1 The Model

1.1 Data Model

$$\left[\boldsymbol{y}_o \middle| \boldsymbol{\beta}, \sigma^2\right] = \frac{1}{\tau} \sum_{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}} \left[\boldsymbol{y}_c \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}\right] \left[\gamma \middle| \boldsymbol{y}_c, \sigma^2\right]$$

where τ is the number of models under consideration.

1.2 Parameter Model

$$\begin{aligned}
[\beta_0] &\propto 1 \\
\left[\beta_j \middle| \sigma^2, \lambda_j, \gamma_j\right] &\overset{iid}{\sim} \begin{cases} 0 & \text{if } \gamma_j = 0 \\ \mathrm{N}\left(0, \frac{\sigma^2}{\lambda_j}\right) & \text{if } \gamma_j = 1 \end{cases} && \text{for } j = 1, \ldots, p \\
\left[\sigma^2\right] &\propto \frac{1}{\sigma^2} \\
[\gamma_j] &\propto \mathrm{Bern}\left(\pi_j\right) && \text{for } j = 1, \ldots, p
\end{aligned}$$

where π_j and λ_j are fixed hyperpriors for $j = 1, \ldots, p$.

1.3 Posterior

For a given model indexed by $\boldsymbol{\gamma}$, the posterior distribution is

$$\begin{aligned}
\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \middle| \boldsymbol{y}_o, \boldsymbol{X}_o, \boldsymbol{X}_a\right] &= \int \left[\boldsymbol{y}_a, \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \middle| \boldsymbol{y}_o, \boldsymbol{X}_o, \boldsymbol{X}_a\right] \, d\boldsymbol{y}_a \\
&= \frac{\int \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_a\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma] \, d\boldsymbol{y}_a}{\sum_{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}} \int\int\int \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_a\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma] \, d\boldsymbol{y}_a \, d\boldsymbol{\beta}_{\boldsymbol{\gamma}} \, d\sigma^2} \\
&= \frac{\left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma] \left\{\int \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_a\right] \, d\boldsymbol{y}_a\right\}}{\sum_{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}} \int\int \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma] \left\{\int \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_a\right] \, d\boldsymbol{y}_a\right\} d\boldsymbol{\beta}_{\boldsymbol{\gamma}} \, d\sigma^2} \\
&= \frac{\left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma]}{\sum_{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}} \int\int \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma] \, d\boldsymbol{y}_a \, d\boldsymbol{\beta}_{\boldsymbol{\gamma}} \, d\sigma^2} \\
&= \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma \middle| \boldsymbol{y}_o, \boldsymbol{X}_o\right]
\end{aligned}$$

which is independent of the augmented data $(\boldsymbol{y}_a, \boldsymbol{X}_a)$. For the Gibbs sampler we use the posterior definition of

$$\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma} \middle| \boldsymbol{y}_o, \boldsymbol{X}_o, \boldsymbol{X}_a\right] \propto \int \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_o\right] \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \gamma, \boldsymbol{X}_a\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \gamma\right] [\gamma] \, d\boldsymbol{y}_a$$

2 Posteriors

2.1 Posterior for $\boldsymbol{\beta_\gamma}$

$$\begin{aligned}
\left[\boldsymbol{\beta_\gamma}\middle|\sigma^2,\gamma,\boldsymbol{y}_o\right] &\propto \left[\boldsymbol{y}_o\middle|\boldsymbol{\beta_\gamma},\sigma^2,\gamma\right]\left[\boldsymbol{\beta_\gamma}\middle|\sigma^2,\gamma\right] \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\boldsymbol{\beta_\gamma}\right)^T\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\boldsymbol{\beta_\gamma}\right)\right\}\exp\left\{-\frac{1}{2\sigma^2}\boldsymbol{\beta}^T_{\boldsymbol{\gamma}}\boldsymbol{\Delta_\gamma}\boldsymbol{\beta_\gamma}\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{\beta}^T_{\boldsymbol{\gamma}}\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)\boldsymbol{\beta_\gamma}-2\boldsymbol{\beta}^T_{\boldsymbol{\gamma}}\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{y}_o\right]\right\}
\end{aligned}$$

which is $\mathrm{MVN}\left(\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}},\boldsymbol{V}_{\boldsymbol{\beta}}\right)$ where $\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}=\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)^{-1}\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{y}_o$ and $\boldsymbol{V}_{\boldsymbol{\beta}}=\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)^{-1}$.

