Web-Based Supporting Material for "A marginalized zero-inflated negative binomial model for spatial data: modeling COVID-19 deaths in Georgia" by Fedelis Mutiso, Hong Li, John L. Pearce, Sara E. Benjamin-Neelon, Noel T. Mueller, and Brian Neelon

Appendix A: Social Vulnerability Index Components Variables

Table S1: Social Vulnerability Index themes and variables. The 15 SVI variables are obtained from the 2016-2020 American Community Survey of the U.S.

SVI Theme	Variables
Socioeconomic Status	Percentage below poverty
	Percentage unemployed
	Per capita income
	Percentage with no high school diploma
Household Composition & Disability	Percentage age 65 and older
	Percentage age 17 or younger
	Percentage age 5 years or older with a disability
	Percentage of single-parent households
Minority Status & Language	Percentage minority
	Percentage who speaks english "less than well"
Housing Type & Transportation	Percentage of multi-unit structures
	Percentage of mobile homes
	Percentage crowding
	Percentage having no vehicle
	Percentage of group quarters

Appendix B: MCMC Algorithm

1. Update the latent at-risk indicators $z_{11}, \ldots, z_{1T}, \ldots, z_{n1}, \ldots, z_{nT}$ for the binary component. The generic form of the spatiotemporal marginalized ZINB model is similar to the model given in equation (1) of the main manuscript. That is $y_{it} \sim (1 - \psi_{it})\mathbb{1}_{(}z_{it} = 0 \wedge y_{it} = 0) + \psi_{it} p(y_{it}; r, \nu_{it})\mathbb{1}_{z_{it}=1}$, where $z_{it} = 1$ (county i is "at-risk" on day t) with prior probability $\psi_{it} = \exp(\boldsymbol{w}_{it}^T\boldsymbol{\gamma} + \phi_{1i} + f_1(t))/[1 + \exp(\boldsymbol{w}_{it}^T\boldsymbol{\gamma} + \phi_{1i} + f_1(t))]$ and $z_{it} = 0$ otherwise. Now, if $y_{it} > 0$, then county i is in the at-risk class on day t but otherwise on the structural class implying that $y_{it} = 0$. Thus, given $y_{it} > 0$, we set $z_{it} = 1$. Conversely, if $y_{it} = 0$, then we observe either a structural zero (implying $z_{it} = 0$) or an at-risk zero (implying $z_{it} = 1$). In this case, we draw z_{it} from a Bernoulli distribution with probability θ_{it}, where θ_{it} is given by

$$\begin{aligned}\theta_{it} = \Pr(z_{it}|y_{it} = 0, \text{rest}) &= \frac{\Pr(z_{it} = 1 \cap y_{it} = 0)}{\Pr(y_{it} = 0)} \\ &= \frac{\Pr(y_{it} = 0|z_{it} = 1)\Pr(z_{it} = 1)}{\Pr(y_{it} = 0|z_{it} = 1)\Pr(z_{it=1}) + \Pr(y_{it} = 0|z_{it} = 0)\Pr(z_{it=0})},\end{aligned}$$

where $\Pr(y_{it} = 0|z_{it} = 1)$ is the probability of observing a zero under the negative binomial model (at-risk zero) which is $(1 - q_{it})^r$, where $q_{it} = r\psi_{it}/(\nu_{it} + r\psi_{it})$ and ν_{it} is defined in equation (7) of the main manuscript. Additionally, $\Pr(y_{it} = 0|z_{it} = 0) = 1$, since, in this case, county i belongs to the structural class on day t and hence $y_{it} = 0$ with probability 1. Thus, we have

$$\begin{aligned}\theta_{it} &= \frac{\Pr(y_{it} = 0|z_{it} = 1)\Pr(z_{it} = 1)}{\Pr(y_{it} = 0|z_{it} = 1)\Pr(z_{it=1}) + \Pr(y_{it} = 0|z_{it} = 0)\Pr(z_{it=0})} \\ &= \frac{(1 - q_{it})^r\psi_{it}}{(1 - q_{it})^r\psi_{it} + 1.(1 - \psi_{it})} \\ &= \frac{\psi_{it}(1 - q_{it})^r}{1 - \psi_{it}[1 - (1 - q_{it})^r]}\end{aligned}$$

