COMPUTATIONAL DEVELOPMENT AND MATLABE CODE

The detailed MCMC sampling process

Step 1: Sampling the auxiliary variable λ_{ijk} and η_{ijk}, given the response $\boldsymbol{Y}$ and $\boldsymbol{\Omega}$. The full conditional distribution can be written as follows1:

$$\lambda_{ijk}\,|\boldsymbol{Y},\boldsymbol{\Omega} \sim \text{Uniform}\left(0, p_{ijk}\right), \text{ if } y_{ijk}=1, \tag{1}$$
$$\eta_{ijk}\,|\boldsymbol{Y},\boldsymbol{\Omega} \sim \text{Uniform}\left(0, \psi_{ijk}\right), \text{ if } y_{ijk}=0. \tag{2}$$

Step 2: Sampling the difficulty parameter b_k; suppose that the prior of the difficulty parameters is $b_k \sim N\left(\mu_b, \sigma_b^2\right)$. According to Equation (1), given item k, $\forall i, j$ when $y_{ijk}=1$, we have $0<\lambda_{ijk}\le p_{ijk}$, and the following inequalities are established:

$$\sum_{q=1}^{Q} a_{kq}\theta_{ijq} - b_k \ge \log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right) \text{ or equivalently}$$
$$b_k \le \sum_{q=1}^{Q} a_{kq}\theta_{ijq} - \log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right). \tag{3}$$

In the same way, when $y_{ijk}=0$, we have $0<\eta_{ijk}\le\psi_{ijk}$, and the inequalities are established:

$$\sum_{q=1}^{Q} a_{kq}\theta_{ijq} - b_k \le \log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right) \text{ or equivalently}$$
$$b_k \ge \sum_{q=1}^{Q} a_{kq}\theta_{ijq} - \log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right). \tag{4}$$

Therefore, given the response variable $\boldsymbol{Y}$, the auxiliary variable $\boldsymbol{\lambda}$, and $\boldsymbol{\eta}$ the parameters $\boldsymbol{\Omega}$, the full conditional distribution can be written as:

$$b_k\,|\boldsymbol{Y},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\Omega} \sim N\left(\mu_b,\sigma_b^2\right) I\left(b_k^L \le b_k \le b_k^U\right), \tag{5}$$

where

$$b_k^L = \max_{(i,j)\in E_k}\left\{\sum_{q=1}^{Q} a_{kq}\theta_{ijq} - \log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right)\right\},$$
$$b_k^U = \min_{(i,j)\in D_k}\left\{\sum_{q=1}^{Q} a_{kq}\theta_{ijq} - \log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right)\right\}.$$

Step 3: Sampling a_{kq}, which is an element of vector $\boldsymbol{a}_k$, $\boldsymbol{a}_k = \left(a_{k1},\cdots,a_{kQ}\right)'$. The prior of a_{kq} is assumed to follow a normal distribution with mean μ_{aq} and variance σ_{aq}^2. Given item k, when $y_{ijk}=1$, $\forall i,j,k$, the inequalities $0<\lambda_{ijk}\le p_{ijk}$, $\theta_{ijq}>0$ are established, and when $y_{ijk}=0$, the inequalities $0<\eta_{ijk}\le\psi_{ijk}$, $\theta_{ijq}<0$ are established. Therefore,

$$a_{kq} \ge \frac{1}{\theta_{ijq}}\left[\log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right], \text{ and}$$
$$a_{kq} \ge \frac{1}{\theta_{ijq}}\left[\log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right].$$

and in the same way, when $y_{ijk}=1$, $\forall i,j,k$, the inequalities $0<\lambda_{ijk}\le p_{ijk}$, $\theta_{ijq}<0$ are established, or when $y_{ijk}=0$, we have $0<\eta_{ijk}\le\psi_{ijk}$, $\theta_{ijq}<0$. Therefore,

$$a_{kq} \le \frac{1}{\theta_{ijq}}\left[\log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right], \text{and}$$
$$a_{kq} \le \frac{1}{\theta_{ijq}}\left[\log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right]. \tag{6}$$

Let

$$\nabla_{kq} = \left\{(i,j)\,|y_{ijk}=1, \lambda_{ijk}\le p_{ijk}, \theta_{ijq}>0\right\},$$
$$F_{kq} = \left\{(i,j)\,|y_{ijk}=0, \eta_{ijk}\le \psi_{ijk}, \theta_{ijq}<0\right\},$$
$$\triangle_{kq} = \left\{(i,j)\,|y_{ijk}=1, \lambda_{ijk}\le p_{ijk}, \theta_{ijq}<0\right\},$$
$$G_{kq} = \left\{(i,j)\,|y_{ijk}=0, \eta_{ijk}\le \psi_{ijk}, \theta_{ijq}>0\right\}.$$

