

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

Zhengwei Song

June 2023

Contents

1	Background	1
2	Data Processing	1
3	Methods	2
3.1	Bayesian Hierarchical Model	2
3.1.1	Hierarchical Structure	2
3.1.2	Choices of Priors	3
3.1.3	Likelihood	3
3.1.4	Posterior	4
3.2	Component-wise MH Algorithm	4
3.3	MCMC Chain Convergence Diagnostics	5
3.3.1	Diagnostic Plots	5
3.3.2	Gelman-Rubin Statistic	6
4	Results	7
4.1	Convergence of MCMC Chain	7
4.2	Posterior Distribution	7
4.3	Credible Intervals	7
5	Discussion	7
5.1	Interpretation	7
5.2	Public Health Policy Recommendations	8
5.3	Suggestions on Further Research	8
6	Conclusion	9
Reference		9
Appendix		10

1 Background

The COVID-19 pandemic has had a profound impact on the world, leading to significant disruptions in public health, economic stability, and social life. In the United States, the pandemic has underscored the importance of comprehending the spread of the virus and evaluating the effectiveness of various government responses in mitigating its impact. Throughout the course of the pandemic, different states have implemented a wide array of containment measures, economic support policies, and public health interventions. Understanding the relationship between these factors and the spread of COVID-19 is essential for informing future policy decisions and improving public health outcomes.

The objective of this project is to utilize Bayesian hierarchical modeling techniques to analyze COVID-19 case data and various government response indexes in the United States during the year 2020, prior to the availability of vaccines. The government response indexes encompass containment measures, economic support policies, and stringency levels. Additionally, this analysis investigates the influence of demographic factors, such as population density and the elderly population, on the spread of COVID-19 and the effectiveness of government responses.

Furthermore, this project aims to gain a comprehensive understanding of the factors that impact the pandemic's effect on public health. Ultimately, the insights derived from this analysis can assist in guiding future policy decisions and developing effective response strategies in the face of ongoing and future public health emergencies.

2 Data Processing

The dataset used for this study was created by merging multiple sources of information to facilitate the Bayesian hierarchical modeling of COVID-19 cases and government responses in the United States. It consists of weekly aggregated data for the year 2020, including state-level COVID-19 case counts, government response indexes (Government Response Index, Containment Health Index, Economic Support Index, and Stringency Index), and mobility changes in retail, parks, and transit stations. Additionally, the dataset incorporates state-level demographic information, such as population density and the proportion of the elderly population. This comprehensive dataset provides a detailed overview of the pandemic landscape in the United States during 2020.

The original dataset comprises 2548 observations and 15 variables, which are as follows: state, week start, weekly cases, pop2019, LandArea, Percentage over 65, infection rate, retail and recreation percent change from baseline, parks percent change from baseline, transit stations percent change from baseline, population density, government response index, containment index, economic support index, and stringency index. Before scaling, the summary statistics of the variables are presented in Table 1.

The data processing step is conducted in three directions, as outlined below:

1. Variable Inspection: The weekly state-level cumulative case counts are transformed into new case counts to align with the assumption of a Poisson model. This transformation allows for more appropriate modeling of the data.
2. Variable Selection: Only four variables are chosen as covariates for analysis. These variables include the Government Response Index and the weekly average percentage change in mobility trends for Retail and Recreation Places, Parks, and Transit Stations in each state. The selection of these variables is based on observed high correlations among several variables, as depicted in Figure 1. Variables such as the Containment Health Index, Economic Support Index, and Stringency Index exhibit high correlation with the Government Response Index. Including highly correlated variables in the analysis can lead to longer computation times and potential convergence issues in Bayesian

hierarchical models. Furthermore, it is discovered that the Containment Health Index, Economic Support Index, and Stringency Index variables can be adequately represented using the Government Response Index.

3. Variable Scaling: The state-level COVID-19 case counts and Population in 2019 are measured in units of millions. To scale the remaining predictors, they are divided by their respective standard deviations. The purpose of variable scaling is to ensure equal weighting of all predictors in the model, enabling a fairer comparison among them. Scaling the predictors allows for a more meaningful assessment of their relative importance, regardless of their original units or scales. Furthermore, scaling can enhance the numerical stability of the model. Many numerical optimization algorithms employed in Bayesian hierarchical models, such as Markov Chain Monte Carlo (MCMC), are sensitive to the scale of predictors. Scaling the predictors mitigates this sensitivity, improving the convergence and efficiency of the optimization process.

3 Methods

3.1 Bayesian Hierarchical Model

A Bayesian hierarchical model is a statistical framework that allows for the modeling of complex data structures by incorporating multiple levels of variation. This type of model assumes that the observed data is generated from a hierarchical structure, where each level of the hierarchy corresponds to a different level of variation in the data.

The parameters of the model are estimated by fitting the model to the observed data using Bayesian inference:

$$g(\lambda|Y) = \frac{f(Y|\lambda)\pi(\lambda)}{f(Y)}$$

This involves specifying a likelihood function $f(Y|\lambda)$ that describes the probability of observing the data given the model parameters, as well as prior distributions $\pi(\lambda)$ that describe the prior knowledge about the model parameters. The posterior distribution of the parameters $g(\lambda|Y)$ is then obtained. This posterior distribution can be used to make predictions about future observations, as well as to estimate the uncertainty in the model parameters.

Bayesian hierarchical models allow for the incorporation of prior information into the analysis, can handle complex data structures with multiple levels of variation, and can provide more accurate estimates of uncertainty in the model parameters.

In this study, I assume that the number of new infections in state i during week j , denoted as Y_{ij} , follows a Poisson distribution:

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

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

where λ_{ij} is the infection rate in state i during week j , and n_{ij} is the population of state i during week j .

3.1.1 Hierarchical Structure

I model the log of infection rates $\log(\lambda_{ij})$ using a linear model that includes state-level random effects and fixed effects for the covariates:

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

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

where α is the overall intercept, β is a vector of fixed effects coefficients for the covariates X_{ij} (e.g., government response index, and mobility changes), γ is the fixed effect coefficient

for the population density P_{ij} , δ is the fixed effect coefficient for the percentage of the elderly population E_{ij} , u_i is the state-level random effect, and ϵ_{ij} is the residual error term.

3.1.2 Choices of Priors

I choose weakly informative priors for the fixed effects coefficients β_k , γ , and δ , such as normal distributions with mean 0 and a large variance:

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

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

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

For the overall intercept α , I use a weakly informative prior:

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

For the state-level random effects u_i , I assume that they follow a normal distribution with mean 0 and a common variance σ_u^2 :

$$u_i \sim Normal(0, \sigma_u^2)$$

I choose a weakly informative prior for the standard deviation σ_u :

$$\sigma_u \sim HalfNormal(0, 10^2) \quad or \quad \sigma_u \sim HalfCauchy(0, 10)$$

For the residual error term ϵ_{ij} , I assume that it follows a normal distribution with mean 0 and a common variance σ_ϵ^2 :

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

I choose a weakly informative prior for the standard deviation σ_ϵ :

$$\sigma_\epsilon \sim HalfNormal(0, 10^2) \quad or \quad \sigma_\epsilon \sim HalfCauchy(0, 10)$$

