Supplementary Materials for "BayesSMILES: Bayesian Segmentation Modeling for Longitudinal Epidemiological Studies"

SHUANG JIANG[1,2], QUAN ZHOU[3], XIAOWEI ZHAN[†2], AND QIWEI LI[*4]

[1] *Department of Statistical Science, Southern Methodist University, Dallas, TX 75205, USA*
[2] *Quantitative Biomedical Research Center, Department of Population and Data Sciences, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA*
[3] *Department of Statistics, Texas A&M University, College Station, TX 77843, USA*
[4] *Department of Mathematical Sciences, The University of Texas at Dallas, Richardson, TX 75080, USA*

[†] Corresponding author. Email: Xiaowei.Zhan@UTSouthwestern.edu

[*] Corresponding author. Email: Qiwei.Li@UTDallas.edu

Received April, 2020; Accepted May, 2020

Table S1: Key notations of the Poisson segmented regression model described in Section 3.3

	Notation	Support	Definition
Data	N	$N \in \mathbb{N}$	The total population size
	$\boldsymbol{C} = \left[C_t\right]_{T\times 1}$	$C_t \in \mathbb{N}, C_t \geq C_{t-1}$	The cumulative confirmed case numbers
	$\boldsymbol{D} = \left[D_t\right]_{T\times 1}$	$D_t \in \mathbb{N}, D_t \geq D_{t-1}$	The cumulative death numbers
	$\boldsymbol{E} = \left[E_t\right]_{T\times 1}$	$E_t \in \mathbb{N}, E_t \geq E_{t-1}$	The cumulative recovery case numbers
	$\boldsymbol{S} = \left[S_t\right]_{T\times 1}$	$S_t \in \mathbb{N}$	The actively susceptible people, $S_t = N - C_t$
	$\boldsymbol{I} = \left[S_t\right]_{T\times 1}$	$I_t \in \mathbb{N}$	The actively infectious case numbers
	$\boldsymbol{R} = \left[R_t\right]_{T\times 1}$	$R_t \in \mathbb{N}, R_t \geq R_{t-1}$	The cumulative removed case numbers
	$\boldsymbol{X} = \left[x_{t,j}\right]_{T\times p}$	$x_{t,j} \in \mathbb{R}, j > 2$	The design matrix, where the first column is an all-one vector for the intercept, the second column is $(1, \ldots, T)^\top$ for the time effect, and all other columns make up the covariate matrix
	γ	$\gamma \in \mathbb{R}^+$	The pre-specified constant removal rate
	$\boldsymbol{\sigma}^2 = \left[\sigma_k^2\right]_{K\times 1}$	$\sigma_k^2 \in \mathbb{R}^+$	The pre-specified variances of regression errors
Parameters	$\boldsymbol{\beta} = \left[\beta_t\right]_{T\times 1}$	$\beta_t \in \mathbb{R}^+$	The time-varying disease transmission rates
	$\boldsymbol{\alpha} = \left[\alpha_t\right]_{T\times 1}$	$\alpha_t \in \mathbb{R}^+$	The adjusted time-varying transmission rates
	$\tilde{\boldsymbol{\alpha}} = \left[\tilde{\alpha}_t\right]_{T\times 1}$	$\tilde{\alpha}_t \in \mathbb{R}$	The adjusted time-varying transmission rates on the logarithmic scale, $\tilde{\alpha}_t = \log \alpha_t$
	$\boldsymbol{\zeta} = \left[\zeta_t\right]_{T\times 1}$	$\zeta_t \in \{0, 1\}$	The change point indicator
	$\boldsymbol{z} = \left[z_t\right]_{T\times 1}$	$z_t \in \{1, \ldots, K\}$	The segmentation indicator
	$\boldsymbol{c} = \left[c_k\right]_{K\times 1}$	$c_k \in \{1, \ldots, T\}$	The locations of change points
	$\boldsymbol{n} = \left[n_k\right]_{K\times 1}$	$n_k \in \mathbb{N}$	The number of time points in each segment
	$\boldsymbol{B} = \left[b_{j,k}\right]_{p\times K}$	$b_{j,k} \in \mathbb{R}$	The regression coefficients for each segment
	$\boldsymbol{\epsilon} = \left[\epsilon_t\right]_{T\times 1}$	$\epsilon_t \in \mathbb{R}$	The regression errors (i.e. the process error)
	$\boldsymbol{H} = \text{Diag}(h_0, \ldots, h_{p-1})$	$h_j \in \mathbb{R}^+$	The diagonal variance-covariance matrix that defines the prior on each column in $\boldsymbol{B}$, i.e. $\boldsymbol{b}_k$
	ω	$\omega \in (0, 1)$	The probability of being a change point *a priori*
	a_ω, b_ω	$a_\omega, b_\omega \in \mathbb{R}^+$	The hyperparameters for ω
Others	$\top$		The matrix transpose operator
	$\lceil \cdot \rceil$		The ceiling function
	$\delta(\cdot)$		The indicator function
	$\mathcal{I}_n = \text{Diag}(1, \ldots, 1)$		The n-by-n identity matrix
	$\boldsymbol{0}_n = \left[0\right]_{n\times 1}$		The n-dimension all-zero column vector

Table S2: Key notations of the stochastic SIR model described in Section 3.4

	Notation	Support	Definition
Data	N	$N \in \mathbb{N}$	The total population size
	$\boldsymbol{C} = \left[C_t\right]_{T\times 1}$	$C_t \in \mathbb{N}, C_t \geq C_{t-1}$	The cumulative confirmed case numbers
	$\dot{\boldsymbol{C}} = \left[\dot{C}_t\right]_{T\times 1}$	$\dot{C}_t \in \mathbb{N}$	The new confirmed case numbers, $\dot{C}_t = C_t - C_{t-1}$
	$\boldsymbol{z} \ = \left[z_t\right]_{T\times 1}$	$z_t \in \{1, \ldots, K\}$	The segmentation indicator, with $z_t = k$ indicating that time point t is in segment k
	γ	$\gamma \in \mathbb{R}^+$	The pre-specified constant removal rate
Parameters	$\dot{\boldsymbol{R}} = \left[\dot{R}_t\right]_{T\times 1}$	$\dot{R}_t \in \mathbb{N}$	The new removed case numbers, which are treated as missing data
	$\boldsymbol{\beta}^* \ = \left[\beta^*_k\right]_{K\times 1}$	$\beta^*_k \in \mathbb{R}^+$	The disease transmission rates for each segment
	$\boldsymbol{\mathcal{R}}_0 = \left[\mathcal{R}_{0k}\right]_{K\times 1}$	$\mathcal{R}_{0k} \in \mathbb{R}^+$	The basic reproduction numbers for each segment
	$\boldsymbol{\phi} \ = \left[\phi_k\right]_{K\times 1}$	$\phi_k \in \mathbb{R}^+$	The dispersion parameters for each segment
	a_ϕ, b_ϕ	$a_\phi, b_\phi \in \mathbb{R}^+$	The hyperparameters for all ϕ_k's

Table S3: Case study: The list of the change points identified by BayesSMILES and the related supporting evidences for the states of New York, Texas, Florida, and California.