When given the response variable $\boldsymbol{Y}$, the auxiliary variable $\boldsymbol{\lambda}$, $\boldsymbol{\eta}$ and other parameters $\boldsymbol{\Omega}_1$ (all of the parameters except a_{kq}), the full conditional distribution is represented by

$$a_{kq}\,|\boldsymbol{Y},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\Omega}_1 \sim N\left(\mu_{aq},\sigma_{aq}^2\right) I\left(a_{kq}^L \le a_{kq} \le a_{kq}^U\right), \tag{7}$$

where

$$a_{kq}^L = \max\left\{\max_{(i,j)\in\triangle_{kq}} \frac{1}{\theta_{ijq}}\left[\log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right], \max_{(i,j)\in F_{kq}} \frac{1}{\theta_{ijq}}\left[\log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right]\right\},$$

$$a_{kq}^U = \min\left\{\min_{(i,j)\in\nabla_{kq}} \frac{1}{\theta_{ijq}}\left[\log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right], \min_{(i,j)\in G_{kq}} \frac{1}{\theta_{ijq}}\left[\log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right) + b_k - \sum_{r\ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right]\right\}.$$

Step 4: Sampling θ_{ijq}, which is an element of the ability vector $\boldsymbol{\theta}_{ij}$, $\boldsymbol{\theta}_{ij} = \left(\theta_{ij1},\cdots,\theta_{ijQ}\right)'$. As we known, a level-2 (individual level) random regression model can be considered the prior of the latent ability, and the latent abilities have relationships between dimensions. Therefore, the conditional independent prior will be considered. Specific derivations are as follows: let $\boldsymbol{\beta}_j = \left(\boldsymbol{\beta}_{j1}^{'}, \cdots, \boldsymbol{\beta}_{jQ}^{'}\right)^{'}$, where $\boldsymbol{\beta}_j$ is a $Q\,(h+1)$-by-1 column vector, $\boldsymbol{\beta}_{jq}^{'} = (\beta_{0jq}, \cdots, \beta_{hjq})$, $q = 1, \cdots, Q$. $\boldsymbol{X}_{ij} = (1, x_{1ij}, \cdots, x_{hij})$ and $\boldsymbol{\mu} = \left(\boldsymbol{X}_{ij}\boldsymbol{\beta}_{j1}, \cdots, \boldsymbol{X}_{ij}\boldsymbol{\beta}_{jQ}\right)^{'}$; we can adjust the order of the concerned dimensional ability θ_{ijq} to the first position of the column vector, and the corresponding mean $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_e$ are also adjusted. Therefore, $\boldsymbol{\theta}_{ij}^{*} = \left(\theta_{ijq}, \boldsymbol{\theta}_{ij(-q)}\right)^{'}$, $\boldsymbol{\mu}^{*} = \left(\boldsymbol{X}_{ij}\boldsymbol{\beta}_{jq}, \boldsymbol{\mu}_1^{(2)}\right)^{'}$, $\boldsymbol{\Sigma}_e^{*} = \begin{pmatrix} \sigma_{e_q}^2 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. The conditional independent prior is as follows:

$$p\left(\theta_{ijq}\left|\boldsymbol{\theta}_{ij(-q)}, \boldsymbol{\beta}_j, \boldsymbol{\Sigma}_e\right.\right) \sim N_q\left(\mu_{ij}^q, \sigma_q^2\right), \tag{8}$$

where

$$\mu_{ij}^q = \boldsymbol{X}_{ij}\boldsymbol{\beta}_{jq} + \Sigma_{12}\Sigma_{22}^{-1}\left(\boldsymbol{\theta}_{ij(-q)} - \boldsymbol{\mu}_1^{(2)}\right), \text{ and}$$
$$\sigma_q^2 = \sigma_{e_q}^2 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

The full conditional distribution of θ_{ijq} can be written as:

(9)

$$\theta_{ijq}\left|\boldsymbol{Y}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\theta}_{ij(-q)}, \boldsymbol{\Omega}_2\right. \sim N_q\left(\mu_{ij}^q, \sigma_q^2\right) I\left(\theta_{ijq}^L \le \theta_{ijq} \le \theta_{ijq}^U\right).$$

where $\boldsymbol{\theta}_{ij(-q)}\left(\boldsymbol{\theta}_{ij} \text{ expect } \theta_{ijq}\right), \boldsymbol{\Omega}_2\left(\boldsymbol{\Omega}_2 \text{ expect } \boldsymbol{\theta}\right)$

$$\theta_{ijq}^L = \max_{k \in C_{ijq}} \left\{\frac{1}{a_{kq}}\left[\log\left(\frac{\lambda_{ijk}}{1-\lambda_{ijk}}\right) + b_k - \sum_{r \ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right]\right\},$$

where $C_{ijq} = \left\{k\left|y_{ijk} = 1, 0 < \lambda_{ijk} \le p_{ijk}\right.\right\}$,

$$\theta_{ijq}^U = \max_{k \in F_{ijq}} \left\{\frac{1}{a_{kq}}\left[\log\left(\frac{1-\eta_{ijk}}{\eta_{ijk}}\right) + b_k - \sum_{r \ne q, r=1}^{Q} a_{kr}\theta_{ijr}\right]\right\},$$

where $F_{ijq} = \left\{k\left|y_{ijk} = 0, 0 < \eta_{ijk} \le \psi_{ijk}\right.\right\}$.

Step 5: Sampling $\boldsymbol{\beta}_j = \left(\boldsymbol{\beta}_{j1}^{'}, \cdots, \boldsymbol{\beta}_{jQ}^{'}\right)^{'}$, the random regression coefficient $\boldsymbol{\beta}_j$ is a $Q\,(h+1)$-by-1 column vector. Let $\boldsymbol{X}_{ij}^{*} = \text{diag}\left(\boldsymbol{X}_{ij}, \cdots, \boldsymbol{X}_{ij}\right)$, where $\boldsymbol{X}_{ij}^{*}$ is a Q-by-$Q\,(h+1)$ matrix, and $\boldsymbol{\theta}_{ij} = (\theta_{ij1}, \cdots, \theta_{ijQ})'$. Given $\boldsymbol{\theta}$, $\boldsymbol{\Sigma}_e$, $\boldsymbol{\gamma}$ and $\boldsymbol{T}$, the full conditional posterior distribution of $\boldsymbol{\beta}_j$ is given by

(10)

$$p\left(\boldsymbol{\beta}_j\left|\boldsymbol{\theta}_{ij}, \boldsymbol{\Sigma}_e, \boldsymbol{\gamma}, \boldsymbol{T}\right.\right) \propto \prod_{i=1}^{n_j} p\left(\boldsymbol{\theta}_{ij}\left|\boldsymbol{\beta}_j, \boldsymbol{\Sigma}_e\right.\right) p\left(\boldsymbol{\beta}_j\left|\boldsymbol{\gamma}, \boldsymbol{T}\right.\right),$$
$$\boldsymbol{\beta}_j\left|\boldsymbol{\theta}_{ij}, \boldsymbol{\Sigma}_e, \boldsymbol{\gamma}, \boldsymbol{T}\right. \sim N\left(\text{Var}_{\boldsymbol{\beta}_j}\left(\Sigma_{\tilde{\boldsymbol{\beta}}_j}^{-1}\tilde{\boldsymbol{\beta}}_j + \boldsymbol{T}^{-1}\boldsymbol{w}_j\boldsymbol{\gamma}\right), \text{Var}_{\boldsymbol{\beta}_j}\right)$$

where $\text{Var}_{\boldsymbol{\beta}_j} = \left(\Sigma_{\tilde{\boldsymbol{\beta}}_j}^{-1} + \boldsymbol{T}^{-1}\right)^{-1}$. As is known, the level-2 likelihood function of $\boldsymbol{\beta}_j$ follows a normal distribution with mean

$$\tilde{\boldsymbol{\beta}}_j = \left(\sum_{i=1}^{n_j}\boldsymbol{X}_{ij}^{*'}\boldsymbol{\Sigma}_e\boldsymbol{X}_{ij}^{*}\right)^{-1}\sum_{i=1}^{n_j}\boldsymbol{X}_{ij}^{*'}\boldsymbol{\Sigma}_e\boldsymbol{\theta}_{ij},$$

and variance

$$\Sigma_{\tilde{\boldsymbol{\beta}}_j} = \left(\sum_{i=1}^{n_j}\boldsymbol{X}_{ij}^{*'}\boldsymbol{\Sigma}_e\boldsymbol{X}_{ij}^{*}\right)^{-1},$$