Let the parameter space Θ to denote the set of parameters. Thus,

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

Let n_s denote the total number of states, the prior can be calculated using:

$$\begin{aligned} \pi(\Theta) &= \pi(\alpha) \prod_{i=1}^4 \pi(\beta_i) \pi(\gamma) \pi(\delta) \prod_{i=1}^{n_s} \pi(u_i | \sigma_u) \pi(\sigma_u) \pi(\epsilon | \sigma_\epsilon) \pi(\epsilon) \\ &\propto \exp\left\{-\frac{\alpha^2 - \sum_{i=1}^k \beta_i^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2}\right\} \frac{1}{\sigma_u \sigma_\epsilon} \exp\left\{-\frac{\sum_{i=1}^{n_s} u_i^2}{2 \sigma_u^2}\right\} \exp\left\{-\frac{\epsilon^2}{2 \sigma_\epsilon^2}\right\} \end{aligned}$$

$$\begin{aligned} log\pi(\Theta) &= log\pi(\alpha) + \sum_{i=1}^4 log\pi(\beta_i) + log\pi(\gamma) + log\pi(\delta) + \sum_{i=1}^{n_s} \pi(u_i | \sigma_u) \\ &\quad + log\pi(\sigma_u) + log\pi(\epsilon | \sigma_\epsilon) + log\pi(\epsilon) \\ &\propto \frac{-\alpha^2 - \sum_{i=1}^k \beta_i^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2} - log(\sigma_u) - log(\sigma_\epsilon) - \frac{\sum_{i=1}^{n_s} u_i^2}{2 \sigma_u^2} - \frac{\epsilon^2}{2 \sigma_\epsilon^2} \end{aligned}$$

3.1.3 Likelihood

Let n_s denote the total number of states and n_w denote the total number of weeks. Since the number of new infections in state i during week j , denoted as y_{ij} , follows a Poisson distribution: $y_{ij} \sim Poisson(\lambda_{ij} n_{ij})$ where λ_{ij} is the infection rate in state i during week j , and n_{ij} is the population of state i during week j , the likelihood can be calculated as:

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

$$\begin{aligned}
L_Y(\Theta) &= \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \frac{(\lambda_{ij}(\Theta)n_{ij})^{Y_{ij}} e^{\lambda_{ij}(\Theta)n_{ij}}}{Y_{ij}!} \\
&= \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \frac{\{\exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}\}^{Y_{ij}} \exp\{\lambda_{ij}(\Theta)n_{ij}\}}{Y_{ij}!} \\
&\propto \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \{\exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}\}^{Y_{ij}} \exp\{\lambda_{ij}(\Theta)n_{ij}\} \\
logL_Y(\Theta) &= \sum_{i=1}^{n_s} \sum_{j=1}^{n_w} Y_{ij} \log(\lambda_{ij}(\Theta)n_{ij}) + \lambda_{ij}(\Theta)n_{ij} - \log(Y_{ij}!) \\
&\propto \sum_{i=1}^{n_s} \sum_{j=1}^{n_w} Y_{ij} \log(\lambda_{ij}(\Theta)n_{ij}) + \lambda_{ij}(\Theta)n_{ij} \\
&\propto \sum_{i=1}^{n_s} \sum_{j=1}^{n_w} Y_{ij} (\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon + \log n_{ij}) + \exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}
\end{aligned}$$

3.1.4 Posterior

The posterior distribution is proportional to the product of the likelihood and prior:

$$\begin{aligned}
g(\Theta|Y) &\propto L_Y(\Theta)\pi(\Theta) \\
&\propto \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \frac{(\lambda_{ij}(\Theta)n_{ij})^{Y_{ij}} e^{\lambda_{ij}(\Theta)n_{ij}}}{Y_{ij}!} \exp\left\{-\frac{\alpha^2 - \sum_{i=1}^k \beta_k^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2}\right\} \frac{1}{\sigma_u \sigma_\epsilon} \exp\left\{-\frac{\sum_{i=1}^{n_s} u_i^2}{2\sigma_u^2}\right\} \exp\left\{-\frac{\epsilon^2}{2\sigma_\epsilon^2}\right\} \\
&\propto \prod_{i=1}^{n_s} \prod_{j=1}^{n_w} \{\exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij}\}^{Y_{ij}} \exp\{\lambda_{ij}(\Theta)n_{ij}\} \\
&\quad \cdot \exp\left\{-\frac{\alpha^2 - \sum_{i=1}^k \beta_k^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2}\right\} \frac{1}{\sigma_u \sigma_\epsilon} \exp\left\{-\frac{\sum_{i=1}^{n_s} u_i^2}{2\sigma_u^2}\right\} \exp\left\{-\frac{\epsilon^2}{2\sigma_\epsilon^2}\right\}
\end{aligned}$$

$$\begin{aligned}
logg(\Theta|Y) &\propto logL_Y(\Theta) + log\pi(\Theta) \\
&\propto \sum_{i=1}^{n_s} \sum_{j=1}^{n_w} Y_{ij} (\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon + \log n_{ij}) + \exp(\alpha + \beta X_{ij} + \gamma P_{ij} + \delta E_{ij} + u_i + \epsilon)n_{ij} \\
&\quad + \frac{-\alpha^2 - \sum_{i=1}^k \beta_k^2 - \sigma_u^2 - \sigma_\epsilon^2}{2 \cdot 10^2} - \log(\sigma_u) - \log(\sigma_\epsilon) - \frac{\sum_{i=1}^{n_s} u_i^2}{2\sigma_u^2} - \frac{\epsilon^2}{2\sigma_\epsilon^2}
\end{aligned}$$

3.2 Component-wise MH Algorithm

The Component-wise Metropolis-Hastings (CMH) algorithm is a specialized version of the Metropolis-Hastings (MH) algorithm, designed specifically for sampling from high-dimensional probability distributions. Unlike the traditional MH algorithm, the CMH algorithm updates each component of a high-dimensional parameter vector individually, rather than updating the entire vector simultaneously.

The core concept of the CMH algorithm involves treating each component of the parameter vector as a separate one-dimensional distribution. In each iteration of the algorithm, a random component is selected, and its value is updated using a one-dimensional MH proposal. This means that only one component is changed at a time while keeping all other components fixed. The acceptance probability for the proposed update is computed based on the ratio of the target density at the proposed value and the current value of the selected component.

The CMH algorithm is particularly useful in situations where updating the entire vector jointly is computationally expensive or impractical, such as in Bayesian inference for complex models with high-dimensional parameters.

In this study, there is a high-dimensional parameter vector

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

The CMH algorithm proceeds as follows:

1. Initialize M chains of length T , with each chain starting from a different initial value of Θ . For each iteration $t = 1, 2, \dots, T$ and for each chain $m = 1, 2, \dots, M$, randomly select a component j from $1, 2, \dots, n$.
2. Propose a new value $\Theta_{j,t}^{(m)}$ for the t -th iteration component j of chain m using a one-dimensional Metropolis-Hastings update. That is, draw a proposal $\Theta_{j,t}^{(m)} \sim q(\cdot | \Theta_{j,t}^{(m)})$, where $q(\cdot | \Theta_{j,t}^{(m)})$ is a proposal distribution centered at the current value $\Theta_{j,t}^{(m)}$ 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

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

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

4. Accept the proposed new value $\Theta_{j,t}^{(m)}$ with probability $a_{j,t}(\Theta_{j,t}^{(m)}, \Theta_{j,t})$, and set $\Theta_{j,t}^{(m+1)} = \Theta_{j,t}^*$ if the proposal is accepted, and $\Theta_{j,t}^{(m+1)} = \Theta_{j,t}^{(m)}$ otherwise.
5. Repeat steps 2-5 until convergence is achieved.

The convergence of the CMH algorithm can be assessed using the same diagnostics as for standard MCMC algorithms, such as examining the trace plots, autocorrelation plots, and the Gelman-Rubin statistic.

3.3 MCMC Chain Convergence Diagnostics

3.3.1 Diagnostic Plots

The trace plot, also known as a chain plot, is a graphical representation of parameter values over iterations in an MCMC chain. It demonstrates how the parameter value changes over time and provides insights into the convergence and mixing of the chain. Ideally, the trace plot should exhibit a sequence of random and independent samples from the posterior distribution, indicating that the chain has successfully converged and is mixing well. However, if the trace plot displays patterns or trends, such as long runs of increasing or decreasing values or oscillations around a specific value, it suggests that the chain has not yet converged or is not mixing well. These patterns can help identify potential issues in the sampling process, such as inappropriate step sizes or high correlations, which can be addressed to improve the efficiency and accuracy of the sampling process.

