Bayesian Hierarchical Modeling of COVID-19 Cases and Government Response in the United States

Tianshu Liu, Jiong Ma, Jiayi Shi, Zhengwei Song, Ziqing Wang

Outline

- Background
  - data description
- Bayesian hierarchical model
  - loglikelihood & logprior
- Component-wise Metropolis-Hastings (CMH) algorithm
  - window length adjustment
  - diagnostic plots
- Results
  - estimates and uncertainty of parameters
- Recommendations

Background

- COVID-19 has disrupted public health, economics, and society globally.
- Understanding virus spread and government responses is crucial, particularly in the US.
- Bayesian hierarchical modeling will analyze COVID-19 data and government response indexes.
- Demographic factors like population density and elderly population will be examined.
- The project aims to improve our understanding of COVID-19's impact on public health.
- The findings can inform future policy decisions and response strategies.

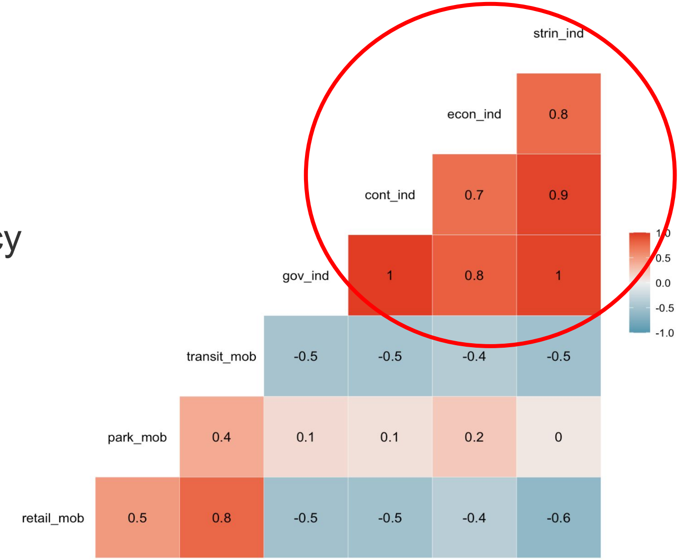
Data description

- Variables
  - weekly state-level COVID-19 case counts in 2020
  - government response indices (0-100): Government Response Index, Containment Health Index, Economic Support Index, and Stringency Index.
  - % mobility changes: retail, parks, and transit stations
  - state-level demographic information: population density, % elderly

Statistic before scaling	N	Mean	St. Dev.	Min	Max
retail_and_recreation_percent_change_from_baseline	2,327	−10.187	14.819	−59.877	23.707
parks_percent_change_from_baseline	2,327	46.095	61.362	−70.286	407.000
transit_stations_percent_change_from_baseline	2,327	−12.803	18.451	−77.714	55.250
population_density	2,327	206.017	270.013	1.280	1,254.244
government_response_index	2,327	54.757	16.550	9.006	80.210
containment_index	2,327	55.242	15.770	10.290	79.640
economic_support_index	2,327	51.358	27.870	0.000	100.000
stringency_index	2,327	57.465	18.656	6.744	93.520

Data processing

- Variable inspection
  - Transformed weekly state-level cumulative case counts to new case counts
- Variable selection
  - High correlation between government response index, containment index, economic support index, and stringency index (circled in red) → kept government response index only to reduce collinearity
- Variable scaling
  - Variables in different units and magnitudes → scaled each by dividing its standard deviation
  - Scaled weekly state-level case counts to per 1,000,000



Bayesian hierarchical model

- Poisson Model:

$$Y_{ij} \sim Poisson(\lambda_{ij} n_{ij})$$

$$P(Y_{ij} = k) = \frac{e^{-\lambda_{ij} n_{ij}} (\lambda_{ij} n_{ij})^k}{k!}$$

$$log(\lambda_{ij}) = \alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon_{ij}$$

