An Economic Analysis of Black-White Disparities in NYPD's Stop and Frisk Program[*,†,‡]

Decio Coviello
HEC Montreal

Nicola Persico
Northwestern University

December 10, 2013

Abstract

We study the possible racial bias of the police involved in NYPD's "stop and frisk program." A model is introduced to explore the identification of two distinct sources of bias: bias at the level of the police officer making the stop decisions, and bias at the level of the police chief allocating manpower across precincts. Previous research offered positive identification results regarding officer bias; this paper adds a new, and negative, identification result for police chief bias. Ten years of data from NYPD's "stop and frisk program" are analyzed in light of this theoretical framework. White pedestrians are found to be slightly less likely than african-american pedestrians to be arrested conditional on being stopped. We interpret this as evidence that the officers

[*]Thanks to Matthew Bloch for sharing with us the NYC electoral data and to Tim Brophy for helping us in mapping the NYC data into police precincts.

[†]This paper is an update of NBER working paper 18803 with the same title, published in February 2013.

[‡]The data used in this paper can be downloaded from: http://www.nyc.gov/html/nypd/html/analysis_and_planning/stop_question_and_frisk_report.shtml. We also offer access to the replication material to academic faculty and graduate students. Please write to decio.coviello@hec.ca if you would like access.

making the stops are *on average* not biased against African Americans relative to whites, because the latter are being stopped despite being a "less productive stop" for a police officer. When the same likelihood is computed at the precinct level, we find heterogeneity in measured bias among precincts. We correlate, without arguing causality, this measure of bias with a number of precinct characteristics. Separately, we find no cross-precinct correlation between this measure and the relative intensity of police pressure on african-american residents (relative to pressure on whites). This absence of correlation seems to go against the presumption that large disparities in police pressure are necessarily the result (or even the correlate) of police officer bias.

JEL codes: J71, K42

1 Introduction

New York City's "stop and frisk program" is a police strategy whereby pedestrians are briefly stopped by police officers, engaged in conversation/questioning, and potentially searched. The procedure is perceived as demeaning for those who are stopped and/or frisked. The program also disproportionately impacts non-whites. The racial impact of the program has given rise to public protests[1] and widespread allegations of racial profiling.[2]

The stop and frisk program has been repeatedly challenged in court. In the most recent such lawsuit, U.S. District Judge Shira Scheindlin said that that the case involved "an issue of great public concern," namely "the disproportionate number of African Americans and Latinos, as compared to whites, who become entangled in the criminal justice system."[3] The case, *Floyd et al.* v. *City of New York*, was decided against the NYPD on August 12, 2013, with the judge finding a violation of equal protection rights.[4] The decision was front page news on the major U.S. newspapers and gave rise to a lively public debate, with Mayor Bloomberg accusing the judge of "deliberately denying the city a fair trial" and vowing to appeal the decision,[5] and the *New York Times* editorial board supporting the decision.[6] This judicial decision is likely to have repercussion on police activity, and perhaps on crime,[7] in

[1]On March 16, 2012 several thousand people marched in New York City protesting the policy, which the organizers say "creates an atmosphere of martial law for the city's African American and Latino residents" (see "Thousands March Silently to Protest Stop-and-Frisk Policies," *New York Times*, June 17, 2012).

[2]According to *The New York Post*, "The stop-and-frisk policy has been under fire by vocal opponents, including many 2013 mayoral contenders, because the vast majority of people stopped are African American or Hispanic." (Quoted from "NYPD issues department-wide memo regarding racial profiling during 'stop-and-frisks'," May 17, 2012). Rev. Al Sharpton, writing on the Huffington Post on June 6, 2012, writes:

> When a majority of those targeted by police are young men of color and when the bulk of them are innocent, what else are we to conclude other than the fact that the NYPD has been implementing a policy of racial profiling and discrimination?

[3]Cited from "Court Strikes Challenge to Stop-and-Frisk Trial," Courthouse News Service, November 14, 2008.

[4]The judge also found that the "stop and frisk" program gave rise to 4th amendment violations (unreasonable searches and seizures). We will not focus on 4th amendment issues.

[5]Quoted from "Judge Rejects New York's Stop-and-Frisk Policy", *New York Times* page A1, August 13, 2013.

[6]See "Racial Discrimination in Stop-and-Frisk". August 13, 2013, page A22 of the New York edition.

[7]According to William Bratton, the former New York City police commissioner and Los Angeles police

other jurisdictions.

In this paper we examine the same "stop and frisk" data that were analyzed in the trial. Our aim is not to "audit" the judge's opinion from a legal standards perspective. Rather, we inquire as to whether the "stop and frisk" program is administered in a racially biased way *from the perspective of social science*. From this perspective, legal tests and rules of evidence are not central; rather, what matters (or should matter) are precise definitions and statistical analyses that speak to them. We believe it is important for social scientists to offer their perspective when troubling allegations of racial bias are made. We also believe it is good for this conversation to take place in academic journals.

In our view the public debate on "stop and frisk" is missing a sharp definition of what it means for the program to be racially biased. We propose one here. Following Becker [5], we say that a program is biased if those who administer it are motivated in part by an (impermissible) preference for a specific racial group. Such a "distortion" in preferences can be present at two levels: at the level of the officers making the stops; and at the level of the police chief allocating officers to precincts. (Both levels of bias are alleged in the public debate). In Section 2 we lay out a game-theoretic model which incorporates these distortions, and we use it to determine which type of statistical evidence can be used to identify bias at either level, given the presence of confounding factors (unobservables). We conclude, consistent with previous literature, that bias at the officer's level can be identified by looking at the success rates of stops. We reach a negative conclusion regarding bias at the police chief level: neither data on the frequency of stops, nor data on their success rates, can identify a latent distortion in the police chief's preferences. Interestingly, data on stop frequency are not helpful for identification at either level, which is at odds with the public focus on such data.

We then turn to the NYPD data. In Section 4 we analyze the prevalence of stops in different racial groups and find that, by any reasonable measure, African Americans are stopped

chief " ... any police department in America that tries to function without some form of stop and frisk, or whatever terminology they use, is doomed to fail." The Associated Press, January 23, 2013.

much more frequently than whites as a proportion of their population. But, as mentioned above, we find it difficult to rule out unobservables, as opposed to officer bias, as potential explanations for this disparity.

In Section 5 we turn to arrest rates and find that the arrest rates of stopped African Americans and whites are essentially identical, at least on average across all precincts. We interpret the latter finding as inconsistent with the hypothesis that officers are biased in their stopping decision, at least on average. When we zoom down to the precinct level, however, we uncover heterogeneity: in some precincts arrest rates are higher for stopped African Americans, while in other precincts they are higher for stopped whites. These differentials are somewhat persistent over the years. We interpret these differentials as indicative of bias on the part of the officers (against whites or African Americans, depending on the precinct). We correlate the cross-precinct variation in measured bias with a number of precinct-specific characteristics. We find that the differential probability of an African American being arrested positively correlates with the Bloomberg margin of victory and the median income of the precincts, while it negatively correlates with the median age in the precincts. Once we introduce precinct-level fixed effects, however, only serious crime is (positively, though only at the 10% significance level) correlated with our measure of bias against African Americans.

Our measure of bias, it turns out, does not correlate with race disparities in police pressure across precincts. This is interesting because disparities in police pressure across races are often taken to strongly indicate bias. The no-correlation finding seems to go against this interpretation, at least in this dataset.

In Section 6 we comment on the results and discuss extensions.

1.1 Related Literature

Two papers are closely related. Gelman *et al.* [10] have analyzed New York stop and frisk data from the years 1988-89. Most of their analysis focuses on documenting disparities in stops; using sophisticated statistical analysis they conclude that "persons of African and Hispanic descent were stopped more frequently than whites, even after controlling for precinct variability and race-specific estimates of crime participation."[8] In their Section 5.3 they also briefly address the disparity in arrest rates conditional on stop, and tentatively conclude that police officers were indeed racially biased against African Americans. This tentative conclusion is based on the statistical fact that African Americans were less likely than whites to be arrested conditional on being stopped. We replicate Gelman *et al.*'s [10] finding in our more recent and extensive data, but show that the finding is overturned when we add precinct-level fixed effects. The second closely related paper is Ridgeway [20], a RAND report sponsored by the New York City Police Foundation. Using data for the year 2006, this report finds that "unjustified" race disparities in stop rates are much smaller than the ones commonly reported in the literature. Key to this finding is the choice of benchmark for what level of disparity is "justifiable."[9] Turning to arrest rates, this report finds arrest rates that are overall higher for whites than for African Americans (Tables 5.1, 5.2). This is the opposite result to Gelman *et al.* [10], and the same result we find in our more extensive dataset. For this part of the analysis, Ridgeway [20] uses a matching procedure which re-weighs observations so as to ensure an equal distribution of several characteristics of the stop.[10] If the matching procedure uses variables that are endogenous to the outcome (racial bias in our case), then the re-weighting procedure might not be innocuous.[11]

[8]Gelman *et al.* [10], p. 813.

[9]Census data about the fraction of residents of a given race is deemed not appropriate. This issue is well explained on p. 15 of the report.

[10]These variables include: crime suspected, precinct, average age, time of day, location, month, sex, day of the week, type of ID (physical or verbal), whether the stop was made on a radio run, the x-y coordinates of the stop location, being reported by witness, being part of an ongoing investigation, being in a high-crime area, being at a high-crime time of day, being close to the scene of an incident, detecting sights and sounds of criminal activity, evasiveness, association with known criminals, changing direction at the sight of an officer, carrying a suspicious object, fitting a suspect description, appearing to be casing, acting as a lookout, wearing clothes consistent with those commonly used in crime, making furtive movements, acting in a manner consistent with a drug transaction or a violent crime, or having a suspicious bulge. See Ridgeway [20], pp. 34-5

[11]A potential source of endogeneity might be the use of location as a matching variable. If the police express their bias by over-policing a location that is mostly frequented by non-whites, the matching procedure "washes out" this channel for bias.

In sum, a comparison of the two papers discussed above indicates that there is no consensus on the key question: is there any bias in NYPD's stop and frisk program? We feel our paper makes progress on this front. First, it organizes the evidence around a model. This is necessary to understand what features of the data can be indicative of bias, and whose bias (the officers', or the chief's) is being identified. Second, it uses a dataset which is more recent and more comprehensive in its time span (ten years as opposed to two for Gelman *et al.* [10] and one for Ridgeway [20]).

Knowles *et al.* [12] first introduced a version of the model in Appendix A and derived Theorem 1. That paper does not feature the analysis in Section 2.2 which is an original contribution of this paper. Also of some relevance, Knowles *et al.* [12] dealt with a setting (highway searches) which did not feature geographic units, such as the precincts which feature prominently both in our theory and in our empirical analysis.

A more broadly related literature is that on hit rate analysis, which develops partly in reaction to Knowles *et al.* [12]. See Ayres [4], Persico and Castleman [16], Todd [23], Whitney [25], and Persico [14] for reviews of this strategy. Ayres and Waldfogel [3] earlier used this strategy to look for racial bias in the judge's decision of the level at which to set bail. The hit rates analysis has been later utilized in the policing context by Persico and Todd [18], Sanga [21], Hernandez-Murillo and Knowles [11], Persico and Todd [19], Childers [6], Gershman [9], Persico [15], Persico and Todd [17]. Anwar and Fang [2] offer an alternative approach to hit rate analysis. Dharmapala and Ross [7], Alpert *et al.* [1], and Smith *et al.* [22] offer critical appraisals of hit rate analysis.

2 The Model

This model is meant to conceptualize what it means for policing to be biased, and to give guidance as to where to look for it in the data. Agents in our model are of three kinds: citizens of race $r \in \{A, W\}$ living in district i, who choose whether to commit a crime which is detected through a stop-and-frisk; a mass P of police officers who stop-and-frisk citizens; and a police chief whose only action is to assign officers to precincts.

