Racial Disparities

Nicolas Delamaide Dahn Youssefi James Bardet

Abstract

We looked at racial disparities in policing in the United States by analyzing over 18 million stops. We were particularly interested in the potential disparities that could arise depending on the race of the officer conducting the stop. We looked at the racial distribution of the population in a given state and compared it to racial makeup of stopped drivers. We found that black drivers were more likely to be stopped. By conducting a veil-of-darkness test, we found that white officers arrested less black drivers after dusk than during daylight, while black officers arrested about the same amount of white drivers during and after daylight. This suggests that there is a bias against black drivers from the part of white officers. We also assessed the potential bias in search decisions with the outcome test. We found that white officers had a bias against minority drivers; for Hispanic and black officers, the data seemed to indicate there was differences in their stop decisions based on the race of the driver but the small amount of data didn't allow us to find statistically significant differences.

1 Introduction

The **rate of white stops** and the **rate of black stops** will be analysed. The assumption is that in an unbiased environment, white people should be stopped at the same rate as black people. Then, we will perform a veil-of-darkness test to see whether the stop rate of white people will be decreasing after sunset. Finally we will assess potential bias in search decisions by carrying out the outcome test.

2 Related Work

This work is an extension of the paper*"A large-scale analysis of racial disparities in police stops across the United States"* [1] published by the Stanford Open Policing Project. In the paper, they assessed racial disparities in policing in the United States. Overall the paper found that there is a bias in stop decisions and search decisions and even suggested some policy interventions to mitigate those disparities. In our work, we put an emphasis on the race of the officer to asses disparities.

3 Data collection

census.gov dataset: An API provided by the US Census Bureau. A request to this API is made to download the data. The number of black and white inhabitants for each US County and States is requested. The API request URL is available here.

The stop data was collected from the Stanford Open Policing Project. We downloaded files containing police stops for states or cities that fitted our criteria, explained in the following section. In order to work with a manageable amount of data we only extracted from the files the columns we were interested in, in particular: *date, time, county_name, subject_race, officer_race, search_conducted, contraband_found*. We also added a column for the state where the stop occurred in case some counties had the same name but were in different states. Moreover, we had to remove missing values (NaN values) as well as drop stops where the race of the officer or driver was unknown, or specified as "asian/pacific islander" and "other".

4 Dataset description

census.gov dataset. Dataset from the US government census data. This dataset contains the Decennial Census data with detailed tables focusing on race. For each US County, the number of black and white living in the County are stored. It is important to notice that each County report census data in different ways, hence comparing County/State census data is not always relevant. Hence the conclusions made with those data must be done carefully. Dataset stats:

- Number of counties or independent cities: 3'221
- Total number of black inhabitants 39'390'817
- Total number of white inhabitants: 226'378'365
- Date of the dataset: 2010

Stanford Open Policing Project. Standardized stop data made available by the Stanford Open Policing Project. We used datasets where columns needed were present, thus statewide data for the states of Florida, South Carolina, Washington and the data from the cities of Chicago (IL), Louisville (KY), Pittsburgh (PA) and Long Beach (CA). However, for some analyses we had to use only a subset of that data as not all datasets had the required information. Specifically, for the veil-of-darkness analysis we had to use locations which the time of the stop was available; for the outcome test we used stop data for which we knew the county as well as whether the search turned up contraband. In the end the dataset we used contained over 18 million stops carried across 7 states and major cities. Among these stops, 16 million were made by white officers, 1.8 million by black officers and almost 1 million by Hispanic officers. On Table 1 we can see for white, black and Hispanic officers the breakdown of stops they made by race of the driver.

TABLE 1: Proportion of drivers stopped from each race depending on the race of the officer.

Officer Race	White drivers	Black drivers	Hispanic drivers
White	69.45%	21.47%	9.08 %
Black	54.03%	33.31%	12.66%
Hispanic	47.56%	18.70%	33.74%

5 Methods

Stop rate. To compute the **rate of white stops** we will use the number of white drivers stopped in a County, and the number of whites registered in the Census data of that County. We will do the same for the **rate of black stops**.

Finally, we will observe the proportion of Counties where the black drivers are more likely to be stopped.

Veil of darkness. We decided to test if there was a significant difference among different officer race between the number of stop per driver race before sunset, where the police officer can recognize driver's race and after sunset, when the officer can't. We only kept in our dataset black and white officers and black and white drivers as it's not always evident to recognize one's race other than those before sunset. We used a library called astral to find out the sunset time at each day of the year in each country to then keep only the data from the stops appearing 0-90 minutes before sunset and 30-120 minutes after. We didn't keep data from 0 to 30 minutes after sunset as it's neither day nor dark. We ended up with 944'036 stops across white and black officers.

