

Exploring Racial Bias in Traffic Policing: An Analysis of Six Eastern States of USA

Mahfuzur Rahman Chowdhury
Matriculation number: 23305223
Friedrich-Alexander-Universität Erlangen-Nürnberg
Erlangen, Germany
mahfuzur.rahman.chowdhury@fau.de

Abstract—This research investigates the potential racial bias in traffic policing across six eastern states in the United States, utilizing data from the Stanford Open Policing Project. The analysis is based on 9,494,312 traffic stops. The search focuses on racial patterns in searches conducted, arrests made, warnings issued and contraband discoveries. Key findings show that the arrest rates are higher for Black drivers in most of the contraband categories. Although the differences in arrest rates are subtle, but the observed patterns suggest racial disparities in traffic policing, warranting further investigation.

I. INTRODUCTION

This research leverages the Stanford Open Policing Project’s dataset [1] to investigate whether racial bias influences traffic policing decisions in the United States of America. The scope of this study is limited to six eastern states of the USA: New Hampshire, Rhode Island, Connecticut, Vermont, Massachusetts and Virginia. While the Stanford Open Policing Project provides 42 datasets for 42 states, the datasets proved too large to compile and analyze comprehensively within the constraints of this research. Therefore, the findings are representative of the selected six states and do not generalize to the entire United States of America.

The motivation for this research is inspired from the ongoing global fight for human rights and equality. Although racism has diminished significantly compared to the past, evidence of racial bias still exists in various domains. When such biases affect law enforcement, they pose a significant threat to societal development and the advancement of human rights.

This research provides the following key findings:

- 1) The proportion of searches conducted varies significantly from state to state.
- 2) Arrest and warning ratios differ across racial groups.
- 3) The number of warnings issued shows a consistent increase over time while the number of arrests remains relatively unchanged.
- 4) The distribution of contraband-related activities is not uniform across races.
- 5) Outcomes of searches, such as discoveries of drugs, weapons, alcohol or other contraband; show varying ratios depending on race.

These findings contribute to the broader understanding of potential racial biases in traffic policing of the six states and

understand the need for further investigation to ensure equity in law enforcement practices.

II. DATA

All six datasets used in this research are sourced from the Stanford Open Policing Project [1]. The datasets are licensed under the ODC-BY 1.0 license [2], which permits to freely share and adapt the data as long as proper attribution is provided.

In Table I the *Total Stops* column represents the number of stop records in each state’s dataset. The last row aggregates the data from all six states, showing a total of 9,494,312 rows across the combined dataset. The columns for each race (*White, Asian or Pacific Islander, Hispanic, Black, Other*) represent the number of stops attributed to each racial group. The *Unidentified* column indicates records where the race of the driver was missing in the datasets.

The combined dataset consists of 22 columns and 9,494,312 rows. Missing values in the dataset were replaced with the value ‘000’. In addition to the *race* column, the following columns are important for this research:

- *arrest_made*
- *warning_issued*
- *contraband_found*
- *contraband_drugs*
- *contraband_weapons*
- *contraband_alcohol*
- *contraband_other*
- *search_conducted*

In these columns, a value of 1 means *Yes*, 0 means *No* and ‘000’ indicates a missing value. These key columns form the foundation for analyzing racial disparities in traffic stops, searches and outcomes in this study.

III. ANALYSIS

This chapter presents a detailed analysis of the findings from this research. The analysis was conducted using Python¹ and for data processing sqlite3² and

¹<https://www.python.org/>

²<https://docs.python.org/3/library/sqlite3.html>

TABLE I
STATE SUMMARY TABLE

State	Total Stops	White	Asian or Pacific Islander	Hispanic	Black	Other	Unidentified
New Hampshire	259822	154037	2520	2629	4128	124	96384
Rhode Island	509681	344734	12826	53125	68579	1344	29073
Connecticut	18439	6057	176	5073	7104	29	0
Vermont	283285	266216	3607	2625	5741	280	4816
Massachusetts	3416238	2529780	166842	338317	351610	11008	18681
Virginia	5006847	3198750	125413	307345	1207078	100780	67481
Total	9494312	6499574	311384	709114	1644240	113565	216435

