Understanding Police Biases in New York City: 1998-1999 Stop-and-Frisk as a Case Study

Damary Gutierrez Hernandez

April 27, 2022

1 Introduction

1.1 Background Information

Stop-and-frisk is a practice that originated out of Terry v. Ohio (1968) that allowed the police to pat "down a stopped suspect's outer clothing, so long as the police officer possesses a reasonable and articulable suspicion both that criminal activity is afoot and that the person with whom the officer was dealing is armed and dangerous" (Meares 2014). While originally known as Terry stops, the term stop-and-frisk became popularized by New York City Police Department the 1990s when the practice became the city's main form of crime control. In the Supreme Court decision, the Court was worried about "the toll that unchecked crime could take if the police were not allowed more discretion to stop and frisk on less than probable cause" (Meares 2014). This decision was later overruled by Floyd v. City of New York (2013) on the basis that "racial composition of an area predicts stop patterns over and above the contribution made by crime," thereby unfairly targeting Black and Hispanic neighborhoods rather than specifically high crime areas (Meares 2014).

While the stop-and-frisk policy had good intentions, it increased racial profiling. As stated in Smith's "Explaining Police Bias," "police could consciously, but without racial animus, take race into account when making stop-related decision; from this perspective, the police may not dislike or hate minority groups, but they may nonetheless base their decisions on beliefs about group criminality and who is most likely to be involved in crime" (Smith and Alpert 2007). While police may think that these are acceptable biases to have, this type of thinking becomes an issue when a person's potential for crime becomes measured only by race, which is likely to happen unconsciously overtime. These are the pitfalls of stop-and-frisk.

In the spirit of analyses such Gelman et. al.'s "An Analysis of the New York City Police Department's 'Stop-and-Frisk' Policy in the Context of Claims of Racial Bias" and Ridgeway's "Stop-and-Frisk Is Essential… and Requires Restraint," I will use similar modeling methods for my own inferences Ridgeway (2017). The objective of this project is to explore how precinct, ethnicity[1], and assumed crime type affect the number of stops made using the stop-and-frisk policy. More specifically, I want to examine the role that race and crime type plays in who police choose to stop and frisk.

In order to determine the how the variables of interest influence the amount of stops made, I will first perform some preliminary exploratory data analysis to see if there are any differences in stops based on precinct, race, or crime type using data gathered in 1998-1999. Since the predictor deals with counts, I will then fit a Poisson regression model to the data. Next, I will test for overdispersion and correct the model if needed; this would involve fitting a negative-binomial model and comparing the fit to that of the Poisson model. I will then select a model based on which performs best on goodness-of-fit measures, and perform inference.

1.2 Data Description

The data used for this analysis is sourced from Gelman and Hill's *Data Analysis Using Regression and Multilevel/Hierarchical Models* textbook (Gelman and Hill 2006). The specific dataset I am using is titled

[1]Ethnicity and race are used interchangeably for the purposes of this analysis.

"frisk_with_noise.dat," and it consists of stop and frisk data (with noise added to protect confidentiality) for New York City over a 15-month period in 1998-1999 (Gelman, n.d.). The original data was collected by the New York City Police Department. In order to get a better idea of the amount of crime that is normal for each precinct, the full dataset also contains the number of arrests within New York City in 1997 as recorded by the Division of Criminal Justice Services of New York State (Gelman and Hill 2006). For sake of simplicity, this full dataset will be referred to as the *frisk* dataset.

In order to get a general idea of the disparity in stops by race and crime type while controlling for precinct, I first got a general sense of stop rates by ethnicity. Looking at Table 1, I can see that when looking only at race, Black or Hispanic individuals make up about 53 percent and 34 percent of all stops, respectively, while only making up about 28 percent and 26 percent of the total population of New York City. White individuals are also less represented than expected, with White people making up about 13 percent of all stops but 47 percent of the population.

Table 1: Stops by ethnicity

Ethnicity	Number of Stops	Percent of All Stops	Percent of Population
Black	69,823	53.129%	27.546%
Hispanic	44,623	33.954%	25.574%
White	16,974	12.915%	46.878%

Looking at Table 2, I can also see that the types of crimes represented in *frisk* are unequal in size. For example, weapon-related offenses make up about 44 percent of the reason for stops, while drug-related offenses only make up about 11 percent. Both violent crimes and property crimes make up about the same proportion of crimes (24 percent and 20 percent, respectively).

Table 2: Stops by crime

Crime Type	Number of Crimes	Percent of All Crimes
Violent	32,078	24.408%
Weapon	57,788	43.972%
Property	26,546	20.199%
Drug	15,008	11.419%

