Supplemental Materials for "Changes in the Geographic Pattern of Heart Disease Mortality in the United States, 1973–2010" by Casper et al.

These are the supplemental materials for "Changes in the Geographic Pattern of Heart Disease Mortality in the United States, 1973–2010". Appendix A contains the details of our Bayesian model and MCMC algorithm. Appendix B presents additional details regarding the various statistics used in the manuscript. Finally, a separate .wmv file presents maps of our expected mortality rates for each time interval.

A Hierarchical Modeling Details

A.1 Hierarchical model

Letting Y_{it} be the observed mortality rate in county i during two-year time interval, t, we will assume

$$Y_{it} \sim N\left(\mathbf{x}_{it}'\boldsymbol{\beta} + Z_{it}, \tau_{it}^2\right) \tag{1}$$

where $\tau_{it}^2 = \tau^2/n_{it}$. Similarly,

$$\mathbf{Y}_i \sim N\left(X_i\boldsymbol{\beta} + \mathbf{Z}_i, \Sigma_i\right) \;\text{ and }\; \mathbf{Y} \sim N\left(X\boldsymbol{\beta} + \mathbf{Z}, \Sigma\right),$$

where Σ_i is a diagonal matrix with τ_{it}^2 on the diagonal and Σ is a block diagonal matrix of the Σ_i. This model is similar to that of Quick et al. [1], where the authors analyzed monthly asthma hospitalization rates in California counties.

Our random effects, $\mathbf{Z}$, will be modeled as arising from a multivariate CAR [conditional autoregressive; e.g., see 2, 3] model. That is, we assume

$$\begin{aligned}
&\pi(\mathbf{Z}_i \,|\, \mathbf{Z}_{(i)}, \rho, \sigma^2) \propto \exp\left[-\frac{1}{2}\sum_{j\sim i}\frac{(\mathbf{Z}_i - \mathbf{Z}_j)'R(\rho)^{-1}(\mathbf{Z}_i - \mathbf{Z}_j)}{\sigma^2}\right] \\
&\implies \mathbf{Z}_i \,|\, \mathbf{Z}_{(i)}, \rho, \sigma^2 \sim N\left(\frac{1}{m_i}\sum_{j\sim i}\mathbf{Z}_j, \frac{\sigma^2}{m_i}R(\rho)\right),
\end{aligned} \tag{2}$$

where $R(\rho)$ is our temporal correlation matrix, σ^2 is our spatiotemporal variance parameter, $j \sim i$ denotes that counties i and j are neighbors, and m_i is the number of neighbors for the ith county. We specify $R(\rho)$ as a autoregressive (AR(1)) model with correlation parameter

ρ; i.e., $\mathrm{Cor}\left(Z_{it}, Z_{it'}\right) = \rho^{|t'-t|}$ and thus

$$R(\rho) = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N_t-1} \\ \rho & 1 & \rho & \dots & \rho^{N_t-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{N_t-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{N_t-1} & \rho^{N_t-2} & \rho^{N_t-3} & \dots & 1 \end{bmatrix}. \tag{3}$$

Our remaining parameters will be assigned the following prior distributions:

$$\begin{aligned} \pi(\boldsymbol{\beta}) &\propto 1 \\ \pi(\tau) &\propto 1 \\ \sigma^2 &\sim IG\left(a_\sigma, b_\sigma\right) \\ \rho &\sim Beta\left(a_\rho, b_\rho\right) \end{aligned}$$

i.e., we assume flat priors for $\boldsymbol{\beta}$ and τ (not τ^2), an inverse gamma prior for σ^2, and a beta prior for ρ. Putting these pieces together, our full hierarchical model is as follows:

$$\begin{aligned} \pi\left(\boldsymbol{\beta}, \mathbf{Z}, \sigma^2, \rho, \tau^2 \,|\, \mathbf{Y}\right) \propto & N\left(\mathbf{Y} \,|\, X\boldsymbol{\beta} + \mathbf{Z}, \Sigma_Y\right) \times MCAR\left(\mathbf{Z} \,|\, \rho\right) \\ & \times IG(\sigma^2 \,|\, 2, 1) \times Beta(\rho \,|\, 9, 1) \times \pi\left(\tau^2\right), \end{aligned} \tag{4}$$

