# MARGINALIZED TRANSITION RANDOM EFFECTS MODELS FOR MULTIVARIATE LONGITUDINAL BINARY DATA: Web appendix

**Ozlem Ilk**[1] **and Michael Daniels**[2]
[1] Department of Statistics, Middle East Technical University, Ankara, Turkey
*email:* oilk@metu.edu.tr
[2] Department of Statistics, University of Florida, Gainesville, FL 32611
*email:* mdaniels@stat.ufl.edu

# Appendix C

*Calculation of $\Delta_{itj}$, $\Delta^*_{itj}$ and their derivatives:*

$\Delta_{itj}$ and $\Delta^*_{itj}$ are determined through equations (4) and (5) by using Newton-Raphson. Numerical solution of the integrals in (5.1) and (5.4) are obtained by a 20-point Gauss-Hermite Quadrature.

• For the special case of p=1, we have two marginal constraint equations; one for $t = 2$, and other for $t > 2$. Specifically, for $t = 2$, we have

$$f(\Delta_{i2j}) = -\frac{e^{\mathbf{X}_{i2j}\beta}}{1+e^{\mathbf{X}_{i2j}\beta}} + \sum_{y_{i1j}=0}^{1} \frac{e^{\Delta_{i2j}+\gamma_{i2j}y_{i1j}}}{1+e^{\Delta_{i2j}+\gamma_{i2j}y_{i1j}}} \frac{e^{y_{i1j}\mathbf{X}_{i1j}\beta^*}}{1+e^{\mathbf{X}_{i1j}\beta^*}} = 0$$

For later time points, $P(Y_{it-1j})$ depends on $\beta$ and not on $\beta^*$. Hence, we have,

$$f(\Delta_{itj}) = -\frac{e^{\mathbf{X}_{itj}\beta}}{1+e^{\mathbf{X}_{itj}\beta}} + \sum_{y_{it-1j}=0}^{1} \frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}} \frac{e^{y_{it-1j}\mathbf{X}_{it-1j}\beta}}{1+e^{\mathbf{X}_{it-1j}\beta}} = 0$$

To use Newton-Raphson, we require the derivatives of these equations. For $t = 2$, this derivative is:

$$\frac{\partial f(\Delta_{i2j})}{\partial \Delta_{i2j}} = \sum_{y_{i1j}=0}^{1} \frac{e^{\Delta_{i2j}+\gamma_{i2j}y_{i1j}}}{(1+e^{\Delta_{i2j}+\gamma_{i2j}y_{i1j}})^2} \frac{e^{y_{i1j}\mathbf{X}_{i1j}\beta^*}}{1+e^{\mathbf{X}_{i1j}\beta^*}}$$

For later time points, this derivative takes the following form:

$$\frac{\partial f(\Delta_{itj})}{\partial \Delta_{itj}} = \sum_{y_{it-1j}=0}^{1} \frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{(1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}})^2} \frac{e^{y_{it-1j}\mathbf{X}_{it-1j}\beta}}{1+e^{\mathbf{X}_{it-1j}\beta}}$$

- From equation (5.4), we have the convolution equation for $\Delta^*_{i1j}$:

$$f(\Delta^*_{i1j}) = -\frac{e^{\mathbf{X}_{i1j}\beta^*}}{1+e^{\mathbf{X}_{i1j}\beta^*}} + \int \frac{e^{\Delta^*_{i1j}+\lambda^*_j\sigma_1 z_i}}{1+e^{\Delta^*_{i1j}+\lambda^*_j\sigma_1 z_i}}\phi(z_i)dz_i = 0$$

For Newton-Raphson, we use

$$\frac{\partial f(\Delta^*_{i1j})}{\partial \Delta^*_{i1j}} = \int \frac{e^{\Delta^*_{i1j}+\lambda^*_j\sigma_1 z_i}}{(1+e^{\Delta^*_{i1j}+\lambda^*_j\sigma_1 z_i})^2}\phi(z_i)dz_i$$

All integrals in functions and derivatives are approximated by using Gauss-Hermite Quadrature.

- The convolution equation of $\Delta^*_{itj}(t \geq 2)$ gives

$$f(\Delta^*_{itj}) = -\frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}} + \int \frac{e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}{1+e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}\phi(z_i)dz_i = 0$$

In Newton-Raphson step, we use

$$\frac{\partial f(\Delta^*_{itj})}{\partial \Delta^*_{itj}} = \int \frac{e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}{(1+e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i})^2}\phi(z_i)dz_i$$

*Full conditional distributions:*

- $f(log(\sigma^2), \lambda|\theta_{-\mathbf{b}}, \theta_{-\sigma}, \theta_{-\lambda}, \mathbf{Y}) \propto \prod_{t=2}^{n}[\Pi(log(\sigma_t^2))] \prod_{j=1}^{r}[\Pi(\lambda_j)] \prod_{t=2}^{n}\prod_{i=1}^{N}[\int \prod_{j=1}^{r}[P(Y_{itj}|Y_{it-1j}, \theta)]f(b_{it}|\sigma_t)db_{it}]$

$$\propto \prod_{t=2}^{n}[\frac{e^{log(\sigma_t^2)}}{(1+e^{log(\sigma_t^2)})^2}] \prod_{j=1}^{r}[e^{\frac{-(\lambda_j-1)^2}{2\sigma_\lambda^2}}] \prod_{t=2}^{n}\prod_{i=1}^{N}[\int \prod_{j=1}^{r}[(\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{y_{itj}}(\frac{1}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{1-y_{itj}}]\frac{1}{\sigma_t}e^{-\frac{b_{it}^2}{2\sigma_t^2}}db_{it}]$$

This integral is also approximated by Gauss-Hermite Quadrature:

$$f(log(\sigma^2),\lambda|\theta_{-\mathbf{b}},\theta_{-\sigma},\theta_{-\lambda},\mathbf{Y}) \propto \prod_{t=2}^{n}[\frac{e^{log(\sigma_t^2)}}{(1+e^{log(\sigma_t^2)})^2}]\prod_{j=1}^{r}[e^{\frac{-(\lambda_j-1)^2}{2\sigma_\lambda^2}}]\prod_{t=2}^{n}\prod_{i=1}^{N}[\sum_l w_l\prod_{j=1}^{r}[\frac{(e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l})^{y_{itj}}}{1+e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}]]$$

$$\equiv \prod_{t=2}^{n}[\frac{e^{log(\sigma_t^2)}}{(1+e^{log(\sigma_t^2)})^2}]\prod_{j=1}^{r}[e^{\frac{-(\lambda_j-1)^2}{2\sigma_\lambda^2}}]\prod_{i=1}^{N}[D_{it}]$$

where $D_{it}=\sum_l w_l\prod_{j=1}^{r}[\frac{(e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l})^{y_{itj}}}{1+e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}]$

- $f(log(\sigma_1^2),\lambda^*|\theta_{-\mathbf{b}},\theta_{-\sigma},\theta_{-\lambda^*},\mathbf{Y}) \propto \Pi(log(\sigma_1^2))\prod_{j=1}^{r}[\Pi(\lambda_j^*)]\prod_{i=1}^{N}[\int\prod_{j=1}^{r}[P(Y_{i1j}|\theta)]f(b_{i1}|\sigma_1)db_{i1}]$

$$\propto \frac{e^{log(\sigma_1^2)}}{(1+e^{log(\sigma_1^2)})^2}\prod_{j=1}^{r}[e^{\frac{-(\lambda_j^*-1)^2}{2\sigma_\lambda^{2*}}}]\prod_{i=1}^{N}[\int\prod_{j=1}^{r}[(\frac{e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})^{y_{i1j}}(\frac{1}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})^{1-y_{i1j}}]\frac{1}{\sigma_1}e^{-\frac{b_{i1}^2}{2\sigma_1^2}}db_{i1}]$$

$$\equiv \frac{e^{log(\sigma_1^2)}}{(1+e^{log(\sigma_1^2)})^2}\prod_{j=1}^{r}[e^{\frac{-(\lambda_j^*-1)^2}{2\sigma_\lambda^{2*}}}]\prod_{i=1}^{N}[D_{i1}]$$

where $D_{i1}=\sum_l w_l\prod_{j=1}^{r}[\frac{(e^{\Delta^*_{i1j}+\sqrt{2e^{log(\sigma_1^2)}}\lambda_j^* z_l})^{y_{i1j}}}{1+e^{\Delta^*_{i1j}+\sqrt{2e^{log(\sigma_1^2)}}\lambda_j^* z_l}}]$

