

Spring 5-19-2016

Stop and Frisk Redux: Analysis of Racial Bias in New York City

Abraham Gutman
CUNY Hunter College

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: http://academicworks.cuny.edu/hc_sas_etds

 Part of the [Other Economics Commons](#), and the [Political Economy Commons](#)

Recommended Citation

Gutman, Abraham, "Stop and Frisk Redux: Analysis of Racial Bias in New York City" (2016). *CUNY Academic Works*.
http://academicworks.cuny.edu/hc_sas_etds/45

This Thesis is brought to you for free and open access by the Hunter College at CUNY Academic Works. It has been accepted for inclusion in School of Arts & Sciences Theses by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

Stop and Frisk Redux: Analysis of Racial Bias in New York City

by

Abraham Gutman

Submitted in partial fulfillment
of the requirements for the degree of
Master of Arts in Economics
Hunter College, City University of New York

2016

Thesis sponsor:

May 17, 2016
Date

Matthew Baker
First Reader

May 17, 2016
Date

Randall Filer
Second Reader

Abstract

With the rise of violent crime in the 1990s the New York City Police Department (NYPD) turned to proactive policing practices. Perhaps the most known of these practices is Stop, Question, and Frisk, known as Stop and Frisk. The practice had a disparate impact on the black and Hispanic community of New York City as the two groups represent more than 80 percent of all stops. Some academics have argued that sheer amount of stops does not reflect bias. Using a hit-rate analysis it has been shown that black and Hispanic pedestrians are arrested in a rate similar to the one white pedestrians are arrested at, and hence no racial bias. In this paper I challenge that conclusion and argue that there is heterogeneity among stops that masks lower arrest rates of black and Hispanic pedestrians. When the sample of all stops is restricted to only stops that involved a frisk or only stops that were required to be reported, black and Hispanic pedestrians are arrested in a lower rate that is both statistically and economically significant. This results leads to the conclusion that Stop and Frisk is an uneven policing practice that carries racial bias.

Keywords: police, hit-rate analysis, racial bias, stop and frisk, multinomial logit, arrest maximization, statistical discrimination, racial prejudice.

“Convictions on all great matters of human interest one must have to a greater or less degree, and they will enter to some extent into the most cold-blooded scientific research as a distributing factor. Nevertheless here are social problems before us demanding careful study, questions awaiting satisfactory answers. We must study, we must investigate, we must attempt to solve ; and the utmost that the world can demand is, not the lack of human interest and moral conviction, but rather the heart-quality of fairness, and an earnest desire for the truth despite its possible unpleasantness”

-W.E.B Du Bois, *The Philadelphia Negro*, 1899

Acknowledgments

I would like to thank Professor Matthew Baker and Professor Randall Filer for sponsoring this paper. I would also like to thank Professor Partha Deb for his mentorship and dedicated advising for the graduate students in the economics department. I would like to acknowledge my student cohort, particularly Andrew Bean, Colin Sharpe, Eden Volkov, and Omari Rahman, for their help and support and for proving that spillover effects are real and important. Finally, I would like to thank my wife and my life's editor-in-chief, Dr. Sarah Gutman, and my parents who gave me the support to pursue liberal arts education in the United States.

Contents

- 1 Introduction** **6**

- 2 Literature Review** **8**
 - 2.1 Race and Policing: Historical Context 8
 - 2.2 Investigation of Racial Bias Using Police Stop Data 11

- 3 Data** **13**

- 4 Empirical Strategy** **15**

- 5 Results** **16**

- 6 Discussion** **19**
 - 6.1 Discussion of Analysis Results 19
 - 6.2 Challenging the Hit-Rate Maximization Assumption 21

- 7 Conclusion** **25**

- References** **26**

- Appendix** **29**

List of Figures

1	Results of the Arrest Hit-Rate Analysis by Stop Characteristic	18
---	--	----

List of Tables

1	Independent Variables	29
2	Dependent Variable	30
3	All Recorded Stops	31
4	Stops with Frisk	32
5	UF-250 Required	33

1 Introduction

In the early 1990s, violent crime in New York City reached an all-time peak. In response to the record high murder rate of 2,245 in the year 1990, the New York Police Department (NYPD) admitted that “whatever we are doing to reduce violent – especially handgun related – crimes is not working.” Local press reported that New Yorkers were “afraid to wear jewelry in public”, and “some citizens reported sprinting to subway exits when train doors opened to avoid victimization” (Bellin, 2014, p. 1503). The NYPD decided to implement a proactive approach to policing and developed the practice of Stop, Question, and Frisk (Stop and Frisk), or “the lawful practice of temporarily detaining, questioning, and, at times, searching civilians on the street” (Spitzer, 1999, p. iv). By the late 1990s, Stop and Frisk became a common practice in New York (Bellin, 2014).

The legal basis for Stop and Frisk was determined in a 1968 United States Federal Supreme Court case. In the case of *Terry v Ohio*, the court ruled that police officers are allowed to stop and frisk, in a reasonable manner, upon reasonable suspicion, “where a police officer observes unusual conduct which leads him reasonably to conclude in light of his experience that criminality may be afoot and that the persons with whom he is dealing may be armed and presently dangerous” (Inbau and Thompson, 1999, p. 333).

In 1999, the New York State Attorney General Eliot Spitzer opened an investigation to assess the impact of Stop and Frisk on minority communities. The investigation assessed 175,000 UF-250 forms between January 1st, 1998 and March 31st, 1999. UF-250 forms are used by police officers to document stops, and will be discussed in detail in the data section. The report notes a disproportionate stop rate of black and Hispanic pedestrians compared to their proportion in the population. There is also an acknowledgement that precincts where blacks and Hispanics are the majority of the population tend to have more stops. However, there is a discussion of the correlation between areas that are predominantly populated by minority communities and higher crime rates, one potential explanation for the discrepancy. When comparing the rationale for a stop, black and Hispanic pedestrians are not more likely

to be stopped without factual bases for reasonable suspicion (Spitzer, 1999).

