Analysis of Law Enforcement Transparency in Fatal Police Shootings*

In the United States from 2015 to 2024

Zhixi(Arno) Qu

November 30, 2025

This study analyzes the relationship between demographic characteristics of victims, case situation, and whether police activated body cameras based on U.S. fatal police shooting data. Data were sourced from fatal police shootings dataset of the Washington, primarily covering cases occurring between 2015 and 2024. Two census datasets from the U.S. Census Bureau were also utilized to analyze annual changes in per-million fatality rates. Most importantly, Bayesian Logistic Regression model was used to analyze trends and factors influencing whether body cameras were activated in each case. Results indicate that the minor victims and victim had mental health illness significantly increases the probability of body camera activation during cases. This study contributes to a more detailed analysis of current unfairness in law enforcement transparency and provides a basis for future targeted interventions and police training.

Table of contents

*Code and data are available at: https://github.com/ZhixiQu/US-Fatal-Police-Shootings.git

1 Introduction

In recent years, fatal shootings by U.S. police officers during law enforcement operations have consistently drawn public attention and have been widely debated. The large number of such cases has prompted society and the public to focus not only on the situations under which police may use fatal force, but also on the extent to which these cases have been recorded and made public. Under this background, body cameras have been used as a key tool for police departments to maintain public trust and provide professional, fair policing services. In social surveys, both citizens and police officers have recognized the importance of body cameras in law enforcement (2020). However, Equipping police officers with body cameras does not guarantee that all police incidents will be recorded and made public, especially in extreme cases such as fatal police shootings. In actual policing, whether officers activate their body cameras is often influenced by individual factors and situations. While body cameras may be more commonly used during certain types of enforcement activities, their usage rates are notably lower in specific scenarios or when interacting with particular groups. This may lead to unequal levels of policing transparency and accountability for police officers or victims across different populations and situations.

Current reports and studies on body-worn cameras are limited to the ability of wearing them to enhance law enforcement transparency (2024). However, what truly concerns society is whether these cameras are genuinely and fairly activated to improve police transparency and public trust. To address this gap, this paper analyzes changes in U.S. police fatal shooting rates using U.S. police fatal shooting data and US Census Data. Most importantly, uing the police fatal shooting data to analyse whether different policing scenarios and demographic groups systematically influence body camera activated status during fatal police shootings. Specifically, I utilize U.S. Police Fatal Shooting Database From The Washington Post (2025) combined with two Census data from U.S. Census Bureau during 2010 and 2025 (2021) (2025). First, I provide an overview of changes in U.S. police fatal shooting rates from 2015 to 2024. Besides, I will treat the probability of body camera was activated as the dependent variable in police-involved fatal incidents and fit a Bayesian Logistic Regression model. This model will include victim race, gender, age grouping variables established based on the victim's age, whether the victim was armed, whether the victim fled, whether the victim had mental health issues, and the year as covariates. Furthermore, an interaction term between victim race and whether the victim was armed will be included to examine systemic associations between different situations and demographic characteristics and body camera activated status. Overall, this study quantifies police transparency in fatal shootings to complement the debates on policing transparency and accountability.

Empirical findings indicate that both the rate of fatal police shootings per million people and the activation rate of body-worn cameras during such incidents showed an overall upward trend from 2015 to 2024. the most important finding is that the activation of body-worn cameras exhibits systemic variation across specific demographic groups. When victims have mental health conditions, body-worn cameras are more likely to be activated. Furthermore, older individuals are less likely to be recorded by body-worn cameras during incidents compared to those who are minors. Incomplete data such as when the victim's race is unknown which correlates with lower body camera activation probality. The interaction between victim race and whether the victim was armed was not statistically significant in this study, it indicates there is no stable and significant systemic differences in body camera usage across racial groups in armed scenarios within currently available public fatality data. Overall, Overall, this study found uneven increases in law enforcement transparency over time and their distribution across specific groups of victims. However, due to limitations of the data, all findings are presented as descriptive correlations rather than causal inferences.

The structure of the paper is divided into four main sections. Section 2 discusses the selected data, data pre-processing, and key variables for analysis. Section 3 discusses the model comparison process that ultimately determined the final model for this study's analysis. Section 4 presents the model results, trends observed in the data, and variables associated with whether body cameras were activated. Section 5 draws conclusions from the models and relates them to real-world scenarios to evaluate the reliability of these conclusions. Additionally, it describes the limitations and weaknesses of this study and suggests directions for future improvements.

2 Data

2.1 Overview

In this analysis, I merged the three datasets into two separate datasets. The preliminary study (1) Change in the rate of fatal force caused by U.S. police per million population. Combined police fatality incidents with U.S. Census data to analyze changes in the rate of fatal police shootings during 2015 and 2024. The main focus (2) Analyzed differences in body camera activation probabilty across different groups of victims and scenarios in U.S. fatal police shootings. All data cleaning and analysis were performed in the statistical programming language R (R Core Team, 2023), primarily utilizing packages such as tidyverse (Wickham et al., 2019), lubridate (2024), readxl (2025), and forcats (Wickham, 2023) for data organization and visualization.

The U.S. police fatality data for this study originates from The Washington Post's "data police shooting" database. Since 2015, researchers at The Washington Post have continuously documented fatal shootings by researchers at The Post. The database has two versions. Version 2 data was migrated from Version 1 in 2022 and is continuously updated, with the latest update occurring in May 2025. After this point, Version 1 data will no longer be updated, so I decided to use Version 2 data. The other two data sources are state population total data from the U.S. Census Bureau: NST-EST2020 (covering 2010–2020) and NST-EST2024 (covering 2020–2024), which provide annual population estimates for each state as of July 1st. After filtering, these population data were used to calculate the fatal police shootings from 2015 to 2024. Similar datasets exist that provide nationwide statistics on police use-of-force incidents through collaboration between the FBI and law enforcement agencies. These include police-involved fatalities (FBI 2024), but crucially, data reporting by local agencies is voluntary, leading to volunteer bias. Besides, the absence of body camera data made this dataset incompatible with my primary research focus, so I finally did not use it.

2.2 US police fatality rate and body camera activated rate

For two population census datasets from the U.S. Census Bureau, I extracted annual population estimates for each state as of July 1st from 2015 to 2019 in NST-EST2020, because the 2020 census data in this dataset is less accurate compared to the NST-EST2024 dataset. From the NST-EST2024 dataset, I extracted annual population estimates for each state as of July 1st from 2020 to 2024. Additionally, I mapped the state names in both files to their two letter state abbreviations. The final dataset is named "state_year_pop" and contains annual population estimates for 50 states and the District of Columbia from 2015 to 2024. It includes three new variables: state, year, and pop.

Each row in the original U.S. police fatality dataset organized by The Washington Post represents a victim shot by police in a fatal shooting case. It includes the date and location of the incident, along with several characteristics of the victim and the situation. First, dates were converted to Gregorian calendar years, and to limit the analysis within the 2015–2024 period. To avoid duplicate counting, only unique incident IDs were kept. Death counts were calculated based on state*year. Finally, the fatal case data and annual population data were merged to create state_year_rates, and the police fatality rate per million people was calculated. This dataset includes the following variables: - state: Two letter state abbreviation. - year: Gregorian calendar year. (2015–2024) - death: Number of fatal police shootings in that state during that year. - pop: Total population estimation of that state during that year. - rate_per_million: Police fatality rate per million residents.