- $f(b_{i1}|\theta_{-\mathbf{b}},\mathbf{Y}) \propto \prod_{j=1}^{r}[P(Y_{i1j}|\theta)]P(b_{i1}|\sigma_1)$

$$\propto \prod_{j=1}^{r}[(\frac{e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})^{y_{i1j}}(\frac{1}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})^{1-y_{i1j}}]e^{-\frac{b_{i1}^2}{2\sigma_1^2}}$$

- $f(b_{it}|\theta_{-\mathbf{b}},\mathbf{Y}) \propto \prod_{j=1}^{r}[P(Y_{itj}|Y_{it-1j},\theta)]P(b_{it}|\sigma_t)$

$$\propto \prod_{j=1}^{r}[(\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{y_{itj}}(\frac{1}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{1-y_{itj}}]e^{-\frac{b_{it}^2}{2\sigma_t^2}}$$

- $f(\beta^*|\theta_{-\beta^*},\mathbf{Y}) \propto \prod_{i=1}^{N}\prod_{j=1}^{r}[P(Y_{i1j}|\theta)P(Y_{i2j}|Y_{i1j},\theta)]\Pi(\beta^*)$

$$\propto \prod_{i=1}^{N}\prod_{j=1}^{r}[(\frac{e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})^{y_{i1j}}(\frac{1}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})^{1-y_{i1j}}(\frac{e^{\Delta^*_{i2j}+\lambda_j b_{i2}}}{1+e^{\Delta^*_{i2j}+\lambda_j b_{i2}}})^{y_{i2j}}(\frac{1}{1+e^{\Delta^*_{i2j}+\lambda_j b_{i2}}})^{1-y_{i2j}}]e^{-\frac{\sum_k \beta_k^{*2}}{2\sigma^2_{\beta^*}}}$$

- $f(\beta|\theta_{-\beta},\mathbf{Y}) \propto \prod_{i=1}^{N}\prod_{j=1}^{r}\prod_{t=2}^{n}[P(Y_{itj}|Y_{it-1j},\theta)]\Pi(\beta)$

$$\propto \prod_{i=1}^{N}\prod_{j=1}^{r}\prod_{t=2}^{n}[(\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{y_{itj}}(\frac{1}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{1-y_{itj}}]e^{-\frac{\sum_k \beta_k^2}{2\sigma^2_\beta}}$$

- $f(\alpha_t|\theta_{-\alpha},\mathbf{Y}) \propto \prod_{i=1}^{N}\prod_{j=1}^{r}[P(Y_{itj}|Y_{it-1j},\theta)]\Pi(\alpha_t)$

$$\propto \prod_{i=1}^{N}\prod_{j=1}^{r}[(\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{y_{itj}}(\frac{1}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{1-y_{itj}}]e^{-\frac{\sum_k \alpha_{t,k}^2}{2\sigma^2_\alpha}}$$

*Derivatives of full conditional distributions:*

For Hybrid MC, we need derivatives of full conditionals, as well as derivatives of $\Delta_{itj}$ and $\Delta^*_{itj}$. Necessary derivatives are obtained by chain rule and implicit differentiation. For simplicity of notation, define,

$$A_{i1j} = \int \frac{e^{\Delta^*_{i1j}+\lambda^*_j\sigma_1 z_i}}{(1+e^{\Delta^*_{i1j}+\lambda^*_j\sigma_1 z_i})^2}\phi(z_i)dz_i$$

For $t > 1$,

$$A_{itj} = \int \frac{e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}{(1+e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i})^2}\phi(z_i)dz_i$$

$$B_{itj} = \frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{(1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}})^2}$$

$$C_{itj} = \sum_{y_{it-1j}} \frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{(1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}})^2}P(Y_{it-1j})$$

- $\frac{\partial log f(log(\sigma^2),\lambda|\theta_{-\mathbf{b}},\theta_{-\sigma},\theta_{-\lambda},\mathbf{Y})}{\partial log(\sigma_t^2)} = \sum_i \frac{\partial D_{it}/\partial log(\sigma_t^2)}{D_{it}} + 1 - 2\frac{e^{log(\sigma_t^2)}}{1+e^{log(\sigma_t^2)}}$ where

$$\frac{\partial D_{it}}{\partial log(\sigma_t^2)} = \sum_l w_l[\prod_{j=1}^r \frac{(e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l})^{y_{itj}}}{1+e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}][\sum_{j=1}^r(\frac{\partial\Delta^*_{itj}}{\partial log(\sigma_t^2)}+\lambda_j z_l\sqrt{e^{log(\sigma_t^2)}/2})\frac{(y_{itj}+e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}(y_{itj}-1))}{1+e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}]$$

$$\frac{\partial\Delta^*_{itj}}{\partial log(\sigma_t^2)} = \frac{\partial\Delta^*_{itj}}{\partial\sigma_t}\frac{\partial\sigma_t}{\partial log(\sigma_t^2)}$$

$$\frac{\partial\Delta^*_{itj}}{\partial\sigma_t} = \frac{-\sqrt{2}\lambda_j\sum_l w_l z_l \frac{e^{\Delta^*_{itj}+\sqrt{2}\lambda_j\sigma_t z_l}}{(1+e^{\Delta^*_{itj}+\sqrt{2}\lambda_j\sigma_t z_l})^2}}{\sum_l w_l \frac{e^{\Delta^*_{itj}+\sqrt{2}\lambda_j\sigma_t z_l}}{(1+e^{\Delta^*_{itj}+\sqrt{2}\lambda_j\sigma_t z_l})^2}}, \quad \frac{\partial\sigma_t}{\partial log(\sigma_t^2)} = \sqrt{e^{log(\sigma_t^2)}}/2$$

- $\frac{\partial log f(log(\sigma^2),\lambda|\theta_{-\mathbf{b}},\theta_{-\sigma},\theta_{-\lambda},\mathbf{Y})}{\partial\lambda_m} = [\sum_i\sum_t \frac{\partial D_{it}/\partial\lambda_m}{D_{it}}] - \frac{\lambda_m}{\sigma_\lambda^2}$ for $m = 2, ..., r$.

where $\frac{\partial D_{it}}{\partial\lambda_m} = \sum_l w_l[\prod_{j=1}^r \frac{(e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l})^{y_{itj}}}{1+e^{\Delta^*_{itj}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}][\frac{(\frac{\partial\Delta^*_{itm}}{\partial\lambda_m}+\sqrt{2e^{log(\sigma_t^2)}}z_l)(y_{itm}+e^{\Delta^*_{itm}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_m z_l}(y_{itm}-1))}{1+e^{\Delta^*_{itm}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_m z_l}}]$

$$\frac{\partial\Delta^*_{itm}}{\partial\lambda_m} = \frac{-\sqrt{2e^{log(\sigma_t^2)}}\sum_l w_l z_l \frac{e^{\Delta^*_{itm}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_m z_l}}{(1+e^{\Delta^*_{itm}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_m z_l})^2}}{\sum_l w_l \frac{e^{\Delta^*_{itm}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_m z_l}}{(1+e^{\Delta^*_{itm}+\sqrt{2e^{log(\sigma_t^2)}}\lambda_m z_l})^2}}$$

