Analyzing Racial Disparities in NYPD 'Level 3' Stop-and-Frisk Civilian Interactions

AJ Strauman-Scott

May 12th, 2024

Introduction

Although a federal court found the New York City Police Department's abusive stop-and-frisk practices unconstitutional over a decade ago, reform attempts have failed and the NYPD is still accused of continuing its discriminatory practices today. Right now, the NYPD can approach anyone they want and ask questions like what their name is or where they're going – what the NYPD classifies as "Level 1 or 2" stops - for *any* reason, or none at all. The person is under no obligation to respond to such requests, and according to the NYPD patrol guide, "may refuse to answer questions or walk or even run away".

These Level 1 or 2 encounters are some of the most common interactions between officers and New Yorkers. And because they don't rise to the level of a "stop-and-frisk" (considered Level 3 stops) – when officers must have "reasonable suspicion" of criminal activity – the NYPD has not been required to track or publicly report on any of these encounters.

The police are already under a federal requirement to collect and share detailed information about 'Level 3' stops, when the police have a stronger suspicion that a person has committed a crime and an officer may use "reasonable force" to stop a person, ask "accusatory" questions and frisk the person. The How Many Stops Act requires that officers now provide similar information for lesser 'Level 1' and 'Level 2' stops, including noting the reason for any inquiry, demographic details about the person being stopped and why the stop was made.

In January 2024, the New York City Council voted to override Mayor Eric Adams's veto to enact the **How Many Stops Act**, which requires police officers to record the race, age and gender of most people they approach. Critics of the bill claim the bill doesn't really combat unbiased policing. "When I was a captain, I was the commanding officer of two precincts. The 81st and the 113th. Both in African-American communities, 90% of the people who were stopped in those communities were African American. If you go to Middle Village and do the stops, 97% of the people there are Caucasian. 'So, what have you proved?'" Wilbur Chapman, a former NYPD chief of patrol and deputy commissioner, said to the NY 1 Sepctrum News in February 2024.

Research Question

Immediately prior the **How Many Stops Act** implementation (which is ongoing), what are the demographic and geospatial allocations of current 'Level 3' or 'Stop-and-Frisk' civilian interactions? Are these numbers significant enough to indicate racial or geographic bias, even before considering the Level 1 and Level 2 interaction data the How Many Stops Act will provide?

Data Sources

Numerical Data: - The NYPD Stop, Question, and Frisk database[1] is a dataset published annually by the NYPD as a Excel spreadsheet of every Level 3 stop, with full details for each - including full address,

and latitude and longitude for each stop. - 2020 Census data[2] of racial demographics by tract is available by Census tract number.

Geospatial Data: NYC Open Data provided GIS data of the NYPD police precincts[3] and the 2020 Census Tracts[3].

Analysis To answer former deputy commissioner Chapman's observation that the an analysis of racial bias in NYPD Stop-and-Frick Level 3 stops must consider the geographic location and race demographic breakdown of the neighborhood when considering statistical significance.

The proposed hypothesis is that observed counts of police level 3 stops are unusually different from what would be expected based on census racial demographic data. A χ^2 Goodness of Fit test to compare the proportions of race arrest in each Census tract determine for which census tracts this is true.

Placing Level 3 Stops in Census Tracts

The first step to determine if the racial breakdown of the NYPD Level 3 stops constitutes bias is to place each stop within a 2020 census tract area, so that the breakdown of stops by race within each census tract can be compared to the neighborhood. `GeoPandas`[5] transcoded the full address, latitude and longitude columns of the NYPD Stop-And-Frisk database into geometric object values, and merged the census tract GeoJSON file into it.

Additionally, both categorical options for 'Race' in both the NYPD data and the census data have been logically standardized so that the proportions can be compared.

Checking Conditions

There are two conditions that must be checking before performing a χ^2 test:

1. **INDEPENDENCE**. Each case that contributes to the table must be independent of all other cases.

- The census counts are absolutely independent from each other by the definition of the census.

- The NYPD Level 3 stops are assumed to be independent. It is possible that the same subject is stopped multiple times within the dataset, but each stop would be a different circumstances and most likely with different officers, and there is no way to know which (if any of the stops) include the same suspect.

2. **SAMPLE SIZE/ DISTRIBUTION**. Each scenario must have 5 expected cases.

- Both datasets are significantly large to fulfill this requirement

Statistical Significance Analysis

Each census tract had a χ^2 Goodness of Fit test performed on the proportions of race appearing in each dataset. If the test output a p-value ≥ 0.05, the proportion of each race that the NYPD engaged in Level 3 stops in that census tract was due to random sampling fluctuations. However, a p-value lower than 0.05 indicates that there is a less than 5% chance of random sampling those proportions of racial demographics in Level 3 stops, and is generally considered to mean that the variation in proportion is *not* due to random chance and there is in fact a selection bias in how NYPD officers select individuals for 'Stop and Frisk' interactions.

"'{python, python.reticulate = FALSE} # Perform X2 GOF test of proportions for all census tracts for tract in nypd.index: # extract rows at matching indices from both dataframes observed_counts = np.array(nypd.loc[nypd.index == tract].values[0]) expected_counts = np.array(censusX2.loc[censusX2.index == tract].values[0])

```
# Calculate observed and expected proportions
observed_prop = observed_counts / np.sum(observed_counts)
expected_prop = expected_counts / np.sum(expected_counts)

# Perform chi-squared test
chi2_stat, p_val = stats.chisquare(f_obs=observed_prop,
                                   f_exp=expected_prop)

# Calculate critical chi2 value
dof = len(categories) - 1
critical_chi2 = np.percentile(np.random.chisquare(dof,
                                                   size=100000), 95)
# update the list with values
results.append({ 'BoroCT2020': tract,
    'p-value': p_val,
    'chi2_statistic': chi2_stat,
    'critical_chi2_value': critical_chi2})
```

"'

About half of the census tracts have p-values of less than 0.05, indicating the variation in proportions between census results and NYPD arrests is *not* due to random chance, but instead statistical significant and influenced by some outside factor - most likely racial bias.

