# Web-based Supplemental Materials for "Bayesian models for multiple outcomes nested in domains" by Sally W. Thurston, David Ruppert, and Philip W. Davidson

## A Derivation of notation for model (3)

We first derive the notation for model (3) using a variant of model (2) that is close to the model we use for the Seychelles data. We will then explain how the specific model we use for our application can also be expressed as in (2). The model we consider is

$$\begin{aligned} y^*_{i,\,j} &= (\beta_x + b_{\mathcal{D},\,d(j),\,x} + b_{\mathcal{O},\,j,\,x})\, x_i + \boldsymbol{\beta}_S^T\,\mathbf{S}_i + r_i + r_{\mathcal{D},\,d(j),\,i} + \epsilon_{i\,j} \qquad \text{(A-1)} \\ &= \begin{pmatrix} x_i & \mathbf{S}_i^T \end{pmatrix} \begin{pmatrix} \beta_x & \boldsymbol{\beta}_S^T \end{pmatrix}^T + x_i\; b_{\mathcal{D},d(j),x} + x_i\; b_{\mathcal{O},j,x} + r_i + r_{\mathcal{D},d(j),i} + \epsilon_{i\,j} \end{aligned}$$

Here we include separate terms for the fixed effects, domain-specific exposure effects, outcome-specific exposure effects, overall subject effects, and domain-specific subject effects, a pattern which we follow in extending this model to all outcomes and all subjects together. In order to be able to write model (A-1) for all outcomes and all subjects together, we need some additional notation. Let $\mathbf{1}_J$ be the $J$ by 1 vector of 1's, and $I_{J\times J}$ be the $J \times J$ identity matrix. Also let $I_{dom}$ be the $J \times D$ indicator matrix in which the $(j, d)$th element is 1 if the $j$th outcome is in the $d$th domain, and 0 otherwise.

We use $\mathbf{F}_i$ to denote exposure and covariates that correspond to the fixed effects for subject $i$. Let

$$\begin{aligned} \mathbf{y}^*_i &= \begin{pmatrix} y^*_{i,1} & \cdots & y^*_{i,J} \end{pmatrix}^T,\ \mathbf{F}_i = \begin{pmatrix} x_i & \mathbf{S}_i^T \end{pmatrix},\ Z_{\mathcal{D},i,j} = x_i,\ Z_{\mathcal{O},i,j} = x_i, \\ \mathbf{r}_{\mathcal{D},i} &= \begin{pmatrix} r_{\mathcal{D},i,1} & \cdots & r_{\mathcal{D},i,D} \end{pmatrix}^T,\ \boldsymbol{\varepsilon}_i = \begin{pmatrix} \epsilon_{i,1} & \cdots & \epsilon_{i,J} \end{pmatrix}^T,\ \mathbf{F}_{(\otimes),i} = \begin{pmatrix} x_i & \mathbf{S}_i^T \end{pmatrix} \otimes \mathbf{1}_J, \\ \boldsymbol{\beta} &= \begin{pmatrix} \beta_x & \boldsymbol{\beta}_S^T \end{pmatrix}^T,\ \mathbf{b}_{\mathcal{D}} = \begin{pmatrix} b_{\mathcal{D},1,x} & \cdots & b_{\mathcal{D},D,x} \end{pmatrix}^T,\ \mathbf{b}_{\mathcal{O}} = \begin{pmatrix} b_{\mathcal{O},1,x} & \cdots & b_{\mathcal{O},J,x} \end{pmatrix}^T. \end{aligned}$$

We need some further definitions in order to write model (2) for all observations together. Let

$$\mathbf{Y}^* = \begin{pmatrix} \mathbf{y}_1^* \\ \vdots \\ \mathbf{y}_n^* \end{pmatrix}, \ \mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ \mathbf{F} = \begin{pmatrix} x_1 & \mathbf{S}_1^T \\ \vdots & \vdots \\ x_n & \mathbf{S}_n^T \end{pmatrix}, \ \mathbf{r} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}, \ \mathbf{r}_{\mathcal{D}} = \begin{pmatrix} \mathbf{r}_{\mathcal{D},1} \\ \vdots \\ \mathbf{r}_{\mathcal{D},n} \end{pmatrix},$$

and $\boldsymbol{\varepsilon} = (\ \boldsymbol{\varepsilon}_1^T \ \ \cdots \ \ \boldsymbol{\varepsilon}_n^T\ )^T$. Also let

$$\mathbf{F}_{(\otimes)} = \mathbf{F} \otimes \mathbf{1}_J,\ \mathbf{Z}_{\mathcal{D}} = \mathbf{X} \otimes I_{dom},\ \mathbf{Z}_{\mathcal{O}} = \mathbf{X} \otimes I_{J\times J},\ \mathbf{R} = \mathbf{r} \otimes \mathbf{1}_J,\ \mathbf{R}_{\mathcal{D}} = (I_{dom} \otimes I_{n\times n})\ \mathbf{r}_{\mathcal{D}}. \tag{A-2}$$

Note that $\mathbf{Y}^*, \mathbf{F}_{(\otimes)}, \mathbf{Z}_{\mathcal{D}}, \mathbf{Z}_{\mathcal{O}}, \mathbf{R}, \mathbf{R}_{\mathcal{D}}$ and $\boldsymbol{\varepsilon}$ all have $n \times J$ rows. With this added notation, model (A-1) can be expresses as shown in model (3).

In our model for the Seychelles data we included fixed domain-specific covariate effects for all covariates (but random effects for exposure, as before). In this case $\mathbf{F}_{(\otimes),i}$ in (A-2) was replaced by $\mathbf{F}_{(\otimes),i} = (\ x_i \otimes \mathbf{1}_J \ \ \mathbf{S}_i \otimes I_{dom}\ )$, and $\boldsymbol{\beta}$ was expanded accordingly. However we also allowed the effect of sex on trailmaking A and B to differ from its effect on outcomes within the motor domain. We did this by introducing an indicator matrix, $I_{sex}$ which is similar to $I_{dom}$, but with an additional column that allows the sex effect for trailmaking A and B to differ from the other outcomes in the motor domain. We then replaced the domain-specific sex effect, $x_{sex,i} \otimes I_{dom}$, within $\mathbf{F}_{(\otimes),i}$ by $x_{sex,i} \otimes I_{sex}$.

