Socioeconomic Factors and Violent Crime Rate Across the United States

Duy Hoang (1023390) - Linh Tran (1022838)

November 2024

Contents

1 Introduction

This project is part of the Aalto University Bayesian Data Analysis 2024 Course.

Ever since its declaration of independence in 1776, the United States of America has put heavy emphasis on the right to bear arms for its citizens. The Second Amendment of the Constitution of the United States declares that "A well regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed." In 2023, about four-in-ten U.S. adults say they live in a household with a gun, including 32 percent who say they personally own one [1]. Personal protection is the most prevalent reason why gun owners own a firearm, with about seven-in-ten gun owners (72 percent) say protection is a major reason they own a gun [2]. However, a majority of Americans (61 percent) also say it is too easy to legally obtain a gun in this country [2]. Gun-related violence has been a heated debate for decades, up to the point where it can be regarded as a public health tragedy. According to a report by Johns Hopkins University [3], in 2022, 48,2041 people died by firearms in the United States, almost half of these are considered homicides. The firearm homicide rate in the U.S. is nearly 25 times higher than other high-income countries and the firearm suicide rate is nearly 10 times that of other high-income countries.

While there have been considerable efforts to regulate guns in the US, the people have also protested the idea of taking away their firearms for safety reasons. Thus, there is a need to determine how much guns contribute to overall crime rates in the US. It is also noteworthy that there are not an abundance of research that performs Bayesian data analysis on this topic, thus we want to contribute by utilizing what we learned from the course to the scene. This project then investigates the effects of shall-carry laws and other socioeconomic factors on crime rate in the US from 1977 to 1999.

To provide some more context, a critical aspect of gun ownership in the United States is concealed carry permits, and there are many approaches to these laws adopted by each state. The most common options are to be either a "may-issue" state or a "shall-issue" state. In a may-issue state, you are not guaranteed a concealed carry permit, meaning the authority can deny a permit based on subjective criteria, such as determining if the applicant has a "good cause" or "proper reason" to carry a firearm [4]. Meanwhile, shall-carry permits require the issuing authority to grant the concealed carry license as long as the applicant meets all predefined legal criteria [4]. In short, may-carry laws are more restrictive than its shall-carry counterparts, and, as seen with the dataset later, were initially adopted by many states. The increasing implementation of shall-carry permits over the years has likely increased the number of guns in the US. Our project aims to investigate what relationship this higher number of guns can have with crime rates. With the violent crime rate as the target variable, we want to explore its relationship with the implementation of shall-carry law and other socioeconomic factors.

The project report consists of the following parts: Introduction, Data Description, Model Description, Results Analysis, Conclusion, and Learnings. Firstly, we will provide some overall descriptions of the data set, coupled with some explanatory graphs. Then, we analyze the performance of the models through convergence analysis, posterior predictive checks, predictive performance assessment, model selection, and prior sensitivity analysis. We will lastly discuss the issues and potential improvements for our models, as well as provide some interesting insights that we have learned while doing the project.

2 Data Description

The dataset explored in this project is `Guns: More Guns, Less Crime?` provided in the R package `AER: Applied Econometrics in R`. The dataset contains various socioeconomic factor indexes collected at 50 US States (and District of Columbia) during 1977-1999. The data collected does not have any NA value.

There are a total of 1173 rows of data and 13 columns corresponding to the variables below:

- `state` (categorical): The name of the state
- `year` (categorical): The year of observations
- `violent` (numerical): The rate of violent crime - number of incidents per 100,000 members of population
- `murder` (numerical): The rate of murder cases - number of incidents per 100,000 members of population
- `robbery` (numerical): The rate of robbery cases - number of incidents per 100,000 members of population
- `prisoners` (numerical): The incarceration rate valued for the previous year - number of sentenced prisoners per 100,000 residents]
- `afam` (numerical): The percentage of American-African in the population (state-level) in the age group of 10-64
- `cauc` (numerical): The percentage of Caucasian in the population (state-level) in the age group of 10-64
- `male` (numerical): The percentage of male population (state-level) in the age group of 10-29
- `population` (numerical): The population of the state in millions people
- `income` (numerical): The per capita personal income in the state
- `density` (numerical): The population per square mile in 1000 units
- `law` (categorical): Whether the state has a Shall-carry law for gun in that year

There have been several reports studying this dataset as it is a concerning problem in the US [$report_1$][$report_2$]. The difference between those reports and ours is that they focus one in-depth pairwise analysis of the variables, while our models follow the Bayesian workflow and study first the impact of `law` on `violent` then explore the relationship of all factors against `violent` to produce comprehensive results and comparison.

As can be seen from Figure 1, the violent crime rate at states that adopt Shall-carry law appears to be lower than those without such law. Although there is a rise during 1988-1993, the phenomenon looks more like nation-wide or due to pre-existing effect rather than consequences of Shall-carry law. Therefore, it still seems that the state of law has some effect on the violent crime rate. Therefore, one of our model is a hierarchical focusing on the relationship between `law` and `violent`.

Looking at the correlation matrix in Figure 2, we see that all the variables are related to

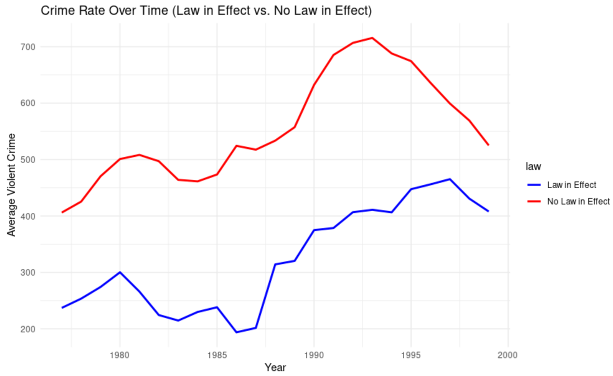

Figure 1: Average Crime Rate through Years

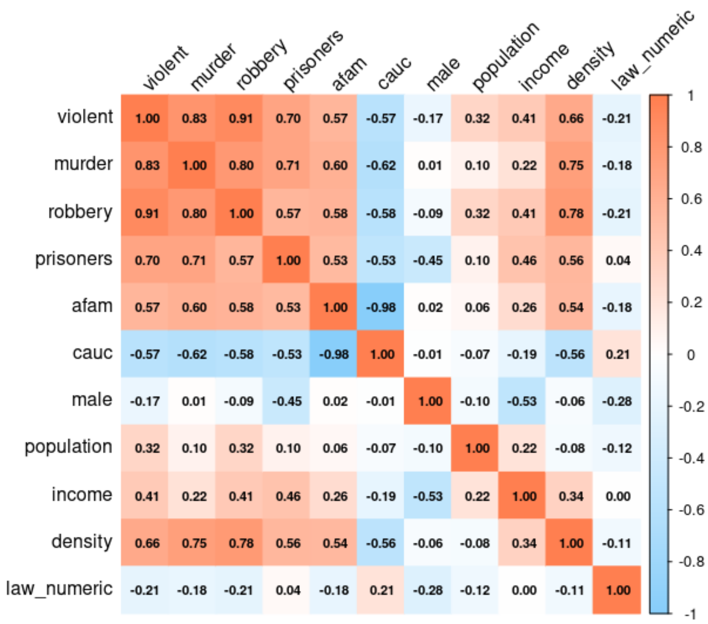

Figure 2: Correlation matrix between variables

`violent` to some extent, either negative or positive. Therefore, it is worth a look at how these variables perform. It can also be seen that `violent` is very correlated to `murder` and `robbery`, hence we will exclude them from our second model of understanding the effects of remaining factors on `violent`. This is also a hierarchical model, where all the variables (except for `murder` and `robbery`) are studied against `violent`.