Following Becker [5], we define bias as a "taste for discrimination." By that we refer to a component of an agent's utility function that is dependent on the race of those with whom the agent interacts. A key modeling choice is whether the officer's possible taste for discrimination is innate, or whether it reproduces the culture of the precinct to which each officer is assigned. In this model we opt for the second possibility, and assume that officers who are assigned to precinct i inherit the bias of that precinct.[12] Therefore, in this model it may be more appropriate to talk about precinct bias as opposed to officer bias.

The game has two stages. In the first stage, a police chief allocates police officers to precincts. In the second stage a game within each precinct is played between officers and pedestrians. The second-stage games (one in each precinct) are assumed to reproduce the one studied by Knowles *et al.* [12].[13] Because the model is somewhat standard, we do not reproduce its description here but we relegate it to Appendix A. The equilibrium of each precinct-level game yields a function $K_i^r(P_i)$, which for each precinct summarizes the fraction of pedestrians of race r who commit a crime in district i when P_i officers are allocated to that district. The function $K_i^r(P)$ is derived in Appendix A.

Let us start with the first stage. What would be a "permissible" objective or motive for the chief which defines unbiased behavior? In the case of a police chief or other central authority allocating resources across districts, it seems reasonable to define this objective as the minimization of crime.[14] If we make this assumption, then we may conceptualize the

[12]This is done for two reasons. First, because we believe there is such a thing as a precinct culture. Second, because otherwise any observed differences in precinct bias would, within the model, result from the police chief's choice of which officers to allocate to which district. This does not seem to us like it would be a realistic assumption given that the personnel being allocated are patrol officers, whose bias in probably not known by the police chief.

[13]More precisely, in the version studied by Persico and Todd [18].

[14]Of note, this objective would be meaningless for the individual police officer in a district, because an individual officer probably has a negligible impact on aggregate crime. Put differently, it would be impractical to reward any police officer based on total crime in New York City or even in her precinct, because that outcome depends only minimally on the officer's behavior.

"legally permissible version" of the police chief's problem as follows:

$$\min_{P_i} \sum_i \left[N_i^A K_i^A (P_i) + N_i^W K_i^W (P_i) \right] \quad \text{s.t.} \ \sum_i P_i \leq P,$$

where P represents the total amount of police officers available to the police chief, and N_i^r denote the number of pedestrians belong to group r. According to this formulation, the legally permissible objective of the police chief is to minimize the sum crimes across all precincts.

Deviations from this behavior can be classified as biased. What would be the impermissible, or biased, version of the police officer's objective function? Perhaps one in which we allow the police chief to "prioritize" the crime rate of certain precincts, which may be objectionable especially if these priorities turn out to be correlated with the race of the precinct residents. This can be conceptualized by assigning "welfare" weights Γ_i to the crime rates of precinct i. In addition, the police chief may single out the crimes commited by pedestrian of a particular race. We can conceptualize this by adding race-specific weights γ^r. Then, a potentially biased police chief would solve the following problem:

$$\min_{P_i} \sum_i \Gamma_i \left[\gamma^A N_i^A K_i^A (P_i) + \gamma^W N_i^W K_i^W (P_i) \right] \quad \text{s.t.} \ \sum_i P_i \leq P. \tag{1}$$

The parameters Γ_i, γ^A and γ^W in the above problem capture the police chief's bias.

2.1 Identifying Precinct-Level Bias

The precinct-level game is described and analyzed in Appendix A. The analysis reproduces that in Persico and Todd [18] and points to the success rate of stops as a key indicator of officer bias. The logic, intuitively, is the following. Suppose the success rate was lower for stops of african-american pedestrians. An officer who was not biased against African Americans, and was motivated by the prospect of making an arrest, should cut down on less-productive stops of African Americans and increase the more-productive stops of whites. This "arbitrage" on the part of individual officers has aggregate effects: as officers shift to policing whites, the crime rate in the other group rises and the crime rate among whites decreases. This arbitrage would continue until, under a perfectly unbiased police force, arrest rates ("hit rates") are equalized between the stops of white and African American pedestrians in the precinct. If the police force were biased, however, this arbitrage would stop earlier, at a point where the differential between the African American and white arrest rates is exactly offset by the officer's bias. This logic gives rise to the following result, which is proved in Appendix A.

Theorem 1 ***(Persico and Todd [18]: positive result on identification of police officer bias)*** *In the equilibrium of the precinct-level game, the arrest rate is the same across all subgroups within a race that are distinguishable by police. Also, if the police are unbiased, then the arrest rate is the same across races. If the police are biased against race r, the arrest rate is lower in race r than in the other race. Thus officer bias can be identified using arrest rates.*

This theorem provides the justification for the hit rates test applied in the next section. The theorem also identifies a significant advantage of the hit rates test: under the behavioral assumptions stipulated in the model, the test is robust to omitted-variable bias. To see this, interpret the subgroups mentioned in the theorem as the set of pedestrians sharing a certain characteristic which is observed by the officer, but perhaps not observed by the econometrician. The econometrician, therefore, only observes the average arrest rate among all subgroups, but not the arrest rates within each subgroup. In principle this poses a problem because arrest rates within a subgroup is what identifies bias against that subgroup. However, the theorem says that this is not a problem, because the *equilibrium* arrest rate must the same across all subgroups within a race that are distinguishable by police, even if these subgroups are not distinguishable by the econometrician. Therefore, the econometrician's lack of ability to discern subgroups has no impact on his ability to infer bias from arrest rates.

2.2 Identifying Bias in Manpower Allocation Across Precincts

Preliminary to the question of identifying the weights Γ_i, γ^A and γ^W in problem (1), a conceptual ambiguity in the interpretation of these weights must be noted. One might be ambivalent about whether the configuration $\Gamma_i > \Gamma_j$ represents bias in favor or against precinct i. On the one hand, $\Gamma_i > \Gamma_j$ means that precinct i's crime rate is more salient than precinct j's, and accordingly, more resources will proportionally be devoted to precinct i, resulting in a lower crime rate in that precinct. On the other hand, $\Gamma_i > \Gamma_j$ means that precinct i will be assigned more police officers, so more stops and more frisks, which some civil liberty advocates object to especially if police pressure correlates with minority population in the neighborhood. So, it is conceptually/normatively ambiguous whether $\Gamma_i > \Gamma_j$ means that precinct i is favored or disfavored relative to precinct j.

Apart from the above conceptual/normative ambiguity, there is also an empirical difficulty in estimating the (unobserved) weights β_i. To see the nature of this difficulty, let us derive the equilibrium predictions which would allow us to estimate the weights. The first order conditions necessary for optimality in problem (1) are:[15]

$$\begin{aligned} &\Gamma_i \left[\gamma^A N_i^A \frac{\partial}{\partial P_i} K_i^A (P_i) + \gamma^W N_i^W \frac{\partial}{\partial P_i} K_i^W (P_i) \right]_{P_i = P_i^*} \\ &= \Gamma_j \left[\gamma^A N_j^A \frac{\partial}{\partial P_j} K_j^A (P_j) + \gamma^W N_j^W \frac{\partial}{\partial P_j} K_j^W (P_j) \right]_{P_j = P_j^*} \quad \text{for all } i, j. \end{aligned} \tag{2}$$

where P_i^* and P_j^* represent the optimal allocation. Conditions (2) represent a system of equations, one for each precinct. In this system, N_i^r and P_i^* are known. The unknowns we seek to solve for are Γ_i, γ^A and γ^W, which together are more than the number of equations. Clearly, identification is not possible from this system alone. Furthermore, even if we restrict

[15]It is convenient to assume that $C_i(\cdot)$ is a concave function. Under this assumption, which we maintain, the first order conditions are also sufficient for optimality.

attention to a subset of parameters (say, we somehow know the Γ_i's and only need to identify the γ^r's), we can't do that either if we do not observe the elasticities of crime to policing, $\frac{\partial}{\partial P_i} K_i^A (P_i)$. This is proved in the next theorem.

Theorem 2 ***(negative result on identification of police chief bias)*** *The parameter Γ_i cannot be identified from system (2) without knowledge of the elasticity of crime to policing$\frac{\partial}{\partial P_i} K_i^r (P_i)$. The parameter γ^r cannot be identified from system (2) without knowledge for at least a pair (i, j) of the elasticities of crime to policing $\frac{\partial}{\partial P_i} K_i^A (P_i)$ and $\frac{\partial}{\partial P_j} K_j^A (P_j)$. Without knowledge of these elasticities neither crime levels (equivalent to hit rates in the model) $K_i^r (P_i)$, nor stop intensities P_i^*, are helpful in identifying any bias parameters in the police chief's problem (1).*

Proof. Let's solve for Γ_1. Write (2) for $i = 1$. Manipulating (2) we get

$$\begin{aligned} &\Gamma_i \left[\gamma^A N_i^A \frac{\partial}{\partial P_i} K_i^A (P_i) + \gamma^W N_i^W \frac{\partial}{\partial P_i} K_i^W (P_i) \right]_{P_i = P_i^*} \\ &= \Gamma_j \left[\gamma^A N_j^A \frac{\partial}{\partial P_j} K_j^A (P_j) + \gamma^W N_j^W \frac{\partial}{\partial P_j} K_j^W (P_j) \right]_{P_j = P_j^*}. \end{aligned}$$

Even if the right hand side is known, identification of Γ_1 separate from the term in brackets by which it is multiplied requires knowedge of the term in brackets. This in turn requires knowedge of $\frac{\partial}{\partial P_i} K_i^A (P_i)$ and $\frac{\partial}{\partial P_i} K_i^W (P_i)$. Thus the first sentence in the proposition is proved. Now let's solve for γ_A. Manipulating (2) we get

$$\gamma^A \left[\Gamma_i N_i^A \frac{\partial}{\partial P_i} K_i^A (P_i) - \Gamma_j N_j^A \frac{\partial}{\partial P_j} K_j^A (P_j) \right] = \gamma^W \left[\Gamma_j N_j^W \frac{\partial}{\partial P_j} K_j^W (P_j) - \Gamma_i N_i^W \frac{\partial}{\partial P_i} K_i^W (P_i) \right].$$

Even if the right hand side is known, identification of γ^A separate from the term in brackets by which it is multiplied requires knowedge of the term in brackets. This in turn requires knowedge of $\frac{\partial}{\partial P_i} K_i^A (P_i)$ and $\frac{\partial}{\partial P_j} K_j^A (P_j)$ for at least one pair (i, j). Thus the second sentence in the proposition is proved. The third sentence follows because neither crime levels/hit rates, nor stop intensities enter system (2) unmediated by the function $\frac{\partial}{\partial P_i} K_i^r(\cdot)$. ■

This theorem proves an important point. Knowledge about the amount of policing directed to precincts, or about hit rates, cannot help identify the bias parameters in the police chief's problem. Instead, the key statistic is an elasticity. Unfortunately, it is generally difficult to get persuasive estimates of elasticities because elasticities captures a counterfactual: what would happen to the crime rate *if the police chief happened to perturb the allocation of manpower from its equilibrium level.* Thus, estimating elasticities requires observing more than simply the level of crime at an equilibrium. This is an empirical challenge.[16]

The takeaway from this section is that identifying bias in the allocation of manpower across precincts is difficult for two reasons. The first difficulty is of a "normative" nature, and it has to do with what it means for an allocation to be biased against a precinct. The second difficulty is that it is difficult to obtain empirical estimates of the weights in problem (1).

3 The Data

We use data collected by the NYPD on individual stops, questionings and frisks in the City of New York between 2003-2012.[17] This appears to be a slightly longer period than the one at issue in the Floyd case. The database contains information on whether the person was frisked, issued a summons or arrested, the type of crime which is suspected by the police making the arrest, the race of the pedestrian, the timing and location of the stop. Unless explicitly mentioned, we restrict the sample to african-american and white pedestrians, setting Hispanics aside because the charge of racial bias seems to have special force with reference to the African American population.[18] In this restricted sample of 2,947,865 stops, approximately 6 percent of the stopped pedestrians were arrested and 84

[16]Typically, exogenous variation of police manpower is necessary to identify the elasticity of crime to policing.