In an MCMC chain, the histogram is another graphical representation that illustrates the distribution of samples drawn from the posterior distribution. It provides a visual depiction of the parameter's posterior distribution and can aid in diagnosing convergence and mixing issues within the MCMC chain. Ideally, the histogram should display a smooth, unimodal, and bell-shaped distribution without noticeable outliers. However, if the histogram reveals multiple modes or significant skewness, it may indicate that the chain has not yet converged or is not mixing well.

The autocorrelation plot is another valuable diagnostic tool that illustrates the correlation between parameter values at one point in the chain and consecutive samples within the chain. An ideal autocorrelation function rapidly decreases to zero as the lag increases, indicating that the parameter values are uncorrelated, and the chain has converged to the target distribution. If the autocorrelation function remains high for several lags, it suggests poor mixing within the chain, with highly correlated samples. In such cases, additional adjustments should be made to address the issue.

3.3.2 Gelman-Rubin Statistic

The Gelman-Rubin statistic, also known as R-hat, is a convergence diagnostic statistic widely used in MCMC simulations. It helps to assess the convergence of multiple chains run in parallel by comparing the within-chain and between-chain variances and conducting a family of tests.

The calculation of statistic is defined as[1]:

Let $x_1^{(j)}, x_2^{(j)}, \dots$ be samples from the j th Markov chain, and suppose there are J chains run in parallel with different starting values.

1. For each chain, first discard D values as "burn-in" and keep the remaining L values, $x^{(j)}D, x^{(j)}D + 1, \dots, x_{D+L-1}^{(j)}$. For example, you might set $D = L$.
2. Formulas

Chain mean:

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

Grand mean:

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

Between chain variance:

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

Within-chain variance:

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

and

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

3. The Gelman-Rubin statistic is then

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

I can see that as $L \rightarrow \infty$ and as $B \rightarrow 0$, R approaches the value of 1. One can then reason that we should run the chains until the value of R is close to 1, say in the neighborhood of 1.1[2].

The Gelman-Rubin statistic is a ratio and hence unit-free, making it a simple summary for any MCMC sampler. In addition, it can be implemented without first specifying a parameter that is to be estimated, unlike Monte Carlo standard errors. Therefore, it can be a useful tool for monitoring a chain before any specific decisions about what kinds of inferences will be made from the model.

4 Results

4.1 Convergence of MCMC Chain

Figure 2-7 show the trace plots and histograms for all parameters using the final window length after adjustment shown in Table 2, indicating good convergence of some parameters, such as β_1 , β_2 , β_3 , β_4 , γ , σs and all the us , which means the samples are likely to be representative of the target distribution and the MCMC chain mixes well.

Figure 8 is the autocorrelation plot for each parameter. In this plot, most parameters have a rapid decrease as the lag increases, such as all the us , which indicates that the samples are independent to the consecutive ones and the Markov chain has mixed well. However, the autocorrelation of some parameters remains horizontal, which suggests that the samples are highly correlated and the chain has not mixed well.

Three chains are generated in parallel using three sets of starting values shown in Table 3. Figure 9-13 shows trace plots for each parameter merged together from three chains. Most of the parameters converge after about 1000 iterations. For these parameters that are able to converge in the end, the selection of starting value does not really influence the final convergence result.

Table 4 shows the Gelman-Rubin statistics calculated based on definitions. Since a statistic below 1.1 indicates convergence of MCMC chain, 57 over 60 parameters converge and only 3 of them do not converge, including α , δ and u_{10} .

4.2 Posterior Distribution

Table 5 gives the posterior summaries for all the parameters of interest. It shows that mobility changes in parks and transit stations and state elderly percentage are negatively associated with the infection rate, while mobility changes in retail and recreation places, government response and state population density are positively associated with the infection rate. The absolute mean value of β_1 , β_2 and β_4 are greater than 1, suggesting that mobility change in retail, recreation places and parks, and the government response index have relatively large effect size compared to the rest.

4.3 Credible Intervals

Bayesian posterior uncertainty intervals, often referred to as credible intervals, are used to approximate the variability of parameters from MCMC draws.

Figure 15 and Table 6 show that the intervals are narrow for β_1 , β_2 , β_3 , β_4 , γ , but wide for δ . It means that the effect of government interventions, mobility changes and population density are more reliable compared with that of the elderly percentage on the infection rate. The increase in the weekly average percentage change in mobility trends for retail and recreation places will increase the number of new infections, while more changes in mobility trends for parks and transit stations tend to lower the infection rate. The credible interval of β_4 and γ contain 0, suggesting that the effects of population density and elderly percentage are not significant. Therefore, we should focus more on government interventions and mobility changes when giving recommendations for public health policies.

5 Discussion

5.1 Interpretation

Regarding model building, I observed that collinearity hampers MCMC parameter convergence. This was evident from the poor convergence of the model when all highly correlated indices were included. However, significant improvement in convergence was achieved when only the government response index was retained. Furthermore, scaling the predictors

proved beneficial in facilitating MCMC chain convergence. In summary, careful inspection and processing of the data are crucial steps in model fitting.

In terms of interpreting the results, I initially noted some counterintuitive findings. However, temporal relationships between the pandemic and the covariates offer potential explanations. For instance, I found variational effects of changes in mobility trends on COVID infection cases. Specifically, I observed a positive association between cases and mobility changes in retail, but a negative association between cases and mobility changes in transit and parks. Analyzing the trends of these mobility changes and COVID cases over weeks sheds light on these associations. Notably, the 2020 COVID pandemic exhibited three waves: the first two were relatively small, occurring approximately from week 8 to week 11 and from week 20 to week 26, while the third wave started around week 36, affecting most states. Examining the corresponding trends in parks and transit mobility, I observed significant decreases in magnitude ($\sim 200\%$ for parks and $\sim 100\%$ for transit) during the third wave. This decrease likely contributed to the negative association between COVID cases and parks and transit mobility. On the other hand, retail mobility experienced substantial increases during the first and second waves, followed by a smaller decrease ($\sim 20\%$) during the third wave. A contextual explanation for this negative association is that during outbreaks, outdoor activities and traveling decreased, and remote work became more prevalent, resulting in reduced parks and transit mobility. Conversely, retail activities like grocery shopping remained essential for most individuals, and some even felt a heightened need to stock up on necessities. In short, the pandemic had varying impacts on essential versus non-essential activities, leading to differential associations between COVID cases and mobility changes.

The positive association between infection and the government response index can be explained in a similar manner. As the outbreak worsened, the level of government response increased, as indicated by the overall rise in the government response index throughout the progression of the pandemic.

Regarding the non-significant association between COVID cases and state-level population density and the percentage of elderly population, other factors such as the state's economy may be confounding the relationship. Adjusting for relevant confounders may help elucidate the relationship between COVID cases and state demographics, providing more informative insights for public health decision-making.

5.2 Public Health Policy Recommendations

Based on the results, I propose several public health policy recommendations for COVID containment. First, it is important to restrict unnecessary mobility at early stages of the outbreak, because waiting until the outbreak has spread may render restrictions less effective, as we have witnessed in 2020. It is also important to improve accessibility to COVID tests, especially in less economically developed states. This is because the recorded number of infections in each state might vary in terms of accuracy because of differential access and distribution of COVID testing. Another recommendation is to research more on confounders between infected cases and relevant indices, as well as state demographics. This helps develop a better statistical model that quantifies the effects of more factors on COVID infection.

5.3 Suggestions on Further Research

The unexpected negative correlation between certain parameter estimates and weekly new case counts suggests the presence of factors that are not fully understood or accounted for in the analysis. To investigate the causal relationships between the disease and the intervention, it is important to consider additional variables that may be influencing the outcome. For instance, examining the time lag in government response could provide valuable insights. Additionally, it is crucial to explore the interactions among different factors to identify potential risks and evaluate the effectiveness of interventions. For example, it would

be beneficial to investigate the relationship between mobility and government response, and how they mutually influence the overall outcome. Overall, further research is necessary to improve public health outcomes and enhance the understanding of disease dynamics and intervention strategies.