Y_{ij} :the number of new infections per million people in state i during week j

n_{ij}: the population(millions) of state i during week j

λ_{ij} : the infection rate in state *i* during week j

X_{ij} : the covariates (government response index, weekly average percentage change in mobility trends for retail and recreation places, parks, and transit stations in each state)

P_{ij}: the population density

E_{ij} : the percentage of the elderly population

Bayesian hierarchical model

- Choice of Priors:

fixed effects: $\beta_k \sim Normal(0, 10^2)$ where k = 1, 2, 3, 4

$\gamma \sim Normal(0, 10^2)$

$\delta \sim Normal(0, 10^2)$

intercept: $\alpha \sim Normal(0, 10^2)$

random effects: $u_i \sim Normal(0, \sigma_u^2)$ where i = 1, 2, …, 50

common variance for the random effects: $\sigma_u \sim HalfNormal(0, 10^2) \quad or \quad \sigma_u \sim HalfCauchy(0, 10)$

residual error term: $\epsilon_{ij} \sim Normal(0, \sigma_\epsilon^2)$

common variance for the residual error term: $\sigma_\epsilon \sim HalfNormal(0, 10^2) \quad or \quad \sigma_\epsilon \sim HalfCauchy(0, 10)$

Bayesian hierarchical model

$$\Theta = \{\alpha, \beta_1, ..., \beta_4, \gamma, \delta, u_1, ..., u_{50}, \epsilon, \sigma_u, \sigma_\epsilon\}$$

- Prior:

$$\pi(\Theta) = \pi(\alpha) \prod_{i=1}^{4} \pi(\beta_i)\pi(\gamma)\pi(\delta) \prod_{i=1}^{n_s} \pi(u_i|\sigma_u)\pi(\sigma_u)\pi(\epsilon|\sigma_\epsilon)\pi(\epsilon)$$

$$\propto exp\{\frac{-\alpha^2 - \sum_{i=1}^{k} \beta_k^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2}\} \frac{1}{\sigma_u \sigma_\epsilon} exp\{-\frac{\sum_{i=1}^{n_s} u_i^2}{2\sigma_u^2}\} exp\{-\frac{\epsilon_i^2}{2\sigma_\epsilon^2}\}$$

- Likelihood:

$$L_Y(\Theta) = \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \frac{(\lambda_{ij}(\Theta)n_{ij})^{Y_{ij}} e^{\lambda_{ij}(\Theta)n_{ij}}}{Y_{ij}!}$$

$$= \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \frac{\{exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}\}^{Y_{ij}} exp\{\lambda_{ij}(\Theta)n_{ij}\}}{Y_{ij}!}$$

$$\propto \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \{exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}\}^{Y_{ij}} exp\{\lambda_{ij}(\Theta)n_{ij}\}$$

- Posterior:

$$g(\Theta|Y) \propto L_Y(\Theta)\pi(\Theta)$$

$$\propto \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \frac{(\lambda_{ij}(\Theta)n_{ij})^{Y_{ij}} e^{\lambda_{ij}(\Theta)n_{ij}}}{Y_{ij}!} exp\{\frac{-\alpha^2 - \sum_{i=1}^{k} \beta_k^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2}\} \frac{1}{\sigma_u \sigma_\epsilon} exp\{-\frac{\sum_{i=1}^{n_s} u_i^2}{2\sigma_u^2}\} exp\{-\frac{\epsilon^2}{2\sigma_\epsilon^2}\}$$

$$\propto \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \{exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}\}^{Y_{ij}} exp\{\lambda_{ij}(\Theta)n_{ij}\}$$

$$\cdot exp\{\frac{-\alpha^2 - \sum_{i=1}^{k} \beta_k^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2}\} \frac{1}{\sigma_u \sigma_\epsilon} exp\{-\frac{\sum_{i=1}^{n_s} u_i^2}{2\sigma_u^2}\} exp\{-\frac{\epsilon^2}{2\sigma_\epsilon^2}\}$$

Component-wise Metropolis-Hastings (CMH) algorithm

- a variant of the Metropolis-Hastings(MH) algorithm, sampling from high-dimensional probability distributions and updates each parameter separately

1. Initialize *M* chains of length *T*, with each chain starting from a different initial value of Θ. For each iteration $t = 1, 2, . . . , T$ and for each chain $m = 1, 2, . . . , M$, randomly select a component *j* from $1, 2, . . . , n$.