State	Change point location (credible interval)	Event	Reference
New York	Mar 28 (Mar 26 - Apr 2)	14-day travel advisory	NBC New York (2020a)
	Apr 9 (Apr 4 - Apr 13)	First negative number for ICU admissions	Jesse Pound (2020)
	Apr 27 (Apr 24 - May 1)		
	Jun 18 (Jun 14 - Jun 23)	Phase II reopening	NBC New York (2020b)
	Jun 29 (Jun 28 - Jul 4)		
Texas	Mar 28 (Mar 28 - Mar 30)	14-day travel advisory	Office of the Texas Governor (2020b)
	Apr 9 (Apr 8 - Apr 11)	Texas disaster declaration extension	Ciara Rouege (2020)
	May 25 (May 22 - May 26)	Additional services and activities allowed to open for phase II reopening	Office of the Texas Governor (2020c)
	Jun 16 (Jun 13 - Jun 17)	The second COVID-19 wave in June	Billy Gates (2020)
	Jun 28 (Jun 24 - Jul 3)	Multiple executive orders issued for mitigating the disease spreading	Office of the Texas Governor (2020d) Office of the Texas Governor (2020e) Office of the Texas Governor (2020a)
Florida	Apr 3 (Apr 2 - Apr 5)	Statewide stay-at-home order	Klas and Contorno (2020)
	Apr 14 (Apr 10 - Apr 19)	Some counties require face coverings in public	Metevia (2020)
	May 13 (May 8 - May 18)	Phase I reopening	Silcox and Turco (2020)
	Jun 7 (Jun 2 - Jun 11)	Phase II reopening	Dobrzyn (2020)
	Jun 27 (Jun 26 - Jul 2)	Mandatory face coverings in public for major cities	Frago (2020) Speck and Sandoval (2020) The City Manager of the City of Miami (2020) Piggott and McLean (2020)
California	Apr 5 (Apr 3 - Apr 6)		
	Jun 17 (Jun 12 - Jun 22)	Higher-risk businesses reopening	Sandhya Kambhampati and Krishnakumar (2020)

S4. Supplement to Section 4: Details of the MCMC Algorithms

This section shows the details of the MCMC algorithms of the proposed Bayesian framework, and the multivariate normal density approximation used to improve the computational efficiency for computing the posterior distributions for the proposed Poisson segmented regression model.

S4.1 MCMC algorithms for detecting change points

We start by writing the full data likelihood and the priors of the change point detection model as

$$\begin{aligned} f(\boldsymbol{I}|\boldsymbol{\alpha}) &= \prod_{t=1}^{T} \mathrm{Poi}(I_t; N\alpha_t) \\ \pi(\tilde{\boldsymbol{\alpha}}|\boldsymbol{\zeta}) &= \prod_{k=1}^{K} \mathrm{MN}\left(\tilde{\boldsymbol{\alpha}}_k; \mathbf{0}_{n_k}, \boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k}\right) \qquad (1) \\ \pi(\boldsymbol{\zeta}) &= \prod_{t=2}^{T} \text{Be-Bern}(\zeta_t; a_\omega, b_\omega). \end{aligned}$$

The full posterior distribution is $\pi(\boldsymbol{\alpha}, \boldsymbol{\zeta}|\boldsymbol{I}) \propto f(\boldsymbol{I}|\boldsymbol{\alpha}, \boldsymbol{\zeta})\pi(\boldsymbol{\alpha}, \boldsymbol{\zeta}) = f(\boldsymbol{I}|\boldsymbol{\alpha})\pi(\tilde{\boldsymbol{\alpha}}|\boldsymbol{\zeta})\pi(\boldsymbol{\zeta})$. Below, we describe the Metropolis-Hastings (MH) algorithms to sample from $\pi(\boldsymbol{\alpha}|\boldsymbol{\zeta}, \boldsymbol{I})$ and $\pi(\boldsymbol{\zeta}|\boldsymbol{\alpha}, \boldsymbol{I})$.

Update the change point indicator ζ: We update the binary latent vector $\boldsymbol{\zeta}$ via an *add-delete-swap* algorithm. We randomly select an entry in $\boldsymbol{\zeta}$, say ζ_t, and change its value to $\zeta_t^{\text{new}} = 1 - \zeta_t$ to form a new $\boldsymbol{\zeta}^{\text{new}}$. This is an *add* step if $\zeta_t^{\text{new}} = 1$ and a *delete* step otherwise. The *swap* step is performed every ten iterations, where we randomly select a change point, say $\zeta_t = 1$, and swap the values between the t and $(t \pm 1)$-th entries in $\boldsymbol{\zeta}$ to form a new $\boldsymbol{\zeta}^{\text{new}}$. We accept the proposed $\boldsymbol{\zeta}^{\text{new}}$ with the probability $\min(1, m_{\text{MH}})$, where the acceptance ratio is

$$\begin{aligned} m_{\text{MH}} &= \frac{\pi(\tilde{\boldsymbol{\alpha}}|\boldsymbol{\zeta}^{\text{new}})}{\pi(\tilde{\boldsymbol{\alpha}}|\boldsymbol{\zeta})} \frac{\pi(\boldsymbol{\zeta}^{\text{new}})}{\pi(\boldsymbol{\zeta})} \frac{J(\boldsymbol{\zeta}^{\text{new}} \leftarrow \boldsymbol{\zeta})}{J(\boldsymbol{\zeta} \leftarrow \boldsymbol{\zeta}^{\text{new}})} \\ &= \frac{\prod_{k=1}^{K} \mathrm{MN}\left(\tilde{\boldsymbol{\alpha}}_k; \mathbf{0}_{n_k}, \boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k}\right)}{\prod_{k=1}^{K} \mathrm{MN}\left(\tilde{\boldsymbol{\alpha}}_k; \mathbf{0}_{n_k}, \boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k}\right)} \times \begin{cases} a_\omega / b_\omega & \text{for } \textit{add} \\ b_\omega / a_\omega & \text{for } \textit{delete} \\ 1 & \text{for } \textit{swap} \end{cases} \end{aligned}$$

Here we use $J(\cdot \leftarrow \cdot)$ to denote the proposal probability distribution for the selected move. Note that the last proposal density ratio equals one. This step simultaneously updates the segmentation vector $\boldsymbol{z}$, as it can be constructed from $\boldsymbol{\zeta}$.