Organizations such as the New York Civil Liberty Union and the Center for Constitutional Justice have argued that Stop and Frisk is a racially biased program. On April 16, 2008 the Center of Constitutional Justice filed a class action suit against the City of New York and the NYPD with the goal to “hold the NYPD accountable for its illegal policing practices.” This case is known as *Floyd, et al v City of New York* (CCR, 2008). In 2013, Judge Shira Scheindlin of the United States Court of Appeals for the Second Circuit ruled that, “officers are conducting stops in a racially discriminatory manner” (New York Times, 2013). Following the ruling, Stop and Frisk underwent internal revisions but continued to play a major role in New York City policing.

Although black and Hispanic pedestrians are stopped at a rate that is higher than their proportion in the population, whether or not Stop and Frisk is a racially biased program continues to be debated. In this paper, I use police records to examine the claim that the program is racially biased. In the second section, I provide historical context on race and police, and review the current literature on racial bias in Stop and Frisk and on the analysis of police stops. In the third section, I review and summarize the data from the 2006-2012 UF-250 forms. I provide summary statistics and discuss the types of stops that were reported. In the fourth section, I lay out the empirical strategy and reasoning behind estimating a multinomial logistic regression model with the interpretation of the coefficients as relative risk ratios and restricting the sample in the various specifications. In the fifth section, I provide results of the empirical model. In the sixth section, I discuss the heterogeneity in arrest rates between black, Hispanic, and white pedestrians in different types of “hits” by stops and pedestrian features. I also challenge the arrest maximizing assumption and provide two alternative assumptions. I end with a brief conclusion.

2 Literature Review

2.1 Race and Policing: Historical Context

Concerns over racial bias in policing practices are not unique to Stop and Frisk, New York City, or recent years. In the seminal essay *The Fire Next Time*, James Baldwin describes his understanding of police as a child in Harlem during the 1940s, “it was absolutely clear that the police would whip you and take you in as long as they could get away with it” (Baldwin, 1962, p. 21). In *Another Country*, Baldwin describes police officers as an integral part of Harlem, “to remember the streets of Harlem, the boys on the stoops... the white policeman who taught him how to hate” (Baldwin, 1960, p.11). Malcolm X chose to start his famous autobiography with the story of a night in 1929 when two white men burned down his house, while he and his family were asleep. The family managed to escape from the flames and Malcolm X recalls that, “the white police and firemen came and stood around watching as the house burned down to the ground” (X and Haley, 1965, p. 81) The lack of trust in police that influenced Malcolm X that night in Omaha, Nebraska resembles descriptions from thirty years before. In a 1891 essay, anti-lynch activist Ida B. Wells describes the lack of trust in the policy by the black community in Memphis, Tennessee, “fundamentally men have an inherent right to defend themselves when lawful authority refuses to do it for them” (Wells, 2014). The notions of police being “arbitrary, ineffective, and unworthy of trust” by communities of color and immigrants persists to this day (Anderson, 2011, p. 27).

The lack of protection by the police is not the only grievance that is raised by communities of color in the United States about police. Another grievance is over-policing of communities of color and racial bias in the administration of justice. Malcolm X describes over-policing in his own home, “I remember that police were always dropping by our house, shoving things around, ‘just checking’ or ‘looking for a gun’” (X and Haley, 1965, p. 82). In the late 1960s and early 1970s, the Black Panther Party in Oakland, California started “police the police” and “defense patrols” campaigns in which members of the Party would follow Oakland law

enforcement officers with tape recorders and cameras to make sure the officers were adhering to police regulation and proper use of force protocols (Nelson, 2011, p. 61). In recent years, with the rise of accessibility to recording devices and social media, incidents of killing of black unarmed men by police led to a renewed debate over the need to “police the police,” and has rekindled a sentiment that “[police] have been endowed with the authority to destroy... [the black] body” (Coates, 2015, p. 11).

In response to these concerns and grievances by black scholars and the community, many emphasize the need to pay attention to the disproportionate rate of participation in crime of communities of color compared to the representation of non-whites in the population. This argument states that communities of color are over-policed because they commit crimes more often. Indeed, this line of thinking has persisted from early research on crime and race. In *The Philadelphia Negro*, one of the first empirical studies of a black community in the U.S., the author, W.E.B Du Bois, shows over-representation of black prisoners in The Eastern State Penitentiary in Philadelphia from the first day it opened in the year 1829. In the first years of the Penitentiary, 8.27 percent of the total population of Philadelphia and 29 percent of the Penitentiary population were black. Between 1835 and 1839 the black population of Philadelphia decreased to 7.39 percent, but the prison population increased to 40.5 percent (Du Bois, 1899, p. 238). At the time of Du Bois’s scholarship, the explanation to the over-representation of the black population in crime was inherent criminality of black and non-white immigrant communities. “Racial inheritance, physical and mental inferiority, barbarian and slave ancestry and culture... [were among] the most serious factors in crime statistics,” concluded University of Chicago social scientist Charles R. Henderson, framing race as a causal variable of criminality (Henderson, 1901, p. 246-247). Henderson was not the only social scientist who argued for a biological relationship between race and criminality, in fact in predominantly white universities of the time this was the mainstream view. In response to such scholarship, W.E.B Du Bois responded with a social structure explanation of crime, “crime is a phenomenon of organized social life, and is the open rebellion of an

individual against his social environment... lack of harmony with social surrounding [will be] leading to crime” (Du Bois, 1899, p. 235). It is the conditions of the communities of color that leads to high rates of participation in crime and not the other way around. Du Bois is also one of the first to describe the inherent endogenous character of crime statistics. Discussing the high representation of black prisoners Du Bois gives an important caveat, “it must be remembered that the discrimination against the Negro... he was arrested for less cause and given longer sentences than whites” (Du Bois, 1899, p. 239).

In the 21st century it seems that the problem of the color-line still exists in the administration of justice and in the criminal justice system. “No other country in the world imprisons so many of its racial or ethnic minorities... the United States imprisons a larger percentage of its black population than South Africa did at the height of Apartheid” (Alexander, 2010, p. 6). Tensions between police and communities of color are on the rise. It is of vast importance to a society to be able to keep itself safe from crime, but at the same time equal and free. Stop and Frisk became a symbol, for some, of a policing practice that perpetuates notions of non-white criminality and increases the disparity in the rates at which non-whites find themselves in contact with the criminal justice system. In the last days of the 19th century, while arguing against notions of black criminality, it was Du Bois who also described the importance of statistical investigation before concluding that racial bias exists in policing practices, “It has been charged by some Negroes that color prejudice plays some part, but there is no tangible proof of this, save perhaps the apt to be a certain presumption of guilt when a Negro is accused on the part of the police” (Du Bois, 1899, p. 249). This call of Du Bois, to investigate the existence of tangible proof of racial prejudice and the police’s presumption of guilt when met with a non-white suspect, is in many ways the motivation for this paper.