[17]The database can be downloaded at the following link:
http://www.nyc.gov/html/nypd/html/analysis_and_planning/

[18]We briefly extend our focus to Hispanics in Section 6.1.

percent of the stops are of African American pedestrians, the rest of whites. Most of the suspected crimes fall into one of these categories: Possession of a Weapon (27%); Robbery (17%); Criminal Trespass (12%); Grand Larceny Auto (9.1%); Burglary (8.9%).[19] Table 1 reports some descriptive statistics.

A possible caveat regarding these data is that NYPD officers are not required to record *all* interactions with private citizens. NYPD policy prescribes the completion of a stop and frisk report (form UF-250) only under the following circumstances: a person is stopped by use of force; a person stopped is frisked or searched; a person is arrested, or a person stopped refused to identify himself (and was later identified by the officer).[20] Stops which occurred but did not give rise to one of these outcomes need not be recorded.[21] Therefore, recorded stops (the ones in our database) may be a selected sample of all stops.[22] Judge Schendlin notes the data limitations, but ultimately accepts the data as a useful, if imperfect, tool to aid her decision. We do the same.[23]

A second, very important caveat must be raised regarding the suitability of arrests as an outcome for hit rates analysis. The ideal outcome is a measure of productivity which the

[19]Other crimes are: Grand Larceny (4.3%); Illegal Possession of Substances (3.6%); Marihuana (3.3%); Assault (4%); Illegal Sales of Substances (2.9%); Petit Larceny (2.5%); Mischief (1.2%); Graffiti (1.1%).

[20]See Chapter 5 of the US Commission for Civili Rights Report [24].

[21]On the other hand, Sara LaPlante in personal communication indicates that "NYPD policy requires officers to fill out a UF-250 form for every Terry stop. From the patrol guide 212-11.6: "Prepare STOP, QUESTION AND FRISK REPORT WORKSHEET (PD344-151A) for EACH person stopped" (emphasis theirs)."

[22]There might be some indication, however, that the sample is somewhat representative of all stops. We estimate that 35% of the stops in our data were recorded despite not resulting in any of the outcomes that legally trigger the requirement to record the stop. (The outcome "refused to identify" is not recorded in the data. We proxy for it using the field "evasive response to questioning.") This suggests that officers may have an incentive to record stops, perhaps as a way of showing effort to their supervisors. To the extent that this incentive results in most stops being recorded, the data set is representative of all the stops in New York City.

[23]It is tempting to address the selective recording concern by restricting the sample to stops that are required by law to be recorded. Within this sample, the problem of selective recording should not exist. The trouble 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. Conditioning our analysis on such ex-post information would mean conditioning on information not possessed by the officer at the time of the stop. Put differently, the outcomes contained this restricted data set cannot be said to fully portray the outcomes generated by any officer's stop behavior. Thus the hit rate analysis cannot properly be applied to such a subsample. For this reason we will utilize the full sample of all stops in the database.

Table 1: Descriptive Statistics

	Mean	sd	n
Outcomes			
Arrest made	5.8	23	2,947,865
Race of the pedestrian			
African American	84	37	2,947,865
Crime details			
Possession of a Weapon	27	44	2,496,267
Robbery	17	37	2,496,267
Criminal Trespass	12	32	2,496,267
Grand Larceny Auto	9.1	29	2,496,267
Burglary	8.9	28	2,496,267
Grand Larceny	4.3	20	2,496,267
Illegal Possession of Substances	3.6	19	2,496,267
Assault	4	20	2,496,267
Marihuana	3.3	18	2,496,267
Illegal Sales of Substances	2.9	17	2,496,267
Petit Larceny	2.5	16	2,496,267
Mischief	1.2	11	2,496,267
Graffiti	1.1	10	2,496,267
Other Crimes	4.3	20	2,496,267

Notes. Variables expressed in percent. *African American* is an indicator variable coding the pedestrian's race. *Crime details* are 13 indicators of the type of crime represent 95% of the crimes recorded in the sample. Years 2003-2005 have missing values in the variable *Crime details*.
Source. Statistics for the City of New York, Years 2003-2012.

officer legitimately maximizes, and which is itself "objective," that is, is not tainted by police bias. Arrests might not be "objective" because they might be subject to police discretion, and thus may themselves be tainted by police bias. For example, all else equal, the police may be more likely to arrest an african-american than a white pedestrian after having stopped either. This is a very valid concern. The best evidence to address this concern would be the rates at which arrests, which are usually warrantless in our sample, are later upheld by a judge.[24] Unfortunately, such data are not available to us. We address the concern, to the extent possible with the data at hand, by looking at the officer's behavior *after the stop*

[24]Because the conversion is done by a judge, it is arguably an "objective" outcome, in the sense that any judicial bias should be uncorrelated with the bias of the police officer making the arrest. Jeff Fagan, in personal communication, indicates that the "conversion rate" of arrests into prolonged detentions is well below 100%.

has been made. We check whether, given the type of suspected crime ascertained by the police officer after the stop, the officer is more likely to arrest a African American than a white pedestrian. If the police use discretion in their decision to arrest, and this discretion is correlated with race, then we would expect to see African Americans being arrested more often than whites for the same type of suspected crime.[25] To implement this test, we check whether the race of the person stopped predicts the probability of arrest *after conditioning on the type of suspected crime* (as recorded by the officer on Form UF-250). Table B.1 presents the results. After controlling for the type of crime suspected, the estimates suggests that there is no evidence of a significant race effect on arrest. We interpret this result as consistent with the hypothesis that, given a crime suspected of a pedestrian of either race, officers are not using discretion in deciding whom to arrest, or at least, that any discretion they use is uncorrelated with race.[26,27] Therefore, we do not find evidence against using the outcome "arrest" as the outcome in the hit rate test.

4 Disparities in Police Pressure

New York City's stop-and-frisk program disproportionally impacts minorities. The New York Civil Liberties Union makes this point forcefully by documenting that, in 2011, 52.9 percent of stops were of African Americans, 33.7 percent were of Latinos, while whites accounted for only 9.3 percent of the stops.[28] After restricting attention to African Americans and whites only, the following figure 1 (left panel) summarizes this striking disparity. The panel depicts the difference in *police pressure* by race of the pedestrian. For each race, police pressure is defined as: average number of stops in a year divided by total population in NYC. In the

[25] The assumption behind this test is that there is no discretionality in the officer's recording of the crime variable.

[26] We also repeat the analysis discussed in the previous section in the subsample of stop and frisk for which we have the data on crimes. The evidence remains consistent with previous findings of no discrimination on average and is available upon request.

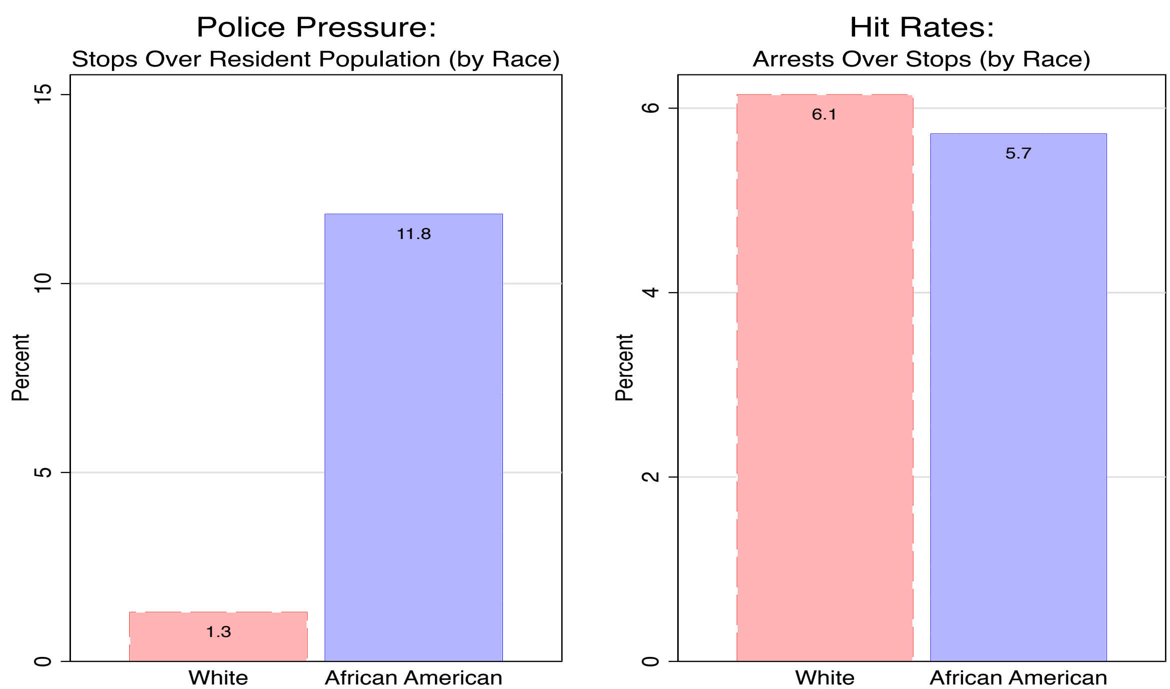
[27] In Section 6.1 we extend our analysis and consider the stops of Hispanics.

[28] See NYCLU [13], pg.5.

whole sample, pressure is about ten times larger for African Americans as it is for whites.

This disparate impact is certainly problematic from a social viewpoint. However, disparate impact can reflect many factors, both observable and unobservable, which affect the stop-and-frisk process. Neither Theorem 1 nor Theorem 2 indicate that police pressure can help identify police bias. This lack of theoretical framework becomes problematic when, as in Table 2, conditioning on observables attenuates the disparity. Let us turn to this table.

Figure 1: **Police Pressure and Hit Rates in New York City**



Notes. The figure reports the yearly average number of stops over resident population (left panel) and the yearly average arrests over yearly average stops (right panel) divided by ethnicity in New York City (in %).
Source. Statistics for the City of New York, Years 2003-2012. Resident population from the 2010 Census data.

Table 2: Correlates of Relative Police Pressure in New York City

Model	OLS	OLS	OLS
Sample	Panel	Panel	Panel
	(1)	(2)	(3)
Fraction of African American	-0.222***	-0.057	-0.063
	(0.073)	(0.055)	(0.043)
Income		0.365***	0.293***
		(0.119)	(0.106)
Constant	23.118***	-1.322	-12.284
	(3.857)	(6.780)	(23.984)
Relative police pressure (average)		17	
Fraction of African Americans (in %) in the average precinct		26.78	
Number of precincts	75	75	75
Observations	750	750	750
Adj. R^2	0.083	0.266	0.465
Precinct Controls	no	no	yes
Time FE	no	yes	yes

Notes. Estimates are on 75 precincts. The dependent variable is (relative) police pressure ($\frac{\frac{Arrests of African Americans}{African American population}}{\frac{Arrests of Whites}{White population}}$) in New York City. *Fraction of African American* is the percentage of the population that is African American in the precinct in 2010. *Income* is the inflation adjusted median income in the precinct in 2010. Column 3 also includes the difference between M. Bloomberg and the first running opponent M. Green, or F. Ferrer, or B. Thompson vote share in the 2001, 2005, 2009 elections, respectively. Missing years are computed using moving averages; the percentage of the population that is African American in the precinct in 2010; the inflation adjusted median income, 2010 precinct average; the median age, 2010 precinct average; the precinct average percentage of the population that is female in 2010; the precinct average percentage of the population aged 15-24 with a college degree in 2010; the number of annual crimes (murders, rapes, robberies, fel. assaults, burglaries, grand larcenies, grand larceny autos) in each precinct divided by the precinct population in 2010 (in 1,000 habitants) for the years 1998, 2001, 2012; the number of annual graffiti in each precinct in 2011 divided by the precinct population in 2010 (in 1,000 habitants); the total number of annual civic initiatives (education, emergency preparedness, environment, helping neighbours in need, strengthening communities) in each precinct, in 2011 divided by the precinct population in 2010 (in 1,000 habitants); and an indicator for African American commanding officers. All variables are described in Appendix C.
To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes" regressions additionally include year fixed effects (9 dummies). Standard errors are clustered at the precinct level. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012. Resident population from the 2010 Census data.