Update the adjusted time-varying transmission rates $\boldsymbol{\alpha}$: For each segment partitioned by $\boldsymbol{\zeta}$, we update α_t within the same segment, say segment k, sequentially by using a random walk Metropolis-Hastings (RWMH) algorithm. We first propose a new $\tilde{\alpha}_t^{\text{new}}$ from $\mathrm{N}\left(\tilde{\alpha}_t, \tau_\alpha^2\right)$. Let $\boldsymbol{\alpha}_k^{(\zeta)\text{new}} = (\alpha_{c_k}, \ldots, \alpha_t^{\text{new}}, \ldots, \alpha_{c_k + n_k - 1})$. Then we accept the proposed value α_t^{new} with probability $\min(1, m_{\text{MH}})$, where the acceptance ratio is

$$\begin{aligned} m_{\text{MH}} &= \frac{f(\boldsymbol{I}|\boldsymbol{\alpha}^{\text{new}})}{f(\boldsymbol{I}|\boldsymbol{\alpha})} \frac{\pi(\tilde{\boldsymbol{\alpha}}^{\text{new}}|\boldsymbol{\zeta})}{\pi(\tilde{\boldsymbol{\alpha}}|\boldsymbol{\zeta})} \frac{J(\tilde{\alpha}^{\text{new}} \leftarrow \tilde{\alpha})}{J(\tilde{\alpha} \leftarrow \tilde{\alpha}^{\text{new}})} \\ &= \frac{\mathrm{Poi}(I_t; \alpha_t^{\text{new}})}{\mathrm{Poi}(I_t; \alpha_t)} \frac{\mathrm{MN}\left(\tilde{\boldsymbol{\alpha}}_k^{\text{new}}; \mathbf{0}_{n_k}, \boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k}\right)}{\mathrm{MN}\left(\tilde{\boldsymbol{\alpha}}_k; \mathbf{0}_{n_k}, \boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k}\right)}. \end{aligned}$$

Note that the proposal density ratio cancels out for this RWMH update. The computation of the multivariate normal (MN) probability density involves matrix inversion, which can be time-consuming, particularly when n_k is large. To significantly improve the computational efficiency, we follow Zhou and Guan (2018) to approximate the exact inversion under an appropriate choice of $\boldsymbol{H}$ that satisfies the asymptotic condition. As mentioned previously, $\boldsymbol{H}$ is a p-by-p diagonal matrix, where the first entry h_0 corresponds to the variance of the normal prior on $b_{1,k}$. Under the asymptotic condition of $h_0 \geq h_j, \forall j \neq 0$, the inversion of an n_k-by-n_k matrix is reduced to an inversion of a p-by-p matrix. In practice, we set $h_0 = 10,000$ and $h_1 = \ldots = h_{p-1} = 10$ to ensure this asymptotic condition. The full details of the approximation method are available in Section S4.3.

S4.2 MCMC algorithms for estimating basic reproduction numbers

We start by writing the full data likelihood and the priors of the stochastic SIR model are written as,

$$\begin{aligned} f(\dot{\boldsymbol{C}}|\boldsymbol{\beta}^*, \boldsymbol{\phi}, \boldsymbol{R}) &= \prod_{k=1}^{K} \prod_{\{t:z_t=k\}} \mathrm{NB}\left(\dot{C}_t; \beta_k^*(N - C_{t-1})\frac{N - C_{t-1} - R_{t-1}}{N}, \phi_k\right) \\ \pi(\boldsymbol{\beta}^*) &= \prod_{k=1}^{K} \mathrm{Ga}(\beta_k^*; a_\beta, b_\beta) \\ \pi(\boldsymbol{\phi}) &= \prod_{k=1}^{K} \mathrm{Ga}(\phi_k; a_\phi, b_\phi), \end{aligned} \tag{2}$$

where $\boldsymbol{\beta}^* = (\beta_1^*, \ldots, \beta_K^*)$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_K)$, i.e. the collections of transmission and dispersion rates of all segments. For the hyperparameters, we set $a_\beta = 1$ and $b_\beta = 1/\gamma$ so that both of the expectation and variance of the basic reproduction number $\mathcal{R}_0 = \beta_k^*/\gamma$ are equal to one. With a pre-defined removal rate γ, we propose the following updates in each MCMC iterations.

Generate $\boldsymbol{R}$ based on $\boldsymbol{C}$: We assume $I_1 = C_1$ and $R_1 = 0$, i.e. all the confirmed cases are capable of passing the disease to all susceptible individuals in a closed population at time point $t = 1$. Then we sample $\dot{R}_2 \sim \mathrm{Poi}(\gamma I_1)$, where γ is a pre-specified tuning parameter and $\dot{R}_2 = R_2 - R_1$ is the new removed case number at time point $t = 2$. Due to the compositional nature of the SIR model, we can compute $I_2 = I_1 + \dot{C}_2 - \dot{R}_2$, where $\dot{C}_2 = C_2 - C_1$ is the new confirmed cases at time point $t = 2$. Next, we repeat this process of sampling $\dot{R}_t \sim \mathrm{Poi}(\gamma I_{t-1})$ and computing $I_t = I_{t-1} + \dot{C}_t - \dot{R}_t, t = 3, \ldots, T$, to generate the sequence $\boldsymbol{R}$ used in every iteration.