Outcome test. To assess the bias in search decisions we carried out an outcome test. We focused especially on finding if there are disparities in these search decisions based on the race of the officer as well as the race of the driver. In particular we wanted to look at if an officer of a given race had a bias in his search decision depending on the race of the driver he stopped. To run the outcome test we defined the number of stops in county c that resulted in a search S as $N_{c,S}$ and the number of searches that resulted in contraband being found (H) in that county as $N_{c,H}$. The hit rate for officers of race R in county c on drivers of race $r = R$ is denoted by $\phi_{c,[r=R]}$ and the hit rate for officers of race R on drivers of race $r \neq R$ is denoted by $\phi_{c,[r\neq R]}$. We then define the two hit rates as :

$$\phi_{c,[r=R]} = \frac{N_{c,H\,[r=R]}}{N_{c,S\,[r=R]}} \qquad \phi_{c,[r\neq R]} = \frac{N_{c,H\,[r\neq R]}}{N_{c,S\,[r\neq R]}}$$

To compare the hit rates on a county basis, we kept in our analysis only counties where an officer made searches on drivers of the same race as him as well as on a different race. In the end, we used the data from 84 counties for white officers, 60 counties for black and 51 for Hispanic officers.After computing the hit races for each county, we did a student t-test under the null hypothesis that the mean hit rate across counties $\mu_{\phi_{[r=R]}}$ for drivers of the same race as the officer was equal to the mean hit rate $\mu_{\phi_{[r\neq R]}}$ for drivers of a different race than the officer. That is H°: $\mu_{\phi_{[r=R]}} = \mu_{\phi_{[r\neq R]}}$ with $\alpha = 0.05$

6 Results

Stop rate. The **rate of white stops** and the **rate of black stops** have been computed for each County, classified by the *police officer race*. rate = number_of_stops / number_of_inhabitants. It can be greater than 1 since it covers stops over many years, and the census data may not contains all the county population. 144 Counties have been analyzed, from 3 different states: Florida, South Carolina and Washington.

It is observed that the probability to be stopped as a black driver is higher than the probability to be stopped as a white driver. On average across all analysed counties, the **rate of white stops** by a *white police officer* is 1.138, and the **rate of black stops** by a *white police officer* is 1.910. This difference is reduced when the police officer is black: 1.910 goes down to 0.140 and 1.138 goes down to 0.137. As expected, the rate to be arrested by a black police officer is lower since there is less black than white police officer.

The stop rate ratio has been computed: ratio = white_stop_rate / black_stop_rate. We found that the *rate of black stops* is often greater than the *rate of white stops* (ratio < 1). Indeed, we found that the proportion of Counties with a ratio < 1 is 3.5 times more important than the counties with a ratio > 1. We also found that the *officer race* does seems not make a significant difference at the County level, even if it tends to decrease the racially disparate treatment

Finally Figure 1 shows the ratios of each County. It is also observed that in many counties, the ratio is lower than 1, hence the *rate of black stops* is greater than the *rate of white stops* in those counties.

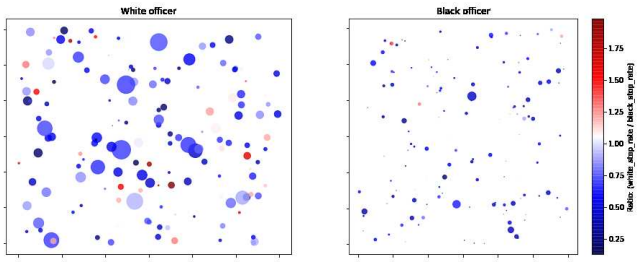

FIGURE 1: Ratio for each County, classified by *police officer race*. Each County is represented by a circle, and the area of the circle is defined by the number of stops. Ratio lower than 1 (in blue) implies that the **rate of black stops** is greater than the **rate of white stops**

Veil of darkness. As we can see in Figure 2 the mean of the number of stops for black people from a white officer seems to decrease after dark. For black officers, the difference doesn't seem to be as big (note : we have more than 4 times more data for white officers than black). This would suggest a bias in stop decision from white officer and as seen in Table. 2 the t-test does shows that the difference in means is significant.

TABLE 2: T-test for Veil of darkness

Officer Race	Driver Race	Mean Before Sunset	Mean After Sunset	T-Statistic	P-Value
White	Black	13.8%	12.7%	-2.921739	0.009980
Black	White	72.4%	72.1%	0.755128	0.461144

Assessing bias in search decisions. After looking at the search rates by race of the officer, we found that white officers on average searched more drivers who were not white than

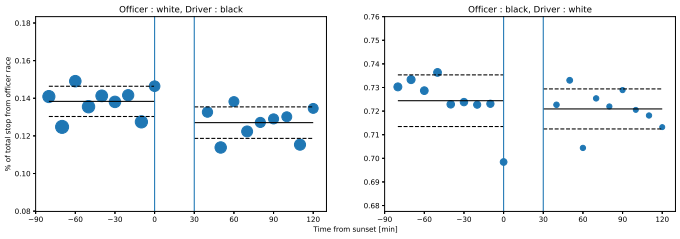

FIGURE 2: We computed the percentage of stops that involved black or white drivers done by white or black officers for a series of 10 min-periods before and after sunset. The figure is based on 944'036 stops with point size according to the total number of stops each bin. The line at t=0 shows sunset time. The dashed horizontal lines are 95% CI.