Along side these two hierarchical model, we also include a pooled linear model to study the population-level effect of all factors on `violent`. This should provide a different perspective from the hierarchical models.

3 Model Description

In this project, we will employ one Pooled Model and two hierarchical Bayesian models to be analyzed: one focusing on gun laws and the other introducing additional predictors like demographic and economic factors.

3.1 Pooled Model

The Pooled Model includes all socioeconomic factors excluding `murder` and `robbery` as the explanatory variables and `violent` as the dependent variable. This studies the population-level impact of the factors without any varying intercept or slope.

The model in `brms` is:

$$violent \sim 1 + prisoners + afam + cauc + male + population + income + density + law$$

This is a simple linear regression model where each of the variables has a slope representing their effect on the dependent variable. This model gives us a comprehensive overview of how a unit change of these variables will be carried over to `violent`. This Pooled Model includes `prisoners`, `afam`, `cauc`, `male`, `population`, `income`, `density`, and `law` as fixed effects; and there is no smaller group level to be considered in this case.

3.1.1 Prior Justification

As expressed in Figure 2, every variable holds a reasonable correlation index towards each other. There is no outstandingly under-correlated and over-correlated one. Therefore, for each of the factors, we use $\beta_j \sim \mathcal{N}(0, 1)$ to assume a weakly informative distribution for the priors. This lets the model to explore on itself as we apply no prior knowledge about the distribution of the parameters. The normal distribution is also advantageous in its flexibility to explore a wide range of beliefs about the factors. The σ (residual error) follows $\mathcal{N}(0, 2)$ for a further exploration range compared to the slopes, with a left bound of 0 to make sure that it will not be negative because it represents the variation of the model.

3.1.2 Model Running

```
gun_pooled_formula <- bf(
  violent ~ 1 + prisoners + afam + cauc + male +
  population + income + density + law,
  family = "gaussian",
  center = FALSE
)

(gun_pooled_priors <- c(
  prior(normal(0, 1), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "prisoners"),
  prior(normal(0, 1), class = "b", coef = "afam"),
  prior(normal(0, 1), class = "b", coef = "cauc"),
  prior(normal(0, 1), class = "b", coef = "male"),
  prior(normal(0, 1), class = "b", coef = "population"),
  prior(normal(0, 1), class = "b", coef = "income"),
  prior(normal(0, 1), class = "b", coef = "density"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb= 0)
))

gun_pooled_fit <- brm(
  formula = gun_pooled_formula,
  prior = gun_pooled_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
```

```
  chains = 4
)
```

The model is run with 4 chains, each of which has 2000 iterations (1000 of them is for warming up).

3.2 Hierarchical Law Model

This model explores the relationship between gun laws (`law`) and violent crime rates (`violent`), accounting for variability of laws across states and a random effect years (varying intercept by year) by utilizing a hierarchical Bayesian framework to model the data.

The model is as follow:

$$violent \sim 1 + law + (1 + law|state) + (1|year)$$

The response variable here is the crime rate. Fixed effects include: the intercept term of `1`, representing the baseline level of violent crime when all predictors are at their reference levels, and `law` estimating the average effect of implementing gun laws on violent crime rates across all states and years. The first random effect (1 + law | state) allows each state to both have its own intercept (baseline level of violent crime) and its own slope for the effect of gun laws, capturing state-level variation in how gun laws impact violent crime. (1 | year) introduces a random intercept for each year, accounting for overall trends or fluctuations in violent crime rates over time like economic conditions or law enforcement policies.

3.2.1 Prior Justification

With the standardized data, we can assume a weakly informative prior $\mathcal{N}(0, 1)$ for fixed effects. This prior is reasonable because it assumes no strong preconception about the scale of violent crime as well as reflecting on the uncertainty of gun law, which can be contentious and their impact on violent crime is often debated with varying empirical results. A standard deviation of 1 ensures the prior does not overly constrain the posterior distribution, allowing large effects if strongly supported by the data.

With the residual variance `sigma`, we assume that the residual standard deviation of the violent crime rate is most likely close to 0 but could plausibly range up to about 2 - $\mathcal{N}(0, 2)$. This allows for a wider range of variability while still regularizing the parameter to avoid extreme values.

For random effects standard deviations, to account for the higher unpredictability, we extend the prior to be between 0 and 5, i.e., $\mathcal{N}(0, 5)$. A standard deviation of 5 can be substantially better at capturing the wide variety in the behavior of these variables. Violent crime rates can vary widely across states due to differences in demographics, culture, and enforcement policies; states may also differ in how they implement and enforce gun laws; and national-level changes like economic conditions or major federal policies can also affect violent crime rates across all states.

3.2.2 Model Running

```
gun_law_formula <- bf(
  violent ~ 1 + law + (1 + law | state) + (1 | year),
  family = "gaussian",
```

```
  center = FALSE
)

get_prior(gun_law_formula, data = Guns)

(gun_law_priors <- c(
  prior(normal(0, 1), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb= 0),
  prior(normal(0, 5), class = "sd", group = "state", coef = "Intercept"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "lawyes"),
  prior(normal(0, 5), class = "sd", group = "year", coef = "Intercept")
))

gun_law_fit <- brm(
  formula = gun_law_formula,
  prior = gun_law_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4
)
```

The model is run with 4 chains, each having 2000 iterations including 1000 warm-up iterations.

3.3 Hierarchical All Model

The third model is an hierarchical model including every socioeconomic factor both at population level and `state` level, including varying intercepts between `state` and `year`.

The model is written in `brms` as:

$$\begin{aligned} violent \sim & 1 + prisoners + afam + cauc + male+ \\ & population + income + density + law+ \\ & (1 + prisoners + afam + cauc + male+ \\ & population + income + density + law|state)+ \\ & (1|year) \end{aligned}$$

Similar to previous models, the dependent variable is still `violent`, but now we have increased the number of predictory variables to include all `prisoners`, `afam`, `cauc`, `male`, `population`, `income`, `density`, and `law`. The general idea is similar to Law Model described above, but now each factor has its own fixed effect on `violent` and a varying slope on the `state` level. Besides, we still keep a varying intercept for `year` factor to explain the baseline level of `violent` through each year.

3.3.1 Prior Justification

For priors, we use the same distribution as above, i.e., $\mathcal{N}(0, 1)$ for predictory factors, but now we have more variables than previously. By choosing these weakly informative priors, we allow the model to explore the parameters itself within a reasonable range.

The residual error σ is still assigned the prior of $\mathcal{N}(0, 2)$ to explain for a bit larger variability and reserves the consistency from previous models.

Regarding random effects' standard deviation, wider-range priors $\mathcal{N}(0, 5)$ are chosen so that the model can capture more variance of these effects, better understanding the relationship between state-level impacts on `violent`.

