Racial Bias Analysis of Investigatory Stop Reports from the Chicago Police Department, 2018-2019

Jay Reiter

June 2020

1 Introduction

1.1 Background on Investigatory Stop Reports

The goal of this project is to quantify the extent of racist bias demonstrated by the Chicago Police Department (CPD) when carrying out investigatory stops. Investigatory stops—synonymous with stop-and-frisk—are notorious catalysts for invoking racial bias in police forces since they allow officers to stop and question people on the street with relatively little restriction.

To detain suspects, officers are only required to demonstrate "reasonable articulable suspicion that the person is committing, is about to commit, or has committed a criminal offense." While reasonable articulable suspicion must "constitute more than a mere hunch," it is not as strong a condition as probable cause. Indeed, no material, corroborated, or explicit evidence needs to be provided in the report; the officer's word suffices.

Investigatory stops differ from consensual encounters by the officer's demonstration of force. A court my use any of the following criteria to determine that an encounter was indeed an investigatory stop:

1. Threatening presence of several officers
2. Display of a weapon by officer
3. Use of language or tone of voice indicating that compliance with the officer's request might be compelled
4. Officer blocks person's path
5. Choice to end the encounter is not available to the person

Officers may choose to conduct a Protective Pat Down of the suspect's outer clothing should they have reason to believe the suspect is in possession of a weapon or contraband. If the officer believes they have found such an item, they can retrieve it and are required to report on it if it is indeed illicit. Pat downs require the same reasonable articulable suspicion threshold as the investigatory stop itself.

As an attempt to ensure the appropriateness of each investigatory stop, officers are required to complete a report (the ISR) that details the encounter. Indeed, in Article III, Section E of CPD's ISR Directive, it states that "Department members will not engage in racial profiling or other bias-based policing when conducting Investigatory Stops."

1.2 Motivation

It is frequently in investigatory stops that police officers resort to excessive violence–particularly against Black people–hence, addressing the racism they facilitate is an extremely salient issue of justice and public safety. The objective of this report is partly to inform police units which may use this data to address issues within their system, but primarily to educate voters, citizens, and taxpayers who have the power to enact systemic change in the way we manage crime in this city.

1.3 Further Information

Each investigatory stop is reported using a form, a sample of which can be found here. Note that the form requires the officer to recall and provide a lot of physical description of the suspect including race, height, sex, eye color, etc. as well as other metadata of the incident.

More information on Investigatory Stop Report policy may be found here.

2 Report Data

The data used in this analysis is taken from the Chicago Police Department's data website and is constituted by the details of more than 290,000 ISRs between 2018 and 2019. Each entry in the data set contains dozens of attributes that specify the officer's unique identification, the suspect's physical appearance, the location, the suspect's residence, what enforcement measures the officer took, a description of any illicit objects found, and more. In this analysis, we focus on the suspect's apparent race, whether a pat down was conducted, whether the officer was suspicious of an object in the suspect's possession, whether any weapon or contraband was found, and whether the officer took any enforcement measures such as making an arrest or issuing a citation. The data set is publicly available, and can be found here along with a complete list of variables.

Demographic estimation data from the United States Census for the city of Chicago in 2018 is also used.

Finally, venue category data from Foursquare is used to analyze the surrounding venues of each ISR. Only the category of each venue is used; other information like the name or corporate affiliation is ignored.

3 Methodology

3.1 Interaction Result Study

The first analysis we run is the Interaction Results plot which relates the investigation techniques reported in each ISR to the race of the suspects. Analysis here consisted of selecting the necessary variables we wished to plot from the ISR data set, as well as from the US Census data. After cleaning the data, it was grouped by race. Though the CPD's ISR form and the US Census specify races not included here (such as Native American), these have been condensed into the Other category since they represent $< 1\%$ of the population each.

Once arranged, the data was plotted in a stacked bar chart to easily convey the differences in representation for each category. The data are sorted in ascending order by the Chicago Population column.

3.2 ISR Vicinity Study

Next, we analyze the most common surrounding businesses around ISRs by suspects' races. Here we use only the location and racial information from the ISR data set. Location data for each ISR is condensed into a single address string, and then the `geopy` library was used to retrieve the geographic coordinates of each.

