NAVAL POSTGRADUATE SCHOOL

Qualified Immunity: An Analysis of Effect on Police Shootings

Erin O'Neil | Ace Padilla | Jason Stisser
Andrew Tamez

OA3801 – Winter 2024

Problem Description

<u>Problem Statement:</u> Does 'qualified immunity' affect the number of police shootings?

What is Qualified Immunity?

- Qualified immunity is the precedent, upheld by the US Supreme Court in the landmark case Pierson vs Ray in 1967, which protects a police officer from civil litigation for violation of a person's constitutional rights.

- In 1982, the Supreme Court established the further guidelines to qualified immunity in Harlow v Fitzgerald. In this case, they established immunity is granted unless the officer's conduct violated a clearly established law.

- Qualified immunity normally impacts day-to-day encounters between the police and civil populace. This project will look at whether there is an effect on the number of police shootings.

- Three states have eliminated 'qualified immunity' for state constitutional violations for police officers: Connecticut, Colorado, and New Mexico. We are interested in whether these states have shown
  a) a change in the number of shootings since the rule change and
  b) the difference compared to other states who still have qualified immunity.

NAVAL POSTGRADUATE SCHOOL

Assumptions

- All shooting treated equally regardless of circumstances. No effect given to whether suspect was armed, signs of mental illness in the suspect, or whether the suspect was fleeing.
    - i.e. Justified vs unjustified

- The only impact we considered on the number of shootings was the invoking of qualified immunity in the respective states. No other laws implemented during this timeframe were considered.

Data Curation

- **Data Sources:**
  - https://www.washingtonpost.com/graphics/investigations/police-shootings-database/
    - Nine Excel spreadsheets for each state of interest
    - Excel sheets contain each occurrence of a police shooting resulting in death
  - https://www.census.gov/
    - Two csv files covering the census data during the time frames of interest
    - One containing census data for each state for the years 2010-2019.
    - One containing census data form 2020-2023
- **States Chosen:**
  - CA, CO, CT, FL, IL, MS, NM, NY, TX

Python Code

Import, Parse Data

```
[1]: # Add Matplotlib inline magic command
%matplotlib inline
# Dependencies and Setup
import matplotlib.pyplot as plt
import pandas as pd
import os
```

```
[2]: state_folder = 'Data' #folder that all of the individual day data is in
state = [file for file in os.listdir(state_folder) if file.endswith('.csv')]  #read the names into a list
states = sorted(state)
states[0:5]
```

```
[2]: ['2024_01_31_CA_police_shootings.csv',
 '2024_01_31_CO_police_shootings.csv',
 '2024_01_31_CT_police_shootings.csv',
 '2024_01_31_FL_police_shootings.csv',
 '2024_01_31_IL_police_shootings.csv']
```

```
[3]: dfs = []

#for loop to loop through the list of each file name, read it into a dataframe, and add it to the df list
for file in states:
    file_path = os.path.join(state_folder, file)
    columns_to_read = ['date', 'name','age', 'gender', 'armed', 'race', 'city', 'state', 'flee', 'body_camera', 'signs_of_mental_illness', '
    df = pd.read_csv(file_path, usecols=columns_to_read)
    dfs.append(df)
#dfs #will display the list of all the dfs, commented out to save space
```

```
[4]: shootings_data_df = pd.concat(dfs, ignore_index=True)
shootings_data_df.head()
```

```
[6]: shootings_data_df = shootings_data_df.rename(columns={'date':'Date'})
shootings_data_df = shootings_data_df.rename(columns={'name':'Name'})
shootings_data_df = shootings_data_df.rename(columns={'age':'Age'})
shootings_data_df = shootings_data_df.rename(columns={'gender':'Gender'})
shootings_data_df = shootings_data_df.rename(columns={'armed':'Armed'})
shootings_data_df = shootings_data_df.rename(columns={'race':'Race'})
shootings_data_df = shootings_data_df.rename(columns={'city':'City'})
shootings_data_df = shootings_data_df.rename(columns={'state':'State'})
shootings_data_df = shootings_data_df.rename(columns={'flee':'Flee'})
shootings_data_df = shootings_data_df.rename(columns={'body_camera':'Body Camera'})
shootings_data_df = shootings_data_df.rename(columns={'signs_of_mental_illness':'Signs of Mental Illness'})
shootings_data_df = shootings_data_df.rename(columns={'police_departments_involved':'Police Department'})


shootings_data_df.head()
```

Visualize

```
sum_by_state = shootings_data_df.groupby('State').agg(
                  Sum_Shootings = ('State',"count"))
sum_by_state.plot(kind='bar')
plt.title('Total Deaths by State (2015-2023)')
plt.xlabel('State')
plt.ylabel('Total Deaths')
plt.show()
```

```
[12]: shootings_data_df['Year'] = pd.to_datetime(shootings_data_df['Date']).dt.year

# Step 4: Group data and calculate total per year for each state

filtered_data = shootings_data_df[shootings_data_df['Year'] < 2024]

# Group data and calculate total per year for each state
grouped_data = filtered_data.groupby(['Year', 'State']).size().reset_index(name='total')

# Step 5: Create a line plot
fig, ax = plt.subplots(figsize=(15, 20))

for state, state_data in grouped_data.groupby('State'):
    ax.plot(state_data['Year'], state_data['total'], label=state)

for state, state_data in grouped_data.groupby('State'):
    last_data_point = state_data.iloc[-1]
    sl_data_point = state_data.iloc[-2]
    first_data_point = state_data.iloc[0]
    ax.text(last_data_point['Year'], last_data_point['total'], str(last_data_point['total']),
            ha='left', va='center', color='black', fontsize=10)
    ax.text(sl_data_point['Year'], sl_data_point['total'], str(sl_data_point['total']),
             ha='center', va='bottom', color='black', fontsize=10)
    ax.text(first_data_point['Year'], first_data_point['total'], str(first_data_point['total']),
             ha='right', va='center_baseline', color='black', fontsize=10)

ax.set_xlabel('Year')
ax.set_ylabel('Total Shootings per Year')
ax.set_xticks(grouped_data['Year'].unique())
ax.legend(title = 'State',loc='upper center', ncols=9)
#ax.legend(title='State', loc='upper right')
plt.title('Total Shootings per Year by State (2015-2023)')
plt.show()
```