Table 2 makes use of variability across precincts. The dependent variable is the ratio of African American/white police pressure.[29] Columns (1) and (2) show that this disparity in police pressure is correlated with the precinct's racial makeup. From an a-theoretical perspective this dependence on race seems troubling. However, column (3) shows that conditioning on income takes away the effect of race (and in addition eliminates the significance of the constant in the regression).[30] Thus, it now appears that police manpower is being allocated disproportionally to precinct with poorer residents – which is arguably less bad than allocating disproportionally to minority precincts. Is this encouraging news? Does this say anything about police bias? In our view, it is difficult to learn much from this a-theoretical approach except at a descriptive level. For this reason we regard the estimates in this section as suggestive, but not dispositive, about the presence of bias. We turn, therefore, to the analysis of arrest rates which has its theoretical underpinning in Theorem 1.

5 Analysis of Arrest Rates

Theorem 1 indicates that a comparison of arrest rates by race can identify bias in the police officers' stop decision. In this section we compute arrest rates and compare them across races.

5.1 Arrest Rates by Race (Aggregate Across Precincts)

We start by noting that, in the aggregate, the probability that a stop translates into an arrest is quite similar across races in our sample. This is shown by regressing an indicator variable coding whether the stopped pedestrian was arrested on another indicator variable

[29] The table shows that, averaging within precincts first, then taking arithmetic average of precinct-level means, the ratio of African American/white police pressure is 17. This figure differs from the ratio (about 10) implied by Figure 1, left panel, because in that figure the average was computed at the city-wide level.

[30] All variables included in this regression are described in Appendix C.

coding the pedestrian's race. Figure 1 (right panel) shows the aggregate arrest rates for stopped pedestrian of either race. Clearly, the large disparity across races that is present in police pressure (left panel) is absent when we look at average arrest rates.

More detailed estimates are reported in Table 3. Depending on the specification, African American pedestrians who are stopped are between 0.42% and 0.44% less likely to be arrested compared to whites (columns 1-3). Thus the probability of a stop resulting in an arrest is about 6% for whites v. 5.6% for African Americans. Although the difference is very small, and perhaps unlikely to be perceived by an officer based on his own experience alone, the difference is significant in two out of three specifications. This pattern is similar to that found by Gelman *et al.* [10] in their more limited sample.

Table 3: Arrest Made

Model	OLS (1)	OLS (2)	OLS (3)	FE (4)	FE (5)	FE (6)	FE (7)
African American	-0.420*** (0.037)	-0.437*** (0.037)	-0.437 (0.469)	0.379*** (0.046)	0.355*** (0.046)	0.355* (0.207)	0.340* (0.204)
Constant	6.140*** (0.034)						
Mean outcome				5.79%			
Fraction of African American				84%			
P-value of $H_0 : u_i = 0$				0.001	0.001	0.001	0.001
Number of precincts				76	76	76	76
Observations	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865
Cluster SE	no	no	yes	no	no	yes	yes
Time FE	no	yes	yes	no	yes	yes	yes
Precincts FE	no	no	no	yes	yes	yes	yes
Time FE · Precincts FE	no	no	no	no	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of being arrested conditional on being stopped in New York City (in %). *African American* is an indicator variable coding the pedestrian's race. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Columns 3, 5-7, shows show clustered standard errors at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

This small difference in arrest rate between races changes sign, however, when we control for precincts. Precincts vary considerably in the likelihood that a stop translates into an arrest (refer to Figure 2).[31] Controlling for precincts is appropriate because precints are, in effect, separate jurisdictions.[32] It is also necessary if baseline arrest rates are correlated with race, to avoid a fallacy in aggregation.[33]

In columns 4-7 of Table 3 we run the same OLS regressions, this time with 76 precincts-level fixed effects. Notably, the coefficient on "African American" changes sign. Now, stopping an african-american pedestrian results in a probability of arrest which is larger by 0.355% compared to a white pedestrian. That is, after accounting for the fact that different precincts have different "baseline" rates of arrest conditional on search, African Americans are no longer less likely to be arrested conditional on being stopped.[34] Based on Theorem 1, this evidence is interpreted as rejecting the hypothesis of a relative bias against African Americans

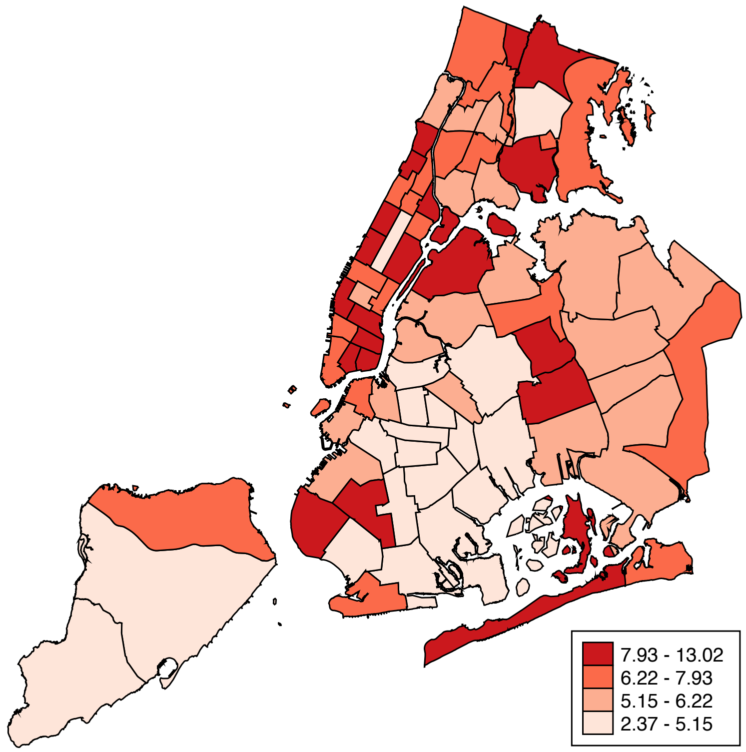
[31]This is not surprising, given the heterogeneity among precincts. For institutional information about precincts, refer to Fyfe and Kane (2006).

[32]The NYPD is organized in 76 precincts, each of which is responsible for a specific geographic area. An officer from one precinct cannot stop pedestrians in another precinct. According to the New York state Criminal Procedure Law (CPL.140.50) , "a police officer may stop a person in a public place located within the geographical area of such officer's employment".

[33]To understand the possible fallacy let's, for the sake of argument, treat precincts as separate jurisdictions. If the police officers in each precinct were unbiased, then *within each precinct* the arrest rates of African American and white pedestrians stopped should be the same. However, the levels of these arrest rates need not all be the same *across precincts*. For example, suppose hypothetically that of all the African Americans and whites stopped in the Bronx 3% were arrested, and 6% of the African Americans and whites stopped in the Financial District were arrested. If we aggregated the data from the two precincts we would mistakenly conclude that the police officers making the stops are biased against African Americans, because in the aggregate sample most African Americans are searched in the Bronx and have a 3% arrest rate, much lower than whites, most of which are searched in the Financial District. Thus the hit rate test carried out without controlling for precincts would be potentially biased, or more precisely, uninformative about the racial bias exhibited by police officers within each precinct. A solution to this aggregation problem is to introduce precinct-level fixed effects in the statistical model that predicts arrest rates. In the above hypothetical example, introducing precinct-level fixed effects into the baseline specification allows the fixed effects to absorb the 3% and 6% baseline arrest rates, while the coefficient on "African American " would be estimated to equal zero. This zero coefficient would properly be interpreted as evidence that the police is not biased. Conversely, if the police were biased then we would observe lower arrest rates on African American searchees in many or all precincts, and this African American-white difference in arrest rates would be picked up by the coefficient on "African American,"after controlling for precinct-level fixed effects. Therefore, controlling for precinct fixed effects is necessary for the hit rates test to function properly.

[34]The precinct-level fixed effects jointly explain, in a statistical sense, the arrest rate (i.e., the p-value is less than 5% for the joint test of all the precincts fixed effects equal to zero, see Table 3).

Figure 2: **Probability of Being Arrested Conditional on Being Stopped in New York City**



Notes. The figure reports the probability of being arrested conditional on being stopped in New York City (in %).
Source. Statistics for the City of New York, Years 2003-2012.

in the officers' stop decisions, *at least on average across precincts*. This is a key message from our analysis.[35]

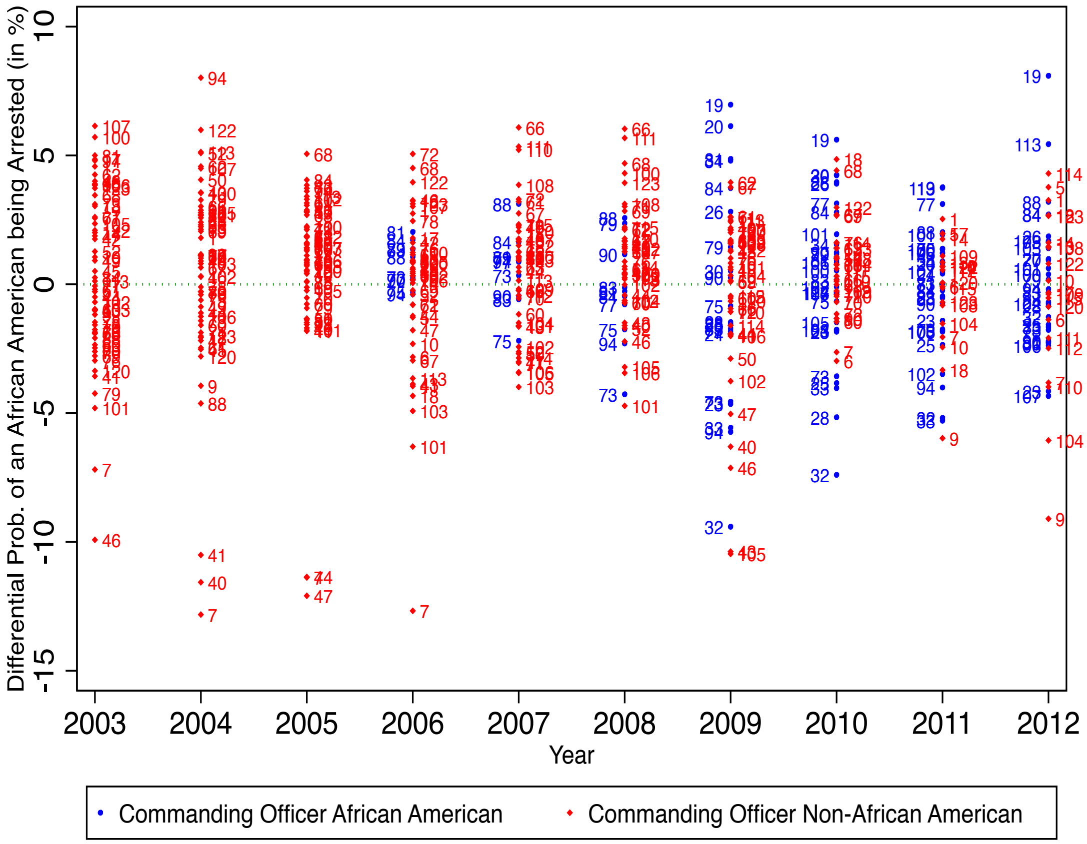
5.2 Cross-Precinct Variation in Arrest Rates by Race

The evidence in the previous section is consistent with the interpretation that, *on average*, there is no bias against African Americans in the police officers' stop decisions. There might, however, be precinct-level biases which cancel each other out. This is confirmed by Figure

[35]For completeness, we ran the various specifications in Table 3 on the subsample of stops that are required by law to be reported – refer to footnote 23. In this subsample the coefficient on "African Americans" does not become positive in the specification with precinct-level fixed effects. As mentioned before, this subsample cannot be a proper sample for a hit rate analysis. Therefore, we disregard the results from this analysis.

3. For each precinct-year, the figure plots the *difference in arrest rates* of African Americans and whites stopped.[36] A value above zero is interpreted as bias against whites.

Figure 3: **Differential Probability of an African American being Arrested**



Notes. The figure reports the differential probability of an African American being arrested by precinct/year. This is the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. Blue dots indicate precincts whose borough commander was african-american (otherwise red diamonds).
Source. Statistics for the City of New York, Years 2003-2012.