2.2 Investigation of Racial Bias Using Police Stop Data

Few academic papers have empirically explored allegations of racial bias in the NYPD's Stop and Frisk Program. Even amongst those that have, the results are conflicting. A common theme is the utilization of a hit-rate analysis, a model described by Knowles et al (2001) to test between statistical discrimination (Arrow, 1973) and racial prejudice (Becker, 1957) by looking at outcomes of law enforcement. Knowles et al (2001) describe the model as applied to police officers' decisions to stop motorists and search for illicit drugs. The main model assumptions are that "the police maximize the number of successful searches, net of the cost of searching motorists" (p. 205-206) (successful could mean either finding drugs or finding a large amount of drugs), and motorists respond to police behavior when deciding whether or not to carry drugs. In equilibrium, if the costs of searching the two groups are the same, the return yielded from the search should be the same for both groups. If one group is discriminated against then the return from searching that group should be below average. The authors tested the model on data from a section of Interstate-95 in Maryland. They analyzed 1,590 incidents of motorists who were stopped and searched between 1995 and 1999. The results showed that if a successful search was defined as finding drugs then there was no racial prejudice against black motorists. In fact, when a successful search was defined as finding large quantities of drugs then the authors found an implied bias against white motorists.

Gelman et al. (2007) addressed concerns about racial bias in Stop and Frisk, "a strategy involving aggressive stops and searches of pedestrians for a wide range of crimes," by utilizing a hit-rate analysis similar to the one performed by Knowles et al (2001). In the context of Stop and Frisk a successful stop is one that ends in an arrest, hence the police officer is arrest maximizing. The authors found that between January 1998 and March 1999, black and Hispanic pedestrians were arrested at a lower rate compared to white pedestrians. It is reasonable to assume, and as follows from economic theory, that there were diminishing marginal returns to stops, and hence in equilibrium the arrest rate among all racial groups

should be equal. In the current state it would therefore be more efficient or productive to stop more white pedestrians until the arrest rates are equalized. Gelman et al (2007) also compared crime rates of various racial groups and the groups' proportion of the population to the rate that they were stopped by the NYPD. In a multilevel model, the authors found that minorities were indeed stopped by the NYPD disproportionately.

Coviello and Persico (2013) used NYPD data of stops from 2003 to 2011 to challenge the conclusion of Gelman et al (2007). The authors estimated a linear probability model to find the effect of being black on the probability of a stop ending in an arrest. A naive hit-rate analysis lead to the same conclusion as Gelman et al., and found that whites are arrested at a higher rate than black pedestrians. However, the authors contended that it is important to recognize the heterogeneity in arrest rates across precincts. When a fixed effects model was estimated, controlling for precinct fixed effects, stops that involved black pedestrians were not more likely to end in an arrest. The authors concluded that this indicates that black pedestrians were not stopped disproportionately and challenges the assertion that Stop and Frisk is a racially biased program.

The precinct fixed effects in the Coviello and Persico (2013) hit-rate analysis model indeed significantly changed the conclusion, but allocation of police officers throughout the city and the decision to stop more in one area or another are endogenously determined. It is important to consider that this allocation may drive the conclusion of the hit-rate analysis. Evans et al (2014) use both precinct level data from the NYPD and 2010 Census data on New York City neighborhoods to estimate the Stop and Frisk rate given the neighborhood characteristics. The authors found a significant non-random spatial distribution associated with the overall and race-specific stop rates. The percentage of black and Hispanic individuals in an area drive up the rate of stops in a neighborhood, all else equal. The presence of Stop and Frisk as a program is therefore also determined by the presence of minorities.

Another question that might arise from the issue of heterogeneity in the Stop and Frisk program is whether results that show racial bias in policing are a department-wide phe-

nomenon or driven by the actions of a few officers. Ridgeway and MacDonald (2009) analyzed 500,000 stops from the year 2005 to attempt to answer this question. Using a doubly robust internal benchmarking model for the activities of police officers from the same groups, the authors showed that out of 2,756 police officers only five stopped black pedestrians in a disproportionately and significantly higher rate, and ten police officers stopped Hispanic pedestrians in disproportionately and significantly higher rate compared to their peers at the same time and place.

The question of racial bias in Stop and Frisk, or any policing program, is not only a question of equity and human rights but also a question of efficiency. Persico (2002) discussed the tension between efficient law enforcement and equal treatment under the law. The court allows the police to continue practices of racial profiling as long as they are not an excuse for racial harassment. The goal of the paper was to study the trade-off between effectiveness and fairness in policing using a stylized theoretical model. The conclusion found that more fairness in policing can increase the effectiveness of crime prevention and a fairness-efficiency trade-off doesn't necessarily exist.

Borooah (2001) additionally discussed the notion of fairness in policing. He derived a model that demonstrated that while statistical discriminations, which might not be optimal due to the community preferences and views on equity, will lead to efficient policing, “bigotry” or taste-based discrimination will lead to inefficient policing. Using data from England and Wales, Borooah showed that the higher rate of minorities stopped and searched by police is due to statistical discrimination.

3 Data

The data utilized was collected by the NYPD and is publicly available through New York City Open Data.¹ For the purpose of this analysis I used the years 2006 to 2012. In the

¹Data is available at http://www.nyc.gov/html/nypd/html/analysis_and_planning/stop_question_and_frisk_report.shtml

database, each stop completed by the NYPD as a part of the Stop and Frisk Program is recorded if the stop included an arrest, a search or a frisk, if the police applied physical force, or if the pedestrian refused to identify him or herself. Out of the sample, approximately 40% of the observations do not meet any of these criteria. Coviello and Persico (2013) addressed this issue stating, “Nevertheless, we think there is reason to trust the sample somewhat as being representative of all stops... officers may have an incentive to record stops, perhaps as a way of demonstrating productivity to their supervisors” (p. 3).