Adjust, Refine

```
merged_df = pd.merge(census_df, ncensus_df, on='State')
merged_df['State'] = merged_df['State'].str.replace('.', '')
years_to_remove = ['2010', '2011', '2012', '2013', '2014']
merged_df = merged_df.drop(columns=years_to_remove)
selected_states = ['CA', 'CO', 'CT', 'FL', 'IL', 'MS', 'NM', 'NY', 'TX']

# Keep only rows with states in the selected_states list
filtered_df = merged_df[merged_df['State'].isin(selected_states)]

filtered_df
#pivot_filtered_df = filtered_df.pivot(index='State', columns='Year', values='total')
```

```
[32]: df_sum = df_filtered.groupby('State').sum('State').reset_index()
grouped_data
pivot_df_sum = grouped_data.pivot(index='State', columns='Year', values='total')
pivot_df_sum
year1 = ['2015', '2016', '2017', '2018', '2019', '2020', '2021', '2022', '2023']
pivot_df_sum.columns = pivot_df_sum.columns.astype(str)
```

```
file_path1 = 'Data/Census_old.xlsx'

# Read the census data from the xlsx file, skipping the first row and setting column names as years
years = ['2010', '2011', '2012', '2013', '2014', '2015', '2016', '2017', '2018', '2019']
census_df = pd.read_excel(file_path1, header=None, names=years)

# Reset the index to make state abbreviations the new row headers
census_df.reset_index(inplace=True)

# Rename the column with state abbreviations
census_df.rename(columns={'index': 'State'}, inplace=True)
census_df = census_df.drop(0)
# Print the first few rows of the dataframe
census_df.head()
```

Computation - Python

- **Cleaning/ Analysis Tools:**
  - Pandas Dataframes
  - Seaborn
  - Plotly.express/Choropleth
  - COLAB Attempt

- **Computations for Analysis:**
  - Summation of total shootings for each selected state over the given timespan
  - Year-to-year totals for each state over the given time span
  - Compute shootings per captia of 100,00 people for each state over the given years
  - Graph the data using our calculations on various graphs (line, heatmap, possible boxplots)
  - Calculate means across years for all states, those without qualified immunity, and those with qualified immunity to provide additional data points to compare in our results.
  - Establish Ho, Ha, and alpha for statistical hypothesis testing
  - Calculate p-value to determine if our statistical hypothesis can be rejected.

Computation - R

Two-sample hypothesis test

μ_1 : Mean of police shootings per capita in states with qualified immunity

μ_2 : Mean of police shootings per capita in states that removed qualified immunity

H_0: $\mu_1 - \mu_2 = 0$ vs. $H_{a:}$ $\mu_1 - \mu_2 \leq 0$

H_0 : No significant difference between the mean of the two groups

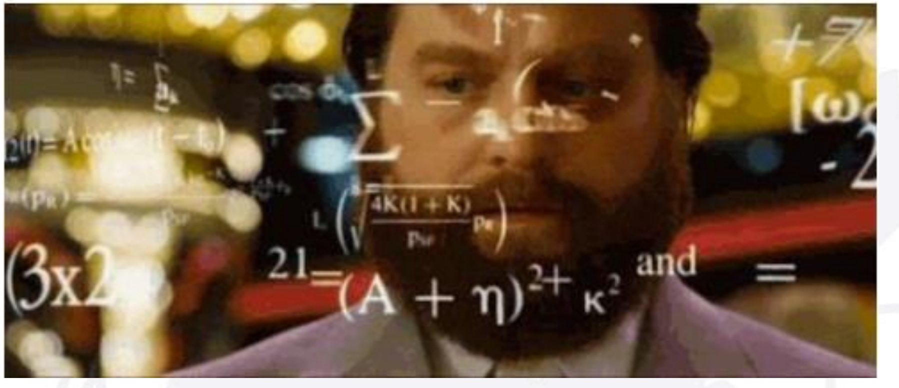
H_a : The mean of police shootings is lower in states that removed qualified immunity.

