

# Investigating the Effects of Demographic, Geographic and Crime Factors on Time-to-Exoneration in the United States

Kanika Chopra

Wrongful convictions can result in catastrophic effects on defendants, separating them from their homes and loved ones and subjecting them to the harsh environment of prison. There are many factors that lead to wrongful convictions, many of which are influenced by prejudicial bias. Previous literature shows that racial and gender bias play a role in wrongful convictions and time-to-exoneration. This paper aims to provide a holistic approach that considers the effects of demographic, geographic and crime factors on time-to-exoneration in the United States. Time-to-exoneration is modelled with a log transformation following a Normal distribution using a Bayesian hierarchical model. The findings of this report illustrate a significant relationship between gender, age, location and time-to-exoneration respectively. Furthermore, it shows that an interaction term between race/ethnicity and crime is important to consider as this impacts time-to-exoneration. Ultimately, this model can be used to better understand demographic bias in the criminal justice system and highlight geographic areas that require more intervention and correction.

## Introduction

Wrongful convictions can take away valuable years from an individual's life, isolate them to the harsh environment of prison and separate them from their loved ones. Fighting to prove their innocence can take years and is both mentally and financially taxing. Not to mention, after proven innocent, there is a lack of support to reintegrate into society

resulting in severe adjustment difficulties [1], [2]. This can result in catastrophic impacts on their life, including feelings of loss of identity, psychological and physical impacts on their health, and difficulties with finances and finding employment [2]. Research has shown that up to 15.4% of convictions are wrongful revealing consequential errors in the justice-system [3].

There are a number of factors that contribute to wrongful convictions including eyewitness misidentification, false confessions, police and prosecutorial misconduct, flawed forensic evidence and perjured testimony [4]. Eyewitness misidentification is the single greatest cause making up 28% of wrongful convictions [4]. After experiencing a stressful event, witnesses are less reliable and influenced by their own personal bias, including prejudicial biases, especially when identifying someone of another race from a lineup or photos [4], [5]. Thus, it is important to be cognizant of these biases and take steps to counteract these errors. Furthermore, identifying bias resulting in wrongful convictions can hold police officers, prosecutors, defense attorneys and judges accountable to avoid future misconduct [6]. This can be a helpful tool for reforming the criminal justice system to reduce these mistakes in the future.

Exonerations are a method of correcting mistakes in the criminal justice system in which “a person who has been convicted of a crime is officially cleared after new evidence of innocence becomes available” [7]. Exonerations can reveal bias in law enforcement through analyzing which groups are more susceptible or frequently wrongfully convicted. Previous literature has investigated the impact of race and ethnicity on time-to-exoneration through a lens of focal concerns [8]. Braden [8] has shown that Black exonerees experienced a longer-time to exoneration than White exonerees while Hispanic exonerees experienced the shortest time-to-exoneration. Furthermore, Itskovich et al. [9] have shown that Black defendants in the United States are at greater risk of being wrongfully convicted for crimes such as murder and rape in comparison to White defendants. Ruesink and Marvin Jr. [10] have revealed gender differences in wrongful convictions, where women were more likely than men to be wrongfully convicted for child abuse or drug violations. They also revealed that African-American women were more likely than White women to be wrongfully convicted for murder [10]. These findings illustrate grave racial and gender biases exhibited in wrongful convictions that must be further investigated.

This motivates a more holistic study investigating the effects of demographic, geographic and crime factors on time-to-exoneration in the United States. In particular, this analysis will use a Bayesian approach to consider factors including race/ethnicity, age, gender, state and the type of crime to determine which significantly affect time-to-exoneration.

## Data

The dataset I am using is from the National Registry of Exonerations from the United States [6]. It is a project of the Newkirk Center for Science & Society at University of California Irvine, the University of Michigan Law School and Michigan State University College of Law. This registry was founded in 2012 and provides comprehensive information on all 3,292 exonerations in the United States since 1989. The purpose of this registry is to analyse trends from past wrongful convictions in order to reduce these errors in the future. This includes thorough information on their First and Last Name, Age, Race/Ethnicity, State, County of Crime, Tags, Official Misconduct (OM) Tags, Crime, Sentence, Convicted Year, Exonerated Year and various columns for the reason they were exonerated including DNA, Mistaken Witness Identification, etc. This dataset is constantly updated every 15 cases; the version used for this project was accessed on Tuesday, March 28, 2023 at 12:03 PM. This data is publicly available on the National Registry of Exoneration [website](#), but requires signing a data agreement to access the spreadsheet.

For this investigation, the dependent variable is the time-to-exoneration (years). This is measured by taking the difference between the year of exoneration and conviction. The independent variables are Race/Ethnicity, Gender, Age, State, and Crime. This will provide us with insights into the demographic, geographic and crime factors.

## Feature Engineering

Two features that are created within our dataset are Time to Exoneration and Gender. In our model, we are interested in modelling  $\log(\text{Time})$  as our time distribution is right-skewed as shown in Figure 5. However, we have some data where time-to-exoneration is 0 so to remedy this issue, we add 1 to every value as an offset. This will then translate to  $\log(1) = 0$  as our smallest Time to Exoneration. Gender is stored in the tags column as “F”, so this character is extracted to create an indicator variable of whether “F” exists in the Tags. In our model, this will be represented as a numeric value where 1 indicates Male and 0 indicates Female.

## Missing Data

The dataset was missing 27 cases of age and 1 case of exonerated and convicted year. Through analyzing the racial/ethnic, gender and crime distributions for the missing data, it does not appear to be correlated with any demographic or crime reason, indicating that these rows are missing at random. These missing rows only make up 0.82% of our data; thus, they are removed as they will not significantly impact the analysis.

## Data Cleaning

In order to clean the dataset further, the distribution of categories within each variable of interest was considered. From the race/ethnicity category, there were cases of multi-racial/multiethnic exonerees, i.e. “Native American/White”. For simplicity purposes, these cases were removed. Furthermore, any unknown race/ethnicity values were treated as missing and removed as well leaving 3,240 exonerees. States were inputted into the data in two different ways. For example, North Carolina was inputted in some cases as “NC” and in other cases as “F-NC”. For consistency, the “F-” was removed from all states in order to keep the geographic data uniform. The number of unique states was 54 so using a list of state abbreviations, this was filtered to the 50 states in the U.S.

## Exploratory Data Analysis (EDA)

### Race and Gender Distributions

In Figure 1, we can see that our data is skewed towards having more Black exonerees, then White exonerees with much lower proportions for the remaining races and ethnicities. We also see that the majority of our data are male exonerees. These imbalances may impact our model results and our effect estimates.

It is important to note that although this data is imbalanced, without having an estimate of the proportion of Black and White convicts in the United States, it is not possible to make conclusions about the proportions of wrongful convictions in each group.

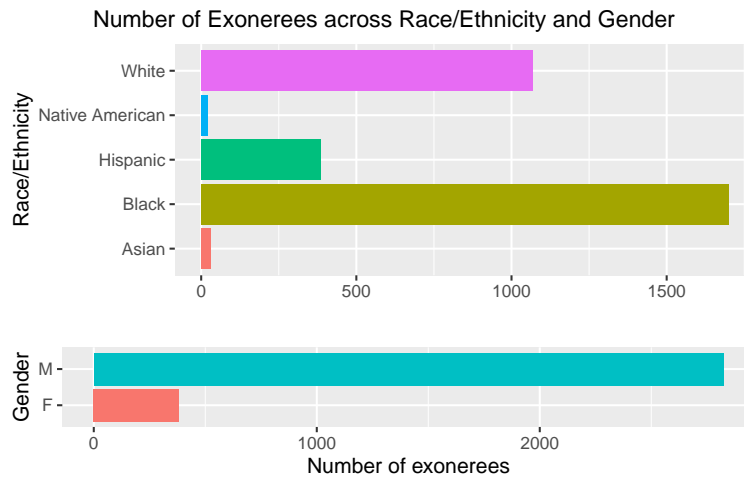


Figure 1: Distribution of exonerees for Race/Ethnicity and Gender. For each category, this is the number of exonerees since 1989.