It will be useful to re-express the model assumptions for the random effects, given in section 2.3. We can write these as $\mathbf{b}_{\mathcal{D}} \sim N(0, \mathbf{D}_{b,\mathcal{D}})$, $\mathbf{b}_{\mathcal{O}} \sim N(0, \mathbf{D}_{b,\mathcal{O}})$, $\mathbf{r} \sim N(0, \mathbf{D}_r)$, $\mathbf{r}_{\mathcal{D}} \sim N(0, \mathbf{D}_{r,\mathcal{D}})$, $\boldsymbol{\varepsilon} \sim N(0, \mathbf{D}_\epsilon)$, where $\mathbf{D}_{b,\mathcal{D}} = \boldsymbol{\Sigma}_{b,\mathcal{D}} \otimes I_{D\times D}$, $\mathbf{D}_{b,\mathcal{O}} = \boldsymbol{\Sigma}_{b,\mathcal{O}} \otimes I_{J\times J}$, $\mathbf{D}_r = \sigma_r^2 I_{n\times n}$, $\mathbf{D}_{r,\mathcal{D}} = I_{n\times n} \otimes \boldsymbol{\Sigma}_{r,\mathcal{D}}$, and $\mathbf{D}_\epsilon = I_{n\times n} \otimes \boldsymbol{\Sigma}_\epsilon$.

# B Conditional posteriors for the multiple domains case

In this appendix we list the conditional posterior of all model parameters. Some derivations of these conditional posteriors are given in the next Web Appendix.

Here we use "rest" to denote all other model parameters and the data. The conditional posterior for $\boldsymbol{\beta}$ is normal with posterior mean and covariance matrix

$$\begin{aligned}
\mathrm{E}(\boldsymbol{\beta} \mid \mathbf{Y}, \text{rest}) &= (\mathbf{F}_{(\otimes)}^T \mathbf{D}_\epsilon^{-1} \mathbf{F}_{(\otimes)} + \boldsymbol{\Sigma}_0^{-1})^{-1} \mathbf{F}_{(\otimes)}^T \mathbf{D}_\epsilon^{-1} (\mathbf{Y}^* - \mathbf{Z}_\mathcal{D} \mathbf{b}_\mathcal{D} - \mathbf{Z}_\mathcal{O} \mathbf{b}_\mathcal{O} - \mathbf{R}_\mathcal{D} - \mathbf{R}) \\
\mathrm{Var}(\boldsymbol{\beta} \mid \mathbf{Y}, \text{rest}) &= (\mathbf{F}_{(\otimes)}^T \mathbf{D}_\epsilon^{-1} \mathbf{F}_{(\otimes)} + \boldsymbol{\Sigma}_0^{-1})^{-1} \\
&= \left\{(\mathbf{F}^T\mathbf{F}) \otimes (\mathbf{1}_J^T \boldsymbol{\Sigma}_\epsilon^{-1} \mathbf{1}_J) + \boldsymbol{\Sigma}_0^{-1}\right\}^{-1} = \{\sum_{j=1}^{J} (1/\sigma_{\epsilon,j}^2) \times \mathbf{F}^T\mathbf{F} + \boldsymbol{\Sigma}_0^{-1}\}^{-1}
\end{aligned}$$

The conditional posterior for the overall subject-specific random effects $\mathbf{r}$, is derived in the next Web Appendix and is

$$\mathbf{r} \mid \mathbf{Y}, \text{rest} \sim N\left(v_r (I_{n\times n} \otimes \mathbf{1}_J^T \boldsymbol{\Sigma}_\epsilon^{-1})(\mathbf{Y}^* - \mathbf{F}_{(\otimes)}\boldsymbol{\beta} - \mathbf{Z}_\mathcal{D} \mathbf{b}_\mathcal{D} - \mathbf{Z}_\mathcal{O} \mathbf{b}_\mathcal{O} - \mathbf{R}_\mathcal{D}),\ v_r I_{n\times n}\right), \quad \text{(A-3)}$$

$$\text{where} \ \ v_r = (\mathbf{1}_J^T \boldsymbol{\Sigma}_\epsilon^{-1} \mathbf{1}_J + 1/\sigma_r^2)^{-1} = \{\sum_j (1/\sigma_{\epsilon,j}^2) + 1/\sigma_r^2\}^{-1} \quad \text{(A-4)}$$

This can be re-expressed as

$$r_i \mid \mathbf{Y}, \text{rest} \sim \ N\left\{\frac{\sum_{j=1}^J (y_{i,j}^* - \mathbf{F}_i\boldsymbol{\beta} - Z_{\mathcal{D},i,j} b_{\mathcal{D},\,d(j),\,x} - Z_{\mathcal{O},i,j} b_{\mathcal{O},\,j,\,x} - r_{\mathcal{D},i})/\sigma_{\epsilon,j}^2}{\sum_{j=1}^J (1/\sigma_{\epsilon,j}^2) + 1/\sigma_r^2},\ v_r\right\}, \quad \text{(A-5)}$$

The conditional posterior for the $d$th subject-specific random effect at the domain level, $\mathbf{r}_{\mathcal{D},d,i}$ is

$$\mathbf{r}_{\mathcal{D},d,i} \mid \mathbf{Y}, \text{rest} \sim \ N\left\{\frac{\sum_{j\in d(j)}^D (y_{i,j}^* - \mathbf{F}_i\boldsymbol{\beta} - Z_{\mathcal{D},i,j} b_{\mathcal{D},\,d(j),\,x} - Z_{\mathcal{O},i,j} b_{\mathcal{O},\,j,\,x} - r_i)/\sigma_{\epsilon,j}^2}{\sum_{j=1}^J (1/\sigma_{\epsilon,j}^2) + 1/\sigma_{r,\mathcal{D},d}^2},\ v_{\mathcal{D},d}\right\},$$

where $v_{\mathcal{D},d} = (\mathbf{1}_J^T \boldsymbol{\Sigma}_\epsilon^{-1} \mathbf{1}_J + 1/\sigma^2_{r,\mathcal{D},d})^{-1} = \{\sum_j (1/\sigma^2_{\epsilon,j}) + 1/\sigma^2_{r,\mathcal{D},d}\}^{-1}$ The conditional posterior for the outcome-specific random effects, $\mathbf{b}_\mathcal{O}$, is normal with posterior mean and variance

