_{ik}\vartheta_{ik} - b(\vartheta_{ik})}{\psi_{\epsilon k}} - \frac{1}{2}(\mathbf{A}_k - \mathbf{A}_{0k})^T \mathbf{H}_{0k}^{-1}(\mathbf{A}_k - \mathbf{A}_{0k}) \Bigg\},$$

$$p(\psi_{\epsilon k}|\mathbf{Y},\mathbf{\Omega},\mathbf{A}_k,\mathbf{\Lambda}_k) \propto \psi_{\epsilon k}^{-(\frac{n}{2}+\alpha_{0\epsilon k}-1)} \exp\Bigg\{ \sum_{i=1}^{n} \Big[\frac{y_{ik}\vartheta_{ik} - b(\vartheta_{ik})}{\psi_{\epsilon k}} + c_k(y_{ik},\psi_{\epsilon k}) \Big] - \frac{\beta_{0k}}{\psi_{\epsilon k}} \Bigg\},$$

$$p(\mathbf{\Lambda}_k|\mathbf{Y},\mathbf{\Omega},\mathbf{A}_k,\psi_{\epsilon k}) \propto \exp\Bigg\{ \sum_{i=1}^{n} \frac{y_{ik}\vartheta_{ik} - b(\vartheta_{ik})}{\psi_{\epsilon k}} - \frac{1}{2}\psi_{\epsilon k}^{-1}(\mathbf{\Lambda}_k - \mathbf{\Lambda}_{0k})^T \mathbf{H}_{0yk}^{-1}(\mathbf{\Lambda}_k - \mathbf{\Lambda}_{0k}) \Bigg\},$$

$$[\psi_{\delta k}^{-1}|\mathbf{\Omega},\mathbf{\Lambda}_{\omega k}] \stackrel{D}{=} Gamma[n/2 + \alpha_{0\delta k}, \beta_{\delta k}],$$

$$[\mathbf{\Lambda}_{\omega k}|\mathbf{\Omega},\psi_{\delta k}] \stackrel{D}{=} N[\boldsymbol{\mu}_{\omega k}, \psi_{\delta k}\mathbf{\Sigma}_{\omega k}],$$

$$[\mathbf{\Phi}|\mathbf{\Omega}] \stackrel{D}{=} IW_{q_2}[(\mathbf{\Omega}_2\mathbf{\Omega}_2^T + \mathbf{R}_0^{-1}), n+\rho_0], \tag{5.A11}$$

where $\mathbf{\Sigma}_{\omega k} = (\mathbf{H}_{0\omega k}^{-1} + \mathbf{G}\mathbf{G}^T)^{-1}$, $\boldsymbol{\mu}_{\omega k} = \mathbf{\Sigma}_{\omega k}(\mathbf{H}_{0\omega k}^{-1}\mathbf{\Lambda}_{0\omega k} + \mathbf{G}\mathbf{\Omega}_{1k})$, and $\beta_{\delta k} = \beta_{0\delta k} + (\mathbf{\Omega}_{1k}^T\mathbf{\Omega}_{1k} - \boldsymbol{\mu}_{\omega k}^T\mathbf{\Sigma}_{\omega k}^{-1}\boldsymbol{\mu}_{\omega k} + \mathbf{\Lambda}_{0\omega k}^T\mathbf{H}_{0\omega k}^{-1}\mathbf{\Lambda}_{0\omega k})/2$, in which $\mathbf{G} = (\mathbf{G}(\boldsymbol{\omega}_1),\cdots,\mathbf{G}(\boldsymbol{\omega}_n))$, $\mathbf{\Omega}_1 = (\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_n)$, $\mathbf{\Omega}_2 = (\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n)$, and $\mathbf{\Omega}_{1k}^T$ is the kth row of $\mathbf{\Omega}_1$.

In simulating observations from $p(\boldsymbol{\omega}_i|\mathbf{y}_i,\boldsymbol{\theta})$ in (5.A10), we choose $N[\cdot,\sigma_\omega^2\mathbf{\Omega}_\omega]$ as the proposal distribution in the MH algorithm, where $\mathbf{\Omega}_\omega^{-1} = \mathbf{\Sigma}_\omega^* + \mathbf{\Lambda}^T\mathbf{\Psi}_\omega\mathbf{\Lambda}$, in which

$$\mathbf{\Sigma}_\omega^* = \begin{bmatrix} \mathbf{\Pi}_0^T\mathbf{\Psi}_\delta^{-1}\mathbf{\Pi}_0 & -\mathbf{\Pi}_0^T\mathbf{\Psi}_\delta^{-1}\mathbf{\Gamma}\mathbf{\Delta} \\ -\mathbf{\Delta}^T\mathbf{\Gamma}^T\mathbf{\Psi}_\delta^{-1}\mathbf{\Pi}_0 & \mathbf{\Phi}^{-1} + \mathbf{\Delta}^T\mathbf{\Gamma}^T\mathbf{\Psi}_\delta^{-1}\mathbf{\Gamma}\mathbf{\Delta} \end{bmatrix},$$