The location information in the ISR data set only provides the city block address, not the specific place where the ISR occurred. For example, an event at 1200 Michigan Ave. would be recorded as 12XX Michigan Ave, indicating only that it occurred on the 1200 block. Since we decided to take data on all the businesses within 50 meters of the ISR, we decided to just pass each address replacing the 'XX' with '00'. In cases where this was not a valid address, the entry was skipped. Since the street number likely does not have much to do with the suspects race (unless certain races are more likely to get stopped on street corners than they are in the middle of the street), this does not represent a bias in our analysis. Moreover, businesses on one end of a block are often similar to those on another end, so even if the ISR occurred at address number 399 and our analysis records it as 300, the businesses retrieved will still be representative of the location.

Since retrieving geographic coordinates and nearby businesses is a time intensive process, we restricted the ISR data set of more than 290,000 entries to a random sample of 3% the size.

Finally, categories of nearby businesses were retrieved from Foursquare, then data were grouped by race and plotted on a multiple horizontal bar chart. Since the size of this sample is significantly smaller than the full set and the Other category of race is very small, it was omitted from this chart.

3.3 Ubiquity of Racial Bias in CPD Study

Finally, using the unique (but anonymous) officer identification number provided with each entry in the ISR data set, reports were grouped by the filing officer, the race of the suspect, and the enforcement method the officer took. The simplest of the studies conducted here, these results were plotted together on a single scatter plot.

For the simplicity of reading the graph, just two races are compared at a time on a two-dimensional scatter plot.

4 Results

4.1 Interaction Result Study

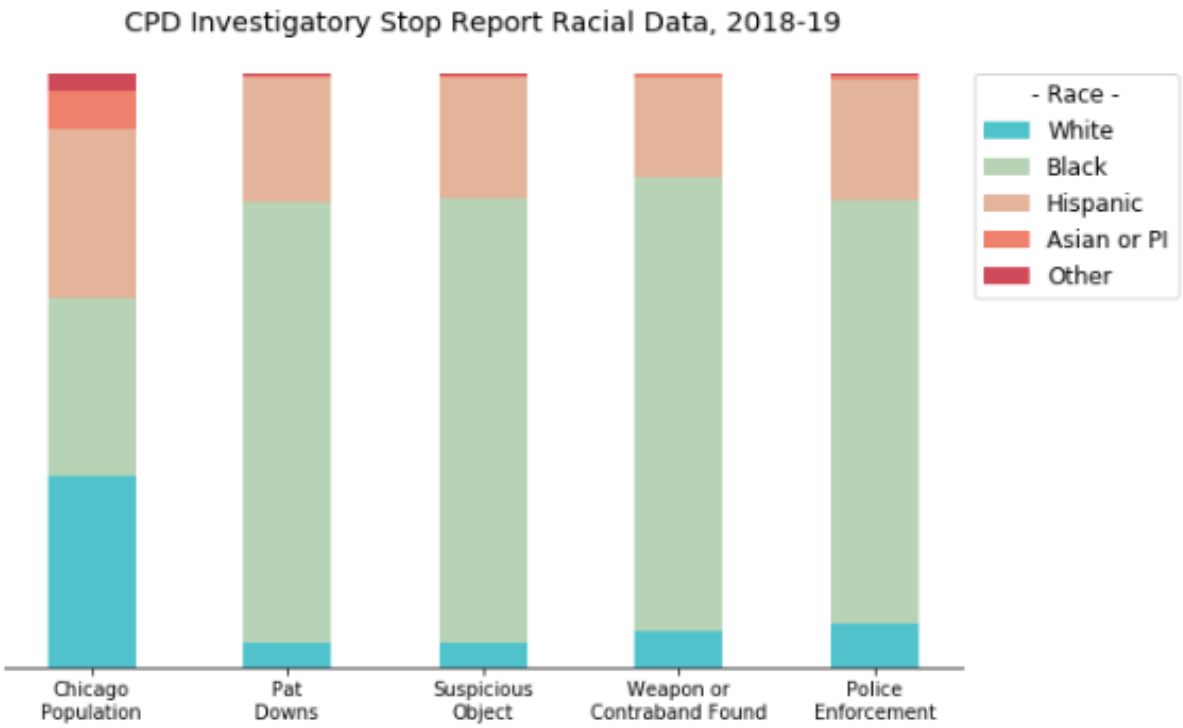

Figure 1: Plot of US Census data and 290,000 entries of ISR data from CPD. The first column is drawn from US Census data, while the latter four are from CPD ISRs. Races are ordered by proportion of the Chicago population. It should be noted that the latter four columns are based off of officers' perception of suspects' races and may not always be accurate.