3.3.2 Model Running

```
gun_all_formula <- bf(
  violent ~ 1 + prisoners + afam + cauc + male +
            population + income + density + law +
            (1 + prisoners + afam + cauc + male +
            population + income + density + law | state) + (1 | year),
  family = "gaussian",
  center = FALSE
)

get_prior(gun_all_formula, data = Guns)

(gun_all_priors <- c(
  prior(normal(0, 1), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "prisoners"),
  prior(normal(0, 1), class = "b", coef = "afam"),
  prior(normal(0, 1), class = "b", coef = "cauc"),
  prior(normal(0, 1), class = "b", coef = "male"),
  prior(normal(0, 1), class = "b", coef = "population"),
  prior(normal(0, 1), class = "b", coef = "income"),
  prior(normal(0, 1), class = "b", coef = "density"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb = 0),

  prior(normal(0, 5), class = "sd", group = "year", coef = "Intercept"),

  prior(normal(0, 5), class = "sd", group = "state", coef = "Intercept"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "prisoners"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "afam"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "cauc"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "male"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "population"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "income"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "density"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "lawyes")
))

gun_all_fit <- brm(
  formula = gun_all_formula,
  prior = gun_all_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4
)
```

Similarly, the model is run with 4 chains and 2000 iterations, within which 1000 are warm-ups

4 Results Analysis

4.1 Convergence Diagnostics

4.1.1 $\hat{R}$ diagnostics

The $\hat{R}$ values for all the models are plotted into these histograms in Figures 3, 4, 5. From here, we can observe that almost all the $\hat{R}$ values for all the models are below 1.05, and they are relatively concentrated around 1.00. This result indicates that, overall, the MCMC chains for All Model have converged quite well.

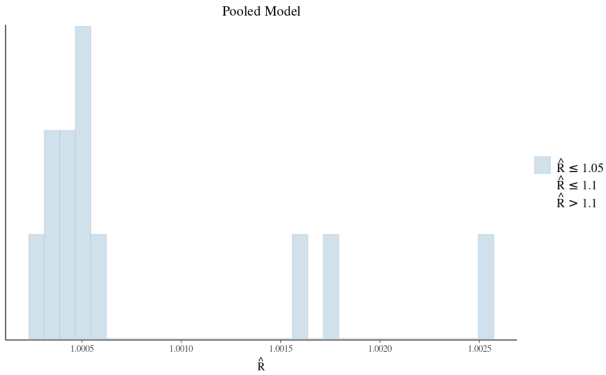

Figure 3: $\hat{R}$ Values of Pooled Model

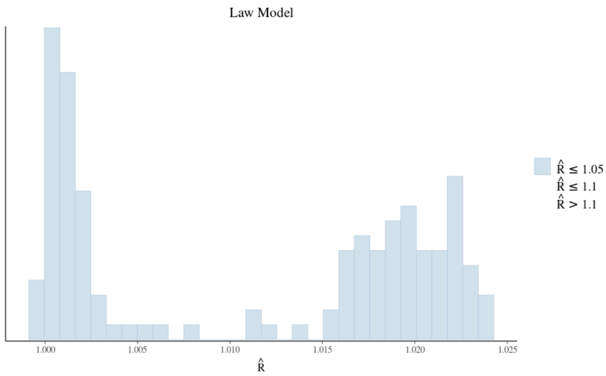

Figure 4: $\hat{R}$ Values of Law Model

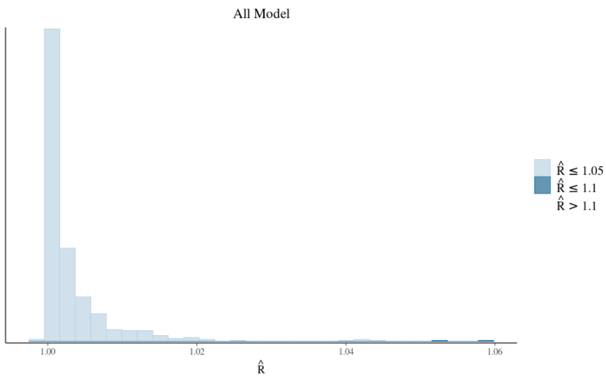

Figure 5: $\hat{R}$ Values of All Model

4.1.2 Effective sample size (ESS) diagnostics

Here, we plot a metric called ESS ratio, which is equal to $\frac{N_{\text{eff}}}{N_{\text{sample}}}$. This is the ratio of an effective sample size to total sample size. This ratio can give us a relative measure that normalizes ESS by the total number of samples, providing additional context on the efficiency of the sampling process. It helps to assess how much of the total sample is effectively contributing independent information.

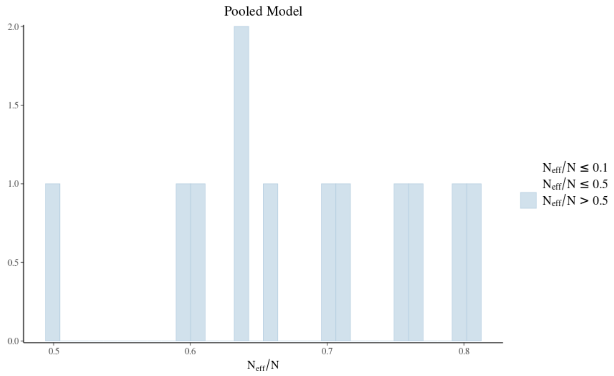

Figure 6: ESS Ratio for Pooled Model

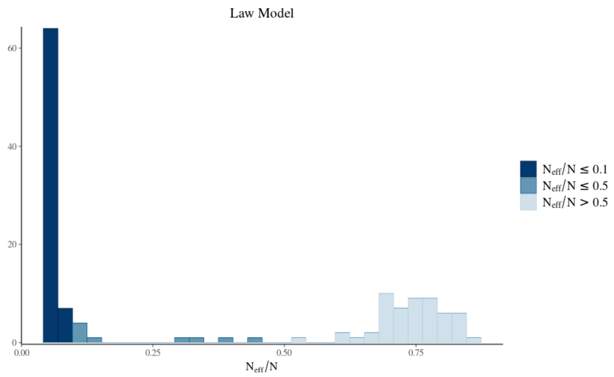

Figure 7: ESS Ratio for Law Model

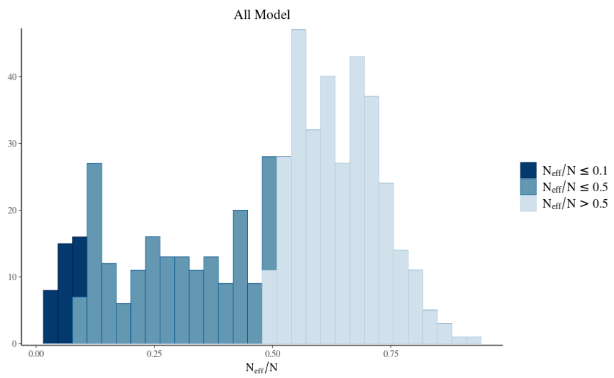

Figure 8: ESS Ratio for All Model

From the Figures 6, 7, 8, we can observe that although Rhat values between 1 and 1.025 indicate convergence, all three models show relatively poor $\frac{N_{\text{eff}}}{N_{\text{sample}}}$ ratio. This means that while the chains may have converged, the sampling process might not be exploring the parameter space effectively and the chains might be highly autocorrelated, suggesting inefficient sampling.

4.1.3 Divergences and tree depth diagnostics

We use the `rstan::check_hmc_diagnostics()` function to learn about the Hamiltonian Monte Carlo (HMC) sampling for each of the models. The results are as follow:

```
> rstan::check_hmc_diagnostics(gun_pooled_fit$fit)

Divergences:
0 of 4000 iterations ended with a divergence.

Tree depth:
0 of 4000 iterations saturated the maximum tree depth of 10.

Energy:
E-BFMI indicated no pathological behavior.


> rstan::check_hmc_diagnostics(gun_law_fit$fit)

Divergences:
0 of 4000 iterations ended with a divergence.

Tree depth:
0 of 4000 iterations saturated the maximum tree depth of 10.

Energy:
E-BFMI indicated no pathological behavior.


> rstan::check_hmc_diagnostics(gun_all_fit$fit)

Divergences:
0 of 4000 iterations ended with a divergence.

Tree depth:
4000 of 4000 iterations saturated the maximum tree depth of 10 (100%).
Try increasing 'max_treedepth' to avoid saturation.

Energy:
E-BFMI indicated no pathological behavior.
```

For divergences, All Model reports 0 divergences out of 4000 iterations, which suggests that these models are rather well-behaved. Meanwhile, both the Pooled and the Law Model report 0 of 4000 iterations that saturated the maximum tree depth (set to 10). This means that the HMC sampler did not need to use excessively deep trees to explore the posterior, suggesting the models are relatively straightforward in terms of their complexity. However, the All Model reports that 100 percent of the iterations saturated the maximum tree depth of 10. Here, when the tree depth is saturated, it means the sampler couldn't efficiently explore certain regions of the posterior and had to truncate the exploration.

