A large-scale analysis of racial disparities in police stops across the United States Extension

Yixin Cheng
321718

Xiangcheng Cao
322240

Jiaqi Wang
322784

***Abstract*—Background: Stanford Team E.Pierson et al.[1] assessed racial disparities in policing in the United States by compiling and analysing a dataset detailing nearly 100 million traffic stops conducted across the country. Their results indicate that police search decisions suffer from persistent racial bias. Aim: Based on their existing dataset, we have introduced a new data set of ethnic proportions and political tendencies in each state to further analyze whether there is racial bias in police search decisions. Methods: We propose three methods to extend the analysis:**

1) **We divide the states in the existing data into diverse states and poor diverse states according to the ethnic composition of each state, and analyze whether there is bias in the police search behavior decision based on the proportion of different races being searched.**
2) **We also divide the states into Democratic states and Republican states, just the blue and red states according to the U.S. presidential election since 1992(The red and blue states in the U.S. have been relatively stable in the past 30 years). We want to know whether the partisanship has effects on the potential racial disparities in traffic stops.**
3) **We extract the states with accessible inferred threshold, and get the population composition of these states. We aim to analyze if inferred threshold of a certain race is related to its population proportion in a state, if it is, which could prove there exists bias in the police search behavior decision.**

Results: Our results shows that police stops and search decisions suffer from persistent racial bias and partisan politics has effects on the traffic stop rates.

I. Introduction

A. Problem

More than 20 million Americans are stopped each year for traffic violations, The huge amount of traffic interception information makes this data very valuable for data analysis[2][3]. However, due to the standardization of American police information, the data format is messy and inconsistent, which brings great difficulties to data analysis[4].

B. Existing work

E.Pierson et.al compiled and analysed a dataset detailing nearly 100 million traffic stops to solve this problem.Their team propose the 'veil of darkness' test which use data from the time interval after dusk to assess potential bias in stopping decisions. Their analysis suggests that there is a bias in stopping the decision.And they apply the threshold test recently developed by Simoiu et al.[5] and refined by Pierson et al[6] to investigate potential bias in the post-stop decision to search drivers for contraband. They found that legalization of recreational marijuana reduced the number of searches of white, black and Hispanic drivers—but the bar for searching black and Hispanic drivers was still lower than that for white drivers post-legalization.

C. Motivation

In the original paper, the author mentioned that drivers may not live in the jurisdictions where they were stopped which makes interpretation of population benchmarks complicated. Part of the reason for this is that part of the data in the original article is based on cities. Due to the frequent traffic between cities, there are indeed drivers being searched that do not belong to the city. When we analyze the entire state as a unit, this attribute of the data is greatly weakened. Relatively speaking, since traffic in the state is more common,drivers are more likely to live in the jurisdictions where they were stopped. It provides the possibility for analysis by the police stops data and local ethic composition data.

In the US presidential election, those who support the Republican Party are called Red State, and those who support the Democratic Party are called Blue State. Comparing the two policies, the Democratic Party's policy is left-wing and supports racial equality and diversity. And policy trends can affect the potential racial discrimination factors in the state to a certain extent. Therefore, we also put forward a new analysis question, whether the political tendency of each state has an impact on the police search decision, so as to analyze whether there are racial biases.

On the basis of original paper, we used the original data set and added ethnic composition and political orientation data we get from official website to further analyze whether the decision-making of the US police in the law enforcement process is affected by the driver's ethnic factor.

II. Data description

Our data consists of four main parts

1) $city.csv$ and $state.csv$: Data set used by the original author, a data set detailing traffic stops carried out by 21 state patrol agencies over almost a decade. Since we are going to use the state as the unit for data analysis, the data provided by the municipal police departments of the city $city.csv$ has been ignored.
   Source:$https://openpolicing.stanford.edu$
2) $races.csv$: Data set contains the ethnic composition ratios of 21 states (white, black, hispanic). In the data from other sources, some data sets include latino as

white people. In this data set, white people specifically refer to White alone, not Hispanic or Latino, percent.The data set is crawled from the target website and stored in csv file.The data distribution is shown below as Fig. 1

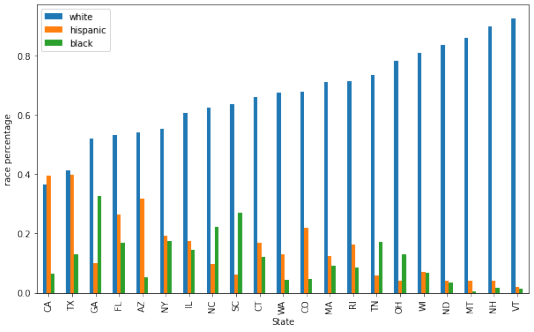

Fig. 1: State population composition

We can see that the ethnic composition of some states is very diverse. In some states, the ethnic composition is relatively single, with a certain ethnicity occupying the majority.
Source:$https://www.census.gov/quickfacts$

