School of Arts & Sciences Theses
Hunter College

Summer 9-1-2016

Can Noncompliant Behavior Explain Racial/Ethnic Disparities in The Use of Force by The NYPD? An Econometric Analysis of New York's Stop-and-Frisk

Omari-Khalid Rahman
CUNY Hunter College

Can Noncompliant Behavior Explain Racial/Ethnic Disparities in The Use of Force by The NYPD? An Econometric Analysis of New York's Stop-and-Frisk

by

Omari-Khalid Rahman

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:

August 2, 2016	Partha Deb
Date	First Reader
August 2, 2016	Matthew Baker
Date	Second Reader

Acknowledgements

Producing this thesis was by far the most challenging thing I've done to date. Many times I faltered, lost focus and questioned my project entirely and it was only through the encouragement and mentorship of others that I was able to see this project through to completion. Foremost, I have to express my sincere gratitude to my adviser Professor Partha Deb for his guidance, insightful comments and continuous support. I also want to thank Professor Matthew Baker who served as my second reader and a valuable resource for all things Stata related. Finally I want to thank the CUNY Pipeline program for providing financial support and a community of like-minded scholars who challenged me to rigorously develop my ideas into a policy relevant research paper.

Contents

1 Introduction

1.1 Background

The Fourth Amendment to the United States Constitution protects individuals from unreasonable searches and seizures by the government by requiring a warrant sanctioned by the judiciary based upon probable cause. In *Terry v. Ohio*, 392 U.S. 1, (1968) the United States Supreme Court ruled that the Fourth Amendment is not violated when a police officer stops, searches and detains a person without probable cause, as long as the officer has reasonable suspicion that the person may be engaging in illegal activity or a potential threat to those in the immediate vicinity. This landmark decision paved the way for the enactment of § 140.50 of the New York State Criminal Procedure Law in 1971, which granted law enforcement officers the ability to stop and potentially search any individual within their jurisdiction based on reasonable suspicion that the person is a threat to public safety.

Just over forty years later § 140.50 has become a significant controversy and public policy issue for residents and observers of New York City, due to claims of racial and ethnic biases exhibited by officers of the New York City Police Department (NYPD) operating under this statute. "Stop-and-Frisk", as it is colloquially known, has gained increasing media attention over the past few years, with the number of monthly New York Times articles (about Stop and Frisk) increasing tenfold during 2010-2012 (Persisco and Coviello, 2013). This tenfold increase coincided with a number of legal and constitutional challenges to Stop and Frisk, most notably *Floyd, et al. v. City of New York* which issued verdict in favor of the plaintiff on August 12th 2013. The court ruled that the NYPD violated the Fourth Amendment by carrying out unreasonable searches and the Fourteenth Amendment by carrying out racially motivated stops.

1.2 Counterfactuals and Causality

Statistical tests of racial bias in policing settings are notoriously difficult to perform due to the inability of the researcher to observe the counterfactual outcome. In a hypothetical laboratory setting, the researcher's ideal experiment would be to observe a police officer being faced with a decision to stop one of two equally suspicious pedestrians, one of whom would have been randomly assigned to the treatment group (i.e. a racial/ethnic minority). We expect that if this experiment was repeated an infinite number of times, an unbiased officer would stop each pedestrian in roughly equal proportion, or with *equal probability.* Only if the officer stops one pedestrian with a greater frequency than the other despite them both being equally suspicious, can we infer racial bias. In every observational dataset on policing stops, this counterfactual (the people who the officer did not stop) are never recorded. It is tempting to use local population data in an attempt to construct the counterfactual outcome, however this approach is flawed as it assumes that everyone in a given locality is equally suspicious and ignores variations in pedestrian activity that the officer uses to assess the likelihood of criminal activity. Furthermore, this approach conflates racial bias with statistical discrimination by assuming that the demographic composition of police stops should mirror the demographic composition of the locality. Drawing on the formulation proposed in Persisco and Coviello (2013), if we assume that by conducting stops the police are attempting to minimize crime then the decision problem of the police chief can be hypothesized as:

$$\min_{p_i} \sum_i \beta_i C_i(p_i) \quad s.t. \quad \sum_i p_i \leq P,$$

where $C_i(\cdot)$ is crime per capita, which is a monotonically decreasing function of p_i, the number of patrol officers conducting stops. P represents the total number of officers under the command of the police chief and β_i represents the rate or the propensity of demographic group i to engage in criminal activity, where $0 \leq \beta_i \leq 1$. If the police have reason to believe that a particular demographic group disproportionately engages in criminal activity ($\beta_i > \bar{\beta}$), then a police chief seeking to minimize crime is justified in targeting that group in an attempt to deter the most crime per unit of policing manpower. Gelman, Fagan and Kiss (2007) attempted to study this decision problem by using the arrest rates for Blacks and Hispanics as a proxy for β in order to estimate the optimal allocation of police stops across racial and ethnic groups if the police were unbiased. However, from an economic perspective this approach is incomplete; what if the crime per captia function does not behave uniformly across all groups? In other words, what if a one unit increase in police manpower does not lead to the same decrease in crime per capita for each $C_i(\cdot)$?[1] To address this concern, it is necessary to impose first-order conditions on the police chief's objective function in order to solve for the optimal allocation of police stops p^* across all demographic groups:

$$\beta_i C_i'(p_i^*) = \beta_j C_j'(p_j^*) \quad for\ all\ i, j$$

where $C_i'(\cdot)$ and $C_j'(\cdot)$ are the elasticites of crime to policing for demographic groups i and j. If these elasticites are not uniform (i.e., $C_i'(\cdot) \neq C_j'(\cdot)$), then a racially unbiased police chief may still disproportionately target a group in excess of their relative crime participation rates. Therefore, one cannot infer racial bias in policing stops without first knowing the crime elasticites of each group. It is difficult to empirically estimate crime elasticites as cross-sectional datasets only capture equilibrium levels of crime and policing, which often vary simultaneously along with other confounding factors that are not directly observed.

1.3 Disparate Impact or Disparate Treatment? The Use of Force in Police Stops

Race and gender discrimination remains a hotly contested public policy issue, which is not surprising given that the 14$^{\text{th}}$ Amendment Equal Protection Clause was established to protect Americans against discrimination. However, very rarely in these policy discussions do we distinguish between *disparate impact* and *disparate treatment*, two very distinct and legally relevant categories. The former refers to policies that are neutral in its treatment of different groups (i.e., not intentionally discriminatory) but ultimately affect groups of people differently. It can be argued that the police chief's decision to allocate police manpower unequally across demographic groups falls in the category of disparate impact as his primary objective is to minimize crime, and certain demographic groups are known to engage in crime disproportionately, consequently receiving a larger share of police stops. If one can show that police stops are being allocated in excess of crime participation and crime elasticities, then one can classify the policy as disparate treatment, since the discrimination is intentional rather than a by-product. Unfortunately, as discussed earlier this hypothesis is not easily falsifiable due to the absence of the relevant statistical information.

[1] As in Persico and Coviello (2013), $C_i(\cdot)$ is assumed to be a concave function for purposes of mathematical tractability. I concede that there is some loss in generality, but the underlying economic intuition remains unchanged.

