Supplementary material: "Alleviating confounding in spatio-temporal areal models with an application on crimes against women in India"

Software

Software in the form of R code, together with a sample input data set and complete documentation will be available at `https://github.com/spatialstatisticsupna/Confounding_article`.

A Supplementary Material A

This section shows that placing constraints is equivalent to an oblique projection. In particular, we focus on the situation studied in the paper where constraints on Gaussian variables (the random effects) are specified by a precision matrix without constraints on the null space. In the paper, the precision matrices of the random effects are rank deficient and usually sum-to-zero constraints corresponding to the null space are required to fit the model. However, as we are making the random effects orthogonal to the fixed effects (intercept included), the usual sum-to-zero constraints are not required and must be replaced with weighted sum-to-zero constraints.

Theorem 1. *Let* $\boldsymbol{Y}$ *be a random variable of length* n *with density*

$$p_Y(\boldsymbol{y}) \propto \exp\left[-\frac{1}{2}\boldsymbol{y}^{'}\mathbf{Q}\boldsymbol{y}\right], \tag{A.1}$$

with $\mathbf{Q}$ *not necessarily of full rank. Let* $\mathcal{A}$ *be a subspace of* $\mathbb{R}^n$ *such that* $\mathcal{A} \cap \mathcal{K}(\mathbf{Q}) =$

$\{\mathbf{0}\}$, and $\mathcal{K}(\mathbf{Q})$ stands for the null space (kernel) of $\mathbf{Q}$. Let $\mathbf{A}$ and $\mathbf{B}$ be matrices with rows that form orthonormal bases for $\mathcal{A}$ and its orthogonal complement $\mathcal{A}^{\perp}$, respectively. Finally, let

$$\mathbf{P}_A = \mathbf{A}^{'}(\mathbf{A}\mathbf{Q}\mathbf{A}^{'})^{-1}\mathbf{A}\mathbf{Q}. \tag{A.2}$$

Then the following distributions are equal:

1. $[\boldsymbol{Y}|\mathbf{B}\boldsymbol{Y} = \mathbf{0}]$;

2. $[\boldsymbol{Y}|\boldsymbol{Y} \in \mathcal{A}]$;

3. $[\mathbf{P}_A\boldsymbol{Y}]$;

4. $\mathcal{N}[\mathbf{0}, \mathbf{A}^{'}(\mathbf{A}\mathbf{Q}\mathbf{A}^{'})^{-1}\mathbf{A}]$.

Proof. Let $\mathbf{M} = (\mathbf{A}^{'}\ \mathbf{B}^{'})^{'}$ be orthogonal where $\mathbf{A}$ and $\mathbf{B}$ are $(n-c) \times n$ and $c \times n$ matrices, and define $\boldsymbol{z} = \mathbf{M}\boldsymbol{y}$. Then

$$\begin{aligned} p_Z(\boldsymbol{z}) &= p_Y(\mathbf{M}^{-1}\boldsymbol{z})|\mathbf{M}^{-1}|, \\ &\propto \exp\left[-\frac{1}{2}\boldsymbol{z}^{'}(\mathbf{M}^{-1})^{'}\mathbf{Q}\mathbf{M}^{-1}\boldsymbol{z}\right], \\ &= \exp\left[-\frac{1}{2}\boldsymbol{z}^{'}(\mathbf{M}\mathbf{Q}\mathbf{M}^{'})\boldsymbol{z}\right]. \end{aligned} \tag{A.3}$$

Letting

$$\begin{pmatrix} \mathbf{K} & \mathbf{L} \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{K}^{'} \\ \mathbf{L}^{'} \end{pmatrix}$$

be the spectral decomposition of $\mathbf{Q}$, where $\mathbf{K}$ and $\mathbf{L}$ are matrices with eigenvectors having non-null and null eigenvalues respectively, and $\mathbf{R}$ is a diagonal matrix with the non-null and positive eigenvalues, note that

$$\mathbf{M}\mathbf{Q}\mathbf{M}^{'} = \begin{pmatrix} \mathbf{A}\mathbf{Q}\mathbf{A}^{'} & \mathbf{A}\mathbf{Q}\mathbf{B}^{'} \\ \mathbf{B}\mathbf{Q}\mathbf{A}^{'} & \mathbf{B}\mathbf{Q}\mathbf{B}^{'} \end{pmatrix}$$
$$= \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{K} & \mathbf{L} \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{K}^{'} \\ \mathbf{L}^{'} \end{pmatrix} \begin{pmatrix} \mathbf{A}^{'} & \mathbf{B}^{'} \end{pmatrix}.$$