Figure 2: This plot depicts the relative frequency of venue catagories near which police investigatory stops occur in Chicago, arranged by the race of suspects. Bars indicate the percent of total ISRs filed against suspects of each race. For example, about 3.9% of ISRs involving Black suspects occur near a sandwich place. This plot is based on approximately 800 events.

4.2 Ubiquity of Racial Bias in CPD Study

In this section, we observe the frequency of different police enforcement measures taken after investigatory stops. On the ISR form, an officer has the option to select "None," indicating that the stop ended with no legal action taken, "Arrest," indicating the supect was arrested, "PSC," (Personal Service Citation), a standard ticket, "ANOV," (Administrative Notice of Ordinance Violation), a slightly less severe ticket, and "Other," which is left unspecified, but is distinct from None.

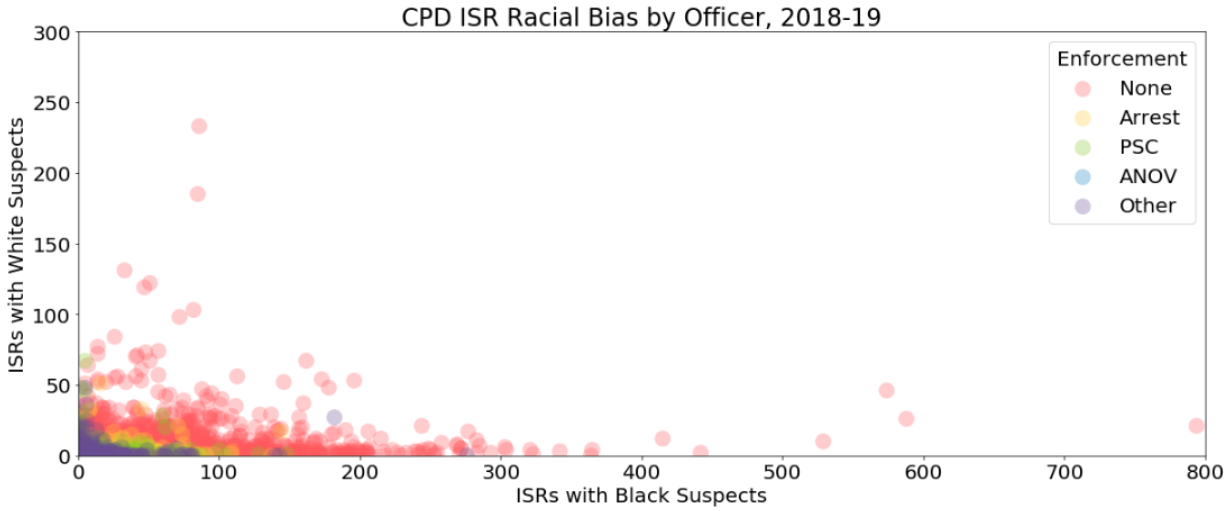

Figure 3: Quantitative results of enforcement actions taken by each officer of the CPD against Black and white suspects.

