```
}

    psd~dgamma(10.0,8.0)
    sgd<-1/psd

    phi1[1:2,1:2]~dwish(R1[1:2,1:2],5)
    phx1[1:2,1:2]<-inverse(phi1[1:2,1:2])

    phi2[1:3,1:3]~dwish(R2[1:2,1:2],5)
    phx2[1:3,1:3]<-inverse(phi2[1:3,1:3])

} # end of model
```

Appendix 6.5: Conditional Distributions: Multisample SEMs

The conditional distributions $[\boldsymbol{\Omega}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \mathbf{Y}, \mathbf{X}, \mathbf{Z}]$, $[\boldsymbol{\alpha}, \mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{X}, \mathbf{Z}]$, and $[\boldsymbol{\theta}|\boldsymbol{\alpha}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}, \mathbf{Z}]$ that are required in the implementation of the Gibbs sampler are presented in this appendix. Note that the results on the first two conditional distributions are natural extension of those given in Section 5.2, but they can be regarded as the special cases of those given in Section 6.2. Also note that we allow common parameters in $\boldsymbol{\theta}$ according to the constraints under the competing models.

Conditional distribution of $[\boldsymbol{\Omega}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \mathbf{Y}, \mathbf{X}, \mathbf{Z}]$ can be obtained as below:

$$p[\boldsymbol{\Omega}|\boldsymbol{\theta}, \boldsymbol{\alpha}, \mathbf{Y}, \mathbf{X}, \mathbf{Z}] = \prod_{g=1}^{G} \prod_{i=1}^{N_g} p(\boldsymbol{\omega}_i^{(g)}|\mathbf{v}_i^{(g)}, \boldsymbol{\theta}^{(g)}),$$

where

$$\begin{aligned} p(\boldsymbol{\omega}_i^{(g)}|\mathbf{v}_i^{(g)}, \boldsymbol{\theta}^{(g)}) \propto \exp\Bigg\{ -\frac{1}{2} \Big[(\mathbf{v}_i^{(g)} - \boldsymbol{\mu}^{(g)} - \boldsymbol{\Lambda}^{(g)}\boldsymbol{\omega}_i^{(g)})^T {\boldsymbol{\Psi}_\epsilon^{(g)}}^{-1} (\mathbf{v}_i^{(g)} - \boldsymbol{\mu}^{(g)} - \boldsymbol{\Lambda}^{(g)}\boldsymbol{\omega}_i^{(g)}) \\ + \ (\boldsymbol{\eta}_i^{(g)} - \boldsymbol{\Lambda}_\omega^{(g)}\mathbf{G}(\boldsymbol{\omega}_i^{(g)}))^T {\boldsymbol{\Psi}_\delta^{(g)}}^{-1} (\boldsymbol{\eta}_i^{(g)} - \boldsymbol{\Lambda}_\omega^{(g)}\mathbf{G}(\boldsymbol{\omega}_i^{(g)})) + {\boldsymbol{\xi}_i^{(g)}}^T {\boldsymbol{\Phi}^{(g)}}^{-1} \boldsymbol{\xi}_i^{(g)} \Big] \Bigg\}. \end{aligned} \tag{6.A10}$$

Since the conditional distribution of (6.A10) is not standard, the Metropolis-Hastings (MH) algorithm can be used to draw random observations from this distribution.

Under the multisample situation, the notation in the conditional distribution $[\boldsymbol{\alpha}, \mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{X}, \mathbf{Z}]$ is very tedious. The derivation is similar to the two-level case as given in Appendix 6.1, Equations (6.A3) and (6.A4). As $(\boldsymbol{\alpha}^{(g)}, \mathbf{Y}^{(g)})$ is independent with $(\boldsymbol{\alpha}^{(h)}, \mathbf{Y}^{(h)})$ for $g \neq h$, and $\boldsymbol{\Psi}_\delta^{(g)}$ is diagonal,