6 Conclusion

In the project, I developed a Bayesian hierarchical Poisson regression model by formulating the log of the prior, likelihood, and posterior density for 60 parameters. To ensure adequate model fitting, I employed diagnostic plots to monitor the convergence of the Markov Chain Monte Carlo (MCMC) chains. By constructing 95% credible intervals for all parameters, I was able to provide insightful interpretations and potential explanations, which can inform public health policy recommendations.

References

- [1] Stephen P. Brooks and Andrew Gelman. “General Methods for Monitoring Convergence of Iterative Simulations”. 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.

Appendix

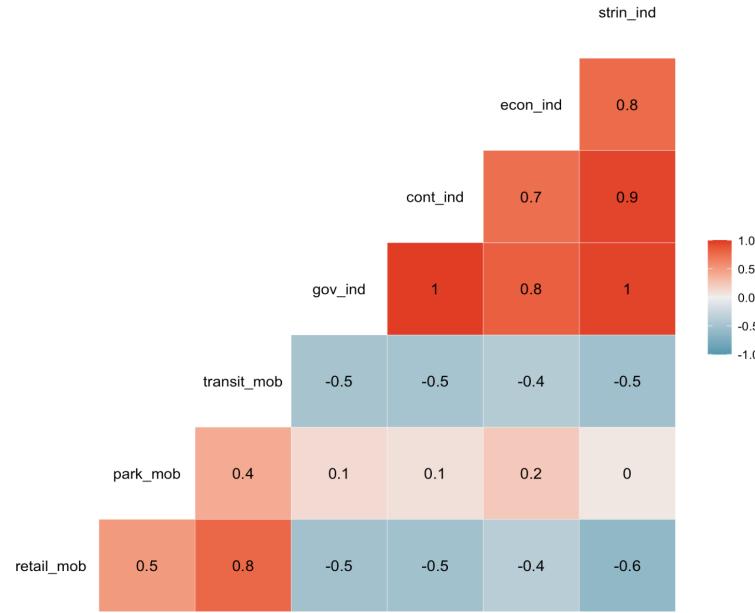


Figure 1: Pair-wise correlations between covariates in the data set

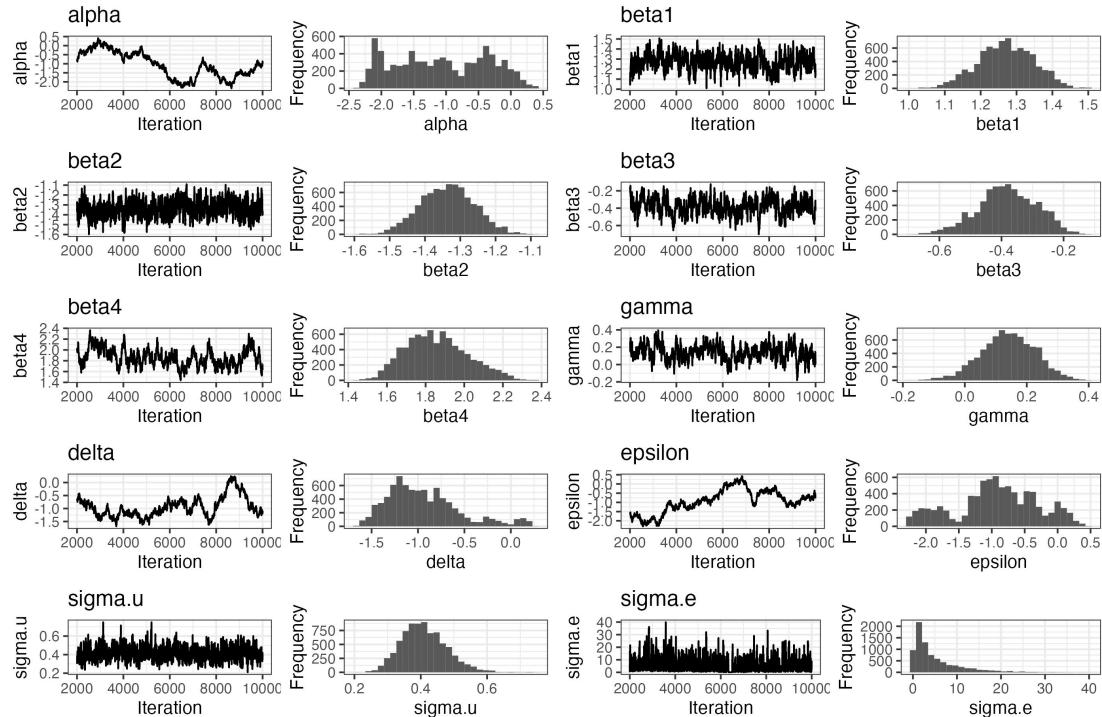


Figure 2: Trace plot for α , β s, γ , δ , ϵ and σ

