Supplementary Material for "Bayesian Variable Selection in Double Generalized Linear Tweedie Spatial Process Models"

Aritra Halder[a,*†], Shariq Mohammed[b†] and Dipak K. Dey[c]

[a]Department of Biostatistics, Dornsife School of Public Health,
Drexel University, Philadelphia, PA, USA
[b]Department of Biostatistics, Boston University School of Public Health,
Boston University, Boston, MA, USA
[b]Rafik B. Hariri Institute for Computing and Computational Science & Engineering,
Boston University, Boston, MA, USA
[c]Department of Statistics, University of Connecticut, Storrs, CT, USA
[*]corresponding author. E-mail: aritra.halder@drexel.edu
[†]equal contribution

June 15, 2023

S1 Posteriors

Under chosen priors on $\{\boldsymbol{\theta}_m, \boldsymbol{\theta}_{pr}\} = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \xi, \sigma^2, \phi_s\}$, the resulting joint posterior is specified by

$$\begin{aligned}\pi\Big(\boldsymbol{\theta}_m, \boldsymbol{\theta}_{pr} \mid \mathbf{y}\Big) \propto U\Big(\xi \mid a_\xi, b_\xi\Big) \times N_p\Big(\boldsymbol{\beta} \mid \mathbf{0}_p, \boldsymbol{\lambda}_\beta^T \mathbf{I}_p\Big) \times N_q\Big(\boldsymbol{\gamma} \mid \mathbf{0}_q, \boldsymbol{\lambda}_\gamma^T \mathbf{I}_q\Big) \times \\ U\Big(\phi_s \mid a_{\phi_s}, b_{\phi_s}\Big) \times \text{Gamma}\Big(\sigma^{-2} \mid a_\sigma, b_\sigma\Big) \times N_L\Big(\mathbf{w} \mid \mathbf{0}_L, \sigma^2 \mathbf{R}(\phi_s)\Big) \times \\ Tw\Big(\mathbf{y} \mid \boldsymbol{\mu} = \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{F}\mathbf{w}), \boldsymbol{\phi} = \exp(\mathbf{Z}\boldsymbol{\gamma}), \xi\Big),\end{aligned} \tag{S1}$$

We list the posteriors for individual parameters in the following equations,

$$\begin{aligned}&\pi\left(\boldsymbol{\beta} \mid \boldsymbol{\gamma}, \xi\right) \propto \exp\left\{-\left(\sum_{i=1}^{L}\sum_{j=1}^{n_i} \frac{1}{\phi_{ij}} d(y_{ij} \mid \mu_{ij}(\boldsymbol{\beta}, \mathbf{w}), \xi) + \frac{\sigma_\beta^{-2}}{2}||\boldsymbol{\beta}||_2^2 + \frac{\sigma^{-2}}{2}\mathbf{w}^T \mathbf{R}^{-1}(\phi_s)\mathbf{w}\right)\right\}, \\ &\pi\left(\boldsymbol{\gamma} \mid \boldsymbol{\beta}, \xi\right) \propto \exp\left\{-\left(\sum_{i=1}^{L}\sum_{j=1}^{n_i} \frac{1}{\phi_{ij}(\boldsymbol{\gamma})} d(y_{ij} \mid \mu_{ij}, \xi) + \frac{\log\phi_{ij}(\boldsymbol{\gamma})}{2} I(y_{ij} > 0) + \frac{\sigma_\gamma^{-2}}{2}||\boldsymbol{\gamma}||_2^2\right)\right\}, \\ &\pi(\xi \mid \boldsymbol{\beta}, \boldsymbol{\gamma}) \propto \prod_{i=1}^{L}\prod_{j=1}^{n_i} c_{ij}(y_{ij} \mid \phi_{ij}, \xi) \exp\left\{-\frac{1}{\phi_{ij}} d(y_{ij} \mid \mu_{ij}, \xi)\right\} I\left(\xi \in (a_\xi, b_\xi)\right), \\ &\pi(\phi_s) \propto |\mathbf{R}(\phi_s)|^{-1/2} \exp\left(-\frac{\sigma^{-2}}{2}\mathbf{w}^T \mathbf{R}^{-1}(\phi_s)\mathbf{w}\right) I\left(\phi_s \in (a_{\phi_s}, b_{\phi_s})\right), \\ &\pi(\sigma^{-2}) = \text{Gamma}\left(a_\sigma + \frac{L}{2}, b_\sigma + \frac{1}{2}\mathbf{w}^T \mathbf{R}^{-1}(\phi_s)\mathbf{w}\right),\end{aligned} \tag{S2}$$

where $|\cdot|$ denotes the determinant and $d(y_{ij} \mid \mu_{ij}, \xi)$ is the deviance function. In the scenario where we have no spatial effect in the DGLM, the reduced set of parameters are $\boldsymbol{\theta}_m = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \xi\}$ having a similar joint posterior after omitting the second row involving spatial process prior specification and setting $\boldsymbol{\mu} = \exp(\mathbf{X}\boldsymbol{\beta})$ in the likelihood, keeping the same prior specifications on the other parameters, the resulting posteriors are as follows,