4.2 Posterior Predictive Checks

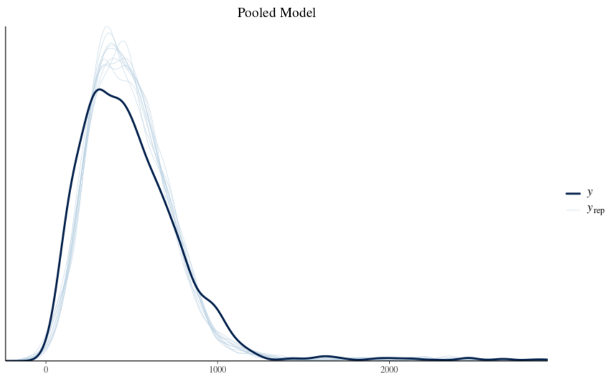

Figure 9: Posterior Predictive Check for Pooled Model

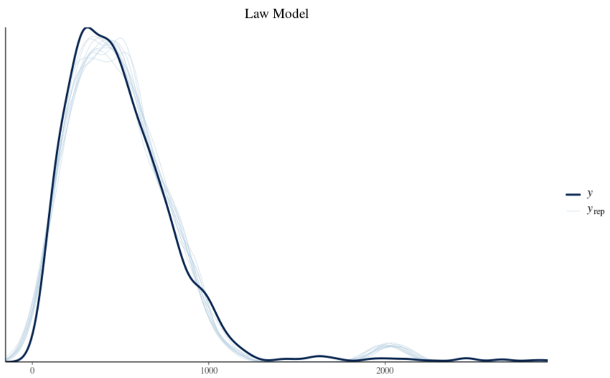

Figure 10: Posterior Predictive Check for Law Model

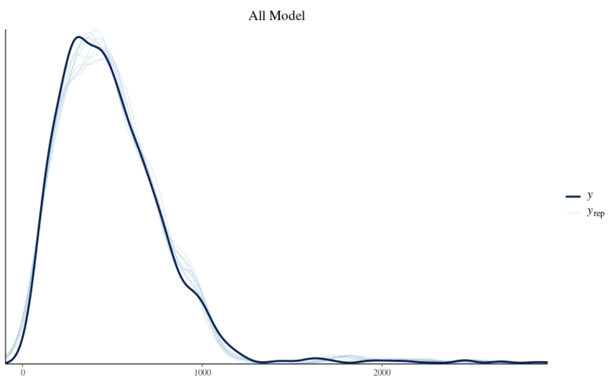

Figure 11: Posterior Predictive Check for All Model

As can be seen from Figure 9, the posterior distribution of the Pooled Model is not the best that we can have. It overestimates the values around the peak of the plot. It also underestimates the values in the right tail. Compared to it, the Law Model has done a visibly better job as presented in Figure 10. The gap in the top has become closer and the tails match better than in Pooled Model case. Even though it still has some flaws and the predictions shift a bit from the observations (especially when there suddenly exists a mode in the right tail), the posterior prediction is able to explain the original distribution to a point. The All Model has performed slightly better than Law Model in a sense that the predictions are not shifting away from the observations, and there is no other sudden rise. Overall, it has matched very well with the observations, better than Pooled Model and Law Model.

4.3 Model Comparison

Model comparison is conducted using Leave-One-Out Cross-Validation (LOO-CV) to estimate the generalization performance of the models. Firstly, Pareto k values are plotted to evaluate the reliability of the estimated expected log predictive density (elpd).

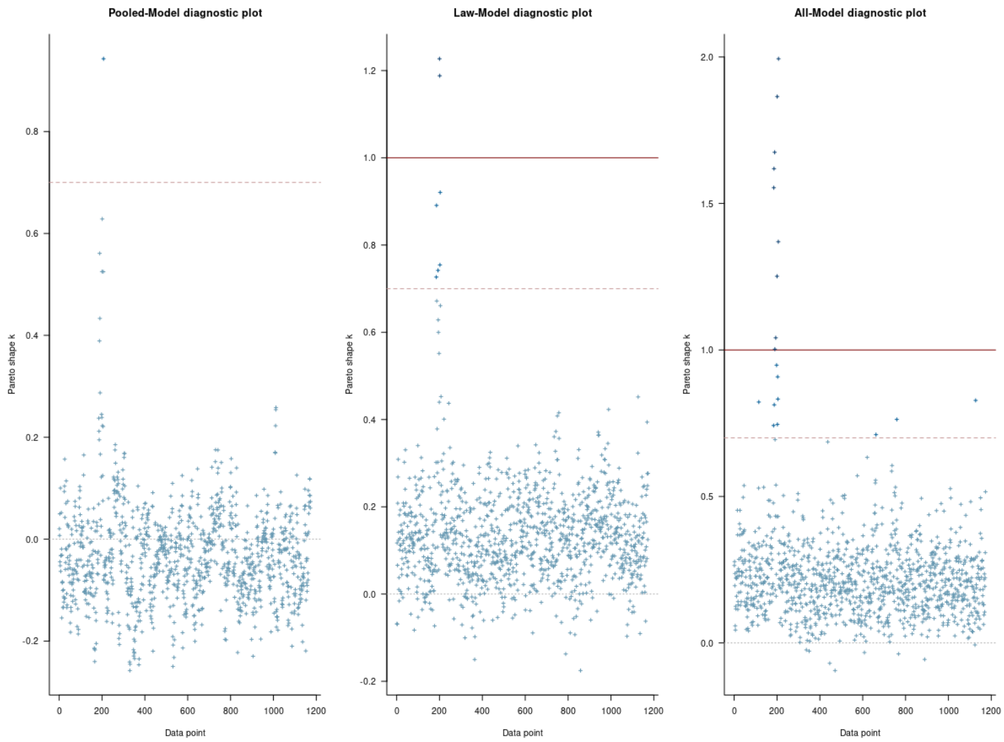

Figure 12: Pareto k diagnostics of different models

While all the Pareto k values for the Pooled Model are below 0.5 (except for one value) - suggesting reliable results, the results for the other two models are not so great. The Law Model reports multiple values over the threshold for bad values of 0.7 and that number for the All Model is even higher. However, it should also be noted that the ratio of these values to all observations is very small. This indicates that the estimates for both hierarchical models can be somewhat overly optimistic, especially the All Model.

The LOO-CV estimation of elpd for All Model is given and compared as follows:

```
> pooled_loo

Computed from 4000 by 1173 log-likelihood matrix.