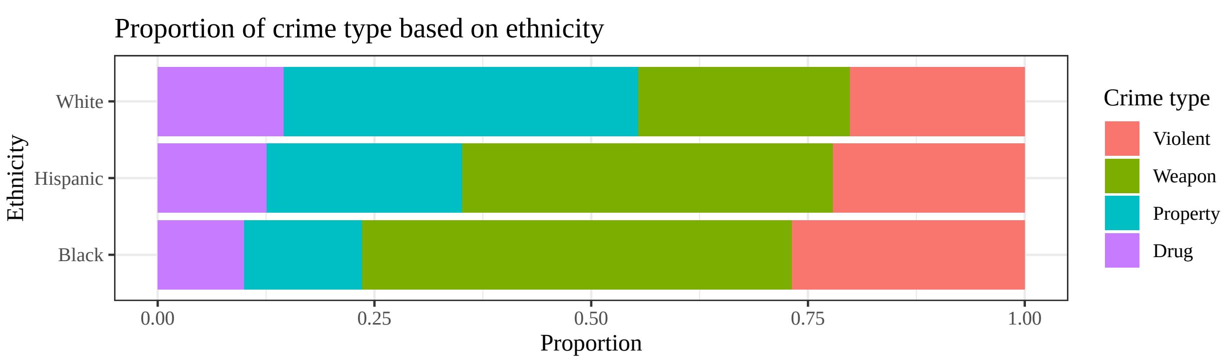


Figure 1

Taking into account both crime type and ethnicity, Figure 1 shows that different races are stopped for different reasons. Of all races, Black individuals are stopped for weapon-related crimes at a much higher rate, while White individuals are stopped the least. In contrast, White individuals are stopped for property-related crimes at a much higher rate, while Black individuals are stopped the least. It is interesting to note that all races are stopped for drug-related or violent offenses at a similar rate. Overall, this implies that both ethnicity and crime type are relevant in understanding the ways in which police bias works.

Such a comparison is a bit simplistic however, so it makes more sense to compare stop rates to a baseline. In this case, the baseline will be the number of past arrests to represent “the frequency of crimes that the police might suspect were committed by members of each group” (Gelman and Hill 2006). Referencing Table 3, I can see that the ratio of stops to previous arrests is high in the case of weapon crimes by Black or Hispanic individuals (about 2.3 and 2.2, respectively). This means that Black and Hispanic people are disproportionately likely to get stopped for weapon-related reasons as compared to what is expected.

Table 3: Ratio of stops to pasts arrests by ethnicity and crime type

Ethnicity	Crime Type	Ratio of Stops to Past Arrests
Black	Violent	0.550
	Weapon	2.294
	Property	0.413
	Drug	0.129
Hispanic	Violent	0.500
	Weapon	2.216
	Property	0.755
	Drug	0.168
White	Violent	0.352
	Weapon	1.324
	Property	0.847
	Drug	0.165

While these ratios involving ethnicity and crime type are useful, it is important to note that they are averaged over all of New York City. It could be the case that police are only targeting high-crime areas and treat everyone within each precinct equally. As such, the ratios could erroneously be picking up differences in ethnicity when the confounding variable is precinct. As such, precinct must also be included in the analysis.

1.3 Variables

The *frisk* dataset consists of six variables and records about 175,000 stops total. All but one are used in the analysis. For specific information regarding each variable, see Table 4 below.

Table 4: Variables in the *frisk* dataset

Variable	Purpose	Description
Number of stops	Response Variable	This variable counts the amount of stops made by police during the previously mentioned time period according to the New York City Police Department. It is a numerical variable consisting of positive integers and 0.
Past arrests	Exposure Variable	This variable counts the number of arrests made in 1997 according to the Division of Criminal Justice Services of New York State. It is a numerical variable consisting of positive integers.
Precint	Predictor Variable	This variable details the precinct in which the stop occurred. It is a nominal categorical variable coding precincts 1 to 75 (all the precinct in New York City) and is a factor variable in the analysis.
Ethnicity	Predictor Variable	This variable details the ethnicity of the stopped individual (as recorded by the officer). It is a nominal categorical variable where the race Black is coded as 1, Hispanic as 2, and White as 3. It is also coded as a factor for the analysis.
Crime type	Predictor Variable	This variable details the type of crime individuals were stopped for. It is a nominal categorical variable where 1 refers to violent crimes, 2 to weapon crimes, 3 to property crimes, and 4 to drug crimes. It is also coded as a factor for the analysis.
Population		This variable counts the total population per precinct. It is a numerical variable consisting of positive integers. For the purposes of the analysis, this variable is not used.

There are no missing values in *frisk*, and the data have already been wrangled. In the case of the ethnicity variable, all other ethnicities/races were excluded "because of sensitivity to ambiguities in classifications"; this consisted of only 4 percent of the original data (Gelman and Hill 2006). As for the crime type variable, the crime types follow the definitions given by the New York City Police Department. In order to be able to use the specific methods of analysis, it was necessary to replace one zero value in the past arrests variable with 1. This had almost no effect on the models, and I am able to proceed with the analysis.

1.4 Hypothesis

Following modern trends, I expect that Black and Hispanic people are disproportionately over-policed when compared to White people, and the reason for the stop (type of crime) varies among ethnic/racial groups by a significant amount. Specifcally, I assert that non-White individuals are disproportionately accused of more serious crimes.

2 Methodology

2.1 Poisson Regression

Since I am interested in modeling the variation in the number of stops–a type of count data that does not have a natural limit, I can first try fitting a Poisson regression model. In the case of the full model, the units i are ethnic groups, precincts, and crime type. The interested outcome y_i is the number of stops of members of that ethnic group in that precinct by crime type. Furthermore, the exposure u_i is the number of arrests by people of that ethnic group in that precinct by crime type in the previous year as recorded by the Department of Criminal Justice Services. Finally, the inputs are the precinct, ethnicity, and crime indexes, and the predictors X_i are the 74 precinct indicators (since 1 is the baseline), the 2 ethnicity indicators (since Black is the baseline), and the 3 crime type indicators (since violent crime is the baseline).