$$\begin{aligned}&\pi\left(\boldsymbol{\beta} \mid \boldsymbol{\gamma}, \xi\right) \propto \exp\left\{-\left(\sum_{i=1}^{N} \frac{1}{\phi_i} d(y_i \mid \mu_i(\boldsymbol{\beta}), \xi) + \frac{\sigma_\beta^{-2}}{2}||\boldsymbol{\beta}||_2^2\right)\right\}, \\ &\pi\left(\boldsymbol{\gamma} \mid \boldsymbol{\beta}, \xi\right) \propto \exp\left\{-\left(\sum_{i=1}^{N} \frac{1}{\phi_i(\boldsymbol{\gamma})} d(y_i \mid \mu_i, \xi) + \frac{1}{2}\log\phi_{ij}(\boldsymbol{\gamma}) I(y_{ij} > 0) + \frac{\sigma_\gamma^{-2}}{2}||\boldsymbol{\gamma}||_2^2\right)\right\}, \\ &\pi(\xi) \propto \prod_{i=1}^{N} c_i(y_i \mid \phi_i, \xi) \exp\left\{-\frac{1}{\phi_i} d(y_i \mid \mu_i, \xi)\right\} I\left(\xi \in (a_\xi, b_\xi)\right).\end{aligned}$$

Updates leveraging MALA for $\boldsymbol{\beta}, \mathbf{w}$ (or $\boldsymbol{\beta}$) and $\boldsymbol{\gamma}$ require the proposals to be specified appropriately using the gradients of log-posterior densities, $\nabla \log \pi\left(\boldsymbol{\beta} \mid \boldsymbol{\gamma}, \xi\right)$ and $\nabla \log \pi\left(\boldsymbol{\gamma} \mid \boldsymbol{\beta}, \xi\right)$ respectively.

Candidate samples are obtained using,

$$
\begin{aligned}
(\boldsymbol{\beta}, \mathbf{w})^{T*} &= (\boldsymbol{\beta}, \mathbf{w})^{T} + \frac{\tau_{\beta,w}^2}{2}\mathbf{A}_{\beta,w}\nabla \log \pi\left(\boldsymbol{\beta} \mid \boldsymbol{\gamma}, \xi\right) + \tau_{\beta,w}\mathbf{A}_{\beta,w}^{1/2} \cdot N_{p+L}(\mathbf{0}, \mathbf{I}_{p+L}), \\
\boldsymbol{\gamma}^* &= \boldsymbol{\gamma} + \frac{\tau_{\gamma}^2}{2}\mathbf{A}_{\gamma}\nabla \log \pi\left(\gamma \mid \boldsymbol{\beta}, \xi\right) + \tau_{\gamma}\mathbf{A}_{\gamma}^{1/2} \cdot N_{q}(\mathbf{0}, \mathbf{I}_{q}),
\end{aligned} \tag{S3}
$$

where $\mathbf{A}_{\beta,w}^{-1} = \mathrm{E}\left[-\nabla^2 \log \pi\left(\boldsymbol{\beta}_{\mathbf{W}} \mid -\right)\right]$ and $\mathbf{A}_{\gamma}^{-1} = \mathrm{E}\left[-\nabla^2 \log \pi\left(\boldsymbol{\gamma} \mid -\right)\right]$.

Under a continuous spike and slab prior on $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ the posteriors for the hierarchical latent DGLM is as follows,