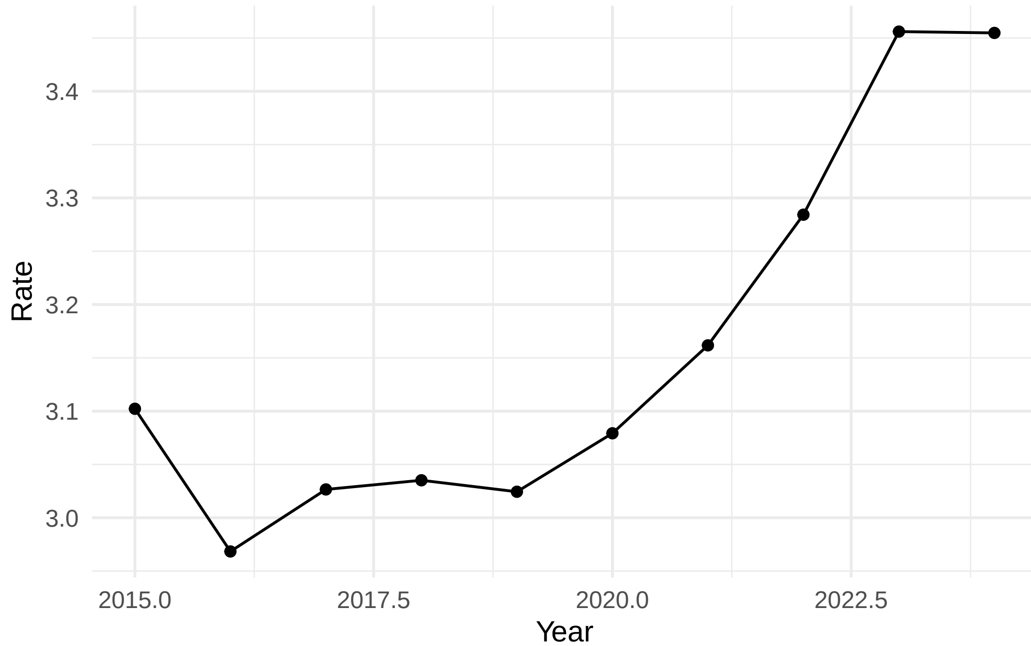


Figure 1: U.S. Fatal police shooting rate over time per million (rate)

Figure 1 shows the time trend of fatal police shootings rate across the United States from 2015 to 2024. It is evident that the rate of fatal police shootings in the U.S. is generally increasing. Although there was a short period of decline between 2015 and 2020, the decrease was not significant and did not appear long. Notably, from 2020 to May 2022, the rate of police fatal shootings experienced rapid and continuous growth and no clear decrease observed after late 2022. This indicates that in high-risk cases the risk of people when interacting with police has increased in recent years and reflects why society continues to focus on fatal police shootings.

2.3 Law enforcement transparency analysis data

The second analytical dataset is also constructed based on the above U.S. police fatality data, used to model whether body cameras were activated during fatal police shooting. When handling missing values, I only deleted entire rows of observations for categorical variables or critical variables. For other variables, I retained the sample by labeling missing values as “Unknown” whenever possible. Specifically, if the missing value pertained to whether the body camera was turned on (the dependent variable) or age (used to construct the crucial covariate of age groups), arbitrary definitions could undermine the analysis’s credibility, so such observations were excluded from the sample. However, for variables such as race, gender, weapon information, and mental health status, missing data were coded as “Unknown.” These observations were retained in charts and models, significantly reducing data loss while ensuring the accuracy and credibility of analytical results. The following variables were constructed:

The second analytical dataset is also constructed based on the above U.S. police fatality data, used to model whether body cameras were activated during fatal police incidents. When processing missing values, I only deleted entire rows of observations for categorical variables or key variables. For other variables, I kept the sample by using “Unknown” where possible. Specifically, if the missing value was whether the body camera was activated (the dependent variable) or age (used to construct the important covariate of age group), arbitrary definitions could undermine the credibility of the analysis, so these observations were excluded from the sample. However, for variables such as race, gender, weapon information, and mental health status, missing data were coded as “Unknown.” These observations were kept in charts and models, significantly reducing data loss while ensuring the accuracy and credibility of analytical results. The following variables were constructed: - body_camera_binary: Assigns 1 when the body camera is activated, and 0 when body camera was off. Since missing data for the dependent variable would cause model fitting to fail, I chose to delete the missing value. - race: The race of victim. In the original data, some records had multiple race categories separated by semicolons. To simplify analysis, I kept only the first race category, treating it as the primary category for that individual. Missing values were defined as “Unknown.” I preserved the original data values such as “W” means the race of victim is white. - gender: Gender of victim. The value of original data are male, female, or non-binary. Missing values were defined as “Unknown”. - age_group: Victim’s age group at the time of case. First, raw ages were converted to numeric values, then grouped into five intervals: under 18, 18–29, 30–44, 45–59, and 60+. Observations with missing age (unable to be grouped) were excluded from the analysis sample. - whether_mental_ill: Assigns “yes” if the victim had a history of mental health issues or was experiencing mental distress at the time of the shooting. Assigns ‘no’ if the original variable indicates no mental illness association. If the value is missing, it is defined as “Unknown” to keep these observations and mark the information as missing. - armed: Whether the victim had a high-threat weapon. If the victim was unarmed during the incident, there was no evidence confirming weapon possession or the victim held only a non-functional firearm, this variable is assigned as “no”. If the victim’s weapon status is missing, it is assigned as “Unknown.” In all other cases, it is assigned as “yes”. - flee: Whether the victim fled. If the victim did not flee assigned “no”. If any form of flight occurred, value is yes. Since there is no missing value so do not need further step. In addition to the above constructor variables, I have kept the following identification and context variables: - id: A unique identifier for each fatal police shooting incident. - state: The two-letter abbreviation of the state where the incident occurred. - year: The year the incident occurred.

Figure 2 shows the proportion of body cameras that were activated during fatal police shootings over the same period. Overall, the activation rate of body cameras has fluctuated but shown a significant increase. Moreover, the peak activation rate over the past decade has been approximately 32%. This indicates that the use of body cameras in U.S. police fatal shootings is increasing but remains unstable, potentially due to insufficient officer training or the environment of the police operation. Furthermore, the fact that the proportion of body camera activated