- $\frac{\partial log f(log(\sigma_1^2),\lambda^*|\theta_{-\mathbf{b}},\theta_{-\sigma},\theta_{-\lambda^*},\mathbf{Y})}{\partial\lambda^*_m} = \sum_i \frac{\partial D_{i1}/\partial\lambda^*_m}{D_{i1}} - \frac{\lambda^*_m}{\sigma^2_{\lambda^*}}$ for $m = 2, ..., r$.

where $\frac{\partial D_{i1}}{\partial\lambda^*_m} = \sum_l w_l[\prod_{j=1}^r \frac{(e^{\Delta^*_{i1j}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_j z_l})^{y_{i1j}}}{1+e^{\Delta^*_{i1j}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_j z_l}}][\frac{(\frac{\partial\Delta^*_{i1m}}{\partial\lambda^*_m}+\sqrt{2}\sigma_1 z_l)(y_{i1m}+e^{\Delta^*_{i1m}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_m z_l}(y_{i1m}-1))}{1+e^{\Delta^*_{i1m}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_m z_l}}]$

$$\frac{\partial\Delta^*_{i1m}}{\partial\lambda^*_m} = \frac{-\sqrt{2e^{log(\sigma_1^2)}}\sum_l w_l z_l \frac{e^{\Delta^*_{i1m}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_m z_l}}{(1+e^{\Delta^*_{i1m}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_m z_l})^2}}{\sum_l w_l \frac{e^{\Delta^*_{i1m}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_m z_l}}{(1+e^{\Delta^*_{i1m}+\sqrt{2e^{log(\sigma_1^2)}}\lambda^*_m z_l})^2}}$$

- $\frac{\partial log f(b_{it}|\theta_{-\mathbf{b}},\mathbf{Y})}{\partial b_{it}} = \sum_{j=1}^r[\lambda_j * (Y_{itj} - \frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})] - \frac{b_{it}}{\sigma_t^2}$

- $\frac{\partial log f(\beta|\theta_{-\beta},\mathbf{Y})}{\partial\beta_k} = \sum_{i=1}^{N}\sum_{j=1}^{r}\sum_{t=2}^{n}[(Y_{itj} - \frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial\Delta^*_{itj}}{\partial\beta_k}] - \frac{\beta_k}{\sigma^2_\beta}$

where $\frac{\partial\Delta^*_{itj}}{\partial\beta_k} = \frac{\partial\Delta^*_{itj}}{\partial\Delta_{itj}}\frac{\partial\Delta_{itj}}{\partial\beta_k}$ for $t \geq 2$.

From convolution equation, we have $\frac{\partial\Delta^*_{itj}}{\partial\Delta_{itj}} = B_{itj}/A_{itj}$.

$\frac{\partial\Delta_{itj}}{\partial\beta_k}$ takes two forms depending on $t$. For $t = 2$,

$$\frac{\partial\Delta_{i2j}}{\partial\beta_k} = \frac{\frac{X_{i2jk}e^{X_{i2j}\beta}}{(1+e^{X_{i2j}\beta})^2}}{\sum_{y_{i1j}=0}^{1}\frac{e^{\Delta_{i2j}+\gamma_{i2j}y_{i1j}}}{(1+e^{\Delta_{i2j}+\gamma_{i2j}y_{i1j}})^2}\frac{e^{X_{i1j}\beta^* y_{i1j}}}{1+e^{X_{i1j}\beta^*}}}$$

From marginal constraint equation for $t > 2$, we have

$$\frac{\partial\Delta_{itj}}{\partial\beta_k} = [\frac{X_{itjk}e^{X_{itj}\beta}}{(1+e^{X_{itj}\beta})^2} - \sum_{y_{it-1j}=0}^{1}\frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}\frac{\partial P(Y_{it-1j})}{\partial\beta_k}]/C_{itj}$$

where $\frac{\partial P(Y_{it-1j}=0)}{\partial\beta_k} = -\frac{X_{it-1jk}e^{\mathbf{X}_{it-1j}\beta}}{(1+e^{\mathbf{X}_{it-1j}\beta})^2}$ and $\frac{\partial P(Y_{it-1j}=1)}{\partial\beta_k} = \frac{X_{it-1jk}e^{\mathbf{X}_{it-1j}\beta}}{(1+e^{\mathbf{X}_{it-1j}\beta})^2}$

- $\frac{\partial log f(\beta^*|\theta_{-\beta^*},\mathbf{Y})}{\partial\beta^*_k} = \sum_{i=1}^{N}\sum_{j=1}^{r}[(y_{i1j} - \frac{e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}}{1+e^{\Delta^*_{i1j}+\lambda^*_j b_{i1}}})\frac{\partial\Delta^*_{i1j}}{\partial\beta^*_k} + (y_{i2j} - \frac{e^{\Delta^*_{i2j}+\lambda_j b_{i2}}}{1+e^{\Delta^*_{i2j}+\lambda_j b_{i2}}})\frac{\partial\Delta^*_{i2j}}{\partial\beta^*_k}] - \frac{\beta^*_k}{\sigma^2_{\beta^*}}$

where $\frac{\partial\Delta^*_{i1j}}{\partial\beta^*_k} = \frac{\frac{X_{i1jk}e^{\mathbf{X}_{i1j}\beta^*}}{(1+e^{\mathbf{X}_{i1j}\beta^*})^2}}{A_{i1j}}$ from convolution equation of $\Delta^*_{i1j}$, and

$$\frac{\partial\Delta^*_{i2j}}{\partial\beta^*_k} = \frac{\partial\Delta^*_{i2j}}{\partial\Delta_{i2j}}\frac{\partial\Delta_{i2j}}{\partial\beta^*_k}$$

$$\frac{\partial\Delta^*_{i2j}}{\partial\Delta_{i2j}} = B_{i2j}/A_{i2j}$$

$$\frac{\partial \Delta_{i2j}}{\partial \beta_k^*} = \frac{\frac{X_{i1jk} e^{\mathbf{X}_{i1j}\beta^*}}{(1+e^{\mathbf{X}_{i1j}\beta^*})^2}\left(\frac{e^{\Delta_{i2j}}}{1+e^{\Delta_{i2j}}} - \frac{e^{\Delta_{i2j}+\gamma_{i2j}}}{1+e^{\Delta_{i2j}+\gamma_{i2j}}}\right)}{\frac{e^{\Delta_{i2j}}}{(1+e^{\Delta_{i2j}})^2}\frac{1}{1+e^{X_{i1j}\beta^*}} + \frac{e^{\Delta_{i2j}+\gamma_{i2j}}}{(1+e^{\Delta_{i2j}+\gamma_{i2j}})^2}\frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}} \text{ from marginal equation of } \Delta_{i2j}.$$

- $\frac{\partial log f(\alpha_t|\theta_{-\alpha},\mathbf{Y})}{\partial \alpha_{tk}} = \sum_{i=1}^{N}\sum_{j=1}^{r}[(Y_{itj} - \frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial \Delta^*_{itj}}{\partial \alpha_{tk}}] - \frac{\alpha_{tk}}{\sigma^2_\alpha}$

where $\frac{\partial \Delta^*_{itj}}{\partial \alpha_{tk}} = \frac{\partial \Delta^*_{itj}}{\partial \gamma_{itj}}\frac{\partial \gamma_{itj}}{\partial \alpha_{tk}} = \frac{\partial \Delta^*_{itj}}{\partial \gamma_{itj}} Z_{itjk}$