Test Statistic: $$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

R Code

Data from Python to R

```r
# HO: p1 - p2 = 0
# Ha: p1 - p2 < 0

#read, parse data and prepare for testing

data = read.csv(file.choose())
tdata = as.data.frame(t(data))
colnames(tdata) = c('CA', 'CO', 'CT', 'FL','IL', 'MS', 'NM', 'NY', 'TX')
tdata # 2015-2023 | rows 1-9 = pop , 10-18 = shootings, 19-27 = shootings/100,000 pop.

#---------------------------------------#
#---------------------------------------#
# 2015-2020
# Shooting proportions per 100,000 with States with immunity (CO, CT, NM)
CO_avg_cap_15 = mean(tdata[19:24,2]) ;
CT_avg_cap_15 = mean(tdata[19:24,3]) ;
NM_avg_cap_15 = mean(tdata[19:24,7]) ;
CO_avg_cap_15 ; CT_avg_cap_15 ; NM_avg_cap_15

data_no_15 = c(CO_avg_cap_15,CT_avg_cap_15,NM_avg_cap_15)
mu_no_15 = mean(data_no_15) ; mu_no_15
sd_no_15 = sd(data_no_15); sd_no_15
n_no_15 = length(data_no_15)

#---------------------------------------#
# 2015-2020
# Shooting proportions per 100,000 with States with qualified immunity (CA, FL, IL, MS, NY, TX)
CA_avg_cap_15 = mean(tdata[19:24,1]) ; CA_avg_cap_15
FL_avg_cap_15 = mean(tdata[19:24,4]) ; FL_avg_cap_15
IL_avg_cap_15 = mean(tdata[19:24,5]) ; IL_avg_cap_15
MS_avg_cap_15 = mean(tdata[19:24,6]) ; MS_avg_cap_15
NY_avg_cap_15 = mean(tdata[19:24,8]) ; NY_avg_cap_15
TX_avg_cap_15 = mean(tdata[19:24,9]) ; TX_avg_cap_15

data_im_15 = c(CA_avg_cap_15,FL_avg_cap_15,IL_avg_cap_15,MS_avg_cap_15,NY_avg_cap_15,TX_avg_cap_15)
mu_im_15 = mean(data_im_15) ; mu_im_15
sd_im_15 = sd(data_im_15); sd_im_15
n_im_15 = length(data_im_15)

#Compare 2015-2020
mu_no_15 ; mu_im_15
```



P-Test

```r
# Ho: mu_no_15 - mu_no_21 = 0
# Ha: mu_no_15 - mu_no_21 > 0   | Got rid of immunity so should be less shooting.
z = ((mu_no_15- mu_no_21)-0) / sqrt( ((sd_no_15)^2/n_no_15) + ((sd_no_21)^2/n_no_21));z
pt(z,n_no_15) # 0.8680375 > 0.05 so reject the null
```

P-value: 0.395 (Less than 0.05)

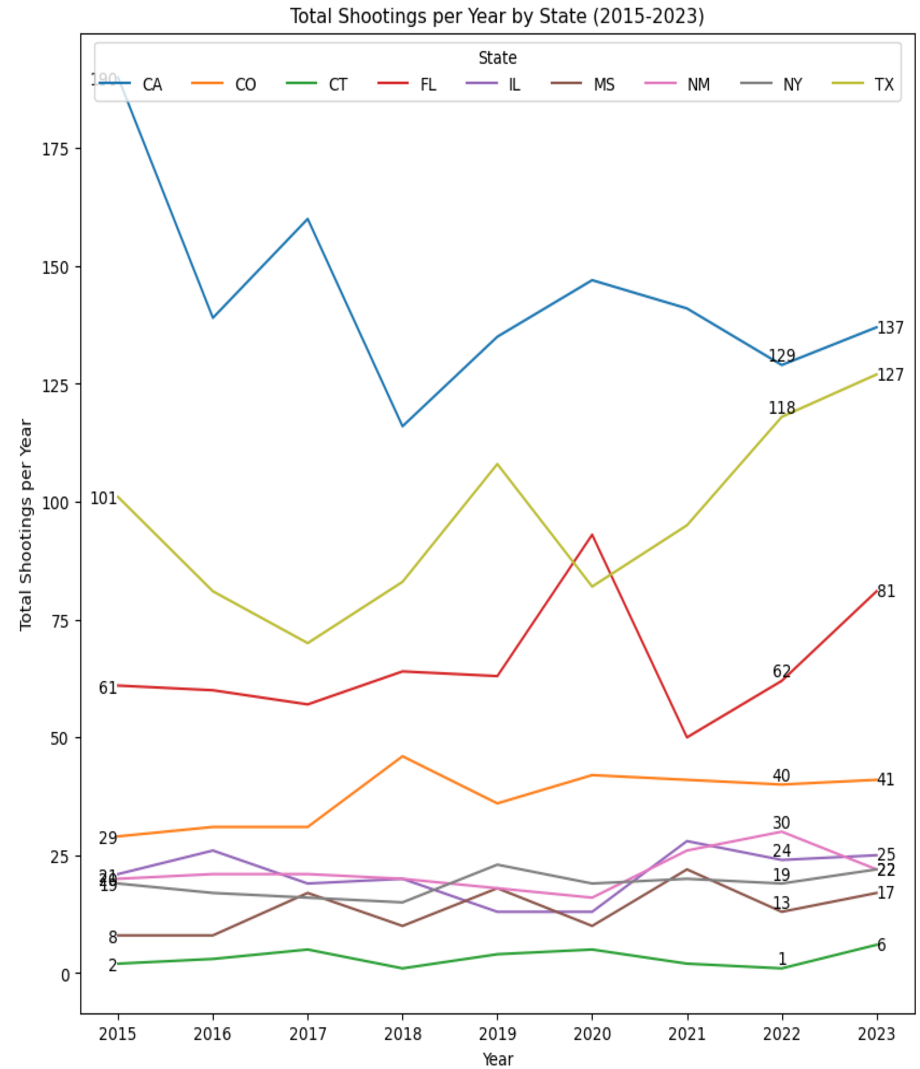
Decision: We failed to reject the null hypothesis.

Results

Does the data suggest qualified immunity has an impact on the number of police shootings in the three states which have eliminated it?

WE FAIL TO REJECT THE NULL

There was not statistically nor practically significant correlation between police shootings and qualified immunity during the selected and with this data sample.