Given econometric limitations, we cannot directly address the question of racial bias in police stops; however if we restrict the universe to only the people who have been stopped then these econometric limitations are quickly dissipated. Racial/Ethnic disparities persist even after a stop has been initiated, with minorities substantially more likely to be frisked, searched and have force used on them. This is particularly puzzling as within the universe of pedestrians who have been stopped, one would assume that they are all relatively "equally suspicious", therefore significant disparities between subpopulations should not exist. This paper seeks to study one such disparity, namely the use of force by the police. By restricting the universe and conditioning on observables, the question of disparate impact or disparate treatment becomes irrelevant as we have removed the criteria of unconditional race neutrality. Put differently, if minorities were disproportionately non-compliant then we would expect them to have a higher rate of stops where force was used–this is disparate impact. If minorities had higher rates of force in excess of their relative rates of noncompliance, then it becomes disparate treatment. If minorities have higher rates of force within the universe of non-compliant stops, then there is no disparate impact or disparate treatment, only bias.

2 Literature Review

2.1 Racial Profiling and Urban Policing

Racial profiling and urban policing have gone hand in hand since the first emancipated slaves from the war-ravaged South turned up on the streets of the urban North. Muhammad (2010) examined the genealogy of Black criminality from its inception in the Gilded Age and traced its transcendence from mere stereotypes to a full-fledged dialectic of social inquiry by the turn of the century. He argued that the tenets of this social science discourse were racist but operated under the guise of being objective by presenting statistical arguments to criminalize Blacks. Muhammad highlighted a series of studies that revealed Black criminality as the result of discriminatory policing, or in Muhammad's words "Jim Crow justice"; the police underreported white crimes and overreported Black crimes and then used disparate crime rates as justification for racial policing.

Racial disparities in the criminal justice system extend well beyond police-citizen interactions. Sampson and Lauritsen (1997) argued for the necessity of researchers to employ an analytic framework that views these disparities as a a cumulative disadvantage over the life course; the concatenation of race and poverty amplifies initial disadvantages into racialized trends and contribute to the overrepresenation of racial/ethnic minorities in all aspects of the criminal justice system. Petit and Western (2004) studied penal inequality by estimating lifetime risks of imprisonment for black and white men at different levels of educational attainment and found that on average, 3% of whites and 20% of blacks had served time in prison by their early thirties. The risks of incarceration were stratified by levels of education; at the most extreme end they found that black men with less than a high school education had a 58.9% risk of incarceration, whereas whites with a similar level of education faced only a 11.2% risk.[2] They concluded that the pervasiveness of imprisonment among young, low-skilled black men has catapulted incarceration from a tragic, unexpected occurrence to an established stage of the life course for the aforementioned demography. The literature establishes the ubiquity of racial profiling across space and time and provides a historical and theoretical lens that can shed light on the initial question of Stop-and-Frisk: why is it that in any given year Blacks and Hispanics comprise roughly 85% of police stops yet represent just over 50% of New Yorkers?

2.2 The Economic Model of Policing and Its Limitations

Knowles, Persisco and Todd (2001) analyzed vehicle search data from Maryland, testing and subsequently rejecting the hypothesis of racial bias against African Americans. They argued that simply analyzing vehicle stop rates and search rates for disparities was insufficient because if the propensity to carry contraband differed across groups, then the police are justified in stopping and searching one group more often than the other(s). Instead they proposed a hit-rate test analyzing disparities in search rates conditional on some *ex-ante* probability that a search will lead to a hit (recover contraband). To carry out the hit-rate test, Knowles et al formalized an economic model of policing behavior in the context of motor vehicle searches by assuming that the police achieved utility through successful searches, and conducted searches until the marginal cost of an additional search exceeded the marginal

[2]Other relevant demographic characteristics were held constant (e.g. income, age, etc.)

benefit. In order to derive equilibrium predictions, the model also allowed for motorists to dynamically alter their probability of carrying contraband based on the current search patterns of the police. In equilibrium, hit-rates are equal across groups because if one group had a higher hit-rate, then the police would search that group more often than the rest; consequently, that group would lower their probability of carrying contraband in response to greater police scrutiny. This lowered probability will result in less hits and the probability will continue to fall until they no longer receive disproportionate attention from the police.[3,4] The paper found that the hit-rates for African Americans and whites were equal, and also found that hit-rates were equal across a number of other driver characteristics, giving credence to the argument that the police are unbiased and primarily seeking to maximize contraband seizures.

Harcourt (2004), among others, have challenged this economic model of policing arguing that it falls short on several accounts. First, it's definition of success is problematic: the proper goal of the police is to minimize crime (and the social costs associated with it), not maximize successful searches or arrests. He further argued that these two objectives are at odds with each other under certain conditions; if the police allocate more stops towards minorities and away from whites then while the offending rate of minorities may fall, the offending rate of whites may increase.[5] Depending on certain factors, the increase in white offending rates may outweigh the decrease in minorities offending rates in an absolute sense, because whites outnumber minorities in general population. Secondly, in Harcourt's view the economic model fails to capture important social costs associated with crime itself and the social cost of carrying out discriminatory searches: a search that recovers small quantities of marijuana should not be considered equally successful as a search that recovers an illegal firearm or large quantities of more illicit substances. Furthermore, social costs of discriminatory policing are not minimized when hit-rates are equal across groups; if, for example, minorities are being searched at higher rates than whites then the social costs of conducting

[3]The level of policing resources is assumed to be fixed, a necessary condition for a unique equilibrium.

[4]In this context it is assumed that the marginal cost of conducting a search is equal across all groups. If the marginal cost of African American searches were lower than for whites, then equilibrium will not be achieved when hit-rates are equal because utility can be further maximized by searching more African Americans. While it is theoretically possible for marginal costs to differ due to taste-based discrimination (see Becker, 1957), this is often difficult to capture empirically and has been supplanted by more concrete measures of discrimination.

minority searches are higher since it requires a greater frequency in order to yield the same hit-rate.

2.3 The Use of Force By Police

Max Weber famously argued in his essay *Politics as a Vocation* (1919), the state as a sociological institution “claims the monopoly of the legitimate use of physical force within a given territory...The state is considered the sole source of the ‘right’ to use violence”. The police are the physical embodiment of the state and exist to enforce and protect its institutions; the application of coercive force is the core function of policing (Bittner 1970). The police exercise discretion in their decision to use force and a considerable amount of research has been done to identify determinants of the use of force. Organizational and psychological approaches have be employed to study this issue and have found that bureaucratic mandates, precinct/policing culture, and social psychological beliefs about racial/ethnic minorities all play a significant role in the officer’s decision to use force (Terrill, Paoline and Manning 2003; Plant and Peruche 2005).