In order to ensure that the full model containing ethnicity, precinct, and crime is actually an improvement on simpler models, I can compare the deviance and see if the decrease is larger than would be expected if the factor were just random noise (Gelman and Hill 2006). After fitting the necessary models, I discover the following deviances listed in Table 5.

Table 5: Poisson models and their deviances

Model	Residual Deviance
(1) Intercept only: number of stops $\sim$ 1	183,986
(2) Addition of ethnicity indicators: number of stops $\sim$ ethnicity	183,300
(3) Addition of precinct indicators: number of stops $\sim$ ethnicity + precinct	141,300
(4) Addition of crime indicators: number of stops $\sim$ ethnicity + precinct + crime type	35,270

Between models 1 and 2, I see that the drop in the residual deviance is 686; this is much more than the 2 that would be expected if ethnicity had no explanatory power in the model. The drop in deviance between models 2 and 3 from 183,300 to 141,300 (42,000), is very large–much larger than the decrease of 74 that would be expected if the precinct factor were random noise. Finally, I find a 106,030 drop in deviance between models 3 and 4; this is an exceedingly large drop considering that I would only expect a drop of 3 if crime type were to be irrelevant in the model. As such, the full model including all the previously mentioned predictors is the main model of interest.

In order to continue with the chosen model, I must check for overdispersion since "Poisson regressions do not supply an independent variance parameter σ" (Gelman and Hill 2006). First, I can study the raw residuals and the standardized residuals. As shown in Figure 2 (and in Figure 6 in the Appendix), the variance of the model increases as the predicted values increase. The standardized residuals should have mean 0 and standard deviation 1, but the large amount of values outside of the 95% error bounds (the red dashed lines) imply that the variance of the standardized residuals is much greater than 1, indicating a large amount of overdispersion. In order to confirm this indication, I calculate the overdispersion ratio and its p-value (Gelman and Hill 2006). This ratio is calculated by dividing the sum of the squared standardized residuals by the number of data points minus the number of linear predictors. I find that the overdispersion ratio is 52.120, and the p-value of the overdispersion test is 1. This means that the probability is essentially zero that a random variable from a χ^2_{820} distribution would be as large as 42,739. In summary, the *frisk* data are overdispersed by a factor of about 52, which is exceedingly large and also statistically significant.

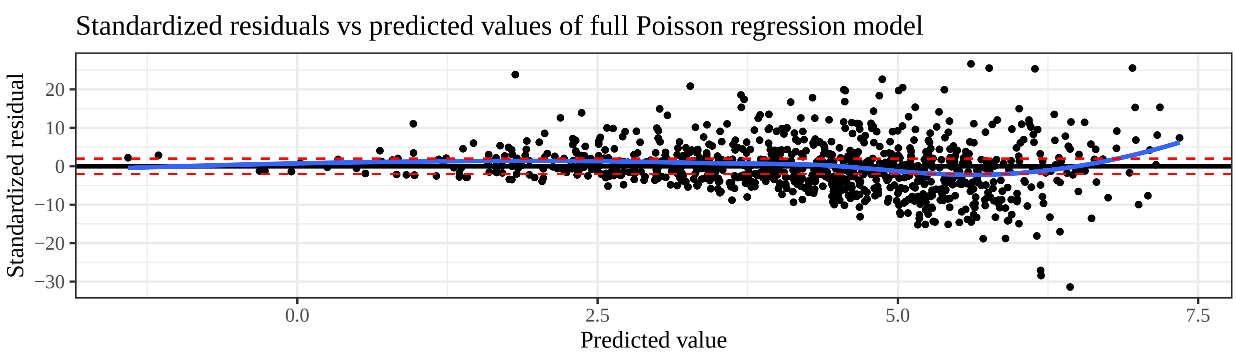


Figure 2

While it is best practice to attempt to fit another model in order to better account for overdispersion, I can fit an overdispersed-Poisson model using the quasi-poisson family. It is important to note that even though this overdispersed-Poisson regression model attempts to correct for overdispersion, inferences made using these adjustments are likely to be unacceptably inaccurate since this model is just a reparameterization of the previous model. As such, I will try another model that is based on the negative-binomial distribution and is likely to improve overdispersion.

2.2 Negative-Binomial Regression

Since the goal is to make a one-to-one comparison of the Poisson and negative-binomial models, I will use the same variables when creating the negative-binomial regression model. Note that in contrast to Poisson regression models, negative-binomial regression models assume the conditional means are not equal to the conditional variances.