where Σ_Y is a diagonal matrix with elements τ_{it}^2, X is the $(N_s N_t \times p)$ matrix of covariates, $\mathbf{Z} = \left(\mathbf{Z}_1', \dots, \mathbf{Z}_{N_s}'\right)'$, $MCAR\left(\mathbf{Z} \,|\, \rho\right)$ denotes the joint distribution induced by (2), and $\pi\left(\tau^2\right)$ is the density for τ^2 which corresponds to a flat prior for τ (equivalent to an improper $IG(-1/2, 0)$).

A.2 MCMC Algorithm

To fit the model in (4), we follow Section 3.4 of the text by Carlin and Louis [4] to construct our MCMC algorithm. During the ℓ-th iteration of our algorithm, we wish to sample from the *full conditional* distribution of each of our model parameters. That is, we proceed as follows:

1. Initialize all model parameters (iteration $\ell = 0$): $\boldsymbol{\beta}^{(0)}, \mathbf{Z}^{(0)}, \left(\sigma^2\right)^{(0)}, \rho^{(0)}, \left(\tau^2\right)^{(0)}$

2. Let $\ell = \ell + 1$

3. During iteration ℓ:

   (a) Draw $\boldsymbol{\beta}^{(\ell)}$ from $\pi\left(\boldsymbol{\beta} \,|\, \mathbf{Y}, \mathbf{Z}^{(\ell-1)}, \left(\sigma^2\right)^{(\ell-1)}, \rho^{(\ell-1)}, \left(\tau^2\right)^{(\ell-1)}\right)$

   (b) Draw $\mathbf{Z}^{(\ell)}$ from $\pi\left(\mathbf{Z} \,|\, \mathbf{Y}, \boldsymbol{\beta}^{(\ell)}, \left(\sigma^2\right)^{(\ell-1)}, \rho^{(\ell-1)}, \left(\tau^2\right)^{(\ell-1)}\right)$

   (c) Draw $\left(\sigma^2\right)^{(\ell)}$ from $\pi\left(\sigma^2 \,|\, \mathbf{Y}, \boldsymbol{\beta}^{(\ell)}, \mathbf{Z}^{(\ell)}, \rho^{(\ell-1)}, \left(\tau^2\right)^{(\ell-1)}\right)$

   (d) Draw $\rho^{(\ell)}$ from $\pi\left(\rho \,|\, \mathbf{Y}, \boldsymbol{\beta}^{(\ell)}, \mathbf{Z}^{(\ell)}, \left(\sigma^2\right)^{(\ell)}, \left(\tau^2\right)^{(\ell-1)}\right)$

(e) Draw $(\tau^2)^{(\ell)}$ from $\pi\left(\tau^2 \,|\, \mathbf{Y}, \boldsymbol{\beta}^{(\ell)}, \mathbf{Z}^{(\ell)}, (\sigma^2)^{(\ell)}, \rho^{(\ell)}\right)$

4. Repeat steps 2 and 3 until convergence is achieved and until a sufficient number of post-convergence samples have been obtained.

This algorithm has been coded in the `R` programming language [5], and the full conditional distributions for each of our model parameters are described in the subsequent subsections. For the sake of brevity, full conditional distributions will be written as $\pi(\theta \,|\, \cdot)$ for each parameter, θ.