## Crimes by Race/Ethnicity and Gender

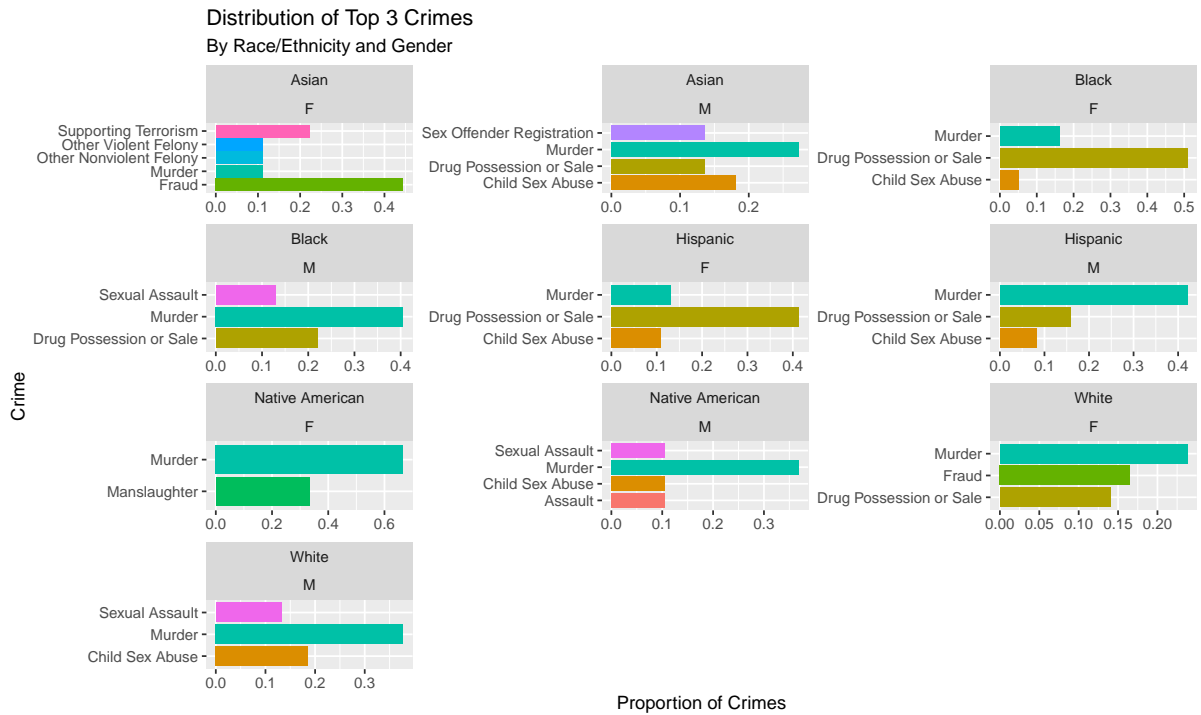


Figure 2: Distribution of Top 3 Crimes per Race, Gender. The proportion of each crime is determined for each race and gender group respectively. The top three crimes are shown per category.

From Figure 2, we observe notable differences in the top 3 crimes between each race/ethnicity and gender group. Firstly, looking at racial differences, we observe that Black and Hispanic exonerees have a higher distribution of drug possession or sale in comparison to White exonerees. We also observe that female exonerees are more likely to be wrongfully convicted for drug possession or sale than male exonerees for Hispanic, White, and Black races/ethnicities; this is in agreement with the previous literature. Other notable differences include supporting terrorism and fraud are in the top 3 crimes for Asian women and fraud for White women whereas the other ethnicities do not have this category. Furthermore, we observe that for Asian, Native American and White men, child sex abuse has a higher proportion than in their female counterparts. We also observe that assault is mainly observed by Native Americans (albeit 2 exonerees). Thus, this seems to be demonstrating a relationship between the wrongfully convicted crimes, race/ethnicity and gender.

## Age Distribution by Race/Ethnicity

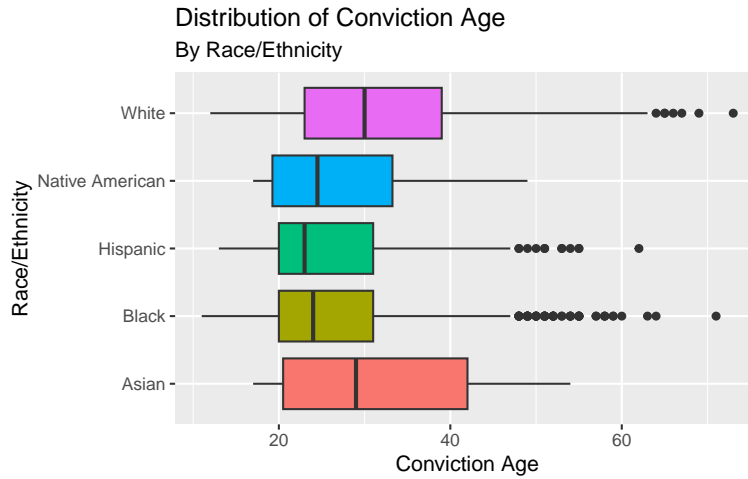


Figure 3: Distribution of Conviction Age by Race/Ethnicity.

Figure 3 plots the distribution of conviction age for each race/ethnicity group. The age range for Asian exonerees is much larger than the remaining groups; this may be attributed to the low number of exonerees within this group. There are more clear differences in data with Black, White and Hispanic with exonerees being as young as 11, 12, 13 respectively. The median age for White and Asian exonerees is closer to 30 in comparison to Black, Hispanic and Native American exonerees being between 23-25.

## Race/Ethnicity by State

Figure 4 shows how the race distribution differs among states. We observe that a higher proportion of Black exonerees come from Illinois, New York, and Texas. The remaining states seem to have lower amounts of data. We can also see some states with more White exonerees than Black exonerees including Florida, Wisconsin and Oregon. As for Hispanic ethnicities, the majority of their data seems to come from Illinois, New York, California and Texas. There could be two reasons for these higher distributions of races from certain states. It could be an indicator of more racially biased law enforcement resulting in a higher distribution of Black exonerees for example from Illinois. The alternative reason could be that more work is being done to exonerate those who were wrongfully convicted based on racial biases and thus there is a higher distribution of exonerees from a specific race/ethnicity. This plot also justifies including state in our model as there are varying distributions of exonerees in each state.