Then, $\mathbf{AQA}^{'} = \mathbf{AKRK}^{'}\mathbf{A}^{'}$. Let $\boldsymbol{x} \neq \mathbf{0}$. Since $\mathbf{A}$ is of full row rank, $\mathbf{A}^{'}\boldsymbol{x} \neq \mathbf{0}$, and $\mathbf{K}^{'}\mathbf{A}^{'}\boldsymbol{x} \neq \mathbf{0}$ as long as $\mathbf{A}^{'}\boldsymbol{x} \notin \mathcal{R}(\mathbf{K}^{'})^{\perp} = \mathcal{R}(\mathbf{L}^{'}) = \mathcal{K}(\mathbf{Q})$, where $\mathcal{R}()$ indicate the row space of a matrix. Since $\mathbf{A}^{'}\boldsymbol{x} \in \mathcal{R}(\mathbf{A})$, and $\mathcal{R}(\mathbf{A}) \cap \mathcal{K}(\mathbf{Q}) = \{\mathbf{0}\}$, $\mathbf{K}^{'}\mathbf{A}^{'}\boldsymbol{x} \neq \mathbf{0}$. Since $\mathbf{R}$ is trivially positive definite, $\boldsymbol{x}^{'}\mathbf{AKRK}^{'}\mathbf{A}^{'}\boldsymbol{x} > 0$, thus $\mathbf{AQA}^{'} = \mathbf{AKRK}^{'}\mathbf{A}^{'}$ is positive definite and therefore invertible. We have

$$\begin{aligned} p(\boldsymbol{z}_{[1,n-c]}|\mathbf{B}\boldsymbol{y} = \mathbf{0}) &= p(\boldsymbol{z}_{[1,n-c]}|\boldsymbol{z}_{[n-c+1,n]} = \mathbf{0}), \\ &\propto \exp\left[-\frac{1}{2}\boldsymbol{z}^{'}_{[1,n-c]}\mathbf{AQA}^{'}\boldsymbol{z}_{[1,n-c]}\right], \end{aligned} \tag{A.4}$$

where $\boldsymbol{z}_{[a,b]} = (z_a, \ldots, z_b)^{'}$. Then, since $\mathbf{AQA}^{'}$ is invertible,

$$[\boldsymbol{Y}|\mathbf{B}\boldsymbol{Y} = \mathbf{0}] \sim \mathcal{N}[\mathbf{0}, \mathbf{V}], \quad \mathbf{V} = \begin{pmatrix} \mathbf{A}^{'} & \mathbf{B}^{'} \end{pmatrix} \begin{pmatrix} (\mathbf{AQA}^{'})^{-1} & \mathbf{0} \\ \mathbf{0}^{'} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} = \mathbf{A}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{A}. \tag{A.5}$$

When the Gaussian random variable $\boldsymbol{Y}$ is specified via the precision matrix $\mathbf{Q}$, $\mathcal{K}(\mathbf{Q})$ contains the unidentified degrees of freedom in the sense that $f_Y(\boldsymbol{y} + \boldsymbol{b}) = f_Y(\boldsymbol{y})$ for all $\boldsymbol{b} \in \mathcal{K}(\mathbf{Q})$. Note that $\mathcal{K}(\mathbf{P}_A) = \mathcal{K}(\mathbf{Q})$, so $\mathbf{P}_A(\boldsymbol{y} + \boldsymbol{b}) = \mathbf{P}_A\boldsymbol{y}$ for all $\boldsymbol{b} \in \mathcal{K}(\mathbf{Q})$, so $\mathbf{P}_A\boldsymbol{Y}$ is identified because changing $\boldsymbol{Y}$ adding $\boldsymbol{b}$ does not change $\mathbf{P}_A\boldsymbol{Y}$. Thus, when computing $\mathrm{Var}[\mathbf{P}_A \boldsymbol{Y}]$, we may restrict $\boldsymbol{Y}$ to $\mathcal{K}(\mathbf{Q})^\perp$ so that $\mathrm{Var}[\boldsymbol{Y}] = \mathbf{Q}^-$. Then,