2. Update ω_{it}: From Polson et al. (2013), the conditional distribution of ω_{it} given $\boldsymbol{\theta}_1$ and $\boldsymbol{\phi}_1$ is $p(\omega_{it}|\boldsymbol{\theta}_1, \boldsymbol{\phi}_1) \stackrel{d}{=} \text{PG}(1, \boldsymbol{l}_{it}^T\boldsymbol{\theta}_1 + \phi_{1i})$ where $\boldsymbol{l}_{it}$ is a $(p + K) \times 1$ vector of fixed and spline effects for the binary part. Thus, draw ω_{it} $(i = 1, \ldots, n; \quad t = 1, \ldots, T)$ independently from $\text{PG}(1, \boldsymbol{l}_{it}^T\boldsymbol{\theta}_1 + \phi_{1i})$ using the accept-reject algorithm described in Polson et al. (2013), which can be implemented using the R package `BayesLogit`.

3. Update $\boldsymbol{\theta}_1$: Given $\boldsymbol{\omega} = (\omega_{11}, \ldots, \omega_{1T}, \ldots, \omega_{n1}, \ldots, \omega_{nT})^T$, $\boldsymbol{\Phi}_1 = (\phi_{11}, \ldots, \phi_{1n})^T$, and

the at-risk indicators $\boldsymbol{z} = (z_{11}, \ldots, z_{1T}, \ldots, z_{n1}, \ldots, z_{nT})^T$, the full conditional for $\boldsymbol{\theta}_1$ is given by

$$
\begin{aligned}
p(\boldsymbol{\theta}_1 | z_{it}, \omega_{it}, \phi_{1i}) &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} p(z_{it} | \boldsymbol{\theta}_1, \omega_{it}, \phi_{1i}) p(\omega_{it} | \boldsymbol{\theta}_1, \phi_{1i}) \\
&\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \frac{\exp(\boldsymbol{l}_{it}^T \boldsymbol{\theta}_1 + \phi_{1i})^{z_{it}}}{1 + \exp(\boldsymbol{l}_{it}^T \boldsymbol{\theta}_1 + \phi_{1i})} p(\omega_{it} | \boldsymbol{\theta}_1, \phi_{1i}) \qquad (1)
\end{aligned}
$$

Next, we make use of the following two properties of PG density detailed in Polson et al. (2013). First, for $a \in \mathfrak{R}$ and $\eta \in \mathfrak{R}$, it follows that

$$
\frac{(\mathrm{e}^{\eta})^a}{(1 + \mathrm{e}^{\eta})^b} = 2^{-b} \mathrm{e}^{\kappa \eta} \int_0^{\infty} \mathrm{e}^{-\omega \eta^2 / 2} p(\omega | b, 0) \mathrm{d}\omega
$$

where $\kappa = a - b/2$ and $p(\omega|b,0)$ denotes a $\mathrm{PG}(b,0)$ density. Second, the conditional distribution $p(\omega|b,c) \sim \mathrm{PG}(b,c)$ follows from an "exponential tilting" of the $\mathrm{PG}(b,0)$ density:

$$
\begin{aligned}
p(\omega | b, c) &= \frac{\exp(-c^2 \omega / 2) p(\omega | b, 0)}{\mathrm{E}_{\omega}[\exp(-c^2 \omega / 2)]} \\
&= \frac{\exp(-c^2 \omega / 2) p(\omega | b, 0)}{\int_0^{\infty} \mathrm{e}^{-c^2 \omega / 2} p(\omega | b, 0) \mathrm{d}\omega}
\end{aligned}
$$

Applying the above properties, we get

$$
\begin{aligned}
p(\boldsymbol{\theta}_1 | z_{it}, \omega_{it}, \phi_{1i}) &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \left[\mathrm{e}^{\kappa_{it} \eta_{it}} \int_0^{\infty} \mathrm{e}^{-\omega_{it} \eta_{it}^2 / 2} p(\omega_{it} | 1, 0) \mathrm{d}\omega_{it} \right] \\
&\times \frac{\mathrm{e}^{-\eta_{it}^2 \omega_{it} / 2} p(\omega_{it} | 1, 0)}{\int_0^{\infty} \mathrm{e}^{-\eta_{it}^2 \omega_{it} / 2} p(\omega_{it} | 1, 0) \mathrm{d}\omega_{it}}
\end{aligned}
$$