3) Red State or Blue State and $states\_info.csv$:Red State or Blue State contains the political orientation of the state to be analyzed in the following analysis (whether it is a red state or a blue state). The source is Wikipedia. $states\_info.csv$ includes the population and ethnic composition of the blue and red states of the 21 states which are given traffic stop information(not including the swing states). This source is Wikipedia and former $races.csv$. The parallel coordinate of the data is just like Fig. 2
Source:$https://en.wikipedia.org/wiki/Red\_states\_and\_blue\_states$

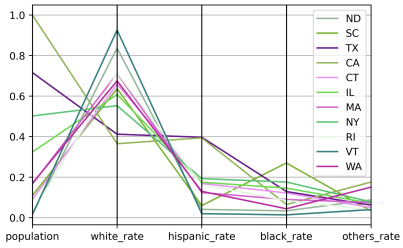

Fig. 2: Data of red and blue states

4) $crime\_rates.csv$:This data set contains the number of offences and crime rate of various races in the United States from 2000 to 2019 released by the US Department of Justice.The data set is crawled from the target website and stored in csv file. And the data is shown in Fig. 3 .
Source:$https://www.ojjdp.gov/ojstatbb/crime/ucr.asp$

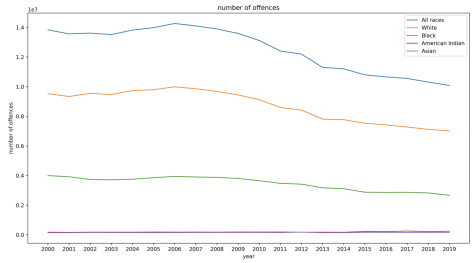

Fig. 3: Number of Offences of U.S.

III. Methods

A. *race search coefficient*

The first analysis will use races.csv and state.csv. As we mentioned in the motivation if we use the state as the unit to calculate the data, because the intra-state traffic flow is definitely much larger than the interstate traffic flow. Drivers who are stopped are likely to belong to this state, so we can approximate the overall driver racial composition of this state based on the proportion of the state's population.

The first analysis is based on a simple assumption. If the police treat drivers of all races equally, then the total number of stoppages for drivers of all races in the state in a year should be approximated by the total number of stoppages in the state multiplied by the race's population proportion in the state proportion.

We define the state with diverse ethnic composition as a diverse state, and a state with an absolute majority of whites as a poor diverse state.The Fig.1 sort by whites as a percentage of the state's population, from highest to lowest.As we can see from the image, the first three states: California, Texas, Georgia are typical diverse states and the last three states: Montana, New Hampshire, Vermont are typical poor diverse state.We specifically consider those 6 states when carrying out this analysis. We use the percentage of each ethnic group in the state multiplied by the total number of searches to get a estimator theoretically unbiased expected value if the police are in the decision-making process of the stop driver. The difference bewtween the real number of searches and our expected value can indicate that whether there is a bias and the magnitude of the bias.

However, since the total number of searches in each state differs in order of magnitude, and we cannot intuitively compare the differences between states through the difference. Assuming that the total number of white residents in a state is n times that of blacks, under other conditions, the total number of stoppages for white drivers should be approximately n times that of black drivers.Based on this simple proportional relationship, we propose a coefficient here to reflect whether there is a difference between the diverse state and the poor diverse state.

race search coefficient is defined as:

$$R1 = \frac{P_{white}}{P_{minority}} \tag{1}$$

$$R2 = \frac{P_s white}{P_s minority} \tag{2}$$

$$coefficient = \frac{R1}{R2} \tag{3}$$

Where P_{white} represent the percentage of whites people in the state,$P_{minority}$ represent the percentage of minority people (hispanic or black)in the state.$P_s white$ represent the real percentage of search number of white people in the state,$P_s minority$ represent the real percentage of search number of black people or hispanic people in the state.According to our assumptions, this coefficient should be around 1. If it is greater than 1, it means that the actual minority searched by the local police is larger than the estimated value, which means that they are more inclined to search for minority drivers.

B. blue state and red state comparsion

Because nearly half of the states in U.S. are swing states, So the red and blue states are not so many. Also, we have only 21 states' traffic stop data. After combining these two elements, we have three red states(ND, SC, TX) and eight blue states(CA, CT, IL, MA, NY, RI, VT, WA) with traffic stop information.