Update the dispersion parameters $\boldsymbol{\phi}$: For each segment, we update ϕ_k by using an RWMH algorithm. We first propose a new ϕ_k^{new}, of which the logarithmic value is generated from $\mathrm{N}\left(\log \phi_k, \tau_\phi^2\right)$. Let $\boldsymbol{\phi}^{\text{new}} = (\phi_1, \ldots, \phi_k^{\text{new}}, \ldots, \phi_K)$, where only the k-th entry is replaced. Then we accept the proposed value ϕ_k^{new} with probability $\min(1, m_{\text{MH}})$, where the acceptance ratio is

$$
\begin{aligned}
m_{\mathrm{MH}} &= \frac{f(\dot{\boldsymbol{C}}|\boldsymbol{\beta}^*, \boldsymbol{\phi}^{\mathrm{new}}, \boldsymbol{R})}{f(\dot{\boldsymbol{C}}|\boldsymbol{\beta}^*, \boldsymbol{\phi}, \boldsymbol{R})} \frac{\pi\left(\phi^{\mathrm{new}}\right)}{\pi\left(\phi\right)} \frac{J\left(\phi \leftarrow \phi^{\mathrm{new}}\right)}{J\left(\phi^{\mathrm{new}} \leftarrow \phi\right)} \\
&= \frac{\prod_{\{t:z_t=k\}} \mathrm{NB}\left(\dot{C}_t; \beta_k^*(N - C_{t-1})\frac{N-C_{t-1}-R_{t-1}}{N}, \phi_k^{\mathrm{new}}\right)}{\prod_{\{t:z_t=k\}} \mathrm{NB}\left(\dot{C}_t; \beta_k^*(N - C_{t-1})\frac{N-C_{t-1}-R_{t-1}}{N}, \phi_k\right)} \frac{\mathrm{Ga}(\phi_k^{\mathrm{new}}; a_\phi, b_\phi)}{\mathrm{Ga}(\phi_k; a_\phi, b_\phi)}.
\end{aligned}
$$

Note that the proposal density ratio cancels out for this RWMH update.

Update the disease transmission rates $\boldsymbol{\beta}^*$: For each segment, we update β_k^* by using an RWMH algorithm. We first propose a new $\beta_k^{*\mathrm{new}}$, of which the logarithmic value is generated from $\mathrm{N}\left(\log \beta_k^*, \tau_\beta^2\right)$. Let $\boldsymbol{\beta}^{*\mathrm{new}} = (\beta_1^*, \ldots, \beta_k^{*\mathrm{new}}, \ldots, \beta_K^*)$, where only the k-th entry is replaced. Then we accept the proposed value $\beta_k^{*\mathrm{new}}$ with probability $\min(1, m_{\mathrm{MH}})$, where the acceptance ratio is

$$
\begin{aligned}
m_{\mathrm{MH}} &= \frac{f(\dot{\boldsymbol{C}}|\boldsymbol{\beta}^{*\mathrm{new}}, \boldsymbol{\phi}, \boldsymbol{R})}{f(\dot{\boldsymbol{C}}|\boldsymbol{\beta}^*, \boldsymbol{\phi}, \boldsymbol{R})} \frac{\pi\left(\beta^{*\mathrm{new}}\right)}{\pi\left(\beta^*\right)} \frac{J\left(\beta^* \leftarrow \beta^{*\mathrm{new}}\right)}{J\left(\beta^{*\mathrm{new}} \leftarrow \beta^*\right)} \\
&= \frac{\prod_{\{t:z_t=k\}} \mathrm{NB}\left(\dot{C}_t; \beta_k^{*\mathrm{new}}(N - C_{t-1})\frac{N-C_{t-1}-R_{t-1}}{N}, \phi_k\right)}{\prod_{\{t:z_t=k\}} \mathrm{NB}\left(\dot{C}_t; \beta_k^*(N - C_{t-1})\frac{N-C_{t-1}-R_{t-1}}{N}, \phi_k\right)} \frac{\mathrm{Ga}(\beta_k^{*\mathrm{new}}; a_\beta, b_\beta)}{\mathrm{Ga}(\beta_k^*; a_\beta, b_\beta)}.
\end{aligned}
$$

Note that the proposal density ratio cancels out for this RWMH update.

S4.3 Approximate the multivariate normal density function

We consider a general setting as follows. Let $\boldsymbol{y}$ be an $n \times 1$ vector, $\boldsymbol{W}$ be an $n \times q$ matrix, $\boldsymbol{U}$ be an $n \times p$ matrix, and $\boldsymbol{X} = (\boldsymbol{W}, \boldsymbol{U})$ (which is an $n \times (p+q)$ matrix). Let $\boldsymbol{\Sigma}$ be a $(q+p) \times (q+p)$ diagonal matrix where the first q diagonal elements are h_0 and the last p diagonal elements are h_1. Suppose $h_0, h_1, \sigma^2 > 0$. By Woodbury identity and Sylvester's determinant identity, we have

$$
\sigma^2 \boldsymbol{y}^\top (\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\top + \sigma^2 \mathcal{I}_n)^{-1} \boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{y} - \boldsymbol{y}^\top \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{X} + \sigma^2 \boldsymbol{\Sigma}^{-1})^{-1} \boldsymbol{X}^\top \boldsymbol{y}, \tag{3}
$$

$$
\left|\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\top + \sigma^2 \mathcal{I}_n\right| = \sigma^{2n} \left|(h_0/\sigma^2)\boldsymbol{W}^\top \boldsymbol{W} + \mathcal{I}_n\right| \cdot \left|h_1 \boldsymbol{U}^\top (h_0 \boldsymbol{W}\boldsymbol{W}^\top + \sigma^2 \mathcal{I}_n)^{-1} \boldsymbol{U} + \mathcal{I}_n\right|, \tag{4}
$$

where $|\cdot|$ denotes the matrix determinant and $\mathcal{I}_n$ is an n-dimensional diagonal matrix. Define

$$
\boldsymbol{P} = \mathcal{I}_n - \boldsymbol{W}(\boldsymbol{W}^\top \boldsymbol{W})^{-1} \boldsymbol{W}^\top, \qquad \tilde{\boldsymbol{U}} = \boldsymbol{P}\boldsymbol{U}, \qquad \tilde{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{y},
$$

where $\tilde{\boldsymbol{U}}$ (respectively $\tilde{\boldsymbol{y}}$) is the residual after regressing out $\boldsymbol{W}$ from $\boldsymbol{U}$ (respectively $\boldsymbol{y}$). Zhou and Guan (2018) showed that the expressions in (3) and (4) can be further simplified when $h_0 \to \infty$. The results are summarized below with the proof available in the supplement of Zhou and Guan (2018).