Sociological theories have largely focused on situationally based determinants such as the socioeconomic characteristics of the suspect, the attributes of the officer and the location of the interaction. Black (1976) developed a sociological theory to explain variations in the application of law and hypothesized that people of marginalized groups will face greater punitive treatment: racial/ethnic minorities, the homeless, the poor, the (mentally or physically) disabled are all more likely to receive greater scrutiny from the police. Terrill and Mastrofski (2002) tested this hypothesis using data collected from an observational study on police stops in Indianapolis, Indiana and St. Petersburg, Florida. In this study they examined police use of nonlethal force in an attempt to gain a better understanding of why the police resort to force. The dependent variable was an ordered set of choices (of increasing severity) ranging from verbal commands/threats to physical force . They found that the police were not more coercive towards disrespectful (or non-compliant) suspects, and that being male, nonwhite, poor and younger all increased the probability of force being used. They did find however, that the level of force used was proportional to the suspect’s resistance, with officers choosing

[5] Again, policing resources are assumed to be fixed.

a more severe level of force when the non-compliant behavior became a risk to the officer's safety. Using the same data but in a different paper, Terrill and Reisig (2003) examined the influence of neighborhood context on the level of force the police exercised; they found that higher levels of force was used on people that were stopped in disadvantaged neighborhoods, and the effect of the person's race was amplified by neighborhood context.

2.4 New York's Stop-and-Frisk

Numerous studies have been done on Stop-and-Frisk, yet a broad consensus on the question of racial bias remains elusive. Gelman et al. (2007) performed a "hit-rate analysis" on Stop-and-Frisk data from 1998-99 and concluded that there was a racial bias against minorities. The subject of this hit-rate analysis was the probability that a given stop would result in an arrest, after being conditioned on race. The conclusion of racial bias stemmed from the fact that white stops were more likely to result in an arrest than minority stops, yet minorities were being disproportionately stopped. Assuming that the goal of Stop-and-Frisk is to minimize crime, and by extension maximize the number of arrests, then a racially unbiased officer should be stopping more whites in order to maximize the probability that a given stop would yield an arrest, the complete opposite of what was observed in the data. Persisco and Coviello (2013) replicated the hit rate analysis on Stop-and-Frisk data from 2003-11 and initially arrived at the same conclusion as Gelman et al (2007). However, after they accounted for precinct level fixed effects they concluded that minorities stops (specifically African-American) were more likely to result in an arrest than white stops; adding further fixed effects for year and suspected crime showed that race was uncorrelated with the probability of being arrested.

Hit-rate analyses may tell us whether or not minorities are being stopped in excess of their relative rates of crime participation (or some other benchmark), however they are relatively uninformative when it comes to assessing another crucial part of economic valuation–the social cost of racial profiling in policing stops. Lerman and Weaver (2014) examined the effect of Stop-and-Frisk on civic engagement in disadvantaged neighborhoods in New York City. Using 311 calls as a measure for civic engagement, they found that a high concentration of police stops was associated with greater civic engagement; however neighborhoods that had a high degree of stops that featured searches or the use of force, particularly when they did not result in an arrest, significantly decreased neighborhood outreach to their local government. Legewie (2016) found strong evidence that NYPD police officers actively discriminate against African Americans in their decision to use force. He employed a natural experiment, relying on publicized shootings of police officers as "exogenous shocks" and using timestamps and geo-coded data to create control (before the shock) and treatment (after the shock) groups. He found that when a police officer was shot by a black suspect, African Americans who were stopped in the following days had a significantly higher probability of force being used compared to African Americans who were stopped in the same location at approximately the same time in the days preceding the shooting.[6] The probability of force for whites and Hispanics remained relatively unchanged. Even more interestingly, this experiment was repeated for a police shooting were the shooter was white, and afterward repeated again for a Hispanic shooter, and the corresponding racial/ethnic groups did not see an increase in the use of force in the days following the shooting.

3 Data and Methods

3.1 Background Information and Descriptive Statistics

The data for this study was obtained from the NYPD Stop, Question and Frisk Report Database and the American FactFinder website, which is managed and maintained by the US Census Bureau. NYPD protocol dictates that stops are recorded by the initiating officer by use of a Stop-and-Frisk report (UF-250 form), and uploaded to a database maintained by the NYPD. Yearly updates are made publicly available, and the database currently holds Stop-and-Frisk data from 2003 to 2015. Given the political turmoil brought about by federal district court ruling in 2013, this study has chosen to exclude the years 2013 - 2015.[7] Due to the absence of detailed demographic data on neighborhoods within the different precinct jurisdictions, the years 2003 and 2004 were also omitted and the analysis was restricted to 2005 - 2012 .[8] The neighborhood profiles were constructed using data collected from

[6]To clarify, 'same location' is a geographic control variable and does not refer to the location that the shooting took place, but rather compares stops across the city that took place in the same location (e.g. specific cross streets) *before* and *after* the shooting

the American Community Survey (ACS); PUMA (Public Use Microdata Area) are nested within precincts, thereby allowing one to construct neighborhood profiles by aggregating and matching the PUMA to their respective police precincts.

Table 1 shows a cross-tabulation between race/ethnicity and the use of force by the police, with stops disaggregated into non-compliant and compliant stops.[9] For the purpose of this analysis, non-compliant behavior was defined by the following activities: changing direction at the sight of a police officer, evasive response to questioning, visibly engaging in criminal activity, making furtive movements, refusing to comply with the officer's directives, verbal threats by the suspect and criminal possession of a weapon. It is immediately clear that minorities are being targeted for stops, however without observing the counterfactual and knowing the crime participation rates and elasticities, we cannot say much more on the issue of who is being stopped. However if we restrict the universe to only people who have been stopped, racial and ethnic disparities persist despite the fact that everyone who was stopped were in some sense "equally suspicious". Whites comprise 10.24% of the people who were stopped, yet were underrepresented in stops were forced was used, both for compliant and non-compliant; we observe the opposite pattern with Blacks and Hispanics, were they are overrepresented in stops where force was used relative to their representation in the total distribution of stops. Figure 1 shows a similar racial/ethnic trend across time and boroughs, with whites consistently claiming the smallest share of stops where any kind of force was used.

[7] Even prior to the court ruling 2013 was a difficult year for the NYPD in general; three unarmed black males aged 16, 18, & 27 were killed by the NYPD in the months preceding the court ruling and racial tensions as well as public distrust in the police were at an all time high. It was also an election year and mayoral candidates openly questioned the authority of the police. It is also evident that the police responded to this increased scrutiny; in 2011 and 2012, over 1,200,000 people were stopped-and-frisked, however in 2014 and 2015 the combined number of stops fell to under 70,000.

[8] This decision was due to the fact that the American Community Survey only goes as far back as 2005

[9] It should be noted that the data was cleaned prior to the construction of this table.The raw dataset (2003-2012) had 4,792,542 observations in total, 214,957 of which were were discarded during the cleaning processes after being deemed unusable; missing data for key explanatory/control variables or data entry errors were the main reasons that led to an observation being discarded.