To make our results more convincing, we have two ways to go at the same time. One is we treat all red states as a whole and all blue states as another whole, after that, we just compare between the two wholes. However, in this way, some information has been ignored, like differences between these states and maybe these differences may have impact. So we propose another method, we find out three pairs of most similar red state and blue state, CA and TX, CT and SC, VT and ND. They are similar in population size, race ratio and so on. We compare each pair and find the result. By applying both methods, the conclusion may be more convincing.

How to find the most similar red state and blue state? After preprocessing(due to states' population varies too much, we should standardize it), a good way is to calculate the distance between every state. We apply multi-dimensional euclidean distance(other distances is the same result) to calculate the difference of every attribute of states and combine them to find the similar one. For example, TX matching CA, they both have large and similar population and various races, no other states are more similar to CA than TX.

C. threshold coefficient

In order to analyze if inferred threshold of a certain race is related to its population proportion in a state, we need to quantify the relationship of inferred threshold with population proportion, we could first compute the population multiple($R1$) and searching threshold multiple($R3$) of whites to minorities, then compute the coefficient of population and searching threshold($R1/R3$).

To compute them, we need to use the datasets races.csv and state.csv, after observing, some states' inferred thresholds are not accessible (NA), therefore, we only compute with the states having specific values of thresholds, after removing the states we won't use, there remain eight states: CT, IL, NC, RI, SC, TX, WA, WI. We only need to compute and analyze $R1$, $R3$, $R1/R3$ of these eight states. Among them, $R1$ is computed by equation (1). The other two are computed as follows(in order to distinguish from the coefficient in equation (3), we use $coefficient_T$ to represent the coefficient here):

$$R3 = \frac{T_{white}}{T_{minority}} \tag{4}$$

$$coefficient_T = \frac{R1}{R3} \tag{5}$$

Here T_{white} represents the inferred threshold of whites in a state, and $T_{minority}$ is of minority(blacks and Hispanics). If the $coefficient_T$ of poor diverse states is much larger than diverse state, it represents that the population ratio is not in harmony with the threshold, which proves there exists bias in the police search behavior decision.

IV. Results and Discussion

A. race search coefficient for diverse state and poor diverse state

We use three diverse states and three poor diverse states to apply out method.The results are shown below in TableI and Fig9, Fig10 We can see that in poor diverse states, the coefficient between blacks and whites is much greater than 1, which proves that police in this state are more inclined to search for black drivers. And we calulate the pearson correlation coefficient bewtween race search coefficient$_{white\ and\ black}$ and P_{white}. The result is 0.44237 and the p value is 0.04463 which is smaller than 0.05.This shows that the proportion of whites in a state has a strong correlation with this coefficient.

This result indicates that police stops and search decisions suffer from racial bias.And for diverse state, the police are more fair in treating drivers of different races. In a poor diverse state, the police are more inclined to search black drivers when deciding whether to stop search.

TABLE I: race search coefficient

state	coefficient$_{white,black}$	coefficient$_{white,hispanic}$	Diverse State
CA	1.067827	0.710639	Yes
GA	0.725888	0.315905	Yes
TX	0.642851	0.721607	Yes
MT	2.399716	0.526377	No
NH	1.341656	0.391117	No
VT	1.495513	0.538128	No

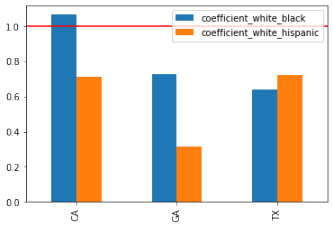

Fig. 5: Diverse State race search coefficient

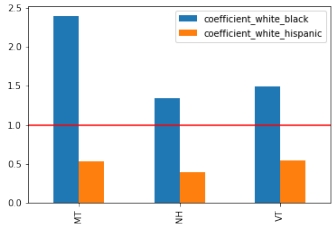

Fig. 6: Poor Diverse State race search coefficient

B. result for red and blue state

If we use race search coefficient to see the result, it is difficult to find obvious differences between the red states and blue states. There are both racial bias in red states and blue states. When comparing three pairs(CA and TX, CT and SC, VT and ND), it is the same that both two classes of states have racial disparities and no obvious gap or difference. We also apply correlation coefficient and clustering method, find the threshold ratio of white and hispanic is the most correlated with the state class but no reference for exploring the racial discrimination.