$$
\begin{aligned}
&\pi\left(\boldsymbol{\beta}_{\mathbf{W}} \mid -\right) \propto \exp\left\{-\left(\sum_{i=1}^{L}\sum_{j=1}^{n_i}\frac{1}{\phi_{ij}}d(y_{ij} \mid \mu_{ij}(\boldsymbol{\beta}, \mathbf{w}), \xi) + \frac{1}{2}\boldsymbol{\beta}^T\boldsymbol{\Gamma}_{\beta}^{-1}\boldsymbol{\beta} + \frac{\sigma^{-2}}{2}\mathbf{w}^T\mathbf{R}^{-1}(\phi_s)\mathbf{w}\right)\right\} \\
&\left.\begin{array}{l}
\boldsymbol{\Gamma}_{\beta} = \mathrm{diag}\{\zeta_{\beta,i_b}\sigma_{\beta,i_b}^2\} \\
\pi(\zeta_{\beta,i_b} \mid -) = \frac{\alpha_{1,\beta}}{\alpha_{1,\beta}+\alpha_{2,\beta}}\delta_{\nu_0}(\cdot) + \frac{\alpha_{2,\beta}}{\alpha_{1,\beta}+\alpha_{2,\beta}}\delta_{\nu_1}(\cdot) \\
\pi(\sigma_{\beta,i_b}^{-2} \mid -) = \mathrm{Gamma}\left(a_{\sigma} + \frac{1}{2}, b_{\sigma} + \frac{\beta_{i_b}^2}{2\zeta_{\beta,i_b}}\right) \\
\pi(\alpha_{\beta} \mid -) = \mathrm{Beta}\left(1 + \#\{i_b : \zeta_{\beta,i_b} = 1\}, 1 + \#\{i_b : \zeta_{\beta,i_b} = \nu_0\}\right)
\end{array}\right\}, i_b = 1, 2, \ldots, p \\
&\pi\left(\boldsymbol{\gamma} \mid -\right) \propto \exp\left\{-\left(\sum_{i=1}^{L}\sum_{j=1}^{n_i}\frac{1}{\phi_{ij}(\boldsymbol{\gamma})}d(y_{ij} \mid \mu_{ij}, \xi) + \frac{1}{2}\log\phi_{ij}(\boldsymbol{\gamma})I(y_{ij} > 0) + \frac{1}{2}\boldsymbol{\gamma}^T\boldsymbol{\Gamma}_{\gamma}^{-1}\boldsymbol{\gamma}\right)\right\}, \\
&\left.\begin{array}{l}
\boldsymbol{\Gamma}_{\gamma} = \mathrm{diag}\{\zeta_{\gamma,i_g}\sigma_{\gamma,i_g}^2\} \\
\pi(\zeta_{\gamma,i_g} \mid -) = \frac{\alpha_{1,\gamma}}{\alpha_{1,\gamma}+\alpha_{2,\gamma}}\delta_{\nu_0}(\cdot) + \frac{\alpha_{2,\gamma}}{\alpha_{1,\gamma}+\alpha_{2,\gamma}}\delta_{\nu_1}(\cdot) \\
\pi(\sigma_{\gamma,i_g}^{-2} \mid -) = \mathrm{Gamma}\left(a_{\sigma} + \frac{1}{2}, b_{\sigma} + \frac{\gamma_{i_g}^2}{2\zeta_{\gamma,i_g}}\right) \\
\pi(\alpha_{\gamma} \mid -) = \mathrm{Beta}\left(1 + \#\{i_g : \zeta_{\gamma,i_g} = 1\}, 1 + \#\{i_g : \zeta_{\gamma,i_g} = \nu_0\}\right)
\end{array}\right\}, i_g = 1, 2, \ldots, q \\
&\pi(\xi \mid -) \propto \prod_{i=1}^{L}\prod_{j=1}^{n_i} c_{ij}(y_{ij} \mid \phi_{ij}, \xi)\exp\left\{-\frac{1}{\phi_{ij}}d(y_{ij} \mid \mu_{ij}, \xi)\right\} I\left(\xi \in (a_{\xi}, b_{\xi})\right), \\
&\pi(\phi_s \mid -) \propto |\mathbf{R}(\phi_s)|^{-1/2}\exp\left(-\frac{\sigma^{-2}}{2}\mathbf{w}^T\mathbf{R}^{-1}(\phi_s)\mathbf{w}\right) I\left(\phi_s \in (a_{\phi_s}, b_{\phi_s})\right), \\
&\pi(\sigma^{-2} \mid -) = \mathrm{Gamma}\left(a_{\sigma} + \frac{L}{2}, b_{\sigma} + \frac{1}{2}\mathbf{w}^T\mathbf{R}^{-1}(\phi_s)\mathbf{w}\right),
\end{aligned} \tag{S4}
$$

where,

$$
\begin{aligned}
\alpha_{1,\beta} &= (1-\alpha_{\beta})\nu_0^{-1/2}\exp\left\{-\frac{\beta_{i_b}^2}{2\nu_0\sigma_{\beta,i_b}^2}\right\}, \;\; \alpha_{2,\beta} = \alpha_{\beta}\exp\left\{-\frac{\beta_{i_b}^2}{2\sigma_{\beta,i_b}^2}\right\}, i_b = 1, 2, \ldots, p; \\
\alpha_{1,\gamma} &= (1-\alpha_{\gamma})\nu_0^{-1/2}\exp\left\{-\frac{\gamma_{i_g}^2}{2\nu_0\sigma_{\gamma,i_g}^2}\right\}, \;\; \alpha_{2,\gamma} = \alpha_{\gamma}\exp\left\{-\frac{\gamma_{i_g}^2}{2\sigma_{\gamma,i_g}^2}\right\}, i_g = 1, 2, \ldots, q.
\end{aligned}
$$

Under no latent specification the resulting posteriors can be obtained similarly, by omitting updates for the latent process and associated process parameters/hyper-parameters. The joint posterior is sampled by leveraging the posteriors in eq. S4 and employing a similar hybrid sampling strategy as earlier, with additional Gibbs updates for the spike and slab parameters, $\{\boldsymbol{\zeta}_\beta, \boldsymbol{\sigma}^2_\beta, \alpha_\beta, \boldsymbol{\zeta}_\gamma, \boldsymbol{\sigma}^2_\gamma, \alpha_\gamma\}$.