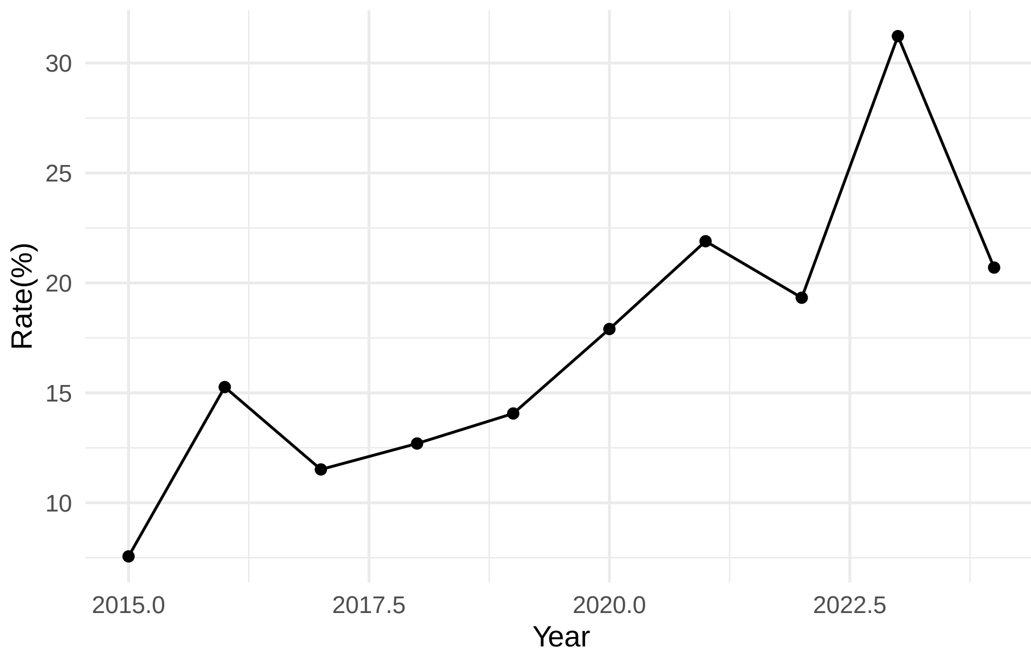
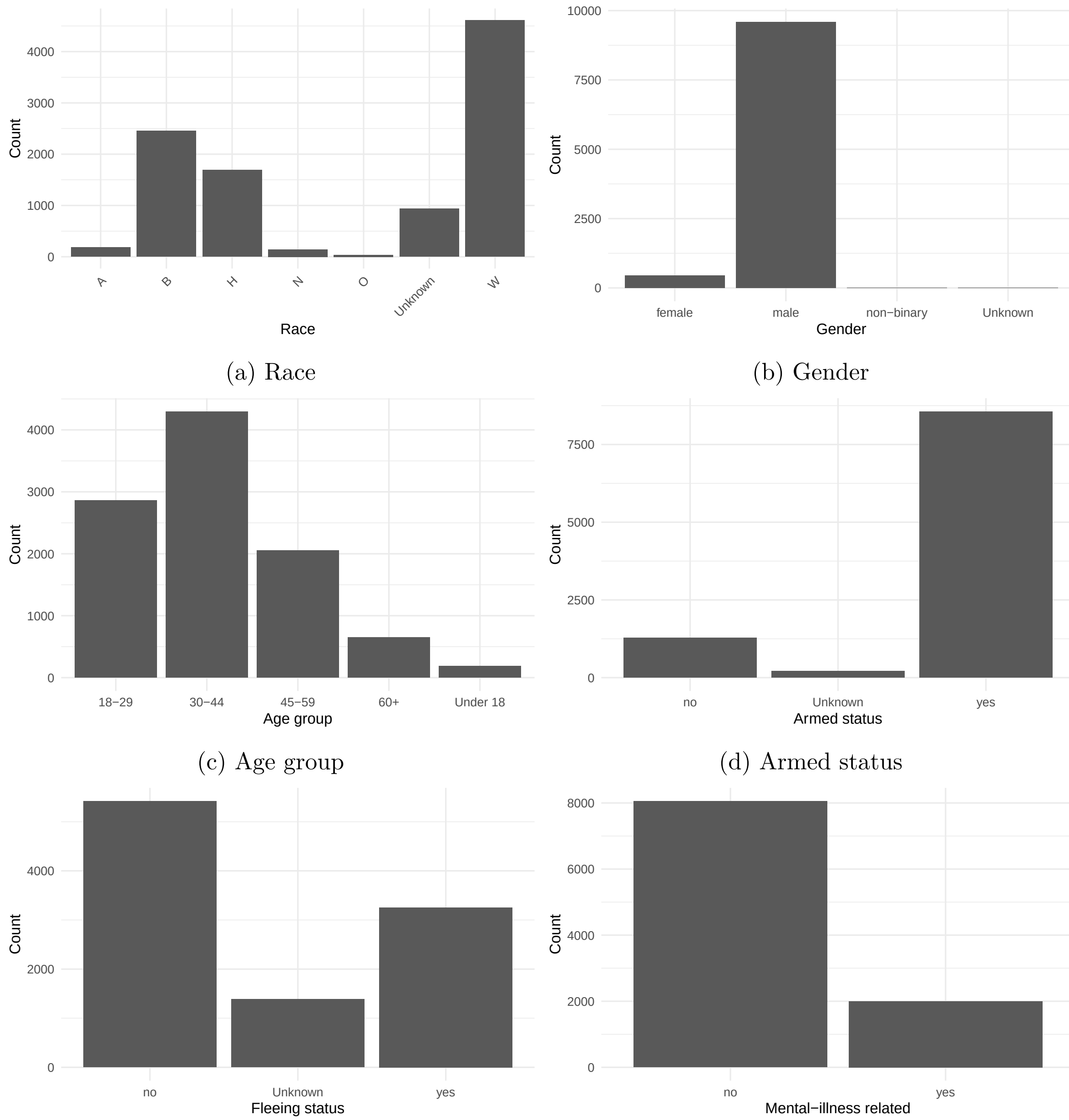


Figure 2: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

in fatal shootings remains below 50% suggests significant debate over police transparency in such cases over the past decade. Notably, the increase in proportion of body camera activated between 2015 and 2017 and the decrease in fatal shootings rate indicate a potential negative correlation between body camera deployment and fatal shooting rates. Therefore, researching factors influencing body camera usage provides empirical evidence to enhance public trust in police and develop more effective regulatory and training strategies.

2.4 Predictor variables

Figure 3 shows the variables I will include in the model. In U.S. police fatal shootings, victims are majority white and male. The age group distribution is close to normal, with victims primarily concentrated between 30 and 44 years old. Most victims were documented as possessing high-threat weapons, and the majority did not attempt to flee. Cases involving mental health issues for the victim were relatively less common, with more cases involving victims without mental health issues. Notably, sample sizes were uneven across groups, but this is common in extreme cases such as fatal police force data. Therefore, this study remains valuable. After gaining an initial understanding of the quantitative differences across variables, it is necessary to further examine the proportion of body camera recordings activated under different victim and case characteristics.

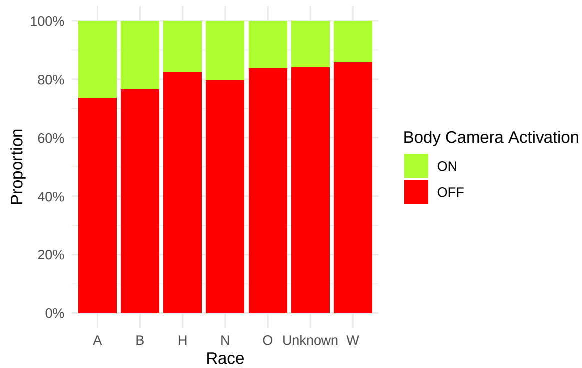


(a) Race (b) Gender

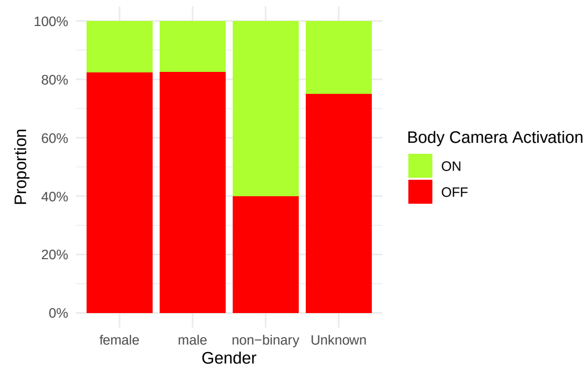
(c) Age group (d) Armed status

(e) Fleeing status (f) Mental-illness related

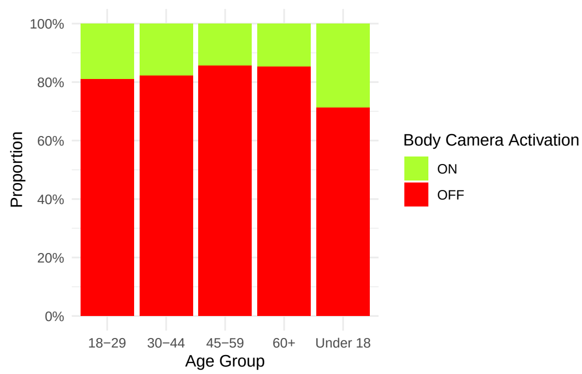
Figure 3: Distributions of victim and incident characteristics-main predictor variable