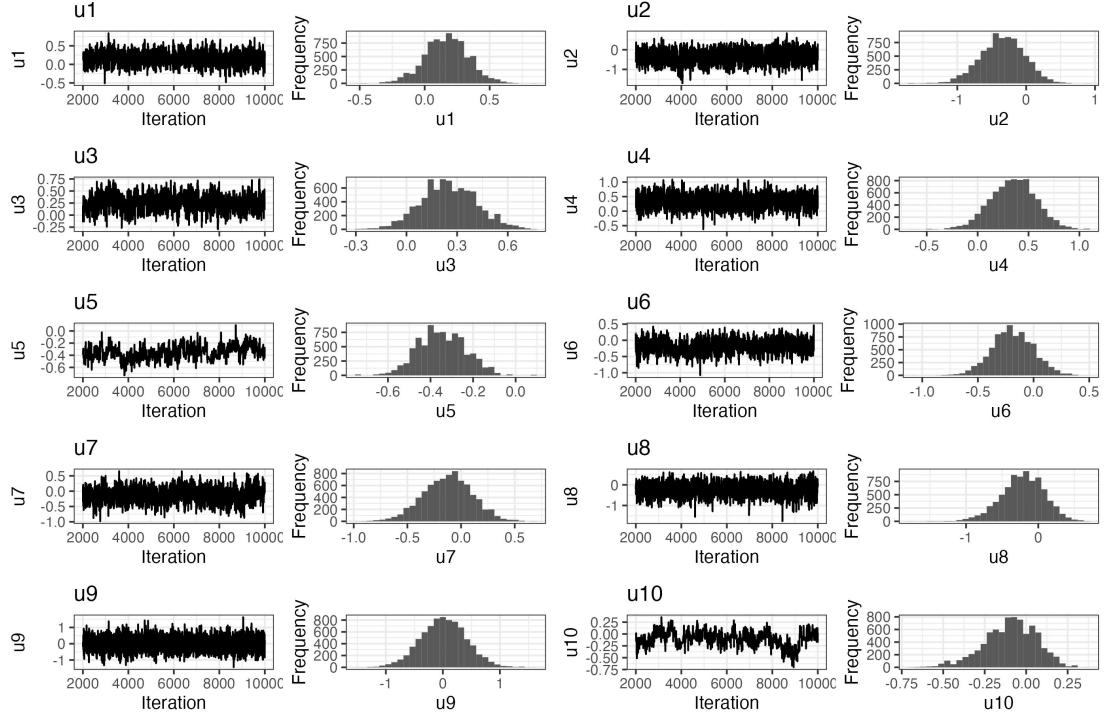


Figure 3: Trace plot for u_1-u_{10}

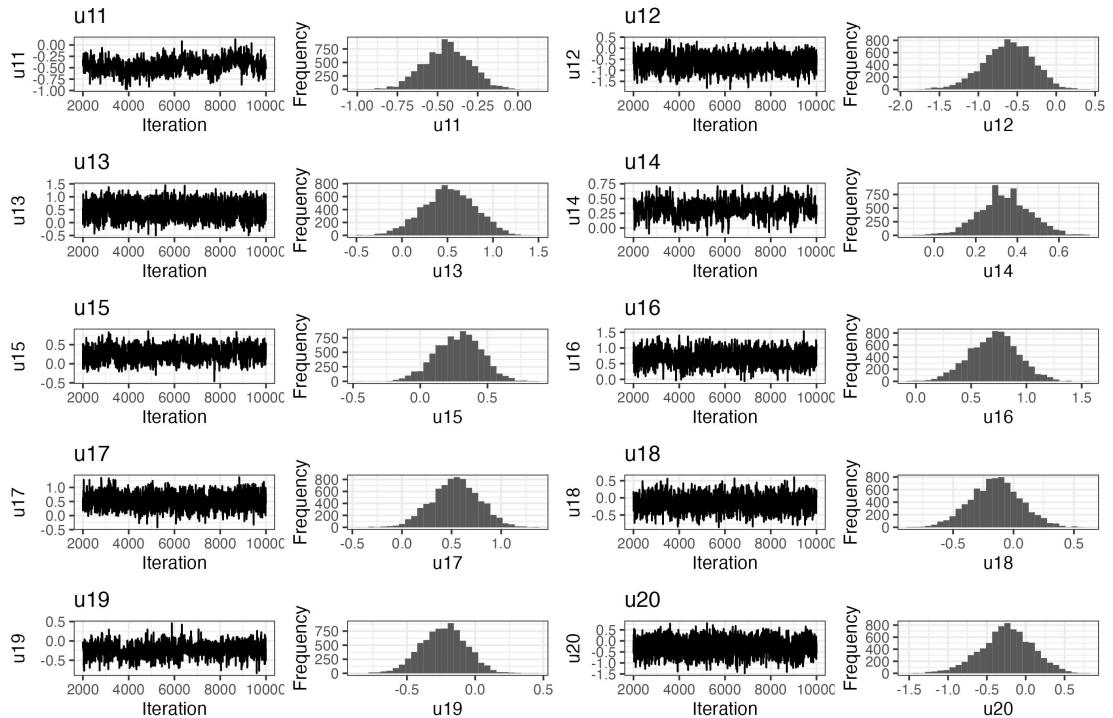


Figure 4: Trace plot for $u_{11}-u_{20}$

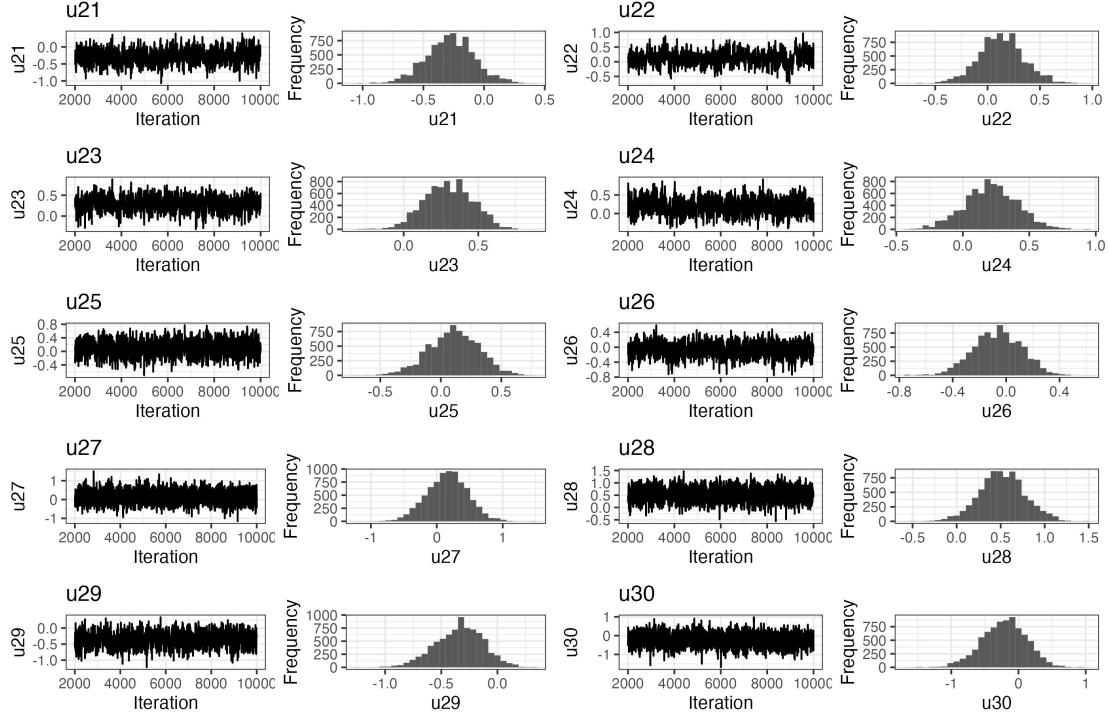


Figure 5: Trace plot for $u_{21}-u_{30}$

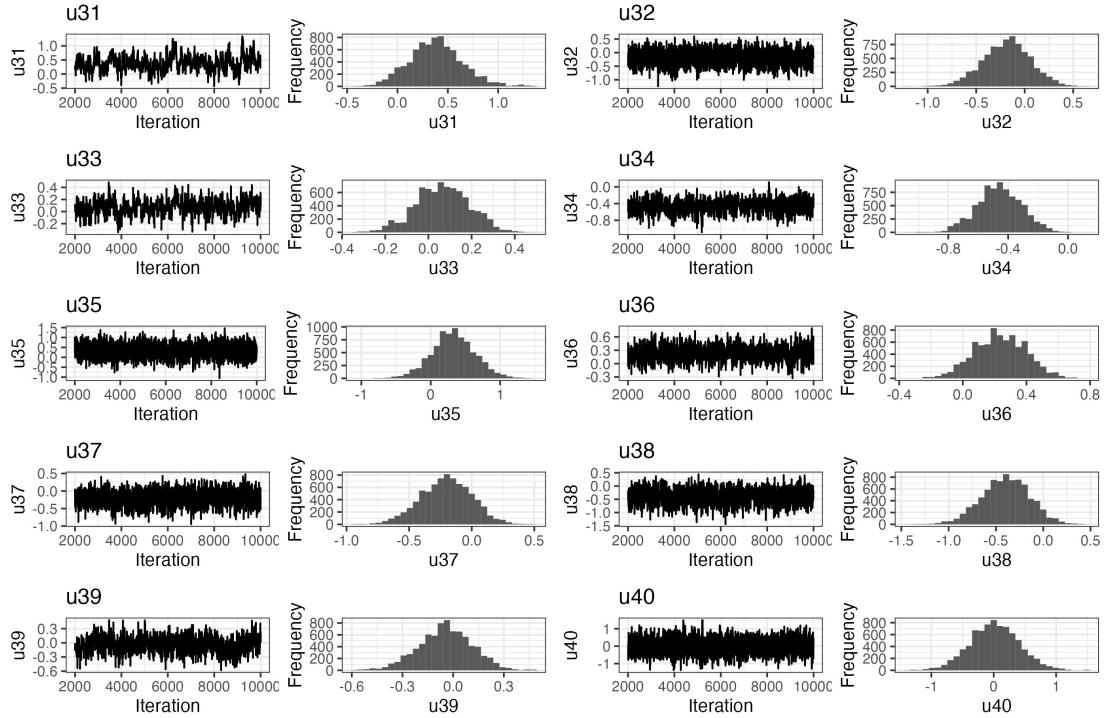


Figure 6: Trace plot for $u_{31}-u_{40}$

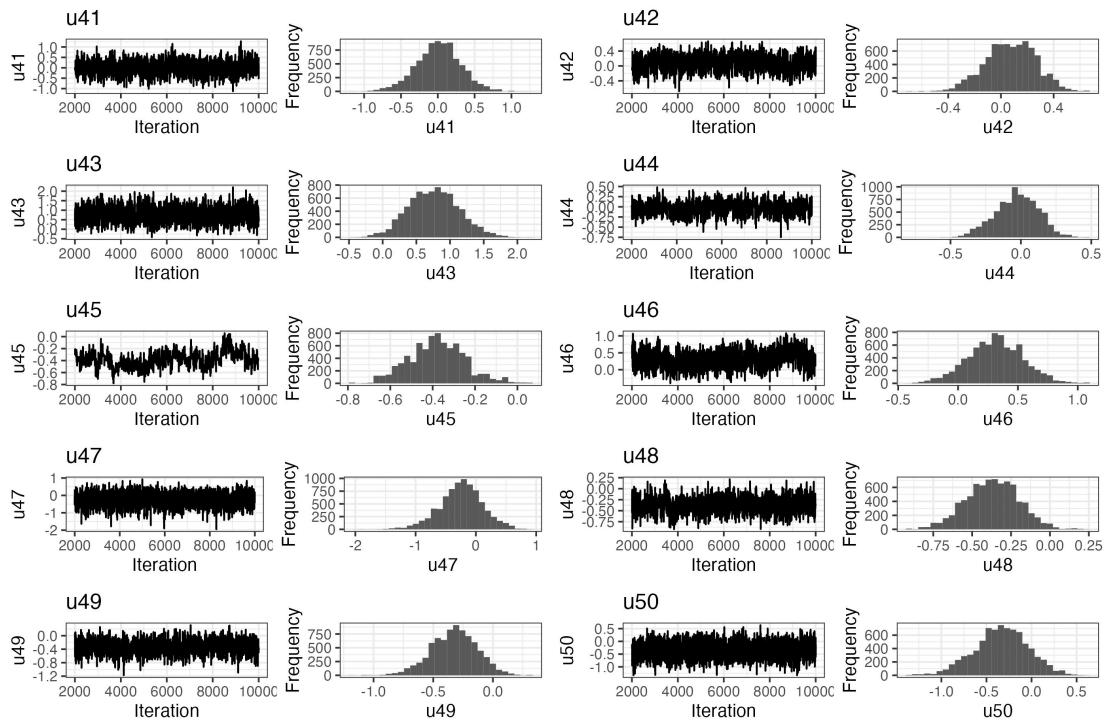


Figure 7: Trace plot for $u_{41}-u_{50}$



Figure 8: Autocorrelation plot for all parameters

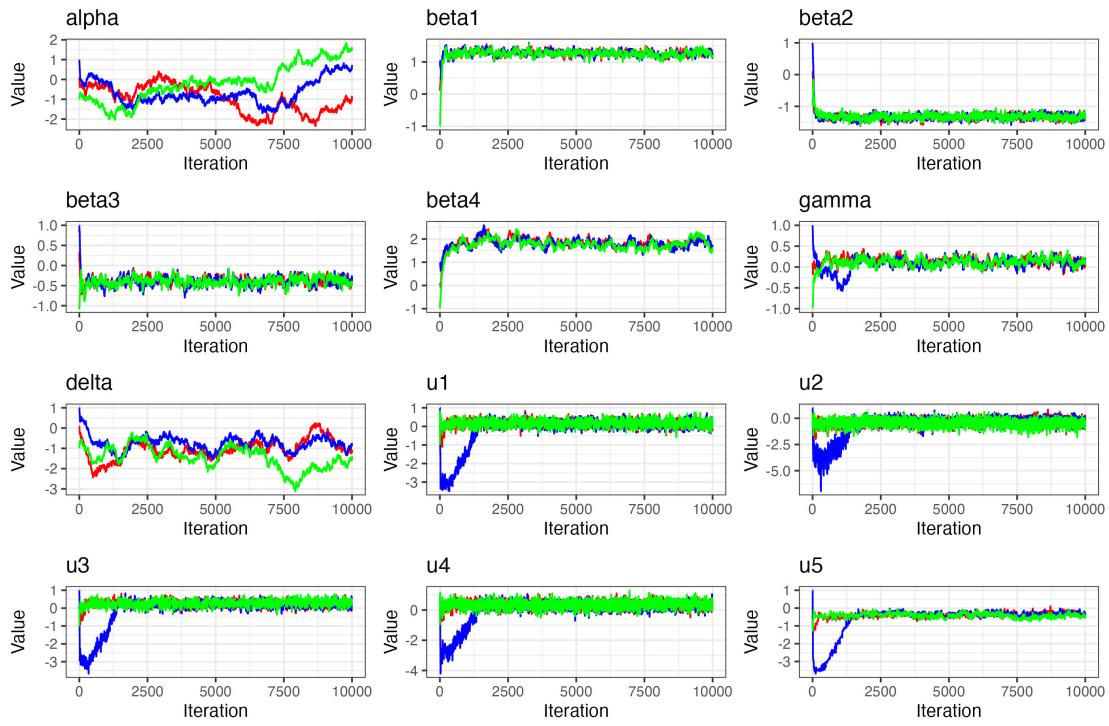


Figure 9: Merged trace plot for α, β s, γ, δ and u_1-u_5

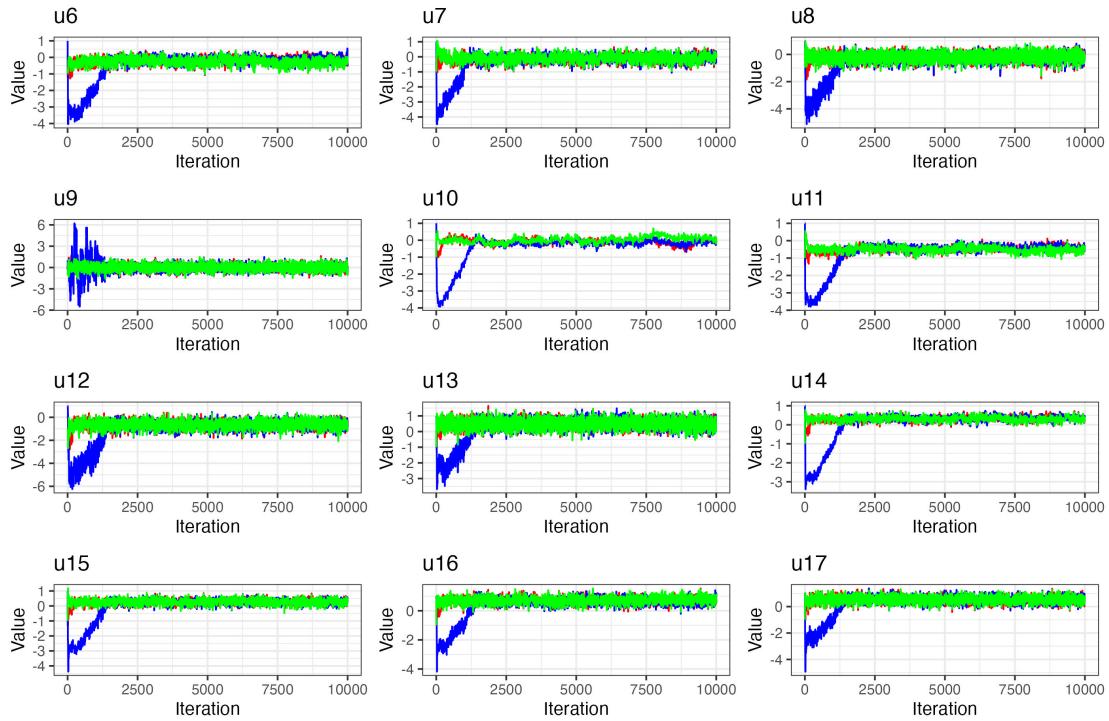


Figure 10: Merged trace plot for u_6-u_{17}

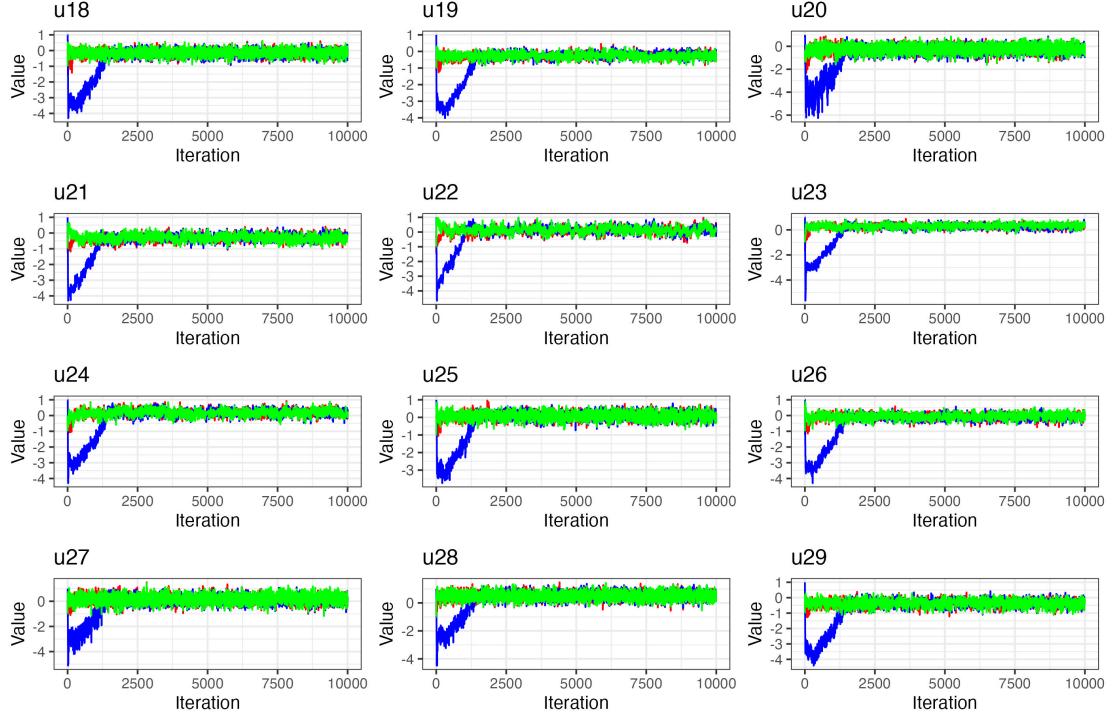


Figure 11: Merged trace plot for u_{18} - u_{29}

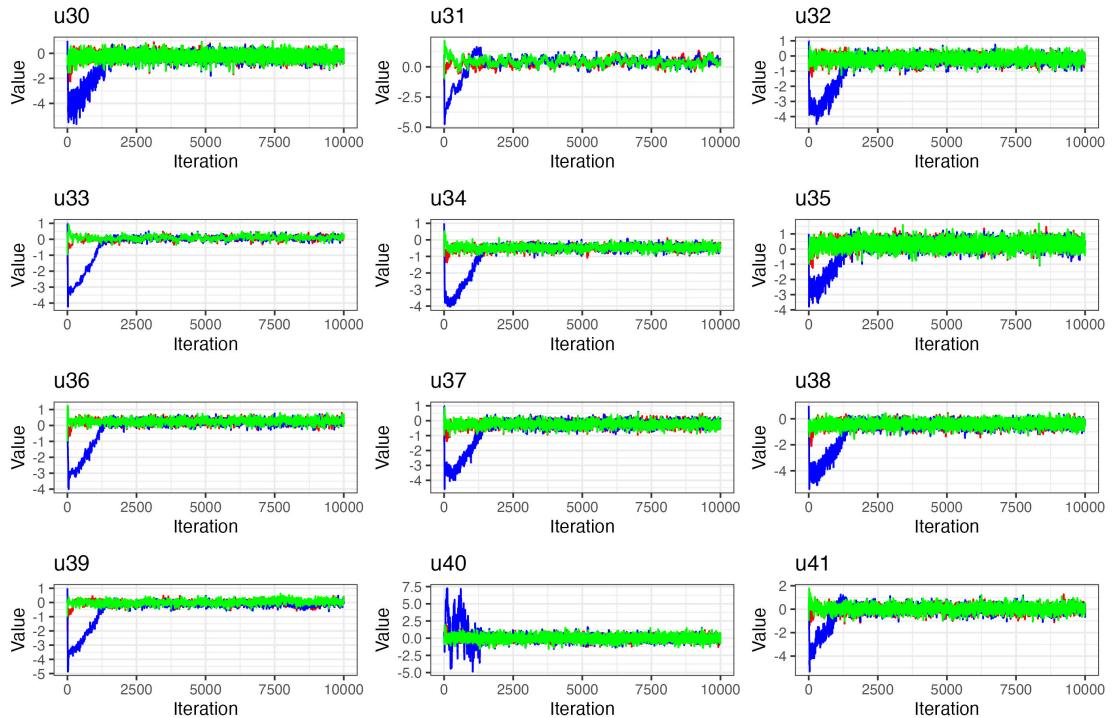


Figure 12: Merged trace plot for u_{30} - u_{41}

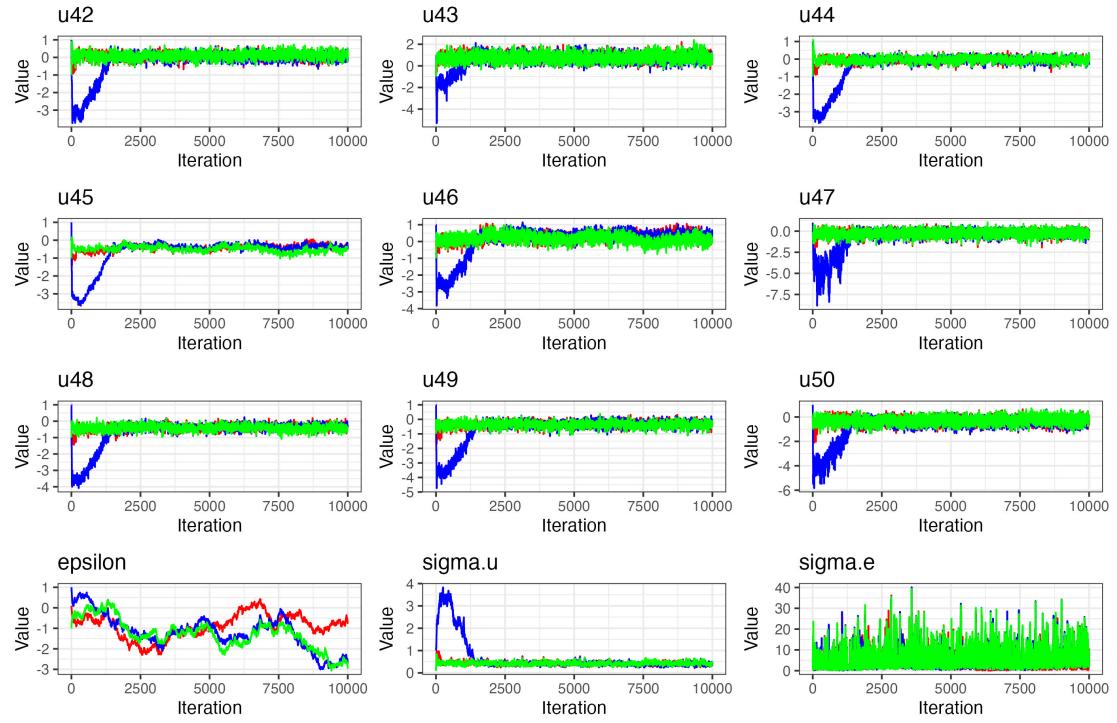