2.2 Posterior for σ^2

$$\left[\sigma^2|\gamma,\boldsymbol{y}_o\right]=\frac{\left[\sigma^2,\boldsymbol{\beta_\gamma}\middle|\gamma,\boldsymbol{y}_o\right]}{\left[\boldsymbol{\beta_\gamma}\middle|\sigma^2,\gamma,\boldsymbol{y}_o\right]}$$

First consider the numerator of the above equation

$$\begin{aligned}
\left[\sigma^2,\boldsymbol{\beta_\gamma}\middle|\gamma,\boldsymbol{y}_o\right] &\propto \left[\boldsymbol{y}_o\middle|\boldsymbol{\beta_\gamma},\sigma^2,\gamma\right]\left[\boldsymbol{\beta_\gamma}\middle|\sigma^2,\gamma\right]\left[\sigma^2\right] \\
&\propto \left(\sigma^2\right)^{-\frac{n_o}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\boldsymbol{\beta_\gamma}\right)^T\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\boldsymbol{\beta_\gamma}\right)\right\}\left(\sigma^2\left|\boldsymbol{\Delta}^+_{\boldsymbol{\gamma}}\right|\right)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}\boldsymbol{\beta}^T_{\boldsymbol{\gamma}}\boldsymbol{\Delta_\gamma}\boldsymbol{\beta_\gamma}\right\}\frac{1}{\sigma^2} \\
&\propto \left(\sigma^2\right)^{-\frac{n_o-1}{2}-1}\exp\left\{-\frac{1}{\sigma^2}\frac{\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\boldsymbol{\beta_\gamma}\right)^T\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\boldsymbol{\beta_\gamma}\right)+\boldsymbol{\beta}^T_{\boldsymbol{\gamma}}\boldsymbol{\Delta_\gamma}\boldsymbol{\beta_\gamma}}{2}\right\}
\end{aligned}$$

Now we average over $\boldsymbol{\beta_\gamma}$ by replacing $\boldsymbol{\beta_\gamma}$ with its posterior mean $\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}=\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)^{-1}\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{y}_o$ to get the posterior distribution

$$\left[\sigma^2|\gamma,\boldsymbol{y}_o\right]\propto\left(\sigma^2\right)^{-\frac{n_o-1}{2}-1}\exp\left\{-\frac{1}{\sigma^2}\frac{\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}\right)^T\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}\right)+\tilde{\boldsymbol{\beta}}^T_{\boldsymbol{\gamma}}\boldsymbol{\Delta_\gamma}\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}}{2}\right\}$$

which is $\mathrm{IG}\left(\frac{n_o-1}{2},\frac{\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}\right)^T\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}\right)+\tilde{\boldsymbol{\beta}}^T_{\boldsymbol{\gamma}}\boldsymbol{\Delta_\gamma}\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}}{2}\right)$. Now consider the quadratic term

$$\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}\right)^T\left(\boldsymbol{y}_o-\boldsymbol{X}_{o}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}\right)=\boldsymbol{y}^T_o\boldsymbol{y}_o-2\tilde{\boldsymbol{\beta}}^T_{\boldsymbol{\gamma}}\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{y}_o+\tilde{\boldsymbol{\beta}}^T_{\boldsymbol{\gamma}}\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}}$$

Note: Somehow $\tilde{\boldsymbol{\beta}}^T_{\boldsymbol{\gamma}}\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{y}_o=\boldsymbol{y}^T_o\boldsymbol{X}_{o}\gamma\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)^{-1}\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{y}_o=0$???