The figure indicates that, in each year, there is a considerable cross-precinct dispersion in arrest rate differentials, which is roughly centered around zero.[37] These deviations from zero

[36]Specifically, for each precinct-year, Figure 3 plots the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians.
[37]Between-precincts standard deviation equals 1.81, within precincts standard deviation is 2.37.

are also somewhat stable over time.[38] Of the 750 estimated coefficients plotted in Figure 3, 72% have 95 % confidence intervals which include zero; 19% have 95 % confidence intervals that are positive; the remaining 9% have negative 95 % confidence intervals. Based on Theorem 1, we interpret these cross-sectional (and somewhat persistent) deviations from zero as evidence of heterogeneity in precinct-level biases.

We would like to explain the heterogeneity uncovered by Figure 3. One might hypothesize that a precinct's police force might be biased for or against whites depending on: the force's characteristics/composition; characteristics of the force's leadership; and/or the precinct's socio-economic conditions, including perhaps the nature and level of crime in that precinct. In this section we collect a number of proxies for these variables and correlate them with our measure of precinct bias. Unfortunately the publicly released data do not include an identifier for the officer making the stop, so our proxies will necessarily be limited to characteristics of the force leadership and of the precinct.

Force leadership information was gathered by focusing on patrol borough commanders. Patrol boroughs are larger units that encompass many precincts (about nine, on average). By searching through news articles, we collected the name, race, and gender of the 27 patrol borough commanders during the years covered by our data.[39] This information is used to create a precinct-year dummy variable which assumes value 1 if a precinct belonged to a patrol borough who was commanded by an african-american officer in that year. We can use this variable to assess whether precincts which are overseen by an african-american borough commander are more or less likely to be biased against pedestrian of a particular race. Figure 3 suggests that any effect, if it exists, is small. The blue dots represent precincts belonging to patrol boroughs commanded by an African American. The distribution of blue dots in each year is virtually indistinguishable from that of red diamonds. We conclude that there is no economically relevant correlation between our measure of bias and the race of the patrol

[38] A regression of the differential probability in a district on its first-year lag, controlling for year dummies, gives a coefficient of .043.

[39] Of these commanders, only 4 are African Americans and they were all appointed after 2005.

borough commander. Curiously, our measure of bias against African American pedestrians is negatively correlated with Mayor Bloomberg's precinct-level margin of victory in his three elections (see Figure B.1). To get a comprehensive picture of these correlations, we correlate our measure of bias with these and other variables measured at the precinct level, including: demographic and economic variables, prevalence of serious crimes (reported), graffiti and other measures of social capital.[40] Table 4 reports the results. In Column 1 the dependent variable is the 2003-2012 average of the differential probability of an African American being arrested in each precinct. In Columns 2 and 3, the dependent variable is the *annual* average of the differential probability of an African American being arrested in each precinct. In Column 3 we control for precincts fixed effects. We find (Table 4 columns 1 and 2) that our measure of bias against african-american pedestrians negatively correlates with the Bloomberg margin of victory and the median income in a precinct, while it positively correlates with the precinct's median age. Once we control for precincts fixed effects, however, we find (column 3) that only the total number of serious crimes (positively) correlates, at 10% significance level, with our measure of bias against african-american pedestrians. We interpret these results as indicating that we are able to account for some of the cross-sectional variation in our measure of bias through our regressors (column 1, with an $R^2 = 0.28$), but little of the within-precinct variation (column 3, $R^2 = 0.01$). We emphasize, however, that these correlations cannot be given a causal interpretation.

[40] Appendix C describes these variables and their sources; Table B.2 reports descriptive statistics; Table B.3 reports the pairwise correlations among these variables, including our measure of bias.

Table 4: Differential Probability of an African American being Arrested and Precinct Characteristics

Model	OLS	OLS	FE
Sample	Cross-section	Panel	Panel
	(1)	(2)	(3)
Margin of Bloomberg victory	0.033***	0.026***	-0.006
	(0.011)	(0.010)	(0.009)
Fraction of African American	0.023	0.013	
	(0.014)	(0.009)	
Median income	0.033***	0.026**	
	(0.012)	(0.012)	
Median age	-0.094*	-0.078	
	(0.054)	(0.056)	
Fraction of female	-0.053	-0.069	
	(0.060)	(0.049)	
Fraction of college degree	0.003	-0.011	
	(0.014)	(0.016)	
Serious crime	0.027	0.029	-0.025*
	(0.019)	(0.022)	(0.014)
Number of graffiti (per capita)	0.074	0.149	
	(0.210)	(0.200)	
Social capital (per capita)	-2.405	-2.706	
	(1.979)	(2.038)	
African American commander	1.282	0.491	-0.188
	(0.776)	(0.474)	(0.583)
Constant	3.416	4.494	-0.808
	(3.397)	(2.828)	(0.636)
Observations	75	750	750
Adj. R^2	0.285	0.107	0.014
Time FE	no	yes	yes

Notes. Estimates are on 75 precincts. The dependent variable is the differential probability of an African American being arrested. This is the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. In Column 1, the dependent variable is the average differential probability between 2003-2012, and *Margin of Bloomb. victory* is computed considering the 2001 elections. In Columns 2 and 3, the dependent variable is the precinct-year estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. Column 3 includes precinct fixed effects. *Margin of Bloomb. victory* is the difference between M. Bloomberg and the first running opponent M. Green, or F. Ferrer, or B. Thompson vote share in the 2001, 2005, 2009 elections, respectively.Missing years are computed using moving averages; *Fr. of African American* is the percentage of the population that is African American in the precinct in 2010; *Income* is the inflation adjusted median income, 2010 precinct average; *Age* is the median age, 2010 precinct average; *Fr. female* is the precinct average percentage of the population that is female in 2010; *Fr. college degree* is the precinct average percentage of the population aged 15-24 with a college degree in 2010; *Serious crimes* is the number of annual crimes (murders, rapes, robberies, fel. assaults, burglaries, grand larcenies, grand larceny autos) in each precinct divided by the precinct population in 2010 (in 1,000 habitants) for the years 1998, 2001, 2012; *Graffiti* is the number of annual graffiti in each precinct in 2011 divided by the precinct population in 2010 (in 1,000 habitants). ; *Soc.cap* is the total number of annual civic initiatives (education, emergency preparedness, environment, helping neighbors in need, strengthening communities) in each precinct, in 2011 divided by the precinct population in 2010 (in 1,000 habitants); *African American comm* is an indicator for African American commanding officers. All variables are described in Appendix C. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes" regressions additionally include year fixed effects (9 dummies). Standard errors are robust. Significance at the 10% (*), at the 5% (**), and at the 1% (***). All variables are described in Appendix C.
Source. Statistics for the City of New York, Years 2003-2012. Resident population from the 2010 Census data.

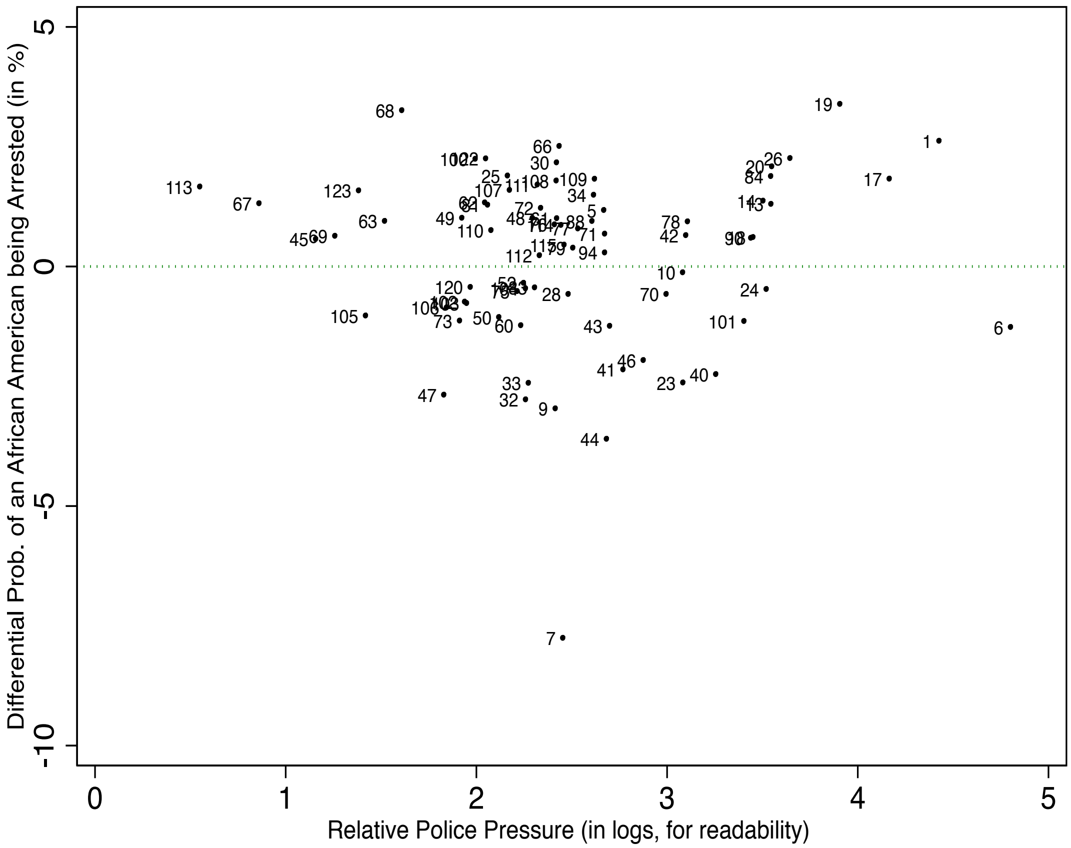
We close this section with an observation regarding the methodology for identifying bias. As mentioned in the introduction, most of the public debate considers the wide racial disparity in police pressure (see Figure 1, left-hand panel) as a strong clue, if not outright damning evidence, that the police act in a racially biased way. Our theoretical analysis (Theorems 1 and 2) has led us to conclude that bias, to the extent that it can be identified with the data at hand, ought to be inferred from disparities in arrest rates (see Figure 1, right-hand panel). One might ask whether the two measures, police pressure and arrest rates, are correlated. In other words, if we take arrest rates differential as a good proxy of bias, are these differentials related to differentials in police pressure? Figure 4 below suggests that this is not the case in our sample. The horizontal axis reports the ratio, computed at the precinct level, of the two columns in Figure 1, left panel. This is a measure of how much higher police pressure is on African American citizens, compared to whites. The vertical axis represents our measure of bias (difference between African American and white arrest rates, 10-year average, by precinct). The figure suggests that there is no correlation between these two measures across precincts. This graphical intuition is supported by the regression in Table B.4. This regression shows that relative police pressure is uncorrelated with our measure of bias. We believe this analysis can be interpreted as undermining the presumption that "where there's smoke there's fire," that is, that large disparities in police pressure are necessarily the result (or even the correlate) of police officer bias.

6 Discussion and Extensions

6.1 Hispanics

So far we have restricted the hit rate analysis to african-american and white pedestrians, setting Hispanics aside. We now extend our analysis to a sample of 4,413,566 stops and frisks of African Americans, Hispanics (black and white) and whites suspects. In this larger

Figure 4: **Differential Probability of an African American being Arrested and Relative Police Pressure**



Notes. The figure plots the differential probability of an African American being arrested (i.e., the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians) against the natural logarithm of relative police pressure: $\frac{\frac{Arrests of African Americans}{African American population}}{\frac{Arrests of Whites}{White population}}$ in each of the New York City police precinct.
Source. Statistics for the City of New York, Years 2003-2012.

sample, African Americans and Hispanics represent 56.1 % and 33.2 % of the stops, and suspects are arrested, on average, 6 % of the times.

