Supplemental material for “Modelling multivariate, overdispersed count data with correlated and non-normal heterogeneity effects”

Iraj Kazemi[1] and Fatemeh Hassanzadeh[2]

December 2020

The material contained herein is supplementary to the article named in the title and published in SORT-Statistics and Operations Research Transactions Volume 44(2).

[1] Department of Statistics, University of Isfahan, Iran. Tel.: +98-31-37934596.
[2] Department of Statistics, University of Khansar, Iran. Tel.: +98-31-57995312.

Appendix A

The mean and variance equations in (10) are derived as follows. By letting $\exp(\mathbf{x}_{it}'\boldsymbol{\beta}+u_{it}) = \mu_{it}\omega_{it}$ an element of vector $E(Y_i)$ is given by $E(Y_{it}) = \mu_{it}\mu_{\omega_{it}}$, where $\mu_{\omega_{it}} = E(\omega_{it})$. Using the usual formual to partition variability the variance equals $var(Y_{it}) = \mu_{it}^2\sigma_{\omega_{tt}} + \mu_{it}\mu_{\omega_{it}}$. Similarly for $t \neq s$,

$$\begin{aligned} cov(Y_{it}, Y_{is}) &= E\left(E\left(cov\left(Y_{it}, Y_{is}|\omega_{it}, \omega_{is}\right)\right)\right) \\ &+ cov\left(E\left(E\left(Y_{it}|\omega_{it}\right)\right), E\left(E\left(Y_{is}|\omega_{is}\right)\right)\right) \\ &+ E\left(cov\left(E\left(Y_{it}|\omega_{it}\right), E\left(Y_{is}|\omega_{is}\right)\right)\right). \end{aligned}$$

It is straightforward to show that the first part is 0, the second part equals $cov(\mu_{it}\mu_{\omega_{it}}, \mu_{is}\mu_{\omega_{is}}) = 0$, and the last part is $E(cov(\mu_{it}\omega_{it}, \mu_{is}\omega_{is})) = \mu_{it}\mu_{is}\sigma_{\omega_{ts}}$. Reorganizing terms leads to (10).

Appendix B

McMC PSN. For the PSN model, after some algebra, we derive the following complete conditional posteriors

$$\begin{aligned}
&i)\ \boldsymbol{\beta}|\mathbf{y}, others \sim N_k\left(\mathbf{C}_{\boldsymbol{\beta}}\left\{\sigma^{-2}\textstyle\sum_{i=1}^n \mathbf{X}_i'(\boldsymbol{\theta}_i - \delta\mathbf{z}_i) + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\boldsymbol{\beta}_0\right\}, \mathbf{C}_{\boldsymbol{\beta}}\right),\\
&ii)\ \sigma^2|\mathbf{y}, others \sim IG\left(\frac{nT}{2}+\nu_0, \frac{1}{2}\textstyle\sum_{i=1}^n\sum_{t=1}^T\left(\theta_{it} - \mathbf{x}_{it}'\boldsymbol{\beta} - \delta z_{it}\right)^2 + \nu_0\right),\\
&iii)\ \delta|\mathbf{y}, others \sim N\left(\frac{c_\delta}{\sigma_\delta^{-2} + \sigma^{-2}\sum_{i=1}^n\sum_{t=1}^T z_{it}^2}, \frac{1}{\sigma_\delta^{-2} + \sigma^{-2}\sum_{i=1}^n\sum_{t=1}^T z_{it}^2}\right),\\
&iv)\ z_{it}|\mathbf{y}, others \overset{ind}{\sim} TN\left(\frac{\delta\left(\theta_{it} - \mathbf{x}_{it}'\boldsymbol{\beta}\right)}{\sigma^2+\delta^2}, \frac{\sigma^2}{\sigma^2+\delta^2}\right) I(0,\infty),\\
&v)\ \pi\left(\theta_{it}|\mathbf{y}, others\right) \propto e^{-e^{\theta_{it}} - \frac{1}{2\sigma^2}\left[\theta_{it} - \left(\mathbf{x}_{it}'\boldsymbol{\beta} - \delta z_{it} + \sigma^2 y_{it}\right)\right]^2},
\end{aligned}$$