and the random regression coefficients $\boldsymbol{\beta}_j$ can be induced by a normal prior with mean $\boldsymbol{w}_j\boldsymbol{\gamma}$ and covariance $\boldsymbol{T}$ on level 3. The group level covariate matrix $\boldsymbol{w}_j$ consists of $w_j = (1, w_{j1}, \cdots, w_{js})$, that is, $\boldsymbol{w}_j = \text{diag}\left(w_j, \cdots, w_j\right)$, where $\boldsymbol{w}_j$ is a $Q\,(h+1)$-by-$Q\,(s+1)\,(h+1)$ matrix. The fixed effect $\boldsymbol{\gamma}_q$ is a 1-by-$(s+1)\times(h+1)$ matrix, $\boldsymbol{\gamma}_q = \left(\boldsymbol{\gamma}_{00q}, \cdots, \boldsymbol{\gamma}_{hsq}\right)$, $\boldsymbol{\gamma} = \left(\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_Q\right)'$. The level-3 covariance structure can be represented by a $Q\,(h+1)$-by-$Q\,(h+1)$ matrix, that is, $\boldsymbol{T} = \text{diag}\left(\boldsymbol{T}_1, \cdots, \boldsymbol{T}_Q\right)$.

Step 6: Sampling $\boldsymbol{\gamma}$, $\boldsymbol{\gamma} = \left(\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_Q\right)^{'}$, the matrix $\boldsymbol{\gamma}$ is the matrix of regression coefficients of regression of $\boldsymbol{\beta}_j$ on $\boldsymbol{w}_j$. The full conditional posterior distribution of is given by

$$p\left(\boldsymbol{\gamma}\left|\boldsymbol{\beta}, \boldsymbol{T}\right.\right) \propto \prod_{j=1}^{J} p\left(\boldsymbol{\beta}_j\left|\boldsymbol{\gamma}, \boldsymbol{T}\right.\right) p\left(\boldsymbol{\gamma}\left|\boldsymbol{\phi}\right.\right), \tag{11}$$

Therefore, the full conditional posterior distribution follows normal distribution with mean

$$\left(\sum_{j=1}^{J}\boldsymbol{w}_j'\boldsymbol{T}^{-1}\boldsymbol{w}_j\right)^{-1}\sum_{j=1}^{J}\boldsymbol{w}_j'\boldsymbol{T}^{-1}\boldsymbol{\beta}_j$$

and variance

$$\left(\sum_{j=1}^{J}\boldsymbol{w}_j'\boldsymbol{T}^{-1}\boldsymbol{w}_j\right)^{-1}$$

Step 7: Sampling the residual variance-covariance structure $\boldsymbol{\Sigma}_e$ (a $Q \times Q$ matrix), a prior for $\boldsymbol{\Sigma}_e$ is an Inverse-Wishart(v_0, Σ_0) distribution. The full conditional posterior distribution of $\boldsymbol{\Sigma}_e$ is given by

$$p\left(\boldsymbol{\Sigma}_e\left|\boldsymbol{\theta}, \boldsymbol{\beta}\right.\right) \propto p\left(\boldsymbol{\theta}\left|\boldsymbol{\beta}, \boldsymbol{\Sigma}_e\right.\right) p\left(\Sigma_0\right). \tag{12}$$

Let $M = \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(\boldsymbol{\theta}_{ij} - \boldsymbol{X}_{ij}\boldsymbol{\beta}_j\right)\left(\boldsymbol{\theta}_{ij} - \boldsymbol{X}_{ij}\boldsymbol{\beta}_j\right)^{'}$ and $N = n_1 + \cdots + n_J$; thus,

$$\boldsymbol{\Sigma}_e\left|\boldsymbol{\theta}, \boldsymbol{\beta}\right. \sim \text{Inverse-Wishart}\left(v_0 + N, M + \Sigma_0\right).$$

Step 8: Sampling $\boldsymbol{T} =$ diag$(\boldsymbol{T}_1, \cdots, \boldsymbol{T}_Q)$. $\boldsymbol{T}_q$ ($q = 1, \cdots, Q$) is a covariance matrix of different dimensional abilities of level 3. First, $\boldsymbol{T}_q$ is drawn. A prior for is an Inverse-Wishart(v_q, Σ_q) distribution. The full conditional posterior distribution of $\boldsymbol{T}_q$ is given by

$$p\left(\boldsymbol{T}_q \left| \boldsymbol{\beta}_{jq}, \boldsymbol{\gamma}_q \right.\right) \propto p\left(\boldsymbol{\beta}_{jq} \left| \boldsymbol{\gamma}_q, \boldsymbol{T}_q \right.\right) p\left(\boldsymbol{T}_q\right).$$