Of those precincts with a significant p-value, more than 90% had p-values that printed as 'zero' because the actual value was so small. The corresponding chi2 statistic values for those census tracts rendered as 'inf' and had to be recoded as a large integer for the values to map. Taken together, these extreme numerical values indicate that there were extreme differences between the expected racial breakdown of stops based on the population demographics of the area, and the race of those individuals the NYPD did stop-and-frisk.

Visualizing Bias

Mapping the census tracts with confirmed significant bias is essential to fully explore the issue. The census tracts are clustered in one neighborhood, or police precinct, the bias becomes a more localized phenomenon. Such a result would indicate that the issue is not perpetrated by the NYPD as a whole.

When mapped the results do not reveal any sort of localization. The figure above show a map of New York City, drawn with the lines of the census tract areas, and shaded to indicate the severity of the bias found in the Level 3 stops in that tract.

Racial bias in the Level 3 stop-and-frisk interactions by the New York City Police Department is undeniable and omnipresent. All areas of the city are affected, and severely.

A overlaid map of police precincts reveals what was already evident. The 'bad apples' of the police force are actually all the apples - or at least, many of them. Enough that every precinct is affected enough to have an purple, 'off the charts' census tract inside of it. And with more than 30,000 Level 3 interactions included in the data, this much widespread bias cannot be perpetrated by just a few individuals, even if they're distributed through the precincts evenly. It reflects a societal and systemic violence that is exacerbated by something in the NYPD officer experience.

Conclusion

The findings of this analysis are consistent with all previous examination of persistent issue of discriminatory policing practices within the New York City Police Department (NYPD). Despite efforts to reform stop-and-frisk procedures following a federal court ruling over a decade ago, there has been no change in racial bias and racial profiling.

Areas of Significant Racial Bias in Level 3 'Stop And Frisk' Police Interactions
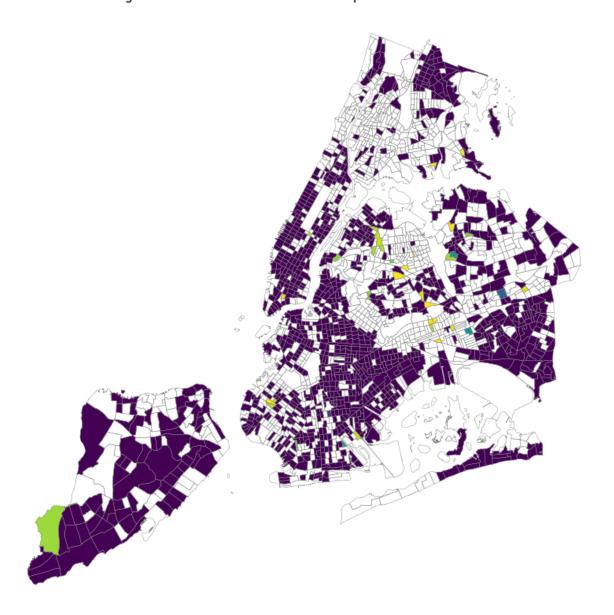

Figure 1: Census Tracts of Significant Racial Bias in 'Stop And Frisk' NYPD Interactions

Census Tracts of Significant Racial Bias in 'Stop And Frisk' NYPD Interactions (with Precinct Boundaries)

Figure 2: Census Tracts of Significant Racial Bias in 'Stop And Frisk' NYPD Interactions (with Precinct Boundaries)

The overlay of police precincts on a map vividly illustrates the widespread nature of bias within the NYPD. Rather than isolated incidents involving a few "bad apples," the data suggests systemic issues that permeate every precinct. The sheer volume of Level 3 interactions documented in the data—exceeding 30,000—underscores the magnitude of the problem. Such pervasive bias cannot be attributed solely to individual officers but reflects broader societal and systemic challenges within the NYPD.

Additional analysis on this issue could include broadening the scope of the data that is analyzed in this way. What are the geospatial dimensions of arrests? Of traffic stops? These could become a broader analysis of the impact on the NYPD on the city, and does whether whatever good the organization does do outweigh the harm caused by the undeniable profiling to the larger NYC population.

References

1. New York City Police Department. (n.d.). *NYPD Stop, Question, and Frisk database.*(2023, 2022) Retrieved from https://www.nyc.gov/site/nypd/stats/reports-analysis/stopfrisk.page

2. U.S. Census Bureau. (2020). *RACE.* Decennial Census, DEC Redistricting Data (PL 94-171), Table P1. Retrieved May 12, 2024, from https://data.census.gov/table/DECENNIALPL2020.P1

3. NYC OpenData. (2024). *Police Precincts [GIS data].* Department of City Planning (DCP). https://data.cityofnewyork.us/Public-Safety/Police-Precincts/78dh-3ptz

4. NYC OpenData. (2024). *2020 Census Tracts [GIS data].* United States Census Bureau. https://data.cityofnewyork.us/City-Government/2020-Census-Tracts-Mapped/weqx-t5xr

5. Jordahl, K., Van den Bossche, J., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Garcia Badaracco, A., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., Albert, M., Bilogur, A., Rey, S., Ren, C., Arribas-Bel, D., Wasser, L., Wolf, L. J., Journois, M., Wilson, J., Greenhall, A., Holdgraf, C., Filipe, & Leblanc, F. (2020). *geopandas/geopandas*: v0.8.1 [Software]. Zenodo. https://doi.org/10.5281/zenodo.3946761