Figure 13: Merged trace plot for u_{42} - u_{50} , ϵ and σ s

Posterior distributions
with medians and 95% credible intervals

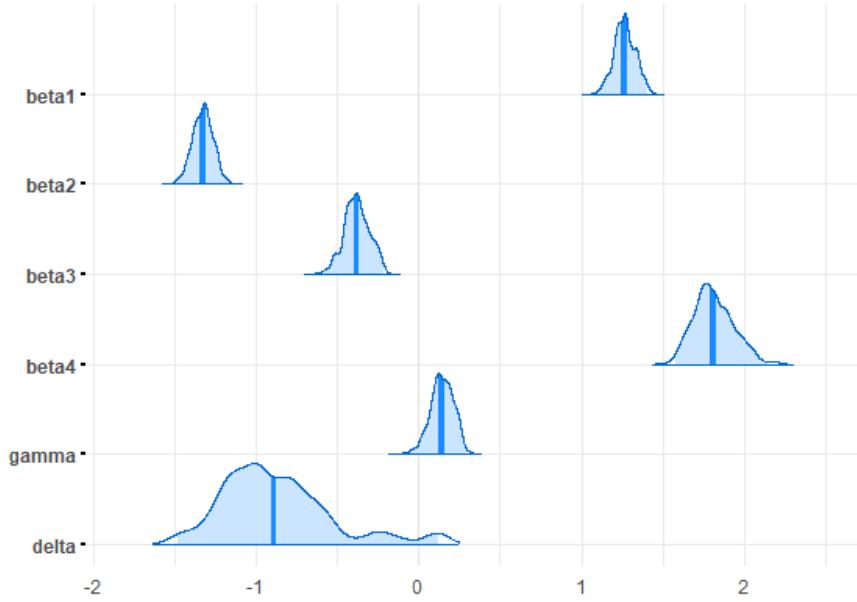


Figure 14: Posterior distributions of parameters of interest

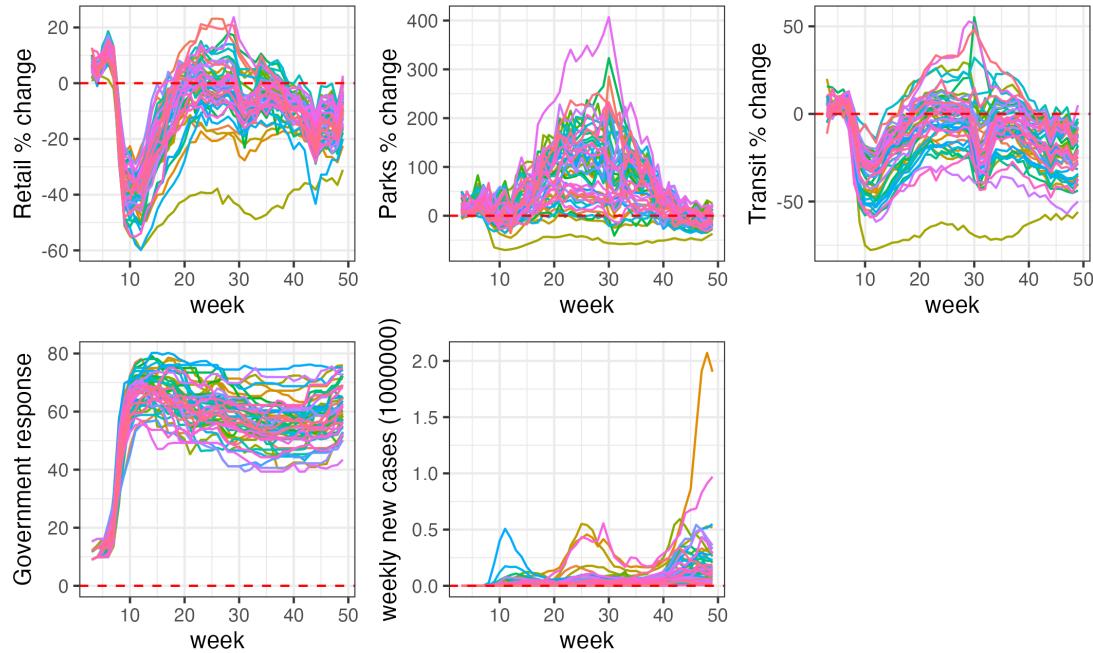


Figure 15: Trends of COVID cases, retail % change from baseline, parks % change from baseline, transit % change from baseline, and government response index over week

Table 1: Summary statistics before scaling

Statistic	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

Table 2: Final window length after adjustment

α	β_1	β_2	β_3	β_4	γ	δ	$u_{1-u_{50}}$	ϵ	σ_u	σ_e
0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	0.1	0.1	10

Table 3: Starting values

α	β_1	β_2	β_3	β_4	γ	δ	$u_{1-u_{50}}$	ϵ	σ_u	σ_e
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
1	1	1	1	1	1	1	1	1	1	1