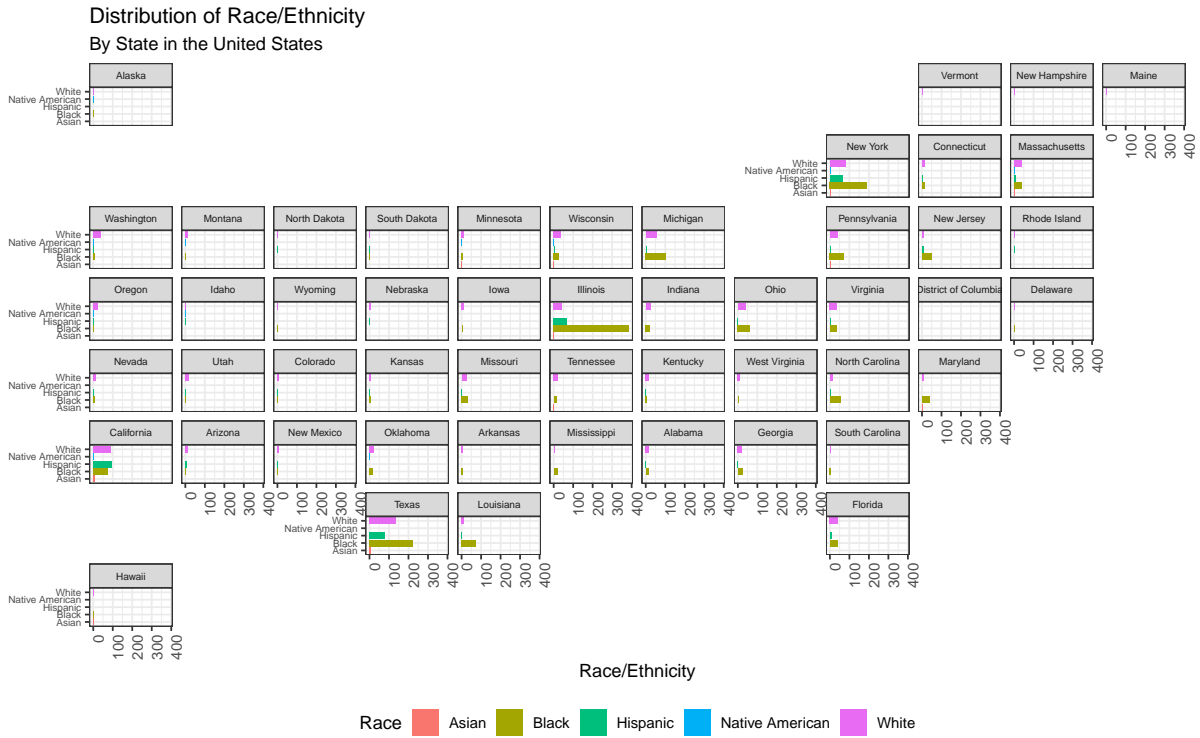


Figure 4

Figure 4: Distribution of Race/Ethnicity across States in the U.S.

## Time-to-Exoneration

Time-to-exoneration on the other hand is a continuous variable so we look at the distribution of time and how it differs based on race, gender and age to determine if any interaction terms are needed. These distributions are displayed on regular time-scale for easier interpretation in Figure 5.

We see more concentration around lower time-to-exoneration values, i.e. around 1-5 years. Our distribution is right-skewed with higher time-to-exoneration values being less common. We do observe that there are some cases where the time-to-exoneration is as high as 60 years which is a very long and detrimental time to spend in prison for a wrongful conviction. As for our comparison with race/ethnicity, we see that the higher values of time seem to be associated with Black exonerees. Furthermore, Native American and Black folk have the highest medians for this time whereas Asian exonerees have the lowest time-to-exoneration. We also observe that Hispanic and White folk have similar distributions for time with very similar median values. With our lower-left plot, we observe that female exonerees typically have lower times-to-exoneration in each race/ethnicity category; thus, an interaction term may be beneficial to consider. Lastly,

we observe higher exoneration times for younger exonerees, which intuitively makes sense since they were convicted younger and spent a longer duration incarcerated. We also observe that females in general seem to experience lower exoneration times than males.

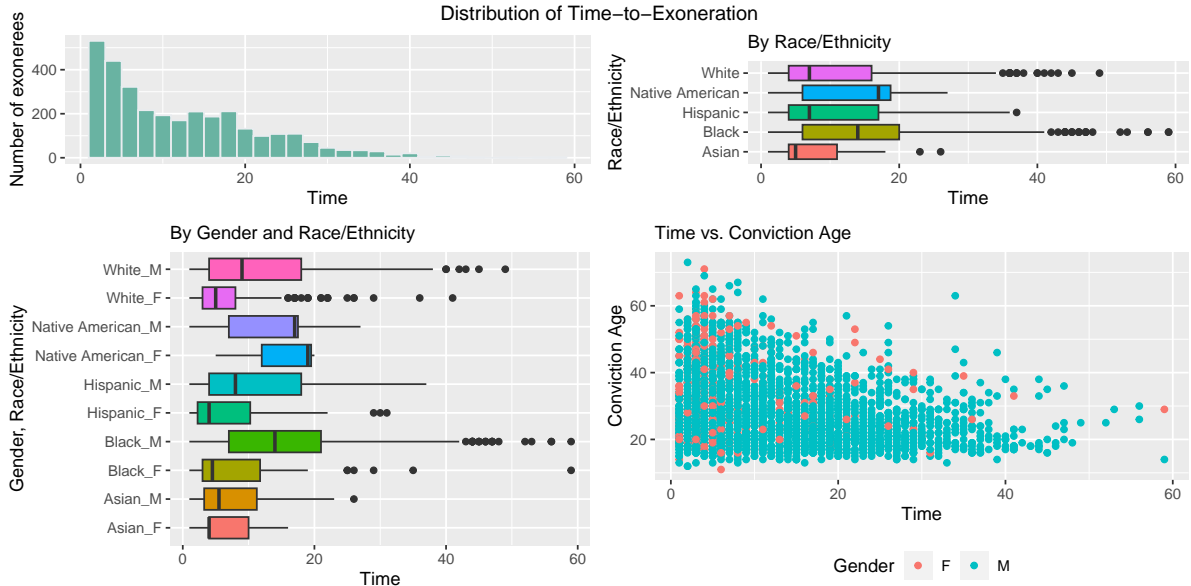


Figure 5: Distribution of Time-to-Exonerees. Top left plot shows the distribution over all exonerees. Top right plot shows the distribution based on Race/Ethnicity. Bottom left plot shows the distribution based on Race/Ethnicity and Gender grouping. Bottom right plot shows the relationship between Conviction Age and Time-to-Exoneration.

## Methods

Since we are working with data with multiple levels, the proposed model is a Bayesian hierarchical model. Specifically, we fit hierarchical variables for race/ethnicity, crime and state. This is because in each of these categories, as shown in our EDA prior, there were low proportions for some groups, i.e. Native Americans are 0.6% of our data and Asians are 0.9% of the data. As for states, we have states with as low as one exoneree and similarly for crime categories. Thus, hierarchical variables are appropriate to capture the information from variables that have limited information for certain levels. The non-hierarchical variables are gender and age. Gender is fit as an indicator variable since there are only two categories and although our data is imbalanced, females still make up 12% of our data so it is not a very low proportion. Age is fit as a continuous variable. As discussed earlier, the time-to-exoneration data is right-skewed so we use



the log transformation on time. This also ensures that our predictive values are always positive as time-to-exoneration will always be positive. Weakly informative priors are set on our regression parameters using a  $N(0, 1)$ . This was chosen to avoid adding any undue influence or information about the parameters to the prior.

The following notation will be used to describe the models: let  $t_i$  represent the time-to-exoneration (years),  $x_{\text{Age}}$  represent the exoneree's conviction age and  $x_{\text{Gender}}$  represent the indicator variable for exoneree's gender. Let  $r[i]$  represent the race/ethnicity group,  $s[i]$  the state and  $c[i]$  the crime group for individual  $i$ . The model will be fit using an R package, `rstan` which allows us the capability to fit a Bayesian model.

Two variations of the model were considered, one with an interaction term. For each model, we validate the results using posterior predictive checks (PPC) to see if our model is able to capture specific test statistics. In particular, we are interested in the distribution of the time-to-exoneration, the 25th percentile, and median. This allows us to compare our observed data to the replicated data from our model to determine which model is the ideal fit. We will also look at these test statistics for race/ethnicity and gender categories as these are variables of interest. Furthermore, for each model, we will calculate the expected log point-wise predictive density (ELPD) and use this metric across the models to choose the preferable model.