$$p(\boldsymbol{\alpha}, \mathbf{Y}|\cdot) = \prod_{g=1}^{G} p(\boldsymbol{\alpha}^{(g)}, \mathbf{Y}^{(g)}|\cdot) = \prod_{g=1}^{G}\prod_{k=1}^{s} p(\boldsymbol{\alpha}_k^{(g)}, \mathbf{Y}_k^{(g)}|\cdot), \tag{6.A11}$$

where $\mathbf{Y}_k^{(g)} = (y_{1k}^{(g)}, \cdots, y_{N_g k}^{(g)})^T$. Let $\boldsymbol{\Psi}_{\epsilon y}^{(g)}$, $\boldsymbol{\Lambda}_y^{(g)}$, and $\boldsymbol{\mu}_y^{(g)}$ be the submatrices and subvector of $\boldsymbol{\Psi}_\epsilon^{(g)}$, $\boldsymbol{\Lambda}^{(g)}$, and $\boldsymbol{\mu}^{(g)}$ corresponding to the ordered categorical variables in $\mathbf{Y}$; let $\psi_{\epsilon yk}^{(g)}$ be the kth diagonal element of $\boldsymbol{\Psi}_{\epsilon y}^{(g)}$, $\mu_{yk}^{(g)}$ be the kth element of $\boldsymbol{\mu}_y^{(g)}$, and $\boldsymbol{\Lambda}_{yk}^{(g)^T}$ be the kth row of $\boldsymbol{\Lambda}_y^{(g)}$, and $I_A(y)$ be an indicator function with value 1 if y in A and zero otherwise, $p(\boldsymbol{\alpha}, \mathbf{Y}|\cdot)$ can be obtained from (6.A11) and

$$p(\boldsymbol{\alpha}_k^{(g)}, \mathbf{Y}_k^{(g)}|\cdot) \propto \prod_{i=1}^{N_g} \phi\{\psi_{\epsilon yk}^{(g)-1/2}(y_{ik}^{(g)} - \mu_{yk}^{(g)} - \boldsymbol{\Lambda}_{yk}{}^{(g)T}\boldsymbol{\omega}_i^{(g)})\} I_{[\alpha_{k,z_{ik}}, \alpha_{k,z_{ik}+1}]}(y_{ik}^{(g)}), \tag{6.A12}$$

where ϕ is the probability density function of $N[0,1]$. Note that in (6.A12) the superscript '(g)' in the threshold is suppressed to simplify notation.

Under the prior distributions of components in $\boldsymbol{\theta}$ as given in Section 6.3, the conditional distribution $[\boldsymbol{\theta}|\boldsymbol{\alpha}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}, \mathbf{Z}]$ is presented. Note that as $\mathbf{Y}$ is given, the model is defined with continuous data; hence, the conditional distribution is independent of $\boldsymbol{\alpha}$ and $\mathbf{Z}$.

The conditional distribution of some components in $\boldsymbol{\theta}^{(g)}$, $g = 1, \cdots, G$, under the situation without any parameter constraints are given as follows. Let $\boldsymbol{\Lambda}_k^{(g)^T}$ be the kth row of $\boldsymbol{\Lambda}^{(g)}$, and $\psi_{\epsilon k}^{(g)}$ be the kth diagonal element of $\boldsymbol{\Psi}_\epsilon^{(g)}$, $\mathbf{V}_k^{*(g)T}$ be the kth row of $\mathbf{V}^{*(g)} = (\mathbf{v}_1^{(g)} - \boldsymbol{\mu}^{(g)}, \cdots, \mathbf{v}_{N_g}^{(g)} - \boldsymbol{\mu}^{(g)})$, and $\boldsymbol{\Omega}_2^{(g)} = (\boldsymbol{\xi}_1^{(g)}, \cdots, \boldsymbol{\xi}_{N_g}^{(g)})$. It can be shown that:

$$\begin{aligned}
&[\boldsymbol{\mu}^{(g)}|\boldsymbol{\Lambda}^{(g)}, \boldsymbol{\Psi}_\epsilon^{(g)}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}] \overset{D}{=} N(\mathbf{a}_\mu^{(g)}, \mathbf{A}_\mu^{(g)}), \\
&[\boldsymbol{\Lambda}_k^{(g)}|\boldsymbol{\Psi}_\epsilon^{(g)}, \boldsymbol{\mu}^{(g)}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}] \overset{D}{=} N(\mathbf{a}_k^{(g)}, \mathbf{A}_k^{(g)}), \\
&[\psi_{\epsilon k}^{(g)-1}|\boldsymbol{\Lambda}_k^{(g)}, \boldsymbol{\mu}^{(g)}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}] \overset{D}{=} Gamma(N_g/2 + \alpha_{0\epsilon k}^{(g)}, \beta_{\epsilon k}^{(g)}), \\
&[\boldsymbol{\Phi}^{(g)}|\boldsymbol{\Omega}_2^{(g)}] \overset{D}{=} IW_{q_2}[(\boldsymbol{\Omega}_2^{(g)}\boldsymbol{\Omega}_2^{(g)T} + \mathbf{R}_0^{(g)-1}), N_g + \rho_0^{(g)}],
\end{aligned} \tag{6.A13}$$

in which

$$\mathbf{a}_\mu^{(g)} = \mathbf{A}_\mu^{(g)}\big[\boldsymbol{\Sigma}_0^{(g)-1}\boldsymbol{\mu}_0^{(g)} + N_g\boldsymbol{\Psi}_\epsilon^{(g)-1}(\bar{\mathbf{v}}^{(g)} - \boldsymbol{\Lambda}^{(g)}\bar{\boldsymbol{\omega}}^{(g)})\big],\ \mathbf{A}_\mu^{(g)} = \big(\boldsymbol{\Sigma}_0^{(g)-1} + N_g\boldsymbol{\Psi}_\epsilon^{(g)-1}\big)^{-1},$$

$$\mathbf{a}_k^{(g)} = \mathbf{A}_k^{(g)}\big[\mathbf{H}_{0yk}^{(g)-1}\boldsymbol{\Lambda}_{0k}^{(g)} + \psi_{\epsilon k}^{(g)-1}\boldsymbol{\Omega}^{(g)}\mathbf{V}_k^{*(g)}\big],\ \mathbf{A}_k^{(g)} = \big[\psi_{\epsilon k}^{(g)-1}\boldsymbol{\Omega}^{(g)}\boldsymbol{\Omega}^{(g)T} + \mathbf{H}_{0yk}^{(g)-1}\big]^{-1}$$

$$\beta_{\epsilon k}^{(g)} = \beta_{0\epsilon k}^{(g)} + \big[\boldsymbol{\Lambda}_k^{(g)T}\boldsymbol{\Omega}^{(g)}\boldsymbol{\Omega}^{(g)T}\boldsymbol{\Lambda}_k^{(g)} - 2\boldsymbol{\Lambda}_k^{(g)T}\boldsymbol{\Omega}^{(g)}\mathbf{V}_k^{*(g)} + \mathbf{V}_k^{*(g)T}\mathbf{V}_k^{*(g)}\big]/2,$$

with $\bar{\mathbf{v}}^{(g)} = \sum_{i=1}^{N_g}\mathbf{v}_i^{(g)}/N_g$ and $\bar{\boldsymbol{\omega}}^{(g)} = \sum_{i=1}^{N_g}\boldsymbol{\omega}_i^{(g)}/N_g$ are the means of $\mathbf{v}_i^{(g)}$ and $\boldsymbol{\omega}_i^{(g)}$ within the gth group.