```

```
          Estimate    SE
elpd_loo   -8654.9 143.7
p_loo         49.4   9.5
looic      17309.8 287.5
------
MCSE of elpd_loo is NA.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.7, 1.5]).

Pareto k diagnostic values:
                         Count Pct.    Min. ESS
(-Inf, 0.7]   (good)     1172  99.9%   169
   (0.7, 1]   (bad)         1   0.1%   <NA>
   (1, Inf)   (very bad)    0   0.0%   <NA>
See help('pareto-k-diagnostic') for details.


> law_loo

Computed from 4000 by 1173 log-likelihood matrix.

          Estimate    SE
elpd_loo   -7104.1 147.5
p_loo        153.1  21.0
looic      14208.3 295.0
------
MCSE of elpd_loo is NA.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.3, 1.8]).

Pareto k diagnostic values:
                         Count Pct.    Min. ESS
(-Inf, 0.7]   (good)     1165  99.3%   165
   (0.7, 1]   (bad)         6   0.5%   <NA>
   (1, Inf)   (very bad)    2   0.2%   <NA>
See help('pareto-k-diagnostic') for details.


> all_loo

Computed from 4000 by 1173 log-likelihood matrix.

          Estimate    SE
elpd_loo   -6633.6 142.1
p_loo        303.0  50.1
looic      13267.3 284.3
------
MCSE of elpd_loo is NA.
MCSE and ESS estimates assume MCMC draws (r_eff in [0.2, 1.4]).

Pareto k diagnostic values:
                         Count Pct.    Min. ESS
(-Inf, 0.7]   (good)     1154  98.4%   95
   (0.7, 1]   (bad)        10   0.9%   <NA>
   (1, Inf)   (very bad)    9   0.8%   <NA>
See help('pareto-k-diagnostic') for details.
> loo_compare(pooled_loo, law_loo, all_loo)
               elpd_diff se_diff
gun_all_fit        0.0       0.0
gun_law_fit     -470.5     111.9
gun_pooled_fit -2021.3     117.9
```

We know that higher elpd values indicate better model predictive performance (on the same dataset). Given this, the All Model has the best predictive performance but also the highest complexity and some reliability concerns in its elpd estimates due to high Pareto k values. The Pooled Model is the most stable and reliable but performs poorly in terms of prediction. The Law Model is more balanced but also outperformed by the

All Model. If interpretability and simplicity are priorities, this model might be preferred.

4.4 Prior Sensitivity Analysis

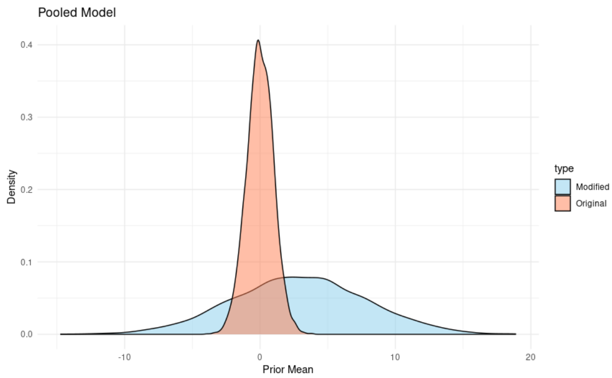

Figure 13: `prior_Intercept`

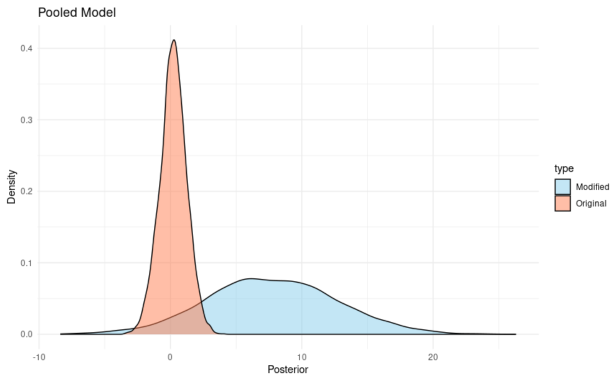

Figure 14: `Intercept`

For the Pooled Model, we change the prior of `Intercept` from $\mathcal{N}(0, 1)$ to $\mathcal{N}(3, 5)$ (Figures 13 and 14)

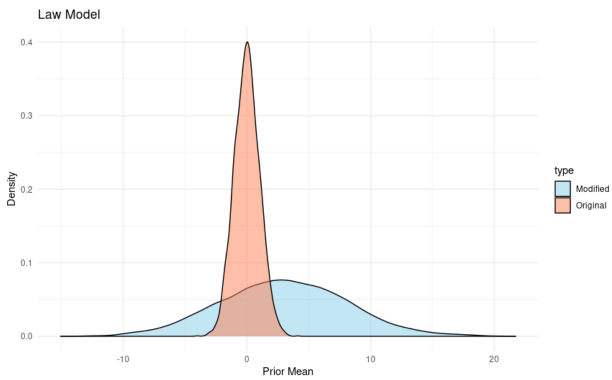

Figure 15: `prior_Intercept`

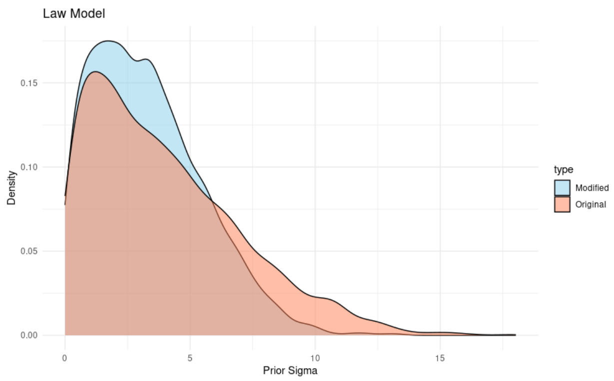

Figure 16: `prior_sd_state__Intercept`

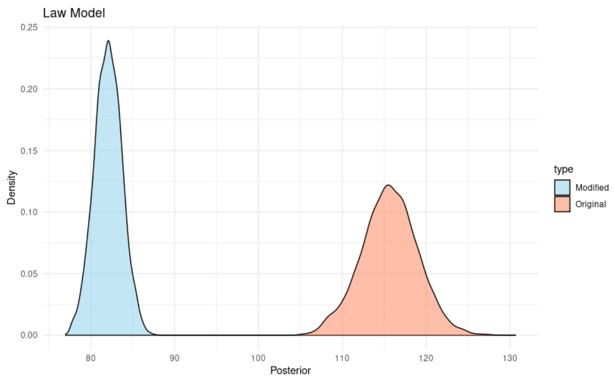

Figure 17: `sd_state__Intercept`

For the Law Model, the prior for `Intercept` is changed from $\mathcal{N}(0,1)$ to $\mathcal{N}(3,5)$ and that for `sd_state__Intercept` is transformed from $\mathcal{N}(0,5)$ to $\mathcal{N}(2,3)$ (Figures 15, 16, 17). .