## Model 1

The proposed model is given below.

$$\log(t_i) \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta_0 + \beta_1 x_{\text{Age},i} + \beta_2 x_{\text{Gender},i} + \alpha_{r[i]}^{\text{Race/Ethnicity}} + \alpha_{s[i]}^{\text{State}} + \alpha_{c[i]}^{\text{Crime}}$$

$$\beta \sim N(0, 1)$$

$$\alpha_r^{\text{Race/Ethnicity}} \sim N(0, \sigma_{\text{Race/Ethnicity}}^2) \quad \text{for } r=1,2,\dots,5$$

$$\alpha_s^{\text{State}} \sim N(0, \sigma_{\text{State}}^2) \quad \text{for } s=1,2,\dots,50$$

$$\alpha_c^{\text{Crime}} \sim N(0, \sigma_{\text{Crime}}^2) \quad \text{for } c=1,2,\dots,44$$

$$\sigma_{\text{Race/Ethnicity}}^2, \sigma_{\text{State}}^2, \sigma_{\text{Crime}}^2 \sim N^+(0, 1)$$

## Model 2

The second model includes an interaction term between race/ethnicity and crime. Let  $rc[i]$  represent the race/ethnicity-crime group. As shown in our EDA, the top 3 crimes between every race group differed and this could be influencing the time-to-exoneration. While gender also plays a role, it will double the number of hierarchical groups and so we opt to keep the model simpler by including it as the indicator function instead.

$$\log(t_i) \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta_0 + \beta_1 x_{\text{Age},i} + \beta_2 x_{\text{Gender},i} + \alpha_{rc[i]}^{\text{Race/Ethnicity,Crime}} + \alpha_{s[i]}^{\text{State}}$$

$$\beta \sim N(0, 1)$$

$$\alpha_{rc}^{\text{Race/Ethnicity,Crime}} \sim N(0, \sigma_{\text{Race/Ethnicity,Crime}}^2) \quad \text{for } rc = 1, 2, \dots, 120$$

$$\alpha_s^{\text{State}} \sim N(0, \sigma_{\text{State}}^2) \quad \text{for } s = 1, 2, \dots, 50$$

$$\sigma_{\text{Race/Ethnicity,Crime}}^2, \sigma_{\text{State}}^2 \sim N^+(0, 1)$$

## Results

To ensure our models have converged we look at the  $\hat{R}$  and  $n_{\text{eff}}$  for each parameter.  $\hat{R}$  compares the between-and within-chain estimates for each parameter; thus, since it is close to 1 for all parameters in each model, this ensures that the chains have mixed well. Furthermore, our measure of effective sample size,  $n_{\text{eff}}$ , is also quite large for each parameter. Lastly, the trace plots for the first 10 parameters in each model appeared to have mixed well.

## Posterior Predictive Checks

Next, to assess the model fit, we compare their ability to capture key test statistics. We will compare the two models using these test statistics grouping by gender and also by race/ethnicity. Since we are modelling time-to-exoneration, we use test statistics that represent the distribution of the data, i.e. median and quantiles.

## Distribution of Time-to-Exoneration

We use 100 samples from our predictive distribution from each model to assess how it fits the distribution of our observed data.

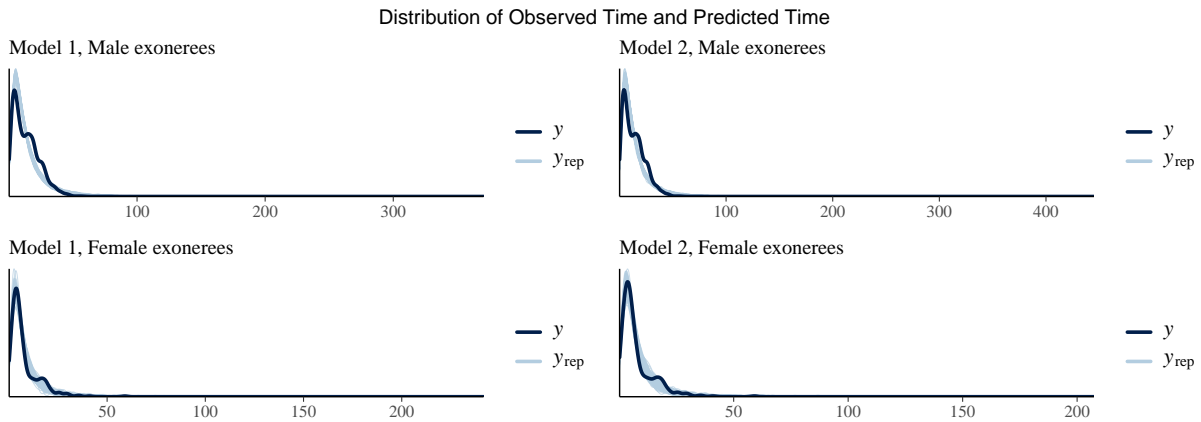


Figure 6: Distribution of Observed Time and Predicted Time based on Gender for Each Model

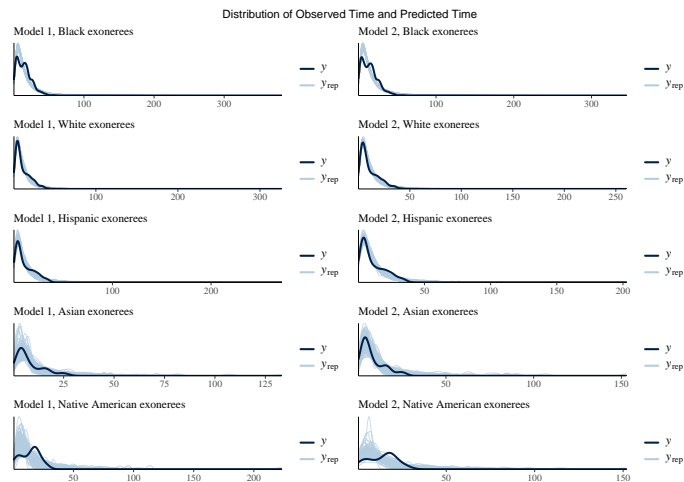


Figure 7: Distribution of Observed Time and Predicted Time based on Race/Ethnicity for Each Model

Figure 6 shows that for both male and female exonerees, both models are able to capture lower time-to-exonerations but our models are not able to capture the higher values. From Figure 7, we observe that both models are able to capture the distribution for White exonerees well. Model 2 captures the distribution for Hispanic exonerees slightly better than Model 1. The distributions for Black, Asian and Native American exonerees

look similar across both models; in both cases, it is not a strong fit. This is likely attributed to the differences in the distribution of data across races/ethnicities.