From convolution equation, $\frac{\partial \Delta^*_{itj}}{\partial \gamma_{itj}} = B_{itj}[\frac{\partial \Delta_{itj}}{\partial \gamma_{itj}} + y_{it-1j}]/A_{itj}$

where, from marginal constraint equations,

for $t = 2$,

$$\frac{\partial \Delta_{i2j}}{\partial \gamma_{i2j}} = \frac{-\frac{e^{\Delta_{i2j}+\gamma_{i2j}}}{(1+e^{\Delta_{i2j}+\gamma_{i2j}})^2}\frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}}{\frac{e^{\Delta_{i2j}}}{(1+e^{\Delta_{i2j}})^2}\frac{1}{1+e^{X_{i1j}\beta^*}} + \frac{e^{\Delta_{i2j}+\gamma_{i2j}}}{(1+e^{\Delta_{i2j}+\gamma_{i2j}})^2}\frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}}$$

for $t > 2$,

$$\frac{\partial \Delta_{itj}}{\partial \gamma_{itj}} = -[\sum_{y_{it-1j}=0}^{1} y_{it-1j}\frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{(1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}})^2}P(Y_{it-1j}) + \sum_{y_{it-1j}=0}^{1}\frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}\frac{\partial P(Y_{it-1j})}{\partial \gamma_{itj}}]/C_{itj}$$

$= -[\sum_{y_{it-1j}} y_{it-1j}\frac{e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}}}{(1+e^{\Delta_{itj}+\gamma_{itj}y_{it-1j}})^2}P(Y_{it-1j})]/C_{itj}$, since $\frac{\partial P(Y_{it-1j})}{\partial \gamma_{itj}} = 0$ for p=1.

*Details on sampling from posterior distribution of parameters in MTREM(2)*

For simplicity of notation, for $t \geq 3$, define:

$$\pi^{(t)}_{y_{it-1j},y_{it-2j}} = P(Y_{it-1j} = y_{it-1j}, Y_{it-2j} = y_{it-2j})$$

Note that,

$$\pi^{(t+1)}_{y_{itj},y_{it-1j}} = P(Y_{itj} = y_{itj}, Y_{it-1j} = y_{it-1j})$$

$$= \sum_{y_{it-2j}} P(Y_{itj} = y_{itj}|Y_{it-1j} = y_{it-1j}, Y_{it-2j} = y_{it-2j})\pi^{(t)}_{y_{it-1j},y_{it-2j}}$$

$$\pi^{(3)}_{y_{i2j},y_{i1j}} = P(Y_{i2j}|Y_{i1j}, \theta)P(Y_{i1j}|\theta)$$

$$= \frac{(e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^{y_{i2j}}}{1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}} \frac{(e^{X_{i1j}\beta^*})^{y_{i1j}}}{1+e^{X_{i1j}\beta^*}}$$

*Calculating $\Delta_{itj}$, $\Delta^*_{itj}$ and derivatives:*

- Marginal constraint for initial state intercept, i.e. for $\Delta_{i2j}$:

$$f(\Delta_{i2j}) = -\frac{e^{X_{i2j}\widetilde{\beta}}}{1+e^{X_{i2j}\widetilde{\beta}}} + \sum_{y_{i1j}} \frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}{1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}} \frac{e^{y_{i1j}X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}} = 0$$

$$\frac{\partial f(\Delta_{i2j})}{\partial \Delta_{i2j}} = \sum_{y_{i1j}} \frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^2} \frac{e^{y_{i1j}X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}$$

$\Delta_{i2j}$ is a function of $\widetilde{\gamma}_{i2j}, \widetilde{\beta}$ and $\beta^*$.

- Convolution equations for initial state intercepts, i.e. for $\Delta^*_{i1j}$ and $\Delta^*_{i2j}$:

$$f(\Delta_{i1j}^*) = -\frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}} + \int \frac{e^{\Delta_{i1j}^*+\lambda_j^*\sigma_1 z_i}}{1+e^{\Delta_{i1j}^*+\lambda_j^*\sigma_1 z_i}}\phi(z_i)dz_i = 0$$

$$\frac{\partial f(\Delta_{i1j}^*)}{\partial \Delta_{i1j}^*} = \int \frac{e^{\Delta_{i1j}^*+\lambda_j^*\sigma_1 z_i}}{(1+e^{\Delta_{i1j}^*+\lambda_j^*\sigma_1 z_i})^2}\phi(z_i)dz_i$$

$\Delta_{i1j}^*$ is a function of $\beta^*, \lambda_j^*, \sigma_1$.

$$f(\Delta_{i2j}^*) = -\frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}{1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}} + \int \frac{e^{\Delta_{i2j}^*+\widetilde{\lambda}_j\sigma_2 z_i}}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j\sigma_2 z_i}}\phi(z_i)dz_i = 0$$

$$\frac{\partial f(\Delta_{i2j}^*)}{\partial \Delta_{i2j}^*} = \int \frac{e^{\Delta_{i2j}^*+\widetilde{\lambda}_j\sigma_2 z_i}}{(1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j\sigma_2 z_i})^2}\phi(z_i)dz_i$$

$\Delta_{i2j}^*$ is a function of of $\widetilde{\gamma}_{i2j}, \widetilde{\lambda}_j, \sigma_2, \Delta_{i2j}$ (hence $\widetilde{\beta}$ and $\beta^*$).

- Marginal constraint for $\Delta_{itj}$ when $t \geq 3$:

$$P(Y_{itj}=1) = \sum_{y_{it-1j},y_{it-2j}} P(Y_{itj}=1|Y_{it-1j}=y_{it-1j}, Y_{it-2j}=y_{it-2j})\pi^{(t)}_{y_{it-1j},y_{it-2j}}$$

$$f(\Delta_{itj}) = -\frac{e^{X_{itj}\beta}}{1+e^{X_{itj}\beta}} + \sum_{y_{it-1j},y_{it-2j}} \frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}\pi^{(t)}_{y_{it-1j},y_{it-2j}} = 0$$

$$\frac{\partial f(\Delta_{itj})}{\partial \Delta_{itj}} = \sum_{y_{it-1j},y_{it-2j}} \frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{(1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}})^2}\pi^{(t)}_{y_{it-1j},y_{it-2j}}$$

$\Delta_{itj}$ is a function of $\beta, \gamma_{itj,1}, \gamma_{itj,2}$ and $\pi^{(t)}_{y_{it-1j},y_{it-2j}}$ (hence $\beta^*, \widetilde{\beta}, \widetilde{\gamma}_{i2j}, (\gamma_{ikj,1}, \gamma_{ikj,2}, 3 \leq k \leq t), (\Delta_{ikj}, 2 \leq k \leq t-1)$).

- Convolution equation for $\Delta_{itj}^*$ when $t \geq 3$:

$f(\Delta^*_{itj}) = -\frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}} + \int \frac{e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}{1+e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}\phi(z_i)dz_i{=}0$

$\frac{\partial f(\Delta^*_{itj})}{\partial \Delta^*_{itj}} = \int \frac{e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i}}{(1+e^{\Delta^*_{itj}+\lambda_j\sigma_t z_i})^2}\phi(z_i)dz_i$

$\Delta^*_{itj}$ is a function of $\lambda_j, \sigma_t, \gamma_{itj,1}, \gamma_{itj,2}, (\Delta_{ikj}, 2 \le k \le t)$ (hence $\beta, \beta^*, \widetilde{\beta}, \widetilde{\gamma}_{i2j}, (\gamma_{ikj,1}, \gamma_{ikj,2}, 3 \le k \le t-1)$).

*Full conditional distributions:*

- $f(log(\sigma_1^2), \lambda^*|\theta_{-\mathbf{b}}, \theta_{-\sigma}, \theta_{-\lambda}, \mathbf{Y})$ takes the same form as in p=1.