A.2.1 Full conditional for $\boldsymbol{\beta}$

$$
\begin{aligned}
\pi(\boldsymbol{\beta} \,|\, \cdot) &\propto \prod_i \prod_t N\left(Y_{it} \,|\, \mathbf{x}_{it}'\boldsymbol{\beta} + Z_{it}, \sigma_{it}^2\right) \\
&\propto \prod_i \prod_t \exp\left[-\frac{1}{2}\frac{(Y_{it} - \mathbf{x}_{it}'\boldsymbol{\beta} - Z_{it})^2}{\sigma_{it}^2}\right] \\
&\propto \exp\left[-\frac{1}{2}\sum_i \sum_t \frac{(Y_{it} - \mathbf{x}_{it}'\boldsymbol{\beta} - Z_{it})^2}{\sigma_{it}^2}\right] \\
&\propto \ldots \\
&\propto \exp\left[-\frac{1}{2}\left\{\boldsymbol{\beta}'\left(\sum_i \sum_t \mathbf{x}_{it}'\mathbf{x}_{it}/\tau_{it}^2\right)\boldsymbol{\beta} - 2\boldsymbol{\beta}'\left(\sum_i \sum_t \mathbf{x}_{it}'\left[Y_{it} - Z_{it}\right]/\tau_{it}^2\right)\right\}\right],
\end{aligned}
$$

yielding

$$
\boldsymbol{\beta} \,|\, \cdot \sim N\left(\left[\sum_i \sum_t \mathbf{x}_{it}'\mathbf{x}_{it}/\tau_{it}^2\right]^{-1}\left[\sum_i \sum_t \mathbf{x}_{it}'\left[Y_{it} - Z_{it}\right]/\tau_{it}^2\right], \left[\sum_i \sum_t \mathbf{x}_{it}'\mathbf{x}_{it}/\tau_{it}^2\right]^{-1}\right). \quad (5)
$$

This can be simplied; e.g., we can write

$$
\sum \sum \mathbf{x}_{it}'\mathbf{x}_{it}/\tau_{it}^2 = \left(\sum \sum n_{it}\mathbf{x}_{it}'\mathbf{x}_{it}\right)/\tau^2,
$$

where $(\sum \sum n_{it}\mathbf{x}_{it}'\mathbf{x}_{it})$ is constant, and thus only needs to be computed once.

A.2.2 Full conditional for $\mathbf{Z}_i$

$$
\begin{aligned}
\pi(\mathbf{Z}_i \,|\, \cdot) &\propto N\left(\mathbf{Y}_i \,|\, X_i\boldsymbol{\beta} + \mathbf{Z}_i, \Sigma_i\right) \times N\left(\mathbf{Z}_i \,|\, \boldsymbol{\mu}_i, K_i\right) \\
&\propto \exp\left[-\frac{1}{2}\left\{(\mathbf{Y}_i - X_i\boldsymbol{\beta} - \mathbf{Z}_i)'\Sigma_i^{-1}(\mathbf{Y}_i - X_i\boldsymbol{\beta} - \mathbf{Z}_i) + (\mathbf{Z}_i - \boldsymbol{\mu}_i)'K_i^{-1}(\mathbf{Z}_i - \boldsymbol{\mu}_i)\right\}\right] \\
&\propto \ldots \\
&\propto \exp\left[-\frac{1}{2}\left\{\mathbf{Z}_i'\left(\Sigma_i^{-1} + K_i^{-1}\right)\mathbf{Z}_i - 2\mathbf{Z}_i'\left(\Sigma_i^{-1}\left[Y_i - X_i\boldsymbol{\beta}\right] + K_i^{-1}\boldsymbol{\mu}_i\right)\right\}\right],
\end{aligned}
$$

where $\boldsymbol{\mu}_i$ and K_i are the conditional mean and variance given in (2). This yields