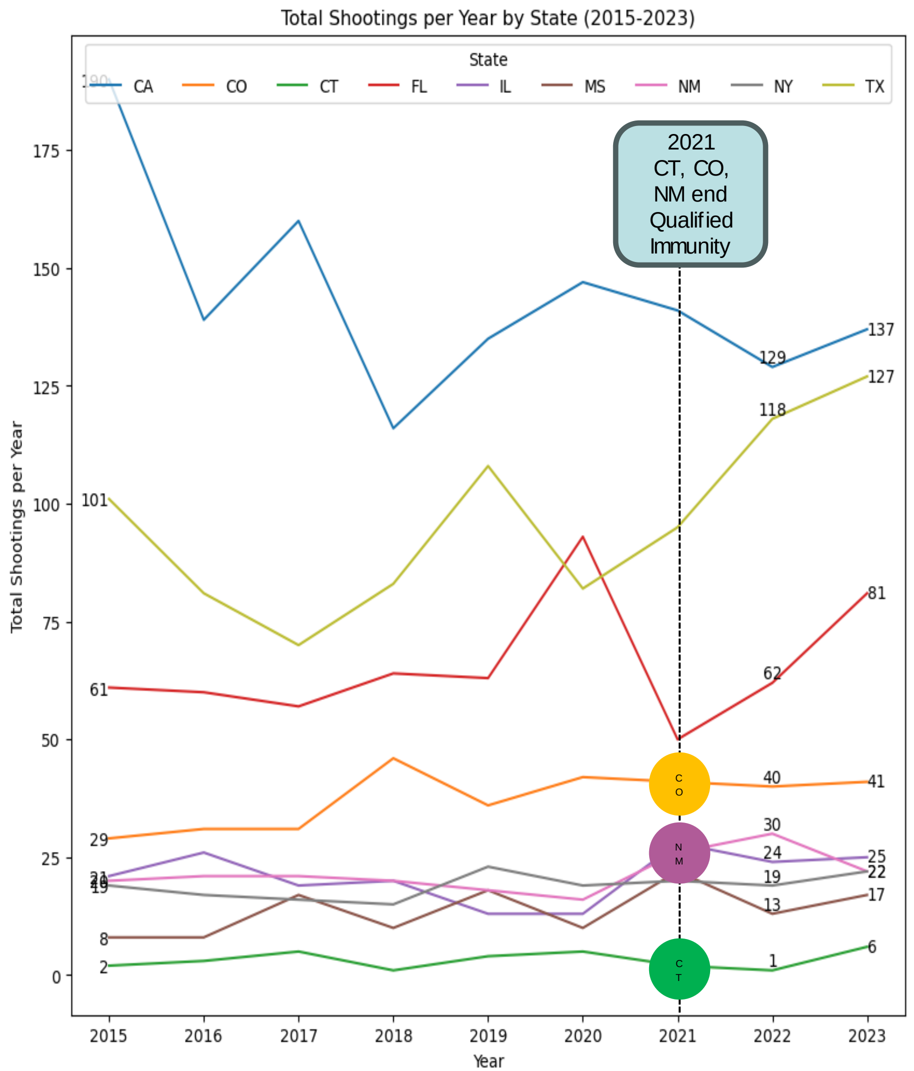


Results

We needed to understand why

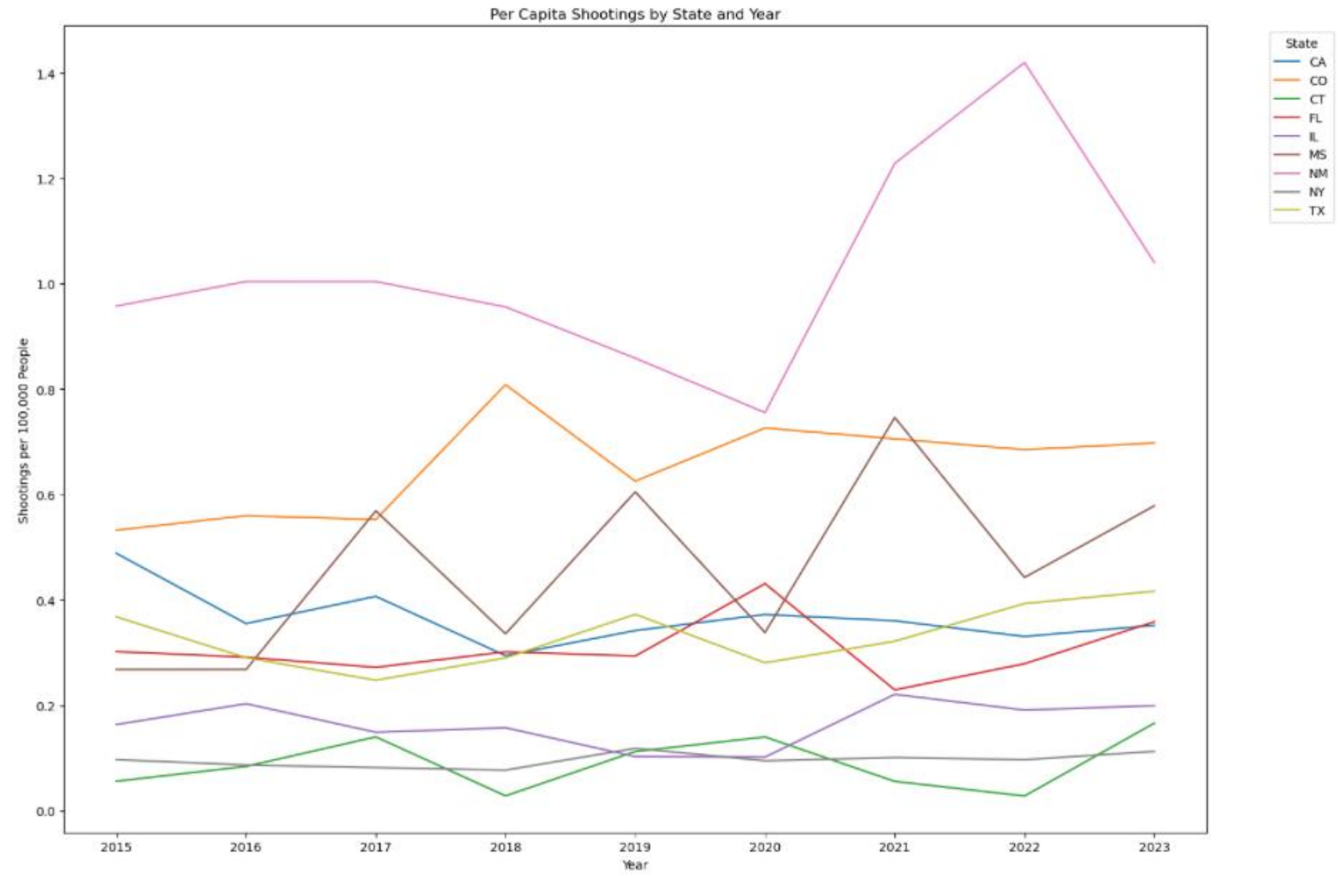
- Challenges:
  - Different States (size, policy)
  - Sample Size
  - Time Frame