3.2 Empirical Strategy

The empirical strategy is two part: the first set of models (fixed-effects logit) focus on identifying the effect of individual (pedestrian) characteristics on the probability of force, and the second set of models (random-intercept logit) focus on identifying "neighborhood effects", or the proportion of variance in the use force that can be attributed to precinct characteristics. In order to identify racial bias in the use of force, it is critical to distinguish between the necessary use of force and *frivolous* use of force; failure to do so will lead to incorrectly attributing racial bias to disparities in the use of force, when in reality the effect of race/ethnicity is being confounded with omitted variables that significantly increase the probability that force will be used during a stop. In order to circumvent this issue, stops were disaggregated into compliant and non-compliant stops (assuming in non-compliant stops officers are more likely to use force) and separate logistic regression models were estimated. The dependent variable for each regression **pforce** is a binary outcome that takes on the value of one when any sort of force is used by the police during a pedestrian stop. Force, if used by the police, is recorded as one (or more) of nine possible categories on the UF-250 form and consists of any of the following actions: baton, weapon drawn, suspect against the wall, hands, handcuffs, pepper spray, weapon pointed, suspect on the ground, other. The main independent variables are factor variables for race/ethnic group (black, hispanic, black hispanic, white hispanic) with white being the omitted category and factor variables for age groups, with 55 and over being the omitted category; body mass index (bmi), whether or not the person was arrested, and a categorical variable denoting "other race" were included as control variables, along with precinct and year fixed effects. The fixed effect models also allowed for clustered standard errors to reflect precinct intracluster correlations. For the random-intercept logit, non-compliance was included as a control variable and stops were pooled instead of being disaggregated, as the goal is to identify the proportion of overall variance in the use of force that is being driven by precinct level differences. As with any random-intercept model there are concerns of possible endogeneity, i.e. omitted variable bias: a quick examination of the point estimates for the individual characteristics in the random-intercept model revealed that they closely mirrored the point estimates in the fixed effects model. To further dispel any fears of endogeneity a scatterplot of the mean-centered fixed effects and predicted random effects was constructed and the points fell almost entirely on the fitted line $y = x$. The point estimates from the random-intercept model were not used in the main analysis and can be found (along with the scatterplot) in Appendix A.

4 Results and Discussion

4.1 The Role of Individual Characteristics

Table 2 shows the exponentiated coefficients (odds-ratio) and standard errors of four logistic regressions with **pforce** as the dependent variable. Regressions (1) & (2) are pairwise identical in specification, as are regressions (3) & (4), with the first regression of each pair being restricted to non-compliant stops and the latter being restricted to compliant stops. The only difference in specification between the pairs of regressions is the treatment of Hispanic pedestrians: the NYPD records Hispanic stops as either Black Hispanic or White Hispanic, and the first pair of regressions ignored this demarcation and aggregated them into a single Hispanic category, whereas the second pair of regressions maintained the original categorization. In the second pair of regressions, Non-Hispanic Black pedestrians where removed from the data in an attempt to isolate the effect of race in the context of Hispanic ethnicity.

Across all regressions the variables for race and ethnicity are statistically significant and greater than one, indicating that minorities are more likely to have force used on them relative to whites, regardless of compliance or non-compliance. The effect of race in the context of Hispanic ethnicity is equally telling: in both regressions (3) (non-compliant) and (4) (compliant) Black Hispanics are more likely to have force used on them relative to both whites and White Hispanics. For age, the effect of age group can only be interpreted relative to the omitted category (55+); we see that younger pedestrians are significantly more likely to force used on them relative to older pedestrians, and the magnitude of the effect decreases as the age group increases. The pattern holds across all four regressions.

4.2 Spatial Disparities and Neighborhood Effects

In order to examine whether spatial disparities in the use of force by the NYPD are correlated with racial demographics, a set of separate random-intercept logistic regressions were estimated. The data was partitioned into distinct 3-year time periods and the model was fitted using the years 2005-2007 and 2010-2012. Table A1 shows the regression output, and figures 2 and 3 show the (normalized) distribution of random intercepts across precincts. These random intercepts represent the contribution of precinct characteristics to the probability that force will be used during a pedestrian stop, and are often referred to as "neighborhood effects" in the literature. This differs from the previous logistic regressions in that it assumes spatial differences to be random draws from a single distribution rather than mere regional idiosyncrasy. The fixed effects logit estimates an idiosyncratic baseline rate or propensity to use force for each precinct and controls for it when identifying the effect of individual characteristics; the random-intercept logit assumes there is only one true propensity to use force and precinct differences represent random draws from a distribution centered on the true propensity. It is possible to impose a functional form on this distribution by estimating a random-coefficients model, however this was avoided due to the erratic behavior and computational complexity of those models. Nevertheless the random-intercept model is quite informative and provides a descriptive, if not inferential, study of precinct (neighborhood) disparities in the use of force.

The distributions of the random-intercepts were normalized to a fixed mean $\mu = 0$ with an unknown variance ψ that was estimated as a model parameter in the regressions; the estimated values of ψ are $\hat{\psi} = 0.2966$ for 2005-2007 and $\hat{\psi} = 0.3833$ for 2010-2012. This rescaling allows the random-intercepts to be interpreted as deviations from the mean, with values less than zero representing deviations below the mean (less likely to use force) and values greater than zero representing deviations above the mean (more likely to use force). The estimated variances tell us that in 2005-2007 29.66% of the variability in the use of force was driven by precinct characteristics rather than individual characteristics; likewise, for 2010-2012 precinct characteristics contributed to 38.33% of the variance in the decision to use force. Precinct characteristics are attributes that vary at the precinct level as opposed to the individual level: demographics of neighborhoods within the precinct jurisdiction, precinct specific mandates, etc. While none of these characteristics were explicitly modeled, the random-intercepts give us a relative measure of how these characteristics affect between-precinct variance. By ranking them we can then identify which precincts have the greatest propensity to use force and which have the lowest propensity to use force.

Table 3 shows an ordered ranking of precincts based on their propensity to use force, with rank 1 corresponding to the precinct with the lowest propensity and rank 76 to the precinct with the highest propensity. The first four precincts with the lowest propensity to use force all have extremely low minority populations, particularly Blacks; we observe the opposite trend for the four precincts with the greatest propensity to use force, where a large minority population corresponds to a greater propensity to use force. The exact trend is also evident (if not more pronounced) in Table 4, suggesting that racial demographics of neighborhoods are highly correlated with an officer's decision to use force. It should be noted that in neighborhoods with both a high white population and a high Hispanic population it is not unusual for a large share of the Hispanic population to also identify as white; however an exact estimate of this share was unavailable in the Public Use Microdata Area.

5 Conclusion

The results of this study provide strong evidence to the claim that NYPD officers exhibit racial bias in Stop-and-Frisk interactions, specifically in stops where force was used. The question of causality is central to econometrics and the methodology employed in this study sought to identify, if any, the causal relationship between race/ethnicity and the use of force by disaggregating stops into compliant and non-compliant pedestrians. Had no demarcation been made between compliant and noncompliant pedestrians the effects of race/ethnicity would have been biased upwards, as minorities are overrepresented in non-compliant stops and hence the effect of noncompliance would have been confounded with the effect of race/ethnicity.

This study adds to the rich literature on racial profiling and policing, specifically studies addressing disparities in the use of force. It also sheds light on another growing area of research in the social sciences that has yet to be fully formalized–neighborhood effects.