The original database includes 3,919,977 observations, with each observation representing a stop. The majority of stops are of black pedestrians (52.2 %), followed by Hispanic pedestrians (31%). White pedestrians represent 9.8% of stops. The racial distribution of stops is consistent throughout the years 2006 to 2012. The vast majority of stops involved males, and female pedestrians make up only 7% of stops. Pedestrians aged 15 to 24 are defined as youth (48%) and aged 25 to 35 are defined as young adults (26.3%). In more than half of the stops the pedestrian was frisked (53.6%). (Table 1) From all the stops, 5.9% ended in an arrest and 6.2% ended in a summons being issued (Table 2).

When reviewing the summary statistics of the data restricted to stops for which the mandatory criteria for filling the UF-250 form were not met – i.e. stops where there was no arrest, or no search or frisk, or no physical force used, or the pedestrian did not refuse to identify him or herself – there is reason to believe that the voluntary reporting is not done randomly. If the choice to report voluntarily were random, then it would be reasonable to expect that the summary statistics of the 1,583,696 stops distributions would not be affected in a significant way. However, the racial composition changes, and white pedestrians represent 12.6% of stops in the restricted sample (9.8% in the full sample). The representation of female pedestrians grows to 10.6% from 7% and the representation of youth decreases from 48% to 42.8%.

4 Empirical Strategy

The focus of the analysis is to extend the hit-rate analysis of Coviello and Persico (2013) to include more nuanced outcomes. The premise of the hit-rate analysis is that police officers are arrest maximizing so that every arrest is an equal “hit.” The argument laid out in this paper is that the hit-rate of different interventions should always be equal among races, such that:

$$P(\text{arrest}|\text{stop}, \text{race}_i) = P(\text{arrest}|\text{stop}, \text{race}_j)$$

I will address the non-statistically significant effect of black and Hispanic that was found by Coviello and Persico (2013) on arrest, and explore the possibility that the effect is due to hidden heterogeneity within hit-rates of different types of stops that on average cancel out in the full sample.

To examine this hypothesis I construct an outcome variable which equals “1” if there was no arrest and no summons issued, “2” if a summons was issued, and “3” if the pedestrian was arrested (Table 2). The statistical framework is of a multinomial logit, which estimates a set of probabilities for a given outcome, J, compared to a base outcome, and given a set of ‘k’ characteristics (the right-hand-side variables) (Greene, 2012). In all the specifications of the multinomial logistic regression the base outcome is ‘nothing’ - no arrest and no summons issued. Due to the suspected heterogeneity across precincts as warned by the existing literature, standard errors are clustered on the precinct level. The interpretation of the coefficient of the multinomial logit is of a relative risk ratio, such that:

$$\text{Relative risk for outcome J} = \exp(\beta_{ij}) \tag{1}$$

The interpretation of the relative risk becomes straightforward as a relative risk of 1 means that there is no difference between the group represented by the coefficient and the base comparison group, if the relative risk is larger than 1 the outcome is more likely for that group, and if the relative risk is smaller than 1 the outcome is less likely (Hardin et al, 2007).

The comparison group in the analysis is “stopped white pedestrians” of the subsample.

In all specifications precinct and year dummies are included to control for time and precinct fixed effects. Aside from these dummies there are no added covariates to the analysis. This is due to the nature of the hit-rate analysis as a model that “can be carried out when race is the only characteristic observed” (Knowles et al., 2001, p. 205). This avoids a discussion on the proper covariates, issues of confounding, and missing information on the stop. It is also important to note that the form is filed from the viewpoint of the stopping police officer. Although police officers are trained professionals, there is a level of subjectivity to many variables (for example the “body-build” variable, which is available in the dataset). The difference among the specifications is the sample of the data set that is used. In each set of specifications the sample is restricted to account for certain stop and pedestrian characteristics.

I estimate three sets of regressions. In each one of the sets the specifications are the same multinomial logit on the full subsample, a restricted subsample of only males, only youth, only young adults, and of only stops in which the police did not wear a uniform. The first set of regressions is estimated on the full sample. The second set of regressions is estimated on a subsample of only stops in which the pedestrian was frisked. The third set of regressions is estimated on a subsample of only stops in which a UF-250 was required. Due to the ambiguous definition of a stop, this approach allows for the hit-rate analysis to change as the stops escalate and exposes heterogeneity.

5 Results

Table 3 shows the result of the hit-rate analysis on the entire sample. The results confirm the Coviello and Persico (2013) results. When including both precinct and year dummies to capture the fixed effects, black pedestrians are 25 percent ($P < 0.001$) less likely to be issued a summons compared to white pedestrians. Black and Hispanic pedestrians are 6.4

percent and 5.7 percent ($P < 0.05$) more likely to be arrested, respectively. When the sample is restricted to males only, there is no major changes in the results. Hispanic pedestrians become 7.9 percent ($P < 0.05$) less likely to be given a summons and the effect of being black on arrest becomes statistically insignificant. When the sample is restricted to youth the likelihood of not being issued a summons increases and the coefficient on Hispanic becomes more significant ($P < 0.001$). Black and Hispanic youth are 11.4 percent ($P < 0.01$) and 13.5 percent ($P < 0.001$) more likely to be arrested. Black young adults, however, are 8.8 percent ($P < 0.05$) less likely to be arrested. When the sample is restricted to only stops in which the police officer was not wearing a uniform then being black reduces the likelihood of the stop ending in an arrest by 13 percent ($P < 0.01$).

Table 4 focuses on stops in which the pedestrian was frisked by the stopping police officer during the stop. In all subsamples, excluding young adults, black and Hispanic pedestrians are less likely to be issued a summons. The effect of Hispanic young adults is statistically insignificant. Given that they were frisked, black pedestrians and Hispanic pedestrians are 17.8 and 11.3 percent ($P < 0.001$), respectively, less likely to be arrested compared to white pedestrians. This effect becomes larger when further restricting the subsample to only males. The effect becomes 19.7 percent and 12.3 percent for black and Hispanic pedestrians respectively. This effect does not exist for Hispanic pedestrians when restricting the sample to youth. Black youth are 9.1 percent ($P < 0.05$) less likely to be arrested compared to white youth. The effect becomes much larger for young adults and for stops in which the police officer did not wear a uniform. Black and Hispanic young adults are 30.8 percent and 18.7 percent ($P < 0.001$) less likely to be arrested given that they were frisked and compared to white young adults. A similar effect, 29 and 15.5 percent ($P < 0.001$) for black and Hispanic respectively, is found when the police officer did not wear a uniform.