$$\mathbf{Z}_i \,|\, \cdot \sim N\left(\boldsymbol{\mu}_{Z_i\,|\,\cdot}, \Sigma_{Z_i\,|\,\cdot}\right), \tag{6}$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{Z_i\,|\,\cdot} &= \left[\Sigma_i^{-1} + K_i^{-1}\right]^{-1}\left[\Sigma_i^{-1}\left[Y_i - X_i\boldsymbol{\beta}\right] + K_i^{-1}\boldsymbol{\mu}_i\right] \\
&= \left[\Sigma_i^{-1} + \frac{m_i}{\sigma^2}R(\rho)^{-1}\right]^{-1}\left[\Sigma_i^{-1}\left[Y_i - X_i\boldsymbol{\beta}\right] + \frac{1}{\sigma^2}R(\rho)^{-1}\sum_{j\sim i}\mathbf{Z}_j\right] \\
\Sigma_{Z_i\,|\,\cdot} &= \left[\Sigma_i^{-1} + K_i^{-1}\right]^{-1} = \left[\Sigma_i^{-1} + \frac{m_i}{\sigma^2}R(\rho)^{-1}\right]^{-1}.
\end{aligned}$$

While this requires looping through $i = 1, \ldots, N_s$, where N_s is large, we are only required to invert an $N_t \times N_t$ matrix, where N_t is relatively small. Finally, due to the impropriety of the CAR model used, we impose a sum-to-zero constraint, $\sum_i \sum_t Z_{it} = 0$, which is implemented each iteration.

A.2.3 Full conditional for σ^2

First, note that if we wish to write out the prior for $\mathbf{Z}$ (as opposed to the conditional prior for $\mathbf{Z}_i$ in (2)), it would take the form

$$\pi\left(\mathbf{Z}\,|\,\rho, \sigma^2\right) \propto \left(\sigma^2\right)^{-(N_s-1)N_t/2}\exp\left[-\frac{1}{2\sigma^2}\mathbf{Z}'\left\{(D-W)\otimes R(\rho)^{-1}\right\}\mathbf{Z}\right], \tag{7}$$

where D is a diagonal matrix with elements m_i, W is an adjacency matrix with elements

$$w_{ij} = \begin{cases} 1 & \text{if } i \sim j \\ 0 & \text{if } i \not\sim j \end{cases},$$

and $\otimes$ denotes the Kronecker product. While we will use this for deriving $\pi\left(\sigma^2\,|\,\cdot\right)$, we haven't used this expression previously because $(D - W)$ is an $N_s \times N_s$ matrix (i.e., it's really big), and thus can be burdensome to manipulate (or even store in a computer's RAM). As such, we will need to do some additional algebra here to avoid storing this. For instance,

$$
\begin{aligned}
\mathbf{Z}'\left\{(D-W)\otimes R(\rho)^{-1}\right\}\mathbf{Z} =& \mathbf{Z}'\left\{D\otimes R(\rho)^{-1}\right\}\mathbf{Z} - \mathbf{Z}'\left\{W\otimes R(\rho)^{-1}\right\}\mathbf{Z} \\
&= \begin{pmatrix}\mathbf{Z}_1' & \ldots & \mathbf{Z}_{N_s}'\end{pmatrix}\left[m_i R(\rho)^{-1}\right]_{ii}\begin{pmatrix}\mathbf{Z}_1\\ \vdots \\ \mathbf{Z}_{N_s}\end{pmatrix} \\
&\quad - \begin{pmatrix}\mathbf{Z}_1' & \ldots & \mathbf{Z}_{N_s}'\end{pmatrix}\left[w_{ij} R(\rho)^{-1}\right]_{ij}\begin{pmatrix}\mathbf{Z}_1\\ \vdots \\ \mathbf{Z}_{N_s}\end{pmatrix} \\
&= \sum_i m_i \mathbf{Z}_i' R(\rho)^{-1}\mathbf{Z}_i - \sum_i \mathbf{Z}_i' \sum_{j\sim i} R(\rho)^{-1}\mathbf{Z}_j \\
&= \sum_i \mathbf{Z}_i' R(\rho)^{-1}\left[m_i\mathbf{Z}_i - \sum_{j\sim i}\mathbf{Z}_j\right]. \qquad (8)
\end{aligned}
$$