with $\mathbf{\Pi}_0 = \mathbf{I}_{q_1} - \mathbf{\Pi}$, $\mathbf{\Delta} = (\partial \mathbf{F}(\boldsymbol{\xi}_i)/\partial \boldsymbol{\xi}_i)^T|_{\boldsymbol{\xi}_i=\mathbf{0}}$, and $\mathbf{\Psi}_\omega = \text{diag}(\ddot{b}(\vartheta_{i1})/\psi_{\epsilon 1}, \cdots, \ddot{b}(\vartheta_{ip})/\psi_{\epsilon p})|_{\boldsymbol{\omega}_i=\mathbf{0}}$.

In simulating observations from the conditional distributions $p(\mathbf{A}_k|\mathbf{Y}, \mathbf{\Omega}, \mathbf{\Lambda}_k, \psi_{\epsilon k})$, $p(\psi_{\epsilon k}\,|\mathbf{Y}, \mathbf{\Omega}, \mathbf{A}_k, \mathbf{\Lambda}_k)$, and $p(\mathbf{\Lambda}_k|\mathbf{Y}, \mathbf{\Omega}, \mathbf{A}_k, \psi_{\epsilon k})$, the proposal distributions are $N[\cdot, \sigma_a^2 \mathbf{\Omega}_{ak}]$, $N[\cdot, \sigma_\psi^2 \Omega_{\psi k}]$, and $N[\cdot, \sigma_\lambda^2 \mathbf{\Omega}_{\lambda k}]$, respectively, where

$$\mathbf{\Omega}_{ak}^{-1} = \sum_{i=1}^n \ddot{b}(\vartheta_{ik}) \mathbf{c}_{ik} \mathbf{c}_{ik}^T / \psi_{\epsilon k} \Big|_{\mathbf{A}_k=\mathbf{0}} + \mathbf{H}_{0k}^{-1},$$

$$\Omega_{\psi k}^{-1} = 1 - n/2 - \alpha_{0\epsilon k} - 2\sum_{i=1}^n [y_{ik}\vartheta_{ik} - b(\vartheta_{ik})] - \ddot{c}_k(y_{ik}, \psi_{\epsilon k}) \Big|_{\psi_{\epsilon k}=1} + 2\beta_{0\epsilon k},$$

$$\mathbf{\Omega}_{\lambda k}^{-1} = \sum_{i=1}^n \ddot{b}(\vartheta_{ik}) \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \Big|_{\mathbf{\Lambda}_k=\mathbf{0}} + \psi_{\epsilon k}^{-1} \mathbf{H}_{0yk}^{-1}.$$

For improving efficiency, we respectively use $N[\boldsymbol{\mu}_{ak}, \mathbf{\Omega}_{ak}]$, $N[\mu_{\psi k}, \mathbf{\Omega}_{\psi k}]$, and $N[\boldsymbol{\mu}_{\lambda k}, \mathbf{\Omega}_{\lambda k}]$ as initial proposal distributions in the first few iterations, where

$$\boldsymbol{\mu}_{ak} = \sum_{i=1}^n \left[y_{ik} - \dot{b}(\vartheta_{ik})|_{\mathbf{A}_k=\mathbf{0}} \right] \frac{\mathbf{c}_{ik}}{\psi_{\epsilon k}} + \mathbf{H}_{0k}^{-1} \mathbf{A}_{0k},$$

$$\mu_{\psi k} = 1 - n/2 - \alpha_{0\epsilon k} - \sum_{i=1}^n \left[y_{ik}\vartheta_{ik} - b(\vartheta_{ik}) \right] + \dot{c}_k(y_{ik}, \psi_{\epsilon k}) \Big|_{\psi_{\epsilon k}=1} + \beta_{0\epsilon k},$$

$$\boldsymbol{\mu}_{\lambda k} = \sum_{i=1}^n \left[y_{ik} - \dot{b}(\vartheta_{ik})|_{\Lambda_k=\mathbf{0}} \right] \frac{\boldsymbol{\omega}_i}{\psi_{\epsilon k}} + \mathbf{H}_{0yk}^{-1} \mathbf{\Lambda}_{0k}.$$