$$\begin{aligned}
\mathrm{Var}[\mathbf{P}_A \boldsymbol{Y}] &= \left[\mathbf{A}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{AQ}\right] \mathbf{Q}^- \left[\mathbf{A}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{AQ}\right]^{'}, \\
&= \mathbf{A}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{AQQ}^-\mathbf{QA}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{A}, \\
&= \mathbf{A}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{AQA}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{A}, \\
&= \mathbf{A}^{'}(\mathbf{AQA}^{'})^{-1}\mathbf{A}.
\end{aligned}$$

□

Notes:

In the situation described in the paper, we need $\mathbf{A}_\xi$ and $\mathbf{B}_\xi$, $\mathbf{A}_\gamma$ and $\mathbf{B}_\gamma$, and $\mathbf{A}_\delta$ and $\mathbf{B}_\delta$. The $\mathbf{B}$ matrices (the constraints matrices) are given by Equation (3.5), and each $\mathbf{A}$ matrix may be constructed by taking its rows to be the eigenvectors of the orthogonal projection matrix $\mathbf{I}_n - \mathbf{B}^{'}(\mathbf{BB}^{'})^{-1}\mathbf{B}$ whose eigenvalues are 1 (equivalently, non-zero). Since the projection is orthogonal, its matrix is symmetric and admits a set of orthonormal eigenvectors forming a basis for $\mathbb{R}^n$, and its null space $\mathcal{R}(\mathbf{B})$ is orthogonal to its row space, the eigenvectors whose eigenvalues are non-zero. The condition $\mathcal{R}(\mathbf{A}) \cap \mathcal{K}(\mathbf{Q}) = \{\mathbf{0}\}$ ensures that the constraint $\mathbf{B}\boldsymbol{Y} = \mathbf{0}$ is sufficient to identify $\boldsymbol{Y}$.

Expressions for constraint matrices $\mathbf{B}_\xi$, $\mathbf{B}_\gamma$, and $\mathbf{B}_\delta$

$$\mathbf{B}_\xi = \mathbf{X}_*^{'}\hat{\mathbf{W}}(\mathbf{1}_T \otimes \mathbf{I}_S) = \begin{pmatrix} \hat{w}_{1.} & \dots & \hat{w}_{S.} \\ (\hat{w}x_1)_{1.} & \cdots & (\hat{w}x_1)_{S.} \\ \vdots & \ddots & \vdots \\ (\hat{w}x_p)_{1.} & \cdots & (\hat{w}x_p)_{S.} \end{pmatrix},$$

$$\mathbf{B}_\gamma = \mathbf{X}_*^{'}\hat{\mathbf{W}}(\mathbf{I}_T \otimes \mathbf{1}_S) = \begin{pmatrix} \hat{w}_{.1} & \dots & \hat{w}_{.T} \\ (\hat{w}x_1)_{.1} & \cdots & (\hat{w}x_1)_{.T} \\ \vdots & \ddots & \vdots \\ (\hat{w}x_p)_{.1} & \cdots & (\hat{w}x_p)_{.T} \end{pmatrix},$$

$$\mathbf{B}_\delta = [(\mathbf{1}_T \otimes \mathbf{I}_S) : (\mathbf{I}_T \otimes \mathbf{1}_S) : \mathbf{X}]^{'}\hat{\mathbf{W}} = \begin{pmatrix} \hat{w}_{11} & \cdots & 0 & \cdots & \hat{w}_{1T} & \cdots & 0 \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{w}_{S1} & \cdots & 0 & \cdots & \hat{w}_{ST} \\ \hat{w}_{11} & \cdots & \hat{w}_{S1} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \hat{w}_{1T} & \cdots & \hat{w}_{ST} \\ x_{111}\hat{w}_{11} & \cdots & x_{1S1}\hat{w}_{S1} & \cdots & x_{11T}\hat{w}_{1T} & \cdots & x_{1ST}\hat{w}_{ST} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ x_{p11}\hat{w}_{11} & \cdots & x_{pS1}\hat{w}_{S1} & \cdots & x_{p1T}\hat{w}_{1T} & \cdots & x_{pST}\hat{w}_{ST} \end{pmatrix}$$