$$
\begin{aligned}
\mathrm{E}(\mathbf{b}_\mathcal{O} \mid \mathbf{Y}, \mathrm{rest}) &= (\mathbf{Z}_\mathcal{O}^T \mathbf{D}_\epsilon^{-1} \mathbf{Z}_\mathcal{O} + \mathbf{D}_{b,\mathcal{O}}^{-1})^{-1} \mathbf{Z}_\mathcal{O}^T \mathbf{D}_\epsilon^{-1} (\mathbf{Y}^* - \mathbf{F}_{(\otimes)} \boldsymbol{\beta} - \mathbf{Z}_\mathcal{D} \mathbf{b}_\mathcal{D} - \mathbf{R} - \mathbf{R}_\mathcal{D}) \\
\mathrm{Var}(\mathbf{b}_\mathcal{O} \mid \mathbf{Y}, \mathrm{rest}) &= (\mathbf{Z}_\mathcal{O}^T \mathbf{D}_\epsilon^{-1} \mathbf{Z}_\mathcal{O} + \mathbf{D}_{b,\mathcal{O}}^{-1})^{-1} \\
&= \Big\{ (\mathbf{X} \otimes I_{n\times n})^T (I_{n\times n} \otimes \boldsymbol{\Sigma}_\epsilon^{-1})(\mathbf{X} \otimes I_{n\times n}) + (\boldsymbol{\Sigma}_{b,\mathcal{O}} \otimes I_{J\times J}) \Big\} \qquad \text{(A-6)} \\
&= \Big\{ (\mathbf{X}^T \mathbf{X} \otimes \boldsymbol{\Sigma}_\epsilon^{-1}) + (\boldsymbol{\Sigma}_{b,\mathcal{O}}^{-1} \otimes I_{J\times J}) \Big\} \qquad \text{(A-7)}
\end{aligned}
$$

In (A-6), the replacement of $\mathbf{Z}_\mathcal{O}$ by $\mathbf{X} \otimes I_{J\times J}$, as given in (A-2) assumes that there are no random outcome-specific covariate effects. If the model includes random outcome-specific covariate effects for all covariates, then $\mathbf{X}$ in (A-6) and (A-7) would be replaced by $\mathbf{F}$.

The conditional posterior for the domain-specific random effects, $\mathbf{b}_\mathcal{D}$, is normal with posterior mean and variance

$$
\begin{aligned}
\mathrm{E}(\mathbf{b}_\mathcal{D} \mid \mathbf{Y}, \mathrm{rest}) &= (\mathbf{Z}_\mathcal{D}^T \mathbf{D}_\epsilon^{-1} \mathbf{Z}_\mathcal{D} + \mathbf{D}_{b,\mathcal{D}}^{-1})^{-1} \mathbf{Z}_\mathcal{D}^T \mathbf{D}_\epsilon^{-1} (\mathbf{Y}^* - \mathbf{F}_{(\otimes)} \boldsymbol{\beta} - \mathbf{Z}_\mathcal{O} \mathbf{b}_\mathcal{O} - \mathbf{R} - \mathbf{R}_\mathcal{D}) \\
\mathrm{Var}(\mathbf{b}_\mathcal{D} \mid \mathbf{Y}, \mathrm{rest}) &= (\mathbf{Z}_\mathcal{D}^T \mathbf{D}_\epsilon^{-1} \mathbf{Z}_\mathcal{D} + \mathbf{D}_{b,\mathcal{D}}^{-1})^{-1}
\end{aligned}
$$

where the variance can be re-expressed in a similar manner to the variance of $\mathbf{b}_\mathcal{O}$.

The posteriors for the variance components are given below. We derive the conditional posterior for the elements of $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_{b,\mathcal{D}}$ in the next Web Appendix.

$$
\begin{aligned}
\sigma^2_{d,\mathcal{D},x} \mid \mathbf{Y}, \mathrm{rest} &\sim \mathrm{IG}(D/2 + A_{0,b,\mathcal{D}}, \quad \sum_{d=1}^{D} b^2_{\mathcal{D},d,x}/2 + B_{0,b,\mathcal{D}}) \\
\sigma^2_{d,\mathcal{O},x} \mid \mathbf{Y}, \mathrm{rest} &\sim \mathrm{IG}(J/2 + A_{0,b,\mathcal{O}}, \quad \sum_{j=1}^{J} b^2_{\mathcal{O},j,x}/2 + B_{0,b,\mathcal{O}}) \\
\sigma^2_{\epsilon,j} \mid \mathbf{Y}, \mathrm{rest} &\sim IG\{n/2 + A_{0,\epsilon}, \\
&\qquad \sum_{i=1}^{n} (y^*_{i,j} - \mathbf{F}_i \boldsymbol{\beta} - Z_{\mathcal{D},i,j} b_{\mathcal{D},d(j),x} - Z_{\mathcal{O},i,j} b_{\mathcal{O},j,x} - r_i - r_{\mathcal{D},d(j),i})^2/2 + B_{0,\epsilon}\}
\end{aligned}
$$

$$
\begin{aligned}
\sigma_r^2 \mid \mathbf{Y}, \text{rest} &\sim IG(n/2 + A_{0,r}, \sum_{i=1}^{n} r_i^2/2 + B_{0,r}) \\
\sigma_{r,\mathcal{D},d}^2 \mid \mathbf{Y}, \text{rest} &\sim IG(n/2 + A_{0,r,\mathcal{D}}, \sum_{i=1}^{n} r_{\mathcal{D},d,i}^2/2 + B_{0,r,\mathcal{D}}).
\end{aligned}
$$

# C Derivation of conditional posteriors

The joint posterior of all model parameters is $p(\boldsymbol{\beta}, \mathbf{Z}_{\mathcal{D}}, \mathbf{Z}_{\mathcal{O}}, \mathbf{r}, \mathbf{r}_{\mathcal{D}}, \boldsymbol{\Sigma}_{\epsilon}, \boldsymbol{\Sigma}_{b,\mathcal{D}}, \boldsymbol{\Sigma}_{b,\mathcal{O}}, \sigma_r^2, \boldsymbol{\Sigma}_{r,\mathcal{D}} \mid \mathbf{Y}) \propto p(\mathbf{D}_{\epsilon})p(\mathbf{D}_{b,\mathcal{D}})p(\mathbf{D}_{b,\mathcal{O}})p(\mathbf{D}_r)p(\mathbf{D}_{r,\mathcal{D}})\exp\{T/2\}$ where

$$
\begin{aligned}
T \;\propto\; & (\mathbf{Y}^* - \mathbf{F}_{(\otimes)}\boldsymbol{\beta} - \mathbf{Z}_{\mathcal{D}}\mathbf{b}_{\mathcal{D}} - \mathbf{Z}_{\mathcal{O}}\mathbf{b}_{\mathcal{O}} - \mathbf{R} - \mathbf{R}_{\mathcal{D}})^T \mathbf{D}_{\epsilon}^{-1} (\mathbf{Y}^* - \mathbf{F}_{(\otimes)}\boldsymbol{\beta} - \mathbf{Z}_{\mathcal{D}}\mathbf{b}_{\mathcal{D}} - \mathbf{Z}_{\mathcal{O}}\mathbf{b}_{\mathcal{O}} - \mathbf{R} - \mathbf{R}_{\mathcal{D}}) \\
& + \mathbf{b}_{\mathcal{D}}^T \mathbf{D}_{b,\mathcal{D}}^{-1} \mathbf{b}_{\mathcal{D}} + \mathbf{b}_{\mathcal{O}}^T \mathbf{D}_{b,\mathcal{O}}^{-1} \mathbf{b}_{\mathcal{O}} + \mathbf{r}^T \mathbf{D}_r^{-1} \mathbf{r} + \mathbf{r}_{\mathcal{D}}^T \mathbf{D}_{r,\mathcal{D}}^{-1} \mathbf{r}_{\mathcal{D}} + \boldsymbol{\beta}^T \boldsymbol{\Sigma}_0 \boldsymbol{\beta}
\end{aligned}
$$