### Median Time-to-Exoneration

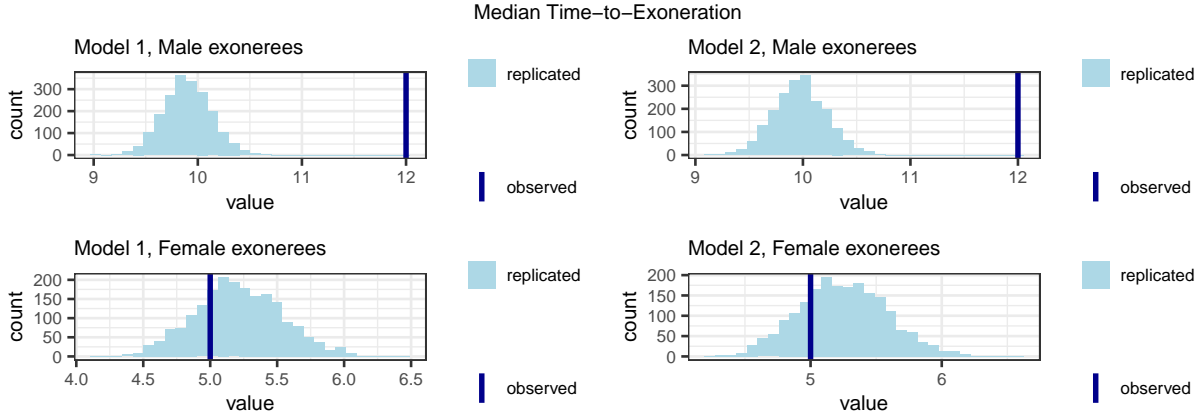


Figure 8: Median Observed Time with Distribution of Predicted Median Time for Each Model. Median time-to-exoneration is calculated across all samples and the histogram is plotted with the observed median time for both Male and Female exonerees

From Figure 8, we observe that we are able to capture the female exonerees median distribution much better than the male exonerees. This seems to be following a trend where we are able to capture the lower time-to-exoneration data in comparison to the higher time-to-exoneration. This is an interesting result as we have much more data on male exonerees than female exonerees but both models are not able to capture this information. Furthermore, Figure 9 shows the differences in the median distributions for both models across race/ethnicity. The models seem to be performing similarly, with slight differences for Native American exonerees, Hispanic and Asian exonerees. Model 2 performs slightly better for Asian exonerees and for Hispanic and Native Americans, we have the opposite. For both models, for Black and Native American exonerees, our model has a lower median time-to-exoneration distribution, whereas it has a higher distribution for the remaining race/ethnicities.

From Figure 10, we again are able to see a similar trend where we are able to capture the 25th percentile for females better than males across both models. Model 2 appears to be slightly better than Model 1. From Figure 11, we can see better results than our median plot. This agrees with previous analyses that our model captures lower time-to-exonerations better. Model 2 appears to be better for capturing the distribution of Black, Hispanic and Asian exonerees. On the other hand, Model 1 appears to be slightly better for capturing the distribution for White and Native American exonerees.

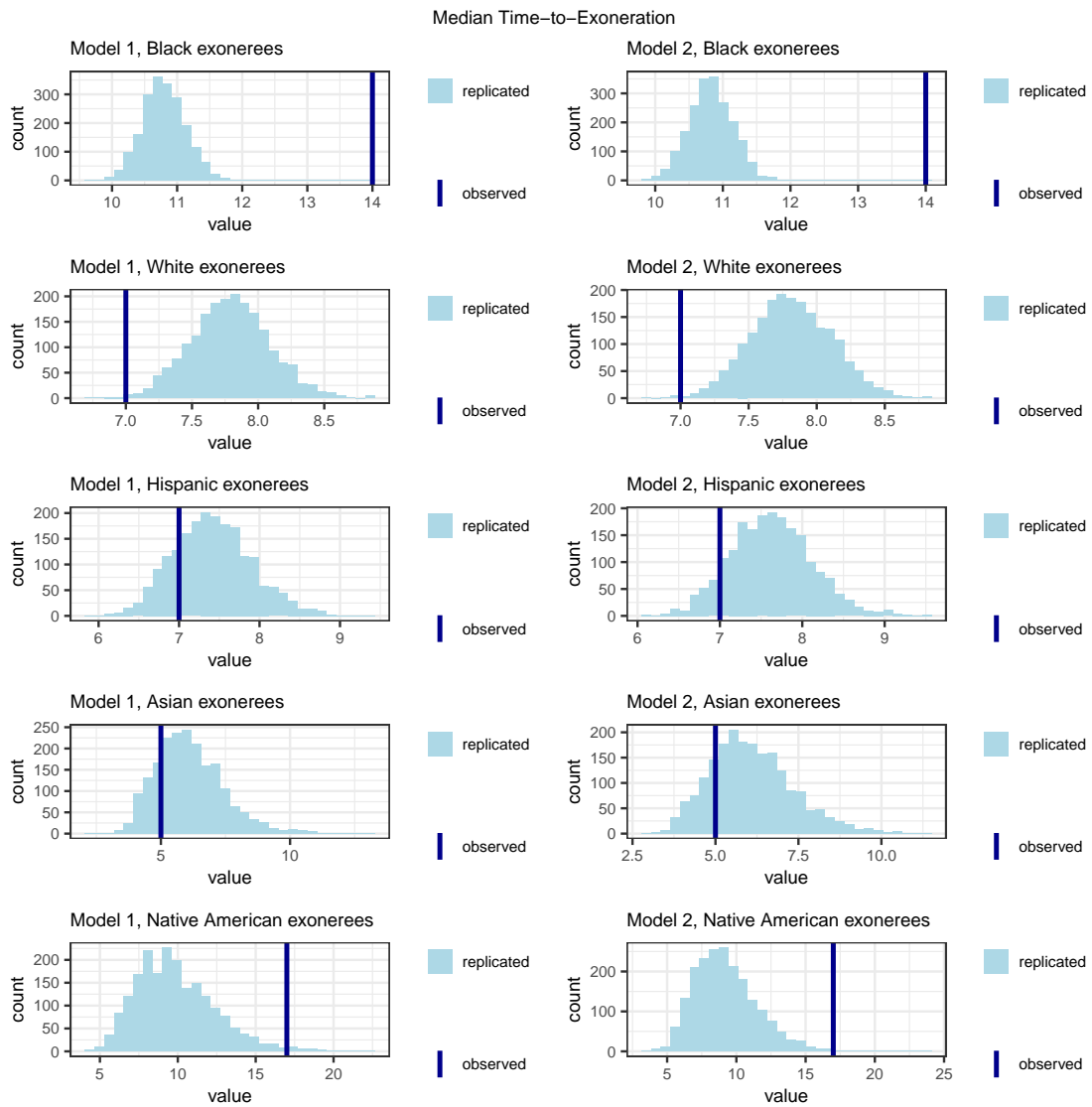


Figure 9: Median Observed Time with Distribution of Predicted Median Time for Each Model. This plot calculates the median time-to-exoneration across all samples and plots the histogram to compare with the actual observed median time for each Race/Ethnicity

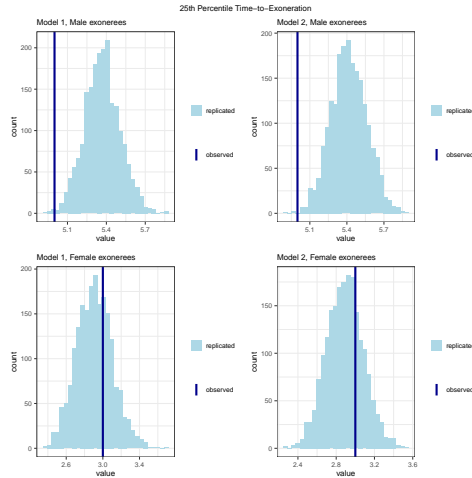


Figure 10: 25th Percentile of Observed Time with Distribution of Predicted 25th Percentile Time for Each Model. This plot calculates the 25th percentile time-to-exoneration across all samples and plots the histogram to compare with the actual observed test statistic for both Male and Female exonerees

These PPCs have shown us how our model performs but determining a better model is difficult as there are not very large differences in the plots for both models. Thus, we will use ELPD to choose our preferable model. Model 2 has a larger ELPD value with a SE difference of 8.8; thus, we can conclude that we prefer Model 2. The remaining interpretations on the estimated effects will be done solely for this model.

### Coefficient Estimates

Using the preferred model, we investigate the coefficient estimates for each of our factors. We will exponentiate our effects so it is in the regular scale for time rather than log of time; thus, these will be multiplicative effects on the time-to-exoneration. We will also be plotting the 95% credible interval; if the credible interval covers 1 then we do not have that the effect is significant at a 5% level.