Geographers have known for some time that spatial patterns are correlated with a variety of social outcomes, and extensive work has been done on studying the formation and implications of these spatial patterns. However, as geographers their primary goal has been the study of the pattern themselves rather than how spatial variations translate into social processes; this paper attempts to go one step further by not simply studying space itself but rather the demography of space and how spatial variations in racial demographics can influence outcomes in a police stop. This interpretation of neighborhood effects, while insightful, is still quite naive: there exists no consensus on a formal definition of a neighborhood, and many papers written claiming to identify neighborhood effects rely on researchers selectively agglomerating people with certain characteristics in a manner that creates the spatial variation that is presumed to be influencing the outcome variable. In essence, the reported neighborhood effects that are construed as causal mechanisms are largely an artifact of the researcher's design. New York City suffers from a similar issue of neighborhoods with no fixed boundaries, with people on the same street sharing the same zip code can claim to be residents of different neighborhoods. In order to circumvent this issue of fuzzy neighborhood boundaries, this paper relied on administrative boundaries that are fixed with sharp discontinuities: precinct boundaries. Put differently, what I've been referring to in this paper as neighborhood effects should be more aptly named *precinct effects*, since the spatial variation occurs at the precinct level rather than the neighborhood level. Given the (somewhat exploratory) findings of this paper, I hope that it may provide some sort of impetus for social science researchers to formally define notions of neighborhood and space in a manner that does not entirely depend on its inhabitants. While it is undeniable that there is reciprocal determinism between these two entities (the neighborhood and its inhabitants), the practice of allowing inhabitant characteristics to shape neighborhood boundaries is problematic as it becomes impossible to distinguish between individual effects and neighborhood effects.

6 Tables

Table 1: Stop-and-Frisk 2005 - 2012

	Non-compliant Stops		Compliant Stops		
Race	**Force**	**No Force**	**Force**	**No Force**	**Total**
White	57,113	210,734	13,682	143,640	425,169
	7.41%	10.20%	8.87%	12.39%	10.24%
Black	426,978	1,135,638	80,779	600,394	2,243,789
	55.37%	54.99%	52.36%	51.78%	54.07%
Hispanic	265,225	644,880	55,107	366,125	1,331,337
	34.40%	31.23%	35.72%	31.58%	32.08%
Other	21,760	73,969	4,694	49,373	149,796
	2.82%	3.58%	3.04%	4.26%	3.61%
Total	771,076	2,065,221	154,262	1,159,502	4,150,061
	100.00%	100.00%	100.00%	100.00%	100.00%

Table 2: Fixed-Effects Logit

	(1)	(2)	(3)	(4)
VARIABLES	**pforce**	**pforce**	**pforce**	**pforce**
Black	1.208***	1.152***		
	(0.0443)	(0.0391)		
Hispanic	1.153***	1.101***		
	(0.0319)	(0.0315)		
Black Hispanic			1.214***	1.177***
			(0.0387)	(0.0390)
White Hispanic			1.124***	1.080***
			(0.0241)	(0.0259)
Age13_18	1.511***	1.591***	1.454***	1.670***
	(0.0422)	(0.0539)	(0.0474)	(0.0685)
Age19_24	1.499***	1.543***	1.491***	1.616***
	(0.0350)	(0.0494)	(0.0433)	(0.0605)
Age25_30	1.409***	1.438***	1.403***	1.481***
	(0.0315)	(0.0375)	(0.0408)	(0.0473)
Age31_36	1.291***	1.303***	1.274***	1.364***
	(0.0292)	(0.0353)	(0.0371)	(0.0440)
Age37_42	1.164***	1.201***	1.165***	1.262***
	(0.0281)	(0.0335)	(0.0367)	(0.0448)
Age43_48	1.095***	1.154***	1.096***	1.200***
	(0.0249)	(0.0306)	(0.0325)	(0.0446)
Age49_54	1.011	1.046*	1.014	1.073**
	(0.0233)	(0.0278)	(0.0310)	(0.0343)
arrested	3.565***	5.620***	3.700***	5.974***
	(0.156)	(0.346)	(0.174)	(0.395)
BMI	1.001**	1.000	1.001**	1.001
	(0.000364)	(0.000240)	(0.000586)	(0.000566)
Constant	0.279***	0.102***	0.263***	0.115***
	(0.0233)	(0.00815)	(0.0178)	(0.00945)
Observations	2,834,038	1,312,743	1,272,527	632,061

Robust Standard Errors in parenthesis * p<0.05 ** p<0.01 *** p<0.001

Table 3: Neighborhood Profile, 2005 - 2007

Rank	Precinct	Black	White	Hispanic	Total Stops	Force Used
1	108	2.40%	51.10%	34.20%	15,493	5.86%
2	123	1.20%	93.00%	8.60%	7,334	5.35%
3	62	0.70%	64.30%	11.20%	13,380	7.88%
4	111	1.70%	60.90%	11.00%	10,063	9.18%
73	70	39.10%	45.30%	12.70%	28,434	38.50%
74	52	20.80%	20.40%	63.70%	10,636	41.49%
75	50	14.60%	47.90%	41.80%	6,867	42.58%
76	44	37.50%	9.60%	62.00%	20,217	63.40%

Neighborhoods within precinct jurisdiction:

- 108 - Long Island City, Sunnyside, Woodside
- 123 - South Shore (Staten Island)
- 62 - Bath Beach, Bensonhurst, Gravesend
- 111 - Bayside, Douglaston, Little Neck, Auburndale, Hollis Hills, Fresh Meadows
- 70 - Flatbush, Midwood, Kensington, Ocean Parkway
- 52 - Bedford Park, Fordham, Kingsbridge, Norwood, University Heights
- 50 - Riverdale, Fieldston, Kingsbridge, Marble Hill, Spuyten Duyvil
- 44 - Highbridge, Concourse, Mount Eden, Concourse Village

Table 4: Neighborhood Profile, 2010 - 2012

Rank	Precinct	Black	White	Hispanic	Total Stops	Force Used
1	111	1.50%	44.00%	12.30%	12,709	4.89%
2	123	1.00%	83.90%	9.40%	6,163	4.99%
3	22				3,164	8.19%
4	62	0.80%	48.00%	14.10%	13,637	7.53%
73	115	5.70%	10.40%	65.00%	38,910	43.55%
74	32	60.60%	10.70%	22.40%	33,840	43.86%
75	46	27.60%	1.50%	68.60%	30,649	48.52%
76	44	33.60%	1.20%	62.40%	43,126	52.43%

Neighborhoods within precinct jurisdiction:

- 123 - South Shore (Staten Island)
- 111 - Bayside, Douglaston, Little Neck, Auburndale, Hollis Hills, Fresh Meadows
- 22 - Central Park
- 62 - Bath Beach, Bensonhurst, Gravesend
- 115 - East Elmhurst, North Corona, Jackson Heights
- 32 - Central Harlem
- 46 - Fordam, University Heights, Morris Heights, Bathgate, Mount Hope
- 44 - Highbridge, Concourse, Mount Eden, Concourse Village