- $f(log(\sigma_2^2), \widetilde{\lambda}|\theta_{-\mathbf{b}}, \theta_{-\sigma}, \theta_{-\widetilde{\lambda}}, \mathbf{Y}) \propto \Pi(log(\sigma_2^2)) \prod_{j=1}^{r} \Pi(\widetilde{\lambda}_j) \prod_{i=1}^{N} [\int \prod_{j=1}^{r} [P(Y_{i2j}|Y_{i1j}, \theta)] f(b_{i2}|\sigma_2)db_{i2}]$

$$\propto \frac{e^{log(\sigma_2^2)}}{(1+e^{log(\sigma_2^2)})^2} \prod_{j=1}^{r} e^{\frac{-(\widetilde{\lambda}_j-1)^2}{2\sigma^2_{\widetilde{\lambda}}}} \prod_{i=1}^{N} [\int \prod_{j=1}^{r} [(\frac{e^{\Delta^*_{i2j}+\widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta^*_{i2j}+\widetilde{\lambda} b_{i2}}})^{y_{i2j}} (\frac{1}{1+e^{\Delta^*_{i2j}+\widetilde{\lambda}_j b_{i2}}})^{1-y_{i2j}}] \frac{1}{\sigma_2} e^{-\frac{b_{i2}^2}{2\sigma_2^2}} db_{i2}]$$

$$\approx \frac{e^{log(\sigma_2^2)}}{(1+e^{log(\sigma_2^2)})^2} \prod_{j=1}^{r} e^{\frac{-(\widetilde{\lambda}_j-1)^2}{2\sigma^2_{\widetilde{\lambda}}}} \prod_{i=1}^{N} [\sum_l w_l \prod_{j=1}^{r} [\frac{(e^{\Delta^*_{i2j}+\sqrt{2e^{log(\sigma_2^2)}}\widetilde{\lambda}_j z_l})^{y_{i2j}}}{1+e^{\Delta^*_{i2j}+\sqrt{2e^{log(\sigma_2^2)}}\widetilde{\lambda}_j z_l}}]]$$

$$\equiv \frac{e^{log(\sigma_2^2)}}{(1+e^{log(\sigma_2^2)})^2} \prod_{j=1}^{r} e^{\frac{-(\widetilde{\lambda}_j-1)^2}{2\sigma^2_{\widetilde{\lambda}}}} \prod_{i=1}^{N} [E_{i2}]$$

where $E_{i2} = \sum_l w_l \prod_{j=1}^{r} [\frac{(e^{\Delta^*_{i2j}+\sqrt{2e^{log(\sigma_2^2)}}\widetilde{\lambda}_j z_l})^{y_{i2j}}}{1+e^{\Delta^*_{i2j}+\sqrt{2e^{log(\sigma_2^2)}}\widetilde{\lambda}_j z_l}}]$

- For $t > 2$ :

f(log $(\sigma^2), \lambda|\theta_{-\mathbf{b}}, \theta_{-\sigma}, \theta_{-\lambda}, \mathbf{Y}) \propto \prod_{t=3}^{n} \Pi(log(\sigma_t^2)) \prod_{j=1}^{r} \Pi(\lambda_j) \prod_{i=1}^{N} [\int \prod_{j=1}^{r} [P(Y_{itj}|Y_{it-1j}, \theta] f(b_{it}|\sigma_t)db_{it}]]$

$$\propto \prod_{t=3}^{n} \frac{e^{log(\sigma_t^2)}}{(1+e^{log(\sigma_t^2)})^2} \prod_{j=1}^{r} e^{\frac{-(\lambda_j-1)^2}{2\sigma^2_{\lambda}}} \prod_{i=1}^{N} [\int \prod_{j=1}^{r} [(\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{y_{itj}} (\frac{1}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})^{1-y_{itj}}] \frac{1}{\sigma_t} e^{-\frac{b_{it}^2}{2\sigma_t^2}} db_{it}]$$

$$\approx \prod_{t=3}^{n} \frac{e^{log(\sigma_t^2)}}{(1+e^{log(\sigma_t^2)})^2} \prod_{j=1}^{r} e^{\frac{-(\lambda_j-1)^2}{2\sigma_\lambda^2}} \prod_{i=1}^{N}[\sum_l w_l \prod_{j=1}^{r}[\frac{(e^{\Delta_{itj}^*+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}]]$$

$$\equiv \prod_{t=3}^{n} \frac{e^{log(\sigma_t^2)}}{(1+e^{log(\sigma_t^2)})^2} \prod_{j=1}^{r} e^{\frac{-(\lambda_j-1)^2}{2\sigma_\lambda^2}} \prod_{i=1}^{N}[E_{it}]$$

where $E_{it} = \sum_l w_l \prod_{j=1}^{r}[\frac{(e^{\Delta_{itj}^*+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\sqrt{2e^{log(\sigma_t^2)}}\lambda_j z_l}}]$

- $f(b_{it}|\theta_{-\mathbf{b}}, \mathbf{Y})$ for $t \geq 1$ takes similar forms to the ones for p=1.

- $f(\beta^*|\theta_{-\beta^*}, \mathbf{Y}) \propto \prod_i \prod_j [(\frac{e^{\Delta_{i1j}^*+\lambda_j^* b_{i1}}}{1+e^{\Delta_{i1j}^*+\lambda_j^* b_{i1}}})^{y_{i1j}} (\frac{1}{1+e^{\Delta_{i1j}^*+\lambda_j^* b_{i1}}})^{1-y_{i1j}} (\frac{e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}})^{y_{i2j}} (\frac{1}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}})^{1-y_{i2j}}]$

$*[\prod_i \prod_j \prod_{t=3} \frac{(e^{\Delta_{itj}^*+\lambda_j b_{it}})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\lambda_j b_{it}}}] e^{-\frac{\sum_k \beta_k^{*2}}{2\sigma_{\beta^*}^2}}$

- $f(\widetilde{\beta}|\theta_{-\widetilde{\beta}}, \mathbf{Y}) \propto \prod_i \prod_j [(\frac{e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}})^{y_{i2j}} (\frac{1}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}})^{1-y_{i2j}}][\prod_i \prod_j \prod_{t=3} \frac{(e^{\Delta_{itj}^*+\lambda_j b_{it}})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\lambda_j b_{it}}}] e^{-\frac{\sum_k \widetilde{\beta}_k^2}{2\sigma_{\widetilde{\beta}}^2}}$

- $f(\beta|\theta_{-\beta}, \mathbf{Y}) \propto [\prod_i \prod_j \prod_{t=3} \frac{(e^{\Delta_{itj}^*+\lambda_j b_{it}})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\lambda_j b_{it}}}] e^{-\frac{\sum_k \beta_k^2}{2\sigma_\beta^2}}$

- $f(\widetilde{\alpha}|\theta_{-\widetilde{\alpha}}, \mathbf{Y}) \propto \prod_i \prod_j [(\frac{e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}})^{y_{i2j}} (\frac{1}{1+e^{\Delta_{i2j}^*+\widetilde{\lambda}_j b_{i2}}})^{1-y_{i2j}}][\prod_i \prod_j \prod_{t=3} \frac{(e^{\Delta_{itj}^*+\lambda_j b_{it}})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\lambda_j b_{it}}}] e^{-\frac{\sum_k \widetilde{\alpha}_{2k}^2}{2\sigma_{\widetilde{\alpha}}^2}}$

- $f(\alpha_t|\theta_{-\alpha}, \mathbf{Y}) \propto [\prod_i \prod_j \frac{(e^{\Delta_{itj}^*+\lambda_j b_{it}})^{y_{itj}}}{1+e^{\Delta_{itj}^*+\lambda_j b_{it}}}] e^{-\frac{\sum_k \sum_{p=1}^2 \alpha_{tk,p}^2}{2\sigma_\alpha^2}}$

*Derivatives of full conditional distributions:*

- $\frac{\partial log f(log(\sigma_1^2),\lambda^*|\theta_{-\sigma_{\mathbf{1}}},\theta_{-\lambda^*},\theta_{-\mathbf{b}},\mathbf{Y})}{\partial log(\sigma_1^2)}$ takes the same form as p=1.