### Age and Gender Effects

For both age and gender effects, Figure 12 shows that the coverage does not include 1 so these are significant effects at a 5% significance level. Our age effect is lower than 1; thus, we have that an increase in age by one year results in a multiplicative effect reducing time-to-exoneration by 0.989 times. This is not a large impact and is very close to 1, but nonetheless a decrease in time-to-exoneration is exhibited when age increases.

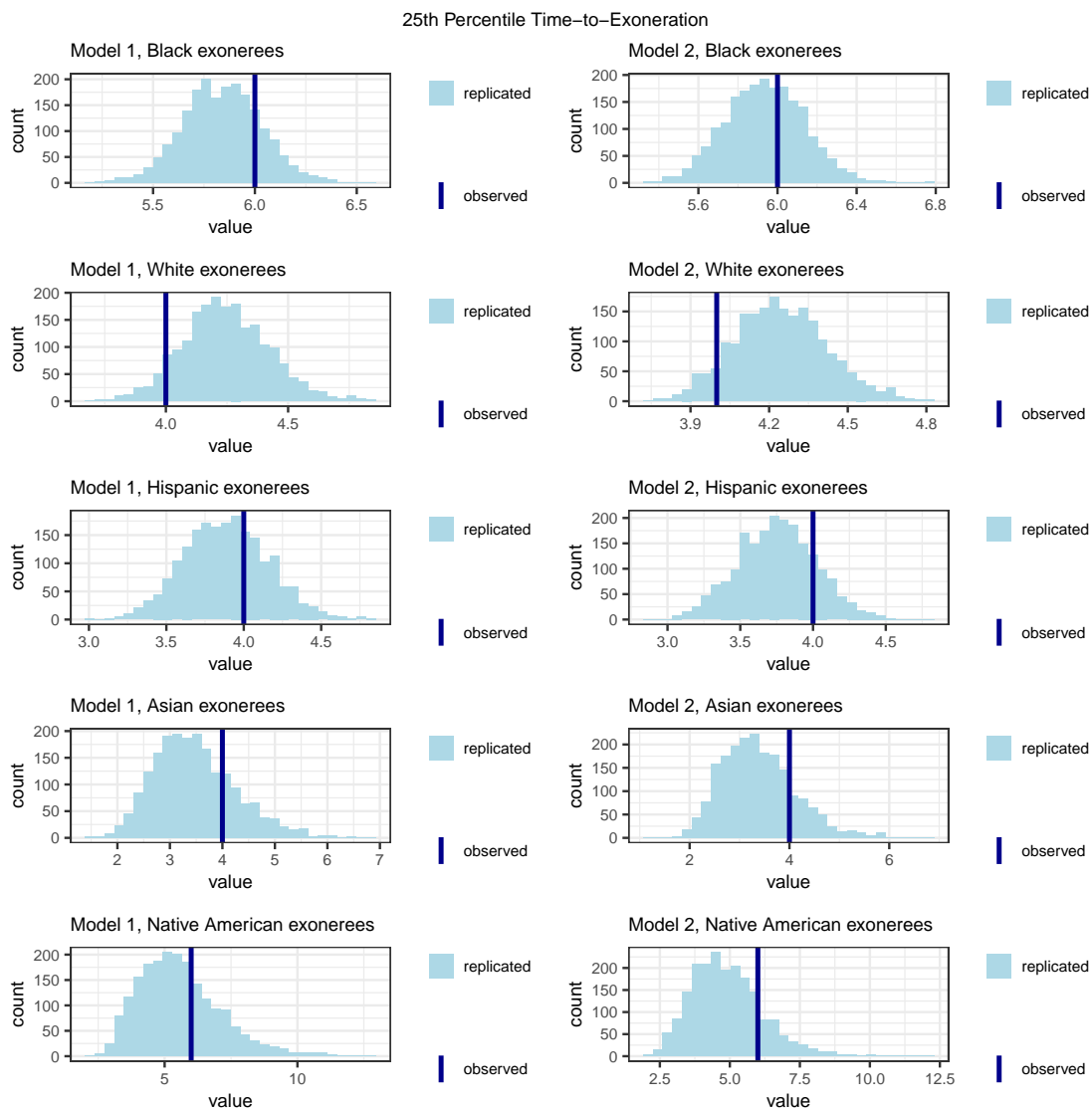


Figure 11: 25th Percentile of Observed Time with Distribution of Predicted 25th Percentile Time for Each Model. The 25th percentile time-to-exoneration is calculated across all samples and plots the histogram to compare with the actual observed test statistic for each Race/Ethnicity

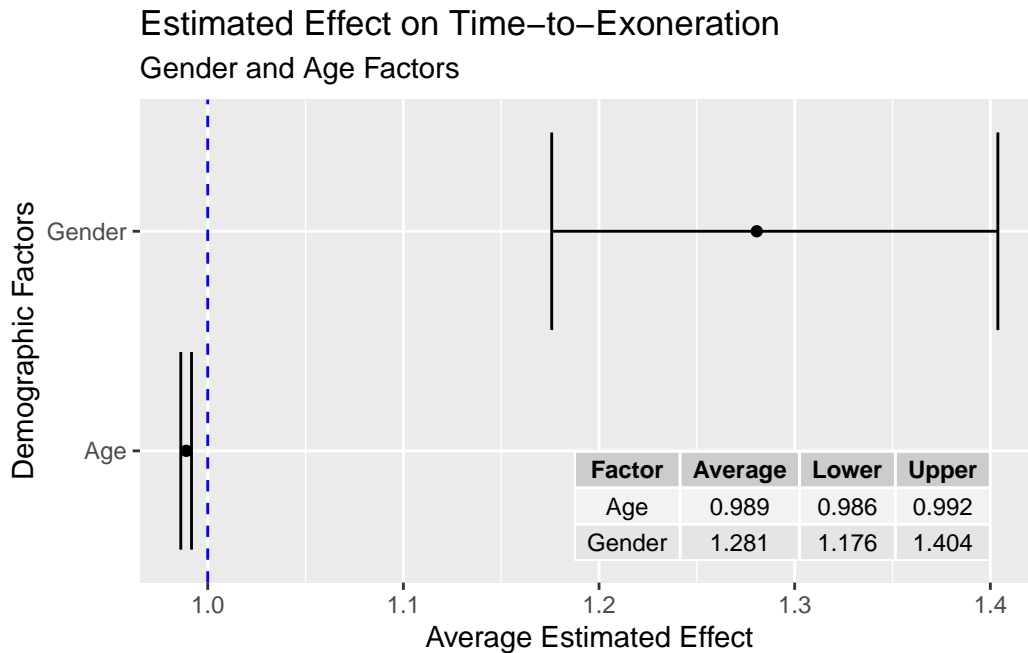


Figure 12: Estimated Effect on Time-to-Exoneration based on Age and Gender. Error bars represent the 95% credible intervals for the average estimate.

For gender, we have that the time-to-exoneration for a male exoneree is 1.281 times longer than a female exoneree. This is expected from our EDA and PPCs as we saw that time-to-exoneration for females was shorter than males.

### Race/Ethnicity and Crime Effects

For this model, race/ethnicity and crime effects are grouped together. Since there are 118 different values, we will choose specific crimes for comparison. Based on previous literature and our EDA, we focus on murder, child sex abuse, drug possession or sale and sexual assault.