S2 Experiments, Diagnostics and Software

We outline supporting simulation experiments, convergence diagnostics and details on available software to supplement Section 3 in the manuscript. We begin with results from additional experiments.

S2.1 Supporting Experiments

We performed simulation experiments to assess performance of model M2—Table S1 shows the results featuring the same settings as outlined in Section 3, only without the spatial effects.

Configuration	Prop. of 0s	$CP(\boldsymbol{\theta})$	False Positive Rate	True Positive Rate
Configuration 1	0.15	1.00 (0.00)	0.00 (0.00)	1.00 (0.00)
	0.30	1.00 (0.00)	0.00 (0.00)	1.00 (0.00)
	0.60	1.00 (0.00)	0.00 (0.00)	1.00 (0.00)
	0.80	1.00 (0.10)	0.04 (0.06)	0.99 (0.03)
Configuration 2	0.15	1.00 (0.00)	0.01 (0.05)	0.99 (0.03)
	0.30	1.00 (0.00)	0.00 (0.00)	1.00 (0.00)
	0.60	1.00 (0.10)	0.00 (0.00)	0.98 (0.04)
	0.80	0.90 (0.10)	0.01 (0.05)	0.99 (0.03)
Configuration 3	0.15	0.90 (0.10)	0.00 (0.00)	1.00 (0.00)
	0.30	1.00 (0.00)	0.01 (0.04)	1.00 (0.00)
	0.60	1.00 (0.00)	0.00 (0.00)	0.99 (0.03)
	0.80	1.00 (0.00)	0.05 (0.09)	0.96 (0.05)

Table S1: Results for synthetic experiments corresponding to Table 3 for model M2 showing average coverage probabilities for the estimated model coefficients.

In what follows, we outline simulation experiments for the scenario where we have spatial covariates in our data. Finally, we focus on showing simulations results for different synthetic spatial patterns.

S2.1.1 Spatial Covariates

Setting up experiments with spatial covariates we are mindful of possible endogeneity issues arising from included spatial covariates being correlated with the true spatial pattern (see, for e.g., Fan and Liao, 2014). We use the following settings and true values: $N = 1 \times 10^4$, $L = 1 \times 10^2$, $\boldsymbol{\beta} = (1.0, 1.5, 1 \times 10^{-5}, 1.4, 1.1, 1 \times 10^{-5}, 2.5)^T$, $\boldsymbol{\gamma} = (1.0, 1 \times 10^{-5}, 1.5, 1.1, 1 \times 10^{-5}, -2.5, 1 \times 10^{-5})^T$. In $\mathbf{X}$ and $\mathbf{Z}$ we include two spatial covariates in the last two columns: $\mathbf{x}_6 = \mathbf{z}_6 \sim N(5(\cos(3\pi s_x) + \sin(3\pi s_y)), 1)$ and $\mathbf{x}_7 = \mathbf{z}_7 \sim N(2(\cos(\pi s_x) + \sin(\pi s_y)), 1)$, while the rest were sampled from standard Gaussian distributions resulting in an average of 50% zeros in the data. Under this setup, we observe that for the mean model $\mathbf{x}_6$ is significant and $\mathbf{x}_7$ is not while for the dispersion model it is the other way around. The true spatial effect is simulated using $\sigma^2 = 1.5$ and $\phi_s = 3$ from an exponential kernel. The true index parameter, $\xi = 1.5$. The design matrices were centered and scaled. The resulting absolute value of the correlations within covariates and between them and the true spatial effect from the setup above was < 0.01. To assess the performance of our models we focus on M3 and and M4 and compute the MSE, coverage probability, FPR and TPR for each of these models. We performed 100 replications under these settings. The results are shown below in Table S2. Models M3 and M4 were able to estimate coefficients β_6, β_7 and γ_6, γ_7 with sufficient accuracy.

Model	MSE ($\boldsymbol{\theta}$)	MSE ($\mathbf{w}$)	CP ($\boldsymbol{\theta}$)	CP ($\mathbf{w}$)	FPR	TPR
M3	0.008	0.072	0.907	0.940	–	–
	(0.019)	(0.186)	(0.111)	(0.092)	–	–
M4	0.003	0.062	0.916	0.959	0.000	1.000
	(0.004)	(0.129)	(0.063)	(0.09)	(0.165)	(0.035)

Table S2: Table showing results for simulation carried using spatial covariates in design matrices $\mathbf{X}$ and $\mathbf{Z}$.

S2.1.2 Spatial Patterns

We show the results for simulation experiments performed for patters shown in Figure 1 of the manuscript in Tables S3 and S4 below.