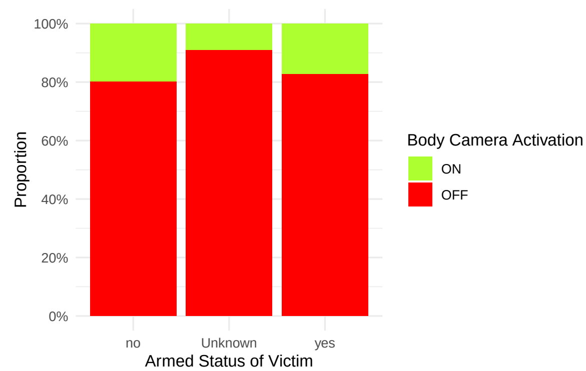
(a) Race



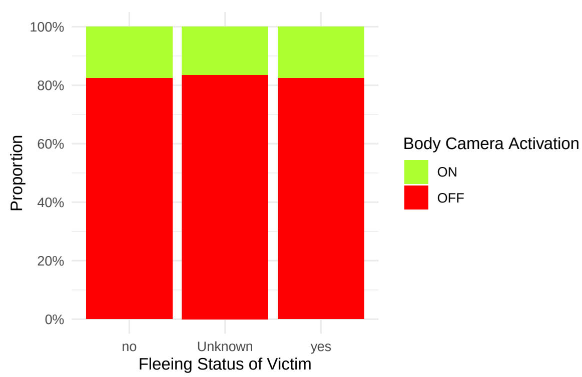
(b) Gender



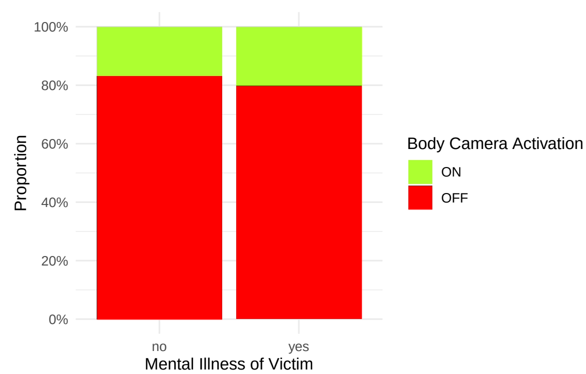
(c) Age group



(d) Armed status



(e) Fleeing status



(f) Mental-illness related

Figure 4: Proportion of body camera in each main predictor variable

Figure 4 shows the proportions of body cameras activated during U.S. police fatal shootings across different variables. Most obviously, the activation rate remains low across all variables, consistent with the annual trends observed earlier. The difference in body camera activation rates across the variables is generally not significant, but the direction of this difference is the primary purpose of this study. For example, regarding the mental health status of the victim, cases involving victims with mental health issues showed a slightly higher proportion of body cameras being activated. Although the difference in the proportion is not large, as society increasingly focuses on interactions between individuals with mental illness and police,(2021) the direction of this difference is very meaningful in real world. This indicates that in fatal cases involving victims with mental health issues, police officers may be more motivated to activate body camera. Overall, there are slight and directional variations in body camera activation proportions across different demographics and enforcement scenarios. These differences may be due to sample quality or the victim and enforcement situation. Visualizing alone cannot fully explain these variations, therefore models are needed to analyze the factors influencing body camera activation.

3 Model

This study will use the Bayesian Logistic Regression model to analyze which situational and individual characteristics are associated with the activation of body cameras in fatal police shootings in the United States. Since the dependent variable is binary (whether the body camera was activated) and police shootings are a socially sensitive issue, predictions cannot be made based on the data alone, and the reliability of the predictions must be ensured. The model employs a binomial distribution as the likelihood function and uses a logit link function to ensure predicted values which is the probability of activation of body camera fall within the range of 0 to 1.

3.1 Model set-up

The observation unit is each fatal police shooting case in The Washington Post dataset, denoted as i for the i_{th} case. If the police officer activated the body camera during the incident, $y_i = 1$; otherwise, $y_i = 0$. pi represents the probability of the body camera being activated in incident i. Given pi, the outcome of each incident follows a Bernoulli distribution:

$$y_i \mid p_i \sim \text{Bernoulli}(p_i)$$

Also, using the logit link function to fit the final model:

$$\begin{aligned}\log\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right) &= \beta_0 + \beta_1\,\text{race}_i + \beta_2\,\text{gender}_i + \beta_3\,\text{age}_i + \beta_4\,\text{armed}_i + \beta_5\,\text{flee}_i \\ &+ \beta_6\,\text{mental\_health}_i + \beta_7\,\text{year}_i + \beta_8\,\text{race}_i \times \text{armed}_i\end{aligned}$$

Prior distribution of parameter:

$$\beta_0 \sim \text{Normal}(0, 2.5)$$
$$\beta_1 \sim \text{Normal}(0, 2.5)$$
$$\beta_2 \sim \text{Normal}(0, 2.5)$$
$$\beta_3 \sim \text{Normal}(0, 2.5)$$
$$\beta_4 \sim \text{Normal}(0, 2.5)$$
$$\beta_5 \sim \text{Normal}(0, 2.5)$$
$$\beta_6 \sim \text{Normal}(0, 2.5)$$
$$\beta_7 \sim \text{Normal}(0, 2.5)$$
$$\beta_8 \sim \text{Normal}(0, 2.5)$$

where $\beta_1, ..., \beta_8$ means that holding other variables constant, on average, how much does the log-odds ratio change when the variable or interaction term corresponding to parameter increases by one unit.

- $\hat{p}$ represents the probability that the body camera were activated in each case.
- β_0 represents the intercept term of this logistical regression. It is the probability that body cameras being activated when an individual belongs to the reference group for each categorical variable and the year is 2015.
- β_1 is the coefficient corresponding to race of the victim.
- β_2 is the coefficient corresponding to gender of the victim.
- β_3 is the coefficients corresponding to age group of the person, Age group is categorical variable grouped by the victim's age. The categories are: Under 18, "18–29" "30–44" "45–59" "60+".
- β_4 is the coefficients corresponding to whether the victim is armed.
- β_5 is the coefficients corresponding to whether the victim is fleeing.
- β_6 is the coefficients corresponding to whether the victim had a history of mental health issues.
- β_7 is the coefficients corresponding to the year the event occurred.
- β_8 is the coefficients corresponding to the interaction term between the race of victim and whether the victim were armed.

For the Bayesian model, I set a weakly informative prior of the same form for $\beta_0, ..., \beta_8$. The prior mean is set to 0 because there is no significant difference in body camera activation rates between groups based on the variable plots. The prior standard deviation is set to 2.5 because I have no strong prior beliefs about the direction or magnitude of the covariates before observing the data. Most importantly, the sample sizes for different categories vary greatly. A moderately appropriate prior distribution provides more stable estimates and prevents overfitting. To avoid extreme or unstable estimates, I moderately shrink the prior distribution rather than using a non-information prior. A non-information prior would allow the posterior distribution

Table 1

Table 2: Table of Compare Models Based on Accuracy and AUC

model	Accurace	Area_under_curve
Bayesian logistic	0.82	0.66
Logistic	0.82	0.66
Simple GLM	0.82	0.66
Complex GLM	0.82	0.66

to be entirely determined by the data, which is inappropriate for uneven sample sizes across categories in this study.

Bayesian Logistic regression requires certain assumptions. First, samples must be independent given the covariates. However, it is very difficult to satisfy data independence in reality. The same state or police department may have similar enforcement requirements and cultural environments, leading to potential correlations between samples. Therefore, given these covariates, events can be approximated as conditionally independent but not strictly independent. Based on my current understanding, only the Linear Mixture Model could better address sample correlations. However, the dependent variable in this study is whether body cameras were activated which is a binary variable, making the Linear Mixture Model unsuitable for this problem. Realistically, data and models may not capture all potential variables, such as some police do not respect with law enforcement guidelines. Most importantly, processing errors may occur during data collection and recording. (2025) Therefore, the findings of this study will be descriptive associations rather than strict causal inferences. Finally, this research focuses on extreme fatality cases, so its conclusions may not be appropriate to generalize ordinary police encounter situations.

The study uses the `rstanarm` package of Goodrich et al. (2022) in R (R Core Team 2023) to fit a Bayesian Logistic Regression Model via Markov Chain Monte Carlo (MCMC) sampling from the posterior distribution of coefficients. After model fitting, I first checked basic convergence diagnostics, such as whether trace plots exhibited uniform shaking and whether the $\widehat{R}$ of each parameter approached 1, to ensure the MCMC chain had converged. To evaluate model performance, data were randomly split into training (70%) and testing (30%) sets. The model was fitted on the training set, and predictions were made on the testing set using the posterior mean. Approximate prediction accuracy and Area Under the ROC Curve (AUC) were calculated. The results indicate that the model has some predictive capability regarding the activation of body cameras.