B Supplementary Material B

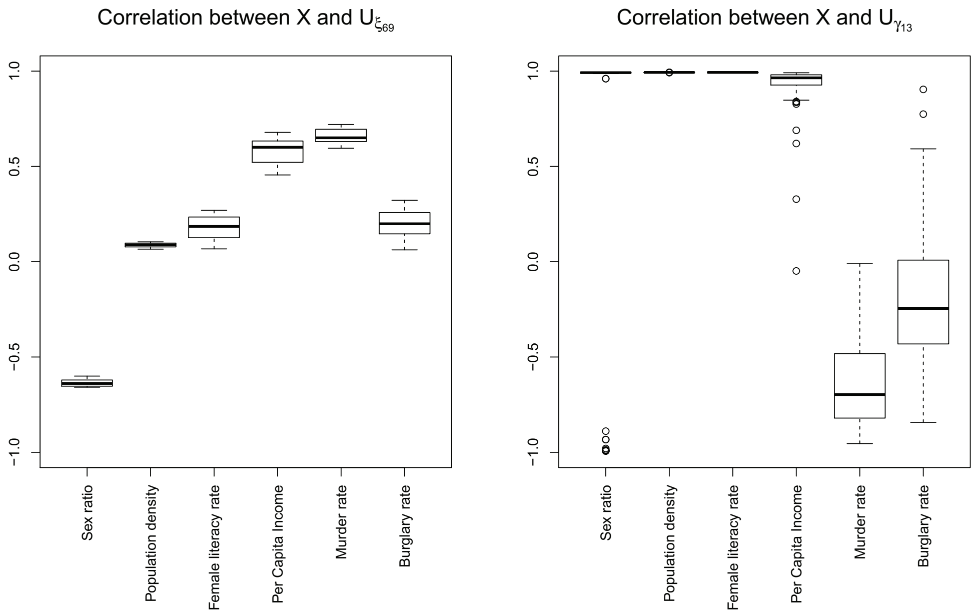

Figure B.1: Boxplots of correlations between the covariates and the spatial eigenvector $\mathbf{U}_{\xi_{69}}$ for each year (left) and correlations between the covariates and the temporal eigenvector $\mathbf{U}_{\gamma_{13}}$ for each area (right).

Figure B.1 displays boxplots of correlations between the covariates and the spatial eigenvector $\mathbf{U}_{\xi 69}$ for each year (left picture), and boxplots of correlations between the covariates and the temporal eigenvector $\mathbf{U}_{\gamma 13}$ for each area (right picture). Sex ratio, per capita income and murder rate exhibit the highest spatial correlations while the other covariates show moderate or low correlations. Regarding temporal correlations, the population-based variables (sex ratio, population density, and female literacy rate) and per capita income show the highest correlations. Population-based covariates exhibit temporal correlations close to 1 or -1 because they are only available at census years and have been linearly interpolated for the other years, and the temporal eigenvector $\mathbf{U}_{\gamma 13}$ is nearly a straight line.

Table B.1: Posterior estimates of the hyperparameters obtained using uniform priors on the positive real line for the standard deviations (INLA), and point estimates obtained with PQL