Table 5 shows specifications on a restricted subsample of the stops that required a UF-250 to be filled. As in the previous sets of specifications, there is a negative race effect on being issued a summons. The effect for black pedestrians ranges from 39.4 percent for

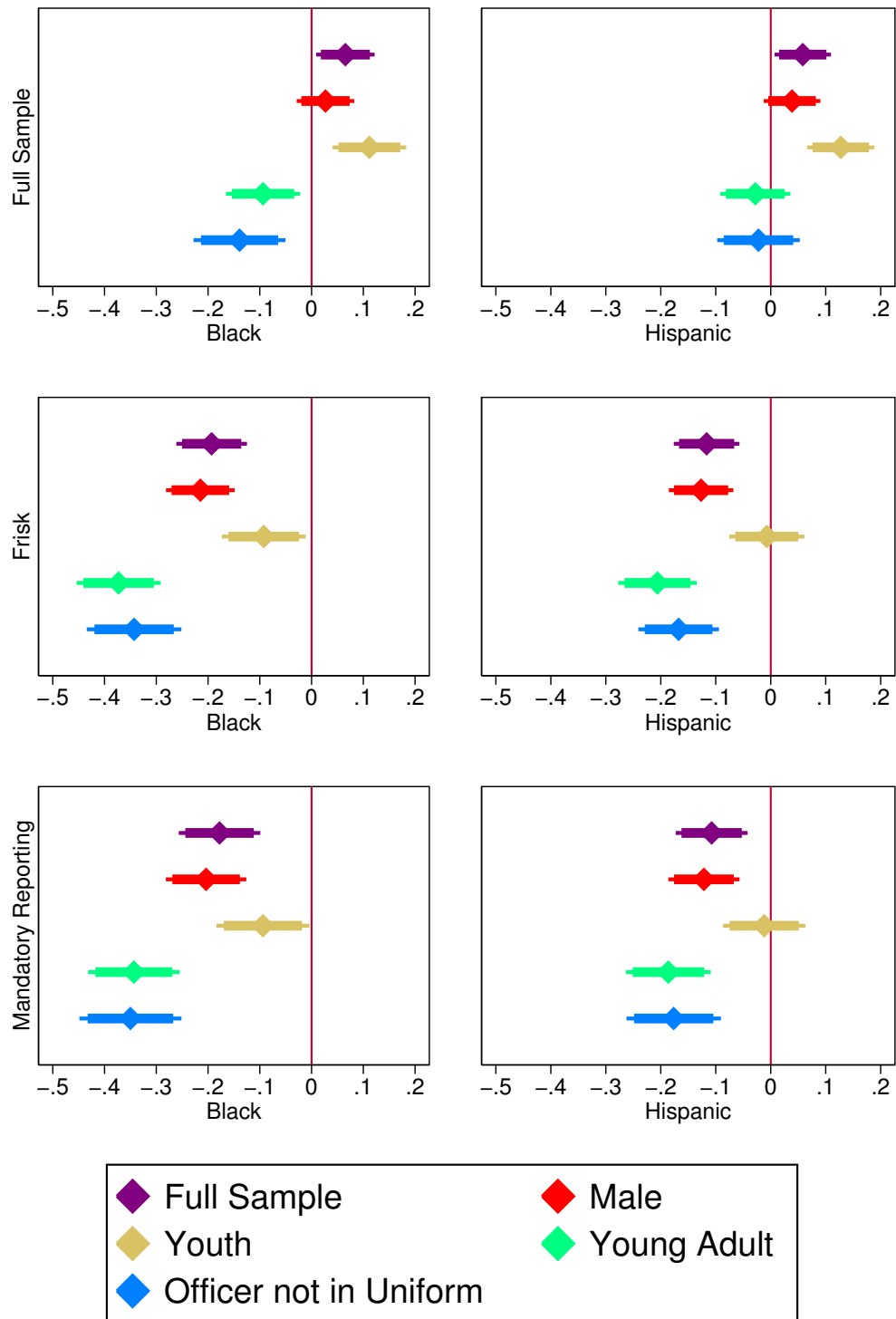


Figure 1: Results of the Arrest Hit-Rate Analysis by Stop Characteristic

young adults to 44.9 percent when the officer was not wearing a uniform ($P < 0.001$). The effect for Hispanic pedestrians ranges from 15.7 percent for young adults ($P < 0.01$) to 24.3 percent ($P < 0.001$) when the stopping officer was not wearing a uniform. In all specifications black pedestrians are less likely to be arrested given that a UF-250 was required. On the full sample the effect is 16.5 percent ($P < 0.001$). The effect for black males increases 18.8 percent ($P < 0.001$). The effect for black youth is 9.3 percent ($P < 0.05$). Black young adults and black pedestrians in stops in which the police officer wasn't wearing a uniform are 28.9 percent and 29.5 percent ($P < 0.001$), respectively, less likely to be arrested compared to white pedestrians for the same subsample. The effect of being Hispanic given that a UF-250 was required is mostly consistent with the effect of being black, although the effect is smaller. On the full sample Hispanics are 10.4 percent less likely to be arrested ($P < 0.001$) the effect slightly increases to 11.7 ($P < 0.001$) when the sample is restricted to males. There is no statistically significant effect when the subsample is of youth, but Hispanic young adults are 17 percent ($P < 0.001$) less likely to be arrested compared to white young adults given that a UF-250 was required. Lastly, Hispanic pedestrians are 16.3 ($P < 0.001$) percent less likely to be arrested when the police officer was not wearing a uniform.

6 Discussion

6.1 Discussion of Analysis Results

The results of the hit-rate analysis show heterogeneity in arrest rates among different types of stops. Although the analysis does not have the ability to determine intent to discriminate, it does show that Stop and Frisk is an uneven program that leads to different policing for different populations in New York City. Under the assumption that police officers in the Stop and Frisk program are arrest maximizing, a hit-rate analysis of the full sample might lead to the conclusion that there is either no racial bias or that the program is slightly biased against white pedestrians. However, some specifications lead to different and concerning results.