2.3 Posterior for $\boldsymbol{y}_a$

This posterior is calculated in the same fashion as a posterior predictive distribution of new observations $\tilde{\boldsymbol{y}}$ given new covariates $\tilde{\boldsymbol{X}}$ for simple linear regression as shown $\left[\tilde{\boldsymbol{y}}\middle|\sigma^2,\boldsymbol{y},\tilde{\boldsymbol{X}}\right]\sim\mathrm{N}\left(\tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}},\sigma^2\left(\boldsymbol{I}+\tilde{\boldsymbol{X}}\boldsymbol{V}_{\boldsymbol{\beta}}\tilde{\boldsymbol{X}}^T\right)\right)$ which is

$$\left[\boldsymbol{y}_A|\sigma^2,\gamma,\boldsymbol{y}_o\right]\sim\mathrm{N}\left(\boldsymbol{X}_{a}\gamma\tilde{\boldsymbol{\beta}}_{\boldsymbol{\gamma}},\sigma^2\left(\boldsymbol{I}_{n_a}+\boldsymbol{X}_{a}\gamma\left(\boldsymbol{X}^T_{o\boldsymbol{\gamma}}\boldsymbol{X}_{o}\gamma+\boldsymbol{\Delta_\gamma}\right)^{-1}\boldsymbol{X}^T_{a\boldsymbol{\gamma}}\right)\right)$$

3 Full Conditionals

3.1 Full Conditional for σ^2

$$
\begin{aligned}
[\sigma^2|\cdot] &\propto \int \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{X}_o\right] \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{X}_a\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \boldsymbol{\gamma}\right] \, d\boldsymbol{y}_a \\
&\propto \int \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{X}_a\right] \, d\boldsymbol{y}_a \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \boldsymbol{\gamma}\right] \\
&\propto \left[\boldsymbol{y}_o, \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \boldsymbol{\gamma}\right] \\
&\propto \left(\sigma^2\right)^{-\frac{n_o}{2}} \exp\left\{-\frac{1}{\sigma^2} \frac{\left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)^T \left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)}{2}\right\} \left(\sigma^2\right)^{-\frac{p_{\boldsymbol{\gamma}}}{2}} \exp\left\{-\frac{1}{\sigma^2} \frac{\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T \boldsymbol{\Lambda}_{\boldsymbol{\gamma}} \boldsymbol{\beta}_{\boldsymbol{\gamma}}}{2}\right\} \left(\sigma^2\right)^{-1} \\
&\propto \left(\sigma^2\right)^{-\frac{n_o + p_{\boldsymbol{\gamma}}}{2} - 1} \exp\left\{-\frac{1}{\sigma^2} \frac{\left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)^T \left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T \boldsymbol{\Lambda}_{\boldsymbol{\gamma}} \boldsymbol{\beta}_{\boldsymbol{\gamma}}}{2}\right\}
\end{aligned}
$$

which is IG $\left(\frac{n_o + p_{\boldsymbol{\gamma}}}{2}, \frac{\left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)^T \left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right) + \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T \boldsymbol{\Lambda}_{\boldsymbol{\gamma}} \boldsymbol{\beta}_{\boldsymbol{\gamma}}}{2}\right)$

3.2 Full Conditional for $\boldsymbol{y}_a$

3.3 Full Conditional for $\boldsymbol{\beta}_{\boldsymbol{\gamma}}$

$$
\begin{aligned}
\left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} \middle| \cdot\right] &\propto \int \left[\boldsymbol{y}_o \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}\right] \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} \middle| \sigma^2, \boldsymbol{\gamma}\right] \, d\boldsymbol{y}_a \\
&\propto \left[\boldsymbol{y}_o \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}} \middle| \sigma^2, \boldsymbol{\gamma}\right] \int \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\gamma}\right] \, d\boldsymbol{y}_a \\
&\propto \exp\left\{-\frac{1}{2\sigma^2} \left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)^T \left(\boldsymbol{y}_o - \boldsymbol{X}_{o\boldsymbol{\gamma}}\boldsymbol{\beta}_{\boldsymbol{\gamma}}\right)\right\} \exp\left\{-\frac{1}{2\sigma^2} \boldsymbol{\beta}_{\boldsymbol{\gamma}}^T \boldsymbol{\Delta}_{\boldsymbol{\gamma}} \boldsymbol{\beta}_{\boldsymbol{\gamma}}\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma^2} \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T \left(\boldsymbol{X}_{o\boldsymbol{\gamma}}^T \boldsymbol{X}_{o\boldsymbol{\gamma}} + \boldsymbol{\Delta}_{\boldsymbol{\gamma}}\right) \boldsymbol{\beta}_{\boldsymbol{\gamma}} - 2\boldsymbol{\beta}_{\boldsymbol{\gamma}}^T \left(\boldsymbol{X}_{o\boldsymbol{\gamma}}^T \boldsymbol{y}_o\right)\right]\right\}
\end{aligned}
$$

