# 1 The Model

## 1.1 Data Model

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

where $\tau$ 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 $\pi_j$ and $\lambda_j$ 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, \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] [\boldsymbol{\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, \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] [\boldsymbol{\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, \boldsymbol{\gamma}, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \boldsymbol{\gamma}\right] [\boldsymbol{\gamma}] \left\{\int \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\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, \boldsymbol{\gamma}, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \boldsymbol{\gamma}\right] [\boldsymbol{\gamma}] \left\{\int \left[\boldsymbol{y}_a \middle| \boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\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, \boldsymbol{\gamma}, \boldsymbol{X}_o\right] \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2 \middle| \boldsymbol{\gamma}\right] [\boldsymbol{\gamma}]}{\sum_{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}} \int \int \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] [\boldsymbol{\gamma}] \, d\boldsymbol{y}_a \, d\boldsymbol{\beta}_{\boldsymbol{\gamma}} \, d\sigma^2} \\
&= \left[\boldsymbol{\beta}_{\boldsymbol{\gamma}}, \sigma^2, \boldsymbol{\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

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

# 2 Full Conditionals

## 2.1 Full Conditional for $\sigma^2$

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

## 2.2 Full Conditional for $\boldsymbol{y}_a$