### Obtain the conditional posterior for r

Let $\mathbf{A}_r = (\mathbf{Y}^* - \mathbf{F}_{(\otimes)}\boldsymbol{\beta} - \mathbf{Z}_{\mathcal{D}}\mathbf{b}_{\mathcal{D}} - \mathbf{Z}_{\mathcal{O}}\mathbf{b}_{\mathcal{O}} - \mathbf{R}_{\mathcal{D}})$. The conditional posterior for $\mathbf{r}$ is proportional to $\exp(-\mathbf{T}_r/2)$, where $\mathbf{T}_r = (\mathbf{R} - \mathbf{A}_r)^T \mathbf{D}_{\epsilon}^{-1} (\mathbf{R} - \mathbf{A}_r) + \mathbf{r}^T \mathbf{D}_r^{-1} \mathbf{r}$. Re-write $\mathbf{R}$ as

$$
\mathbf{R} = \mathbf{r} \otimes \mathbf{1}_J = \begin{bmatrix} r_1 \mathbf{1}_J \\ \dots \\ r_n \mathbf{1}_J \end{bmatrix} = \begin{bmatrix} \mathbf{1}_J & 0 & 0 & 0 \\ 0 & \mathbf{1}_J & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \mathbf{1}_J \end{bmatrix} \mathbf{r} = (I_{n \times n} \otimes \mathbf{1}_J)\mathbf{r}
$$

Then $\mathbf{T}_r$ becomes

$$
\begin{aligned}
\mathbf{T}_r &= [(I_{n\times n} \otimes \mathbf{1}_J)\mathbf{r} - \mathbf{A}_r]^T \mathbf{D}_{\epsilon}^{-1} [(I_{n\times n} \otimes \mathbf{1}_J)\mathbf{r} - \mathbf{A}_r] + \mathbf{r}^T \mathbf{D}_r^{-1} \mathbf{r} \\
&= [\mathbf{r}^T (I_{n\times n} \otimes \mathbf{1}_J^T) - \mathbf{A}_r^T](I_{n\times n} \otimes \boldsymbol{\Sigma}_{\epsilon}^{-1})[(I_{n\times n} \otimes \mathbf{1}_J)\mathbf{r} - \mathbf{A}_r] + \mathbf{r}^T \mathbf{D}_r^{-1} \mathbf{r} \\
&\propto \mathbf{r}^T [(I_{n\times n} \otimes \mathbf{1}_J^T)(I_{n\times n} \otimes \boldsymbol{\Sigma}_{\epsilon}^{-1})(I_{n\times n} \otimes \mathbf{1}_J) + \mathbf{D}_r^{-1}]\mathbf{r}
\end{aligned}
$$

$$
\begin{aligned}
& -2\mathbf{r}^T(I_{n\times n}\otimes \mathbf{1}_J)(I_{n\times n}\otimes \mathbf{\Sigma}_\epsilon^{-1})\mathbf{A}_r \\
\propto\ & \mathbf{r}^T[(I_{n\times n}\otimes \mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1}\mathbf{1}_J)+\mathbf{D}_r^{-1}]\mathbf{r}-2\mathbf{r}(I_{n\times n}\otimes \mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1})\mathbf{A}_r \\
=\ & \mathbf{r}^T[I_{n\times n}\otimes(\mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1}\mathbf{1}_J+1/\sigma_r^2)]\mathbf{r}-2\mathbf{r}(I_{n\times n}\otimes \mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1})\mathbf{A}_r \\
=\ & \mathbf{r}^T(v_r^{-1}I_{n\times n})\mathbf{r}-2\mathbf{r}(I_{n\times n}\otimes \mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1})\mathbf{A}_r
\end{aligned}
$$

where as given in (A-4), $v_r=(\mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1}\mathbf{1}_J+1/\sigma_r^2)^{-1}=\{\sum_j(1/\sigma_{\epsilon,j}^2)+1/\sigma_r^2\}^{-1}$ so as given in (A-3)

$$
\mathbf{r}\mid\mathbf{Y},\mathbf{b},\boldsymbol{\beta}\sim N\{v_r(I_{n\times n}\otimes \mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1})(\mathbf{Y}^*-\mathbf{F}_{(\otimes)}\boldsymbol{\beta}-\mathbf{Z}_{\mathcal{D}}\mathbf{b}_{\mathcal{D}}-\mathbf{Z}_{\mathcal{O}}\mathbf{b}_{\mathcal{O}}-\mathbf{R}_{\mathcal{D}}),\ v_rI_{n\times n}\}
$$

We can re-write this in scalar format. Let $\mathbf{A}_{r,i}$ be the $j$-dimensional column vector corresponding to the $i$th subject: $A_{r,i}=\mathbf{y}_i^*-\mathbf{F}_{(\otimes),i}\boldsymbol{\beta}-\mathbf{Z}_{\mathcal{D},i}\mathbf{b}_{\mathcal{D}}-\mathbf{Z}_{\mathcal{O},i}\mathbf{b}_{\mathcal{O}}-\mathbf{R}_{\mathcal{D},i}$, where

$$
\mathbf{Z}_{\mathcal{D},i}=Z_{\mathcal{D},i,j}I_{dom},\ \mathbf{Z}_{\mathcal{O},i}=Z_{\mathcal{O},i,j}I_{J\times J},\ \mathbf{R}_i=r_i\mathbf{1}_J,\ \mathbf{R}_{\mathcal{D},i}=I_{dom}\ \mathbf{r}_{\mathcal{D},i}.
$$