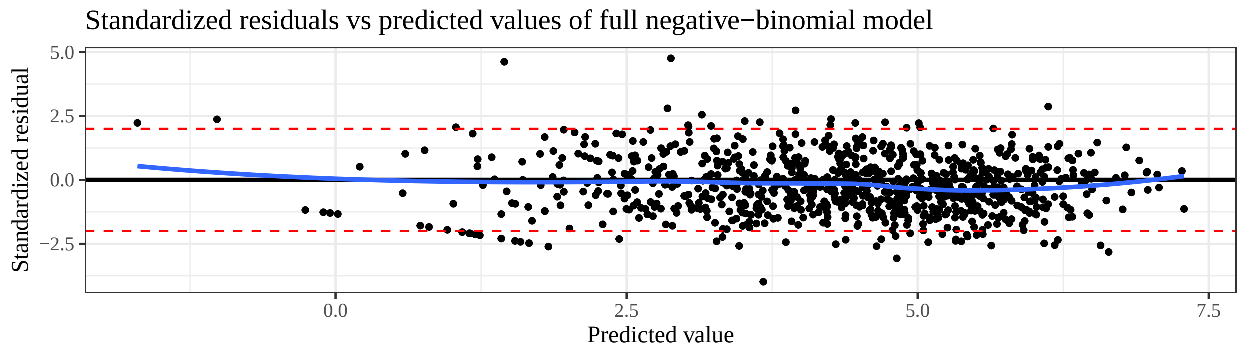


Figure 3

Looking at Figure 3 (and at Figure 7 in the Appendix), I can see that the variance of the model does not seem to increase as the predicted values increase. Furthermore, many of the values lie within the 95% error bounds, which implies that the variance of the standardized residuals is close to 1, which is as desired. In order to confirm this indication, I calculate the overdispersion ratio and its p-value. I find that the overdispersion ratio is 1.342, and the p-value of the overdispersion test is 1. As such, the probability is essentially zero that a random variable from a χ^2_{820} distribution would be as large as 1,100. In summary, the *frisk* data are overdispersed by a factor of about 1.3, which is close to 1 as desired and statistically significant.

2.3 Model Comparisons

In order to determine which model performs best, I can compare the full Poisson regression model with overdispersion and the full negative-binomial regression using a variety of metrics. One measure of goodness-of-fit is comparing deviance statistics from a chi-squared test (University, n.d.). At 820 degrees of freedom, I see that both the p-value for the Poisson model and the p-value for the negative-binomial model is 0. Since the p-value is so small, there is a statistically significant lack of fit for both models; this lack of fit is likely due to the amount of overdispersion that is present, even in the negative-binomial model. This lack of fit can be seen in the rootograms in Figure 4 and Figure 5; while there seems to be issues with over- or under-fitting in both models, the rootogram for the negative-binomial model shows an improvement (Kleiber and Zeileis 2016). I can also compare the distribution of standardized residuals vs predicted values. Using this metric, I find that the negative-binomial model performs more closely to ideal due to the smaller amount of overdispersion. Furthermore, I can also use a likelihood ratio test to compare the two models and test that the conditional means are not equal to the conditional variances (Group, n.d.). The resulting p-value is 0 with 81 degrees of freedom; this strongly suggests the negative-binomial model, estimating the dispersion parameter, is more appropriate than the Poisson model.