Let $H_q = \sum_{j=1}^{J} \left(\boldsymbol{\beta}_{jq} - \boldsymbol{w}_j \boldsymbol{\gamma}_q\right) \left(\boldsymbol{\beta}_{jq} - \boldsymbol{w}_j \boldsymbol{\gamma}_q\right)^{'}$; thus,

$$\boldsymbol{T}_q \left| \boldsymbol{\beta}_{jq}, \boldsymbol{\gamma}_q \right. \sim \text{Inverse-Wishart}\left(v_q + J, H_q + \Sigma_q\right).$$

Matlab code

```
Main Code:
Sample b:
    D=sum(a1.*theta1,3);
    lambda0=unifrnd(0,prostar);
    psi0=unifrnd(0,1-prostar);
    z1=D-log(lambda0./(1-lambda0));
    z2=D-log((1-psi0)./psi0);
    bleft=max(z2);
    bright=min(z1);
    u1=unifrnd(0,1,1,n);
    y1=normcdf(bleft,0,1)
      +(normcdf(bright,0,1)-normcdf(bleft,0,1)).*u1;
    y1(y1>0.9999)=0.9999;
    y1(y1<0.0001)=0.0001;
    b0=norminv(y1,0,1);
    b0(1:Ndim)=0;
    Rb(i,:)=b0;
    b1=M*b0;
    prostar=exp(D-b1)./(1+exp(D-b1));
Sample theta:
     for i0=1:Ndim
        muability(:,i0)=sum(beta1(:,:,i0).*x,2);
     end
    sigmaability=sigma0;
     for i0=1:Ndim
        indic0=indic~=i0;
        z3(:,:,i0)=(1./a1(:,:,i0)).*(log(lambda0./(1-lambda0))
                  +b1 sum(a1(:,:,indic0).*theta1(:,:,indic0),3));
        z4(:,:,i0)=(1./a1(:,:,i0)).*(log((1-psi0)./psi0)
                  +b1-sum(a1(:,:,indic0).*theta1(:,:,indic0),3));
        thetaleft=max(z3(:,:,i0),[],2);
        thetaright=min(z4(:,:,i0),[],2);
        s1=sigma0(indic==i0,indic==i0);
        s2=sigma0(indic~=i0,indic==i0);
        s3=sigma0(indic~=i0,indic~=i0);
        s4=s1-s2'*s3^(-1)*s2;
        u3=unifrnd(0,1,m,1);
        y3=normcdf(thetaleft,muability(:,i0),s4)
          +(normcdf(thetaright,muability(:,i0),s4)
          -normcdf(thetaleft,muability(:,i0),s4)).*u3;
        y3(y3>0.9999)=0.9999;
        y3(y3<0.0001)=0.0001;
        theta0=norminv(y3,muability(:,i0),s4);
        theta0(theta0>=4)=4;
        theta0(theta0<=-4)=-4;
        theta0(:,i0)=theta0;
        Rtheta(:,i0,i)=theta0(:,i0);
        theta1(:,:,i0)=theta0(:,i0)*N;
    end
    D=sum(a1.*theta1,3);
    prostar=exp(D-b1)./(1+exp(D-b1));
Sample a:
    for i0=1:Ndim
        indic0=indic~=i0;
        z5=(1./theta1(:,:,i0)).*(log(lambda0./(1-lambda0))
          +b1-sum(a1(:,:,indic0).*theta1(:,:,indic0),3));
        z6=(1./theta1(:,:,i0)).*(log((1-psi0)./psi0)
          +b1-sum(a1(:,:,indic0).*theta1(:,:,indic0),3));
        indic21=theta1(:,:,i0)>0;
        F1=z5.*indic21;
        F1(F1==0)=NaN;
        aleft1=max(F1);
        indic22=theta1(:,:,i0)<0;
        F2=z6.*indic22;
        F2(F2==0)=NaN;
        aleft2=max(F2);
        aleft=max(aleft1,aleft2);
        F3=z5.*indic22;
        F3(F3==0)=NaN;
        aright1=min(F3);
        F4=z6.*indic21;
        F4(F4==0)=NaN;
        aright2=min(F4);
        aright=min(aright1,aright2);
        u2=unifrnd(0,1,1,n);
        y2=normcdf(aleft,0,1)
          +(normcdf(aright,0,1)-normcdf(aleft,0,1)).*u2;
        y2(y2>0.9999)=0.9999;
        y2(y2<0.0001)=0.0001;
        a0(i0,:)=norminv(y2,0,1);
        a0(1:Ndim,1:Ndim)=eye(Ndim);
        Ra(i0,:,i)=a0(i0,:);
        a1(:,:,i0)=M*a0(i0,:);
    end
    D=sum(a1.*theta1,3);
    prostar=exp(D-b1)./(1+exp(D-b1));
```