which is MVN $\left(\boldsymbol{A}^{-1}\boldsymbol{b}, \boldsymbol{A}^{-1}\right)$ where $\boldsymbol{A}^{-1} = \left(\boldsymbol{X}_{o\boldsymbol{\gamma}}^T \boldsymbol{X}_{o\boldsymbol{\gamma}} + \boldsymbol{\Delta}_{\boldsymbol{\gamma}}\right)^{-1}$ and $\boldsymbol{b} = \boldsymbol{X}_{o\boldsymbol{\gamma}}^T \boldsymbol{y}_o$

3.4 Full Conditional for γ_j

For $j = 1, \dots, p$ and using the fact that $\beta_j = \left(\boldsymbol{X}_{cj}^T \boldsymbol{X}_{cj}\right)^{-1} \boldsymbol{X}_{cj}^T \boldsymbol{y}_c$ and $\boldsymbol{X}_{cj}^T \boldsymbol{X}_{cj} = \delta_j$,

$$
\begin{aligned}
[\gamma_j|\cdot] &\propto \left[\boldsymbol{y}_c, |\beta_{\gamma_j}, \sigma^2, \gamma_j, \boldsymbol{X}_o, \boldsymbol{X}_a, \boldsymbol{y}_a\right] \left[\beta_{\gamma_j}, \sigma^2 | \gamma_j\right] [\gamma_j] \\
&\propto \exp\left\{-\frac{1}{2\sigma^2} \left(\boldsymbol{y}_c - \boldsymbol{X}_{cj}\gamma_j\beta_j\right)^T \left(\boldsymbol{y}_c - \boldsymbol{X}_{cj}\gamma_j\beta_j\right)\right\} \left(\frac{\lambda_j}{\sigma^2}\right)^{\frac{\gamma_j}{2}} \exp\left\{-\frac{\gamma_j \lambda_j \beta_j^2}{2\sigma^2}\right\} \pi^{\gamma_j} \left(1 - \pi\right)^{1-\gamma_j} \\
&\propto \left[\left(\frac{\lambda_j}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} \left[\beta_j^2 \left(\boldsymbol{X}_{cj}^T \boldsymbol{X}_{cj} + \lambda_j\right) - 2\beta_j \boldsymbol{X}_{cj}^T \boldsymbol{y}_c\right]\right\} \frac{\pi}{1 - \pi}\right]^{\gamma_j} \\
&\propto \left[\left(\frac{\lambda_j}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} \left[\beta_j^2 \left(\delta_j + \lambda_j\right) - 2\delta_j \beta_j^2\right]\right\} \frac{\pi}{1 - \pi}\right]^{\gamma_j} \\
&\propto \Psi^{\gamma_j}
\end{aligned}
$$

which is Bern$\left(\frac{\Psi}{1+\Psi}\right)$ where $\Psi = \left(\frac{\lambda_j}{\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\beta_j^2\left(\delta_j + \lambda_j\right) - 2\delta_j\beta_j^2\right]\right\}\frac{\pi}{1-\pi}$

4 Data Augmentation

To perform the model selection and averaging, the "complete" design matrix

$$\boldsymbol{X}_c = \begin{bmatrix} \boldsymbol{X}_o \\ \boldsymbol{X}_a \end{bmatrix}$$

which has orthogonal columns, hence $\boldsymbol{X}_c^T\boldsymbol{X}_c = \boldsymbol{I}$. The matrix $\boldsymbol{X}_a$ is chosen to be the Cholesky decomposition of $\boldsymbol{D} - \boldsymbol{X}_o^T\boldsymbol{X}_o$ where $\boldsymbol{D}$ is a diagonal matrix with $\delta + \varepsilon$ on the diagonal where δ is the largest eigenvalue of $\boldsymbol{X}_o^T\boldsymbol{X}_o$ and $\varepsilon = 0.001$ is added to avoid computationally unstable solutions.