Using this, we can now find the full conditional distribution for σ^2:

$$
\begin{aligned}
\pi\left(\sigma^2 \,|\, \cdot\right) \propto& IG\left(\sigma^2 \,|\, a_\sigma, b_\sigma\right) \times \pi\left(\mathbf{Z} \,|\, \rho, \sigma^2\right) \\
&\propto \left(\sigma^2\right)^{-a_\sigma - 1}\exp\left[-\frac{b}{\sigma^2}\right]\left(\sigma^2\right)^{-(N_s-1)N_t/2}\exp\left[-\frac{1}{2\sigma^2}\mathbf{Z}'\left\{(D-W)\otimes R(\rho)^{-1}\right\}\mathbf{Z}\right] \\
&\propto \left(\sigma^2\right)^{-([N_s-1]N_t/2+a_\sigma)-1}\exp\left[-\frac{1}{\sigma^2}\left(b_\sigma + \frac{\sum_i \mathbf{Z}_i' R(\rho)^{-1}\left[m_i\mathbf{Z}_i - \sum_{j\sim i}\mathbf{Z}_j\right]}{2}\right)\right],
\end{aligned}
\qquad (9)
$$

which yields

$$
\sigma^2 \,|\, \cdot \sim IG\left(\left(N_s - 1\right)N_t/2 + a_\sigma, \left\{b_\sigma + \frac{\sum_i \mathbf{Z}_i' R(\rho)^{-1}\left[m_i\mathbf{Z}_i - \sum_{j\sim i}\mathbf{Z}_j\right]}{2}\right\}\right). \qquad (10)
$$

A.2.4 Full conditional for τ^2

$$
\begin{aligned}
\pi\left(\tau^2 \,|\, \cdot\right) \propto& N\left(\mathbf{Y} \,|\, X\boldsymbol{\beta} + \mathbf{Z}, \Sigma\right) \times \pi\left(\tau^2\right) \\
&\propto \left(\tau^2\right)^{-N_sN_t/2}\exp\left[-\frac{1}{2}\left(\mathbf{Y} - X\boldsymbol{\beta} - \mathbf{Z}\right)'\Sigma^{-1}\left(\mathbf{Y} - X\boldsymbol{\beta} - \mathbf{Z}\right)\right]\left(\tau^2\right)^{-1/2} \\
&\propto \left(\tau^2\right)^{-(N_sN_t-1)/2-1}\exp\left[-\frac{1}{\tau^2}\sum_i\sum_t \frac{\left(Y_{it} - \mathbf{x}_{it}\boldsymbol{\beta} - Z_{it}\right)^2 n_{it}}{2}\right], \qquad (11)
\end{aligned}
$$

which is the form of an $IG\left(\left[N_sN_t - 1\right]/2, \sum_i\sum_t \frac{(Y_{it} - \mathbf{x}_{it}\boldsymbol{\beta} - Z_{it})^2 n_{it}}{2}\right)$. Note, a uniform prior on τ implies $\pi\left(\tau^2\right) \propto \left(\tau^2\right)^{-1/2}$.

A.2.5 Metropolis update for ρ

As there is not a conjugate prior for ρ, updates for this parameter require a little more work. First, we will let $\phi = \text{logit}(\rho) = \log(\rho/(1-\rho))$, and define $\text{expit}(\cdot)$ such that $\rho = \text{expit}(\phi)$; this transformation allows us to work with a parameter with domain $(-\infty, \infty)$, rather than $(0, 1)$. Later, we will require the Jacobian of this transformation,

$$J_\phi(\phi) = \frac{d\,\text{expit}(\phi)}{d\phi} = \frac{1}{1+\exp(\phi)}\frac{\exp(\phi)}{1+\exp(\phi)} = \rho(1-\rho) = J_\rho(\rho)\,. \tag{12}$$

For our Metropolis updates, we will use a candidate density of the form

$$Q_\phi\left(\phi^{(*)} \mid \phi^{(t-1)}\right) = N\left(\phi^{(t-1)}, \gamma^2\right), \tag{13}$$

where $\phi^{(*)}$ (and similarly $\rho^{(*)} = \text{expit}\left(\phi^{(*)}\right)$) is a candidate value, $\phi^{(t-1)}$ ($\rho^{(t-1)}$) is the value of ϕ from the $(t-1)$-th iteration of our sampler, and γ^2 is the variance of our proposal density, which we can tune this in order to achieve a desired acceptance rate.