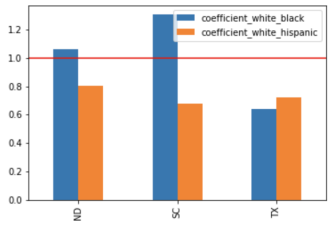

Fig. 7: Red State race search coefficient

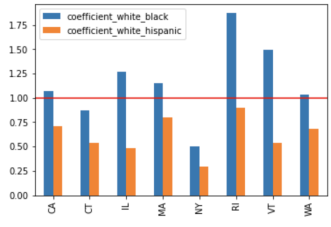

Fig. 8: Blue State race search coefficient

C. threshold coefficient for diverse and poor diverse states

In order to distinguish diverse and poor diverse states, in this part, we define the state with $R1 \leq 5$ is diverse state, while with $R1 \geq 5$ is poor diverse state. After classification, the diverse states are IL, NC, SC, TX; the poor diverse states are CT, RI, WA, WI. The computed $coefficient_T$ are shown in Fig.9 and Fig.10:

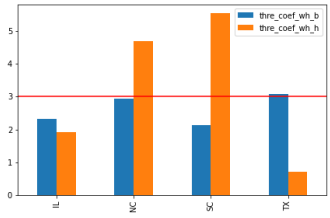

Fig. 9: Diverse States threshold coefficients

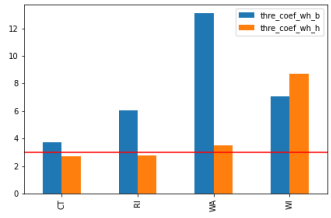

Fig. 10: Poor Diverse States threshold coefficients

Observing the results, we could find that in four poor diverse states, the $coefficient_T$ are all larger than 3, while only TX is larger than 3 in diverse state. This proves that in poor diverse state, the population composition is not commensurate with threshold, that is, the population ratio multiple is much greater than the threshold multiple, which proves that there exists bias in the police search behavior decision.

V. Conclusion

Through the above methods, whether through the threshold coefficient or the race coefficient, black drivers are more likely to be stopped by the police than white drivers in a poor diverse state.In the diverse state, black drivers are treated fairly, but they are still more likely to be searched by the police than white drivers.Meanwhile in the existing data, there is no obvious data to support that hispanic drivers are affected by racial factors in the decision-making process.We can conclude that the demographic composition of a state will have an impact on the police stop decision, and a state with a higher percentage of white people is more likely to be affected by racial discrimination.

For red states and blue states, the existing data are too few, and since most of the trafiic stop data set are swing states, no obvious difference has been found to prove that political tendencies affect police decisions in the process of searching for drivers.

VI. Limitations and Future Work

Because there are some problems in the original data set: some attributes of many states are missing, such as inferred threshold, and the state data does not include the precise time when searched. And for most states and cities, the data only contains information before 2016. These problems limit us cannot try some more intersting ideas and methods.For example, use economic data to distinguish between states and explore whether racial factors will affect the police's decision-making during the search process.

During the blue and red states and race search coefficient analysis, it seems that the black are more likely to be searched than the hispanic. As we can see it from Fig.7 and Fig.8. To analyze this, we want to get some prior experience from the crime rate data, although it shows higher black_crime_rate than black_population_ratio, but there are reasons to question its reliability, like too many offences included, not distinguishing hispanic from white and so on. So we need more detailed official data of crime rate. At present, we have not found better data. Because the existing data contains a small number of states, and most of them are swing states, if the data is sufficient, we can further explore whether political factors will affect the police search behavior decision through the clustering method.

References

[1] E. Pierson, C. Simoiu, J. Overgoor, S. Corbett-Davies, D. Jenson, A. Shoemaker, V. Ramachandran, P. Barghouty, C. Phillips, R. Shroff *et al.*, "A large-scale analysis of racial disparities in police stops across the united states," *Nature human behaviour*, pp. 1–10, 2020.

[2] D. M. Provine, "Suspect citizens: What 20 million traffic stops tell us about policing and race. by frank r. baumgartner, derek a. epp, and kelsey shoub. cambridge: Cambridge university press, 2018. 292p. 99.99*cloth*,24.99 paper." *Perspectives on Politics*, vol. 17, no. 3, pp. 892–893, 2019.

[3] L. Langton and M. R. Durose, *Police behavior during traffic and street stops, 2011*. US Department of Justice, Office of Justice Programs, Bureau of Justice ..., 2013.

[4] E. Davis, A. Whyde, and L. Langton, "Contacts between police and the public, 2015."

[5] C. Simoiu, S. Corbett-Davies, S. Goel *et al.*, "The problem of infra-marginality in outcome tests for discrimination," *The Annals of Applied Statistics*, vol. 11, no. 3, pp. 1193–1216, 2017.

[6] C.-D. S. . G. S. Pierson, E., *International Conference on Artifcial Intelligence and Statistics*, 2018.