7 Figures

Figure 1: Racial/Ethnic Trends In The Use Of Force Across Boroughs

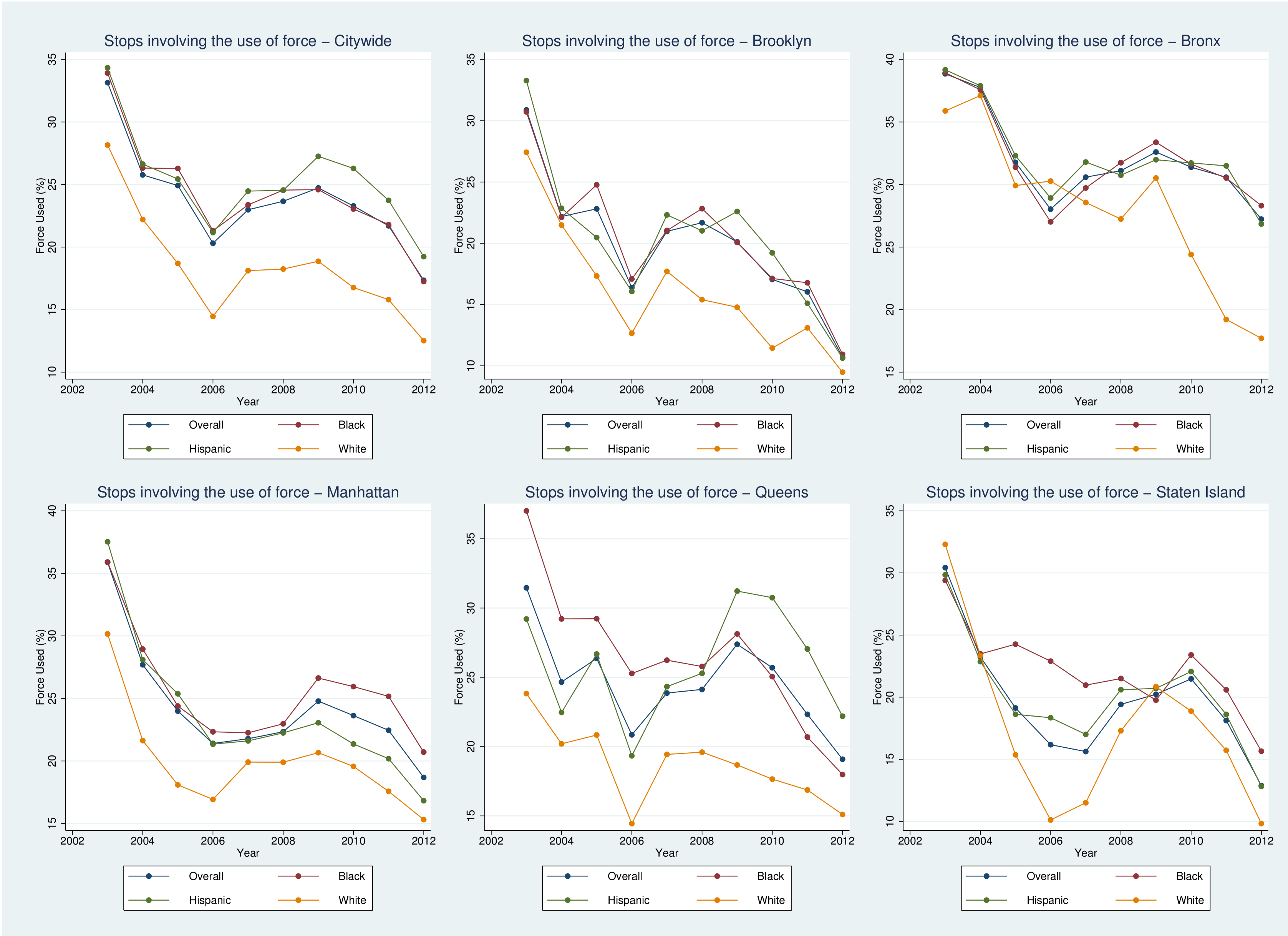


Figure 2: Neighborhood Effects 2005 - 2007

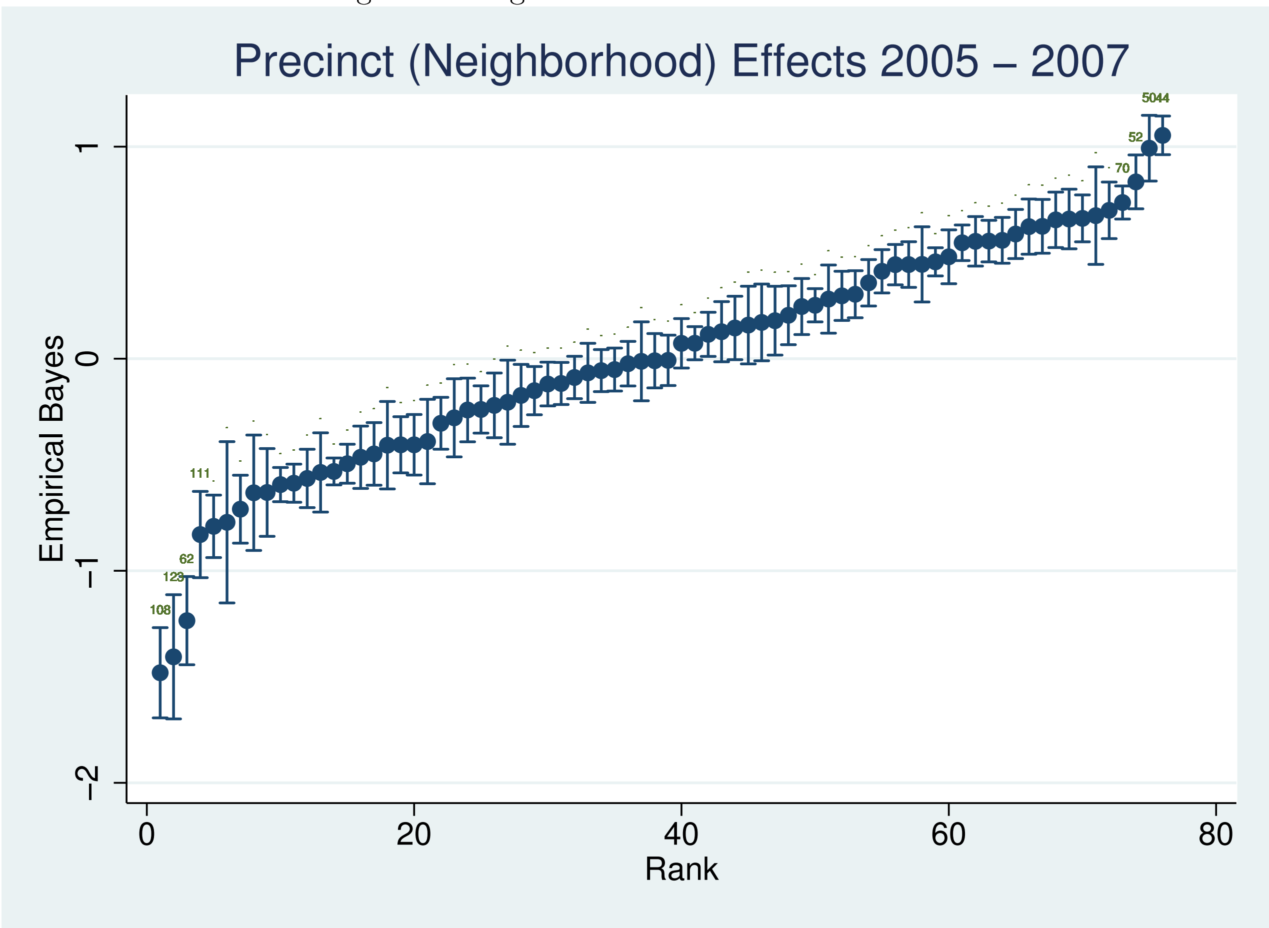
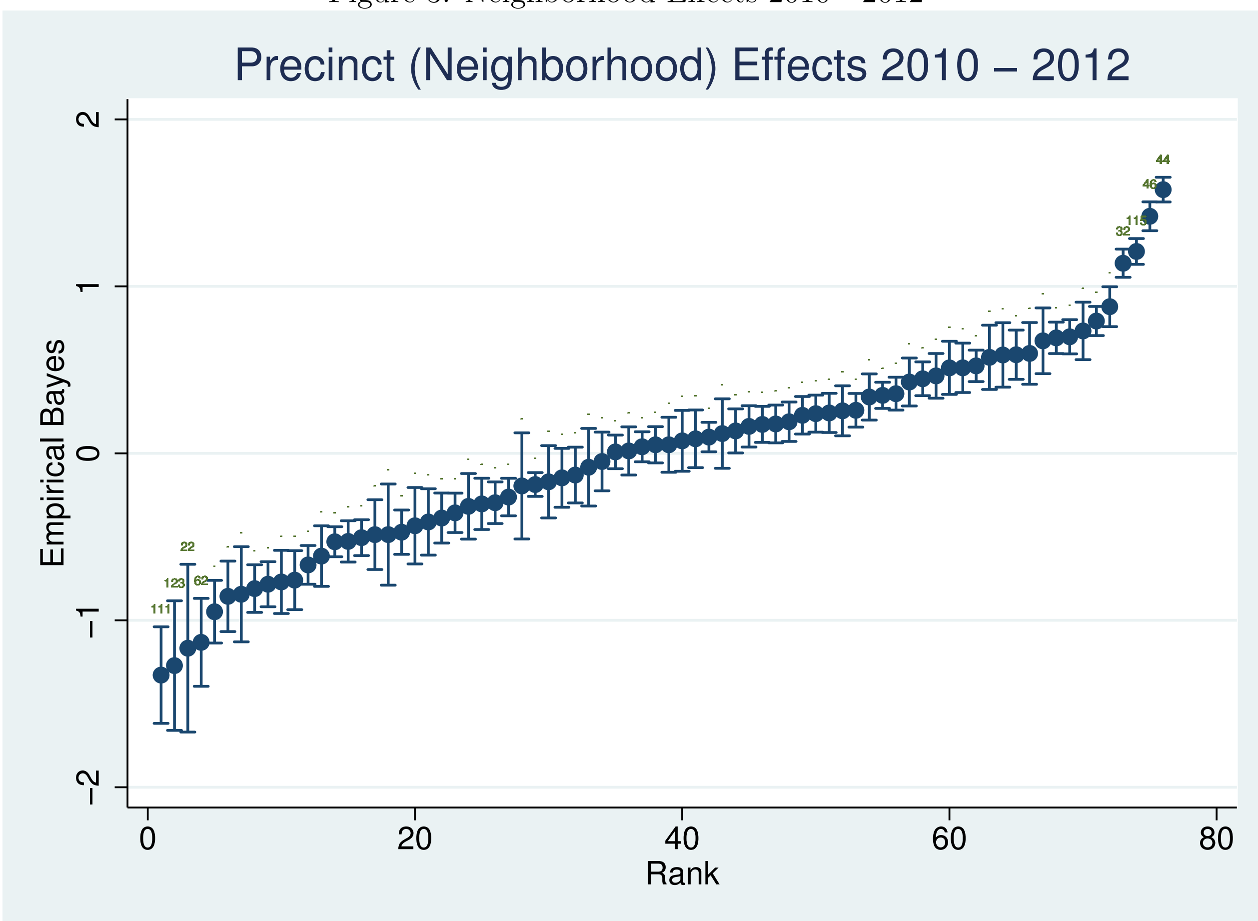