Throughout the analysis black and Hispanic individuals are less likely to be issued a summons. It is hard to know what to make of that result in regard to racial bias. One possible explanation is the fact that blacks and Hispanics constitute together 85 percent of those who refuse to show ID, and thus cannot be issued a summons. It could be interesting to see if in the years following the popular New York City ID Card the rate of summons amongst black and Hispanic pedestrians will increase due of their ability to provide proper identification.

The group that appears to be most impacted by bias in the Stop and Frisk program are black and Hispanic young adults. In all subsamples, excluding the full sample for Hispanic young adults, black and Hispanic young adults are arrested in a rate lower than white young adults. The difference between white and black young adults reaches 30.8 percent and between Hispanic young adults reaches 18.7 percent. The theory of decreasing marginal returns to arrests suggests that the NYPD is over-stopping minorities aged 25 to 35 in New York City. Another consistent result across the specifications is the lower rate at which black and Hispanic pedestrians are arrested when the police officer is not in uniform. Further research on the possible explanation to this correlation is needed.

When allowing for escalation from stops in which there was a mere verbal interaction between the police officer and the pedestrian to stops that involved a frisk, the racial bias in the arrest rates becomes economically and statistically significant. Given that the pedestrian was frisked, black and Hispanic pedestrians are much less likely to be arrested. This effect increases when the sample is restricted to males and increases more when the sample is restricted to young adults. These results lead to the conclusion that black and Hispanic pedestrians are stopped more “severely” and are policed more heavily in the nature of the stop itself than a general hit-rate would suggest.

Perhaps the most concerning result is that when the subsample is limited to only the stops that required a UF-250 to be filled, the hit-rate analysis shows that black and Hispanic pedestrians are arrested in a both statistically and economically significant lower rate, 16.5

percent and 10.4 percent ($P < 0.001$), respectively, than white pedestrians. Coviello and Persico (2013) argue that the sample should not be restricted to only the stops that require a UF-250 because, “the problem with this strategy is that, at the time of choosing whom to stop, the officer cannot distinguish whether the stop will develop into one that has to be recorded or not” (p. 4). As shown before, the full sample and the sample of the voluntarily recorded stops differ in their characteristics and summary statistics. If indeed the decision to record the stops occurs before the stop itself, i.e. the decision is random, on a large sample the summary statistics of the demographics of the voluntarily reported sample and the mandatorily recorded sample should converge to very similar summary statistics. The difference in the samples could arise either if the police officers that over-report for some reason have a different stop pattern than officers that do not over report, or if the decision of which stop to voluntarily report is related to stop characteristics.

Racial bias in NYPD’s Stop and Frisk program is masked by stops that perhaps should have never been reported and that give the pretense of equal arrest rates between white, black, and Hispanic pedestrians. When taking out of the sample stops in which the pedestrian was not frisked or stops in which a UF-250 was voluntarily filled then only stops that had more than a verbal interaction between the police officer and the pedestrian are left. In those stops the arrest rate of black and Hispanic pedestrians is lower, demonstrating that the Stop and Frisk is racially discriminatory.

6.2 Challenging the Hit-Rate Maximization Assumption

The main model assumption behind the hit-rate analysis is that police officers’ decision to stop a pedestrian is with the purpose of maximizing an arrest function subject to the cost of the stop. Since the marginal cost to the police officer to stop pedestrians of different races is equal, the first order condition comes down to:

$$P(\text{arrest}|\text{stop}, \text{race}_i) = P(\text{arrest}|\text{stop}, \text{race}_j)$$

Under this first order condition an inequality will suggest racial prejudice that cannot be explained as statistical discrimination. In this analysis we have shown that indeed:

$$P(arrest|stop, X, race_i) < P(arrest|stop, X, race_j)$$

Where X represents stop characteristics for various characteristics. But if the decision making process is not a result of maximizing an arrest function, the argument inherently changes. In this section two alternative assumptions are presented: maximization of benefit to society from crime prevented and minimization of crime.

The first alternative assumption is that police officers are not arrest maximizing, they are attempting to prevent crime such that the benefit to society will be maximized. Under the original arrest maximizing assumption, there is no account for the type of crime prevented. Would a police officer prefer to forgo a possible opportunity to prevent two very minor non-violent offenses to possibly prevent a heinous crime with multiple casualties? If the opportunity cost, the benefit lost to society by the two minor offenses, is smaller than the benefit of preventing the heinous crime, the police officer should forgo the opportunity to prevent the petty crimes. Under the benefit maximizing assumption, the first order condition changes to be an average of the probability of the arrest weighted by the probability of the expected value of the benefit from the crime prevented to society:

$$E(B(X, race_i))P(arrest|stop, race_i, X) = E(B(X, race_j))P(arrest|stop, race_j, X)$$

Where $B(X, race_i)$ is a function of the benefit to society of the crime prevented and X is a vector of stop characteristics. The results of the hit-rate analysis then might be currently skewed, depending on the relationship between $E(B(race_i, X))$ and $E(B(race_j, X))$. Hence, although the hit-rate analysis shows that some racial groups are over stopped, if the benefit to society on average is higher from arrests that are made from that specific racial group,

the over-representation could be explained as statistical discrimination.

The estimation of any benefit function of such sort is a challenge that could be impossible to overcome. One might argue that such function exists in the form of the punishment the law sets to each crime. The higher the punishment, the higher the benefit to society if that crime is prevented. The endogenous nature of crime, law-making, and attitudes toward the perpetrators of crimes, makes this a weak and biased measure. One example for the correlation that might exist between the perpetrators of a crime and the sentencing could be found in drug laws. Many legal specialists and drug experts claim that crack-cocaine and powder-cocaine are in essence the same substance in different forms, and hence abuse of either of them would have the same negative effect on society. However, “federal law... punish crack offenders one hundred times more severely than offenses involving powder-cocaine” and “the majority of those charged with crimes involving crack at the time were black (approximately 93 percent of convicted crack offenders were black, 5 percent was white), as powder-cocaine offenders were predominantly white,” argues Michelle Alexander (2010, p. 112-113). Could the 100-to-1 ratio in punishment reflect a 100-to-1 ratio of benefit to society in crime prevented of crack-cocaine compared to powdered-cocaine, which are essentially the same substance? Probably not. Another problem with the benefit maximization assumption is the hypothesis behind the benefit function that has race as a variable. In a world where the idea of biological criminality was scientifically rejected, it seems very difficult to generate an hypothesis that would necessitate race as a determining variable in such benefit function and not only as a confounding variable to some social measure.