Lemma 1. *Let $\boldsymbol{y}, \boldsymbol{W}, \boldsymbol{U}, \boldsymbol{X}, \boldsymbol{\Sigma}, \tilde{\boldsymbol{y}}, \tilde{\boldsymbol{U}}, \sigma^2, h_0, h_1$ be defined as above. Then,*

$$
\lim_{h_0 \uparrow \infty} \sigma^2 \boldsymbol{y}^\top (\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\top + \sigma^2 \mathcal{I}_n)^{-1} \boldsymbol{y} = \tilde{\boldsymbol{y}}^\top \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{y}}^\top \tilde{\boldsymbol{U}} (\tilde{\boldsymbol{U}}^\top \tilde{\boldsymbol{U}} + (\sigma^2/h_1)\mathcal{I}_n)^{-1} \tilde{\boldsymbol{U}}^\top \tilde{\boldsymbol{y}}, \tag{5}
$$

and

$$
\lim_{h_0 \uparrow \infty} \frac{\left|\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}^\top + \sigma^2 \mathcal{I}_n\right|}{\left|h_0 \boldsymbol{W}\boldsymbol{W}^\top + \sigma^2 \mathcal{I}_n\right|} = \left|(h_1/\sigma^2)\tilde{\boldsymbol{U}}^\top \tilde{\boldsymbol{U}} + \mathcal{I}_n\right|. \tag{6}
$$

The conclusions in Lemma 1 can be used to improve the computational efficiency for approximating the multivariate normal probability density function, shown in Equation (1), in our model. Within each segment k we assumed $\tilde{\boldsymbol{\alpha}}_k \sim \mathrm{MN}(\boldsymbol{X}_k\boldsymbol{b}_k, \sigma_k^2\mathcal{I}_{n_k})$. Under the prior specification discussed in Section 3.3, we have $\tilde{\boldsymbol{\alpha}}_k \sim \mathrm{MN}(0, \boldsymbol{X}_k\boldsymbol{H}\boldsymbol{X}_k^\top + \sigma_k^2\mathcal{I}_{n_k})$, and the corresponding p.d.f. for $\tilde{\boldsymbol{\alpha}}_k$ is:

$$\pi(\tilde{\boldsymbol{\alpha}}_k) = (2\pi)^{-n_k/2}\left|\boldsymbol{X}_k\boldsymbol{H}\boldsymbol{X}_k^\top + \sigma_k^2\mathcal{I}_{n_k}\right|^{-1/2}\exp\left\{-\frac{1}{2}\tilde{\boldsymbol{\alpha}}_k^\top(\boldsymbol{X}_k\boldsymbol{H}\boldsymbol{X}_k^\top + \sigma_k^2\mathcal{I}_{n_k})^{-1}\tilde{\boldsymbol{\alpha}}_k\right\}, \tag{7}$$

where n_k is the segment length. Next, we simplify the calculation of $\left|\boldsymbol{X}_k\boldsymbol{H}\boldsymbol{X}_k^\top + \sigma_k^2\mathcal{I}_{n_k}\right|$ and $(\boldsymbol{X}_k\boldsymbol{H}\boldsymbol{X}_k^\top + \sigma_k^2\mathcal{I}_{n_k})^{-1}$ by using Lemma 1. Consider $\boldsymbol{U} = (t_1, \ldots, t_{n_k})^\top$ and let $\boldsymbol{W}$ be a column vector of 1's with length n_k. The vector $\boldsymbol{y}$ in our case matches $\tilde{\boldsymbol{\alpha}}_k$, and $\boldsymbol{\Sigma} = \boldsymbol{H} = \left(\begin{smallmatrix} h_0 & 0 \\ 0 & h_1 \end{smallmatrix}\right)$. Lemma 1 states that the inverse and determinant calculation of an $n_k \times n_k$ matrix can be reduced to that of a $p \times p$ matrix, and in our case $p = 1$ (since the regression model only includes "time" as a covariate except the intercept term). Therefore, the computational benefit could be significant when n_k was large. We first derive the formula of $\boldsymbol{P}$ as follows,

$$\boldsymbol{P} = \mathcal{I}_{n_k} - \boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top = \begin{pmatrix} 1-\frac{1}{n_k} & -\frac{1}{n_k} & \cdots & -\frac{1}{n_k} \\ -\frac{1}{n_k} & 1-\frac{1}{n_k} & \cdots & -\frac{1}{n_k} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n_k} & -\frac{1}{n_k} & \cdots & 1-\frac{1}{n_k} \end{pmatrix}.$$

Next, according to Equation (5), $\left|\boldsymbol{X}_k\boldsymbol{H}\boldsymbol{X}_k^\top + \sigma_k^2\mathcal{I}_{n_k}\right|$ can be approximated by $\left|(h_1/\sigma_k^2)\tilde{\boldsymbol{U}}^\top\tilde{\boldsymbol{U}} + \mathcal{I}_{n_k}\right| \times \left|h_0\boldsymbol{W}\boldsymbol{W}^\top + \sigma_k^2\mathcal{I}_{n_k}\right|$ when $h_0 \to +\infty$. In particular, we can derive

$$\tilde{\boldsymbol{U}}^\top = (\boldsymbol{P}\boldsymbol{U})^\top = \begin{pmatrix} t_1 - \bar{t} & t_2 - \bar{t} & \ldots & t_{n_k} - \bar{t} \end{pmatrix},$$

where $\bar{t} = n_k^{-1}\sum_{i=1}^{n_k} t_i$ denotes the average. Therefore,

$$\left|(h_1/\sigma_k^2)\tilde{\boldsymbol{U}}^\top\tilde{\boldsymbol{U}} + \mathcal{I}_{n_k}\right| = (h_1/\sigma_k^2)\sum_{i=1}^{n_k}(t_i - \bar{t})^2 + 1. \tag{8}$$