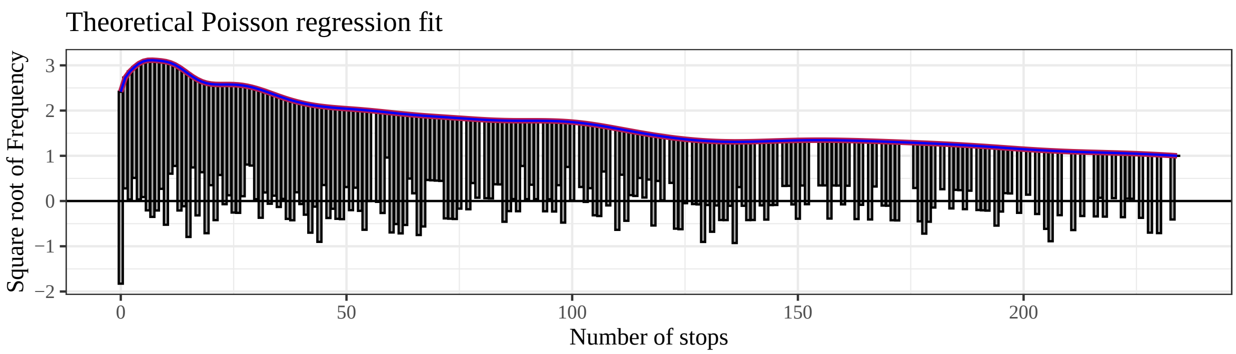


Figure 4

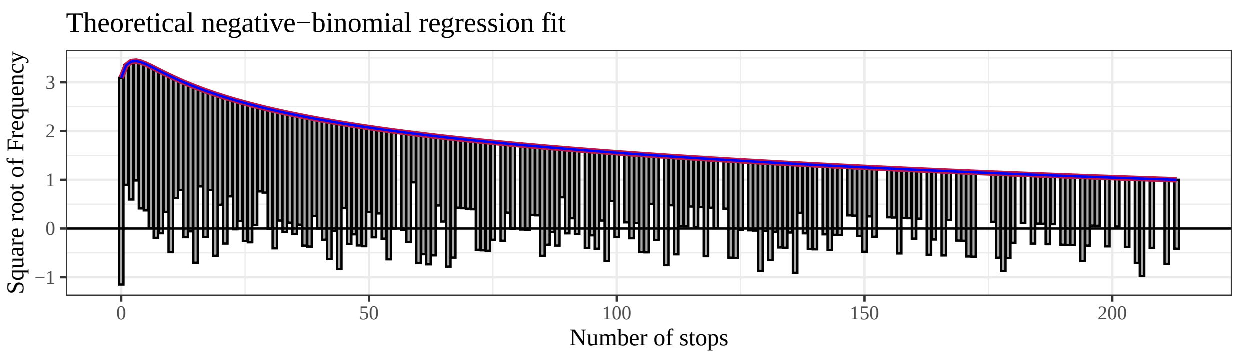


Figure 5

Therefore, I should choose the negative-binomial model over the Poisson model, and recognize that overdispersion may cause some inaccuracies. As such, the final model is as follows, with the baselines being Black, Precinct 1, and violent crimes:

$$
\begin{aligned}
log(\widehat{\text{number of stops}_i}) = &-1.276 + 0.025I(\text{ethnicity}_i = Hispanic) - 0.302I(\text{ethnicity}_i = White) \\
&- 0.281I(\text{precinct}_i = 2) + 0.783I(\text{precinct}_i = 3) + 1.212I(\text{precinct}_i = 4) \\
&+ 0.287I(\text{precinct}_i = 5) + 1.027I(\text{precinct}_i = 6) + 0.263I(\text{precinct}_i = 7) \\
&- 0.159I(\text{precinct}_i = 8) + 0.103I(\text{precinct}_i = 9) + 0.235I(\text{precinct}_i = 10) \\
&+ 0.384I(\text{precinct}_i = 11) + 1.581I(\text{precinct}_i = 12) + 0.799I(\text{precinct}_i = 13) \\
&+ 0.453I(\text{precinct}_i = 14) + 1.158I(\text{precinct}_i = 15) + 0.570I(\text{precinct}_i = 16) \\
&+ 0.192I(\text{precinct}_i = 17) + 0.162I(\text{precinct}_i = 18) - 0.055I(\text{precinct}_i = 19) \\
&+ 0.017I(\text{precinct}_i = 20) + 0.314I(\text{precinct}_i = 21) + 0.870I(\text{precinct}_i = 22) \\
&+ 0.453I(\text{precinct}_i = 23) + 0.573I(\text{precinct}_i = 24) + 0.456I(\text{precinct}_i = 25) \\
&- 0.713I(\text{precinct}_i = 26) + 1.823I(\text{precinct}_i = 27 - 1.225I(\text{precinct}_i = 28) \\
&+ 0.672I(\text{precinct}_i = 29) + 0.259I(\text{precinct}_i = 30) + 1.549I(\text{precinct}_i = 31) \\
&+ 1.201I(\text{precinct}_i = 32) + 0.664I(\text{precinct}_i = 33) + 1.143I(\text{precinct}_i = 34) \\
&+ 0.538I(\text{precinct}_i = 35) + 1.368I(\text{precinct}_i = 36) + 0.782I(\text{precinct}_i = 37) \\
&+ 1.372I(\text{precinct}_i = 38) + 0.153I(\text{precinct}_i = 39) + 1.103I(\text{precinct}_i = 40) \\
&+ 1.962I(\text{precinct}_i = 41) + 0.605I(\text{precinct}_i = 42) - 0.129I(\text{precinct}_i = 43) \\
&+ 0.523I(\text{precinct}_i = 44) + 0.268I(\text{precinct}_i = 45) + 0.045I(\text{precinct}_i = 46) \\
&+ 1.009I(\text{precinct}_i = 47) + 0.211I(\text{precinct}_i = 48) + 0.926I(\text{precinct}_i = 49) \\
&+ 0.472I(\text{precinct}_i = 50) - 0.472I(\text{precinct}_i = 51) + 0.183I(\text{precinct}_i = 52) \\
&+ 0.472I(\text{precinct}_i = 53) + 0.088I(\text{precinct}_i = 54) + 0.073I(\text{precinct}_i = 55) \\
&+ 0.467I(\text{precinct}_i = 56) + 1.107I(\text{precinct}_i = 57) + 1.256I(\text{precinct}_i = 58) \\
&+ 0.740I(\text{precinct}_i = 59) + 0.209I(\text{precinct}_i = 60) + 1.007I(\text{precinct}_i = 61) \\
&+ 0.804I(\text{precinct}_i = 62) + 0.910I(\text{precinct}_i = 63) + 1.601I(\text{precinct}_i = 64) \\
&+ 1.434I(\text{precinct}_i = 65) + 1.942I(\text{precinct}_i = 66) + 0.774I(\text{precinct}_i = 67) \\
&+ 1.740I(\text{precinct}_i = 68) + 1.620I(\text{precinct}_i = 69) - 0.028I(\text{precinct}_i = 70) \\
&+ 1.325I(\text{precinct}_i = 71) + 1.120I(\text{precinct}_i = 72) + 0.633I(\text{precinct}_i = 73) \\
&+ 0.933I(\text{precinct}_i = 74) + 0.932I(\text{precinct}_i = 75) + 1.449I(\text{crime type}_i = Weapon) \\
&+ 0.348I(\text{crime3}_i = Property) - 0.860I(\text{crime4}_i = Drug)
\end{aligned}
$$

3 Results

For a full table of results from the model, please visit https://github.com/damarygh/STA-440-Final/blob/53cbb9312114bb33a1c8d5fb9b7fa11588589b30/full_model.pdf. Otherwise, refer to the model in the methodology section.