Another alternative assumption to the hit-rate maximization is crime minimization. Police officers do not only choose to stop pedestrians to actively prevent a crime, but also to deter future criminal activity. Dominitz and Knowles (2006) lay out a stylized model of crime-rate minimization in which the decision to attempt to commit a crime is made given information on the probability of being stopped, i.e. there is a deterrence effect. Under the hit-rate maximization assumption, deterrence is not a part of the model. The estimation of

the model proposed by Dominitz and Knowles (2006) would not be possible as, “it is extremely unlikely that it will be possible to observe the marginal deterrence effects” (Dominitz and Knowles, 2006, F375).

In March of 1982 George L. Kelling and James Q. Wilson wrote an essay titled Broken Windows that was published in *The Atlantic Monthly*. In the essay, Kelling and Wilson describe the importance of perception of order in the prevention of crime, “vandalism can occur anywhere once communal barriers — the sense of mutual regard and the obligations of civility — are lowered by actions that seem to signal that ‘no one cares.’” The Theory of Broken Windows Policing puts emphasis on deterrence through “cracking down” on petty crime, such as breaking a window, to prevent the signal that in that area crime is allowed, to prevent, “the process whereby one broken window becomes many.” (Kelling and Wilson, 1982) Is Stop and Frisk Broken Windows Policing? Ignoring the legal implications (the original *Terry* stop was permitted when crime was afoot and not to prevent future crime) the data seems to support this notion. Black and Hispanic pedestrians are less likely to be arrested given that they were frisked. The results could lead to one of two conclusions: (1) following the hit-maximizing assumption there is racial bias due to lower hit-rate, or, (2) the objective function is crime minimization and over-frisking is one of many police interventions that increases the contact between the community and the police and dispels the notion that “no one care.” The second explanation requires a renewed investigation as to why the increased contact with the community, over-policing, is primarily with black and Hispanic pedestrians, if not due to racial bias. Indeed, during the peak years of Stop and Frisk incarceration in New York City was declining. Under hit-rate maximizing this is a somewhat anomalous result. However, the crime minimizing approach rationalizes this result, “If successful, NYC Stop and Frisk’s low – level interventions minimize subsequent serious offenses (and thus incarceration), while ratcheting up the degree of non-incarcerative coercion applied to innocent civilians” (Bellin, 2014, p. 1533). Constructing the proper model to estimate disparate impact in deterrence and crime minimizing policing is an important task

for future research.

7 Conclusion

Times have changed since W.E.B Du Bois in 1899 demanded statistical investigation of the lives of communities of color in the United States. Policing issues has always been at the heart of the grievances of these communities, which is true to this day. But since Du Bois wrote *The Philadelphia Negro*, arguing for the social determinants of crime as opposed to the biological criminality of non-whites, this view has become the mainstream. Economists such as Garry Becker and Kenneth Arrow developed the notions of statistical discrimination and racial prejudice to be able to differentiate between disparate impact due to confounding variables and bigotry. With this tradition of scholarship in mind the literature about policing in general, and specifically of the practice of Stop, Question, and Frisk, in New York City, attempted to investigate the existence of racial bias.

The data available from Stop and Frisk is of the stops themselves. Without the counterfactual, those who were not stopped, the proportions of stops by race do not indicate racial bias, even though they are skewed compared to the proportions in the population. The hit-rate analysis model allows the argument of whether the over-representation is due to statistical discrimination, using post-stop outcomes. Although previous papers have shown an equal hit-rate across racial groups, the analysis of this paper suggests otherwise. When restricting the sample by stop characteristics, the hit-rate analysis shows lower arrest hit-rates for black and Hispanic pedestrians as the stop escalates, i.e. as there is more contact between the police officer and the pedestrian (frisk, mandatory reporting). Even though on average, on the full sample of stops, the hit-rates are equal across racial groups, there is a need for further investigation of the cause that leads to the increased contact and over-policing of black and Hispanic pedestrians, which cannot be explained as statistical discrimination. This results suggest that Stop and Frisk is a racially biased practice.

References

- Alexander, M. (2010). *The new jim crow : Mass incarceration in the age of colorblindness*. New York : New Press.
- Anderson, E. (2011). *The cosmopolitan canopy: Race and civility in everyday life*. W.W Norton Company, Inc.
- Arrow, K. J. (1973). Discrimination in labor markets. In (chap. The Theory of Discrimination). Princeton University Press.
- Baldwin, J. (1963). *The fire next time*. Dial Press.
- Becker, G. S. (1957). *The economics of discrimination*. Chicago University Press.
- Bellin, J. (2014, May). *The Inverse Relationship between the Constitutionality and Effectiveness of New York City 'Stop and Frisk'* (SSRN Scholarly Paper No. ID 2413935). Rochester, NY: Social Science Research Network. Retrieved 2016-03-03, from <http://papers.ssrn.com/abstract=2413935>
- Borooah, V. K. (2001, March). Racial bias in police stops and searches: an economic analysis. *European Journal of Political Economy*, 17(1), 17–37. Retrieved 2016-03-03, from <http://www.sciencedirect.com/science/article/pii/S0176268000000264> doi: 10.1016/S0176-2680(00)00026-4
- Coates, T.-N. (2015). *Between the world and me*. Spiegel Grau.
- Coviello, D., & Persico, N. (2013, February). *An Economic Analysis of Black-White Disparities in NYPD's Stop and Frisk Program* (Working Paper No. 18803). National Bureau of Economic Research. Retrieved 2016-03-03, from <http://www.nber.org/papers/w18803>
- Dominitz, J., & Knowles, J. (2006, November). Crime minimisation and racial bias: What can we learn from police search data? *The Economic Journal*, 116(515), F368-F384.
- DuBois, W. (1899). *The philadelphia negro*. University of Pennsylvania Press.
- Evans, D., Maragh, C.-L., & Porter, J. (2014, March). What Do We Know About NYC's Stop and Frisk Program?: A Spatial and Statistical Analysis. *Advances in*