where $\mathbf{C}_{\boldsymbol{\beta}} = \left(\sigma^{-2}\sum_{i=1}^n \mathbf{X}_i'\mathbf{X}_i + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\right)^{-1}$, $c_\delta = \sigma_\delta^{-2}\delta_0 + \sigma^{-2}\sum_{i=1}^n\sum_{t=1}^T z_{it}\left(\theta_{it} - \mathbf{x}_{it}'\boldsymbol{\beta}\right)$, $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \cdots, \mathbf{x}_{in})'$ and $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n)'$ with $\mathbf{y}_i = (\mathbf{y}_{i1}, \mathbf{y}_{i2}, \cdots, \mathbf{y}_{iT})'$. It is seen that all conditional posteriors, except for θ_{it}, are analytically in closed forms of known distributions and thus simulation is easy. To generate θ_{it} for subject $i = 1, \cdots, n$ and at time $t = 1, \cdots, T$, a Metropolis-within-Gibbs algorithm, which is available in the OpenBugs software, is used.
The Gibbs sampling to fit mixed Poisson models with multivariate skew-normal mixing priors are given as follow.

- **McMC PMSN1**. Using the hierarchical form for PMSN1 and adopting the priors, we derive the related complete conditional posteriors as follow

$$\begin{aligned}
&i)\ \boldsymbol{\beta}|\mathbf{y}, others \sim N_k\left(\mathbf{C}_{\boldsymbol{\beta}}\left\{\textstyle\sum_{i=1}^n \mathbf{X}_i'\mathbf{V}^{-1}\left(\boldsymbol{\theta}_i - \boldsymbol{\delta} z_i\right) + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\boldsymbol{\beta}_0\right\}, \mathbf{C}_{\boldsymbol{\beta}}\right),\\
&ii)\ \mathbf{V}|\mathbf{y}, others \sim IW_T\left(\mathbf{C}_{\mathbf{V}}, n+m\right),\\
&iii)\ \boldsymbol{\delta}|\mathbf{y}, others \sim N\left(\mathbf{C}_{\boldsymbol{\delta}}\left(\mathbf{V}_{\boldsymbol{\delta}}^{-1}\boldsymbol{\delta}_0 + \textstyle\sum_{i=1}^n z_i\mathbf{V}^{-1}\left(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta}\right)\right), \mathbf{C}_{\boldsymbol{\delta}}\right),\\
&iv)\ z_i|\mathbf{y}, others \sim TN\left(\frac{\boldsymbol{\delta}'\mathbf{V}^{-1}\left(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta}\right)}{\boldsymbol{\delta}'\mathbf{V}^{-1}\boldsymbol{\delta}+1}, \frac{1}{\boldsymbol{\delta}'\mathbf{V}^{-1}\boldsymbol{\delta}+1}\right) I(0,\infty),\\
&v)\ \pi\left(\boldsymbol{\theta}_i|\mathbf{y}, others\right) \propto \left\{\textstyle\prod_{t=1}^T e^{-e^{\theta_{it}} + \theta_{it}y_{it}}\right\} e^{-\frac{1}{2}c_{\boldsymbol{\theta}_i}},
\end{aligned}$$

where $\mathbf{C}_{\boldsymbol{\delta}} = \left(\mathbf{V}_{\boldsymbol{\delta}}^{-1} + \mathbf{V}^{-1}\sum_{i=1}^{n} z_i^2\right)^{-1}$, $c_{\boldsymbol{\theta}_i} = (\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i)' \mathbf{V}^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i)$, $\mathbf{C}_{\mathbf{V}} = \boldsymbol{\Omega} + \sum_{i=1}^{n}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i)(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i)'$, and $\mathbf{C}_{\boldsymbol{\beta}} = \left(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{V}^{-1}\mathbf{X}_i + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\right)^{-1}$, and $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \cdots, \mathbf{x}_{iT})'$.

- **McMC PMSN2**. Using the hierarchical form for PMSN2 and adopting similar priors given in PMSN1 for $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$, $IW_T(\boldsymbol{\Omega}_\varepsilon, m_\varepsilon)$ for matrix $\mathbf{V}_\varepsilon$, the inverse-Gamma $IG(\gamma, \gamma)$ for the variance component σ_α^2, the complete conditional posteriors are derived as