- $\frac{\partial log f(log(\sigma_1^2),\lambda^*|\theta_{-\sigma_{\mathbf{1}}},\theta_{-\lambda^*},\theta_{-\mathbf{b}},\mathbf{Y})}{\partial \lambda_m^*}$ takes the same form as p=1.

- $\frac{\partial log f(\beta^*|\theta_{-\beta^*},\mathbf{Y})}{\partial \beta_k^*} = \sum_i \sum_j [(Y_{i1j} - \frac{e^{\Delta_{i1j}^* + \lambda_j^* b_{i1}}}{1+e^{\Delta_{i1j}^* + \lambda_j^* b_{i1}}}) \frac{\partial \Delta_{i1j}^*}{\partial \beta_k^*} + (Y_{i2j} - \frac{e^{\Delta_{i2j}^* + \widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta_{i2j}^* + \widetilde{\lambda}_j b_{i2}}}) \frac{\partial \Delta_{i2j}^*}{\partial \beta_k^*}] +$

$\sum_i \sum_j \sum_{t=3} [(Y_{itj} - \frac{e^{\Delta_{itj}^* + \lambda_j b_{it}}}{1+e^{\Delta_{itj}^* + \lambda_j b_{it}}}) \frac{\partial \Delta_{itj}^*}{\partial \beta_k^*}] - \frac{\beta_k^*}{\sigma_{\beta}^{2\,*}}$

where $\frac{\partial \Delta_{i1j}^*}{\partial \beta_k^*} = \frac{X_{i1jk} e^{X_{i1j}\beta^*}/(1+e^{X_{i1j}\beta^*})^2}{\int \frac{e^{\Delta_{i1j}^* + \lambda_j^* \sigma_1 z_i}}{(1+e^{\Delta_{i1j}^* + \lambda_j^* \sigma_1 z_i})^2} \phi(z_i) dz_i}$

$$\frac{\partial \Delta_{i2j}^*}{\partial \beta_k^*} = \frac{\partial \Delta_{i2j}^*}{\partial \Delta_{i2j}} \frac{\partial \Delta_{i2j}}{\partial \beta_k^*}$$

$$\frac{\partial \Delta_{i2j}^*}{\partial \Delta_{i2j}} = \frac{\frac{e^{\Delta_{i2j} + \widetilde{\gamma}_{i2j} y_{i1j}}}{(1+e^{\Delta_{i2j} + \widetilde{\gamma}_{i2j} y_{i1j}})^2}}{\int \frac{e^{\Delta_{i2j}^* + \widetilde{\lambda}_j \sigma_2 z_i}}{(1+e^{\Delta_{i2j}^* + \widetilde{\lambda}_j \sigma_2 z_i})^2} \phi(z_i) dz_i}$$

$$\frac{\partial \Delta_{i2j}}{\partial \beta_k^*} = \frac{\frac{X_{i1jk} e^{X_{i1j}\beta^*}}{(1+e^{X_{i2j}\beta^*})^2} [\frac{e^{\Delta_{i2j}}}{1+e^{\Delta_{i2j}}} - \frac{e^{\Delta_{i2j} + \widetilde{\gamma}_{i2j}}}{1+e^{\Delta_{i2j} + \widetilde{\gamma}_{i2j}}}]}{[\frac{e^{\Delta_{i2j}}}{(1+e^{\Delta_{i2j}})^2} \frac{1}{1+e^{X_{i1j}\beta^*}} + \frac{e^{\Delta_{i2j} + \widetilde{\gamma}_{i2j}}}{(1+e^{\Delta_{i2j} + \widetilde{\gamma}_{i2j}})^2} \frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}]}$$

For $t \geq 3$,

$$\frac{\partial \Delta_{itj}^*}{\partial \beta_k^*} = \frac{\partial \Delta_{itj}^*}{\partial \Delta_{itj}} \frac{\partial \Delta_{itj}}{\partial \beta_k^*}$$

$$\frac{\partial \Delta_{itj}^*}{\partial \Delta_{itj}} = \frac{\frac{e^{\Delta_{itj} + \gamma_{itj,1} y_{it-1j} + \gamma_{itj,2} y_{it-2j}}}{(1+e^{\Delta_{itj} + \gamma_{itj,1} y_{it-1j} + \gamma_{itj,2} y_{it-2j}})^2}}{\int \frac{e^{\Delta_{itj}^* + \lambda_j \sigma_t z_i}}{(1+e^{\Delta_{itj}^* + \lambda_j \sigma_t z_i})^2} \phi(z_i) dz_i}$$

$$\frac{\partial \Delta_{itj}}{\partial \beta_k^*} = \frac{- \sum_{y_{it-1j}, y_{it-2j}} \frac{e^{\Delta_{itj} + \gamma_{itj,1} y_{it-1j} + \gamma_{itj,2} y_{it-2j}}}{1+e^{\Delta_{itj} + \gamma_{itj,1} y_{it-1j} + \gamma_{itj,2} y_{it-2j}}} \frac{\partial \pi_{y_{it-1j}, y_{it-2j}}^{(t)}}{\partial \beta_k^*}}{\sum_{y_{it-1j}, y_{it-2j}} \frac{e^{\Delta_{itj} + \gamma_{itj,1} y_{it-1j} + \gamma_{itj,2} y_{it-2j}}}{(1+e^{\Delta_{itj} + \gamma_{itj,1} y_{it-1j} + \gamma_{itj,2} y_{it-2j}})^2} \pi_{y_{it-1j}, y_{it-2j}}^{(t)}}$$

- $\frac{\partial logf(\widetilde{\beta}|\theta_{-\widetilde{\beta}},\mathbf{Y})}{\partial\widetilde{\beta}_k} = \sum_i\sum_j[(Y_{i2j} - \frac{e^{\Delta^*_{i2j}+\widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta^*_{i2j}+\widetilde{\lambda}_j b_{i2}}})\frac{\partial\Delta^*_{i2j}}{\partial\widetilde{\beta}_k}] + \sum_i\sum_j\sum_{t=3}[(Y_{itj} - \frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial\Delta^*_{itj}}{\partial\widetilde{\beta}_k}] - \frac{\widetilde{\beta}_k}{\sigma^2_{\widetilde{\beta}}}$

where $\frac{\partial\Delta^*_{i2j}}{\partial\widetilde{\beta}_k} = \frac{\partial\Delta^*_{i2j}}{\partial\Delta_{i2j}}\frac{\partial\Delta_{i2j}}{\partial\widetilde{\beta}_k}$

$$\frac{\partial\Delta^*_{i2j}}{\partial\Delta_{i2j}} = \frac{\frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^2}}{\int\frac{e^{\Delta^*_{i2j}+\widetilde{\lambda}_j\sigma_2 z_i}}{(1+e^{\Delta^*_{i2j}+\widetilde{\lambda}_j\sigma_2 z_i})^2}\phi(z_i)dz_i}$$

$$\frac{\partial\Delta_{i2j}}{\partial\widetilde{\beta}_k} = \frac{\frac{X_{i2jk}e^{X_{i2j}\widetilde{\beta}}}{(1+e^{X_{i2j}\widetilde{\beta}})^2}}{\sum_{yi1j}\frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^2}\frac{(e^{X_{i1j}\beta^*})^{y_{i1j}}}{1+e^{X_{i1j}\beta^*}}}$$