		INLA (simplified Laplace)				PQL (tol=1e-5)			
	Model	Mean	SD	$q_{0.025}$	$q_{0.975}$	Estimate	SE	$q_{0.025}$	$q_{0.975}$
σ_s^2	ST2	0.2072	0.0420	0.1391	0.3021	0.1961	0.0370	0.1236	0.2685
	ST3	0.2072	0.0420	0.1391	0.3021	0.1961	0.0370	0.1237	0.2686
	ST4	0.2349	0.0466	0.1595	0.3403	0.2218	0.0422	0.1391	0.3044
σ_t^2	ST2	0.0174	0.0086	0.0065	0.0390	0.0131	0.0056	0.0022	0.0241
	ST3	0.0174	0.0086	0.0065	0.0390	0.0131	0.0056	0.0022	0.0241
	ST4	0.0164	0.0135	0.0036	0.0519	0.0093	0.0057	0.0000	0.0204
σ_{st}^2	ST2	0.0212	0.0038	0.0147	0.0294	0.0207	0.0035	0.0139	0.0275
	ST3	0.0212	0.0038	0.0147	0.0294	0.0207	0.0035	0.0139	0.0275
	ST4	0.0222	0.0039	0.0155	0.0306	0.0213	0.0036	0.0142	0.0283

Table B.1 displays posterior estimates of the standard deviations of the random effects obtained with INLA, and point estimates of the standard deviations obtained with PQL. The estimates are rather similar in all models and in general, INLA estimates do not differ much from those obtained with PQL. Here uniform priors on the real line have been used for the standard deviations, but similar results were obtained with logGamma(1,0.00005) priors on the log-precisions.

Figure B.2 shows scatter plots of the estimated relative risks from Models ST2, ST3, and ST4 fitted with INLA (posterior means, top row) and PQL (point estimates, bottom row). For both fitting techniques, Model ST2 (spatio-temporal model with no correction for confounding) shows the same fit as Model ST3 (accounting for con-

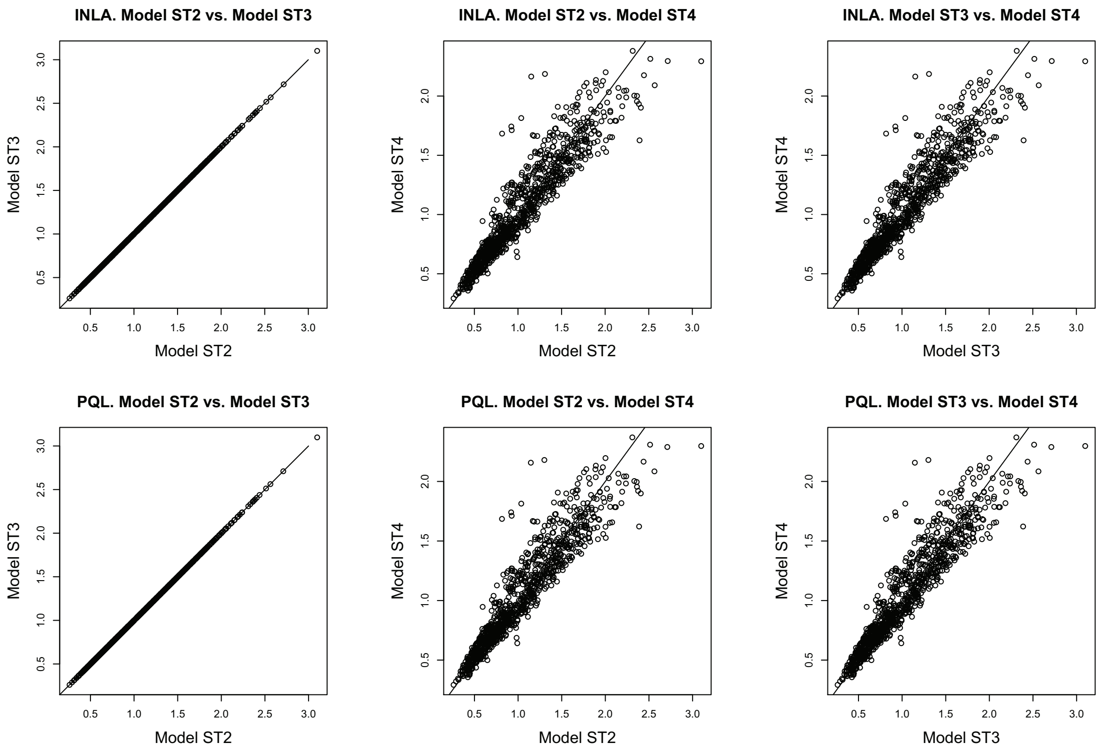

Figure B.2: Scatter plots of relative risk estimates obtained from Models ST2, ST3, and ST4. Top row: posterior means estimated with INLA; bottom row: point estimates estimated with PQL.

founding using restricted regression). However, comparing Models ST2 and ST3 with Model ST4 (accounting for confounding using constraints) shows notable differences: the two methods that deal with confounding give different fits.