As we mentioned, slight modifications are required to handle models with parameter constraints, see Section 6.3. Under the constraints $\boldsymbol{\Lambda}_k^{(1)} = \cdots = \boldsymbol{\Lambda}_k^{(G)} = \boldsymbol{\Lambda}_k$; the conjugate prior distribution of $\boldsymbol{\Lambda}_k$ is $N[\boldsymbol{\Lambda}_{0k}, \mathbf{H}_{0yk}]$, and the conditional distribution is

$$[\boldsymbol{\Lambda}_k|\psi_{\epsilon k}^{(1)}, \cdots, \psi_{\epsilon k}^{(G)}, \boldsymbol{\mu}^{(1)}, \cdots, \boldsymbol{\mu}^{(G)}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}] \overset{D}{=} N[\mathbf{a}_k, \mathbf{A}_k], \tag{6.A14}$$

where $\mathbf{a}_k = \mathbf{A}_k(\mathbf{H}_{0yk}^{-1}\boldsymbol{\Lambda}_{0k} + \sum_{g=1}^{G}\ \psi_{\epsilon k}^{(g)-1}\ \boldsymbol{\Omega}^{(g)}\mathbf{V}_k^{*(g)})$, and $\mathbf{A}_k = (\sum_{g=1}^{G}\psi_{\epsilon k}^{(g)-1}\boldsymbol{\Omega}^{(g)}\boldsymbol{\Omega}^{(g)T} + \mathbf{H}_{0yk}^{-1})^{-1}$. Under the constraints $\psi_{\epsilon k}^{(1)} = \cdots = \psi_{\epsilon k}^{(G)} = \psi_{\epsilon k}$, the conjugate prior distribution of $\psi_{\epsilon k}^{-1}$ is $Gamma(\alpha_{0\epsilon k}, \beta_{0\epsilon k})$, and the conditional distribution is

$$[\psi_{\epsilon k}^{-1}|\boldsymbol{\Lambda}_k^{(1)}, \cdots, \boldsymbol{\Lambda}_k^{(G)}, \boldsymbol{\mu}^{(1)}, \cdots, \boldsymbol{\mu}^{(G)}, \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{X}] \overset{D}{=} Gamma(N^*/2 + \alpha_{0\epsilon k}, \beta_{\epsilon k}), \tag{6.A15}$$

where $N^* = N_1 + \cdots + N_G$,

$$\beta_{\epsilon k} = \beta_{0\epsilon k} + \sum_{g=1}^{G}\big[\boldsymbol{\Lambda}_k^{(g)T}(\boldsymbol{\Omega}^{(g)}\boldsymbol{\Omega}^{(g)T})\boldsymbol{\Lambda}_k^{(g)} - 2\boldsymbol{\Lambda}_k^{(g)T}\boldsymbol{\Omega}^{(g)}\mathbf{V}_k^{*(g)} + \mathbf{V}_k^{*(g)T}\mathbf{V}_k^{*(g)}\big]/2.$$

Under the constraints $\boldsymbol{\Phi}^{(1)} = \cdots = \boldsymbol{\Phi}^{(G)} = \boldsymbol{\Phi}$, the conjugate prior distribution of $\boldsymbol{\Phi}^{-1}$ is $W_{q_2}[\mathbf{R}_0, \rho_0]$, and the conditional distribution is

$$[\boldsymbol{\Phi}|\boldsymbol{\Omega}_2^{(1)}, \cdots, \boldsymbol{\Omega}_2^{(G)}] \overset{D}{=}\ IW_{q_2}[(\sum_{g=1}^{G}\boldsymbol{\Omega}_2^{(g)}\boldsymbol{\Omega}_2^{(g)T} + \mathbf{R}_0^{-1}),\ N^* + \rho_0]. \tag{6.A16}$$

The conditional distributions of $\boldsymbol{\Lambda}_{\omega k}^{(g)}$ and $\psi_{\delta k}^{(g)}$ are similar, and hence not presented.

As the conditional distributions involved in (6.A13) or (6.A14)-(6.A16) are standard distributions, drawing observations from them is straightforward. Simulating observations from the conditional distributions that are given in (6.A12) involves the univariate truncated normal distribution, and this is done by the inverse distribution method proposed by Robert (1995). A Metropolis-Hastings (MH) algorithm is used to simulate observations from the more complex conditional distribution (6.A10).