In Table B.5 we augment our baseline specification regressing the indicator variable coding whether the stopped pedestrian was arrested on an indicator variable for african-american pedestrians and on an indicator for Hispanic pedestrians. Depending on the specification, Hispanic pedestrians who are stopped are between 0.12% and 0.15% less likely to be arrested compared to whites. This small difference in arrest rate between races, however, vanishes if we control for precincts fixed effects.[41]

6.2 Other Characteristics

So far we have focused on race and ethnicity. Other characteristics of the stop are available in our data such as gender, time of the stop, etc. To the extent that these are known to the officer before the stop, the theory predicts that the return to these searches should be similar provided that the search cost is similar across characteristics. The characteristics we focus on are: pedestrian gender, height, age, and time of the stop. In Table B.7 we compare arrest rates across these characteristics.

Across most of the characteristics the differentials in probability of arrest rates is smaller than 1% (though statistically significant in all cases) and thus, we believe, not easily detectable by the officers based on their own experience. Gender is the exception, with a differential probability of arresting a female exceeding 2% (7.6% compared to the 5.3% probability of arresting a male, see column 3). Thus stops of female pedestrians are on average substantially more productive than those of male pedestrians. If so, then why don't officers stop and frisk more women? (Currently, only 7% of all stops are of women). We conjecture that part of the reason may be that frisks of female suspects by male officers, though lawful, are even more controversial than those of male suspects.[42] Some male patrol officers might therefore shy away, at the margin, from searching women, or they may be advised by their supervisors to exercise restraint.[43] However this is only speculation; an analysis of this gender disparity is deferred to future work.

It is worth noting that conditioning on all these characteristics does not affect the main

[41] In Table B.6 we control for the type of crimes and find similar results.

[42] See "For Women in Street Stops, Deeper Humiliation," by Wendy Ruderman. *The New York Times,* published August 6, 2012.

[43] Male officers are a majority of officers. The news article cited in the previous footnote indicates that in 2011 more than 80% of patrol officers were men, and that it is deemed unsafe and often impractical for male officers to summon a female officer in order to conduct a frisk of a female pedestrian.

estimate of interest in this paper. Columns 11 and 12 in Table B.7 indicate that, even after controlling for these characteristics the estimated coefficients on "black" parallel those in columns 1 and 2, which in turn are in line with those in Table 3.[44] Therefore the main message of this paper is robust to including controls for other characteristics of the stop.

6.3 Alternative Outcome: Summons

Presumably, issuing a summons is a lesser or secondary goal for a police officer compared to an arrest. Nevertheless, issuing a summons does make the stop to some extent successful, or productive. Therefore, in Appendix B, Table B.8 we perform the hit rate test on the outcome "summons issued." The results are the opposite of Table 3: after controlling for precincts, the sign on "African American" switches and becomes *negative*.[45] The interpretation, according to the hit rate analysis, would be that officers are biased against African Americans in their decision to stop *if officers only cared about issuing a summons*. But, probably it is proper for officers to care both about issuing a summons and about making an arrest.

This reasoning leads us to add a dimension to the model: the rate at which police officers trade off arrests and summons. Mathematically, the officer's payoff from a stop can be conceptualized as follows:

$$\pi(\alpha) = \alpha \cdot \mathrm{I}_{\text{arrest}} + (1-\alpha) \cdot \mathrm{I}_{\text{summons}},$$

where $\mathrm{I}_{\text{arrest}}$ and $\mathrm{I}_{\text{summons}}$ are indicators taking value 1 if the pedestrian is arrested or issued a summons, respectively. If α is close to zero then the payoff $\pi(\alpha)$ will closely mimic the variable "summons;" vice versa, when α is close to one then the payoff $\pi(\alpha)$ will be close to

[44] The estimates of Table B.7 differ slightly from those of Table 3 because of the smaller sample in the latter table, which is due to missing values in pedestrians characteristics.

[45] Table B.9 in Appendix B corresponds to Table B.1 and gives a similar result: given a certain crime committed by a pedestrian of either race, officers are not using discretion in favor of whites when deciding whom to issue a summons to. In fact, African Americans appear to be issues summons less often than whites. Therefore, we see no evidence that any police discretion which might affect the outcome "summons" is biased against African Americans.

the variable "arrests."

Once the model is extended in this way, there are two unobserved parameters in the officer's objective function: one is the officer's possible bias against African Americans; the other is the parameter α. For the purpose of this paper, both parameters are taken to reflect the police officer's tastes or values.[46] Both parameters, arguably, have a normatively correct value: for bias this value is obviously zero, that is, no bias; for α we don't know it, so we will call it α^*. If the parameters in the officer's utility function differs from their "normatively correct" values, then the officer's behavior will lead to a disparity in hit rates, where the hit rates are computed on the variable $\pi(\alpha^*)$. But it will be difficult to determine whether the disparity is due to racial bias or to a normatively incorrect α (or both). Given this ambiguity, in the rest of the section we perform the following analysis. We stipulate as part of our null hypothesis that the police is unbiased, which implies that in equilibrium the hit rates on African American and whites need to be equalized on the variable $\pi(\alpha)$, whatever α is in the officer's mind (this prediction comes from the reasoning in Section 2). We then we look for the value $\overline{\alpha}$ which equalizes the hit rates on the variable $\pi(\alpha)$. The resulting $\overline{\alpha}$ is interpreted as the taste parameter which, given the null of no bias, rationalizes the officers' observed stop behavior. After this exercise, which is purely an exercise in the identification of unobserved taste parameters, we go back and ask the question of whether the estimated parameter $\overline{\alpha}$ is normatively acceptable. If it is not, then we conclude that the police is behaving in a way that departs from the norm—even though we will not be able to distinguish whether the departure is that the officer is racially biased, or that the officer's α differs from the normatively acceptable α^*, or both.

Let's proceed. Fix any α which we interpret as the known taste parameter in the police utility function. What happens if we run the hit rate test on the outcome variable $\pi(\alpha)$? This depends on the chosen value of α. If α is close to zero then the payoff $\pi(\alpha)$ will closely mimic the variable "summons," and we know from Table B.8 in Appendix B that the hit rates

[46] In a broader treatment, one could argue that α might be determined by career incentives, for example.

are not equalized. Conversely, if α is close to one then the payoff $\pi(\alpha)$ will closely replicate the variable "arrests," and we know from Table 3 that hit rates are not equalized either. We have performed a search for the threshold value $\overline{\alpha}$ that equalizes the arrest rates on $\pi(\alpha)$ across races (analysis not reported). This threshold value $\overline{\alpha}$ turns out to be around 0.8.[47] This means that the observed police behavior is consistent with that of a police force which is unbiased and uses an $\alpha \approx 0.8$. Also, the observed police behavior is consistent with that of a police force which is not biased against African Americans and uses an $\alpha > 0.8$.[48] A value of 0.8 implies that the police values each arrest equal to about four summons. We regard this "conversion rate" as normatively not unacceptable—although of course readers are free to make their own judgments. If that rate of 4 to 1 is found normatively acceptable, then we conclude that the police is behaving in a way that is observably equivalent to one whose tastes are in line with the normatively correct values, that is, zero bias and a normatively acceptable $\alpha^* \approx 0.8$.

6.4 Using Information About Frisks

Until now the analysis has proceeded on the assumption that the stop is the unit of analysis. In this section we use the information about what happens within the stop, and more precisely, about the fraction of stops which develop into frisks. About 53.7% of stops of African Americans develop into frisks, as opposed to 39.3% of whites.[49] Thus, stops of African American pedestrians are more likely to develop into a frisk. Frisks of african-american pedestrians are less likely to be "productive," compared to a white.[50] It is tempting to use this disparity in the productivity of frisks to infer bias in the frisk choice. To pursue this line of inquiry, we need a a model that distinguishes stops from frisks.

We propose the following heuristic model. Let us assume that the police officer first de-

[47] Tables B.10 and B.11 report the main results when the dependent variable is $\pi(0.8)$. By construction, the estimated coefficients on "African American" in columns 6 and 7 of Tables B.10 are not significantly different from zero, which means that arrest rates are equalized.

[48] This should be expected: if α is close to one then we are back to the analysis in Section 5.

[49] Table B.12 reports some descriptive statistics.

[50] About 9% of frisks of African Americans are associated with an arrest, compared with 13% of whites.

cides whether to expend cost t_1 and stop a pedestrian. Then, after having interviewed the pedestrian and learned some unobserved (to us) characteristic u, the officer will be able to immediately make an arrest, or may choose to frisk the pedestrian at an additional cost t_2, or may choose to let the pedestrian go. Let us focus on the frisk decision. In this model, the officer uses knowledge of u, which is gained after the stop, in his decision of whether to frisk. Only those pedestrians with u's suggesting a sufficiently high likelihood of a crime will be frisked. Note that in the equilibrium of this model it is possible to have two characteristics u_1 and u_2, both of which lead to a frisk, but with u_2 yielding a slightly higher probability of a successful frisk. Such small differences in the returns to frisk are not be arbitraged away in equilibrium because officers need to pay t_1 in order to learn whether a pedestrian has characteristic u_1 or u_2.

We have just argued that, in a model of sequential interdiction of this type, some types u are "inframarginal" on the success rate of frisks. In such a case, it is well known (see e.g. Persico and Todd [19]) that hit rates test cannot be performed on frisks. However, the test can still be performed on stops. This is because the decision to stop a pedestrian is still subject to the usual arbitrage argument, because this decision it is based on outwardly observable characteristics which are learned at zero cost. Therefore the hit rate test developed in Section 2 applies *to the productivity of stops*. We now develop this test within a simplified version of the model in Section 2, where for expositional simplicity we omit the heterogeneity in outwardly visible characteristics c and assume that all officers have the same bias $B(p) \equiv \beta_A - \beta_W$.[51] However, the analysis will extend the one carried out in Section 2 in a key dimension: it allows for frisks.

In the model sketched out above, the stop of a white pedestrian has an expected payoff of

$$-t_1 + \Pr(\text{arrest}|\text{W}) \cdot \beta_W - t_2 \Pr(\text{frisk}|\text{W}), \tag{3}$$

where the parameter β_W captures the officer's satisfaction from arresting a white pedestrian.

[51] The results would not change if we introduced all the richness of parameters allowed for in Section 2.

The stop of an african-american pedestrian has an expected payoff of

$$-t_1 + \Pr(\text{arrest}|\text{A}) \cdot \beta_A - t_2 \Pr(\text{frisk}|\text{A}), \tag{4}$$

where the parameter β_A captures the officer's satisfaction from arresting an african-american pedestrian. Within this simple model, an officer is deemed to be biased against African Americans if $\beta_A > \beta_W$.

In equilibrium the officers need to be indifferent between stopping a white and an African American. Therefore expressions (3) and (4) must be equal. Under the null that $\beta_A \leq \beta_W$ which means that officers are not biased against African Americans, the equality of (3) and (4) implies

$$\Pr(\text{arrest}|\text{W}) \cdot \beta_W - t_2 \Pr(\text{frisk}|\text{W}) \leq \Pr(\text{arrest}|\text{A}) \cdot \beta_W - t_2 \Pr(\text{frisk}|\text{A}),$$

with equality holding if $\beta_A = \beta_W$, that is, when officers are unbiased. This inequality can be rewritten as follows:

$$\frac{t_2}{\beta_W} \leq \frac{[\Pr(\text{arrest}|\text{A}) - \Pr(\text{arrest}|\text{W})]}{[\Pr(\text{frisk}|\text{A}) - \Pr(\text{frisk}|\text{W})]}. \tag{5}$$

An additional implication of the same model is that in equilibrium the officers must be indifferent between stopping a pedestrian (of either race) or not (refer to Section 2). Equating expression (3) to zero yields

$$\frac{t_1}{\beta_W} = \Pr(\text{arrest}|\text{W}) - \frac{t_2}{\beta_W} \Pr(\text{frisk}|\text{W}). \tag{6}$$

We can use the data to put some numbers to these conditions. For equation (5), we estimated in Table 3 that $\Pr(\text{arrest}|\text{A}) - \Pr(\text{arrest}|\text{W}) \approx 0.003$; moreover, from Table B.12 we have

$\Pr(\text{frisk}|\text{A}) - \Pr(\text{frisk}|\text{W}) \approx 0.10$. Therefore equation (5) reads

$$\frac{t_2}{\beta_W} \leq \frac{0.003}{0.10} = 0.03, \tag{7}$$

and equality holds if officers are unbiased. Turning to equation (6), we have

$$\frac{t_1}{\beta_W} = 0.06 - \frac{t_2}{\beta_W} 0.39 \geq 0.06 - 0.03 \cdot 0.39 = 0.04, \tag{8}$$

where 0.06 is the constant from Table 3, 0.39 represents the average probability that a stop of a white pedestrian develops into a frisk, and the inequality comes from (7). Combining equations (7) and (8) yieds

$$\frac{t_2}{t_1} \leq \frac{3}{4}, \tag{9}$$

with equality holding if police are unbiased. This means that the additional cost of a frisk is perceived by the officer as smaller than the cost of a stop without frisk. Although this is not a direct test of bias, this calculation could provide one if we had reason to believe, for example, that condition (9) does not hold in reality. Then the model would imply that police officers are biased against African Americans.