In order to decide whether we want to accept a candidate value $\rho^{(*)}$, we need to compute the acceptance ratio, r, which can be expressed as:

$$\begin{aligned} r =& \frac{\pi\left(\mathbf{Z} \mid \rho^{(*)}, \sigma^2\right) \times \pi\left(\rho^{(*)}\right) \times J\left(\rho^{(*)}\right)}{\pi\left(\mathbf{Z} \mid \rho^{(t-1)}, \sigma^2\right) \times \pi\left(\rho^{(t-1)}\right) \times J\left(\rho^{(t-1)}\right)} \\ =& \left\{\left(\frac{|R\left(\rho^{(*)}\right)|}{|R\left(\rho^{(t-1)}\right)|}\right)^{-\frac{(N_s-1)}{2}}\right\} \times \left\{\left(\frac{\rho^{(*)}}{\rho^{(t-1)}}\right)^{a_\rho}\left(\frac{1-\rho^{(*)}}{1-\rho^{(t-1)}}\right)^{b_\rho}\right\} \\ & \times \left\{\exp\left[-\frac{1}{2\sigma^2}\left(\mathbf{Z}'\left\{(D-W)\otimes\left[R\left(\rho^{(*)}\right)^{-1} - R\left(\rho^{(t-1)}\right)^{-1}\right]\right\}\mathbf{Z}\right)\right]\right\} \\ =& r_1 \times r_2 \times r_3. \end{aligned} \tag{14}$$

In order to evaluate r_1 and r_3, we should first note two properties for $R(\rho)$. First and foremost,

$$R(\rho)^{-1} = \begin{bmatrix} \frac{1}{1-\rho^2} & -\frac{\rho}{1-\rho^2} & 0 & \cdots & \cdots & 0 \\ -\frac{\rho}{1-\rho^2} & \frac{1+\rho^2}{1-\rho^2} & -\frac{\rho}{1-\rho^2} & \ddots & \cdots & 0 \\ 0 & -\frac{\rho}{1-\rho^2} & \frac{1+\rho^2}{1-\rho^2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \frac{1+\rho^2}{1-\rho^2} & -\frac{\rho}{1-\rho^2} \\ 0 & \cdots & \cdots & 0 & -\frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} \end{bmatrix}, \tag{15}$$

a tridiagonal matrix. Using the properties of tridiagonal matrices, it can be shown that

$$|R(\rho)^{-1}| = 1/\left(1-\rho^2\right)^{N_t-1} \implies |R(\rho)| = \left(1-\rho^2\right)^{N_t-1},$$

thus avoiding the computational burdensome task of finding the determinant of a matrix.

Using this, we find

$$r_1 = \left(\frac{|R\left(\rho^{(*)}\right)|}{|R\left(\rho^{(t-1)}\right)|} \right)^{-\frac{(N_s-1)}{2}} = \left(\frac{1-\left(\rho^{(*)}\right)^2}{1-\left(\rho^{(t-1)}\right)^2} \right)^{-\frac{(N_s-1)(N_t-1)}{2}}. \tag{16}$$

We can then efficiently construct $R(\rho)^{-1}$ and use the algebra in (8) to compute

$$\begin{aligned} r_3 &= \left\{ \exp\left[-\frac{1}{2\sigma^2} \left(\mathbf{Z}' \left\{ (D-W) \otimes \left[R\left(\rho^{(*)}\right)^{-1} - R\left(\rho^{(t-1)}\right)^{-1} \right] \right\} \mathbf{Z} \right) \right] \right\} \\ &= \exp\left[-\frac{1}{2\sigma^2} \sum_i \mathbf{Z}_i' \left\{ R\left(\rho^{(*)}\right)^{-1} - R\left(\rho^{(t-1)}\right)^{-1} \right\} \left(m_i \mathbf{Z}_i - \sum_{j\sim i} \mathbf{Z}_j \right) \right]. \end{aligned} \tag{17}$$