2. Propose a new value Θ_j for component *j* of chain *m* using a one-dimensional Metropolis Hastings update. That is, draw a proposal $\Theta_j \sim q(\cdot|\Theta_{(jm)})$, where $q(\cdot|\Theta_{(jm)})$ is a proposal distribution centered at the current value $\Theta_{(jm)}$ of component *j*:

$$q(\cdot|\Theta_{j,t}^{(m)}) = q(\cdot|\Theta_{j-1,t}^{(m)}) + w_j * 2 * (r_{j,t} - 0.5)$$

where w_j is the window length, and, $r_{j,t}$ is a random number follows *Uniform*(0, 1).

3. Compute the acceptance probability

$$a_{j,t}(\Theta_{j,t}^{(m)}, \Theta_{j,t}) = \min\{1, \frac{p(\Theta_{1,t}^{(m)}, \ldots, \Theta_{j-1,t}^{(m)}, \Theta_{j,t}, \Theta_{j+1,t}^{(m)}, \ldots, \Theta_{n}^{(m)})}{p(\Theta_{1,t}^{(m)}, \ldots, \Theta_{n,t}^{(m)})} \frac{q(\Theta_{j,t}^{(m)}|\Theta_{j,t})}{q(\Theta_{j,t}|\Theta_{j,t}^{(m)})}\}$$

where $p(\cdot)$ is the target density of the parameters Θ.

4. Accept the proposed new value Θ_j with probability $a_j(\Theta_{(j,m)}, \Theta_j)$, and set $\Theta_{(j,m+1)} = \lambda_{*j}$ if the proposal is accepted, and $\Theta_{(j,m+1)} = \Theta_{(j,m)}$ otherwise.

5. Repeat steps 2-5 until convergence is achieved.

MCMC Chain

Adjust the window length of random walk:

- final window length:

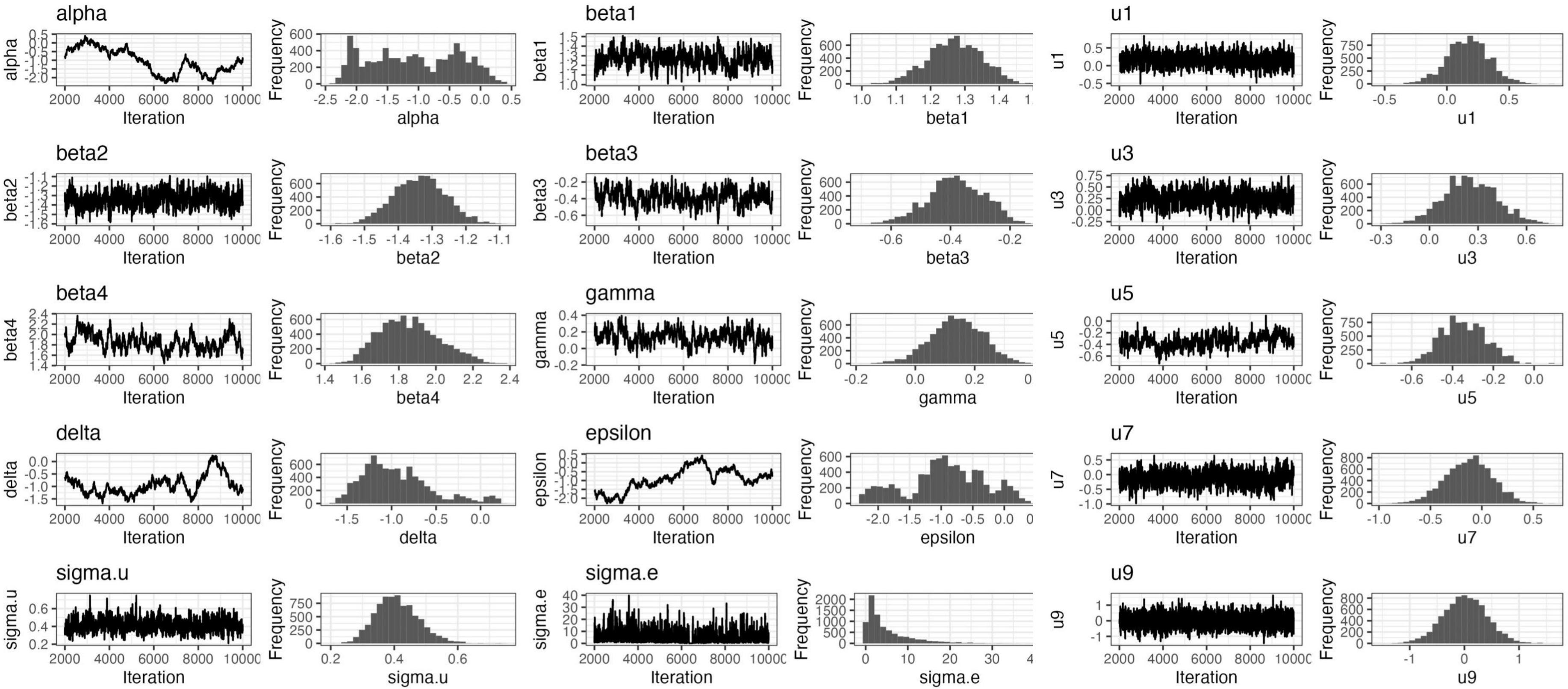
```
> a
 [1]  0.1  0.1  0.1  0.1  0.1  0.1  0.1  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0
[16]  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0
[31]  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0
[46]  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0  0.1  0.1 10.0
```

- most of the acceptance rates are between 30% and 60%

```
> 1-n_reject(chain11)
 [1] 0.3836616 0.4177582 0.5742426 0.3791621 0.3742626 0.3460654 0.3941606
 [8] 0.2562744 0.4522548 0.2160784 0.3278672 0.1122888 0.2685731 0.3230677
[15] 0.4297570 0.5642436 0.1343866 0.1958804 0.4385561 0.4212579 0.1748825
[22] 0.2599740 0.3115688 0.3549645 0.3102690 0.2379762 0.4880512 0.2609739
[29] 0.2428757 0.2292771 0.2701730 0.3150685 0.2551745 0.4767523 0.3993601
[36] 0.3081692 0.4732527 0.2017798 0.3750625 0.1361864 0.2182782 0.4471553
[43] 0.2220778 0.2976702 0.3908609 0.2143786 0.5699430 0.4199580 0.2715728
[50] 0.4975502 0.2146785 0.1271873 0.3090691 0.5129487 0.2534747 0.2882712
[57] 0.4279572 0.3807619 0.5727427 0.3634637
```