N	Prop. of 0s	MSE						CP($\boldsymbol{\theta}$)	
		$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\mathbf{w}$	ξ	σ^2	ϕ_s	$\boldsymbol{\theta}_m$	$\boldsymbol{\theta}_{pr}$
2000	0.15	0.00	0.00	0.02	0.00	0.04	1.45	1.00	0.95
	(0.06)	(0.00)	(0.00)	(0.01)	(0.00)	(0.05)	(2.98)	(0.00)	(0.01)
	0.30	0.00	0.00	0.06	0.00	0.11	4.48	1.00	0.94
	(0.07)	(0.00)	(0.00)	(0.02)	(0.00)	(0.07)	(6.12)	(0.00)	(0.03)
	0.60	0.00	0.00	0.08	0.00	0.08	6.46	0.98	0.94
	(0.05)	(0.00)	(0.00)	(0.02)	(0.00)	(0.12)	(14.51)	(0.05)	(0.03)
	0.80	0.01	0.00	0.21	0.00	0.09	10.00	0.96	0.92
	(0.04)	(0.00)	(0.00)	(0.04)	(0.00)	(0.08)	(21.00)	(0.06)	(0.03)
5000	0.15	0.00	0.00	0.01	0.00	0.06	1.89	0.99	0.93
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(3.07)	(0.04)	(0.02)
	0.30	0.00	0.00	0.03	0.00	0.09	5.07	0.94	0.93
	(0.05)	(0.00)	(0.00)	(0.01)	(0.00)	(0.08)	(9.23)	(0.11)	(0.02)
	0.60	0.00	0.00	0.04	0.00	0.04	2.05	0.99	0.94
	(0.07)	(0.00)	(0.00)	(0.01)	(0.00)	(0.08)	(4.86)	(0.04)	(0.02)
	0.80	0.00	0.00	0.10	0.00	0.10	12.86	0.98	0.94
	(0.05)	(0.00)	(0.00)	(0.03)	(0.00)	(0.08)	(27.54)	(0.05)	(0.02)
10000	0.15	0.00	0.00	0.00	0.00	0.13	1.30	0.99	0.95
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.21)	(2.72)	(0.04)	(0.01)
	0.30	0.00	0.00	0.01	0.00	0.10	2.70	0.99	0.94
	(0.08)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(3.88)	(0.04)	(0.03)
	0.60	0.00	0.00	0.02	0.00	0.07	4.16	0.95	0.95
	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)	(9.62)	(0.06)	(0.02)
	0.80	0.00	0.00	0.06	0.00	0.05	1.34	0.99	0.93
	(0.04)	(0.00)	(0.00)	(0.02)	(0.00)	(0.04)	(1.61)	(0.04)	(0.03)

Table S3: Results for Pattern 1, with number of locations, $L = 100$.

N	Prop. of 0s	MSE				CP($\boldsymbol{\theta}$)	
		$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\mathbf{w}$	ξ	$\boldsymbol{\theta}_m$	$\boldsymbol{\theta}_{pr}$
2000	0.15	0.00	0.00	0.06	0.00	0.98	0.93
	(0.01)	(0.00)	(0.00)	(0.03)	(0.00)	(0.05)	(0.03)
	0.30	0.00	0.00	0.34	0.00	0.96	0.95
	(0.01)	(0.00)	(0.00)	(0.10)	(0.00)	(0.06)	(0.02)
	0.60	0.00	0.00	0.92	0.00	0.96	0.94
	(0.01)	(0.00)	(0.00)	(0.31)	(0.00)	(0.08)	(0.03)
	0.80	0.00	0.00	2.74	0.00	1.00	0.91
	(0.01)	(0.00)	(0.00)	(0.26)	(0.00)	(0.00)	(0.02)
5000	0.15	0.00	0.00	0.03	0.00	0.95	0.95
	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)	(0.11)	(0.02)
	0.30	0.00	0.00	0.13	0.00	0.98	0.93
	(0.01)	(0.00)	(0.00)	(0.04)	(0.00)	(0.05)	(0.03)
	0.60	0.00	0.00	0.46	0.00	1.00	0.95
	(0.01)	(0.00)	(0.00)	(0.15)	(0.00)	(0.00)	(0.02)
	0.80	0.00	0.00	1.68	0.00	0.91	0.91
	(0.01)	(0.00)	(0.00)	(0.28)	(0.00)	(0.10)	(0.03)
10000	0.155	0.00	0.00	0.01	0.00	0.92	0.93
	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.06)	(0.03)
	0.30	0.00	0.00	0.09	0.00	0.99	0.93
	(0.01)	(0.00)	(0.00)	(0.03)	(0.00)	(0.04)	(0.03)
	0.60	0.00	0.00	0.30	0.00	0.96	0.93
	(0.01)	(0.00)	(0.00)	(0.07)	(0.00)	(0.06)	(0.02)
	0.80	0.00	0.00	1.37	0.00	0.98	0.93
	(0.01)	(0.00)	(0.00)	(0.19)	(0.00)	(0.05)	(0.03)

Table S4: Results for Pattern 2, with number of locations, $L = 100$.

S2.2 Convergence

To assess convergence we resort to monitoring trace plots, posterior density and auto-correlation (ACF) plots for model parameters and coefficients. We show them particularly for the index parameter ξ in Figure S1 below. They were produced using parameter settings outlined in the examples within the R-package. Additional plots for $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, σ^2 and ϕ_s are included in examples within the R-package described in the next subsection.

