APPENDIX

**Priors and Bayesian Inference.** To complete the Bayesian model specifications, priors have to be assigned. For the direct effects $\beta_{0j}, \beta_j$ and the indirect parametric effects $\alpha$, we assume diffuse priors

$$p(\beta_{0j}) \propto \text{const}, \quad p(\beta_j) \propto \text{const}, \quad p(\alpha) \propto \text{const}.$$ 

For the factor loadings, we specify informative Gaussian priors

$$p(\lambda_j) \propto N(0, \sigma_j^2),$$

with $\sigma_j^2 = 1$ as the standard choice, to avoid the so-called Heywood cases (see e.g. Heywood). 

**Priors for functions.** For a function $f(w)$ of a continuous covariate $w$, we assume Bayesian P-spline-priors as in Brezger and Lang (2005). More specifically, the $f(w)$ approximated by a spline function and represented as a linear combination

$$f(w) = \sum_{i=1}^{L} \alpha_i B_i(w),$$

where $B_i(w)$ are B-spline basis functions. The vector $\alpha = (\alpha_1, ..., \alpha_L)$ follows a Gaussian random walk model of first or second orders, i.e the prior is Gaussian with
\[ p(\alpha) \propto \exp\left( -\frac{1}{2\tau} \alpha'K\alpha \right) \]  

(12)

**Priors for spatial effects.** We usually split \( f_{\text{geo}}(s) \) into a smooth structured effect and an unstructured effect, i.e.

\[
 f_{\text{geo}}(s) = f_{\text{str}}(s) + f_{\text{unstr}}(s)
\]

where \( f_{\text{str}}(s) \) models global spatial trends while \( f_{\text{unstr}}(s) \) captures local effects. For \( f_{\text{str}}(s) \) we assume a Markov random field prior (Beseg, York and Mallie, 1990). The spatial smoothness prior of function evaluations \( f_{\text{str}}(s) \) is

\[
 f_{\text{str},t} \mid f_{\text{str},t} \neq s, \tau_{\text{str}}^2 \sim N\left( \sum_{t \in \delta_s} \frac{f_{\text{str},t}}{N_s}, \frac{\tau_{\text{str}}^2}{N_s} \right)
\]

(13)

where \( N_s \) is the number of adjacent sites and \( t \in \delta_s \) denotes, that site \( f_s \) is a neighbor of site \( f_t \). Thus the (conditional) mean of \( f_s \) is an unweighted average of function evaluations of neighboring sites. Note that for spatial data conditioning is undirected since there is no natural ordering of different sites \( f_s \) as in the case for metrical covariates.

For the uncorrelated effect, we assume i.i.d. Gaussian random effects, i.e.

\[
 f_{\text{unstr}}(s) \sim N(0, \tau_{\text{unstr}}^2) \quad s = 1, \ldots, S
\]

The variance parameter \( \tau^2 \) controls the amount of the smoothness. It is estimated automatically by the inverse gamma IG\((a,b)\) with known hyperparameters \( a \) and \( b \). Standard choices for the hyperparameters are \( a = 1 \) and \( b = 0.005 \) or \( a = b = 0.001 \).