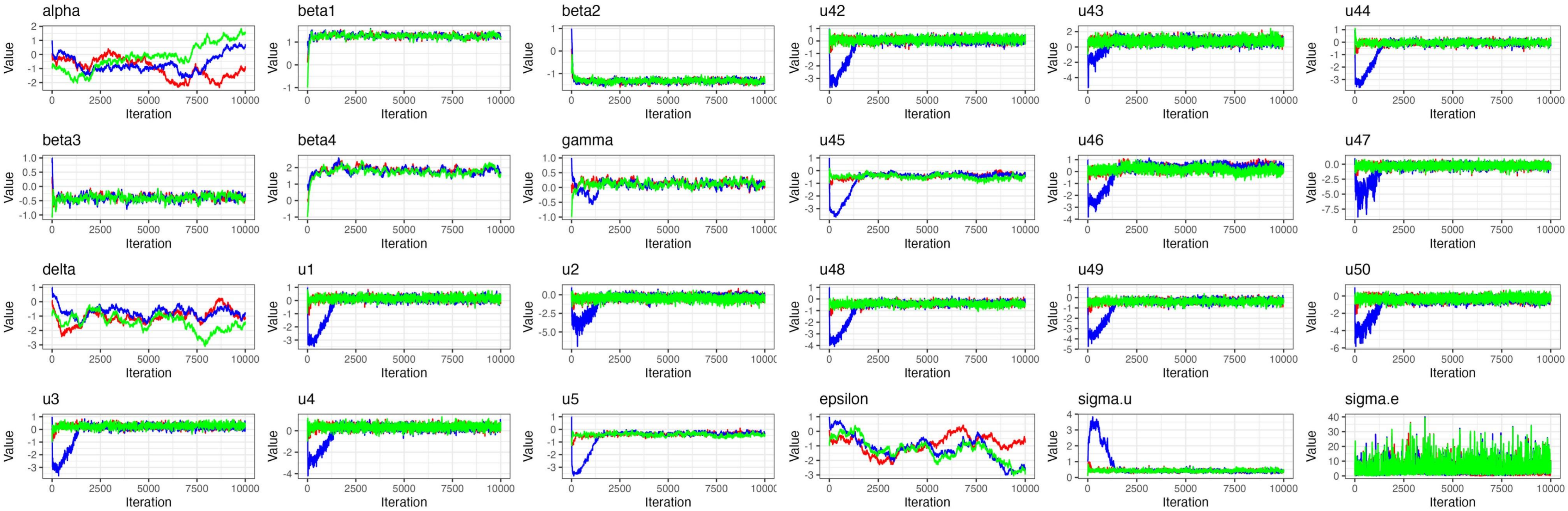
MCMC Chain Convergence

diagnostic plot – trace plot and histogram



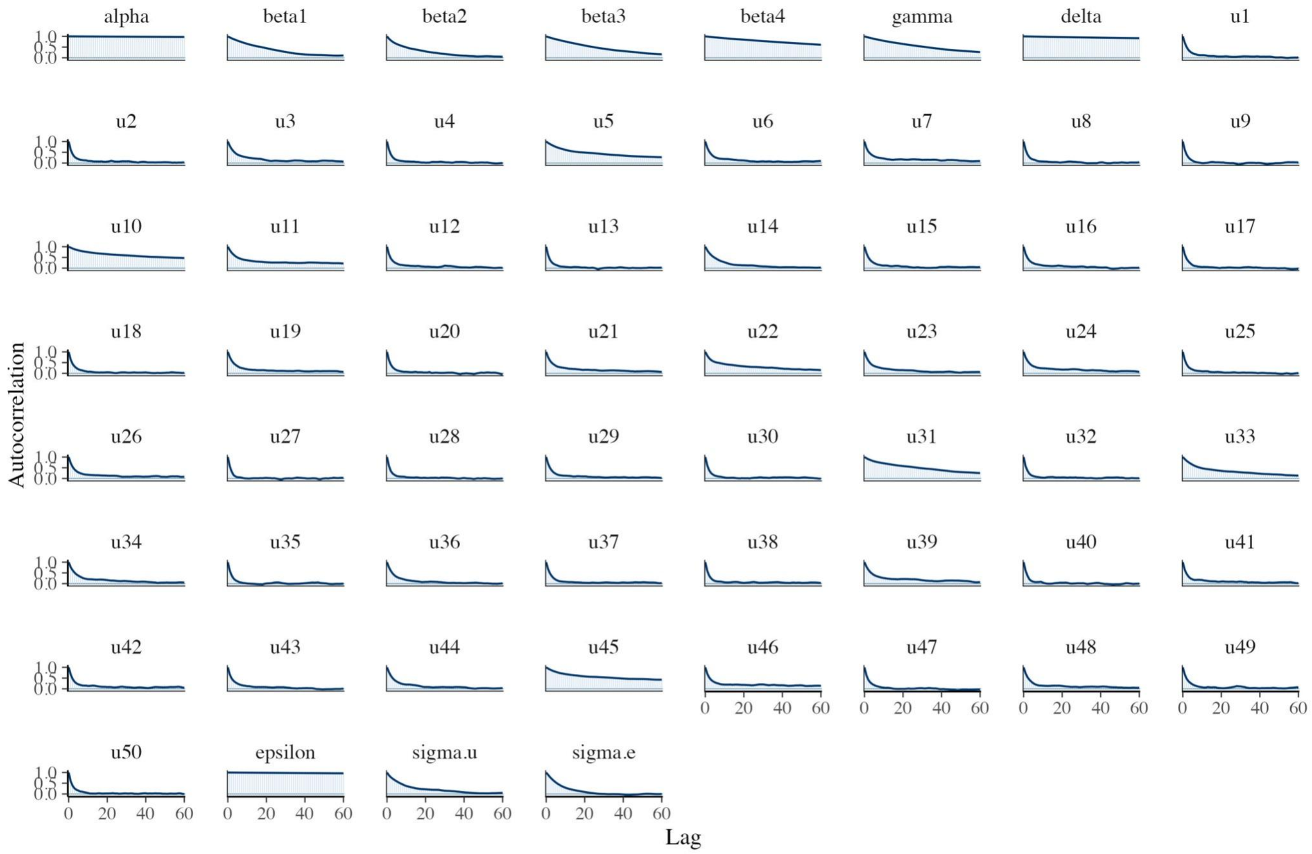
MCMC Chain Convergence

diagnostic plot – merged trace plot of 3 chains



MCMC Chain Convergence

diagnostic plot – autocorrelation plot



Gelman-Rubin Statistics (R-hat)

- The Gelman-Rubin statistic compares the variance of the pooled samples from multiple chains to the average of the variances of each chain separately.

$$\bar{x}_j = \frac{1}{L} \sum t = 1^L x_t^{(j)}$$

$$s_j^2 = \frac{1}{L-1} \sum t = 1^L (x_t^{(j)} - \bar{x}_j)^2$$

$$\bar{x}. = \frac{1}{J} \sum j = 1^J \bar{x}_j$$

$$W = \frac{1}{J} \sum_{j=1}^{J} s_j^2$$

$$B = \frac{L}{J-1} \sum_{j=1}^{J} (\bar{x}_j - \bar{x}.)^2$$

$$R = \frac{\frac{L-1}{L} W + \frac{1}{L} B}{W}$$

MCMC Chain Convergence

Gelman-Rubin Statistics: generate 3 chains from different starting values

parameter	statistics	converge
alpha	1.20277939	FALSE
beta1	1.00573816	TRUE
beta2	1.00323769	TRUE
beta3	1.00275761	TRUE
beta4	1.01825063	TRUE
gamma	1.03939982	TRUE

parameter	statistics	converge
delta	1.3598521	FALSE
u1-50	<1.1	TRUE
u10	1.13927314	FALSE
epsilon	1.04892468	TRUE
sigma.u	1.07219144	TRUE
sigma.e	1.00337833	TRUE

alpha: overall intercept

beta1: change in mobility trend in retail and recreation places

beta2: change in mobility trend in parks

beta3: change in mobility trend in transit stations

beta4: government response index

gamma: coefficient for state population density

delta: coefficient for state % elderly

u: state-level random effect

epsilon: residual error term

sigma.u: standard deviation of u

sigma.e: standard deviation of epsilon

Results - Posterior distribution

Table 1: Posterior summaries

Statistic	N	Mean	St. Dev.	Min	Max
beta1	5,001	1.264	0.077	1.008	1.500
beta2	5,001	−1.330	0.076	−1.576	−1.090
beta3	5,001	−0.386	0.095	−0.702	−0.124
beta4	5,001	1.823	0.149	1.437	2.301
gamma	5,001	0.136	0.089	−0.182	0.377
delta	5,001	−0.824	0.414	−1.631	0.246