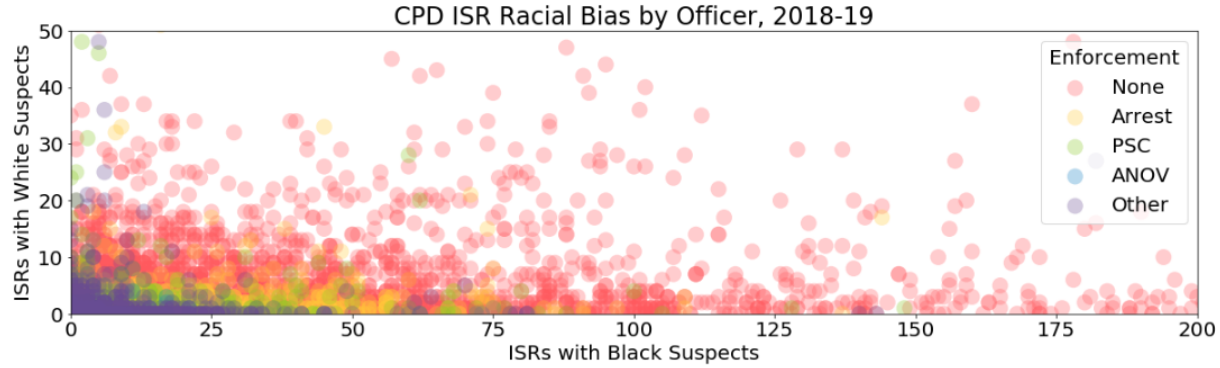

Figure 4: Same as Figure 3, but enlarged around the origin to show detail.

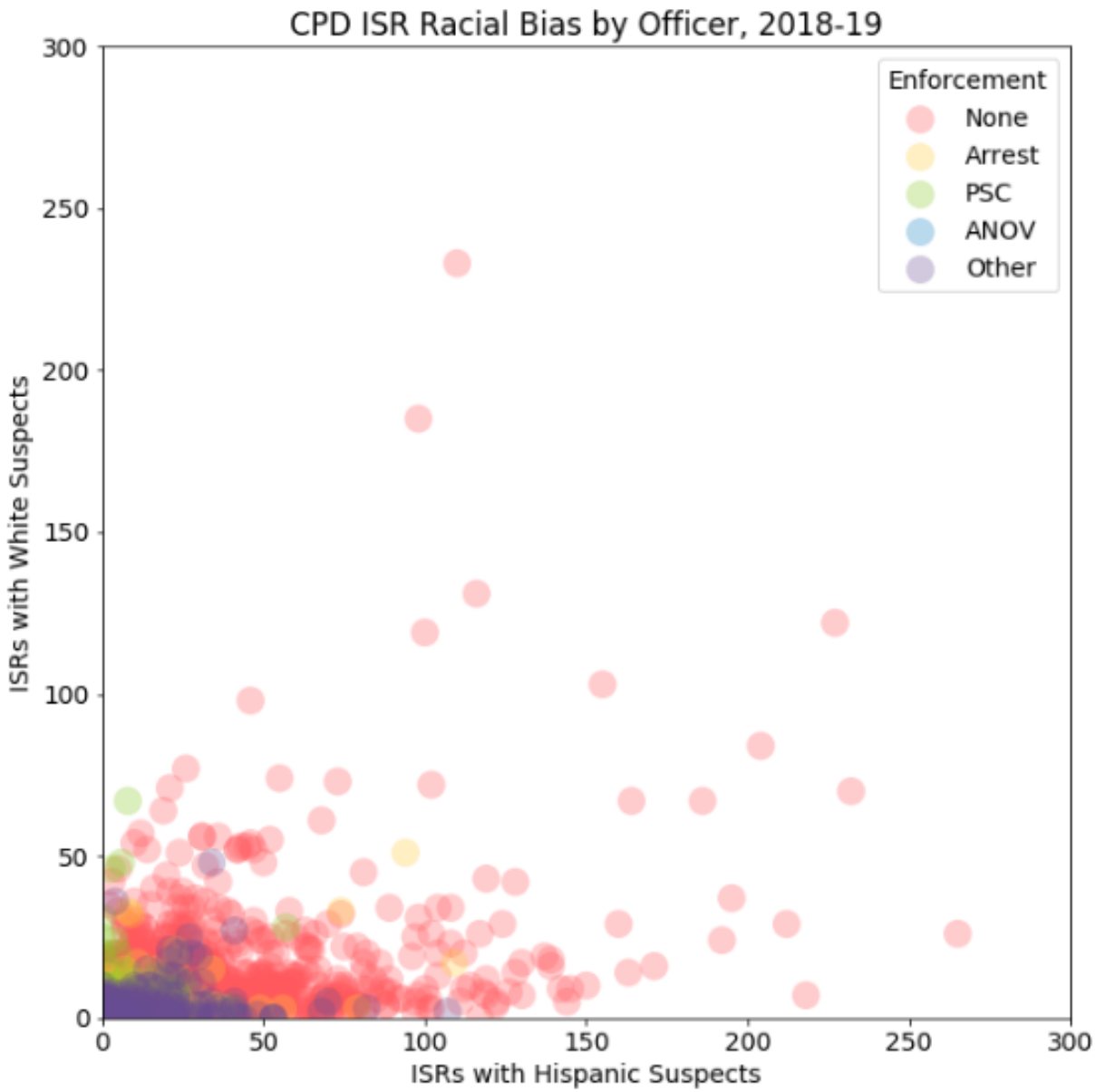

Figure 5: Same as Figure 3, but comparing results of Latinx and White ISRs.