For $t \geq 3$,

$$\frac{\partial\Delta^*_{itj}}{\partial\widetilde{\beta}_k} = \frac{\partial\Delta^*_{itj}}{\partial\Delta_{itj}}\frac{\partial\Delta_{itj}}{\partial\widetilde{\beta}_k}$$

$$\frac{\partial\Delta_{itj}}{\partial\widetilde{\beta}_k} = \frac{-\sum_{y_{it-1j},y_{it-2j}}\frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}\frac{\partial\pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial\widetilde{\beta}_k}}{\sum_{y_{it-1j},y_{it-2j}}\frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{(1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}})^2}\pi^{(t)}_{y_{it-1j},y_{it-2j}}}$$

- $\frac{\partial logf(\beta|\theta_{-\beta},\mathbf{Y})}{\partial\beta_k} = \sum_i\sum_j\sum_{t=3}[(Y_{itj} - \frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial\Delta^*_{itj}}{\partial\beta_k}] - \frac{\beta_k}{\sigma^2_\beta}$

$$\frac{\partial\Delta^*_{itj}}{\partial\beta_k} = \frac{\partial\Delta^*_{itj}}{\partial\Delta_{itj}}\frac{\partial\Delta_{itj}}{\partial\beta_k}$$

$$\frac{\partial\Delta_{itj}}{\partial\beta_k} = \frac{X_{itjk}\frac{e^{X_{itj}\beta}}{(1+e^{X_{itj}\beta})^2} - \sum_{y_{it-1j},y_{it-2j}}\frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}\frac{\partial\pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial\beta_k}}{\sum_{y_{it-1j},y_{it-2j}}\frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{(1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}})^2}\pi^{(t)}_{y_{it-1j},y_{it-2j}}}$$

- $\frac{\partial log f(\widetilde{\alpha}_2|\theta_{-\widetilde{\alpha}_{\mathbf{2}}},\mathbf{Y})}{\partial\widetilde{\alpha}_{2k}} = \sum_i\sum_j[(Y_{i2j}-\frac{e^{\Delta^*_{i2j}+\widetilde{\lambda}_j b_{i2}}}{1+e^{\Delta^*_{i2j}+\widetilde{\lambda}_j b_{i2}}})\frac{\partial\Delta^*_{i2j}}{\partial\widetilde{\alpha}_{2k}}]+\sum_i\sum_j\sum_{t=3}[(Y_{itj}-\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial\Delta^*_{itj}}{\partial\widetilde{\alpha}_{2k}}]-\frac{\widetilde{\alpha}_{2k}}{\sigma^2_{\widetilde{\alpha}_2}}$

$$\frac{\partial\Delta^*_{i2j}}{\partial\widetilde{\alpha}_{2k}} = \frac{\partial\Delta^*_{i2j}}{\partial\widetilde{\gamma}_{i2j}}\frac{\partial\widetilde{\gamma}_{i2j}}{\partial\widetilde{\alpha}_{2k}}$$

$$\frac{\partial\Delta^*_{i2j}}{\partial\widetilde{\gamma}_{i2j}} = \frac{(\frac{\partial\Delta_{i2j}}{\partial\widetilde{\gamma}_{i2j}}+y_{i1j})\frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^2}}{\int\frac{e^{\Delta^*_{i2j}+\widetilde{\lambda}_j\sigma_2 z_i}}{(1+e^{\Delta^*_{i2j}+\widetilde{\lambda}_j\sigma_2 z_i})^2}\phi(z_i)dz_i}$$

$$\frac{\partial\Delta_{i2j}}{\partial\widetilde{\gamma}_{i2j}} = \frac{\frac{-e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}}}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}})^2}\frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}}{[\frac{e^{\Delta_{i2j}}}{(1+e^{\Delta_{i2j}})^2}\frac{1}{1+e^{X_{i1j}\beta^*}}+\frac{e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}}}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}})^2}\frac{e^{X_{i1j}\beta^*}}{1+e^{X_{i1j}\beta^*}}]}$$

$$\frac{\partial\widetilde{\gamma}_{i2j}}{\partial\widetilde{\alpha}_{2k}} = Z_{i2jk}$$

- for $t=3$:

$$\frac{\partial log f(\alpha_3|\theta_{-\alpha},\mathbf{Y})}{\partial\alpha_{3k,l}} = [\sum_i\sum_j\sum_{t=3}(y_{itj}-\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial\Delta^*_{itj}}{\partial\alpha_{3k,l}}]-\frac{\alpha_{3k,l}}{\sigma^2_\alpha}$$

for $t\geq 4$

$$\frac{\partial log f(\alpha_t|\theta_{-\alpha},\mathbf{Y})}{\partial\alpha_{tk,l}} = [\sum_i\sum_j(y_{itj}-\frac{e^{\Delta^*_{itj}+\lambda_j b_{it}}}{1+e^{\Delta^*_{itj}+\lambda_j b_{it}}})\frac{\partial\Delta^*_{itj}}{\partial\alpha_{tk,l}}]-\frac{\alpha_{tk,l}}{\sigma^2_\alpha}$$

where

$$\frac{\partial\Delta^*_{itj}}{\partial\alpha_{tk,l}} = \frac{\partial\Delta^*_{itj}}{\partial\gamma_{itj,l}}\frac{\partial\gamma_{itj,l}}{\partial\alpha_{tk,l}} = \frac{\partial\Delta^*_{itj}}{\partial\gamma_{itj,l}}Z_{itjk,l}$$

$$\frac{\partial\Delta^*_{itj}}{\partial\gamma_{itj,l}} = \frac{(\frac{\partial\Delta_{itj}}{\partial\gamma_{itj,l}}+y_{it-lj})\frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{(1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}})^2}}{\frac{1}{\sqrt{\pi}}\sum_k w_k\frac{e^{\Delta^*_{itj}+\sqrt{2}\lambda_j\sigma_t z_k}}{(1+e^{\Delta^*_{itj}+\sqrt{2}\lambda_j\sigma_t z_k})^2}}$$

$$\frac{\partial \Delta_{itj}}{\partial \gamma_{itj,l}} = \frac{-\sum_{y_{it-1j},y_{it-2j}} y_{it-lj} \frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{(1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}})^2}\pi_{y_{it-1j},y_{it-2j}}}{\sum_{y_{it-1j},y_{it-2j}} \frac{e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}}}{(1+e^{\Delta_{itj}+\gamma_{itj,1}y_{it-1j}+\gamma_{itj,2}y_{it-2j}})^2}\pi_{y_{it-1j},y_{it-2j}}}$$

*Derivatives of joint distributions:*

Unlike p=1, we need to calculate and update bivariate probabilities. These probabilities are used in the calculation of marginal constraints, and derivatives are required for the Hybrid step of some parameters.

Recall that, for $t \geq 4$,

$$\pi^{(3)}_{y_{i2j},y_{i1j}} = P(Y_{i2j}|Y_{i1j},\theta)P(Y_{i1j}|\theta) = \frac{(e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^{y_{i2j}}}{1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}\frac{(e^{X_{i1j}\beta^*})^{y_{i1j}}}{1+e^{X_{i1j}\beta^*}}$$

- $\frac{\partial \pi^{(3)}_{y_{i2j},y_{i1j}}}{\partial \widetilde{\gamma}_2} = \frac{(e^{X_{i1j}\beta^*})^{y_{i1j}}}{1+e^{X_{i1j}\beta^*}}\frac{[(\frac{\partial \Delta_{i2j}}{\partial \widetilde{\gamma}_{i2j}}+y_{i1j})(e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^{y_{i2j}}(y_{i2j}+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}(y_{i2j}-1))]}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^2}$

- $\frac{\partial \pi^{(3)}_{y_{i2j},y_{i1j}}}{\partial \widetilde{\beta}_k} = \frac{(e^{X_{i1j}\beta^*})^{y_{i1j}}}{1+e^{X_{i1j}\beta^*}}\frac{[(\frac{\partial \Delta_{i2j}}{\partial \widetilde{\beta}_k})(e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^{y_{i2j}}(y_{i2j}+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}(y_{i2j}-1))]}{(1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}})^2}$