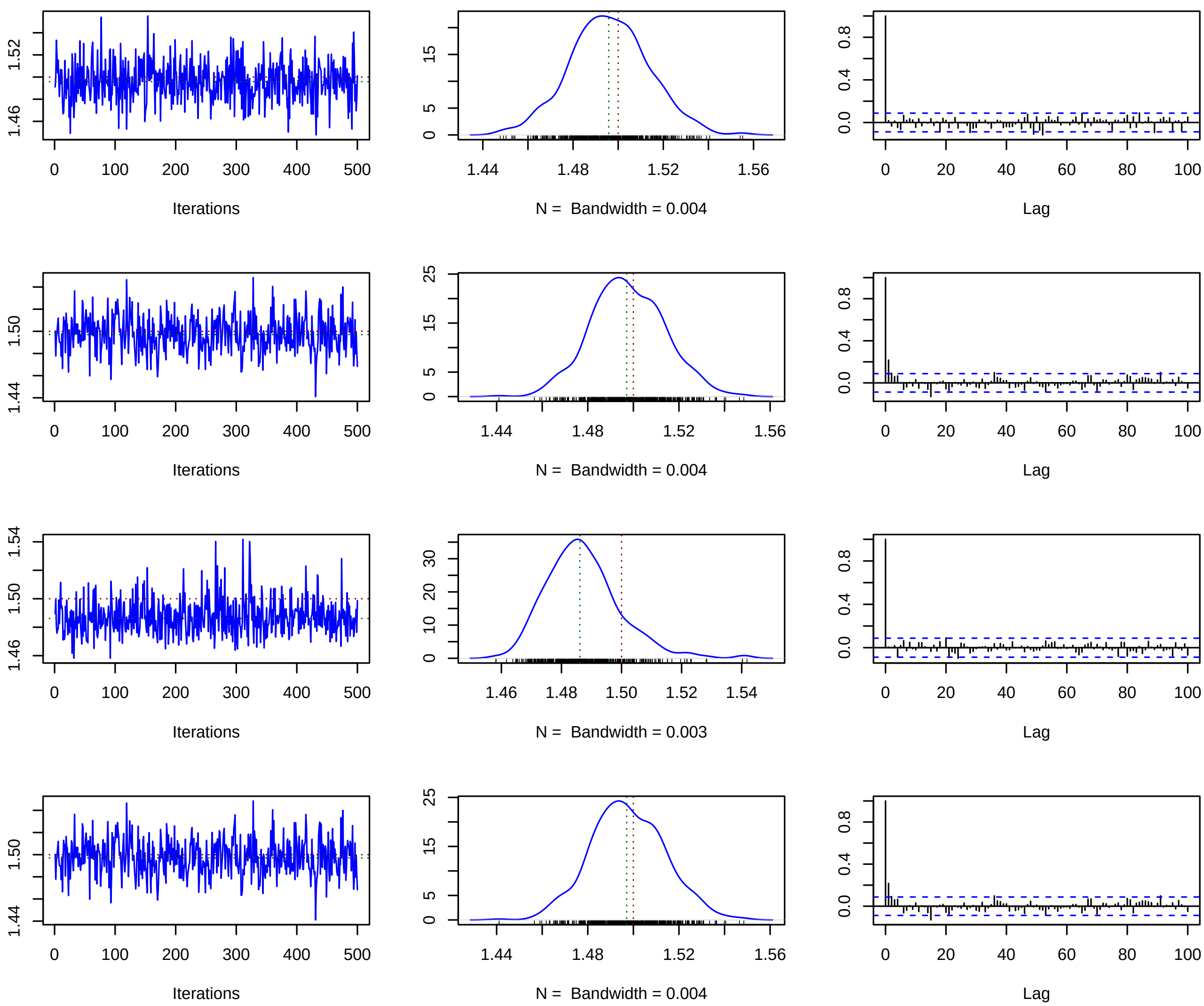

Figure S1: Convergence diagnostics for posterior samples of the index parameter, ξ. Each row corresponds to a model—(first row) M1 (second row) M2 (third row) M3 (fourth row) M4. In each row we show (left) trace, (center) posterior probability density and (right) ACF plots.

S2.3 R-package: `sptwdglm`

Algorithms featuring posteriors in Section S1 constitute our implementation that is available in the publicly accessible open-source repository: https://github.com/arh926/sptwdglm. It is composed of four Markov chain Monte Carlo samplers as described below. Each function is accompanied with examples in the package.

1. `dglm.autograd.R`: Fits a DGLM to the data. It requires the response as a vector, y, covariates as matrices $\mathbf{X}$ and $\mathbf{Z}$, upper and lower bounds for the index parameter ξ as input at the minimum to produce an output. Other optional parameters are also provided for custom fine-tuning which have optimized defaults for ease of use. This is model M1 in the manuscript.