- Is there a discernable pattern?



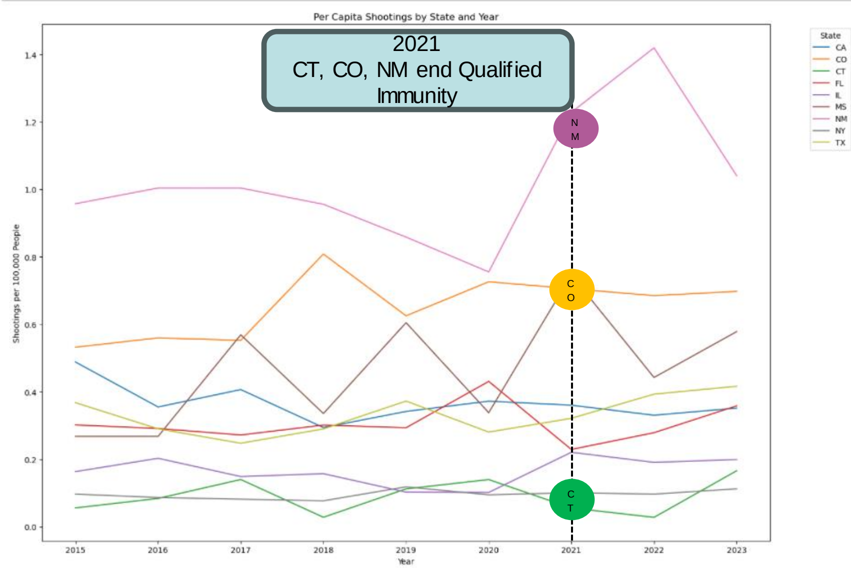
Results

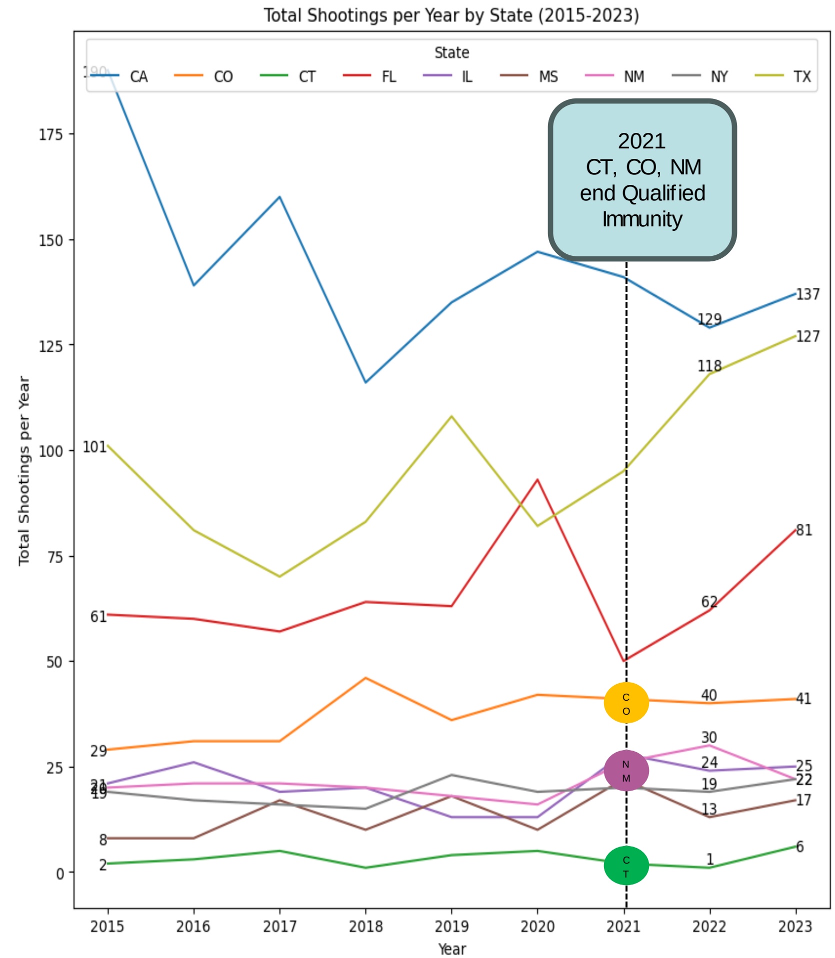
Shootings per 100,000 for each state from 2015-2023