7 Conclusions

New York City's "stop and frisk" program disproportionately impacts minorities. At the same time, former New York City's police commissioner William Bratton said: "Stop-and-frisk is not something you can stop. It is an absolutely basic tool of American policing."[52] If "stop and frisk" cannot be stopped, then it becomes especially important to ensure that this program is carried out in a racially unbiased way. To this end, we have argued that a strong theoretical foundation is needed which can rigorously identify two distinct sources of bias: bias at the level of the police officer making the stop decisions, and bias at the level of the police chief allocating manpower across precincts. Previous research offered positive

[52] *The Wall Street Journal*, "The Real Cures for Gun Violence" by David Feith. January 19-20, 2013.

identification results regarding officer bias; this paper adds a new, and negative, identification result for police chief bias.

Ten years of data from NYPD's "stop and frisk program" are analyzed in light of this theoretical framework. White pedestrians are found to be slightly less likely than african-american pedestrians to be arrested conditional on being stopped. We interpret this fact as evidence that the officers making the stops are *on average* not biased against African Americans relative to whites, because the latter are being stopped despite being a "less productive stop" for a police officer. However, the average masks heterogeneity among precincts. In some precincts the police force is, according to our measure, biased against african-american pedestrians. The opposite is true in other precincts. We have correlated this measure of bias with precinct characteristics and, at least in the cross section, found some significant correlations with median income and Bloomberg's margin of victory in mayoral elections.

Our measure of bias turns out to be uncorrelated, across precincts, with the intensity of relative police pressure on african-american residents (relative to pressure on whites). This is interesting because the public, and many experts, consider the wide racial disparity in police pressure as a strong clue, if not outright evidence, that the police who stop and frisk are acting in a racially biased way. However, the lack of correlation suggests to us that this presumption may be incorrect. Large disparities in police pressure need not correlate with police officer bias; put more colorfully, there can be smoke without fire.

An important caveat: our analysis is based on the assumption that the decision to arrest (as opposed to the stop-frisk decision) is not tainted by police bias. We have tested this assumption to the extent possible with the data at hand, and found no evidence pointing to its rejection. However, we feel that this assumption deserves further scrutiny.

Our results cannot be interpreted as proving that the stop and frisk program is lawful. If is possible that the program may be unlawful in other ways, for example, that many of its searches may not arise from a reasonable suspicion.

References

[1] Alpert, G., MacDonald, J. H., and Dunham, R. G. (2005), "Police Suspicion and Discretionary Decision Making During Citizen Stops." *Criminology*, 43, 407–434.

[2] Anwar, Shamena and Hanming Fang (2006), "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence. "*American Economic Review* 96(1): 127-151.

[3] Ayres, Ian, and Joel Waldfogel (1994), "A Market Test for Race Discrimination in Bail Setting." *Stanford Law Review* 46(5): 987-1047.

[4] Ayres, Ian (2002), "Outcome tests of racial disparities in police practices." *Justice Research and Policy* 4, pp. 131-142.

[5] Becker, Gary (1957), "The economics of discrimination." The University of Chicago Press, Chicago.

[6] Childers, Sean (2012), "Discrimination During Traffic Stops: How an Economic Account Justifying Racial Profiling Falls Short." *New York University Law Review* 87(4): 1025-59.

[7] Dharmapala Dhammika & Ross Stephen L, 2004, "Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence." *The B.E. Journal of Economic Analysis & Policy*, vol. 3(1), pages 1-23, September.

[8] Fyfe, James J. and Robert Kane (2006), "Bad Cops: A Study of Career-Ending Misconduct Among New York City Police Officers." Document No. 215795, Final report on National Institutio of Justice Award 96-IJ-CX-0053, September 2006.

[9] Gershman, Bennett L. (2000), "Use of Race in "Stop-and-Frisk": Stereotypical Beliefs Linger, But How Far can the Police Go?" *Pace Law Faculty Publications*. Paper 60. http://digitalcommons.pace.edu/lawfaculty/60

[10] 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(479), 813-823.

[11] Hernández-Murillo, Rubén and John Knowles (2004), "Racial Profiling or Racist Policing? Testing in Aggregated Data." *International Economic Review* 45(3): 959-989.

[12] Knowles, John, Persico, Nicola, and Todd, Petra (2001), "Racial bias in motor vehicle searches: theory and evidence." *Journal of Political Economy* 109(1): 203–229.

[13] NYCLU (2011) "Stop-and-Frisk, NYCLU BRIEFING." *http://www.nyclu.org/files/publications/NYCLU_2011_Stop-and-Frisk_Report.pdf.*

[14] Persico, Nicola (2009), "Racial Profiling? Detecting Bias Using Statistical Evidence." *Annual Review of Economics*, Vol. 1, pp. 229-254.

[15] Persico, Nicola (2002), "Racial Profiling, Fairness, and Effectiveness of Policing." *American Economic Review* 92(5): 1472-97.

[16] Persico, Nicola and David Castleman (2005), "Detecting Bias: Using Statistical Evidence to Establish Intentional Discrimination in Racial Profiling Cases." *University of Chicago Legal Forum*, pp. 217-35.

[17] Persico, Nicola and Petra Todd (2005), "Passenger Profiling, Imperfect Screening, and Airport Security." *American Economic Review Papers and Proceedings* 95(2): 127-131.

[18] Persico, Nicola and Petra Todd (2006), "Using Hit Rates to Test for Racial Bias in Law Enforcement: Vehicle Searches in Wichita." *The Economic Journal* 116 : F351-F367.

[19] Persico, Nicola and Petra Todd (2008), "The hit rates test for racial bias in motor-vehicle searches." *Justice Quarterly*. 25(1): 37-53.

[20] Ridgeway, Greg (2007), "Analysis of Racial Disparities in the New York Police Department's Stop, Question, and Frisk Practices. Santa Monica." CA: RAND Corporation, 2007. http://www.rand.org/pubs/technical_reports/TR534.

[21] Sanga, Sarath (2009), "Reconsidering Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 117(6): 1155-1159

[22] Smith, M. R., Makarios, M., and Alpert, G. P. (2006), "Differential Suspicion: Theory Specification and Gender Effects in the Traffic Stop Context." Justice Quarterly, 23, 271–295.

[23] Todd, Petra (2006), "Testing for Racial Bias in Law Enforcement." *Palgrave Dictionary of Economics.*

[24] USCCR (2010), "Police Practices and Civil Rights in New York City." *http://www.usccr.gov/pubs/nypolice/ch5.htm*

[25] Whitney, Melissa (2008), "Statistical Evidence of Racial Profiling in Traffic Stops and Searches: Rethinking the Use of Statistics to Prove Discriminatory Intent." *Boston College Law Review* 49, pp. 263-99

Appendices

A The Theory of Pedestrian and Officer Behavior, and a Test for Officer Bias

Condition the analysis to precinct i, so that all variables in this section are precinct-specific. We want to allow for officer heterogeneity within the district, and at the same time we want to allow for additions (or subtractions) to the mass of police officers allocated to the precinct. This raises the question of how the differential officers should affect the pre-existing distribution of officer characteristics within the precinct. For simplicity, we assume that officers are blank slates, and that they take on their characteristics (possible bias, cost of searching, etc.) by drawing from a precinct-specific distribution. This modelling device avoids the need for keeping track of changing officer characteristics in a precinct as the precinct's manpower is changed.

Let r denote the race of the pedestrian, which is assumed to be observable by the police. Without loss of generality, assume that there are pedestrians of two races, either African American (A) or white (W). Other characteristics that are costlessly observable by the police are represented by the number $c \in \{1, \ldots, C\}$. These characteristics might represent such things as the age or gender of the pedestrian. From the police viewpoint, a pedestrian is characterized by two variables, r and c. Let $N^{r,c}$ denote the number of pedestrians belong to group (r, c).

We assume that the police can distinguish between pedestrian groups (r, c), but cannot detect pedestrian heterogeneity within (r, c) groups. (We revisit this assumption in Section 6.4). Two sources of unobserved pedestrian heterogeneity within groups are the values to an individual of committing a crime and the costs of being detected. Let v represent the value of committing a crime. If the crime is detected, the payoff to the pedestrian is $v - j$, where

j captures the cost of being detected. We allow v and j to vary across individuals within a (r, c) group and denote the joint conditional distribution of v and j by cdf $F_{r,c}(v, j)$.

A pedestrian in group (r, c) makes a binary decision: commit a crime or not. Just as pedestrians may differ in their costs and benefits, we allow police officers to be heterogeneous in three respects: their search capacity, their per-search cost, and their racial bias. We assume that there is a mass P of police officers, which after being assigned to the precinct, draw a type p from a uniform distribution on $[0, 1]$. Each police officer p is endowed with a search capacity of S_p and a per-search cost t_p. If a search does not yield any evidence of crime, then we term the search unsuccessful and assume that the police officer incurred the cost of search without any benefit. We introduce the potential for police bias by allowing the benefit that the police derives from a successful search to depend on the race of the pedestrian. Suppose the benefit to a police officer p of finding a criminal of race W is y_p^W and the benefit of finding criminal of race A is $y_p^A = y_p^W + B(p)$. We say that police are biased against African American pedestrians if $B(p) > 0$ for all p, against whites if $B(p) < 0$ for all p, and unbiased if $B(p) = 0$ for all p. If no search is conducted, there is a zero payoff. As described, this setup accommodates police heterogeneity in intensity of bias. However, we rule out environments in which $B(p)$ changes sign as p varies, i.e., where some policemen are biased against whites and some are biased against African Americans. Below, we propose a test for inferring the sign of $B(p)$.

A member of group (r, c) with given v, j, who commits a crime and expects σ members of his group to be searched receives an expected payoff

$$u_{r,c}(v, j, \sigma) = v - j \cdot \frac{\sigma}{N^{r,c}}.$$

When this payoff exceeds zero, the individual will choose to commit a crime. Let $K_{r,c}(v, j, \sigma)$ be an indicator function that equals 1 if the individual chooses to commit a crime. The fraction of pedestrians within each group (r, c) who commit a crime is given by

$$K^{r,c}(\sigma) = \int K_{r,c}(v, j, \sigma)\ dF_{r,c}(v, j).$$

The function $K^{r,c}(\sigma)$ summarizes the crime rate in group (r, c) when the police search that group with intensity σ. One can think of this function as a response function or as the supply of crime.

Denote by $S_p(r, c)$ the number of searches that officer p decides to devotes to group (r, c). The total number of searches of members of group (r, c) is obtained by aggregating the behavior of all police officers:

$$S(P, r, c) = P \int_0^1 S_p(r, c)\ dp.$$

Officer p's expected payoff is the sum of the expected payoffs of all his searches, given by

$$\sum_{r,c} S_p(r, c) \left[y_p^r \cdot K^{r,c}(S(P, r, c)) - t_p \right], \tag{10}$$

which depends on the officers perceived benefit from apprehending someone of race r (y_p^r) as well as the officer's costs of search (t_p).