To calculate $\left|h_0\boldsymbol{W}\boldsymbol{W}^\top + \sigma_k^2\mathcal{I}_{n_k}\right|$, upon noting that $\boldsymbol{W}\boldsymbol{W}^\top$ is an $n_k \times n_k$ matrix, we find

$$\left|h_0\boldsymbol{W}\boldsymbol{W}^\top + \sigma_k^2\mathcal{I}_{n_k}\right| = h_0^{n_k}\left|\boldsymbol{W}\boldsymbol{W}^\top + \frac{\sigma_k^2}{h_0}\mathcal{I}_{n_k}\right|. \tag{9}$$

According to Sylvester's determinant lemma, we have

$$\left|\boldsymbol{W}\boldsymbol{W}^\top + \frac{\sigma_k^2}{h_0}\mathcal{I}_{n_k}\right| = \left|\mathcal{I}_{n_k} + \frac{h_0}{\sigma_k^2}\boldsymbol{W}^\top\boldsymbol{W}\right| \times \left(\frac{\sigma_k^2}{h_0}\right)^{n_k} = \left(1 + \frac{n_k h_0}{\sigma_k^2}\right) \times \left(\frac{\sigma_k^2}{h_0}\right)^{n_k}.$$

Therefore, the formula in Equation (9) equals the following,

$$\left|h_0\boldsymbol{W}\boldsymbol{W}^\top + \sigma_k^2\mathcal{I}_{n_k}\right| = h_0^{n_k} \times \left(1 + \frac{n_k h_0}{\sigma_k^2}\right) \times \left(\frac{\sigma_k^2}{h_0}\right)^{n_k} = (\sigma_k^2)^{n_k} \times \left(1 + \frac{n_k h_0}{\sigma_k^2}\right). \tag{10}$$

Combining the results in Equations (8) and (10), we can approximate the matrix determinant in (7) as follows when $h_0 \to \infty$,

$$\left|\boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k}\right|^{-1/2} \approx \frac{\sigma^{-n_k}}{\sqrt{\frac{n_k h_0}{\sigma_k^2}+1}\sqrt{\frac{h_1}{\sigma_k^2}\sum_{i=1}^{n_k}(t_i-\bar{t})^2+1}}. \tag{11}$$

Next, it is straightforward to derive the formula in the exponent part in Equation (7) using the result in Equation (5). In particular, we can derive:

$$\begin{aligned}
\tilde{\boldsymbol{y}}^\top &= (\boldsymbol{P}\boldsymbol{y})^\top = \tilde{\boldsymbol{\alpha}}_k^\top \boldsymbol{P} = \left(\tilde{\alpha}_1 - \overline{\tilde{\alpha}_k} \ \ \ldots \ \ \tilde{\alpha}_{n_k} - \overline{\tilde{\alpha}_k}\right), \\
\left(\tilde{\boldsymbol{U}}^\top \tilde{\boldsymbol{U}} + \frac{\sigma_k^2}{h_1}\mathcal{I}_{n_k}\right)^{-1} &= \left(\sum_{i=1}^{n_k}(t_i-\bar{t})^2 + \frac{\sigma_k^2}{h_1}\right)^{-1}, \\
\tilde{\boldsymbol{y}}^\top \tilde{\boldsymbol{U}} &= \sum_{i=1}^{n_k}(t_i-\bar{t})(\tilde{\alpha}_i - \overline{\tilde{\alpha}_k}).
\end{aligned}$$

Then we approximate the exponent part in Equation (7) as follows under the condition that $h_0 \to \infty$,

$$-\frac{1}{2}\tilde{\boldsymbol{\alpha}}_k^\top (\boldsymbol{X}_k \boldsymbol{H} \boldsymbol{X}_k^\top + \sigma_k^2 \mathcal{I}_{n_k})^{-1} \tilde{\boldsymbol{\alpha}}_k \approx -\frac{1}{2\sigma_k^2}\left(\sum_{i=1}^{n_k}(\tilde{\alpha}_i - \overline{\tilde{\alpha}_k})^2 - \frac{\left(\sum_{i=1}^{n_k}(t_i-\bar{t})(\tilde{\alpha}_i - \overline{\tilde{\alpha}_k})\right)^2}{\sum_{i=1}^{n_k}(t_i-\bar{t})^2 + \frac{\sigma_k^2}{h_1}}\right). \tag{12}$$

We further introduce the following notation,

$$S_{tt} = \sum_{i=1}^{n_k}(t_i-\bar{t})^2, \quad S_{\alpha\alpha} = \sum_{i=1}^{n_k}(\tilde{\alpha}_i - \overline{\tilde{\alpha}_k})^2, \quad S_{\alpha t} = \sum_{i=1}^{n_k}(\tilde{\alpha}_i - \overline{\tilde{\alpha}_k})(t_i-\bar{t}).$$

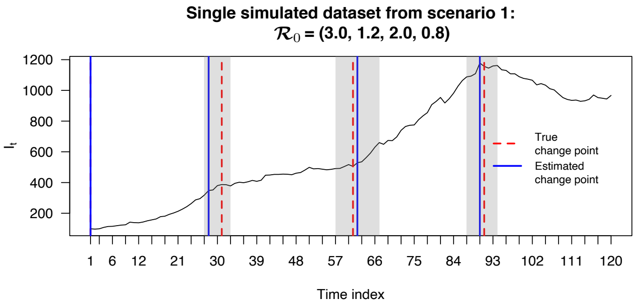
By combining the approximations in (11) and (12), we obtain

$$p(\tilde{\boldsymbol{\alpha}}_k) \approx \frac{(2\pi)^{-n_k/2}\sigma_k^{-n_k}}{\sqrt{\frac{n_k h_0}{\sigma_k^2}+1}\sqrt{\frac{h_1}{\sigma_k^2}S_{tt}+1}} \exp\left\{-\frac{1}{2\sigma_k^2}\left(S_{\alpha\alpha} - \frac{S_{\alpha t}^2}{S_{tt} + \frac{\sigma_k^2}{h_1}}\right)\right\}.$$

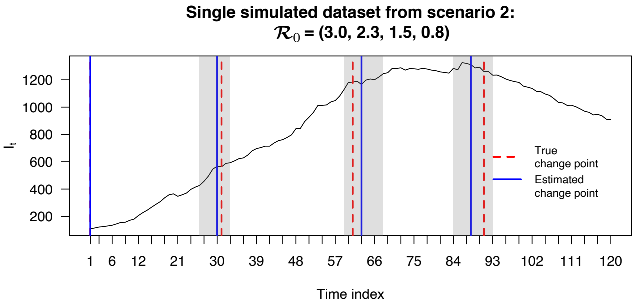
In practice, assuming that all covariates are standardized, this approximation works well if h_0/h_1 is large enough. We suggest using $h_0 = 10,000$ and $h_1 = 10$, which yielded highly accurate approximations in our simulation.