Also let $\mathbf{p}_\epsilon^T=\mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1}=\begin{pmatrix}1/\sigma_{\epsilon,1}^2 & \ldots & 1/\sigma_{\epsilon,J}^2\end{pmatrix}$, so $I_{n\times n}\otimes\mathbf{1}_J^T\mathbf{\Sigma}_\epsilon^{-1}=\text{blockdiag}\{\mathbf{p}_\epsilon\}$, where

$$
\text{blockdiag}\{\mathbf{p}_\epsilon\}=\begin{pmatrix}
\mathbf{p}_\epsilon^T & \mathbf{0} & \cdots & \mathbf{0}\\
\mathbf{0} & \mathbf{p}_\epsilon^T & \cdots & \mathbf{0}\\
\vdots & \vdots & \vdots & \vdots\\
\mathbf{0} & \mathbf{0} & \cdots & \mathbf{p}_\epsilon^T
\end{pmatrix}
$$

Then it follows that the posterior mean of $\mathbf{r}$ can be re-expressed as

$$v_r(I_{n\times n}\otimes \mathbf{1}_J^T\boldsymbol{\Sigma}_\epsilon^{-1})\mathbf{A}_r \;=\; v_r \text{ blockdiag}\{\mathbf{p}_\epsilon\}\mathbf{A}_r = \begin{pmatrix} v_r\mathbf{p}_\epsilon^T A_{r,1} \\ v_r\mathbf{p}_\epsilon^T A_{r,2} \\ \vdots \\ v_r\mathbf{p}_\epsilon^T A_{r,n} \end{pmatrix}$$

Substituting in for $\mathbf{p}_\epsilon$ and $\mathbf{A}_r$, we can see that the $i$th element is

$$\begin{aligned} & v_r(\begin{array}{ccc} 1/\sigma^2_{\epsilon,1} & \cdots & 1/\sigma^2_{\epsilon,J} \end{array})(\mathbf{y}_i^* - \mathbf{F}_{(\otimes),i}\boldsymbol{\beta} - \mathbf{Z}_{\mathcal{D},i}\mathbf{b}_{\mathcal{D}} - \mathbf{Z}_{\mathcal{O},i}\mathbf{b}_{\mathcal{O}} - \mathbf{R}_{\mathcal{D},i}) \\ &= \sum_{j=1}^{J}\frac{v_r}{\sigma^2_{\epsilon,j}}(y^*_{i,j} - \mathbf{F}_i\boldsymbol{\beta} - Z_{\mathcal{D},i,j}b_{\mathcal{D},\,d(j),\,x} - Z_{\mathcal{O},i,j}b_{\mathcal{O},\,j,\,x} - r_{\mathcal{D},d(j),i}) \end{aligned}$$

from which the posterior of $\mathbf{r}$ as given in (A-5), follows.

**Obtain the conditional posterior for $\boldsymbol{\Sigma}_\epsilon$**

Using properties of determinants (Harville, 1997), we can rewrite the determinant of $\mathbf{D}_\epsilon$ as

$$|\mathbf{D}_\epsilon| \;=\; |I_{n\times n}\otimes\boldsymbol{\Sigma}_\epsilon| = |I_{n\times n}|^J|\boldsymbol{\Sigma}_\epsilon|^n = \left(\prod_j \sigma^2_{\epsilon,j}\right)^n$$

It follows that $|\mathbf{D}_\epsilon|^{-1/2} = \left(\prod_j \sigma^2_{\epsilon,j}\right)^{-n/2}$. For a given $j$, the only part of $|\mathbf{D}_\epsilon|^{-1/2}$ that involves $\sigma^2_{\epsilon,j}$ is $(\sigma^2_{\epsilon,j})^{-n/2}$. Therefore the contribution to the likelihood from a particular $\sigma^2_{\epsilon,j}$ is just

$$(\sigma^2_{\epsilon,j})^{-n/2}\exp\{-(1/2)\sum_{i=1}^{n}(y^*_{i,j} - \mathbf{F}_i\boldsymbol{\beta} - Z_{\mathcal{D},i,j}b_{\mathcal{D},\,d(j),\,x} - Z_{\mathcal{O},i,j}b_{\mathcal{O},\,j,\,x} - r_i - r_{\mathcal{D},i})^2/\sigma^2_{\epsilon,j}\}$$

Combining this with the prior, $\sigma^2_{\epsilon,j}\sim IG(A_{0,\epsilon},B_{0,\epsilon})$, the posterior follows.

**Obtain the conditional posterior for $\boldsymbol{\Sigma}_{b,\mathcal{D}}$**

We can re-write $|\mathbf{D}_{b,\mathcal{D}}|$ as $|\mathbf{D}_{b,\mathcal{D}}| = |\boldsymbol{\Sigma}_{b,\mathcal{D}}\otimes I_{D\times D}| = |\boldsymbol{\Sigma}_{b,\mathcal{D}}|^D = (\sigma^2_{d,\mathcal{D},x}\prod_{k=1}^{K}\sigma^2_{d,\mathcal{D},S_k})^D$ The contribution to the posterior from the likelihood involves the exponentiation of the term $-(1/2)\mathbf{b}_{\mathcal{D}}^T\mathbf{D}_{b,\mathcal{D}}^{-1}\mathbf{b}_{\mathcal{D}}$. Since $\mathbf{\Sigma}_{b,\mathcal{D}}$ is diagonal,

$$\mathbf{b}_{\mathcal{D}}^T\mathbf{D}_{b,\mathcal{D}}^{-1}\mathbf{b}_{\mathcal{D}} = \frac{1}{\sigma^2_{b,\mathcal{D},x}}\sum_{d=1}^{D} b_{\mathcal{D},d,x} + \frac{1}{\sigma^2_{b,\mathcal{D},S_1}}\sum_{d=1}^{D} b_{\mathcal{D},d,S_1} + \cdots + \frac{1}{\sigma^2_{b,\mathcal{D},S_p}}\sum_{d=1}^{D} b_{\mathcal{D},d,S_p}$$

Since we have also assumed inverse gamma priors for each component of $\mathbf{\Sigma}_{b,\mathcal{D}}$ we can consider each element separately. Consider the posterior for $\sigma^2_{d,\mathcal{D},x}$, and using what we have derived above,

$$p(\sigma^2_{d,\mathcal{D},x} \mid \mathbf{Y}, \text{rest}) \quad \propto \quad (\sigma^2_{d,\mathcal{D},x})^{-(D/2+A_{0,b,\mathcal{D}})}\exp\left(-\frac{\sum_{d=1}^{D} b_{\mathcal{D},x}/2 + B_{0,b,\mathcal{D}}}{\sigma^2_{d,\mathcal{D},x}}\right)$$

from which the conditional posterior follows. The posteriors for the other elements of $\mathbf{\Sigma}_{\mathcal{D}}$ factor in a similar manner.