where from equation (1) and property 1, $\kappa_{it} = z_{it} - 1/2$. Continuing, we have

$$\begin{aligned} p(\boldsymbol{\theta}_1|z_{it}, \omega_{it}, \phi_{1i}) &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \mathrm{e}^{\kappa_{it}\eta_{it}} \mathrm{e}^{-\omega_{it}\eta_{it}^2/2} p(\omega_{it}|1,0) \\ &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \mathrm{e}^{\kappa_{it}\eta_{it}} \mathrm{e}^{-\omega_{it}\eta_{it}^2/2} \text{ since } p(\omega_{it}|1,0) \text{ is constant w.r.t } \boldsymbol{\theta}_1 \\ &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \mathrm{e}^{-\frac{\omega_{it}}{2}\left[\eta_{it}^2 - 2\eta_{it}\frac{\kappa_{it}}{\omega_{it}}\right]} \\ &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \mathrm{e}^{-\frac{\omega_{it}}{2}\left[\eta_{it}^2 - 2\eta_{it} z_{it}^*\right]}, \end{aligned}$$

where $z_{it}^* = \frac{\kappa_{it}}{\omega_{it}} = \frac{z_{it}-1/2}{\omega_{it}}$. Completing the square we get,

$$\begin{aligned} p(\boldsymbol{\theta}_1|z_{it}, \omega_{it}, \phi_{1i}) &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \mathrm{e}^{\left\{-\frac{\omega_{it}}{2}\left[\eta_{it}^2 - 2\eta_{it} z_{it}^* + z_{it}^{*2}\right]\right\}} \\ &\propto \pi(\boldsymbol{\theta}_1) \prod_{i=1}^{n} \prod_{t=1}^{T} \mathrm{e}^{\left\{-\frac{\omega_{it}}{2}\left(z_{it}^* - \eta_{it}\right)^2\right\}} \\ &\propto \pi(\boldsymbol{\theta}_1) \exp\left[-\frac{1}{2}(\boldsymbol{z}^* - \boldsymbol{\eta})^T \boldsymbol{\Omega} (\boldsymbol{z}^* - \boldsymbol{\eta})\right] \end{aligned}$$

where $\boldsymbol{z}^*$ is an $N \times 1$ vector with it-th element z_{it}, where $N = nT$ is the total number of observations; $\boldsymbol{\eta}$ is an $N \times 1$ mean vector with it-th element $\eta_{it} = \boldsymbol{l}_{it}^T \boldsymbol{\theta}_1 + \phi_{1i}$; $\boldsymbol{\Omega} = \mathrm{diag}(\boldsymbol{\omega})$ is an $N \times N$ diagonal matrix of PG precisions. The last expression is kernel of a $\mathrm{N}_N(\boldsymbol{\eta}, \boldsymbol{\Omega}^{-1})$ density. Thus, assuming a $\mathrm{N}_{p+K}(\boldsymbol{\theta}_0, \boldsymbol{V}_0)$ prior for $\boldsymbol{\theta}_1$ and applying standard Bayesian linear regression results, the conjugate full conditional for $\boldsymbol{\theta}_1$ given $\boldsymbol{z}^*$, $\boldsymbol{\Phi}_1$, and $\boldsymbol{\omega}$ is $\mathrm{N}_{p+K}(\boldsymbol{\mu}, \boldsymbol{V})$ where

$$\begin{aligned} \boldsymbol{V} &= \left(\boldsymbol{V}_0^{-1} + \boldsymbol{L}^T \boldsymbol{\Omega} \boldsymbol{L}\right)^{-1} \\ \boldsymbol{\mu} &= \boldsymbol{V}\left[\boldsymbol{V}_0^{-1}\boldsymbol{\theta}_0 + \boldsymbol{L}^T \boldsymbol{\Omega}(\boldsymbol{z}^* - \boldsymbol{L}^* \boldsymbol{\Phi}_1)\right], \end{aligned}$$

and $\boldsymbol{L}$ is an $N \times (p+K)$ design matrix for the fixed-effect covariates and B-spline basis functions for the binary part and

$$\underset{N\times n}{\boldsymbol{L}^*} = \begin{pmatrix} 1_1 & 0 & 0 & \dots & 0 & 0 \\ 1_2 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1_J & 0 & 0 & \dots & 0 & 0 \\ 0 & 1_1 & 0 & \dots & 0 & 0 \\ 0 & 1_2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1_J & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1_1 \\ 0 & 0 & 0 & \dots & 0 & 1_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1_J \end{pmatrix}$$

is the random effects design matrix.