To evaluate the quality of model fit and determine whether the model is significantly overfitted or underfitted, I compared the Bayesian logistic regression used in the main analysis with several alternative models. These included Logistic Regression using the same covariates, a simpler Bayesian Logistic Regression containing only year and primary demographic variables, and a more complex Bayesian Logistic Regression with additional interaction terms. Table 1 shows the prediction accuracy and area under the ROC curve (AUC) for each model on the independent test set. Results show that all four models achieved accuracy around 0.82 and AUC around 0.66. This indicates that in the current analysis data, the impact of any alternative models or variants on overall predictive performance is very limited. However, Bayesian logistic regression can directly provide posterior distributions and credible intervals for coefficients, making it more beneficial for quantifying and interpreting uncertainty. Therefore, Bayesian Logistic Regression is the final model choice in this study. More diagnostics and graphs will be included in the Section A.

4 Results

Table 3 presents the coefficient estimates and 95% credible intervals for the log-odds of whether body camera was activated in the Bayesian logistic regression model. `conf.low` and `conf.high` represent the lower and upper bounds of the 95% credible intervals, respectively. The 95% credible intervals indicate that there is 95% probability that the true value lies within this range, given that the model assumptions hold. In the model, categorical variables are coded using a reference group: race uses `Asian` as the reference group, gender uses `female` as the reference, age uses `under 18` as the reference, armed uses `unarmed` as the reference group, flee uses `no`(not flee) as the reference group, and mental health variables use `no` (no mental illness) as the reference. Thus, each coefficient represents the average difference between corresponding group in the log odds of activating the body camera relative to the reference group, holding all other variables constant. If the 95% credible interval is entirely below 0, this indicates strong evidence that the probability of police activated body cameras is lower for this category relative to the reference group. If the interval is entirely above 0, it corresponds to a higher activation probability. An interval crossing 0 indicates weaker evidence or no clear evidence.

To provide a more visual representation of how predictor variables influence the log-odds of whether body camera was activated, I created a visualization based on the previous summary table of the model. The vertical axis displays different race, gender, age group, and situational variables. The horizontal axis shows the strength of the relationship between the predictor variable and log-odds of body camera activation. Specifically, it indicates the different in the predictor variables relative to the reference group regarding body camera activation, holding all other variables constant. Each point represents the estimated coefficient for a predictor variable, with the horizontal line indicating the 95% credible interval around the estimate.

Regarding victim age, with those under 18 as the reference group, Figure 5 shows that coefficients and intervals for all higher age groups fall to the left of zero. This indicates that the probability of activating body cameras is generally lower when victims are older individuals. Particularly within the 18-59 age group, the proportion of cameras activated statistically decreases with increasing age. The coefficient for victims with mental illness is positive, with intervals entirely above zero. This indicates that when victims have mental illness, the probability of the body

Table 3: Coefficient Estimates and 95% Credible Intervals

term	estimate	conf.low	conf.high
(Intercept)	-318.37	-360.37	-277.78
gendermale	0.02	-0.23	0.28
gendernon-binary	1.41	-0.46	3.27
genderUnknown	0.19	-2.43	2.23
age_group30–44	-0.10	-0.22	0.03
age_group45–59	-0.26	-0.42	-0.09
age_group60+	-0.23	-0.49	0.02
age_groupUnder 18	0.47	0.13	0.80
fleeUnknown	-0.38	-0.55	-0.22
fleeyes	-0.09	-0.21	0.04
whether_mental_illyes	0.40	0.27	0.53
year	0.16	0.14	0.18
raceB	0.19	-0.67	1.09
raceH	-0.47	-1.37	0.43
raceN	-0.77	-2.10	0.49
raceO	0.14	-1.78	1.82
raceUnknown	-0.97	-1.93	0.02
raceW	-0.52	-1.37	0.35
armedUnknown	-0.37	-2.41	1.40
armedyes	0.02	-0.86	0.93
raceB:armedUnknown	-0.61	-2.57	1.53
raceH:armedUnknown	0.59	-1.35	2.74
raceN:armedUnknown	-0.61	-5.14	3.56
raceUnknown:armedUnknown	-1.46	-5.69	1.70
raceW:armedUnknown	-0.16	-2.11	1.92
raceB:armedyes	-0.39	-1.34	0.50
raceH:armedyes	-0.07	-1.03	0.89
raceN:armedyes	0.58	-0.76	2.03
raceO:armedyes	-1.17	-3.15	0.96
raceUnknown:armedyes	0.13	-0.90	1.14
raceW:armedyes	-0.28	-1.19	0.62

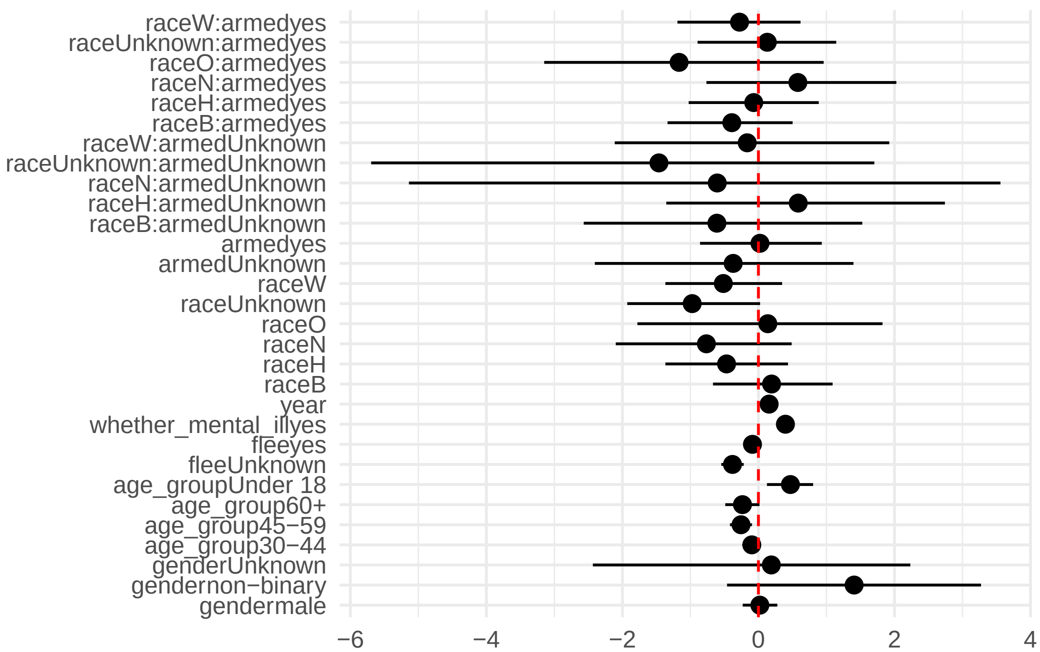


Figure 5: Coefficient Estimates and 95% Credible Intervals

camera being activated is significantly higher compared to victims without mental illness. This aligns with the previously mentioned societal focus on interactions between individuals with mental illness and police. Furthermore, the usage rate of body cameras has increased significantly over the years. However, under this dataset, the model did not find highly stable and significant associations between victim gender, whether the victim was armed, or whether the victim fled, and the activation of body-worn cameras. For example, compared to Asian victims, the 95% credible intervals for the coefficients of other race of victim included zero, indicating that after controlling for other variables, the probability of activating body cameras in these cases differed from that of Asian victims but was not significantly different. Finally, I added an interaction term between victim race and whether the victim was armed (race × armed) to the main model to examine whether the effect of whether the victim was armed differs across races. Specifically, this allows us to investigate whether there are systematic differences in the probability of body cameras activation across races in U.S. fatal police shooting, when the victim's armed status is the same. The 95% credible intervals for most interaction coefficients cross zero and were relatively wide, indicating that in this data, controlling for other variables, there was no strong evidence to conclude that there is a systemic difference in the impact of armed status of victim on body camera activation across race groups.