Results

Shootings per 100,000 vs.
Total Shootings per State

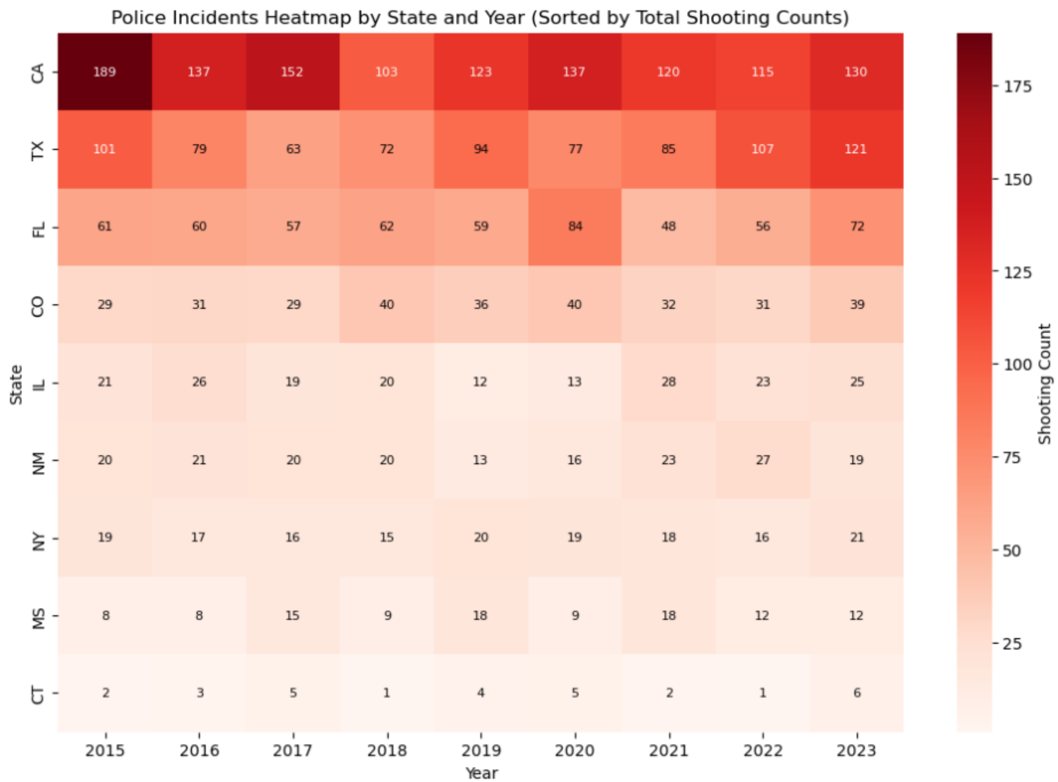




A different picture…

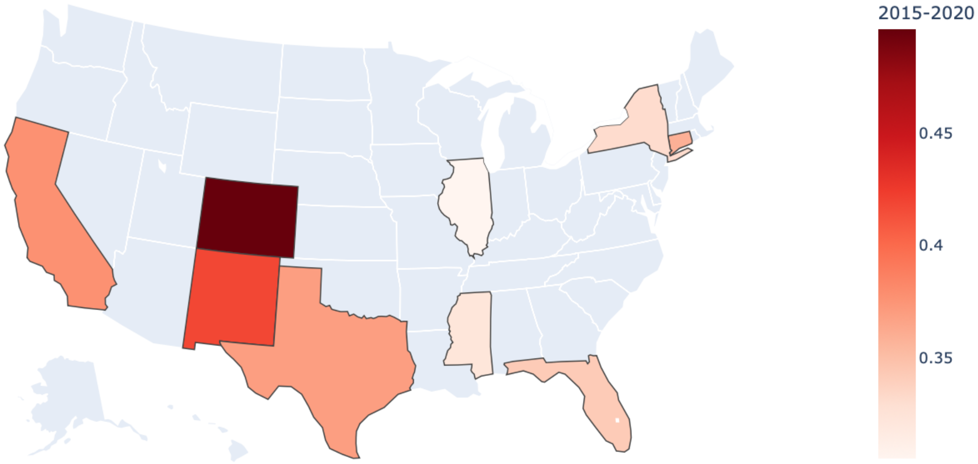
Results

Total Shootings for each state from 2015-2023



Results

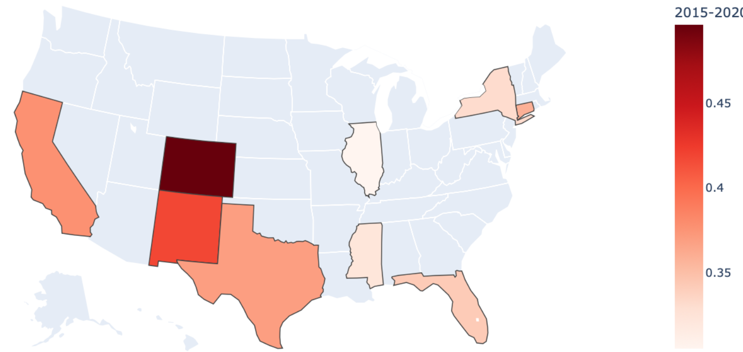
- How can we better understand this picture and make sure we're testing the right alternate hypothesis?