4. Update of $\boldsymbol{\Phi}_1$: We update the $n \times 1$ vector of random effects for the binary component, $\boldsymbol{\Phi}_1 = (\phi_{11}, \dots, \phi_{1n})^T$, conditional on the $n \times 1$ random effect vector, $\boldsymbol{\Phi}_2 = (\phi_{21}, \dots, \phi_{2n})^T$, for the mean component. Recall that the bivariate CAR prior for the 2×1 vector of spatial effects, $\boldsymbol{\phi}_i = (\phi_{1i}, \phi_{2i})^T$, for county i is

$$\boldsymbol{\phi}_i | \boldsymbol{\phi}_{(-i)}, \boldsymbol{\Sigma} = \mathrm{N}_2 \left(\frac{1}{m_i} \sum_{l \in \partial_i} \boldsymbol{\phi}_l, \frac{1}{m_i} \boldsymbol{\Sigma} \right),$$

where m_i is the number of neighboring counties and $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{\phi_1} & \rho\sigma_{\phi_1}\sigma_{\phi_2} \\ \rho\sigma_{\phi_1}\sigma_{\phi_2} & \sigma^2_{\phi_2} \end{bmatrix}$ is the bivariate CAR scale matrix. By Brook's lemma, the joint multivariate intrinsic CAR prior for the $2n \times 1$ vector $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1^T, \boldsymbol{\Phi}_2^T)^T$ is proportional to a mean-zero, singular (i.e., rank deficient) multivariate normal density:

$$\begin{pmatrix} \boldsymbol{\Phi}_1 \\ \boldsymbol{\Phi}_2 \end{pmatrix} \propto \exp\left[-\frac{1}{2} \boldsymbol{\Phi}^T \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{Q} \right) \boldsymbol{\Phi} \right],$$

where $\boldsymbol{Q} = \boldsymbol{M} - \boldsymbol{A}$ is the $n \times n$ intrinsic CAR structure matrix of rank $n-1$; $\boldsymbol{M} = \mathrm{diag}(m_1, \dots, m_n)$ with diagonal elements equal to the number of neighbors for each spatial unit; $\boldsymbol{A}$ is an $n \times n$ adjacency matrix with $a_{ii} = 0$, $a_{il} = 1$ if spatial units i and l are neighbors, and $a_{il} = 0$ otherwise. Thus, from the properties of the multivariate (singular) normal distribution, the conditional prior for $\mathbf{\Phi}_1$ given $\mathbf{\Phi}_2$ is:

$$\begin{aligned} p(\mathbf{\Phi}_1|\mathbf{\Phi}_2, \mathbf{\Sigma}) &\propto \exp\left[-\frac{1}{2}\left(\mathbf{\Phi}_1 - \boldsymbol{\mu}_{1|2}\right)^T \mathbf{\Sigma}_{1|2}^{-1}\left(\mathbf{\Phi}_1 - \boldsymbol{\mu}_{1|2}\right)\right], \text{ where} \\ \mathbf{\Sigma}_{1|2}^{-1} &= \left[\sigma_{\phi_1}^2(1-\rho^2)\right]^{-1}\boldsymbol{Q}, \\ \boldsymbol{\mu}_{1|2} &= \rho\frac{\sigma_{\phi_1}}{\sigma_{\phi_2}}\boldsymbol{Q}^+\boldsymbol{Q}\mathbf{\Phi}_2 \to \rho\frac{\sigma_{\phi_1}}{\sigma_{\phi_2}}\mathbf{\Phi}_2 \text{ for large } n, \end{aligned}$$

and $\boldsymbol{Q}^+$ is the Moore-Penrose generalized inverse of the rank-deficient structure matrix $\boldsymbol{Q}$. The approximation in the last line follows from Corollary 2.3 in Neelon et al. (2023), which states that the expression $\boldsymbol{Q}^+\boldsymbol{Q} \to \boldsymbol{I}_n$ as $n \to \infty$, where $\boldsymbol{I}_n$ is the n-dimensional identity matrix. Hence, the conditional prior mean of $\mathbf{\Phi}_1$ is closely approximated by $\boldsymbol{\mu}_{1|2} \approx \rho\frac{\sigma_{\phi_1}}{\sigma_{\phi_2}}\mathbf{\Phi}_2$ for moderate to large n.