- **beta1**: change in mobility trend in retail and recreation places
- **beta2**: change in mobility trend in parks
- **beta3**: change in mobility trend in transit stations
- **beta4**: government response index
- **gamma**: coefficient for state population density
- **delta**: coefficient for state elderly percentage

Results - Credible intervals

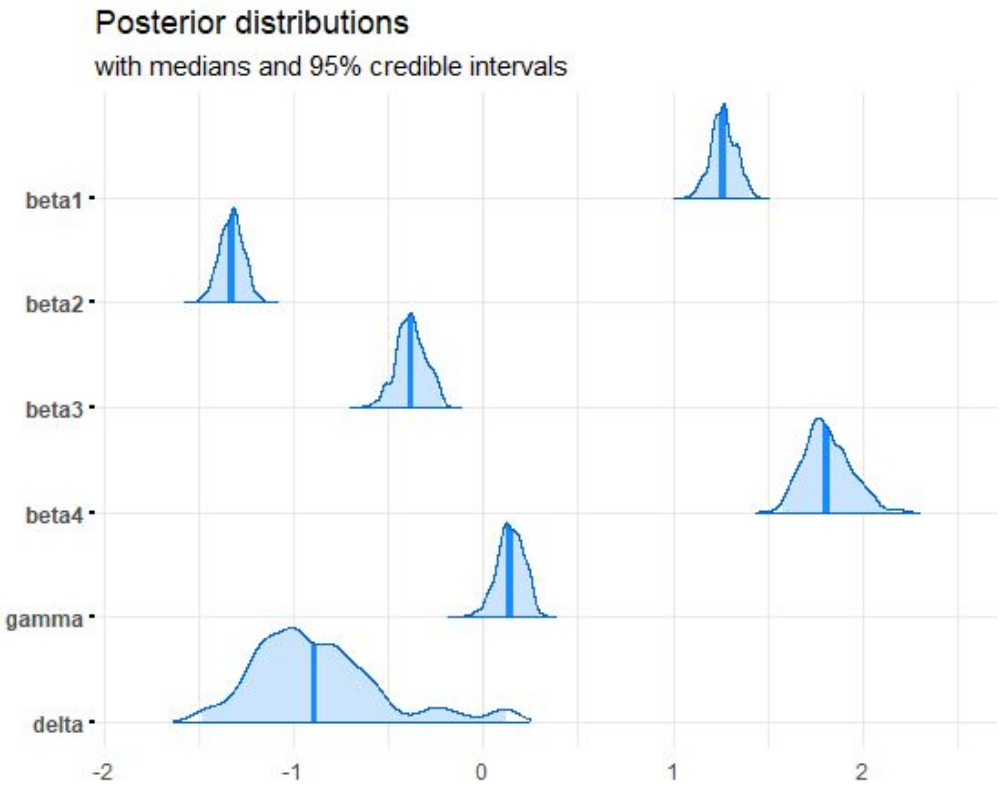


Table 2: 95% credible intervals

	2.5%	97.5%
beta1	1.111	1.411
beta2	−1.478	−1.178
beta3	−0.579	−0.217
beta4	1.575	2.146
gamma	−0.053	0.296
delta	−1.478	0.123