$$\begin{aligned}
&i)\ \boldsymbol{\beta}|\mathbf{y}, others \sim N_k(\mathbf{d}_{\boldsymbol{\beta}}, \mathbf{C}_{\boldsymbol{\beta}}),\\
&ii)\ \mathbf{V}_\varepsilon|\mathbf{y}, others \sim IW_T(\mathbf{C}_{\mathbf{V}_\varepsilon}, n + m_\varepsilon),\\
&iii)\ \sigma_\alpha^2|\mathbf{y}, others \sim IG\left(\frac{n}{2} + \gamma, \frac{1}{2}\sum_{i=1}^{n} \alpha_i^2 + \gamma\right),\\
&iv)\ \boldsymbol{\delta}|\mathbf{y}, others \sim N_T(\mathbf{d}_{\boldsymbol{\delta}}, \mathbf{C}_{\boldsymbol{\delta}}),\\
&v)\ z_i|\mathbf{y}, others \sim TN\left(\frac{c_{z_i}}{\boldsymbol{\delta}'\mathbf{V}_\varepsilon^{-1}\boldsymbol{\delta} + 1}, \frac{1}{\boldsymbol{\delta}'\mathbf{V}_\varepsilon^{-1}\boldsymbol{\delta} + 1}\right) I(0, \infty),\\
&vi)\ \alpha_i|\mathbf{y}, others \sim N\left(\frac{\mathbf{1}_T'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i)}{\mathbf{1}_T'\mathbf{V}_\varepsilon^{-1}\mathbf{1}_T + \sigma_\alpha^{-2}}, \frac{1}{\mathbf{1}_T'\mathbf{V}_\varepsilon^{-1}\mathbf{1}_T + \sigma_\alpha^{-2}}\right),\\
&vii)\ \pi(\boldsymbol{\theta}_i|\mathbf{y}, others) \propto \left\{\prod_{t=1}^{T} e^{-e^{\theta_{it}} + \theta_{it} y_{it}}\right\} e^{-\frac{1}{2} c_{\boldsymbol{\theta}_i}},
\end{aligned}$$

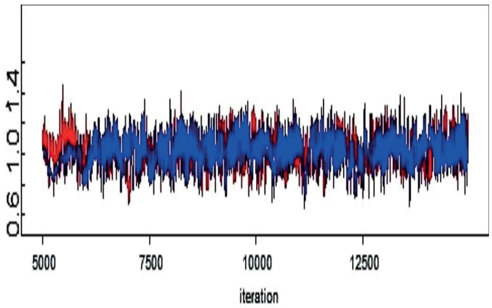
where $\mathbf{C}_{\boldsymbol{\beta}} = \left(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{V}_\varepsilon^{-1}\mathbf{X}_i + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\right)^{-1}$, $\mathbf{C}_{\boldsymbol{\delta}} = \left(\mathbf{V}_{\boldsymbol{\delta}}^{-1} + \mathbf{V}_\varepsilon^{-1}\sum_{i=1}^{n} z_i^2\right)^{-1}$, $c_{z_i} = \boldsymbol{\delta}'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \alpha_i\mathbf{1}_T)$, $c_{\boldsymbol{\theta}_i} = (\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i - \alpha_i\mathbf{1}_T)'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i - \alpha_i\mathbf{1}_T)$, $\mathbf{d}_{\boldsymbol{\beta}} = \mathbf{C}_{\boldsymbol{\beta}}\left\{\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \boldsymbol{\delta} z_i - \alpha_i\mathbf{1}_T) + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\boldsymbol{\beta}_0\right\}$, $\mathbf{d}_{\boldsymbol{\delta}} = \mathbf{C}_{\boldsymbol{\delta}}\left(\mathbf{V}_{\boldsymbol{\delta}}^{-1}\boldsymbol{\delta}_0 + \sum_{i=1}^{n} z_i\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \alpha_i\mathbf{1}_T)\right)$, and $\mathbf{C}_{\mathbf{V}_\varepsilon} = \boldsymbol{\Omega}_\varepsilon + \sum_{i=1}^{n}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i - \alpha_i\mathbf{1}_T)(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \boldsymbol{\delta} z_i - \alpha_i\mathbf{1}_T)'$.

- **McMC PMSN3**. Using the hierarchical form for PMSN3 and letting the underlying priors being similar to those given in model PMSN2 except that for $\boldsymbol{\delta}$ the prior $\delta \sim N(\delta_0, \sigma_\delta^2)$ is replaced. The complete conditional posteriors are derived as