Putting these pieces together, we can compute our acceptance ratio

$$\begin{aligned} r = & \left(\frac{1-\left(\rho^{(*)}\right)^2}{1-\left(\rho^{(t-1)}\right)^2} \right)^{-\frac{(N_s-1)(N_t-1)}{2}} \times \left(\frac{\rho^{(*)}}{\rho^{(t-1)}} \right)^{a_\rho} \left(\frac{1-\rho^{(*)}}{1-\rho^{(t-1)}} \right)^{b_\rho} \\ & \times \exp\left[-\frac{1}{2\sigma^2} \sum_i \mathbf{Z}_i' \left\{ R\left(\rho^{(*)}\right)^{-1} - R\left(\rho^{(t-1)}\right)^{-1} \right\} \left(m_i \mathbf{Z}_i - \sum_{j\sim i} \mathbf{Z}_j \right) \right]. \end{aligned}$$

We then accept a move to $\rho^{(*)}$ with probability $\min(r, 1)$.

B Summary Statistics

All analyses presented in the manuscript were conducted using the posterior distribution of the modeled mortality rate, $\widehat{Y}_{it} = \beta_t + Z_{it}$, which offers further stability by borrowing strength across both space and time. Note that while we obtain samples $\widehat{Y}_{it}$, denoted as $\widehat{Y}_{it}^{(\ell)}$, the ℓ superscript is omitted from this and subsequent expressions for the sake of notation.

For ease of visualizing spatial patterns, we present maps of both the expected mortality rates and the local indicators of spatial association (LISA) statistics [6] for the beginning and end of the study period. The LISA statistic is a local version correlate of the Moran's I identifying the degree to which counties with higher (or lower) than expected death rates tend to cluster with neighboring counties which also have higher (lower) than expected death rates. Thus the LISA statistics highlight patterns of local and regional spatial clustering in the heart disease death rates. The LISA statistic is computed as:

$$I_{it} = \frac{(\widehat{Y}_{it} - \bar{Y}_t) \sum_{j\sim i} (\widehat{Y}_{jt} - \bar{Y}_t)/m_i}{\sum_i (\widehat{Y}_{it} - \bar{Y}_t)^2 / N_s},$$

where $\bar{Y}_t = \sum_{i=1}^{N_s} \widehat{Y}_{it}/N_s$.

In addition to estimating the percent decline for each county, we also estimate the posterior distribution for the coefficient of variation (CV) for each time interval. The coefficient of variation in time interval t can be expressed as

$$CV_t = \frac{\sum_i (\widehat{Y}_{it} - \bar{Y}_t)^2/(N_s - 1)}{\bar{Y}_t}.$$

Changes in the coefficient of variation measure the changing magnitude of geographic disparity among the counties [7].

References

[1] H. Quick, S. Banerjee, and B. P. Carlin. Modeling temporal gradients in regionally aggregated California asthma hospitalization data. *Annals of Applied Statistics*, 7:154–176, 2013.

[2] J. Besag, J. York, and A. Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43:1–59, 1991.

[3] A. E. Gelfand and P. Vounatsou. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics*, 4:11–25, 2003.

[4] B. P. Carlin and T. A. Louis. *Bayesian Methods for Data Analysis*. Chapman & Hall/CRC, 2008.

[5] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL `http://www.R-project.org/`.

[6] L. Anselin. Local indicators of spatial association. *Geographical Analysis*, 27:93–115, 1995.

[7] C. W. Harper, J. M. Blair, P. A. Fay, A. K. Knapp, and J. D. Carlisle. Increased rainfall variability and reduced rainfall amount decreases soil CO_2 flux in a grassland ecosystem. *Global Change Biology*, 11:322–334, 2005.