-1	-1	-1	-1	-1	-1	-1	-1	-1	0.1	0.1

Table 4: Gelman-Rubin Statistics

Para	Stat	Converge	Para	Stat	Converge	Para	Stat	Converge
α	1.20	FALSE	u_{14}	1.06	TRUE	u_{34}	1.08	TRUE
β_1	1.01	TRUE	u_{15}	1.07	TRUE	u_{35}	1.05	TRUE
β_2	1.00	TRUE	u_{16}	1.08	TRUE	u_{36}	1.08	TRUE
β_3	1.00	TRUE	u_{17}	1.06	TRUE	u_{37}	1.07	TRUE
β_4	1.02	TRUE	u_{18}	1.07	TRUE	u_{38}	1.08	TRUE
γ	1.04	TRUE	u_{19}	1.07	TRUE	u_{39}	1.10	TRUE
δ	1.36	FALSE	u_{20}	1.07	TRUE	u_{40}	1.01	TRUE
u_1	1.09	TRUE	u_{21}	1.04	TRUE	u_{41}	1.04	TRUE
u_2	1.04	TRUE	u_{22}	1.04	TRUE	u_{42}	1.09	TRUE
u_3	1.10	TRUE	u_{23}	1.09	TRUE	u_{43}	1.04	TRUE
u_4	1.08	TRUE	u_{24}	1.07	TRUE	u_{44}	1.08	TRUE
u_5	1.05	TRUE	u_{25}	1.07	TRUE	u_{45}	1.04	TRUE
u_6	1.05	TRUE	u_{26}	1.09	TRUE	u_{46}	1.03	TRUE
u_7	1.05	TRUE	u_{27}	1.07	TRUE	u_{47}	1.06	TRUE
u_8	1.07	TRUE	u_{28}	1.06	TRUE	u_{48}	1.06	TRUE
u_9	1.00	TRUE	u_{29}	1.07	TRUE	u_{49}	1.07	TRUE