Let $\mathbf{y}_k^{*T}$ be the kth row of $\mathbf{Y}$ that is not directly observable, $\mathbf{z}_k$ be the corresponding ordered categorical vector, and $\boldsymbol{\alpha}_k = (\alpha_{k,2}, \cdots, \alpha_{k,b_k-1})$. It can be shown by similar derivation as in Appendix 5.1 that

$$p(\boldsymbol{\alpha}_k, \mathbf{y}_k^* | \mathbf{z}_k, \mathbf{\Omega}, \boldsymbol{\theta}) = p(\boldsymbol{\alpha}_k | \mathbf{z}_k, \mathbf{\Omega}, \boldsymbol{\theta}) p(\mathbf{y}_k^* | \boldsymbol{\alpha}_k, \mathbf{z}_k, \mathbf{\Omega}, \boldsymbol{\theta}) \propto$$
$$\prod_{i=1}^n \exp\left\{ [y_{ik}^* \vartheta_{ik} - b(\vartheta_{ik})]/\psi_{\epsilon k} + c_k(y_{ik}^*, \psi_{\epsilon k}) \right\} I_{[\alpha_{k,z_{ik}}, \alpha_{k,z_{ik}+1})}(y_{ik}^*), \quad (5.A12)$$

where $I_A(y)$ is an indicator function which takes 1 if $y \in A$, and 0 otherwise. The treatment of dichotomous variables is similar.

A multivariate version of the MH algorithm is used to simulate observations from $p(\boldsymbol{\alpha}_k, \mathbf{y}_k^* |\, \mathbf{z}_k, \mathbf{\Omega}, \boldsymbol{\theta})$ in (5.A12). Following Cowles (1996), for the joint proposal distribution of $\boldsymbol{\alpha}_k$ and $\mathbf{y}_k^*$ given $\mathbf{z}_k$, $\boldsymbol{\Omega}$, and $\boldsymbol{\theta}$ can be constructed according to the factorization $p(\boldsymbol{\alpha}_k, \mathbf{y}_k^*|\mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta}) = p(\boldsymbol{\alpha}_k|\mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta})p(\mathbf{y}_k^*|\boldsymbol{\alpha}_k, \mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta})$. At the jth iteration, we generate a candidate vector of thresholds $(\alpha_{k,2}, \cdots, \alpha_{k,b_k-1})$ from the following univariate truncated normal distribution

$$\alpha_{k,m} \sim N[\alpha_{k,m}^{(j)}, \sigma_{\alpha_k}^2]I_{(\alpha_{k,m-1}, \alpha_{k,m+1}^{(j)}]}(\alpha_{k,m}), \quad \text{for} \quad m = 2, \cdots, b_k - 1,$$

where $\alpha_{k,m}^{(j)}$ is the current value of $\alpha_{k,m}$, and $\sigma_{\alpha_k}^2$ is chosen to obtain an average acceptance rate of approximately 0.25 or greater. The acceptance probability for a candidate vector $(\boldsymbol{\alpha}_k, \mathbf{y}_k^*)$ as a new observation $(\boldsymbol{\alpha}_k^{(j+1)}, \mathbf{y}_k^{*(j+1)})$ is $\min\{1, R_k\}$, where

$$R_k = \frac{p(\boldsymbol{\alpha}_k, \mathbf{y}_k^*|\mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta})p(\boldsymbol{\alpha}_k^{(j)}, \mathbf{y}_k^{*(j)}|\boldsymbol{\alpha}_k, \mathbf{y}_k^*, \mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta})}{p(\boldsymbol{\alpha}_k^{(j)}, \mathbf{y}_k^{*(j)}|\mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta})p(\boldsymbol{\alpha}_k, \mathbf{y}_k^*|\boldsymbol{\alpha}_k^{(j)}, \mathbf{y}_k^{*(j)}, \mathbf{z}_k, \boldsymbol{\Omega}, \boldsymbol{\theta})}.$$

For an accepted $\boldsymbol{\alpha}_k$, a new $\mathbf{y}_k^*$ is simulated from the following univariate truncated distribution:

$$[y_{ik}^*|\boldsymbol{\alpha}_k, z_{ik}, \boldsymbol{\omega}_i, \boldsymbol{\theta}] \stackrel{D}{=} \exp\{[y_{ik}^*\vartheta_{ik} - b(\vartheta_{ik})]/\psi_{\epsilon k} + c_k(y_{ik}^*, \psi_{\epsilon k})\}I_{[\alpha_{k,z_{ik}}, \alpha_{k,z_{ik}+1})}(y_{ik}^*),$$

where y_{ik}^* and z_{ik} are the ith components of $\mathbf{y}_k^*$ and $\mathbf{z}_k$, respectively, and $I_A(y)$ is an indicator function which takes 1 if y in A and zero otherwise.

Appendix 5.3: WinBUGS Code Related to Section 5.3.4

```
model {
    for(i in 1:N){
        #Measurement equation model
        for(j in 1:3){
            y[i,j]~dnorm(mu[i,j],1)I(low[z[i,j]+1],high[z[i,j]+1])
        }
        for(j in 4:P){
            z[i,j]~dbin(pb[i,j],1)
            pb[i,j]<-exp(mu[i,j])/(1+exp(mu[i,j]))
        }
        mu[i,1]<-uby[1]+eta[i]
        mu[i,2]<-uby[2]+lam[1]*eta[i]
```