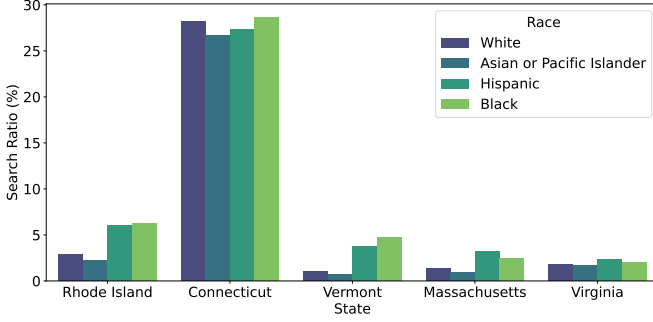


Fig. 1. Search Conducted Ratio by State and Race (Conducted vs Not Conducted).

pandas³ had been used. For visualization matplotlib⁴ and seaborn⁵ libraries had been used. The complete jupyter notebook file used for analysis and visualization is available in the project repository [3].

The `search_conducted` column in the dataset indicates whether the police conducted a search on the vehicle or checked the driver's alcohol level. Figure 1 presents the percentage of searches conducted by the police grouped by race for each of the five out of six states.

One notable observation is the absence of state for New Hampshire. This is because the `search_conducted` column is entirely missing in the New Hampshire dataset; making search conduct analysis for this state impossible.

From Figure 1, it is evident that the percentage of searches conducted for Black individuals in Vermont is higher than for other races, although the difference is not large. An interesting observation is that in Connecticut the search conducted ratio is between 26% and 29% for all races. In contrast, the other four states have search conducted ratios below 10% for all races. This indicates that searches are significantly more common in Connecticut, where police conduct a search approximately once in every four stops.

However, the key question this research aims to address is whether racial bias influences traffic policing. Figure 1 shows that the percentage of searches conducted for Black individuals in Vermont, Connecticut and Rhode Island are higher than other races, although the difference is not large. Because of

that, from Figure 1 there is no conclusive evidence of racial bias based solely on the search conducted data. Consequently, the research delves deeper into the dataset to investigate arrest patterns and their correlation with race.

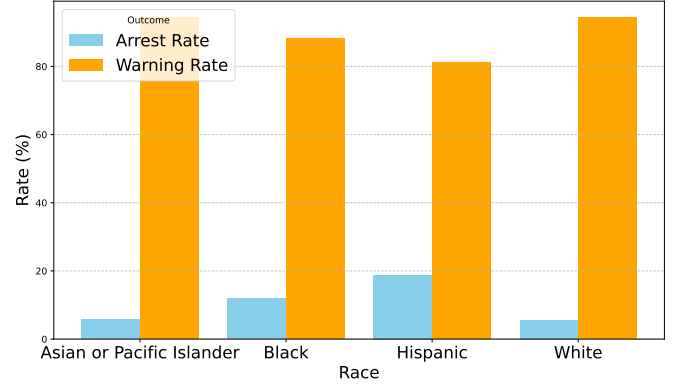


Fig. 2. Arrest vs. Warning Rates by Race.

Figure 2 provides an analysis of the differences between arrest and warning rates across racial groups. When a driver is found in possession of contraband: drugs, weapons, alcohol or other prohibited items; the police issue either an arrest or a warning. Depending on the severity of the situation. The goal of this analysis is to determine whether arrest rates disproportionately affect certain racial groups or not.

To achieve this, the combined dataset which was the output of the data pipeline was used to calculate the total number of arrests and warnings issued for each race. The ratios of arrests and warnings were then computed and visualized as percentages in Figure 2. The objective was to identify whether certain racial groups are more likely to receive warnings rather than arrests or vice-versa.