From step (2) above, $\boldsymbol{z}^*|\boldsymbol{\theta}_1, \mathbf{\Phi}_1 \sim \mathrm{N}_N(\boldsymbol{\eta}, \mathbf{\Omega}^{-1})$. Therefore,

$$\begin{aligned} p(\mathbf{\Phi}_1|\boldsymbol{z}^*, \boldsymbol{\theta}_1) &\propto \pi(\boldsymbol{z}^*|\boldsymbol{\theta}_1, \mathbf{\Phi}_1)\pi(\mathbf{\Phi}_1|\mathbf{\Phi}_2, \mathbf{\Sigma}_{1|2}) \\ &\propto \exp\left[\frac{-1}{2}\left(\boldsymbol{z}^* - \boldsymbol{L}\boldsymbol{\theta}_1 - \boldsymbol{L}^*\mathbf{\Phi}_1\right)^{\mathrm{T}}\mathbf{\Omega}\left(\boldsymbol{z}^* - \boldsymbol{L}\boldsymbol{\theta}_1 - \boldsymbol{L}^*\mathbf{\Phi}_1\right)\right] \\ &\times \exp\left[\frac{-1}{2}(\mathbf{\Phi}_1 - \boldsymbol{\mu}_{1|2})^{\mathrm{T}}\mathbf{\Sigma}_{1|2}^{-1}(\mathbf{\Phi}_1 - \boldsymbol{\mu}_{1|2})\right] \\ &\propto \exp\left\{\frac{-1}{2}\left[\mathbf{\Phi}_1^T\left(\boldsymbol{L}^{*T}\mathbf{\Omega}\boldsymbol{L}^* + \mathbf{\Sigma}_{1|2}^{-1}\right)\mathbf{\Phi}_1 - 2\mathbf{\Phi}_1^T\left\{\boldsymbol{L}^{*T}\mathbf{\Omega}(\boldsymbol{z}^* - \boldsymbol{L}\boldsymbol{\theta}_1) + \mathbf{\Sigma}_{1|2}^{-1}\boldsymbol{\mu}_{1|2}\right\}\right]\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[\mathbf{\Phi}_1^T\boldsymbol{V}_{\mathbf{\Phi}_1}^{-1}\mathbf{\Phi}_1 - 2\mathbf{\Phi}_1^T\boldsymbol{\eta}_{\mathbf{\Phi}_1}\right]\right\} \end{aligned}$$

Completing the square in n dimensions, we have $\mathbf{\Phi}_1|\boldsymbol{z}^*, \boldsymbol{\theta}_1 \sim \mathrm{N}_n(\boldsymbol{\mu}_{\mathbf{\Phi}_1}, \boldsymbol{V}_{\mathbf{\Phi}_1})$ where

$$\begin{aligned} \boldsymbol{V}_{\mathbf{\Phi}_1} &= \left(\boldsymbol{L}^{*T}\mathbf{\Omega}\boldsymbol{L}^* + \mathbf{\Sigma}_{1|2}^{-1}\right)^{-1} \\ \boldsymbol{\mu}_{\mathbf{\Phi}_1} &= \boldsymbol{V}_{\mathbf{\Phi}_1}\boldsymbol{\eta}_{\mathbf{\Phi}_1} = \boldsymbol{V}_{\mathbf{\Phi}_1}\left[\mathbf{\Sigma}_{1|2}^{-1}\boldsymbol{\mu}_{1|2} + \boldsymbol{L}^{*T}\mathbf{\Omega}(\boldsymbol{z}^* - \boldsymbol{L}\boldsymbol{\theta}_1)\right], \end{aligned}$$

where $\mathbf{\Sigma}_{1|2}^{-1}$ and $\boldsymbol{\mu}_{1|2}$ are, respectively, the conditional prior precision matrix and (the approximated) conditional prior mean for $\mathbf{\Phi}_1|\mathbf{\Phi}_2$ given above.