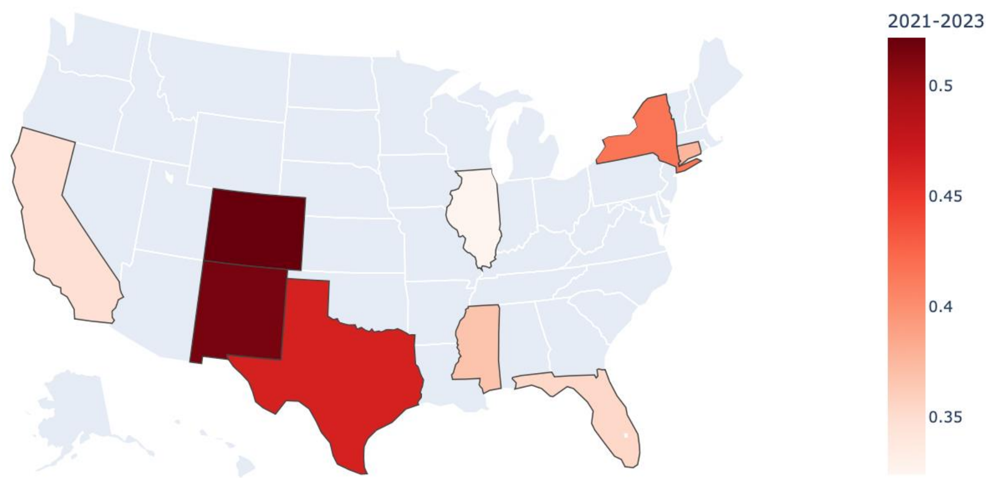
2015-2020

Results

- How can we better understand this picture and make sure we're testing the right alternate hypothesis?



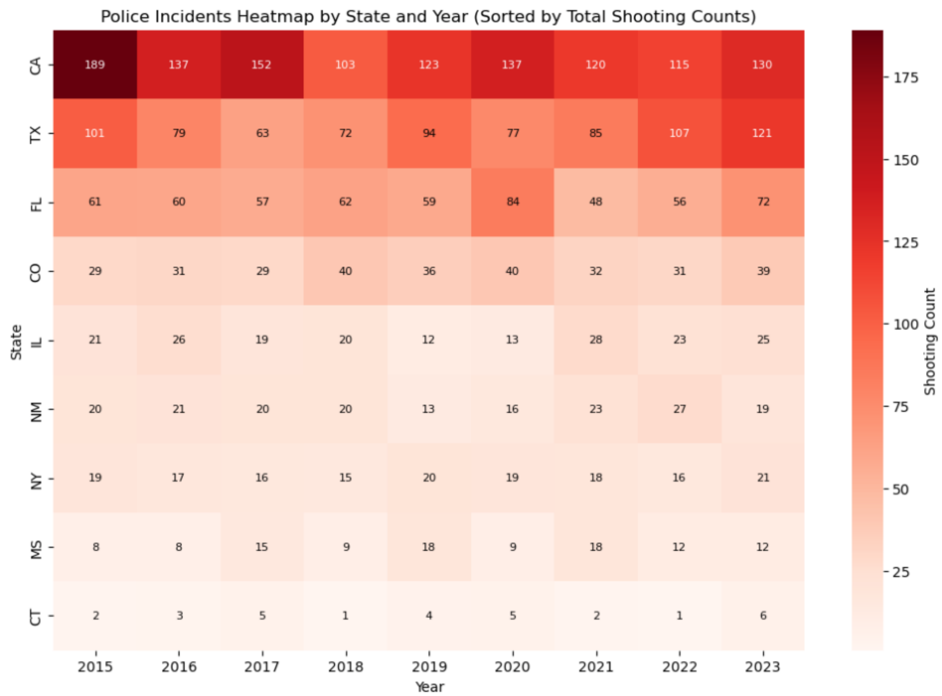
2015-2020

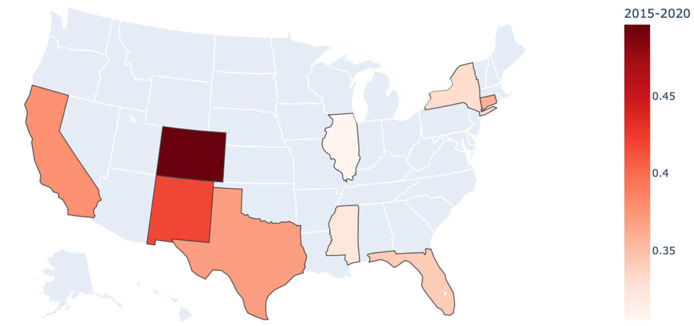


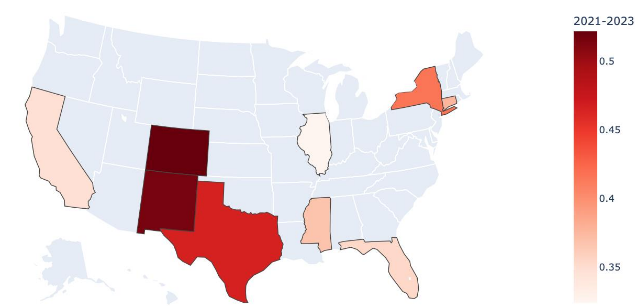
2021-2023

Results

"More is not always better…" –Yoda (maybe)