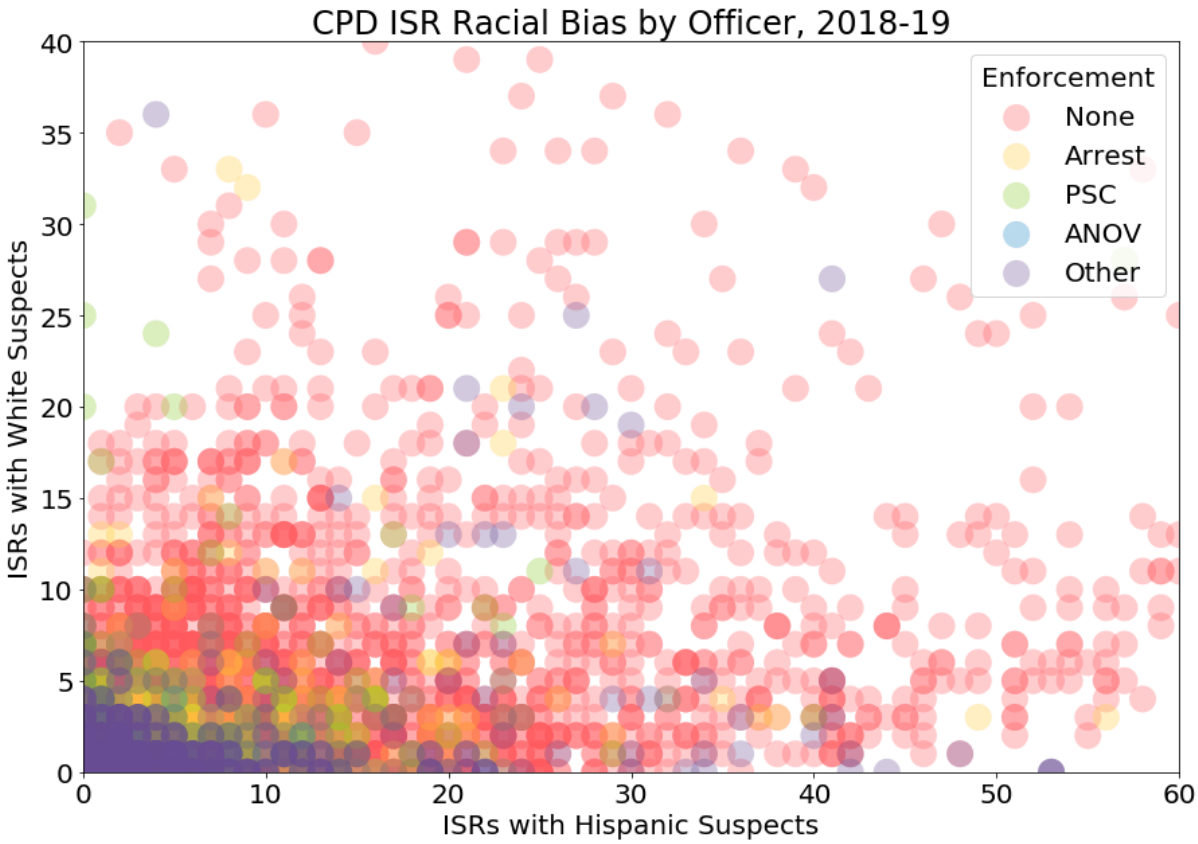

Figure 6: Same as Figure 5, but enlarged around the origin to show detail.

5 Discussion

5.1 Interaction Result Study

From *Fig. 1*, it is readily apparent that the rate of invasive investigative techniques used against Black Chicagoans is vastly steeper than against any other race. Indeed, we see that while every other race category experiences proportionally less of these invasive measures taken against them than their demographic percentage would indicate, Black citizens experience proportionately more.