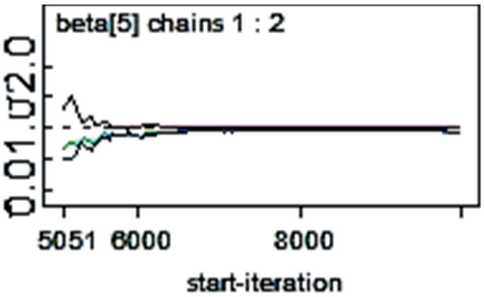
$$\begin{aligned}
&i)\ \boldsymbol{\beta}|\mathbf{y}, others \sim N_k\left(\mathbf{C}_{\boldsymbol{\beta}}\left\{\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \alpha_i\mathbf{1}_T) + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\boldsymbol{\beta}_0\right\}, \mathbf{C}_{\boldsymbol{\beta}}\right),\\
&ii)\ \mathbf{V}_\varepsilon|\mathbf{y}, others \sim IW_T(\mathbf{C}_{\mathbf{V}_\varepsilon}, n + m_\varepsilon)\\
&iii)\ \sigma_\alpha^2|\mathbf{y}, others \sim IG\left(\frac{n}{2} + \gamma, \frac{1}{2}\sum_{i=1}^{n}(\alpha_i - \delta z_i)^2 + \gamma\right).\\
&iv)\ \delta|\mathbf{y}, others \sim N\left(\frac{\sigma_\delta^2\delta_0 + \sigma_\alpha^{-2}\sum_{i=1}^{n} z_i\alpha_i}{\sigma_\delta^{-2} + \sigma_\alpha^{-2}\sum_{i=1}^{n} z_i^2}, \frac{1}{\sigma_\delta^{-2} + \sigma_\alpha^{-2}\sum_{i=1}^{n} z_i^2}\right),\\
&v)\ z_i|\mathbf{y}, others \sim TN\left(\frac{\delta\alpha_i}{\sigma_\alpha^2 + \delta^2}, \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \delta^2}\right) I(0, \infty),\\
&vi)\ \alpha_i|\mathbf{y}, others \sim N\left(\frac{c_{\alpha_i}}{\mathbf{1}_T'\mathbf{V}_\varepsilon^{-1}\mathbf{1}_T + \sigma_\alpha^{-2}}, \frac{1}{\mathbf{1}_T'\mathbf{V}_\varepsilon^{-1}\mathbf{1}_T + \sigma_\alpha^{-2}}\right),\\
&vii)\ \pi(\boldsymbol{\theta}_i|\mathbf{y}, others) \propto \left\{\prod_{t=1}^{T} e^{-e^{\theta_{it}} + \theta_{it} y_{it}}\right\} e^{-\frac{1}{2} c_{\boldsymbol{\theta}_i}},
\end{aligned}$$

where $\mathbf{C}_{\boldsymbol{\beta}} = \left(\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{V}_\varepsilon^{-1}\mathbf{X}_i + \mathbf{V}_{\boldsymbol{\beta}}^{-1}\right)^{-1}$, $c_{\boldsymbol{\theta}_i} = (\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \alpha_i\mathbf{1}_T)'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \alpha_i\mathbf{1}_T)$, $c_{\alpha_i} = \mathbf{1}_T'\mathbf{V}_\varepsilon^{-1}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta}) + \sigma_\alpha^{-2}\delta z_i$ and $\mathbf{C}_{\mathbf{V}_\varepsilon} = \boldsymbol{\Omega}_\varepsilon + \sum_{i=1}^{n}(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \alpha_i\mathbf{1}_T)(\boldsymbol{\theta}_i - \mathbf{X}_i\boldsymbol{\beta} - \alpha_i\mathbf{1}_T)'$.

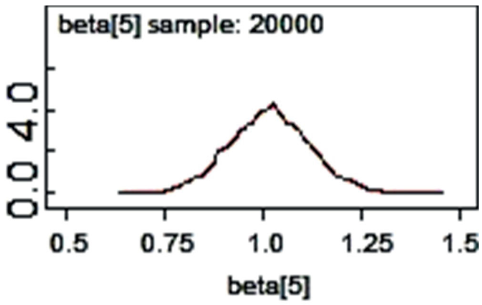
Posterior plots in the health reform data analysis: Figure 1 shows the history, Gelman-Rubin diagnostic, kernel density, and autocorrelation plots for the PMSN2 and β_5. By using two different sequences of starting points, the Gelman-Rubin diagnostic plot shows the same behaviour. Also, we observe a low correlation between successive samples in the autocorrelation plot. Moreover, the history plot moves up and down around the mode of the distribution. Thus the samples will reach a stationary distribution.



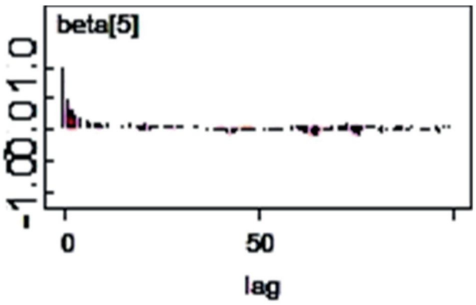
(a) History plot



(b) Gelman-Rubin diagnostic



(c) Kernel density plot



(d) Autoregressive plot

Figure 1: *The posterior plots of β_5 in the simulation study.*