- $\frac{\partial \pi^{(3)}_{y_{i2j},y_{i1j}}}{\partial \beta^*_k} = \frac{(e^{X_{i1j}\beta^*})^{y_{i1j}}}{1+e^{X_{i1j}\beta^*}}\frac{(e^{\Delta^*_{i2j}+\widetilde{\gamma}_2 y_{i1j}})^{y_{i2j}}}{1+e^{\Delta^*_{i2j}+\widetilde{\gamma}_2 y_{i1j}}}[\frac{X_{i1jk}}{1+e^{X_{i1j}\beta^*}}(y_{i1j}+e^{X_{i1j}\beta^*}(y_{i1j}-1))+\frac{\frac{\partial \Delta_{i2j}}{\partial \beta^*_k}(y_{i2j}+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}(y_{i2j}-1))}{1+e^{\Delta_{i2j}+\widetilde{\gamma}_{i2j}y_{i1j}}}]$

The derivative of $\pi^{(3)}_{y_{i2j},y_{i1j}}$ with respect to any other parameter is zero. After calculating $(\pi^{(3)}, \frac{\partial \pi^{(3)}}{\partial \theta})$, we need to update $(\pi^{(t)}, \frac{\partial \pi^{(t)}}{\partial \theta})$ for $t \geq 4$.

Recall that,

$$\pi^{(t)}_{y_{it-1j},y_{it-2j}} = P(Y_{it-1j} = y_{it-1j}, Y_{it-2j} = y_{it-2j})$$

$$= \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}}\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

- $$\frac{\partial\pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial\widetilde{\gamma}_2} = \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}} \frac{\partial\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}}{\partial\widetilde{\gamma}_2} +$$

$$\sum_{y_{it-3j}} \frac{(\frac{\partial\Delta_{it-1j}}{\partial\widetilde{\gamma}_2})(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}[y_{it-1j}+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}(y_{it-1j}-1)]}{(1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^2} *$$

$$\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

- $$\frac{\partial\pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial\widetilde{\beta}_k} = \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}} \frac{\partial\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}}{\partial\widetilde{\beta}_k} +$$

$$\sum_{y_{it-3j}} \frac{(\frac{\partial\Delta_{it-1j}}{\partial\widetilde{\beta}_k})(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}[y_{it-1j}+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}(y_{it-1j}-1)]}{(1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^2} *$$

$$\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

- $$\frac{\partial\pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial\beta^*_k} = \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}} \frac{\partial\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}}{\partial\beta^*_k} +$$

$$\sum_{y_{it-3j}} \frac{(\frac{\partial\Delta_{it-1j}}{\partial\beta^*_k})(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}[y_{it-1j}+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}(y_{it-1j}-1)]}{(1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^2} *$$

$$\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

- $$\frac{\partial\pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial\beta_k} = \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}} \frac{\partial\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}}{\partial\beta_k} +$$

$$\sum_{y_{it-3j}} \frac{(\frac{\partial \Delta_{it-1j}}{\partial \beta_k})(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}[y_{it-1j}+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}(y_{it-1j}-1)]}{(1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^2} *$$

$$\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

- $\frac{\partial \pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial \gamma_{it-1j,1}} = \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}} \frac{\partial \pi^{(t-1)}_{y_{it-2j},y_{it-3j}}}{\partial \gamma_{it-1j,1}} +$

$$\sum_{y_{it-3j}} \frac{(\frac{\partial \Delta_{it-1j}}{\partial \gamma_{it-1j,1}}+y_{it-2j})(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}[y_{it-1j}+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}(y_{it-1j}-1)]}{(1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^2} *$$

$$\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

- $\frac{\partial \pi^{(t)}_{y_{it-1j},y_{it-2j}}}{\partial \gamma_{it-1j,2}} = \sum_{y_{it-3j}} \frac{(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}}{1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}} \frac{\partial \pi^{(t-1)}_{y_{it-2j},y_{it-3j}}}{\partial \gamma_{it-1j,1}} +$

$$\sum_{y_{it-3j}} \frac{(\frac{\partial \Delta_{it-1j}}{\partial \gamma_{it-1j,2}}+y_{it-3j})(e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^{y_{it-1j}}[y_{it-1j}+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}}(y_{it-1j}-1)]}{(1+e^{\Delta_{it-1j}+\gamma_{it-1j,1}y_{it-2j}+\gamma_{it-1j,2}y_{it-3j}})^2} *$$

$$\pi^{(t-1)}_{y_{it-2j},y_{it-3j}}$$

The derivative of $\pi^{(t)}_{y_{it-1j},y_{it-2j}}$ with respect to any other parameter is zero.

# Appendix D

*Form and parameterization of the model for the missing covariates:*

$logitP(x_{t,3} = 1|x_{t,2}, \mathbf{X}_{t-1}, \psi_3) = \psi_{31}+\psi_{32}x_{t,2}+\psi_{33}x_{t-1,3}+\psi_{34}x_{t-1,4}+\psi_{35}x_{t-1,5}+\psi_{36}x_{t-1,6}+\psi_{37}x_{t-1,7} + \psi_{38}x_{t,10} + \psi_{39}x_{t,11}$

$logitP(x_{t,4} = 1|x_{t,2}, \mathbf{X}_{t-1}, \psi_3) = \psi_{3,10} + \psi_{32}x_{t,2} + \psi_{33}x_{t-1,3} + \psi_{34}x_{t-1,4} + \psi_{35}x_{t-1,5} + \psi_{36}x_{t-1,6} + \psi_{37}x_{t-1,7} + \psi_{38}x_{t,10} + \psi_{39}x_{t,11}$

where $\psi_3 = (\psi_{31}, \psi_{32}, \psi_{33}, \psi_{34}, \psi_{35}, \psi_{36}, \psi_{37}, \psi_{38}, \psi_{39}, \psi_{3,10})$.

$logitP(x_{t,5} = 1|x_2, x_{t,3}, x_{t,4}, \mathbf{X}_{t-1}, \psi_5) = \psi_{51} + \psi_{52}x_{t,2} + \psi_{53}x_{t-1,3} + \psi_{54}x_{t-1,4} + \psi_{55}x_{t-1,5} + \psi_{56}x_{t-1,6} + \psi_{57}x_{t-1,7} + \psi_{58}x_{t,10} + \psi_{59}x_{t,11} + \psi_{5,10}x_{t,3} + \psi_{5,11}x_{t,4}$

$logitP(x_{t,6} = 1|x_{t,2}, x_{t,3}, x_{t,4}, x_{t,5}, \mathbf{X}_{t-1}, \psi_6) = \psi_{61} + \psi_{62}x_{t,2} + \psi_{63}x_{t-1,3} + \psi_{64}x_{t-1,4} + \psi_{65}x_{t-1,5} + \psi_{66}x_{t-1,6} + \psi_{67}x_{t-1,7} + \psi_{68}x_{t,10} + \psi_{69}x_{t,11} + \psi_{6,10}x_{t,3} + \psi_{6,11}x_{t,4} + \psi_{6,12}x_{t,5}$

$logitP(x_{t,7} = 1|x_{t,2}, x_{t,3}, x_{t,4}, x_{t,5}, \mathbf{X}_{t-1}, \psi_6) = \psi_{6,13} + \psi_{62}x_{t,2} + \psi_{63}x_{t-1,3} + \psi_{64}x_{t-1,4} + \psi_{65}x_{t-1,5} + \psi_{66}x_{t-1,6} + \psi_{67}x_{t-1,7} + \psi_{68}x_{t,10} + \psi_{69}x_{t,11} + \psi_{6,10}x_{t,3} + \psi_{6,11}x_{t,4} + \psi_{6,12}x_{t,5}$