The trends between columns reveal something more, though. If we take the Police Enforcement column to indicate the rate actual crimes are committed (which in itself is a large assumption considering the racial bias in police enforcement that we discuss in *Figs. 3-5*), we notice that the rate white (Black) people face enforcement is greater (less) than the rate at which they receive pat downs. From this we conclude that police officers are generally over-suspicious of Black people and, moreover, under-suspicious of white people. So, not only is the policing system disproportionately harming Black people, but it is also specifically aiding white people.

Though not to the same extreme as with Black people, we notice that though white and Latinx people represent similar fractions of the Chicago population (in fact, Latinx people represent 3.8% less), Latinx people face more invasive investigatory stops than white people do. Interestingly, it seems that Latinx people are the only group where the proportion of the Weapon or Contraband Found column to be significantly less than the Suspicious Object column. This perhaps indicates that as with Black people, Chicago police officers are over-suspicious of people in the Latinx community.

5.2 ISR Vicinity Study

It seems there is little surprising information to be gained from this portion of the study. Since this plot is based on a small random sample of the data set (800 events of the total 290,000), many differences in location popularity may just be a result of noise. Additionally, the most popular places on this plot likely only represent the relative popularity of those sorts of venues in Chicago. For example, it is not surprising that more investigatory stops occur near sandwich places than near cosmetic shops since sandwich places are undoubtedly more prevalent.

There seem to be two unique pieces of information to take away from this. First that people marked "Hispanic" are more likely than any other group to be stopped near a Mexican restaurant–not surprising since neighborhoods with greater Latinx populations are likely to have more Mexican restaurants.

Second, Asian and Pacific Islander Chicagoans are much more likely to be stopped near a Coffee Shop. This may be an interesting insight, but given the relatively small Asian population in Chicago (6.4% in 2018), this plot is represents about 51 ISRs involving Asian or Pacific Islander suspects. Thus, it is not unlikely that this odd result is simply due to noise.

5.3 Ubiquity of Racial Bias in CPD Study

Finally, the scatter plots of the bias ubiquity study present our most striking results. In *Figs. 3-4*, the tendency for officers to stop Black people more than white people is obvious. We see here that this bias is department-wide, not just the result of a few bad actors. Only a relative handful of officers plotted at or above the line $y = x$ with more that 50 total events, indicating that the majority do indeed show racial preferences.

What is perhaps most interesting about these charts are the data points plotted in red that indicate investigatory stops where no final police enforcement was taken. Such ISRs indicate that the reporting officer did not deem the situation severe enough to warrant legal punishment–most likely meaning the suspect was innocent. We see that the trend of these no-enforcement stops is just as heavily skewed towards Black suspects. Therefore Black suspects are not only arrested and cited more often, but also stopped much more often without an enduring reason; officers are more suspicious of Black people.

Comparing this result to *Figs. 5-6* that relate stops with Hispanic and White suspects, our results from section *5.1: Interaction Result Study* are largely confirmed. Though not to the same extent as Black people, Hispanic people also seem to be inconsequentially stopped (stops without police enforcement) at a greater rate than white people. Here the data seems to form a little closer to the line $y = x$, yet still lies mostly below. It is also worth noting that officers seem to favor using "Other" methods of enforcement much more with Hispanic people than with white people. (The "Other" category on the ISR form is left unspecified.)

6 Conclusions

The results of our study confirm what was suspected all along: that the CPD is disproportionately suspicious of Black people and stops them much more frequently than Latinx and especially white people. Considering the additional investigatory stops this extreme bias fosters, it is obvious why significantly more Black people face police enforcement than white people; since Black people are stopped far more frequently, Black criminals are much more likely to be prosecuted than white ones. Moreover, because this level of suspicion is so incredibly disproportionate, it is likely that even Black people who have not committed crimes will face prosecution for less severe offences for which a white person would be released.

Indeed, it is not only the case that Black people are treated with more suspicion than necessary and frequently stopped without enduring reason, but white people do not seem to be treated with enough suspicion and thus investigatory stops involving white people are disproportionately more likely to end in enforcement.

Though it is beyond the scope of this project, our further research may seek to include data on police funding and crime rates in Chicago. From the results of this study it seems likely that as police budget increases linearly, suspicion and thus disproportionate inconsequential stops will increase at a more than linear rate. This would be an interesting topic to investigate with predictive modeling.