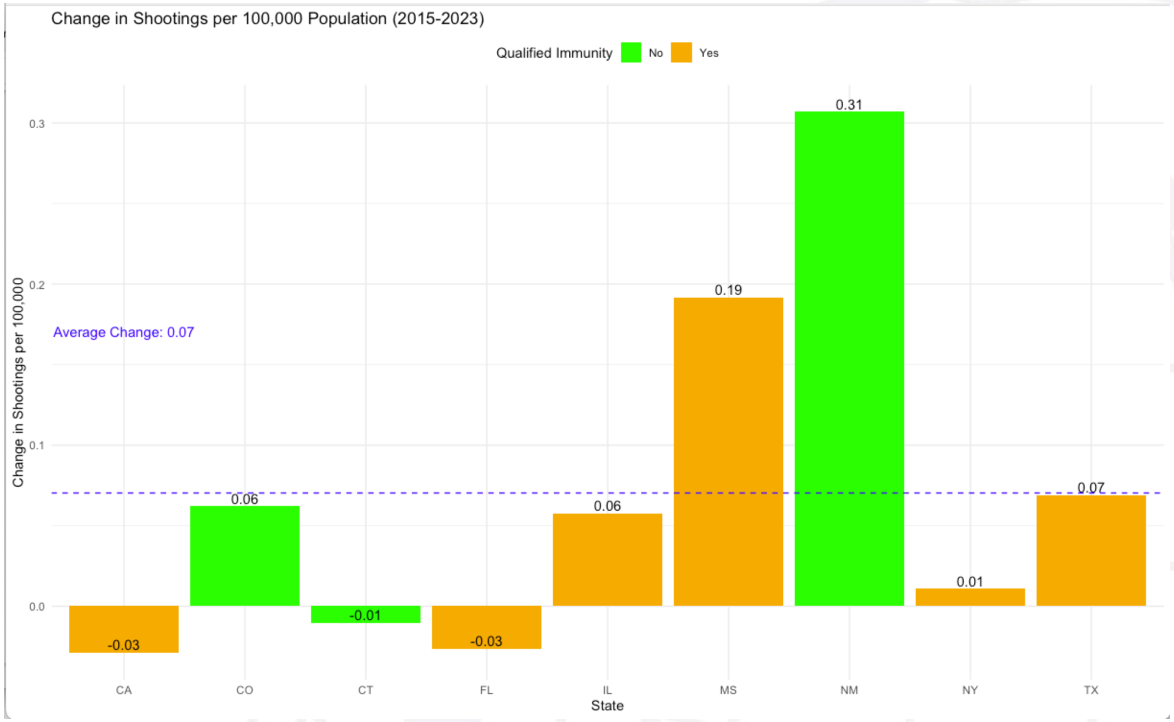
Was there a change?

Results

What was the change from 2013 to 2023?

There is not sufficient evidence to suggest qualified immunity decreases police shootings in CO, CT or NM.

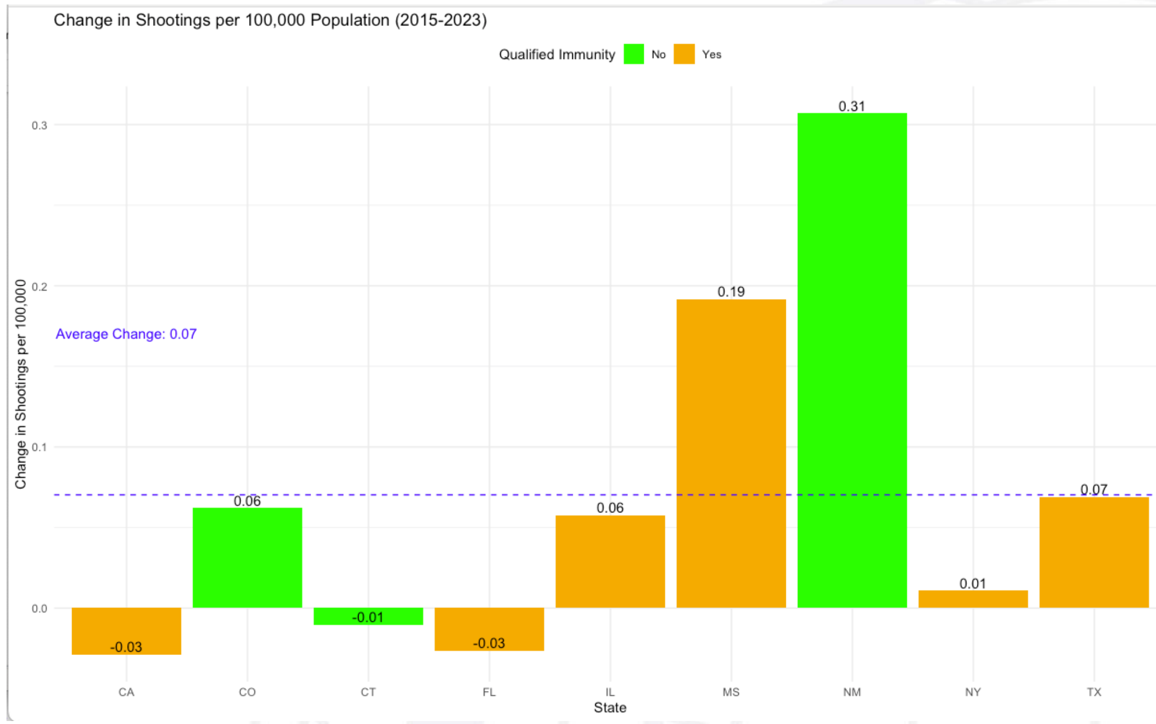
Factors:
-Data
-Time
-Variance



Results

Areas for Future Reasearch:

1. Policy Differences
2. Dept. Differences
3. Rural vs. Urban
4. Did no immunity make it worse?



Questions?