5. Update $\boldsymbol{\theta}_2$: To update the $(p+K) \times 1$ vector of fixed and spline effects for the overall mean part, $\boldsymbol{\theta}_2 = (\boldsymbol{\beta}^T, \boldsymbol{\zeta}_2^T)^T$, we use a Metropolis-Hastings (MH) step with symmetric multivariate t proposal density centered at the previous value of $\boldsymbol{\theta}_2$ with acceptance

ratio

$$\rho_{\boldsymbol{\theta}_2} = \frac{p(\boldsymbol{\theta}_2^{(p)}|\boldsymbol{y}^*, \boldsymbol{\theta}_1, \boldsymbol{\Phi}_2^*, r)}{p(\boldsymbol{\theta}_2^{(s)}|\boldsymbol{y}^*, \boldsymbol{\theta}_1, \boldsymbol{\Phi}_2^*, r)} = \frac{\prod_{i=1}^{n}\prod_{t=1}^{n_i^*} \mathrm{NB}(y_{it}^*|\boldsymbol{\theta}_2^{(p)}, \boldsymbol{\theta}_1, \boldsymbol{\Phi}_2^*, r)}{\prod_{i=1}^{n}\prod_{t=1}^{n_i^*} \mathrm{NB}(y_{it}^*|\boldsymbol{\theta}_2^{(s)}, \boldsymbol{\theta}_1, \boldsymbol{\Phi}_2^*, r)} \times \frac{\mathrm{N}_p(\boldsymbol{\theta}_2^{(p)}; \boldsymbol{\theta}_0, \boldsymbol{V}_0)}{\mathrm{N}_p(\boldsymbol{\theta}_2^{(s)}; \boldsymbol{\theta}_0, \boldsymbol{V}_0)},$$

where $\boldsymbol{\theta}_2^{(p)}$ and $\boldsymbol{\theta}_2^{(s)}$ are the proposed and current values of $\boldsymbol{\theta}_2$ at iteration s, respectively; $\boldsymbol{y}^*$ is a vector of $N^* = \sum_{i=1}^{n} n_i^* \leq N$, where $n_i^* = \sum_{t=1}^{T} z_{it}$ is the number of at-risk observations for county i and z_{it} is the latent at-risk indicator for county i on day t defined in equation (7) of the manuscript; and $\mathrm{NB}(y_{it}^*|\boldsymbol{\theta}_2^{(p)}, \boldsymbol{\theta}_1, \boldsymbol{\Phi}_2^*)$ and $\mathrm{N}_{p+K}(\boldsymbol{\theta}_2^{(p)}; \boldsymbol{\theta}_0, \boldsymbol{V}_0)$ are the probability distribution functions for negative binomial and the $p+K-$variate normal prior distribution with mean $\boldsymbol{\theta}_0$ and covariance $\boldsymbol{V}_0$ evaluated at $\boldsymbol{\theta}_2^{(p)}$.

6. Update ϕ_{2i}: Similar to $\boldsymbol{\theta}_2$, to update ϕ_{2i} $(i = 1, \ldots, n)$, we use a random walk MH step with a symmetric univariate t proposal density centered at the previous ϕ_{2i} and acceptance ratio

$$\rho_{\phi_{2i}} = \frac{p(\phi_{2i}^{(p)}|\boldsymbol{y}^*, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, r)}{p(\phi_{2i}^{(s)}|\boldsymbol{y}^*, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, r)} = \frac{\prod_{t=1}^{n_i^*} \mathrm{NB}(y_{it}^*|\phi_{2i}^{(p)}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, r)}{\prod_{t=1}^{n_i^*} \mathrm{NB}(y_{it}^*|\phi_{2i}^{(s)}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, r)} \times \frac{\pi(\phi_{2i}^{(p)}|\phi_{1i})}{\pi(\phi_{2i}^{(s)}|\phi_{1i})},$$

where $\phi_{2i}^{(p)}$ and $\phi_{2i}^{(s)}$ are the proposed and current values of ϕ_{2i} at current iteration s for the n_i^* "at-risk" observations, and $\pi(\phi_{2i}|\phi_{1i})$ is the conditional univariate CAR prior distribution of $\phi_{2i}|\phi_{1i}$ analogous to equation (10) in the main manuscript. Note that the update for ϕ_{2i} only depends on the "at-risk" observations for county i.