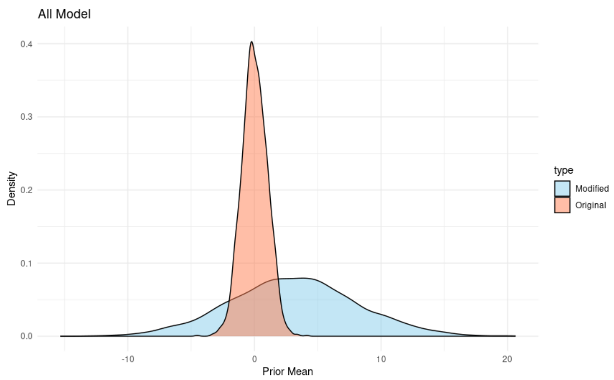

Figure 18: `prior_Intercept`

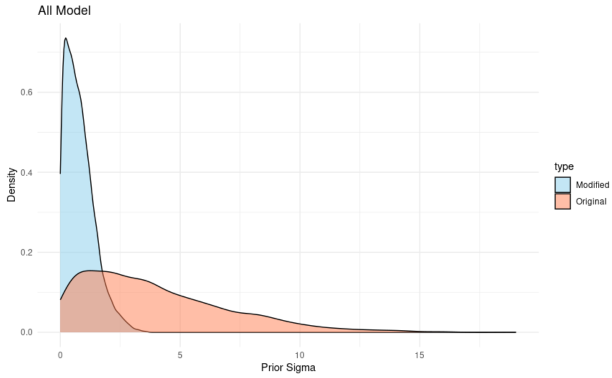

Figure 19: `prior_sd_state__Intercept`

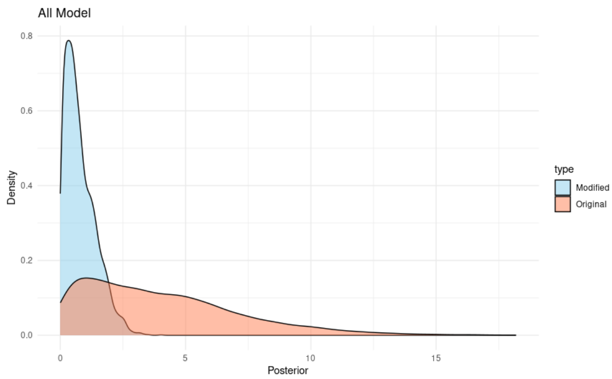

Figure 20: `sd_state__Intercept`

For the All Model, the prior for `Intercept` is also changed from $\mathcal{N}(0, 1)$ to $\mathcal{N}(3, 5)$ and that of `sd_state__Intercept` is transformed from $\mathcal{N}(0, 5)$ to $\mathcal{N}(0, 1)$ (Figures 18, 19, 20)

As can be seen from these plots, all of the models are very sensitive to prior choices. Whether a more informative or general prior will still affect the posterior results. Therefore, proper prior choices and profound domain knowledge are of much importance in studying the effects of socioeconomic factors on `violent`. The models depend very much on the informative priors about the distribution of variables and relationship between them. A generic prior will not be the best option for our case.

5 Conclusion and Improvement Discussion

5.1 Conclusion

The summary of regression coefficients for all three models is included below:

```
> summary(gun_pooled_fit)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: violent ~ 1 + prisoners + afam + cauc + male + population +
    income + density + law
   Data: Guns (Number of observations: 1173)
```

```
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Regression Coefficients:
           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept      0.20      0.99    -1.72     2.14 1.00     3911     2551
prisoners      1.01      0.02     0.96     1.05 1.00     3898     2475
afam           6.76      0.73     5.34     8.21 1.00     2749     2720
cauc          -3.01      0.30    -3.59    -2.42 1.00     2240     2488
male           9.20      0.81     7.57    10.79 1.00     2774     2916
population    11.31      0.53    10.24    12.33 1.00     3650     2709
income         0.02      0.00     0.01     0.02 1.00     3367     3154
density        9.04      0.97     7.11    10.93 1.00     4038     2790
lawyes        -2.12      1.00    -4.06    -0.18 1.00     4227     3061


> summary(gun_law_fit)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: violent ~ 1 + law + (1 + law | state) + (1 | year)
   Data: Guns (Number of observations: 1173)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept     0.97      1.00    -0.97     2.94 1.00     5770     2969
lawyes       -0.18      0.99    -2.11     1.76 1.00     5861     2973


> summary(gun_all_fit)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: violent ~ 1 + prisoners + afam + cauc + male + population +
    income + density + law + (1 + prisoners + afam + cauc + male +
    population + income + density + law | state) + (1 | year)
   Data: Guns (Number of observations: 1173)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Regression Coefficients:
           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept      0.01      0.99    -1.91     2.01 1.00     2626     2882
prisoners     -0.02      0.12    -0.24     0.21 1.00      460      886
afam           0.25      0.96    -1.65     2.22 1.00     2551     2542
cauc           0.66      0.80    -0.92     2.21 1.00      988     1697
male           0.39      0.97    -1.56     2.27 1.00     2106     2261
population     0.28      1.00    -1.65     2.26 1.00     2419     2742
income         0.03      0.00     0.02     0.04 1.00      860     1406
density       -0.12      1.02    -2.08     1.90 1.00     2446     2303
lawyes         0.06      0.99    -1.88     1.98 1.00     2829     2751
```

This project focuses on identifying the effects of gun laws and other socioeconomic factors on the violent crime rates in the US through three models: one Pooled Model and two hierarchical models. Looking at the summary table, we observe the following:

- For Pooled Model, `law` does seem to have a small effect on reducing the violent rate. However, compared to other factors and to the values of `violent` itself, this is just a very small number that cannot prove the importance of `law`. Also, as we pointed above in posterior predictive checks, the resulting values of Pooled Model do not explain much the actual observations.

- For Law Model, the coefficient suggests that `law` does not have much impact on `violent`. Even when it is the sole consideration of this model, the number does not show any significance.

- For All Model, the situation is now that all the socioeconomic factors are quite insignificant towards `violent`. This suggests that including all variables might not be the best idea to learn about `violent`.

And from all the analyses conducted, our models are not the best in stability and reliability. They can show us some aspects about these socioeconomic factors, but letting them to predict the `violent` rate given these variables is not the best choice.

5.2 Issues and Improvement Space

The biggest issue when we first handled the models is that we have no prior knowledge about the domain. Therefore, we stick to the safest options for priors. Certainly, this would lead to sub-optimal results as we can already see from above analysis. If the priors can be chosen more wisely and informatively, they might prove to be more useful. As we have stated previously in Section 4.4, a slight change in priors can lead to a different posterior distribution. Misspecified priors most likely lead to the poor results presented in the code snippet right above.

The second issue is that variable selection. Initially, we want to study the impact of Shall-carry Law on controlling the crime rate, therefore we have a model just for `law`. But the results don't turn out to be great. Therefore, we continue putting other factors in the model to see if the situation can get better. Eventually, none of these models can actually cover the `violent` rate. We think that, if the model size is reduced to just several very impactful factors, such as `prisoners` and `male` (as studied in [$report_1$]), the models can return better results.

The third thing stems from the fact that we have included quite a few variables in our model and since there is a hierarchical model dependent on `state` (which has 51 values), the computation and waiting time is also a hindrance for us. Just waiting the All Model to complete sampling takes several hours. This has prevented us from exploring the effective priors for it. Could we keep the model smaller, it would be a big step further already.

While experimenting with the models, we have tried several options such as increasing iterations for sampling, using another family for the formula (`lognormal`, `gamma`) for a positively skewed distribution, but none of them turns out to be better than the current version. By choosing relatively general options while we don't have much domain knowledge, we can let the models explore on their own. One possible improvement that we think might prove beneficial is to clip the range of `violent` so that outliers won't be considered and create bias in our models.

6 Learnings

Throughout this project, we have better understood how to design and draw insights from Bayesian data analysis models. Finding our own data and determining what approaches to adopt have been both challenging and rewarding at the same time, unlike when we were given data and starter codes for assignments. The topic of choice was also some thing that we both care about: we wanted to better understand the situation of guns and violent crime rates in the US. The model building process was also problematic at first since we needed to decide on what models would be more effective in addressing our topic. The two hierarchical models were chosen primarily, with the addition of the Pooled Model later to show better variety. Beside all the technical knowledge, this project has also taught us to better handle time, organize a more readable and professional report, and problem-solving skills. This has been a very valuable and insightful experience, and we are satisfied that we have created this project on our own.