2. `ssdglm.autograd.R`: Fits a DGLM featuring variable selection via the spike and slab prior to the data. It requires the response as a vector, y, covariates as matrices $\mathbf{X}$ and $\mathbf{Z}$, upper and lower bounds for the index parameter ξ as input at the minimum to produce an output. Other optional parameters are also provided for custom fine-tuning which have optimized defaults for ease of use. This is model M2 in the manuscript.

3. `spdglm.autograd.R`: Fits a spatial DGLM to the data. It requires the coordinates as a matrix, the response as a vector, y, covariates as matrices $\mathbf{X}$ and $\mathbf{Z}$, upper and lower bounds for the index parameter ξ as input at the minimum to produce an output. Other optional parameters are also provided for custom fine-tuning which have optimized defaults for ease of use. This is model M3 in the manuscript.

4. `spssdglm.autograd.R`: Fits a spatial DGLM featuring variable selection via the spike and slab prior to the data. It requires the coordinates as a matrix, the response as a vector, y, covariates as matrices $\mathbf{X}$ and $\mathbf{Z}$, upper and lower bounds for the index parameter ξ as input at the minimum to produce an output. Other optional parameters are also provided for custom fine-tuning which have optimized defaults for ease of use. This is model M4 in the manuscript.

5. `FDR.R`: Computes local false discovery rates (see page 10 of manuscript) for posterior samples of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ arising from M2 or M4. Optional parameters to be specified are threshold defaulting at 0.05 and percentage at which FDR is computed defaulting to 5%.

S3 Automobile Insurance Premiums, Connecticut, 2008: Additional results

Parameters	Levels	Median	Mean	SD	Lower HPD	Upper HPD
(Intercept)	–	5.039	4.994	0.372	4.286	5.676
age.car	0	-0.373	-0.344	0.224	-0.787	0.111
	1	-0.306	-0.271	0.218	-0.641	0.224
	2	-0.196	-0.164	0.219	-0.519	0.333
	3	-0.371	-0.331	0.262	-0.740	0.262
	4	-0.187	-0.156	0.225	-0.528	0.369
	5	-0.111	-0.078	0.227	-0.454	0.437
	6	-0.350	-0.320	0.230	-0.696	0.213
	7	-0.401	-0.367	0.252	-0.855	0.138
risk	S	-0.176	-0.173	0.060	-0.292	-0.054
agec	2	-0.014	-0.013	0.097	-0.208	0.163
	3	-0.292	-0.292	0.101	-0.504	-0.105
	4	-0.150	-0.152	0.114	-0.347	0.096
	5	-0.430	-0.424	0.230	-0.881	-0.022
gender	F	0.451	0.436	0.220	0.015	0.838
	M	0.279	0.268	0.219	-0.141	0.668
marital	M	-4.096	-3.941	1.031	-5.532	-1.829
	S	4.954	4.893	1.751	2.285	7.588
deductible	B	2.613	2.595	0.832	0.987	4.065
	C	1.025	0.999	0.439	0.049	1.779
	D	1.519	1.487	0.432	0.605	2.322
	E	1.626	1.579	0.429	0.651	2.375
	F	1.691	1.678	0.424	0.695	2.412
	G	1.733	1.717	0.433	0.879	2.626
	H	1.804	1.754	0.630	0.552	2.885
genderMarital	A	-4.998	-4.940	1.725	-7.626	-2.515
	B	3.756	3.622	1.057	1.413	5.219
	C	-4.521	-4.532	1.823	-7.360	-1.766
	D	3.850	3.685	0.986	1.814	5.412

Table S5: Model coefficients for fixed effects in the mean model for M1. We show the median, mean, standard deviation and HPDs.

Parameters	Levels	Median	Mean	SD	Lower HPD	Upper HPD
(Intercept)	–	5.871	5.867	0.052	5.784	5.948
age.car	0	-0.814	-0.804	0.088	-0.915	-0.663
	1	-1.494	-1.471	0.136	-1.657	-1.217
	2	-1.122	-1.112	0.118	-1.299	-0.918
	3	0.720	0.737	0.162	0.485	1.009
	4	-0.827	-0.826	0.072	-0.937	-0.709
	5	-0.045	-0.033	0.122	-0.203	0.170
	6	-0.502	-0.489	0.057	-0.571	-0.392
	7	0.672	0.673	0.128	0.485	0.887
risk	S	1.688	1.693	0.048	1.620	1.772
agec	2	1.333	1.343	0.025	1.317	1.400
	3	1.384	1.388	0.010	1.374	1.409
	4	1.394	1.396	0.013	1.370	1.417
	5	1.157	1.160	0.124	0.981	1.351
gender	F	0.168	0.151	0.053	0.055	0.214
	M	-0.074	-0.075	0.027	-0.118	-0.024
marital	M	0.022	0.044	0.055	-0.034	0.131
	S	2.308	2.277	0.243	1.829	2.665
deductible	B	-0.301	-0.280	0.052	-0.349	-0.184
	C	-0.122	-0.131	0.093	-0.306	-0.011
	D	-0.519	-0.509	0.154	-0.746	-0.241
	E	-0.297	-0.287	0.183	-0.563	0.017
	F	0.446	0.444	0.051	0.358	0.515
	G	1.664	1.662	0.039	1.596	1.720
	H	0.513	0.516	0.042	0.437	0.583
genderMarital	A	-0.912	-0.904	0.166	-1.175	-0.624
	B	0.270	0.270	0.005	0.257	0.279
	C	-0.660	-0.654	0.126	-0.856	-0.456
	D	-0.222	-0.236	0.042	-0.309	-0.176
ξ	–	1.674	1.674	0.007	1.664	1.687

Table S6: Model coefficients for fixed effects in the dispersion model for M1. We show the median, mean, standard deviation and HPDs.