Discussion

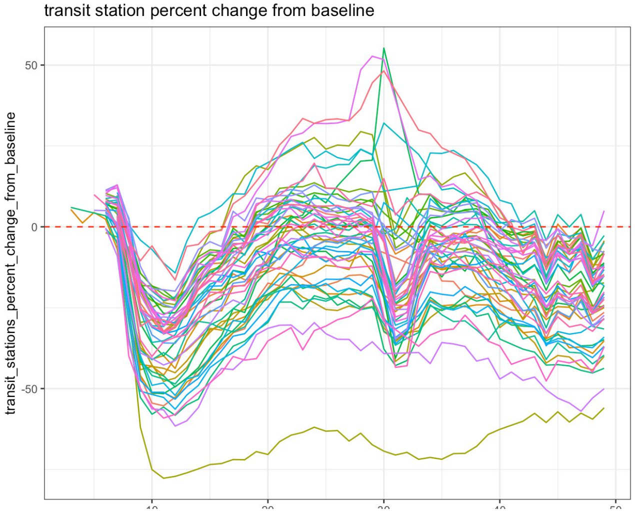
Modeling - What affects model convergence?

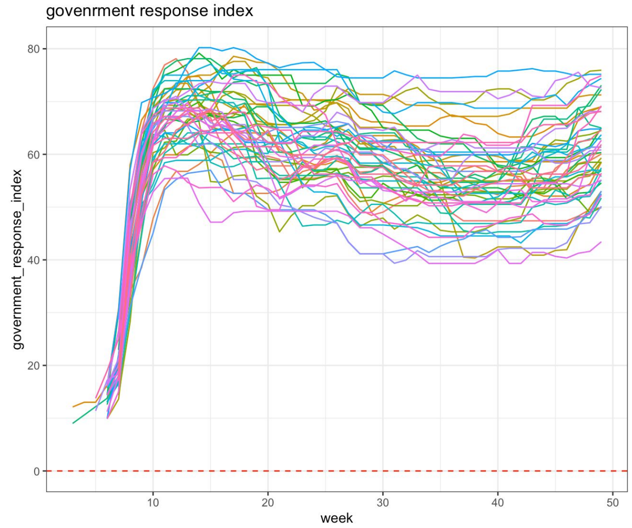
- Collinearity weakens MCMC parameter convergence
    a. Poor convergence of fixed effects when all highly correlated indices were included in the model
- Scaling the predictors helps the MCMC chains to converge

Discussion

Interpretation - Some counterintuitive final results?

- Fitted the lmer() equivalent of the model and got similar parameter estimates
- Variational effects of changes in mobility trends?
  - Association != causation: Outbreak -> decreasing outdoor activities/travelling and increasing remote work, resulting in negative association
  - However, retail activities, e.g., grocery shopping, are essential for most, even during an outbreak, resulting in positive association
- Positive association between infected cases and government response index?
  - Potential explanation: The worse the outbreak, the higher level of government response.
- Non-significant association between infected cases and population density & % elderly?
  - Potential explanation: The relationship between infected cases and these two variables may be confounded by other factors, such as the economy of the state.





Public Health Policy Recommendation

1. Enhance government stringency, restrict unnecessary mobility at early stages of the outbreak
    a. waiting until the outbreak has spreaded may render restrictions less effective
2. Apply COVID test, especially to less economically developed states, to get more accurate COVID case count data
    a. recorded number of infections might not reflect the reality accurately because of differential access/distribution of COVID tests among different states
3. Reasearch on more confounders between infected cases and relevant indices, population density, and demongraphics
    a. To develop a better statistical model to infer associations
    b. To quantify the effect of more factors on the infection rate

Conclusion

1. Write down log posterior density function for 60 parameters.
2. Monitor and diagnosis MCMC chain convergence.
3. Construct 95% credible intervals for parameters and give interpretation
4. Share public health policy recommendations.

Reference

[1] Stephen P. Brooks and Andrew Gelman. “General Methods for Monitoring Convergence of IterativeSimulations”. In: *Journal of Computational and Graphical Statistics* 7.4 (1998), pp. 434–455.

[2] Andrew Gelman and Donald B. Rubin. “Inference from Iterative Simulation Using Multiple Sequences”.In: *Statistical Science* 7.4 (1992), pp. 457–472.