7 References

[1] Pew Research Center. "For most U.S. gun owners, protection is the main reason they own a gun." Accessed: 2024-11-25. (Aug. 2023), [Online]. Available: https://www.pewresearch.org/politics/2023/08/16/for-most-u-s-gun-owners-protection-is-the-main-reason-they-own-a-gun/.

[2] Pew Research Center. "Key facts about americans and guns." Accessed: 2024-11-25. (Jul. 2024), [Online]. Available: https://www.pewresearch.org/short-reads/2024/07/24/key-facts-about-americans-and-guns/.

[3] Johns Hopkins University, Center for Gun Violence Solutions. "Firearm violence in the united states." Accessed: 2024-11-25. (2024), [Online]. Available: https://publichealth.jhu.edu/center-for-gun-violence-solutions/research-reports/firearm-violence-in-the-united-states.

[4] USCCA (U.S. Concealed Carry Association). "May-issue vs. shall-issue concealed carry states." Accessed: 2024-11-25. (2024), [Online]. Available: https://www.usconcealedcarry.com/blog/may-issue-vs-shall-issue-concealed-carry-states/.

8 Appendix

The *R* script for this project is appended below.

```r
library(ggplot2)
library(posterior)
library(bayesplot)
library(brms)
library(ggpubr)
library(tidyr)
library(dplyr)
library(ggcorrplot)
library(gridExtra)
library(dplyr)
library(loo)
library(invgamma)
library(AER)
library(viridis)
library(corrplot)
library(tidyverse)

data("Guns", package = "AER")

Guns$year <- as.numeric(as.character(Guns$year))
guns_yes <- Guns %>% filter(law == "yes")
guns_no <- Guns %>% filter(law == "no")

guns_yes_mean <- guns_yes %>%
  group_by(year) %>%
  summarise(total_violent = mean(violent))

guns_no_mean <- guns_no %>%
  group_by(year) %>%
  summarise(total_violent = mean(violent))

# Average crime incidents through year with Yes law
plot_yes <- ggplot(guns_yes_mean, aes(x = year, y = total_violent)) +
  geom_line(color = "blue") +
  labs(title = "Crime Rate Over Time (Law in effect)", x = "Year", y = "Average Violent Crime") +
  theme_minimal()
plot_yes

# Average crime incidents through year with No law
plot_no <- ggplot(guns_no_mean, aes(x = year, y = total_violent)) +
  geom_line(color = "red") +
  labs(title = "Crime Rate Over Time (No Law in effect)", x = "Year", y = "Average Violent Crime") +
  theme_minimal()
plot_no

guns_yes_mean$law <- "Law in Effect"
guns_no_mean$law <- "No Law in Effect"

combined_data <- rbind(guns_yes_mean, guns_no_mean)

plot_mix <- ggplot(combined_data, aes(x = year, y = total_violent, color = law)) +
  geom_line(size = 1) +
  labs(
    title = "Crime Rate Over Time (Law in Effect vs. No Law in Effect)",
    x = "Year",
    y = "Average Violent Crime"
  ) +
  scale_color_manual(values = c("blue", "red")) +
  theme_minimal()
plot_mix

# Boxplot of Crime Incidents over Yes/No Shall-carry Laws
ggplot(Guns) +
  aes(x = law, y = violent, fill = law) +
  geom_boxplot() +
  labs(x = "Concealed-carry laws", y = "Crime Rate(per100,000)", fill = "Concealed-carry laws")

Guns$law_numeric <- ifelse(Guns$law == "yes", 1, 0)

# Normalized relevant variables
Guns_normalized <- Guns %>%
  select(-state, -year, -law) %>%
  mutate(across(where(is.factor), as.numeric)) %>%
  scale() %>%
  as.data.frame()
```

```r
# Correlation matrix
correlation_matrix <- cor(Guns_normalized)

# Correlation matrix
corrplot(correlation_matrix,
         method = "color",
         type = "full",
         col = colorRampPalette(c("lightskyblue", "white", "coral"))(200),
         addCoef.col = "black",
         tl.col = "black",
         tl.srt = 45,
         number.cex = 0.7)


# Pooled model
gun_pooled_formula <- bf(
  violent ~ 1 + prisoners + afam + cauc + male +
            population + income + density + law,
  family = "gaussian",
  center = FALSE
)

(gun_pooled_priors <- c(
  prior(normal(0, 1), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "prisoners"),
  prior(normal(0, 1), class = "b", coef = "afam"),
  prior(normal(0, 1), class = "b", coef = "cauc"),
  prior(normal(0, 1), class = "b", coef = "male"),
  prior(normal(0, 1), class = "b", coef = "population"),
  prior(normal(0, 1), class = "b", coef = "income"),
  prior(normal(0, 1), class = "b", coef = "density"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb= 0)
))

gun_pooled_fit <- brm(
  formula = gun_pooled_formula,
  prior = gun_pooled_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "yes"
)


# Law model
gun_law_formula <- bf(
  violent ~ 1 + law + (1 + law | state) + (1 | year),
  family = "gaussian",
  center = FALSE
)

get_prior(gun_law_formula, data = Guns)

(gun_law_priors <- c(
  prior(normal(0, 1), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb= 0),
  prior(normal(0, 5), class = "sd", group = "state", coef = "Intercept"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "lawyes"),
  prior(normal(0, 5), class = "sd", group = "year", coef = "Intercept")
))

gun_law_fit <- brm(
  formula = gun_law_formula,
  prior = gun_law_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "yes"
)
```

```
# All model
gun_all_formula <- bf(
  violent ~ 1 + prisoners + afam + cauc + male + population + income + density + law +
    (1 + prisoners + afam + cauc + male + population + income + density + law | state) + (1 | year),
  family = "gaussian",
  center = FALSE
)

get_prior(gun_all_formula, data = Guns)

(gun_all_priors <- c(
  prior(normal(0, 1), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "prisoners"),
  prior(normal(0, 1), class = "b", coef = "afam"),
  prior(normal(0, 1), class = "b", coef = "cauc"),
  prior(normal(0, 1), class = "b", coef = "male"),
  prior(normal(0, 1), class = "b", coef = "population"),
  prior(normal(0, 1), class = "b", coef = "income"),
  prior(normal(0, 1), class = "b", coef = "density"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb = 0),

  prior(normal(0, 5), class = "sd", group = "year", coef = "Intercept"),

  prior(normal(0, 5), class = "sd", group = "state", coef = "Intercept"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "prisoners"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "afam"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "cauc"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "male"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "population"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "income"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "density"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "lawyes")
))

gun_all_fit <- brm(
  formula = gun_all_formula,
  prior = gun_all_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "yes"
)


# Load models
gun_law_fit <- readRDS("notebooks/bda2024/gun_law_model.rds")
gun_all_fit <- readRDS("notebooks/bda2024/gun_all_model.rds")


# Convergence diagnostics
mcmc_rhat_hist(rhat(gun_pooled_fit)) + labs(title = "Pooled Model") + theme(plot.title = element_text(hjust = 0.5))
mcmc_rhat_hist(rhat(gun_law_fit)) + labs(title = "Law Model") + theme(plot.title = element_text(hjust = 0.5))
mcmc_rhat_hist(rhat(gun_all_fit)) + labs(title = "All Model") + theme(plot.title = element_text(hjust = 0.5))

mcmc_neff_hist(neff_ratio(gun_pooled_fit)) + labs(title = "Pooled Model") + theme(plot.title = element_text(hjust = 0.5))
mcmc_neff_hist(neff_ratio(gun_law_fit)) + labs(title = "Law Model") + theme(plot.title = element_text(hjust = 0.5))