Parameters	Levels	Median	Mean	SD	Lower HPD	Upper HPD
(Intercept)	–	5.325	5.343	0.102	5.173	5.542
age.car	0	0.023	0.021	0.080	-0.109	0.163
	1	0.084	0.081	0.080	-0.066	0.215
	2	0.194	0.192	0.080	0.051	0.330
	3	-0.036	-0.040	0.080	-0.187	0.094
	4	0.021	0.020	0.080	-0.126	0.155
	5	0.072	0.068	0.080	-0.074	0.209
	6	-0.022	-0.025	0.081	-0.168	0.116
	7	-0.072	-0.078	0.083	-0.224	0.057
risk	S	-0.195	-0.195	0.024	-0.238	-0.143
agec	2	-0.135	-0.135	0.032	-0.200	-0.075
	3	-0.479	-0.480	0.031	-0.540	-0.421
	4	-0.626	-0.626	0.036	-0.696	-0.559
	5	-0.694	-0.694	0.059	-0.814	-0.581
gender	F	0.602	0.605	0.090	0.438	0.789
	M	0.721	0.726	0.094	0.552	0.900
marital	M	1.527	1.623	0.286	1.197	2.152
	S	-0.354	-0.339	0.258	-0.763	0.222
deductible	B	1.371	1.366	0.222	0.921	1.815
	C	0.147	0.088	0.253	-0.353	0.495
	D	0.535	0.465	0.247	0.010	0.829
	E	0.718	0.637	0.243	0.207	1.010
	F	0.680	0.598	0.243	0.178	0.979
	G	0.680	0.599	0.240	0.188	0.970
	H	0.766	0.791	0.331	0.143	1.436
genderMarital	A	0.226	0.215	0.223	-0.240	0.623
	B	-2.005	-2.065	0.257	-2.551	-1.664
	C	0.281	0.275	0.273	-0.278	0.720
	D	-2.023	-2.107	0.299	-2.604	-1.680

Table S7: Model coefficients for fixed effects in the mean model for M3.

Parameters	Levels	Median	Mean	SD	Lower HPD	Upper HPD
(Intercept)	–	6.655	6.645	0.093	6.473	6.842
age.car	0	-0.837	-0.838	0.049	-0.926	-0.734
	1	-1.053	-1.051	0.044	-1.139	-0.967
	2	-0.916	-0.915	0.042	-0.992	-0.828
	3	-0.889	-0.889	0.041	-0.966	-0.805
	4	-0.875	-0.873	0.041	-0.955	-0.793
	5	-0.824	-0.824	0.040	-0.908	-0.750
	6	-0.798	-0.796	0.040	-0.869	-0.712
	7	-0.764	-0.763	0.040	-0.851	-0.693
risk	S	0.105	0.104	0.015	0.075	0.134
agec	2	-0.027	-0.027	0.022	-0.067	0.018
	3	-0.042	-0.042	0.014	-0.072	-0.015
	4	-0.078	-0.078	0.014	-0.105	-0.049
	5	-0.280	-0.279	0.029	-0.338	-0.225
gender	F	0.004	0.003	0.038	-0.071	0.077
	M	0.005	0.005	0.043	-0.076	0.089
marital	M	-3.145	-3.154	0.180	-3.460	-2.839
	S	-3.280	-3.281	0.205	-3.642	-2.900
deductible	B	-1.226	-1.224	0.133	-1.455	-0.943
	C	-0.751	-0.758	0.082	-0.940	-0.620
	D	-0.808	-0.815	0.080	-0.973	-0.664
	E	-0.753	-0.759	0.078	-0.926	-0.626
	F	-0.598	-0.604	0.078	-0.774	-0.473
	G	-0.236	-0.240	0.076	-0.405	-0.107
	H	0.224	0.224	0.138	-0.049	0.473
genderMarital	A	3.172	3.169	0.202	2.800	3.536
	B	3.233	3.247	0.175	2.948	3.540
	C	3.307	3.305	0.203	2.909	3.639
	D	3.265	3.280	0.176	2.980	3.575
ξ	–	1.671	1.671	0.004	1.664	1.679

Table S8: Model coefficients for fixed effects in the dispersion model for M3.

References

Fan, J. and Liao, Y. (2014). Endogeneity in high dimensions. *Annals of statistics*, 42(3):872.