The main goal is to make inferences about differences in stop rates based on race while taking into account the type of crime individuals are being stopped for and where those individuals are being stopped. In doing so, I hope to show the disparities in how the New York City Police Department policed its Black and Hispanic residents as compared to its White residents and how such a disparity changes depending on the crime type. I also hope to identify precincts where the disparity is particularly apparent in order to show which areas need the most improvement. While it is possible to make predictions using the model, any predictions made would not make sense using more recent data because the stop-and-frisk policy was declared unconstitutional in 2013 (Ridgeway 2017).

For the purposes on the analysis, the ethnicity baseline is Black, and all inferences are in comparison to the Division of Criminal Justice Services of New York State arrest rates which are used as a baseline. As such, I can say that Hispanic individuals are 1.025 ($e^{0.025}$) times more likely to be stopped than Black individuals, after controlling for crime type and precinct; this means that both Black and Hispanic people get stopped at about the same rate because there is only a 2.5% increase. In contrast, the model shows that White residents are 0.739 ($e^{-0.302}$) times less likely to be stopped by police when compared to Black residents of New York City, holding all else constant. This amounts to a 30.2% decrease in stops, which is an alarmingly high number. Such a statistic gives evidence to community concerns that the stop-and-frisk policy in New York City unfairly targeted Black and Hispanic residents.

When comparing the effect of crime types, it is important to note that the baseline is violent crimes. I can then say that weapon crimes are 4.259 ($e^{1.449}$) times more likely to be the reason for the stop than violent crimes, holding race and precinct constant. Property crimes are 1.416 ($e^{0.348}$) times more likely and drug crimes are 0.423 ($e^{-0.860}$) times less likely than violent crimes, after controlling for other variables[2]. This again confirms some of the differences noted in the exploratory data analysis, with weapon crimes being the most represented and drug crimes being the least represent in police stops.

Since the final model consists of only categorical variables, I can combine the effects of each category to get an idea of how different situations differ from the baseline. For example, the stop rate (per Division of Criminal Justice Services of New York State arrest) increases by a factor of 8.671 ($e^{-1.276+0.025+1.449+1.962}$) when in the 41st precinct and the individual is Hispanic and thought to be committing a weapon crime when compared to Black individuals in the 1st precinct who are thought to be committing a violent crime. The stop rate in the scenario mentioned prior increases by a similar factor when the individuals are Black. In contrast, the stop rate decreases by a factor of 0.026 ($e^{-1.276-0.302-0.860-1.225}$) when the individual is White, thought to be committing a drug crime, and is located in the 28th precinct when compared to the baseline of Black individuals in the 1st precinct who are thought to be committing a violent crime.

Using these sorts of inferences, I can recommend that individuals avoid precincts with a higher rate of stops when compared to the 1st precinct, especially if the individual is Black or Hispanic. These precincts include the 41st precinct, the 66th precinct, the 27th precinct, and other precincts with similarly large coefficients. While an individual cannot control what crime an officer thinks they are committing, by avoiding precincts with a larger stop rate factor, residents can avoid getting stopped because of their ethnicity. Residents cannot always avoid certain precincts however, especially if they work or reside in that precinct. As such, it is important to hold police accountable for targeting certain groups of people, regardless of whether or not it is on purpose.

4 Discussion

4.1 Conclusions

As stated in the literature, this analysis found that it is generally the case that Black and Hispanic people are disproportionately over-policed when compared to White people, and the reason for the stop (type of crime) varies among ethnic/racial groups by a significant amount. This analysis also found that non-White individuals are disproportionately accused of more serious crimes.

Since the final model consists of only categorical variables, I can combine the effects of each category to get an idea of how different situations differ from the baseline. For example, the stop rate (per Division of Criminal Justice Services of New York State arrest) increases by a factor of 8.671 ($e^{-1.276+0.025+1.449+1.962}$) when in the 41st precinct and the individual is Hispanic and thought to be committing a weapon crime when compared to Black individuals in the 1st precinct who are thought to be committing a violent crime. The stop rate in the scenario mentioned prior increases by a similar factor when the individuals are Black. In contrast, the stop rate decreases by a factor of 0.026 ($e^{-1.276-0.302-0.860-1.225}$) when the individual is White, thought to be committing a drug crime, and is located in the 28th precinct when compared to the baseline of Black individuals in the 1st precinct who are thought to be committing a violent crime.

Using these sorts of inferences, I can recommend that individuals avoid precincts with a higher rate of stops when compared to the 1st precinct, especially if the individual is Black or Hispanic. These precincts include the 41st precinct, the 66th precinct, the 27th precinct, and other precincts with similarly large coefficients. While an individual cannot control what crime an officer thinks they are committing, by avoiding precincts

[2] All in comparison to the Division of Criminal Justice Services of New York State arrest rates.

with a larger stop rate factor, residents can avoid getting stopped because of their ethnicity. Residents cannot always avoid certain precincts however, especially if they work or reside in that precinct. As such, it is important to hold police accountable for targeting certain groups of people, regardless of whether or not it is on purpose.