7. Update $\boldsymbol{\Sigma}$: Assuming an $\mathrm{IW}(\nu_0, \boldsymbol{S}_0)$ prior, we update the random effects covariance matrix, $\boldsymbol{\Sigma}$, from a conjugate IW distribution given by

$$\boldsymbol{\Sigma}|\boldsymbol{\Phi} \sim \mathrm{IW}(\nu_0 + n - 1, \boldsymbol{S}_0 + \boldsymbol{S}_{\boldsymbol{\Phi}^*}),$$

where $\boldsymbol{S}_{\boldsymbol{\Phi}^*} = \boldsymbol{\Phi}^{*T}\boldsymbol{Q}\boldsymbol{\Phi}^*$ and $\boldsymbol{\Phi}^* = [\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2]$ is the $n \times 2$ random effects matrix centered at its mean.

8. Update r: To update the NB dispersion parameter, r, we use a MH step with a zero-truncated normal proposal centered at the current value of r and acceptance ratio

$$\rho_r = \frac{p(r^{(p)}|\boldsymbol{y}^*, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)}{p(r^{(s)}|\boldsymbol{y}^*, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)} = \frac{\mathrm{NB}(\boldsymbol{y}^*|r^{(p)}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)}{\mathrm{NB}(\boldsymbol{y}^*|r^{(s)}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2)} \times \frac{\mathrm{N}^+(r^{(s)}, r^{(p)}, \sigma_r)}{\mathrm{N}^+(r^{(p)}, r^{(s)}, \sigma_r)},$$

where $\boldsymbol{y}^*$ denotes the N^* total "at-risk" observations, N^+ is the proposal density function for normal distribution truncated at zero with mean set to the current value of r

at iteration s and standard deviation σ_r tuned to achieve good mixing. We assume a diffuse prior with positive support for r.

Appendix C: Sensitivity Analysis with a Reduced Model for the Binary Component

Model Component	Variable	Parm	Posterior Mean (95% CrI)
Binary	SVI	γ_1	$0.33\ (0.10, 0.57)$
	% of adult smokers	γ_2	$-0.45\ (-0.75, -0.13)$
	CVD Hospitalizations	γ_3	$0.17\ (-0.07, 0.40)$
	Population density	γ_4	$0.10\ (-0.10, 0.34)$
	$PM_{2.5}$	γ_5	$0.03\ (-0.12, 0.17)$
	Temperature	γ_6	$-3.21\ (-3.95, -2.39)$
Mean	SVI	β_1	$0.08\ (0.00, 0.15)$
	% of adult smokers	β_2	$0.26\ (0.08, 0.41)$
	No. of physicians per 100K	β_3	$0.04\ (-0.01, 0.09)$
	% fair or poor health	β_4	$-0.08\ (-0.23, 0.08)$
	CVD Hospitalizations	β_5	$-0.02\ (-0.07, 0.04)$
	Population density	β_6	$-0.07\ (-0.14, -0.02)$
	$PM_{2.5}$	β_7	$0.00\ (-0.05, 0.04)$
	Temperature	β_8	$-0.14\ (-0.20, -0.10)$
	Precipitation	β_9	$0.02\ (-0.01, 0.05)$
	Dispersion	r	$1.64\ (1.45, 1.87)$
Random Effects	$\text{var}(\phi_{1i})$	Σ_{11}	$1.75\ (0.78, 3.08)$
	$\text{cov}(\phi_{1i}, \phi_{2i})$	Σ_{12}	$0.33\ (0.14, 0.55)$
	$\text{var}(\phi_{2i})$	Σ_{22}	$0.21\ (0.15, 0.30)$

Table S2: Posterior mean estimates and 95% credible intervals (CrIs) for the COVID-19 study from the spatiotemporal MZINB model with few covariates for the binary component

References

1. Polson, N. G., Scott, J. G. and Windle, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. In *Journal of the American Statistical Association*, 108(504), 1339–1349.

2. Neelon, B. (2019). Bayesian zero-inflated negative binomial regression based on Pólya-Gamma mixtures. *Bayesian Analysis*.

3. Neelon, B., Wen, C.C., Benjamin-Neelon, S.E. (2023). A multivariate spatiotemporal model for tracking COVID-19 incidence and death rates in socially vulnerable populations. *Journal of Applied Statistics*, 1–24.

4. Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. CDC/ATSDR social vulnerability index 2018 database US 2018.