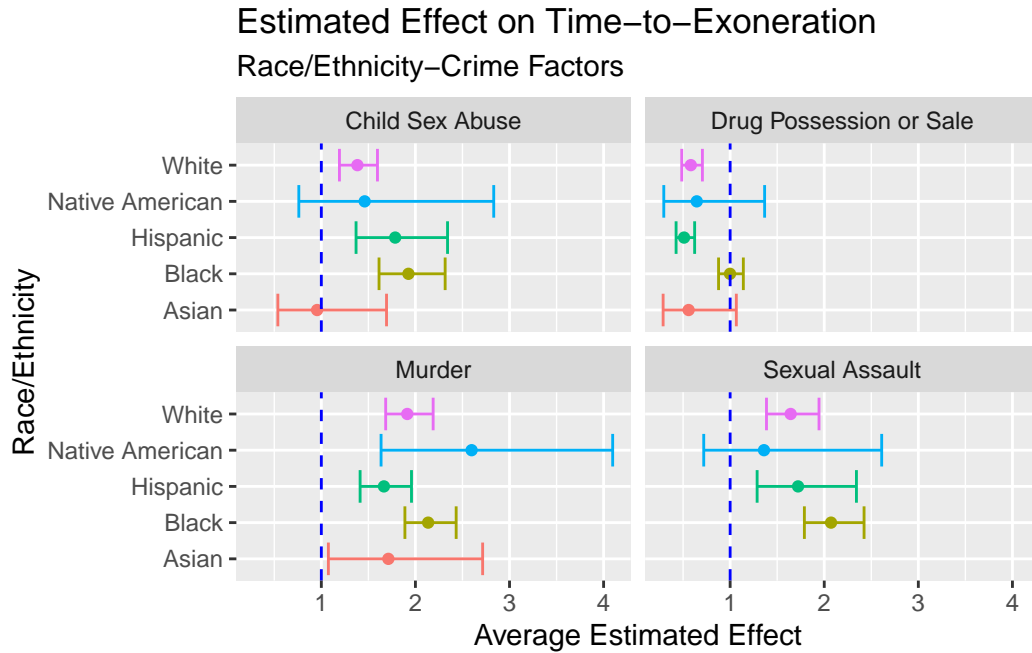


Figure 13: Estimated Effect on Time-to-Exoneration for Child Sex Abuse, Drug Possession or Sale, Murder and Sexual Assault based on Race/Ethnicity. Error bars represent the 95% credible intervals for the average estimate.

From Figure 13, we can see that our significance of the estimates depends on the crime and the race/ethnicity. For child sex abuse, the estimates are significant if White, Hispanic or Black. This may be due to the imbalanced data as there are less Native American and Asian exonerees wrongfully convicted of child sex abuse. Of the three significant estimates, we have that Black exonerees have approximately 1.93 times longer time-to-exonerations than non-Black exonerees and similarly 1.46 times for Hispanic exonerees. This is greater than the effect for a White exoneree showing that for child sex abuse, Black and Hispanic exonerees experience longer times-to-exoneration.

For drug possession or sale, we do not have significant results for Native American, Black or Asian exonerees as our credible interval includes 1. Of our significant results, Hispanic exonerees experience 0.51 times shorter times-to-exoneration than non-Hispanic exonerees and 0.58 times shorter for White exonerees; thus, their effects are less than 1. This shows that Hispanic and White exonerees experience lower times-to-exoneration for this crime.

For murder, all estimated effects are significant. Native Americans have the largest average effect with 2.6 times longer times-to-exoneration. Black exonerees have 2.13 times longer times-to-exoneration. Of the races/ethnicities, Asian and Hispanics experience shorter times-to-exoneration.

Lastly, for sexual assault, we have significant results for White, Black and Hispanic exonerees. Black exonerees experience 2.07 times longer times-to-exoneration than non-Black exonerees. Hispanic exonerees experience 1.72 times longer times-to-exoneration and 1.64 times for White exonerees. Thus, we observe that a Black defendant has a longer time-to-exoneration when wrongfully convicted of sexual assault. We do not have any information on Native Americans wrongfully convicted of sexual assault.

### State Effects

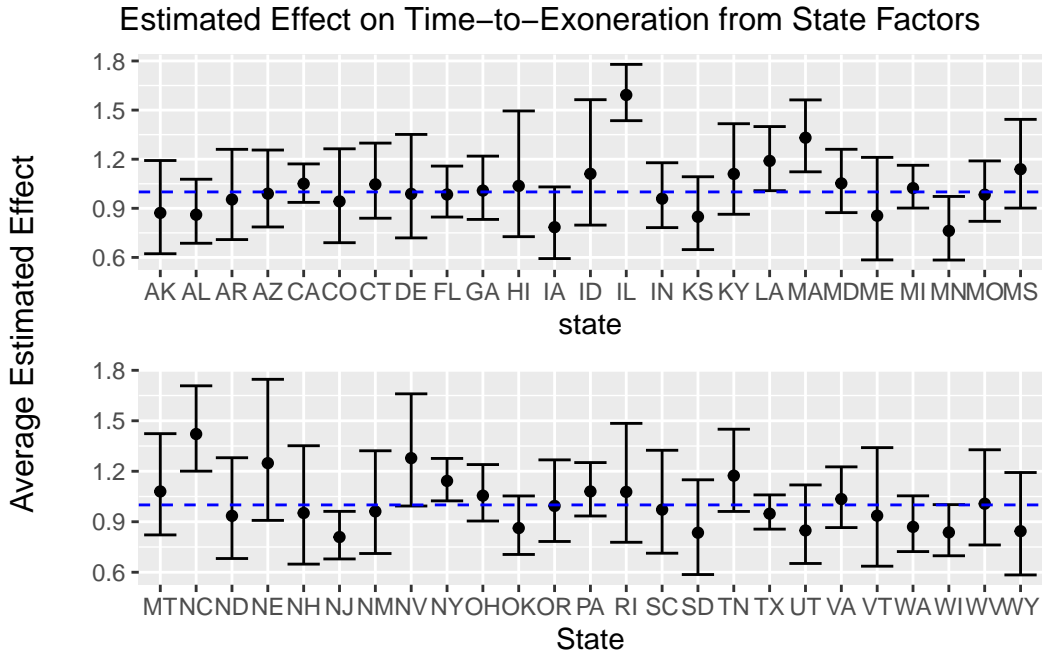


Figure 14: Estimated Effect on Time-to-Exoneration based on State. Error bars represent the 95% credible intervals for the average estimate.

Firstly, we observe from Figure 14 that the majority of states do include 1 in their effects; thus, they are not significant at a 5% level. The states that have a significant effect include Illinois, Massachusetts, Minnesota, North Carolina and New York. There appear to be varying effects for each of the states as well. For example, Illinois has a much higher effect and smaller credible interval than the remaining other states; thus, time-to-exoneration is approximately 1.6 times higher if convicted in Illinois than another state. Similarly, time-to-exoneration is approximately 1.42 times higher if wrongfully convicted in North Carolina than another state and 1.33 times higher in Massachusetts. New York's estimated effect is slightly higher than 1 (1.14) so it has a higher time-to-exoneration but does not appear to be largely substantial. On the other hand, if

wrongfully convicted in Minnesota, the time-to-exoneration is 0.76 times shorter than other states.

## **Additional Models**

Alternative models with an interaction term between race/ethnicity and gender and an interaction term with race/ethnicity, gender and crime were also considered. However, the former did not account for the differences based on crime which seemed to be an important factor. The latter had too many hierarchical terms and had a lower ELPD than our Model 2 so these two models were left out of the analysis.

## **Discussion**

The focus of this analysis was to investigate the effects on time-to-exoneration from demographic, geographic and crime factors. This adds to the previous literature by showing the magnitude and significance of these effects on the time it takes to be exonerated when wrongfully convicted; thus, highlighting any prejudicial bias in the justice system. From the model analysis and validation, we found that an interaction term between race/ethnicity and crime was able to capture the data better; however, both models were not able to capture the data well. Thus, these results are taken with a grain of salt as the models were not a strong fit to the data.