4.2 Limitations and Future Directions

While this data is from 1997-1999 New York City and stop-and-frisk was made unconstitutional in 2013, this project and its analysis can be used to examine similar policies that are still in effect today. For example, while stop-and-frisk is no longer legal, the New York City Police Department is still notorious for over-policing Black and Hispanic neighborhoods. As such, this type of analysis can be used to add credibility to those claims. Furthermore, in the case that the Supreme Cthet deliberates a similar court case that has the potential to overturn Floyd v. City of New York, this can be used to maintain the current decision. Moreover, while this analysis is limited in scope (it only consists of a few variables), it can be expanded upon with the addition of data from other sources.

Those who most stand to benefit from this project are police departments and citizens. Police departments can use this to measure how they have shifted the way they police into something more effective and less discriminatory, and citizens can use this evidence to lobby for better practices at a local level if enough improvements cannot be determined. Either way, this project–and others like it–can serve as a measure of progress in policing practices.

4.3 Summary

By using the *frisk* dataset which details police stops in New York City from 1998 to 1999, I created a negative-binomial regression model to model the amount of stops based on ethnicity, crime type, and precinct while taking into account arrests made the previous year. Recognizing that overdispersion may slightly skew theresults, I can conclude that the type of crime an individual is being stopped for, the individual's race, and where the stop occurred all affect the amount of stops police make–sometimes by striking amounts.

5 References

Gelman, Andrew. n.d. "Stop and Frisk Data." http://www.stat.columbia.edu/~gelman/arm/examples/police/frisk_with_noise.dat.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. 2007. "An Analysis of the New York City Police Department's 'Stop-and-Frisk' Policy in the Context of Claims of Racial Bias." *Journal of the American Statistical Association* 102 (479): 813–23. https://doi.org/10.1198/016214506000001040.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. 1st edition. Cambridge University Press.

Group, UCLA Statistical Consulting. n.d. "Negative Binomial Regression | R Data Analysis Examples." https://stats.idre.ucla.edu/r/dae/negative-binomial-regression/.

Kleiber, Christian, and Achim Zeileis. 2016. "Visualizing Count Data Regressions Using Rootograms." *The American Statistician* 70 (3): 296–303. https://doi.org/10.1080/00031305.2016.1173590.

Meares, Tracey L. 2014. "The Law and Social Science of Stop and Frisk." *Annual Review of Law and Social Science* 10 (1): 335–52. https://doi.org/10.1146/annurev-lawsocsci-102612-134043.

Ridgeway, Greg. 2017. "Stop-and-Frisk Is Essential … and Requires Restraint." *Journal of Policy Analysis and Management* 36 (3): 683–89. https://doi.org/10.1002/pam.21990.

Smith, Michael R., and Geoffrey P. Alpert. 2007. "Explaining Police Bias: A Theory of Social Conditioning and Illusory Correlation." *Criminal Justice and Behavior* 34 (10): 1262–83. https://doi.org/10.1177/0093854807304484.

University, Pennsylvania State. n.d. "12.3 - Poisson Regression | STAT 462." STAT 462 | Applied Regression Analysis. https://online.stat.psu.edu/stat462/node/209/.