5 Discussion

5.1 Paper Overview

This paper examines the outcome of "whether body cameras were activated" in U.S. fatal police shootings. Using The Washington Post's 2015–2024 database of fatal police shootings and combining it with U.S. Census Bureau census data from 2015 to 2024.It constructs trends in both the annual rate of fatal police shootings and the annual activation rate of body cameras during such cases. This demonstrates the purpose of this study and why the standardization and transparency of police conduct in fatal shootings have become enduring topics of societal debate in the United States. Most importantly, based on the data structure and after comparing different models, the final choice is that using Bayesian logistic regression model to analyze whether different enforcement scenarios and demographic categories have a systematic impact on enforcement transparency. The model treats body camera activation as the dependent variable, and uses covariates including victim race, gender, age group, whether armed, whether fleeing, whether the victim had mental health issues, and year of cases. Besides, an interaction term between race and whether armed of victim was also included. Because the observational origin of the fatal police shooting data and some assumptions are not strictly satisfied such as the example is independent with each other, the final results represent descriptive associations rather than causal inferences.

5.2 Rising Patterns of Law-Enforcement Transparency

First, after controlling for other variables, the year and the activation status of body cameras showed a positive correlation. The final analytical results of the model align with the visual data. This indicates that the probability of body cameras being activated between 2015 and 2024 is overall increasing. Although fluctuations are visible in the visual data of body camera activation rates over time, the overall trend is upward. This finding is consistent with the context of U.S. states continuously increasing the purchase and use of body cameras since 2016 (2022). Although this finding superficially supports the continued improvement of police transparency after 2015, this increase in transparency is not homogeneous across all groups and scenarios.

Second, body cameras are more likely to be activated in cases associated with mental health issues. After controlling for victim race, gender, age, whether armed, and whether fled, the coefficient of whether the victim had a mental health illness is significantly positive. This indicates that given the same enforcement situation, if the victim had mental health illness, the body camera activation probability in fatal police shooting is significantly higher than in other victim who did not have mental health illness. This result relates to the growing societal attention to mental health and policing responses mentioned in the Section 2. But more importantly, police interactions with individual has mental illness are more protective than their interactions with the general individual. (2021) This indicates law enforcement agencies are increasingly aware that mishandling incidents involving victims with mental health conditions often leads to intense public debate and subsequent accountability. Therefore, there are stronger motivations for police and their agencies to ensure transparency in the handling of such cases. This indirectly leads to a significantly higher body camera activation probability in extreme case when the victim has mental health issues compared to other cases.

Third, probability of body camera activation has significant systemic differences across victim age groups in fatal police shooting case. Using victims under 18 as the reference group, all coefficients for older age groups are negative with 95% credible intervals excluding zero. This indicates that controlling for other variables, the probability of camera activation generally lower with larger age of victim. This does not imply that police is unwilling to record interactions with victims over 18 during enforcement. Rather, it is more likely that in fatal police shooting case, polce tend to protect the legal rights of juveniles by preferring to activate body cameras to ensure transparency and accountability. This leads in a significantly lower overall activation rate across other age groups.

Body cameras not only standardize police conduct but also protect the fairness of police accountability. During interactions involving vulnerable groups of prople such as under 18, camera activation probability are often significantly higher. From perspective of relative social fairness, this is not always bad, and vulnerable groups often receive greater attention and consideration in public debate. However, from the perspective of law enforcement agencies, cases involving vulnerable groups are often more sensitive in the public debate. Police accountability in such cases often requires more and stronger evidence . On the other hand, in cases involving vulnerable groups, even minor misconduct by police could lead to serious public trouble after intense social discussion. At the same time, this reminds us that current law enforcement transparency is not fair but is focused on certain social issues.

5.3 Complex Race–Armed Patterns and Overlapping Missing Information

Results on the race of victim is more complex. Overall, this study did not find any single race consistently less likely to be recorded across all scenarios. Fryer's analysis of police force usage similarly found that while Blacks and Hispanics faced significantly higher rates of non-lethal force compared to Whites. However, after controlling for situational and behavioral variables, his data did not reveal similarly clear racial disparities in the fatal shooting. (2017) Although this paper and its data are not the most recent, and its findings pertain to whether victims died. However, from the perspective of police decision making in fatal cases, it aligns with the results of this study. Interestingly, cases where race was recorded as "Unknown" were at a more significant disadvantage in terms of body camera activation. This combination is highly alarming, as information missing in real world scenarios is often overlapping. Specifically, in fatal police shooting, some victims lack not only complete names and race documentation but also video evidence of the enforcement process. This reflects a serious social unfairness: certain individuals' deaths in such incidents are more easily forgotten. This may not always be due to data collection issues but also indicates that for some groups in society, even the right to be recorded cannot be guaranteed.

In contrast, this study found no consistent or significant systemic differences in victims' gender, whether they fled, or whether they were armed during police-involved fatalities. This does not mean these factors are not important in the incidents, but rather indicates that the small differences in who is more likely to be recorded in cases with lethal outcomes cannot be explained by a single behavioral label.

In fatal police incidents, whether the victim was armed does not systematically differ based on race in terms of its impact on body camera activation. We have no strong statistical evidence to support the conclusion that a particular race is significantly less likely to be recorded when carrying a weapon. However, racial stereotypes present in the United States may lead police to respond differently to individuals of different races in the same law enforcement situation during rapid-response cases, such as fatal police shootings. Psychological research on racial stereotypes and weapon perception has found that when Black individuals hold daily tools, participants may initially feel these objects more like weapons in the first time. (2025) If similar stereotypes play a role in extreme case, police may have systematic biases in their subjective judgments about "whether a suspect is armed" and "the level of threat" when faced with different races, even when the situation is similar. This could influence decisions of police, including whether to activate body cameras. However, the results of this study were not significant, likely reflecting limitations in data. It is possible that in fatal cases, current information is severely insufficient, such as internal agency policies and the emotional states of different police in enforcement situation, which may affect the activation of body cameras.

5.4 Weaknesses and Limitations

Although this study provides valuable insights for law enforcement agencies in subsequent policy development, it is essential to recognize its weaknesses and limitations.

First, the data only cover fatal police shooting and exclude general police-citizen interactions, which means there is selection bias within the sample. This may cause the model to underestimate or overestimate the true correlation coefficient between the two variables. Public policy research typically is more interested in the generalizability of conclusions. Due to the selection bias of the sample, our conclusions are only valid for the fatal police shooting in the United States. These findings apply specifically to fatal cases and cannot be generalized to general police-citizen interactions.

Second, missing information and sample unevenness further limit interpretability. Given the limited sample size and potential associations between body camera usage and victim characteristics, arbitrarily deleting data potentially relevant to the research question could skew correlation estimates. Therefore, I kept a subset of data labeled "Unknown." However, the presence of missing values impacts the final conclusion's description. For example, if certain "Unknown" values are significantly higher or lower than the reference group, the final conclusion becomes less interpretable. Furthermore, groups with smaller sample sizes—such as female victims or victims under 18—may yield unstable and inaccurate final coefficient estimates due to sample size limitations.

Third, the Bayesian Logistic Regression model used in this study is relatively simple and may not strictly satisfy the assumption of independence in real world situations which affects the model's predictive ability. For example, in fatal cases, whether officers activate body cameras may be influenced by differing policies across agencies or correlations among officers, making activation rates non-independent. When samples violate the independence assumption, model fitting may produce biased results, leading to unstable and inaccurate conclusions. Currently, the only model I understand to handle correlated samples is the Linear Mixed Model. However, since body camera activation is a binary variable, it cannot be applied. Through research and exploration, I discovered that the Generalized Linear Mixed Model can handle data with sample correlations and binary dependent variables.