- Social Sciences Research Journal*, 1(2), 130–144. Retrieved 2016-03-03, from <http://scholarpublishing.org/index.php/ASSRJ/article/view/66> doi: 10.14738/assrj.12.66
- Floyd, et al. v. City of New York, et al. (2008). Retrieved 2016-03-03, from <https://ccrjustice.org/node/1765>
- Gelman, A., Fagan, J., & Kiss, A. (2007, September). 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), 813–823. Retrieved 2016-03-03, from <http://amstat.tandfonline.com/doi/abs/10.1198/016214506000001040> doi: 10.1198/016214506000001040
- Greene, W. H. (2012). *Econometric Analysis: International Edition* (7th ed.). Pearson Education.
- Hardin, J. W., Hilbe, J. M., & Hilbe, J. (2007). *Generalized Linear Models and Extensions, Second Edition*. Stata Press.
- Henderson, C. R. (1901). *An introduction to the study of the dependent, defective, and delinquent classes*. Boston: D.C. Heath and Company.
- Inbau, F. E., & Thompson, J. R. (1999). Stop and Frisk: The Power and the Obligation of the Police. *The Journal of Criminal Law and Criminology (1973-)*, 89(4), 1445–1448. Retrieved 2016-03-03, from <http://www.jstor.org/stable/1144193> doi: 10.2307/1144193
- Kelling, G. L., & Wilson, J. Q. (1982). Broken windows. *The Atlantic*.
- Knowles, J., Persico, N., & Todd, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy*, 109(1), 203–229. Retrieved 2016-03-03, from <http://www.jstor.org/stable/10.1086/318603> doi: 10.1086/318603
- Nelson, A. (2011). *Body and soul: The black panther party and the fight against medical discrimination*. University of Minnesota Press.
- Persico, N. (2002). Racial Profiling, Fairness, and Effectiveness of Policing. *Amer-*

ican Economic Review, 92(5), 1472–1497. Retrieved 2016-03-03, from <http://www.aeaweb.org/articles.php?doi=10.1257/000282802762024593> doi: 10.1257/000282802762024593

Spitzer, E. (1999). *The new york city police department's "stop and frisk" practices* (Tech. Rep.). Office of the New York State Attorney General.

The Editorial Board. (2013, August). Racial Discrimination in Stop-and-Frisk. *The New York Times*. Retrieved 2016-03-03, from <http://www.nytimes.com/2013/08/13/opinion/racial-discrimination-in-stop-and-frisk.html>

Wells, I. B. (2014). The light of truth: Writing of an anti-lynching crusader. In (p. 33-35). Penguin Classics.

X, M., & Haley, A. (1965). *The autobiography of malcolm x*. New York: Grove Press.

Appendix

Table 1: Independent Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Full	Male	Youth	Young Adult	No Uniform	Voluntarily Recorded
black	0.522	0.531	0.523	0.504	0.510	0.502
hispanic	0.310	0.316	0.315	0.334	0.309	0.292
white	0.098	0.096	0.093	0.095	0.126	0.126
male	0.910	1.000	0.916	0.915	0.945	0.873
female	0.070	0.000	0.066	0.065	0.044	0.106
youth	0.480	0.483	1.000	0.000	0.504	0.428
yadult	0.263	0.264	0.000	1.000	0.267	0.263
frisk	0.536	0.557	0.591	0.533	0.685	0.000
uf250	0.596	0.613	0.640	0.596	0.722	0.000
<i>N</i>	3919977	3568854	1882123	1029751	1052495	1583696

Table 2: Dependent Variable

	(1)	(2)	(3)	(4)	(5)
	Full	Male	Young Adult	Youth	No Uniform
nothing	0.880	0.882	0.893	0.880	0.898
arrest	0.059	0.057	0.054	0.056	0.085
summons	0.062	0.062	0.054	0.066	0.018
<i>N</i>	3919977	3568854	1882123	1029751	1052495

Table 3: All Recorded Stops

	(1) Full	(2) Male	(3) Youth	(4) Young Adult	(5) No Uniform
summons					
black	0.750*** (0.034)	0.715*** (0.034)	0.690*** (0.034)	0.772*** (0.041)	0.679*** (0.042)
hispanic	0.956 (0.030)	0.926* (0.031)	0.891*** (0.030)	0.980 (0.042)	0.884 (0.056)
arrest					
black	1.064* (0.030)	1.023 (0.029)	1.114** (0.040)	0.912* (0.033)	0.870** (0.039)
hispanic	1.057* (0.027)	1.036 (0.027)	1.135*** (0.035)	0.972 (0.031)	0.977 (0.037)
<i>N</i>	3919975	3568852	1882123	1029750	1052493

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Stops with Frisk

	(1) Full	(2) Male	(3) Youth	(4) Young Adult	(5) No Uniform
summons					
black	0.719*** (0.033)	0.692*** (0.032)	0.701*** (0.031)	0.732*** (0.045)	0.675*** (0.043)
hispanic	0.938* (0.029)	0.910** (0.030)	0.892*** (0.027)	0.956 (0.046)	0.858** (0.050)
arrest					
black	0.822*** (0.028)	0.803*** (0.027)	0.909* (0.037)	0.692*** (0.028)	0.710*** (0.033)
hispanic	0.887*** (0.027)	0.877*** (0.026)	0.992 (0.034)	0.813*** (0.029)	0.845*** (0.031)
<i>N</i>	2102329	1989232	1112665	549298	721381

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: UF-250 Required

	(1) Full	(2) Male	(3) Youth	(4) Young Adult	(5) No Uniform
summons					
black	0.594*** (0.038)	0.571*** (0.037)	0.569*** (0.035)	0.606*** (0.043)	0.551*** (0.037)
hispanic	0.822*** (0.033)	0.797*** (0.033)	0.787*** (0.031)	0.844** (0.044)	0.757*** (0.050)
arrest					
black	0.835*** (0.033)	0.812*** (0.032)	0.907* (0.041)	0.711*** (0.032)	0.705*** (0.035)
hispanic	0.896*** (0.030)	0.883*** (0.029)	0.987 (0.037)	0.830*** (0.032)	0.837*** (0.037)
<i>N</i>	2336281	2186686	1205029	613832	759496

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$