S5. Supplement to Section 5: Simulation Study Results

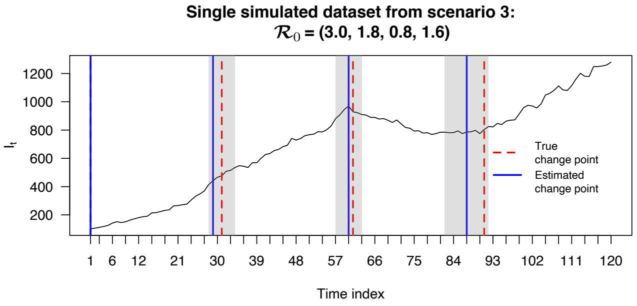
This section provides additional simulation results for Scenarios 1, 2, and 3 described in Section 5.1. For each scenario, we randomly selected a single simulated dataset from the 50 replicates. Figure S1 shows the PPM estimates on the change point indicator $\boldsymbol{\zeta}$. The red dashed and blue solid lines represent the true and estimated change point locations, respectively, while the gray ribbons indicate the corresponding 95% credible intervals. As we can see, BayesSMILES correctly pointed out the true change points. The resulting ARI values were 0.87, 0.85, and 0.95, respectively, while the MI values were 1.21, 1.18, and 1.31, respectively. Figure S2 shows our estimates on the basic reproduction number $\mathcal{R}_{0k}$ for each segment partitioned by the identified change points. The red dashed and blue solid line pinpoint the true and posterior mean of $\mathcal{R}_{0\hat{k}}$'s, while the two black solid lines mark the boundary of 95% credible intervals. In all scenarios, the true values were within their corresponding 95% credible intervals. The $\mathcal{R}_0$ RMSEs for the single datasets from the three scenarios were 0.33, 0.36, and 0.14, respectively.



(a) Change point detection performance for Scenario 1

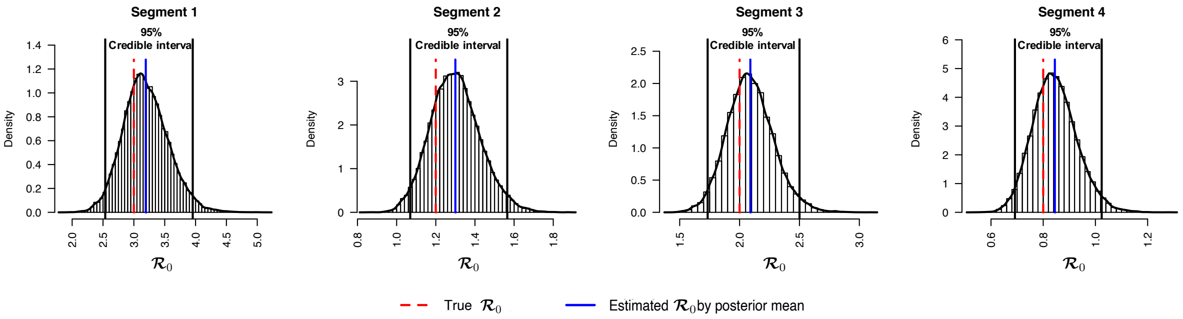


(b) Change point detection performance for Scenario 2

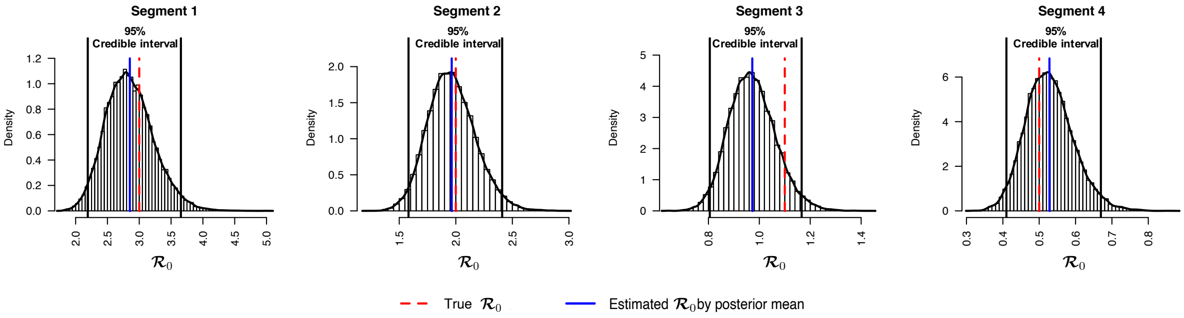


(c) Change point detection performance for Scenario 3

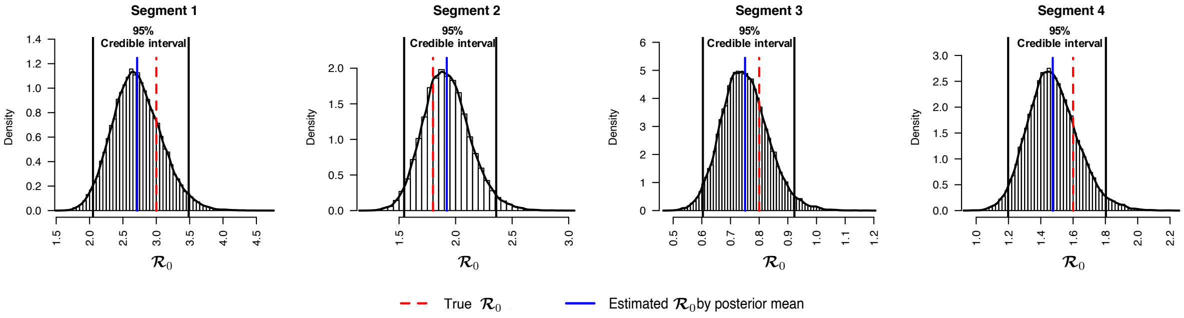
Figure S1: Simulation study: The locations of change points (blue solid lines) estimated from the posterior pairwise probability matrix (PPM) and their credible intervals (gray ribbons). The red dashed lines mark the true change point locations.