# D Drawing missing outcome values

Let $\theta$ denote all model parameters, superscript $(t)$ denote the sampled value at iteration $t$, and subscripts of *mis* and *obs* denote missing data and observed data respectively. At the $(t+1)$st iteration of the MCMC we alternate between the imputation step (I-step) in which we draw $y^{*(t+1)}_{mis} \sim P(y^*_{mis} \mid y^*_{obs}, \theta^{(t)})$, and the posterior step (P-step) in which we draw $\theta^{(t+1)} \sim P(\theta \mid y^*_{obs}, y^{*(t+1)}_{mis})$[24]. The P-step is the iterative step for drawing from the posterior distribution for the model parameters given complete data, as already discussed.

Given all model parameters, the $y^*_{i,j}$ are all independent. Thus the I-step involves draws from

$$y^{*(t+1)}_{mis,i,j} \mid y^*_{obs}, \text{rest} \sim N\left(\mathbf{F}_i\boldsymbol{\beta}^{(t)} + Z_{\mathcal{D},i,j}b^{(t)}_{\mathcal{D},d(j),x} + Z_{\mathcal{O},i,j}b^{(t)}_{\mathcal{O},j,x} + r_i^{(t)} + r^{(t)}_{\mathcal{D},d(j),i},\ \sigma^{2(t)}_{\epsilon,j}\right)$$

If we did not recenter draws of $y^{*(t+1)}$ and were to fit our model with intercepts, the intercepts would not be exactly zero. This is because centering took place on the observed data $y_{obs}$ and not on the complete data $(y_{obs}, y^*_{mis})$. To avoid a possible bias in other model parameters for our no-intercept model, we took an additional step to ensure that the intercepts are exactly zero. Specifically, after each draw of $y^*_{mis}$ we rescaled this draw to the original outcome scale using the mean and SD from the observed data on the $j$th outcome, then re-centered and rescaled the entire vector of $(y^{(t)}_{mis}, y_{obs})$.

## E Details of prior choice and model checking

The $IG(1, 1)$ prior for variance components is a more typical choice for a non-informative prior than our priors. Although a shape hyperparameter of 1 is not overly informative in some applications, here only $D = 4$ values contribute to estimation of $\sigma^2_{d,\mathcal{D},x}$, motivating our adoption of a smaller shape parameter. Likewise, a scale hyperparameter of 1 may be uninformative when the relevant posterior sum of squares is large. However in our application the posterior sum of squares for nearly all variance components were substantially smaller than 1, motivating our choice of a much smaller shape parameter. For "prior A" we visually compared the scale hyperparameter for each variance component to the draws of the relevant posterior sums of squares. Our prior hyperparameters were smaller than nearly all draws of the posterior sums of squares, indicating that our hyperparameter choices were not overly influential.

The MCMC for each prior ("A", "B", and "C") was run with 10 different sets of starting values for the variance components. Six chains started from variance components of either all 100, all $10^{-7}$ or a mixture of 100 and $10^{-7}$. Three chains used starting values that were sampled from uniform distributions on $[0, 1]$, and one chain used reasonable values as determined from previous MCMC samples. Starting values for all fixed and random effects were always set to zero, which affects only the initial imputation step for missing outcomes. Starting values for missing data are not needed because our algorithm first draws missing data (which initially depend on starting values for model parameters), and these draws are then used in sampling model parameters.

The MCMC for each of the 10 chains was run for 6000 iterations for prior “A” and “B”, and for $11,000$ iterations for “prior C”. Based on model diagnostics we used a burn-in of 1000 for priors “A” and “B”, and a burn-in of 6000 for “prior C”, thereby keeping $50,000$ draws (5000 from each chain) for every prior. Final posterior estimates were based on $10,000$ draws in which we took every 5th draw.

As discussed in Section 2.4, in addition to a visual examination of the traceplots our diagnostics included the Gelman-Rubin diagnostic $\hat{R}$ for the multiple chains [11], `mcgibbsit`, and the Raftery-Lewis diagnostic [18]. `Mcgibbsit` [28] in `R` is a generalization of the Raftery-Lewis diagnostic [18] for multiple chains. We used `mcgibbsit` to determine the number of draws needed to estimate the $q = 0.025$th quantile of each model parameter within $r = \pm 0.0125$ of the model parameter quantile, with $s = 95\%$ probability. For single chains, we calculated the Raftery-Lewis diagnostic [18] using the same values of $q, r$, and $s$ as we used for `mcgibbsit`.

Using “prior A”, trace plots for $\sigma_r^2$ showed that in 3 chains the draws were essentially zero for several hundred iterations, before these chains found the more likely mode already reached by the remaining chains, at about 0.17. Draws of most elements of $\mathbf{\Sigma}_\epsilon$ from these 3 chains were substantially larger than the other chains for several hundred iterations, before these 3 chains found the other mode with smaller values of $\mathbf{\Sigma}_\epsilon$ already reached by the remaining chains. Trace plots for “prior B” and “prior C” showed the same patterns, but with the switches occurring after several thousand iterations for “prior C”. In all cases the earlier modes were never revisited.

Under “prior A”, the 97.5th percentile for the Gelman-Rubin $\hat{R}$ was 1.00 for nearly all parameters, 1.01 for a few others, 1.02 for one, and was 1.12 for $\sigma^2_{d,\mathcal{D},x}$. `Mcgibbsit` indicated that all model parameters could be estimated within the desired level of precision in fewer than 6000 draws. Using the final sample with every 5th draw from all 10 chains, the Raftery-Lewis diagnostic for "prior A" indicated that 1400 draws were sufficient to reach the desired level of precision for all model parameters. Although 1400 draws would be sufficient, we report results from $10,000$ draws (every 5th draw from all 10 chains) for increased precision.

The model diagnostic results for "prior B" and "prior C" were similar to, but not quite as good, as those from "prior A". The 97.5th percentile of the Gelman-Rubin $\hat{R}$ for $\sigma^2_{d,\mathcal{D},x}$ was 1.39 for "prior B" and 1.08 for "prior C". The Raftery-Lewis diagnostic on the final chain indicated that all parameters could be estimated with the desired level of precision within 3000 draws for prior "B" and $13,000$ draws for prior "C". Under "prior C", only $\sigma^2_{d,\mathcal{O},x}$ could not be estimated with the desired level of precision within 5000 draws.