Demographic analysis can be done by examining the effects from age and gender. We observed that as defendants get older, their time-to-exoneration decreases. However, this could be due to the fact that those who have experienced longer times-to-exoneration were convicted younger. More work has been done in the last two decades with Innocence Projects in various states working to free those wrongfully convicted; thus, those who were younger prior to these groups would have waited a longer time before being exonerated. Hence, this could play a role in the age effect. As for gender effects, male exonerees are likely to experience longer times-to-exoneration than female exonerees which agrees with previous literature.

Our effects based on race/ethnicity and crime are grouped together in our interaction hierarchical variable. The analysis was focused on child sex abuse, drug possession or sale, murder and sexual assault. For child sex abuse, we observe that Black exonerees have approximately double the times-to-exoneration than non-Black exonerees and a higher time-to-exoneration for Hispanic exonerees similarly. For drug possession or sale, we have that Hispanic and White exonerees experience a shorter time-to-exoneration. For murder, we have significant effects from all race/ethnicity groups. Native American and Black exonerees have the largest times-to-exoneration with their effect being more

than double than if an exoneree was of another race/ethnicity. Asian and Hispanic exonerees experience shorter times-to-exoneration than these groups. Lastly, for sexual assault, we have that Black exonerees experience around double times-to-exoneration than if they were not Black. Hispanic and White exonerees also experience longer times-to-exoneration but by a smaller scale. This shows that race/ethnicity and the crime wrongfully convicted of make an impact on the time-to-exoneration. These crimes investigated are more serious crimes and thus the process to obtain proof of innocence would be more strenuous than being wrongfully convicted for a less severe crime, i.e. a traffic offense. Thus, the effects on time-to-exoneration may be influenced by this process. In addition to this, these results are similar to previous literature that showed that Black defendants were more likely to be wrongfully convicted of murder than White defendants; hence, more data and longer times-to-exoneration.

Geographic factors were investigated by illustrating the estimated effect for each state on time-to-exoneration. These results are not surprising as Illinois was found to have the highest impact on time-to-exoneration. It has also been cited as being the state that has the most exoneration cases [[11]]. There was a spike in exoneration in 2022 which was largely attributed to the state of Illinois. From the EDA, we saw that Illinois had a skewed racial/ethnic distribution in exoneration where the majority were Black exonerees. This can be attributed to the corruption in law enforcement resulting in many of these exoneration. Specifically, in Cook County, Chicago, there were more than 90 exoneration where a corrupt police officer framed defendants from a public-housing neighborhood between 2003-2011 [12]. Thus, this reveals bias within law enforcement on a lower-income neighborhoods which tend to have more people of color; thus, racial/ethnic bias as well. On the other hand, we observe that Minnesota has a lower time-to-exoneration effect than other states; this may be due to work being done by the Great North Innocence Project to ensure wrongful convictions are overturned quickly. These geographic effects are able to capture some notable information regarding time-to-exoneration and key areas that may need more focus on fixing the bias.

A very evident limitation was the lack of data; this dataset was comprised of roughly 3000 exonerees across all states. This was a large contributor to the poor model fit as shown through the Posterior Predictive Checks. Except for the case with female exonerees, our model failed to capture key test statistics regarding the time-to-exoneration. Additionally, the race/ethnic categories could be more specific, i.e. there are different associations and implications likely associated with those from different regions of Asia rather than grouping them under the same ethnicity. However, these are difficult limitations to overcome as more detailed data may not be available and dataset contained all of the exonerees since 1989; hence, in the United States this is the complete dataset. As the data continues to grow, this model can become more accurate. More data from other countries could be used to investigate an overall effect and capture countries using

hierarchical variables; however, this may limit a more granular investigation geographically.

Furthermore, the model used for this project was a Bayesian Hierarchical model for time-to-exoneration. A more appropriate model would be using various time-to-event models to capture the time from conviction to exoneration; this can include Accelerated Time Failure (AFT) Regression models, or approaches of Bayesian models for time-to-event data. In addition, more suitable priors could be investigated as this could also have an impact on model fit and performance. Currently, weakly informative priors are being used, so more informative priors using information from previous literature on the effects could provide a better fit. Future work could also consider breaking down states to look at counties as well; it would be interesting to see if there is a larger effect captured from Cook County, Illinois. Our EDA revealed that there was a relationship between race/ethnicity, gender and crime so an alternative approach could look at these factors in more depth rather than solely the relationship between race/ethnicity and crime. Lastly, it would be interesting to look at proportions of those exonerated in each race/ethnicity across states in comparison to the number of convicts in prison; this can give insight regarding whether similar proportions of groups are exonerated or if there is a certain group that has a higher proportion of wrongful exonerations.

In conclusion, although the model had room for improvement, this analysis has provided insights into the effects on time-to-exoneration using demographic, geographic and crime factors. We were able to observe that male exonerees experience longer time-to-exonerations and for more serious crimes, Black exonerees experience almost double the amount of time-to-exoneration. Similarly, for certain crimes, Hispanic exonerees experience lower time-to-exonerations. This analysis showed that the time-to-exoneration is influenced significantly by an interaction between race/ethnicity and crime. Furthermore, the geographic analysis revealed that certain States have a higher time-to-exoneration; thus, the need for correction. This study has highlighted the impact that can come from bias, i.e. racial bias, on the time-to-exoneration and geographic factors and revealed areas where more work can be done to reduce time-to-exonerations and correct any wrongful convictions.

## References

- [1] H. Weigand, “Rebuilding a life: The wrongfully convicted and exonerated,” *BU Pub. Int. Lj*, vol. 18, p. 427, 2008.
- [2] S. K. Brooks and N. Greenberg, “Psychological impact of being wrongfully accused of criminal offences: A systematic literature review,” *Medicine, Science and the Law*, vol. 61, no. 1, pp. 44–54, 2021.
- [3] T. G. Poveda, “Estimating wrongful convictions,” *Justice Quarterly*, vol. 18, no. 3, pp. 689–708, 2001.
- [4] “Why do wrongful convictions happen?” University of Colorado Boulder. Available: <https://www.colorado.edu/outreach/korey-wise-innocence-project/our-work/why-do-wrongful-convictions-happen>
- [5] G. L. Wells and E. A. Olson, “Eyewitness testimony,” *Annual review of Psychology*, vol. 54, no. 1, pp. 277–295, 2003.
- [6] “Our mission.” National Registry of Exonerations. Available: <https://www.law.umich.edu/special/exoneration/Pages/mission.aspx>
- [7] “Research resources.” Benjamin N. Cardozo School of Law, Yeshiva University, Apr. 2023. Available: <https://innocenceproject.org/research-resources/>
- [8] V. E. Braden, “In pursuit of innocence: A study of race and ethnicity differences in time to exoneration.” 2021.
- [9] E. Itskovich, R. Factor, and D. Ohana, “Haven’t they suffered enough? Time to exoneration following wrongful conviction of racially marginalized minority-vs. Majority-group members,” *Punishment & Society*, p. 14624745221148318, 2023.
- [10] M. Ruesink and F. D. Marvin Jr, “Wrongful convictions among women: An exploratory study of a neglected topic,” *Women & Criminal Justice*, vol. 16, no. 4, pp. 1–23, 2005.
- [11] D. Selby, “These 8 states had the most exonerations in 2019,” *Innocence Project*. Apr. 2020. Available: <https://innocenceproject.org/these-8-states-had-the-most-exonerations-in-2019/>
- [12] “Race and wrongful convictions.” Available: <https://www.law.umich.edu/special/exoneration/Pages/Race-and-Wrongful-Convictions.aspx>