Figure 2 shows that Hispanic and Black individuals have a higher arrest-to-warning ratio compared to Asian or White individuals. This finding suggests that Hispanic and Black drivers are more likely to be arrested than warned during traffic stops involving contraband. Although, a counterargument could be made that Hispanic and Black individuals may be involved in more serious contraband related incidents, leading to a higher arrest ratio. To evaluate this hypothesis, further analysis of the contraband distribution across races was conducted.

Besides that, Figure 2, a significant gap is evident between

³<https://pandas.pydata.org/>

⁴<https://matplotlib.org/>

⁵<https://seaborn.pydata.org/>

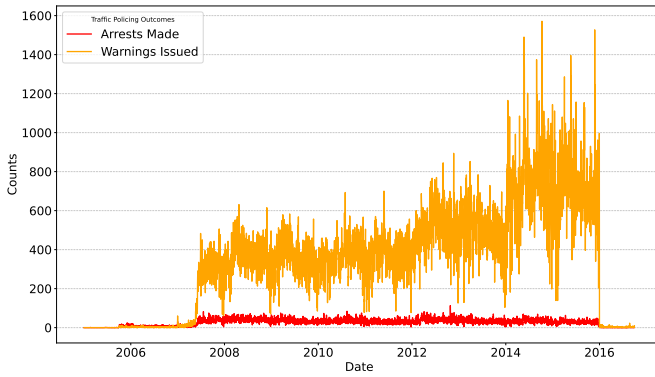


Fig. 3. Timeline of Arrests and Warnings Issued.

the number of arrests made and warnings issued. Observing Figure 3, it becomes apparent that over the years the number of warnings issued has consistently increased; while on the other hands the number of arrests has remained relatively stable. This shift in policing behavior likely contributes to the disparity observed in the arrest-to-warning ratio highlighted in Figure 2.

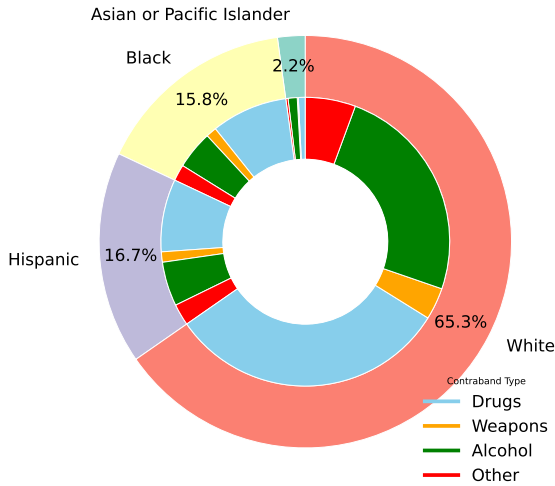


Fig. 4. Contraband Distribution.

Now, moving on to contraband distribution across races Figure 4 illustrates the distribution of contraband incidents among the four racial groups across the six states. To create this visualization, all contraband related events were aggregated and the percentage of incidents attributed to each race was calculated. Missing or ambiguous values in the race column such as '000' or 'other' were excluded from the analysis. The outer circle of the pie chart represents the overall distribution of contraband incidents by race.

As shown in Figure 4 White individuals account for the majority of contraband related incidents at 65.3%, followed by Hispanic with 16.7%, Black with 15.8% and Asian or Pacific Islander individuals with only 2.2%. Furthermore, the inner

circle of Figure 4 breaks down the distribution of contraband categories: Drugs, Weapons, Alcohol and Other for each race. It reveals that White individuals have the largest share across all categories of contraband.

Despite the results of Figure 4, it is observed that in Figure 2 Hispanic and Black individuals have a higher arrest-to-warning ratio than White and Asian individuals. While the differences in ratios are not extreme, they are notable. More detailed analysis was conducted of the outcomes associated with contraband incidents, broken down by race; to arrive at a more definitive conclusion regarding the research question of racial bias in traffic policing.