In order to see whether our model was consistent with the observed data, at every 50 iterations of the final $10,000$ draws we took a sample of outcomes (size: $533 \times 20$) from the posterior predictive distribution, $p(\mathbf{y}^{rep} \mid \mathbf{y})$, conditional on the observed values of MeHg and covariates. We then compared the distribution of the following summary statistics from the posterior predictive distribution to the corresponding value in the SCDS data: marginal means and variances of the outcomes, maximum and minimum correlations between an outcome and all other outcomes, regression coefficients and their corresponding p-values for MeHg and each covariate from separate regressions for each relevant outcome, and ratio of the residual variance from the separate regression to the marginal variance of the outcome. Because of the large number of outcomes, we only examined these statistics from the first outcome in each domain.

In nearly all respects the draws from the posterior predictive distribution were similar to their corresponding observed value in the SCDS data. P-values for some covariates were occasionally very different from the observed value, but these occasional differences were not consistent across outcomes or covariates. However the maximum correlations between outcome pairs from the posterior predictive distributions were consistently smaller, and the minimum correlations between outcome pairs consistently larger than was observed in the SCDS data. This suggests that our treatment of outcomes within a domain as exchangeable is an oversimplification of the true situation, as discussed briefly in Section 3.2.

# F Proof that $\mathbf{Z}^T\mathbf{Z}$ is singular when combining $\mathbf{Z}_{\mathcal{O}}$ and $\mathbf{Z}_{\mathcal{D}}$

In this appendix we prove that when we use a single $\mathbf{Z}$ matrix for both outcome-specific and domain-specific effects, the $\mathbf{Z}^T\mathbf{Z}$ matrix is singular. The implication of this is that the outcome-specific and domain-specific effects must be considered separately. We prove this for the situation in which we allow outcome-specific and deviation-specific covariates effects for the same set of covariates. In our implementation, $\mathbf{F} = \mathbf{X}$, i.e. we did not have random effects for covariates. In the general case

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_{\mathcal{D}} & \mathbf{Z}_{\mathcal{O}} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \otimes I_{dom} & \mathbf{F} \otimes I_{J\times J} \end{pmatrix}$$

where, as defined on page i, $I_{dom}$ is the $J \times D$ matrix in which the $(j, d)$th element is 1 if the $j$th outcome is in the $d$th domain, and 0 otherwise.

First, $\mathbf{Z} = \mathbf{F} \otimes (I_{dom} \quad I_{J\times J})$ by (1.12) in Chapter 16 of Harville (1997). Next, rank$(I_{dom} \quad I_{J\times J})$ $= J$ since $(I_{dom} \quad I_{J\times J})$ has $J$ rows. Then, by (1.26) in Chapter 16 of Harville (1997), rank$(\mathbf{Z})$ $=$ rank$(\mathbf{F})$rank$(I_{dom} \quad I_{J\times J})$, which is less than the number of columns in $\mathbf{Z}$. Therefore, the columns of $\mathbf{Z}$ are linearly dependent, so there is a non-zero vector $\boldsymbol{\alpha}$ such that $\mathbf{Z}\boldsymbol{\alpha} = 0$. Therefore, $\mathbf{Z}^T\mathbf{Z}\boldsymbol{\alpha} = 0$ so that $\mathbf{Z}^T\mathbf{Z}$ is singular.

# G Additional references

Harville DA (1997). Matrix Algebra from a Statistician's Perspective. Springer, New York.

| | | Model results | | | | Separate regressions | | | |
|---|---|---|---|---|---|---|---|---|---|
| | true | mean | .025 | .975 | MSE | mean | .025 | .975 | MSE |
| $\beta_x$ | −0.0500 | −0.0471 | −0.0953 | 0.0045 | 0.0010 | | | | |
| $b_{\mathcal{D},1}$ | −0.0300 | −0.0130 | −0.0624 | 0.0183 | 0.0009 | | | | |
| $b_{\mathcal{D},2}$ | 0.0300 | 0.0134 | −0.0209 | 0.0560 | 0.0008 | | | | |
| $b_{\mathcal{D},3}$ | −0.0150 | −0.0072 | −0.0498 | 0.0252 | 0.0005 | | | | |
| $b_{\mathcal{D},4}$ | 0.0150 | 0.0052 | −0.0333 | 0.0465 | 0.0006 | | | | |
| slope 1 | −0.0800 | −0.0623 | −0.1164 | −0.0104 | 0.0015 | −0.0749 | −0.1499 | 0.0001 | 0.0029 |
| slope 2 | −0.0800 | −0.0634 | −0.1166 | −0.0122 | 0.0014 | −0.0790 | −0.1492 | −0.0088 | 0.0024 |
| slope 3 | −0.0800 | −0.0624 | −0.1164 | −0.0099 | 0.0014 | −0.0776 | −0.1562 | 0.0009 | 0.0028 |
| slope 4 | −0.0200 | −0.0304 | −0.0819 | 0.0231 | 0.0013 | −0.0085 | −0.0905 | 0.0736 | 0.0036 |
| slope 5 | −0.0200 | −0.0320 | −0.0833 | 0.0202 | 0.0013 | −0.0204 | −0.0985 | 0.0578 | 0.0030 |
| slope 6 | −0.0200 | −0.0317 | −0.0829 | 0.0208 | 0.0013 | −0.0185 | −0.0961 | 0.0591 | 0.0031 |
| slope 7 | −0.0200 | −0.0331 | −0.0842 | 0.0188 | 0.0013 | −0.0267 | −0.1038 | 0.0504 | 0.0026 |
| slope 8 | −0.0200 | −0.0327 | −0.0844 | 0.0202 | 0.0013 | −0.0224 | −0.1084 | 0.0637 | 0.0034 |
| slope 9 | −0.0200 | −0.0325 | −0.0843 | 0.0206 | 0.0014 | −0.0201 | −0.1070 | 0.0668 | 0.0035 |
| slope 10 | −0.0650 | −0.0547 | −0.1054 | −0.0043 | 0.0011 | −0.0575 | −0.1384 | 0.0235 | 0.0033 |
| slope 11 | −0.0650 | −0.0543 | −0.1051 | −0.0034 | 0.0011 | −0.0572 | −0.1399 | 0.0256 | 0.0037 |
| slope 12 | −0.0650 | −0.0547 | −0.1058 | −0.0039 | 0.0011 | −0.0597 | −0.1440 | 0.0247 | 0.0034 |
| slope 13 | −0.0650 | −0.0540 | −0.1045 | −0.0035 | 0.0011 | −0.0548 | −0.1365 | 0.0269 | 0.0033 |
| slope 14 | −0.0650 | −0.0561 | −0.1077 | −0.0056 | 0.0011 | −0.0649 | −0.1470 | 0.0171 | 0.0039 |
| slope 15 | −0.0650 | −0.0559 | −0.1069 | −0.0061 | 0.0011 | −0.0658 | −0.1443 | 0.0127 | 0.0035 |
| slope 16 | −0.0650 | −0.0540 | −0.1053 | −0.0026 | 0.0012 | −0.0560 | −0.1434 | 0.0313 | 0.0043 |
| slope 17 | −0.0650 | −0.0562 | −0.1077 | −0.0061 | 0.0010 | −0.0696 | −0.1546 | 0.0154 | 0.0036 |
| slope 18 | −0.0350 | −0.0415 | −0.0956 | 0.0135 | 0.0014 | −0.0381 | −0.1247 | 0.0484 | 0.0044 |
| slope 19 | −0.0350 | −0.0407 | −0.0944 | 0.0145 | 0.0013 | −0.0322 | −0.1179 | 0.0536 | 0.0035 |
| slope 20 | −0.0350 | −0.0405 | −0.0943 | 0.0148 | 0.0013 | −0.0273 | −0.1148 | 0.0603 | 0.0036 |