(a) Basic reproduction number estimation performance for Scenario 1



(b) Basic reproduction number estimation performance for Scenario 2



(c) Basic reproduction number estimation performance for Scenario 3

Figure S2: Simulation study: The posterior distributions of $\mathcal{R}_{0k}$'s for $k = 1, 2, 3, 4$ estimated from the segmented time-series data, given the three identified change points as shown in Figure S1. The red dashed and blue solid lines are the true and estimated values of $\mathcal{R}_{0k}$'s, respectively. The two black solid lines are the lower and upper bounds of the 95% credible intervals.

References

Billy Gates AGYC Matthew Prendergast (2020). Austin health leaders concerned about possible 2nd COVID-19 wave in June. *kxan.* URL: `https://www.kxan.com/news/austin-public-health-to-give-covid-19-update-at-9-a-m/`, accessed 2020-05-27.

Ciara Rouege (2020). Abbott extends Texas disaster declaration | Here's what that means. *KHOU.* URL: `https://www.khou.com/article/news/health/coronavirus/gov-abbott-extends-disaster-declaration-another-30-days/285-9e494ee3-5ea5-48d9-bf35-b8c4b88f5126`, accessed 2020-04-12.

Dobrzyn E (2020). Gov. Ron DeSantis says most of Florida will enter phase 2 of reopening Friday. *Click Orlando.* URL: `https://www.clickorlando.com/news/local/2020/06/03/gov-ron-desantis-says-most-of-florida-will-enter-phase-2-of-reopening-friday/`, accessed 2020-06-03.

Frago C (2020). Tampa's mask order went into effect Friday. Here's what you need to know. *Tampa Bay News.* URL: `https://www.tampabay.com/news/tampa/2020/06/19/tampas-mask-order-goes-into-effect-today-heres-what-you-need-to-know/`, accessed 2020-06-19.

Jesse Pound (2020). New York posts negative net change in ICU admissions for first time since coronavirus outbreak. *CNBC.* URL: `https://www.cnbc.com/2020/04/10/ny-has-first-negative-net-change-in-icu-admissions-since-coronavirus-outbreak.html`, accessed 2020-04-10.

Klas ME, Contorno S (2020). Florida Gov. Ron DeSantis issues statewide stay-at-home order. *Tampa Bay News.* URL: `https://www.tampabay.com/news/health/2020/04/01/florida-gov-ron-desantis-issues-statewide-stay-at-home-order/`, accessed 2020-04-01.

Metevia T (2020). Osceola County now requires face coverings in public. *Click Orlando.* URL: `https://www.clickorlando.com/news/local/2020/04/10/watch-live-osceola-county-officials-provide-update-on-covid-19-pandemic/`, accessed 2020-04-13.

NBC New York (2020a). CDC Issues 14-Day Travel Advisory for New York, New Jersey, Connecticut. *NBC New York.* URL: `https://www.nbcnewyork.com/news/local/even-with-relief-bill-passed-no-rest-for-ny-as-cuomo-says-peak-of-crisis-still-yet-to-come/2348306/`, accessed 2020-03-28.

NBC New York (2020b). NYC Moves to Phase II in Biggest Reopening Step Yet; State Daily Virus Deaths Plunge to New Low. *NBC New York.* URL: `https://www.nbcnewyork.com/news/local/nyc-reopening-hits-highest-gear-yet-as-shops-outdoor-dining-and-playgrounds-return-monday/2477627/`, accessed 2020-06-22.

Office of the Texas Governor (2020a). A proclamation amending Executive Order GA-28 by Office of the Texas Governor. URL: `https://open.texas.gov/uploads/files/organization/opentexas/DISASTER-proclamation-amending-GA-28-mass-gatherings-IMAGE-07-02-2020.pdf`, accessed 2020-07-02.

Office of the Texas Governor (2020b). Executive Order GA-11 by Office of the Texas Governor. URL: `https://www.nbcnewyork.com/news/local/nyc-reopening-hits-highest-gear-yet-as-shops-outdoor-dining-and-playgrounds-return-monday/2477627/`, accessed 2020-03-26.

Office of the Texas Governor (2020c). Executive Order GA-24 by Office of the Texas Governor. URL: `https://gov.texas.gov/uploads/files/press/DISASTER_Adding_Covered_Services_to_GA-23_No_2_COVID-19.pdf`, accessed 2020-05-26.

Office of the Texas Governor (2020d). Executive Order GA-28 by Office of the Texas Governor. URL: `https://gov.texas.gov/uploads/files/press/EO-GA-28_targeted_response_to_reopening_COVID-19.pdf`, accessed 2020-06-26.

Office of the Texas Governor (2020e). Executive Order GA-29 by Office of the Texas Governor. URL: `https://open.texas.gov/uploads/files/organization/opentexas/EO-GA-29-use-of-face-coverings-during-COVID-19-IMAGE-07-02-2020.pdf`, accessed 2020-07-02.

Piggott J, McLean J (2020). Jacksonville changes course, issues face mask mandate. *News4JAX*. URL: `https://www.news4jax.com/news/local/2020/06/29/jacksonville-issues-face-mask-mandate/`, accessed 2020-06-29.

Sandhya Kambhampati MM, Krishnakumar P (2020). Which California counties are reopening? *Los Angeles Times*. URL: `https://www.latimes.com/projects/california-coronavirus-cases-tracking-outbreak/reopening-across-counties/`.

Silcox F, Turco R (2020). Florida Businesses Ready for Full Phase One Reopening. *Bay News*. URL: `https://www.baynews9.com/fl/tampa/coronavirus/2020/05/18/ongoing-coronavirus-coverage`, accessed 2020-05-18.

Speck E, Sandoval E (2020). Orange County residents will be required to wear face masks under new executive order. *Click Orlando*. URL: `https://www.clickorlando.com/news/local/2020/06/18/orange-county-residents-will-be-required-to-wear-face-masks-under-new-executive-order/`, accessed 2020-06-18.

The City Manager of the City of Miami (2020). City Mandates Facial Coverings in Public; Civil Penalties Approved. URL: `https://www.miamigov.com/files/sharedassets/public/news/2020/0625-emergency-order-20-16.pdf`, accessed 2020-06-25.

Zhou Q, Guan Y (2018). On the null distribution of bayes factors in linear regression. *Journal of the American Statistical Association*, 113(523): 1362–1371.