Figure 3: Neighborhood Effects 2010 - 2012



8 Appendix A: Endogeneity Concerns & Robustness Tests

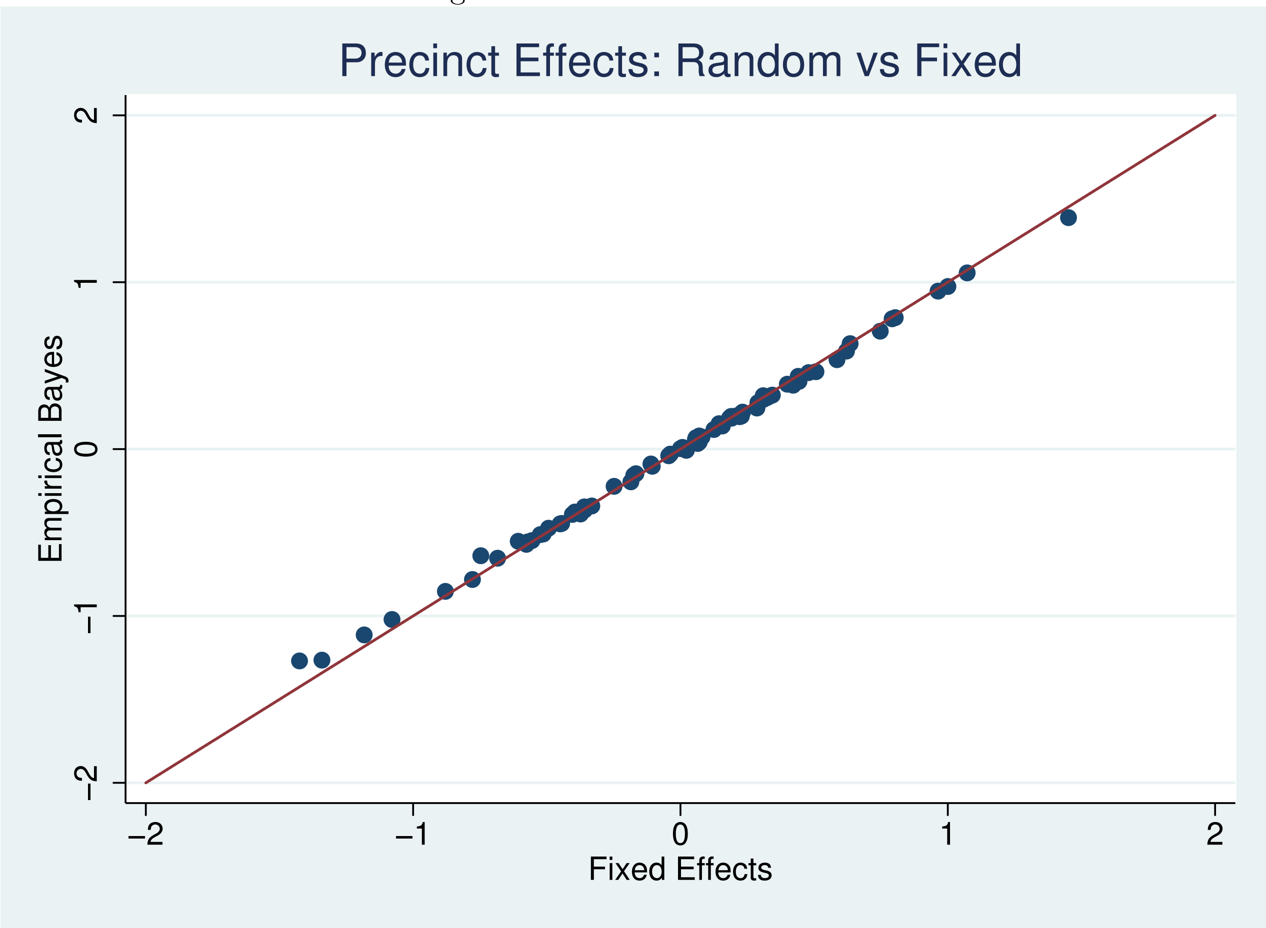
Precinct Heterogeneity: Random Effects (Intercepts) versus Fixed Effects