Table Web-1: Simulation results from a model with relatively large domain-specific deviations, and outcome-specific deviations of zero. Rows labeled "slope 1" to "slope 20" refer to the estimated exposure effects on outcomes 1 to 20. Column headings under "Model results" and "Separate regressions" give the mean, 0.025 and 0.975 posterior quantiles, and mean squared error (MSE).

| | | Model results | | | | Separate regressions | | | |
|---|---|---|---|---|---|---|---|---|---|
| | true | mean | .025 | .975 | MSE | mean | .025 | .975 | MSE |
| $\beta_x$ | $-0.0500$ | $-0.0482$ | $-0.0918$ | $-0.0051$ | 0.0008 | | | | |
| $b_{\mathcal{D},1}$ | $-0.0100$ | $-0.0028$ | $-0.0334$ | 0.0233 | 0.0003 | | | | |
| $b_{\mathcal{D},2}$ | 0.0100 | 0.0031 | $-0.0222$ | 0.0321 | 0.0003 | | | | |
| $b_{\mathcal{D},3}$ | $-0.0100$ | $-0.0032$ | $-0.0320$ | 0.0216 | 0.0003 | | | | |
| $b_{\mathcal{D},4}$ | 0.0100 | 0.0029 | $-0.0241$ | 0.0341 | 0.0003 | | | | |
| slope 1 | $-0.0700$ | $-0.0533$ | $-0.1044$ | $-0.0040$ | 0.0013 | $-0.0649$ | $-0.1399$ | 0.0101 | 0.0029 |
| slope 2 | $-0.0550$ | $-0.0514$ | $-0.1013$ | $-0.0024$ | 0.0010 | $-0.0540$ | $-0.1242$ | 0.0162 | 0.0024 |
| slope 3 | $-0.0550$ | $-0.0509$ | $-0.1014$ | $-0.0011$ | 0.0010 | $-0.0526$ | $-0.1312$ | 0.0259 | 0.0028 |
| slope 4 | $-0.0450$ | $-0.0437$ | $-0.0928$ | 0.0062 | 0.0010 | $-0.0335$ | $-0.1155$ | 0.0486 | 0.0036 |
| slope 5 | $-0.0450$ | $-0.0452$ | $-0.0941$ | 0.0037 | 0.0010 | $-0.0454$ | $-0.1235$ | 0.0328 | 0.0030 |
| slope 6 | $-0.0450$ | $-0.0451$ | $-0.0945$ | 0.0042 | 0.0011 | $-0.0435$ | $-0.1211$ | 0.0341 | 0.0031 |
| slope 7 | $-0.0350$ | $-0.0447$ | $-0.0935$ | 0.0043 | 0.0011 | $-0.0417$ | $-0.1188$ | 0.0354 | 0.0026 |
| slope 8 | $-0.0350$ | $-0.0443$ | $-0.0938$ | 0.0057 | 0.0011 | $-0.0374$ | $-0.1234$ | 0.0487 | 0.0034 |
| slope 9 | $-0.0350$ | $-0.0441$ | $-0.0933$ | 0.0058 | 0.0011 | $-0.0351$ | $-0.1220$ | 0.0518 | 0.0035 |
| slope 10 | $-0.0650$ | $-0.0522$ | $-0.1020$ | $-0.0036$ | 0.0011 | $-0.0575$ | $-0.1384$ | 0.0235 | 0.0033 |
| slope 11 | $-0.0650$ | $-0.0519$ | $-0.1013$ | $-0.0031$ | 0.0011 | $-0.0572$ | $-0.1399$ | 0.0256 | 0.0037 |
| slope 12 | $-0.0650$ | $-0.0521$ | $-0.1015$ | $-0.0033$ | 0.0011 | $-0.0597$ | $-0.1440$ | 0.0247 | 0.0034 |
| slope 13 | $-0.0650$ | $-0.0516$ | $-0.1010$ | $-0.0031$ | 0.0011 | $-0.0548$ | $-0.1365$ | 0.0269 | 0.0033 |
| slope 14 | $-0.0550$ | $-0.0522$ | $-0.1018$ | $-0.0033$ | 0.0010 | $-0.0549$ | $-0.1370$ | 0.0271 | 0.0039 |
| slope 15 | $-0.0550$ | $-0.0517$ | $-0.1009$ | $-0.0032$ | 0.0010 | $-0.0558$ | $-0.1343$ | 0.0227 | 0.0035 |
| slope 16 | $-0.0550$ | $-0.0504$ | $-0.0998$ | $-0.0006$ | 0.0010 | $-0.0460$ | $-0.1334$ | 0.0413 | 0.0043 |
| slope 17 | $-0.0550$ | $-0.0524$ | $-0.1024$ | $-0.0035$ | 0.0009 | $-0.0596$ | $-0.1446$ | 0.0254 | 0.0036 |
| slope 18 | $-0.0500$ | $-0.0463$ | $-0.0986$ | 0.0056 | 0.0012 | $-0.0531$ | $-0.1397$ | 0.0334 | 0.0044 |
| slope 19 | $-0.0350$ | $-0.0438$ | $-0.0949$ | 0.0085 | 0.0012 | $-0.0322$ | $-0.1179$ | 0.0536 | 0.0035 |
| slope 20 | $-0.0350$ | $-0.0437$ | $-0.0950$ | 0.0089 | 0.0012 | $-0.0273$ | $-0.1148$ | 0.0603 | 0.0036 |

Table Web-2: Simulation results from a model with small domain-specific deviations, and small outcome-specific deviations. Rows labeled "slope 1" to "slope 20" refer to the estimated exposure effects on outcomes 1 to 20. Column headings under "Model results" and "Separate regressions" give the mean, 0.025 and 0.975 posterior quantiles, and mean squared error (MSE).