u_{10}	1.14	FALSE	u_{30}	1.06	TRUE	u_{50}	1.08	TRUE
u_{11}	1.05	TRUE	u_{31}	1.02	TRUE	ϵ	1.05	TRUE
u_{12}	1.07	TRUE	u_{32}	1.08	TRUE	σ_u	1.07	TRUE
u_{13}	1.06	TRUE	u_{33}	1.07	TRUE	σ_e	1.00	TRUE

Table 5: Posterior summaries

Statistic	N	Mean	St. Dev.	Min	Max
β_1	5,001	1.264	0.077	1.008	1.500
β_2	5,001	-1.330	0.076	-1.576	-1.090
β_3	5,001	-0.386	0.095	-0.702	-0.124
β_4	5,001	1.823	0.149	1.437	2.301
γ	5,001	0.136	0.089	-0.182	0.377
δ	5,001	-0.824	0.414	-1.631	0.246

β_1 : change in mobility trend in retail and recreation places; β_2 : change in mobility trend in parks; β_3 : change in mobility trend in transit stations; β_4 : government response index; γ : coefficient for state population density; δ : coefficient for state elderly percentage.

Table 6: 95% credible intervals

	2.5%	97.5%
β_1	1.111	1.411
β_2	-1.478	-1.178
β_3	-0.579	-0.217
β_4	1.575	2.146
γ	-0.053	0.296
δ	-1.478	0.123