Figure B.3 shows the posterior spatial patterns (top row), the posterior temporal patterns (middle row) obtained from Models ST2, ST3, and ST4 fitted with INLA (see Adin et al., 2017), and posterior spatio-temporal patterns for three districts, Agra, Balrampur, and Gautam Buddha Nagar (bottom row). While the posterior spatial patterns are quite similar for all models (top row), the posterior temporal and spatio-temporal patterns differ. The temporal patterns obtained with Models ST2 and ST3 are identical, while the temporal pattern obtained with Model ST4

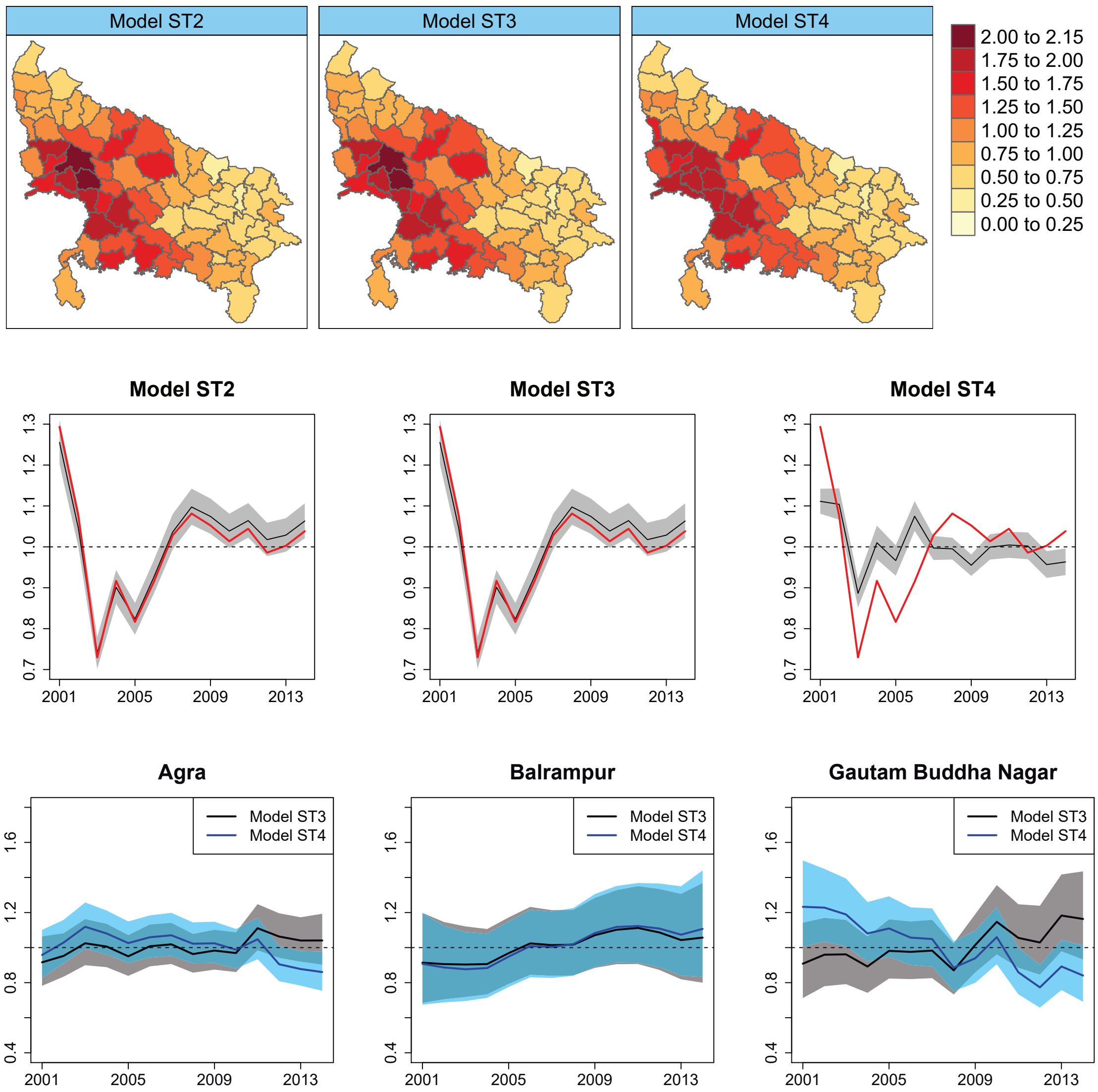

Figure B.3: Maps of posterior spatial patterns (top row) and posterior temporal patterns (middle row) obtained with models ST2, ST3, and ST4. Red lines (middle row) are the global standardized mortality ratios. Posterior spatio-temporal patterns (bottom row) obtained with Models ST3 and ST4 are shown for three districts (Agra, Balrampur and Gautam Buddha Nagar). Results are from the INLA fit.

is clearly different and does not track the global standardized mortality ratios (red line). Regarding posterior spatio-temporal patterns (space-time interactions), some

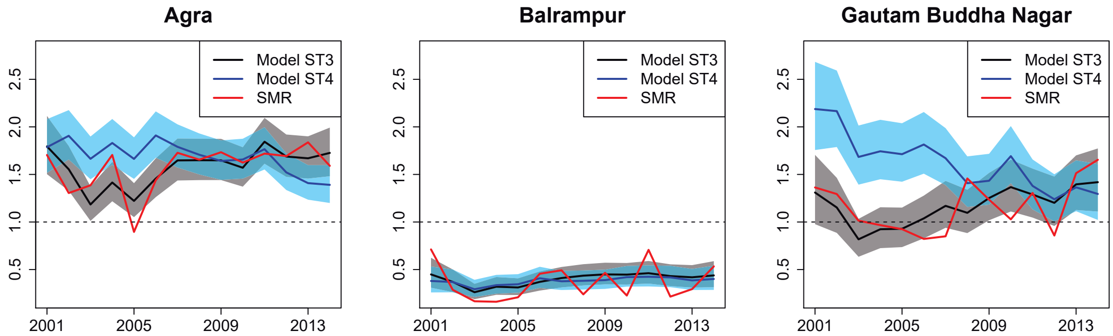

Figure B.4: Final risk estimates obtained with models ST3 and ST4 and INLA in three districts, Agra, Balrampur, and Gautam Buddha Nagar. Black lines and grey credible intervals correspond to Model ST3, blue lines and credible intervals to Model ST4. Red lines represent the crude standardized mortality ratios.

areas present mild differences between Models ST3 and ST4 (e.g., Agra) and others exhibit negligible differences (Balrampur), but some districts show striking differences (Gautam Buddha Nagar). In general, most districts have modest differences in the spatio-temporal component (not shown).

Figure B.4 displays the INLA relative risk estimates (posterior means) obtained with models ST3 and ST4 in the same three districts shown in Figure B.3, Agra, Balrampur, and Gautam Buddha Nagar. Black lines and grey credible intervals are from Model ST3, while blue lines and blue credible intervals are from Model ST4. Standardized mortality ratios are shown in red. The differences in risks between Models ST3 and ST4 in Agra and Gautam Buddha Nagar are due to both the temporal and spatio-temporal components, while the differences in Balrampur are due to the temporal component. Given that the temporal pattern is common to all districts, it seems striking that the differences in risk in Balrampur are very small in comparison to Agra and Gautam Buddha Nagar. The reason is that the risk estimate is the product of the spatial, temporal, and spatio-temporal components. In Balrampur, the spatial component is small (between 0.25 and 0.50) whereas in Agra and Gautam Buddha Nagar the spatial relative risk is greater than one. Consequently, differences in risk are softened in Balarampur and accentuated in Agra and Gautam Buddha Nagar.