Endogeneity is an important issue in econometrics and plagues many applied econometric models: it can significantly bias regression coefficients, making it difficult to properly identify parameters and estimate the true treatment effect(s). Endogeneity often takes the form of an omitted variable bias and can be solved using proxy variables or an instrumental variable design; in the context of a panel or hierarchical dataset (such as Stop-and-Frisk), a common solution is to use fixed effects dummy variables as a proxy to control for unobserved differences across clusters. This was the approach adopted in the fixed effect logit regressions, where the goal of the analysis was to identify the average effect of race/ethnicity of the probability of force. The fixed effect dummy variable controls for precinct differences, but doesn't explicitly tell us precinct specific contributions to the probability of force. The random-intercept logit does the opposite: it provides a relative measure of precinct specific contributions, but does not control for them in the same manner as the fixed effects therefore leaving it prone to bias. These approaches are considered to be orthogonal with the latter approach being less favored due to the increased potential of bias; however there are times when they can deliver close results. Table A1 and Figure A1 shows that this is indeed the case in the context of Stop-and-Frisk, as the point estimates for the random effects (RE) and fixed effects (FE) are close across all time periods. Figure A1 shows a scatterplot of the mean-centered fixed effects versus the empirical bayes random intercepts (from the 2005 - 2012 pair of regressions), with the trend line $y = x$ superimposed. The estimates sit almost entirely on the line with the tails of the distribution being pulled towards zero (the mean), which is consistent with the empirical bayes shrinkage properties.

Table A1: Random-Intercept (RE) & Fixed-Effects (FE) regressions

	2005 - 2007	2005 - 2007	2010 - 2012	2010 - 2012	2005- 2012	2005 - 2012
VARIABLES	RE	FE	RE	FE	RE	FE
black	1.210***	1.206***	1.229***	1.222***	1.208***	1.204***
	(0.0611)	(0.0606)	(0.0505)	(0.0496)	(0.0520)	(0.0510)
hispanic	1.149***	1.140***	1.185***	1.175***	1.128***	1.116***
	(0.0485)	(0.0478)	(0.0483)	(0.0473)	(0.0451)	(0.0441)
age13_18	1.555***	1.604***	1.459***	1.457***	1.608***	1.642***
	(0.0763)	(0.0789)	(0.0696)	(0.0694)	(0.0727)	(0.0705)
age19_24	1.429***	1.471***	1.421***	1.419***	1.597***	1.633***
	(0.0649)	(0.0641)	(0.0638)	(0.0635)	(0.0717)	(0.0699)
age25_30	1.310***	1.350***	1.401***	1.395***	1.508***	1.533***
	(0.0521)	(0.0541)	(0.0706)	(0.0712)	(0.0639)	(0.0641)
age31_36	1.186***	1.218***	1.299***	1.301***	1.420***	1.454***
	(0.0557)	(0.0560)	(0.0678)	(0.0674)	(0.0665)	(0.0651)
age37_42	1.028	1.056	1.171***	1.167***	1.243***	1.253***
	(0.0517)	(0.0503)	(0.0595)	(0.0594)	(0.0579)	(0.0556)
age43_48	1.032	1.062	1.073	1.066	1.168***	1.181***
	(0.0497)	(0.0471)	(0.0607)	(0.0597)	(0.0604)	(0.0611)
age49_54	0.936	0.962	0.988	0.991	1.100	1.130**
	(0.0570)	(0.0574)	(0.0560)	(0.0556)	(0.0661)	(0.0650)
ncompli	2.622***	2.589***	2.656***	2.674***	2.591***	2.627***
	(0.108)	(0.104)	(0.134)	(0.134)	(0.111)	(0.115)
arrested	4.214***	4.207***	4.044***	4.035***	3.857***	3.913***
	(0.240)	(0.239)	(0.213)	(0.217)	(0.186)	(0.193)
bmi	0.998*	0.999*	1.002	1.002	1.001	1.002*
	(0.000795)	(0.000768)	(0.00180)	(0.00181)	(0.000885)	(0.000854)
Constant	0.0894***	0.0500***	0.0567***	0.123***	0.0675***	0.0825***
	(0.00874)	(0.00374)	(0.00630)	(0.0106)	(0.00572)	(0.00771)
Cons(Precinct)	1.345***		1.467***		1.323***	
	(0.0664)		(0.0911)		(0.0653)	
Var(Precinct)	0.2966		0.3833		0.2798	
	(0.0493)		(.0621)		(.0494)	
Observations	129,953	129,953	123,458	123,458	124,396	124,396

Figure A1: Robustness Check



9 References

[1] Black, D. (2010). *The Behavior of Law*. New York: Academic Press.

[2] Bittner, E. (1970). *The Functions of Police in Modern Society*. Washington, DC: Government Printing Office.

[3] Gelman, A., Fagan, J., & Kiss, A. (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* 102:813-823.

[4] Harcourt, B. (2004). Unconstitutional Police Searches and Collective Responsibility. *University of Chicago Public Law Legal Theory Working Paper*. No. 66

[5] Knowles, J., Persico, N., & Todd, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy* 109:203-229.

[6] Legewie, J. (2016). Racial Profiling and Use of Force in Police Stops: How Local Events Trigger Periods of Increased Discrimination. *American Journal of Sociology*. (forthcoming)

[7] Lerman, A., & Weaver, V. (2014). Staying Out of Sight? Concentrated Policing and Local Political Action. *The Annals of the American Academy of Political and Social Science* 651:202-219

[8] Muhammad, K. (2010). *The Condemnation of Blackness: Race, Crime and the Making of Modern Urban America*. Cambridge: Harvard University Press.

[9] Persisco, N., & Coviello, D. (2013). An Economic Analysis of Black-White Disparities in NYPD's Stop and Frisk Program. *NBER Working Paper*. No. 18803

[10] Pettit, B. & Western, B. (2004). Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration. *American Sociological Review* 69:151-169.

[11] Plant, E. & Peruche, M. (2005). The Consequences of Race for Police Officers' Responses to Criminal Suspects. *Psychological Science* 16:180-183

[12] Sampson, R. & Lauritsen, J. (1997). Racial and ethnic disparities in crime and criminal justice in the United States. *Crime and Justice* 21:311-374.

[13] Terrill, W., Paoline, E., & Manning, P. (2003). Police Culture and Coercion. *Criminology* 41:1003-1034

[14] Terrill, W. & Mastrofski, S. (2002). Situational and officer-based determinants of police coercion. *Justice Quarterly* 19:215-248

[15] Terrill, W. & Reisig, M. (2003). Neighborhood Context and Police Use of Force. *Journal of Research in Crime and Delinquency* 40:291-321.

[16] Weber, M. (2014). *Politics as a Vocation*. Philadelphia: Fortress Press.