Figure 5 provides a detailed analysis of contraband related outcomes across different races. The figure consists of four subgraphs. Each subgraph represents one of the racial groups. Each subgraph visualizes the arrest-to-warning ratio for four categories of contraband: Drugs, Weapons, Alcohol and Other.

To compute this visualization, race specific data was extracted from the combined dataset. The data was then further categorized by contraband type. For each category, the columns `arrest_made` and `warning_issued` were aggregated then the arrest-to-warning ratio was calculated as a percentage. This process was repeated for each contraband type and for each racial group.

A comparison of the bar charts in Figure 5 reveals notable disparities. For contraband categories Drugs and Weapons: Black individuals have arrest rates exceeding 29% and 28% respectively. In contrast; Asian individuals have arrest rates of 23% and 18%, Hispanic individuals 23% and 22% and White individuals 25% and 24%. This indicates that Black individuals face at least a 4% or higher arrest rate than other racial groups when involved in the same type of contraband incident. A similar pattern is observed for Alcohol related contraband. Here the arrest rate for Black individuals is 19% while other racial groups have rates of 15% or lower. However, for Other types of contraband, Asian individuals exhibit the highest arrest rate at 32%.

While Figure 2 presents an overall view of arrest-to-warning ratios across races, it also could be argued that racial groups with higher arrest rates are simply committing more crimes. However, Figure 4 counters this argument by demonstrating that contraband distribution is not higher for the racial group that has more arrest rate. But further counterargument could be that the distribution of contraband related incidents is influenced by the population size of each racial group. This is a valid point. However, Figure 5 addresses this by normalizing the arrest-to-warning ratios for each race based on contraband type.

Despite accounting for these factors, the results consistently show a 4% to 5% difference in arrest rates for certain racial groups, particularly Black individuals in most of the cases. This disparity suggests the presence of a similar level of racial bias across police interactions in the six states analyzed in this study. While the differences may appear subtle, their consistency highlights the need for further investigation.