drivers who were white. In contrast, black and Hispanic officers searched more stopped drivers who were the same race as them than those of a different race. In particular, the gap in search rates for drivers of the same race as the officer versus those of a different race was -1% (t-statistic -6.7, P < 0.001) for white officers, 1% (t-statistic 2.0, P < 0.001) for black officers and 1.7% (t-statistic 2.7, P < 0.001) for Hispanic officers. Though the disparities in search rates are not necessarily due to discrimination. For example, black and Hispanic drivers could carry contraband at higher rates than white drivers and in consequence the search rate for these two group of drivers would be higher even with no racial bias [1].

To assess whether the gap was the result of biased decisions by the officers we applied the outcome test [1]. We performed the test using a subset of our data for which we had the name of the county where the stop occurred and whether the search turned up contraband. This amounts to the statewide data for Washington and South Carolina for a total of 361 923 searches. In Figure 3 the hit rates of drivers of the same race as the officer are plotted against the hit rates of drivers of a different race. We found that when it was a white officer conducting the search, the hit rate was higher on white drivers than on drivers who weren't white, suggesting that search decisions of white officers may be biased against minorities. In contrast, the data seems to indicate that Hispanic officers have a tendency to search more Hispanic drivers than other drivers, though we can't draw definitive conclusions as we only had enough data on 6 720 searches conducted by Hispanic officers. For black officers, we found that the hit rate is higher on black drivers suggesting a bias against non-black drivers, but again as for Hispanic officers the difference was not statistically significant probably due to the small data we had on black officers (28 099 searches). Overall, the gap in hit rates for drivers of the same race as the officer versus those of a different race was 4.3% for white officers, 4% for black officers and -7.7% for Hispanic officers, as summarized on Table 3. Our outcome test for searches made by white officers thus showed a similar bias against minority drivers as the outcome test conducted in the reference paper. For searches conducted by black and Hispanic officers though we didn't find any statistically significant difference in hit rates even though the data seems to suggest otherwise.

TABLE 3: Outcome test statistics under the null hypothesis that the mean hit rate for a driver of the same race as the officer is the same as the mean hit rate of a driver of a different race.

Officer Race	Search Count	Mean Same Race	Mean Different race	T-Statistic	P-Value
White	327 104	23.8%	19.5%	2.88	0.004
Black	28 099	18.2%	14.2%	1.39	0.166
Hispanic	6 720	8.5%	16.2%	-1.86	0.066

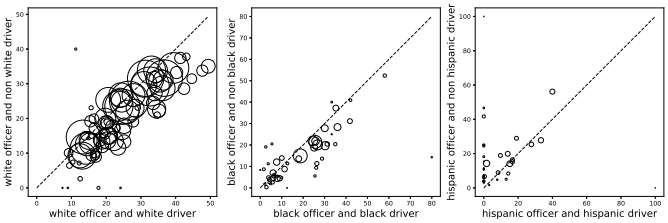

FIGURE 3: Hit rates by race of officer and location. The data represents a total of 361 923 searches in 3 states. Each point corresponds to a county, scaled by the number of searches. The left panel represents searches made by white officers, the middle one by black officers and the right one by Hispanic officers.

7 Conclusion

The observed disparities in stops rate are a starting point for understanding racial disparities in traffic stops. The officer race is also related to the racial stop rate disparities. But they do not provide strong evidence of racially disparate treatment. In particular, stop rates do not account for possible race-specific differences in driving behaviour (i.e. amount of time spent on the road). With the outcome test we showed that white officers have a bias against minority drivers in their search decisions. For Hispanic and black officers, there seems to be a bias in treatment based on the race of the driver but the amount of data we had wasn't enough to say that these differences are statistically significant. Also, the veil-of-darkness test showed a significant difference in stop decisions for white officers against black drivers. But this conclusion must be used sparingly as other factors are playing a role in stop decisions. The analyses should be conducted again with more data on stops conducted by black and Hispanic officers in particular in order to obtain more robust results.

References

[1] Pierson, E., Simoiu, C., Overgoor, J. et al. A large-scale analysis of racial disparities in police stops across the United States. Nat Hum Behav 4, 736–745 (2020).

[2] Aarts, H., & Dijksterhuis, A. (1999). How often did I do it? Experienced ease of retrieval and frequency estimates of past behavior. *Acta Psychologica, 103*(3), 77–89. http://someurl.html.

[3] Barberis, N. & Thaler, R. (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris & R. Stultz (Eds.), *Handbook of the Economics of Finance,* pp. 1053–1123. Elsevier Science, North Holland, Amsterdam

[4] Grogger, J. Ridgeway, G. Testing for racial profiling in traffic stops from behind a veil of darkness. J. Am. *Stat. Assoc. 101*, 878–887 (2006).