mcmc_neff_hist(neff_ratio(gun_all_fit)) + labs(title = "All Model") + theme(plot.title = element_text(hjust = 0.5))

rstan::check_hmc_diagnostics(gun_pooled_fit$fit)
rstan::check_hmc_diagnostics(gun_law_fit$fit)
rstan::check_hmc_diagnostics(gun_all_fit$fit)


# Posterior predictive check
pp_check(gun_pooled_fit) + labs(title = "Pooled Model") + theme(plot.title = element_text(hjust = 0.5))

pp_check(gun_law_fit) + labs(title = "Law Model") + theme(plot.title = element_text(hjust = 0.5))

pp_check(gun_all_fit) + labs(title = "All Model") + theme(plot.title = element_text(hjust = 0.5))


# Model comparison
pooled_loo <- loo(gun_pooled_fit)
law_loo <- loo(gun_law_fit)
```

```
all_loo <- loo(gun_all_fit)
pooled_loo
law_loo
all_loo
loo_compare(pooled_loo, law_loo, all_loo)

# k diagnostics
plot(
  pooled_loo,
  diagnostic = c("k"),
  label_points = FALSE,
  main = "Pooled-Model diagnostic plot"
)

plot(
  law_loo,
  diagnostic = c("k"),
  label_points = FALSE,
  main = "Law-Model diagnostic plot"
)

plot(
  all_loo,
  diagnostic = c("k"),
  label_points = FALSE,
  main = "All-Model diagnostic plot"
)


# Prior sensitivity analysis
# Pooled model change
gun_pooled_c_formula <- bf(
  violent ~ 1 + prisoners + afam + cauc + male +
    population + income + density + law,
  family = "gaussian",
  center = FALSE
)

(gun_pooled_c_priors <- c(
  prior(normal(3, 5), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "prisoners"),
  prior(normal(0, 1), class = "b", coef = "afam"),
  prior(normal(0, 1), class = "b", coef = "cauc"),
  prior(normal(0, 1), class = "b", coef = "male"),
  prior(normal(0, 1), class = "b", coef = "population"),
  prior(normal(0, 1), class = "b", coef = "income"),
  prior(normal(0, 1), class = "b", coef = "density"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb= 0)
))

gun_pooled_c_fit <- brm(
  formula = gun_pooled_c_formula,
  prior = gun_pooled_c_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "yes"
)

m_before <- tibble(value = posterior_samples(gun_pooled_fit)$prior_b_Intercept)
m_after <- tibble(value = posterior_samples(gun_pooled_c_fit)$prior_b_Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "Pooled Model",
    x = "Prior Mean",
    y = "Density"
  ) +
  theme_minimal()

m_before <- tibble(value = posterior_samples(gun_pooled_fit)$b_Intercept)
```

```
m_after <- tibble(value = posterior_samples(gun_pooled_c_fit)$b_Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "Pooled Model",
    x = "Posterior",
    y = "Density"
  ) +
  theme_minimal()




# Law model change
gun_law_c_formula <- bf(
  violent ~ 1 + law + (1 + law | state) + (1 | year),
  family = "gaussian",
  center = FALSE
)

get_prior(gun_law_formula, data = Guns)

(gun_law_c_priors <- c(
  prior(normal(3, 5), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb= 0),
  prior(normal(2, 3), class = "sd", group = "state", coef = "Intercept"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "lawyes"),
  prior(normal(0, 5), class = "sd", group = "year", coef = "Intercept")
))

gun_law_c_fit <- brm(
  formula = gun_law_c_formula,
  prior = gun_law_c_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "yes"
)

gun_law_fit_prior <- brm(
  formula = gun_law_formula,
  prior = gun_law_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "only"
)

m_before <- tibble(value = posterior_samples(gun_law_fit_prior)$b_Intercept)
m_after <- tibble(value = posterior_samples(gun_law_c_fit)$prior_b_Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "Law Model",
    x = "Prior Mean",
    y = "Density"
  ) +
  theme_minimal()

m_before <- tibble(value = posterior_samples(gun_law_fit_prior)$sd_state__Intercept)
m_after <- tibble(value = posterior_samples(gun_law_c_fit)$prior_sd_state__Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
```

```r
    title = "Law Model",
    x = "Prior Sigma",
    y = "Density"
  ) +
  theme_minimal()

m_before <- tibble(value = posterior_samples(gun_law_fit)$sd_state__Intercept)
m_after <- tibble(value = posterior_samples(gun_law_c_fit)$sd_state__Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "Law Model",
    x = "Posterior",
    y = "Density"
  ) +
  theme_minimal()




# All model change
gun_all_c_formula <- bf(
  violent ~ 1 + prisoners + afam + cauc + male + population + income + density + law +
    (1 + prisoners + afam + cauc + male + population + income + density + law | state) + (1 | year),
  family = "gaussian",
  center = FALSE
)

get_prior(gun_all_c_formula, data = Guns)

(gun_all_c_priors <- c(
  prior(normal(3, 5), class = "b", coef = "Intercept"),
  prior(normal(0, 1), class = "b", coef = "prisoners"),
  prior(normal(0, 1), class = "b", coef = "afam"),
  prior(normal(0, 1), class = "b", coef = "cauc"),
  prior(normal(0, 1), class = "b", coef = "male"),
  prior(normal(0, 1), class = "b", coef = "population"),
  prior(normal(0, 1), class = "b", coef = "income"),
  prior(normal(0, 1), class = "b", coef = "density"),
  prior(normal(0, 1), class = "b", coef = "lawyes"),
  prior(normal(0, 2), class = "sigma", lb = 0),

  prior(normal(0, 5), class = "sd", group = "year", coef = "Intercept"),

  prior(normal(0, 1), class = "sd", group = "state", coef = "Intercept"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "prisoners"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "afam"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "cauc"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "male"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "population"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "income"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "density"),
  prior(normal(0, 5), class = "sd", group = "state", coef = "lawyes")
))

gun_all_c_fit <- brm(
  formula = gun_all_c_formula,
  prior = gun_all_c_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "yes"
)

gun_all_fit_prior <- brm(
  formula = gun_all_formula,
  prior = gun_all_priors,
  data = Guns,
  iter = 2000,
  warmup = 1000,
  chains = 4,
  sample_prior = "only"
```

```r
)

m_before <- tibble(value = posterior_samples(gun_all_fit_prior)$b_Intercept)
m_after <- tibble(value = posterior_samples(gun_all_c_fit)$prior_b_Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "All Model",
    x = "Prior Mean",
    y = "Density"
  ) +
  theme_minimal()

m_before <- tibble(value = posterior_samples(gun_all_fit_prior)$sd_state__Intercept)
m_after <- tibble(value = posterior_samples(gun_all_c_fit)$prior_sd_state__Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "All Model",
    x = "Prior Sigma",
    y = "Density"
  ) +
  theme_minimal()

m_before <- tibble(value = posterior_samples(gun_all_fit)$sd_state__Intercept)
m_after <- tibble(value = posterior_samples(gun_all_c_fit)$sd_state__Intercept)
m_before$type <- "Original"
m_after$type <- "Modified"
ggplot(rbind(m_before, m_after), aes(x = value, fill = type)) +
  geom_density(alpha = 0.5) +
  scale_fill_manual(values = c("skyblue", "coral")) +
  labs(
    title = "All Model",
    x = "Posterior",
    y = "Density"
  ) +
  theme_minimal()


# Save models
saveRDS(gun_law_fit, file = "gun_law_model.rds")
saveRDS(gun_all_fit, file = "gun_all_model.rds")
```