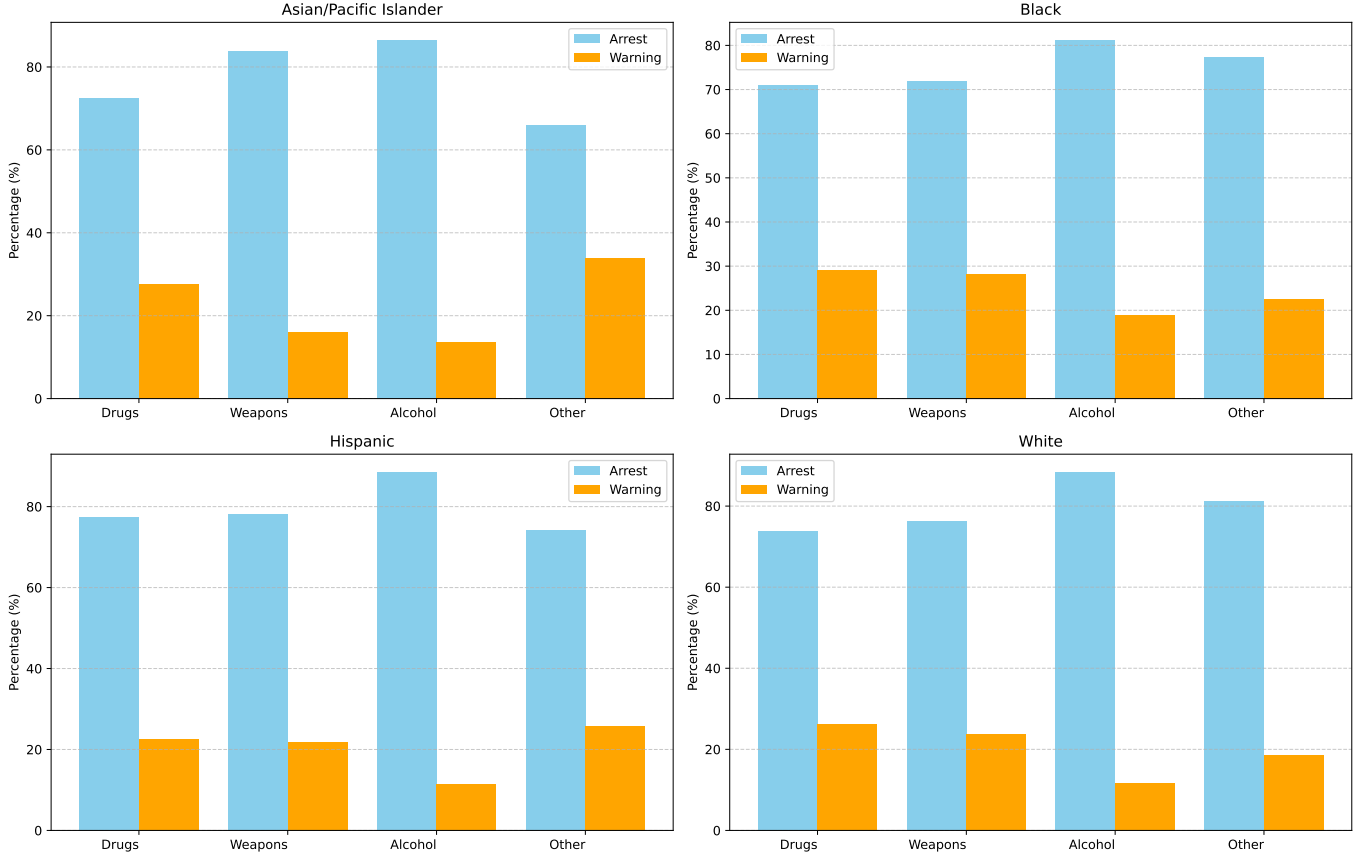


Fig. 5. Contraband Outcome Analysis by Race.

IV. CONCLUSION

This research set out to explore whether racial bias influences traffic policing decisions in six eastern states of the United States of America using data from the Stanford Open Policing Project [1]. The analysis focused on key features such as searches conducted, arrests made, warnings issued and contraband outcomes, aiming to identify disparities across racial groups.

The findings suggest evidence of racial disparities in certain aspects of traffic policing. Black and Hispanic drivers are disproportionately subjected to higher arrest rates compared to warnings. These disparities indicate the potential presence of racial bias in traffic policing practices. However, when analyzing searches conducted, the results appear relatively balanced across racial groups in most states, suggesting that racial bias may not be universal in these regions. Additionally, it is observed that the number of warnings issued has been increasing over time, while the number of arrests has remained constant.

Despite these findings, the research question cannot be answered completely due to several limitations. First, the analysis is restricted to six states and the results may not generalize to the entire United States of America. Secondly, the datasets

lack contextual variables such as level of Alcohol or amount of drugs; which could provide a deeper understanding of the observed patterns. Additionally, missing or incomplete data across all the datasets may have introduced biases into the analysis. Lastly, incorporating datasets of all remaining states for the research introduced computational limitations.

In conclusion, while this research provides valuable insights into the racial dynamics of traffic policing, the results highlight the need for further investigation to fully understand and address potential biases. Expanding the analysis to include more states, incorporating additional variables and addressing data gaps are essential steps toward answering the question comprehensively.

REFERENCES

- [1] Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., Ramachandran, V., Barghouty, P., Phillips, C., Shroff, R., Goel, S.: A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behaviour* **4**(7), 736–745 (2020). <https://doi.org/10.1038/s41562-020-0858-1>
- [2] Open Data Commons: Open Data Commons Attribution License (ODC-BY) Summary — Open Data Commons: legal tools for open data. <https://opendatacommons.org/licenses/by/summary/>
- [3] ThreeSwordAI: made-project-W24/project at main · ThreeSwordAI/made-project-W24. GitHub Repository. <https://github.com/ThreeSwordAI/made-project-W24/tree/main/project> (accessed 2023).