Finally, inefficient observation variables may cause the model's predictive performance to not meet expectations. For example, while I could analyze changes in police fatality rates using U.S. Census data, obtaining accurate racial population data is extremely challenging due to data limitations. With respect to predictive performance, the AUC of this study's model on the test set is approximately 0.66. This indicates that even after controlling for variables such as race, age, and mental health illness, our ability to distinguish whether a body camera was activated remains limited. However, variable selection is constrained by the data itself. The activation of body cameras may be influenced by unobserved institutional and situational factors, and using unnecessary variables could lead to model underrfitting. This suggests that only use fatal police shooting data is insufficient for causal inference regarding which groups are more likely to be recorded during fatal cases. Future research should focus on addressing these weaknesses and limitations.

5.5 Future Directions

This analysis only uses police-involved fatality data and cannot compare body camera activation probability across common contacts, non-fatal force cases. Future research should add more levels of data, such as routine traffic pulls, supermarket theft cases and so on. This would not only better help analyze systemic differences in recording rates across different groups during police-public interactions but also better explore consistency in law enforcement transparency throughout the entire system.

Data analysis primarily involves three steps: data collection and cleaning, model and variable selection, and final conclusions. Data missingness is unavoidable, but I will first evaluate its impact on conclusions. If missing data significantly affects model results, I will supplement it using other administrative data or news reports. Model application requires satisfying model assumptions. In practice, sample correlations are common. Therefore, I will employ a Generalized Linear Mixed Model to estimate random intercepts and random slopes, which will help analyze whether systemic differences exist in police transparency. After selecting an appropriate model, variable choice significantly determines analysis quality. Given the limited variables in this dataset, the Bayesian logistic regression model primarily shows associative patterns rather than causal effects on body camera activation across different groups or scenarios in fatal police shooting. Future research will include variables such as departmental body camera policies, disciplinary mechanisms, and body camera usage in police departments. To address sample imbalance, I will employ weighting methods without increasing noise significantly.

The findings of this study indicate that younger victims and victim with mental health illness have more consistent difference in body camera activation probability. However, the study has not explored the intersection of victim race, gender, age, and mental health status. Future research could be designed to analyze these vulnerable groups more precisely, such as focusing on systemic differences in body camera activation probability among black minors with mental health issues. This would not only enable a more detailed analysis of inequalities in police transparency but also provide a basis for targeted interventions and police training.

A Appendix

A.1 Markov chain Monte Carlo (MCMC) Mixing and Convergence Diagnostics

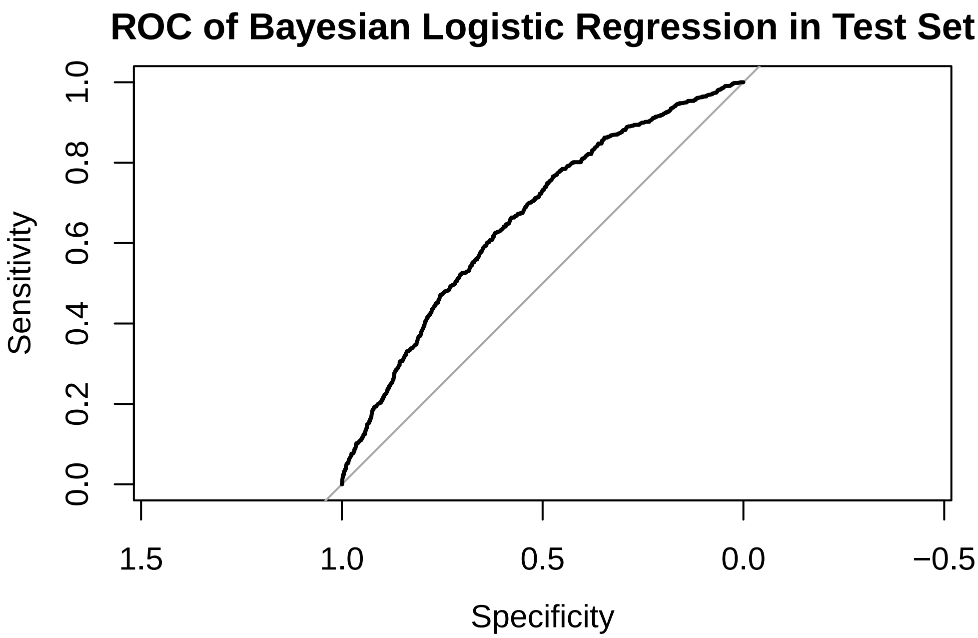


Figure 6: ROC Plot

Figure 6 shows that the ROC curve lies entirely above the diagonal line, with an AUC of approximately 0.66 on the test set. This indicates that after controlling for covariates, the model's predictions regarding "whether body-worn cameras were activated" were better than random guessing. However, it only captures partial systematic patterns, and its predictive capability remains limited.

Figure 7, Figure 8 and Figure 9 show that perform MCMC sampling on the model using 4 chains, each with 2000 iterations (including 1000 warm-up iterations). The trace plots for each parameter show stable, non-trending fluctuations, with good mixing across chains.

Figure 10 shows that The R-values for all parameters are very close to 1.

Therefore, Figure 7, Figure 8, Figure 9 and Figure 10 indicating that the MCMC chains have mixed well and converged to the same posterior distribution.