6 Appendix

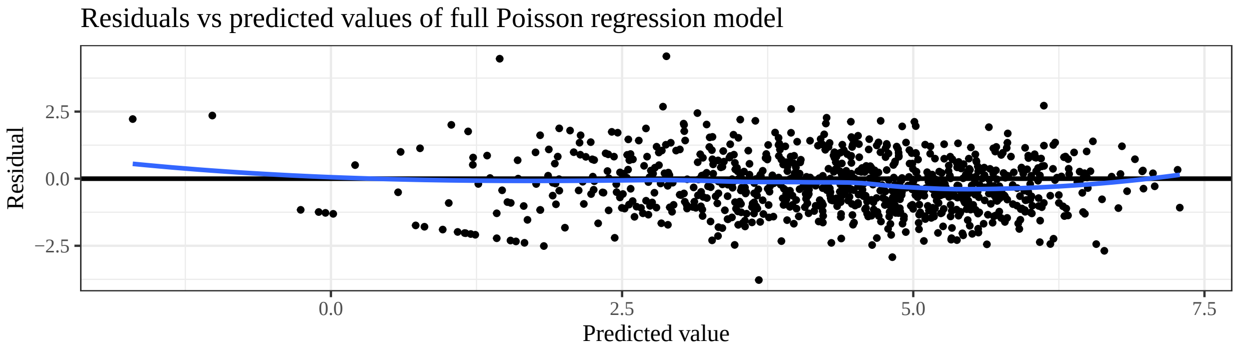


Figure 6

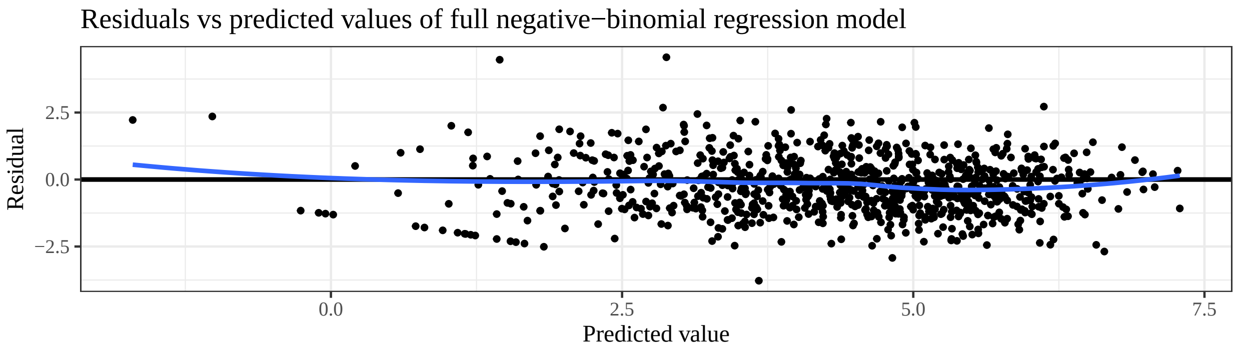
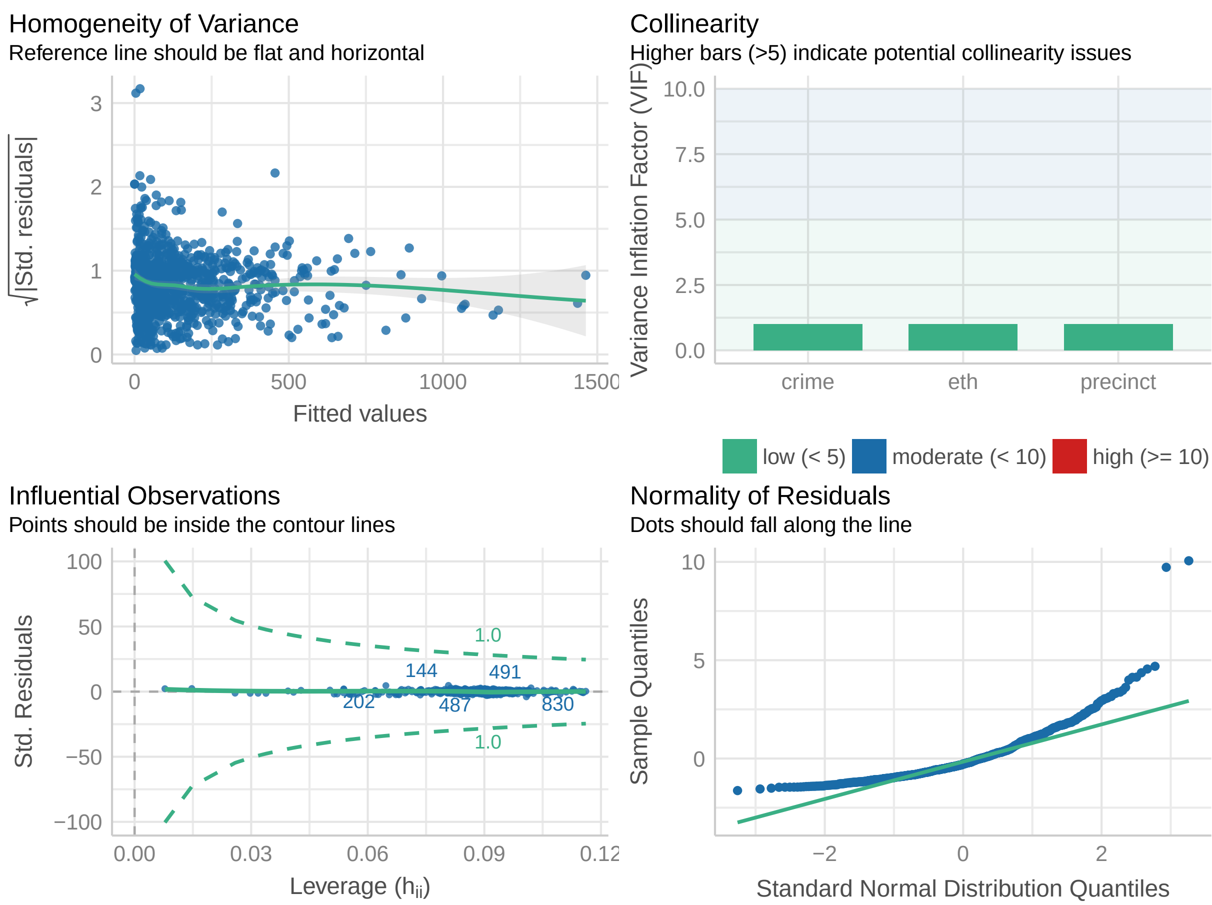


Figure 7



Homogeneity of Variance
Reference line should be flat and horizontal
√|Std. residuals|
Fitted values
Collinearity
Higher bars (>5) indicate potential collinearity issues
Variance Inflation Factor (VIF)
crime
eth
precinct
low (< 5)
moderate (< 10)
high (>= 10)
Influential Observations
Points should be inside the contour lines
Std. Residuals
Leverage (h_{ii})
Normality of Residuals
Dots should fall along the line
Sample Quantiles
Standard Normal Distribution Quantiles

Figure 8