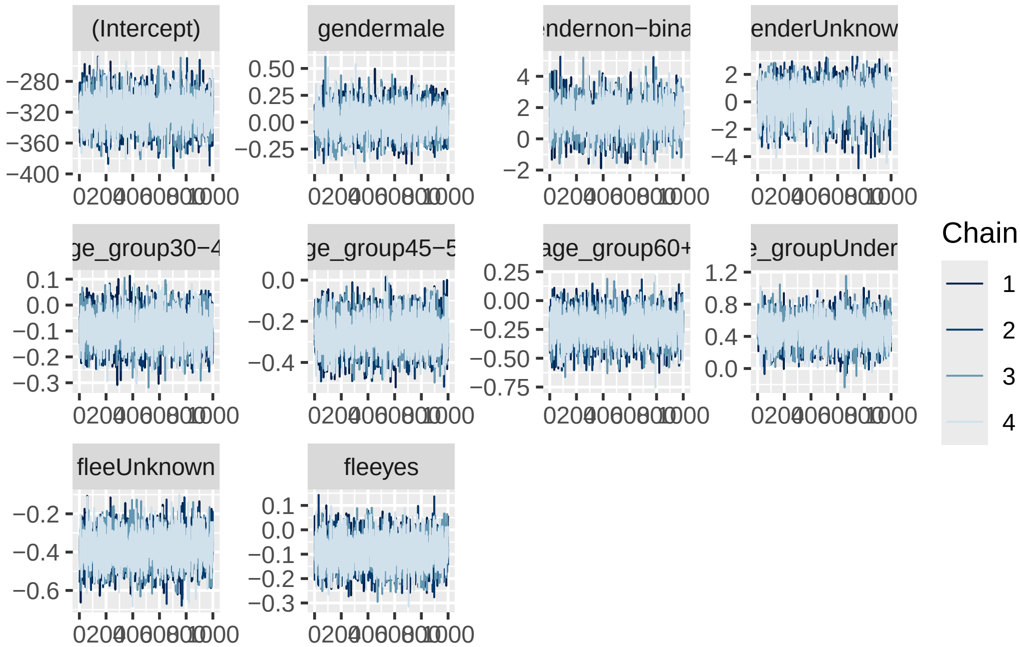


Figure 7: MCMC Trace Plot for Predictor Variable

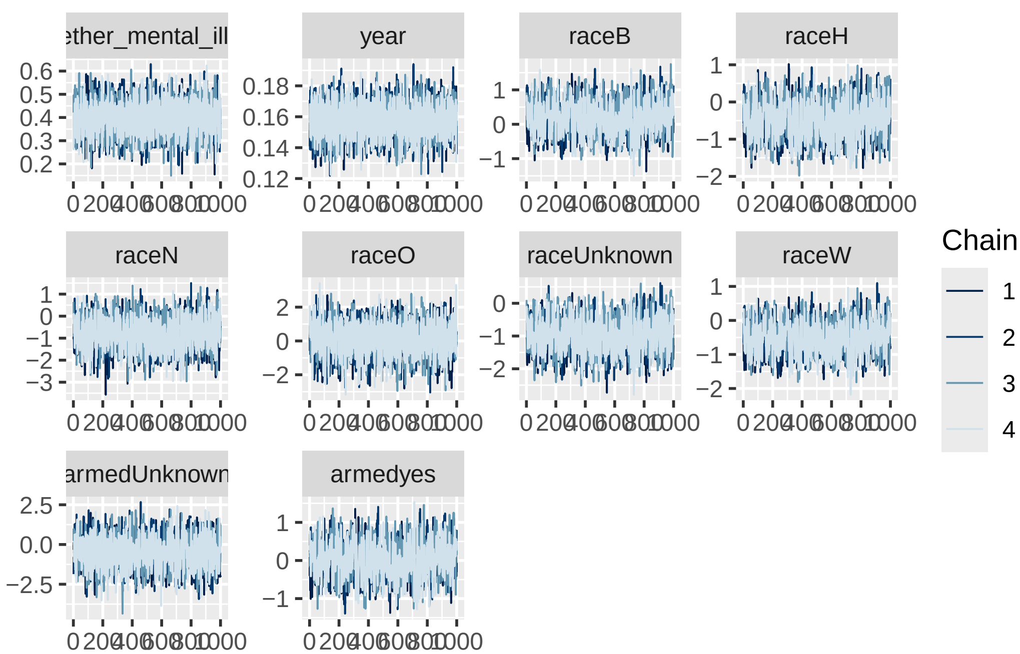


Figure 8: MCMC Trace Plot for Predictor Variable

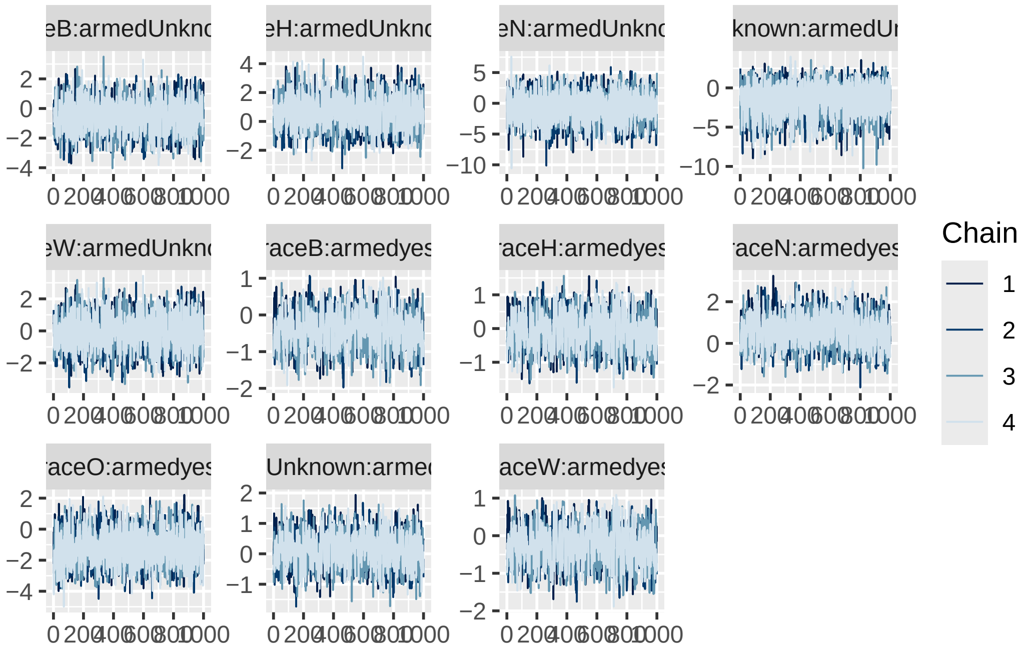


Figure 9: MCMC Trace Plot for Predictor Variable

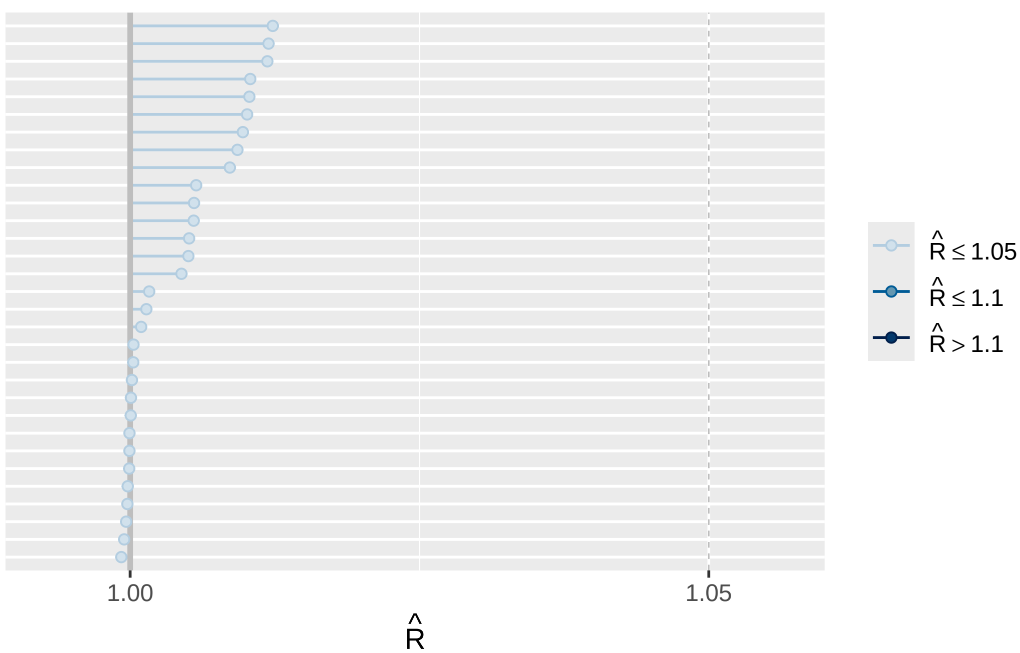


Figure 10: Rhat Plot

B References

Body-Worn Cameras-Toronto Police Service.(2020). https://www.tps.ca/use-technology/body-worn-cameras/

Sponsored content: Body-Worn cameras build transparency and trust for law enforcement across the nation. (2024). POLITICO. https://www.politico.com/sponsored/2024/06/body-worn-cameras-build-transparency-and-trust-for-law-enforcement-across-the-nation/

washingtonpost/data-police-shootings: The Washington Post is compiling a database of every fatal shooting in the United States by a police officer in the line of duty since 2015. (2025). https://github.com/washingtonpost/data-police-shootings

US Census Bureau. (2021, October 8). State population totals: 2010-2020. Census.gov. https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-state-total.html

US Census Bureau. (2025, May 28). State Population Totals and Components of Change: 2020-2024. Census.gov. https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html

Use-of-Force. (2024, June 10). Federal Bureau of Investigation. https://www.fbi.gov/how-we-can-help-you/more-fbi-services-and-information/ucr/use-of-force

Laniyonu, A., & Goff, P. A. (2021). Measuring disparities in police use of force and injury among persons with serious mental illness. BMC Psychiatry, 21(1), 500. https://doi.org/10.1186/s12888-021-03510-w

Canada, S. (2025, October 30). Guide to the Census of Population, 2021, Chapter 9 – Data quality evaluation. https://www12.statcan.gc.ca/census-recensement/2021/ref/98-304/2021001/chap9-eng.cfm

Research on Body-Worn Cameras and Law Enforcement | National Institute of Justice. (2022). National Institute of Justice. https://nij.ojp.gov/topics/articles/research-body-worn-cameras-and-law-enforcement

Fryer, R. G., Jr. (2017). An Empirical analysis of racial differences in police use of force. https://fryer.scholars.harvard.edu/sites/g/files/omnuum5986/files/fryer/files/empirical_analysis_tables_figures.pdf

Racial stereotypes can make us see weapons where they don't exist. (2025). Columbia News. https://news.columbia.edu/news/racial-stereotypes-can-make-us-see-weapons-where-they-dont-exist

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.