Persico and Todd [18] show that a Nash equilibrium of this game exists and is generically unique. Let $[S^*(P, r, c)]_{r,c}$ be a vector denoting the search intensities at the Nash equilibrium. Suppose groups (r, c) and (r', c') are searched in equilibrium. Then, there must be a p and a p' such that

$$y_p^r K^{r,c}(S^*(P, r, c)) - t_p \geq y_p^{r'} K^{r',c'}(S^*(P, r', c')) - t_p,$$
$$y_{p'}^r K^{r,c}(S^*(P, r, c)) - t_{p'} \leq y_{p'}^{r'} K^{r',c'}(S^*(P, r', c')) - t_{p'}.$$

If $r = r'$, or if the police are unbiased then $y_p^r = y_p^{r'}$ for all p's, and so the two inequalities can only be simultaneously satisfied if

$$K^{r,c}\left(S^*\left(P,r,c\right)\right)=K^{r',c'}\left(S^*\left(P,r',c'\right)\right). \tag{11}$$

If the police are biased against race r then $y^r_{p'} > y^{r'}_{p'}$, and so the second inequality can only be satisfied if the crime rates are such that

$$K^{r,c}\left(S^*\left(P,r,c\right)\right)<K^{r',c'}\left(S^*\left(P,r',c'\right)\right).$$

Note that in our model the hit rate, i.e., the likelihood that a search of group (r,c) yields a find, coincides with that group's crime rate $K^{r,c}$.Thus, the implications on the crime rate translate into testable implications on the hit rates. This observation yields Theorem 1.

For use in Section 2 we now define the function

$$K^r_i\left(P\right)=K^{r,c}\left(S^*\left(P,r,c\right)\right).$$

Note that the left hand side lacks a c argument, which makes sense because, by (11), we have $K^{r,c}\left(S^*\left(P,r,c\right)\right)=K^{r,c}\left(S^*\left(P,r,c'\right)\right)$ for all c,c'. By the same argument, it is proper to omit reference to characteristics c in the model of Section 2.

B Extra Tables and Figures

Table B.1: Arrest Made, Controlling for the Type of Crime

Model	OLS	FE	FE
	(1)	(2)	(3)
African American	-0.551	0.233	0.225
	(0.418)	(0.217)	(0.214)
Mean outcome		5.76%	
Fraction of African American pedestrians		84%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	2,496,267	2,496,267	2,496,267
Cluster SE	yes	yes	yes
Time FE	yes	yes	yes
Precincts FE	no	yes	yes
Crime FE	yes	yes	yes
Time FE · Precincts FE	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of being arrested conditional on being stopped in New York City (in %). *African American* is an indicator variable coding the pedestrian's race. All the regressions include 13 indicators of the type of crime representing 95% of the crimes (as recorded by the officer on Form UF-250). To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Standard errors are clustered at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.2: Descriptive Statistics of Precinct Characteristics

	Mean	sd	p25	p50	p75	n
Differential Probability of an African American being Arrested	.21	1.8	-1	.67	1.4	75
Margin of Bloomberg victory	-7.3	34	-33	-16	18	75
Fraction of African American	27	26	5.2	16	45	75
Income	56	26	36	52	68	75
Age	36	4.3	32	35	38	75
Fraction of female	52	2.6	51	53	54	75
Fraction of college degree	21	16	9.7	16	25	75
Serious crime	22	27	14	18	22	75
Graffiti	1.3	.93	.66	1.1	1.6	75
Social capital	.12	.32	.0057	.026	.073	75
African American commander	.18	.25	0	0	.42	75

Notes. *Differential Probability of an African American being Arrested* is the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. *Margin of Bloomb. victory* is the difference between M. Bloomberg and the first running opponent M. Green, or F. Ferrer, or B. Thompson vote share in the 2001, 2005, 2009 elections, respectively.Missing years are computed using moving averages; *Fr. of African American* is the percentage of the population that is African American in the precinct in 2010; *Income* is the inflation adjusted median income, 2010 precinct average; *Age* is the median age, 2010 precinct average; *Fr. female* is the precinct average percentage of the population that is female in 2010; *Fr. college degree* is the precinct average percentage of the population aged 15-24 with a college degree in 2010; *Serious crimes* is the number of annual crimes (murders, rapes, robberies, fel. assaults, burglaries, grand larcenies, grand larceny autos) in each precinct divided by the precinct population in 2010 (in 1,000 habitants) for the years 1998, 2001, 2012; *Graffiti* is the number of annual graffiti in each precinct in 2011 divided by the precinct population in 2010 (in 1,000 habitants). ; *Soc.cap* is the total number of annual civic initiatives (education, emergency preparedness, environment, helping neighbors in need, strengthening communities) in each precinct, in 2011 divided by the precinct population in 2010 (in 1,000 habitants); *African American comm* is an indicator for African American commanding officers. All variables are described in Appendix C.
Source. Statistics for the City of New York, Years 2003-2012.

Table B.3: Cross-correlation table of Differential Probability of an African American being Arrested and Precinct Charactheristics

Variables	Diff. Prob. of. Americ. Arrest.	Margin of Bloomb. victory	Fr. of A.A.	Income	Age	Fr. female	Fr. degree	Serious crime	Graffiti	Soc. cap.
Diff. Prob. of. A. A.	1.00									
Margin of Bloomb. victory	0.35	1.00								
	(0.00)									
Fr. of A.A.	-0.21	-0.75	1.00							
	(0.07)	(0.00)								
Income	0.39	0.38	-0.45	1.00						
	(0.00)	(0.00)	(0.00)							
Age	0.15	0.53	-0.26	0.44	1.00					
	(0.19)	(0.00)	(0.02)	(0.00)						
Fr. female	-0.20	-0.26	0.34	-0.24	-0.04	1.00				
	(0.09)	(0.03)	(0.00)	(0.04)	(0.72)					
Fr. college degree	0.24	0.26	-0.53	0.74	0.31	-0.17	1.00			
	(0.04)	(0.02)	(0.00)	(0.00)	(0.01)	(0.15)				
Serious crimes	0.06	-0.08	-0.04	0.08	-0.01	-0.09	0.20	1.00		
	(0.60)	(0.51)	(0.75)	(0.50)	(0.91)	(0.46)	(0.08)			
Graffiti	-0.09	-0.28	0.02	-0.21	-0.37	-0.02	0.05	0.16	1.00	
	(0.44)	(0.02)	(0.90)	(0.07)	(0.00)	(0.87)	(0.66)	(0.17)		
Soc. cap.	-0.05	-0.01	-0.20	0.26	0.04	-0.12	0.32	0.80	0.12	1.00
	(0.64)	(0.95)	(0.08)	(0.03)	(0.74)	(0.29)	(0.01)	(0.00)	(0.32)	
A.A. comm.	0.03	-0.43	0.33	-0.16	-0.30	0.16	-0.16	-0.05	0.19	-0.13
	(0.78)	(0.00)	(0.00)	(0.16)	(0.01)	(0.18)	(0.17)	(0.69)	(0.10)	(0.26)

Notes. Correlations (and *p-values* for statistical significance in parenthesis) for 75 precincts. *Diff. Prob. of A. A.* is the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. *Margin of Bloomb. victory* is the difference between M. Bloomberg and the first running opponent M. Green, or F. Ferrer, or B. Thompson vote share in the 2001, 2005, 2009 elections, respectively.Missing years are computed using moving averages; *Fr. of A.A.* is the percentage of the population that is African American in the precinct in 2010; *Income* is the inflation adjusted median income, 2010 precinct average; *Age* is the median age, 2010 precinct average; *Fr. female* is the precinct average percentage of the population that is female in 2010; *Fr. college degree* is the precinct average percentage of the population aged 15-24 with a college degree in 2010; *Serious crimes* is the number of annual crimes (murders, rapes, robberies, fel. assaults, burglaries, grand larcenies, grand larceny autos) in each precinct divided by the precinct population in 2010 (in 1,000 habitants) for the years 1998, 2001, 2012; *Graffiti* is the number of annual graffiti in each precinct in 2011 divided by the precinct population in 2010 (in 1,000 habitants). ; *Soc.cap* is the total number of annual civic initiatives (education, emergency preparedness, environment, helping neighbors in need, strengthening communities) in each precinct, in 2011 divided by the precinct population in 2010 (in 1,000 habitants); *African Americans comm* is an indicator for African American commanding officers. All variables are described in Appendix C.
Source. Statistics for the City of New York, Years 2003-2012.

Table B.4: Differential Probability of an African American being Arrested and Relative Police Pressure

Model	OLS	OLS	FE
Sample	Cross-section	Panel	Panel
	(1)	(2)	(3)
Relative Police Pressure	0.010	0.006	-0.021
	(0.013)	(0.005)	(0.017)
Constant	0.078	0.143	1.059**
	(0.303)	(0.144)	(0.463)
Differential Probability of an African American being Arrested		0.242	
Relative police pressure		17.19	
Observations	75	750	750
Adj. R^2	0.009	0.001	0.029
Time FE	no	no	yes

Notes. Estimates are on 75 precincts. The dependent variable is the differential probability of an African American being arrested. This is the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. *Relative police pressure* is $(\frac{\frac{ArrestsofAfricanAmericans}{AfricanAmericanpopulation}}{\frac{ArrestsofWhites}{Whitepopulation}})$ in New York City. Column 3 includes precinct fixed effects. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes" regressions additionally include year fixed effects (9 dummies). Standard errors are robust. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012. Resident population from the 2010 Census data.

Table B.5: Arrest Made Adding Hispanics

Model	OLS	OLS	OLS	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
African American	-0.420***	-0.436***	-0.436	0.279***	0.247***	0.247	0.235
	(0.037)	(0.037)	(0.469)	(0.043)	(0.043)	(0.187)	(0.185)
Hispanic	-0.119***	-0.149***	-0.149	-0.004	-0.029	-0.029	-0.000
	(0.039)	(0.039)	(0.344)	(0.043)	(0.043)	(0.164)	(0.161)
Constant	6.140***						
	(0.034)						
Mean outcome				5.86%			
Fraction of African American				56.1%			
Fraction of Hispanic				33.2%			
P-value of $H_0 : u_i = 0$				0.001	0.001	0.001	0.001
Number of precincts				76	76	76	76
Observations	4,413,566	4,413,566	4,413,566	4,413,566	4,413,566	4,413,566	4,413,566
Cluster SE	no	no	yes	no	no	yes	yes
Time FE	no	yes	yes	no	yes	yes	yes
Precincts FE	no	no	no	yes	yes	yes	yes
Time FE · Precincts FE	no	no	no	no	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of being arrested conditional on being stopped in New York City (in %). *African American (Hispanic)* is an indicator variable for African American (Hispanic) pedestrian. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Columns 3, 5-7, shows show clustered standard errors at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.6: Arrest Made, Controlling for the Type of Crime Adding Hispanics

Model	OLS	FE	FE
	(1)	(2)	(3)
African American	-0.582	0.138	0.140
	(0.424)	(0.196)	(0.194)
Hispanic	-0.095	0.046	0.069
	(0.322)	(0.172)	(0.169)
Mean outcome		5.84%	
Fraction of African American pedestrians		56.4%	
Fraction of Hispanic pedestrians		33.1%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	3,733,833	3,733,833	3,733,833
Cluster SE	yes	yes	yes
Time FE	yes	yes	yes
Precincts FE	no	yes	yes
Crime FE	yes	yes	yes
Time FE · Precincts FE	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of being arrested conditional on being stopped in New York City (in %). *African American (Hispanic)* is an indicator variable for African American (Hispanic) pedestrian. All the regressions include 13 indicators of the type of crime representing 95% of the crimes (as recorded by the officer on Form UF-250). To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Standard errors are clustered at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.7: Other Pedestrians Characteristics

Model	OLS (1)	FE (2)	OLS (3)	FE (4)	OLS (5)	FE (6)	OLS (7)	FE (8)	OLS (9)	FE (10)	OLS (11)	FE (12)
African American	-0.427	0.359*									-0.332	0.461**
	(0.475)	(0.210)									(0.468)	(0.202)
Female			2.315***	2.265***							2.358***	2.343***
			(0.287)	(0.259)							(0.290)	(0.258)
Above 6 feet					0.477***	0.348***					0.634***	0.507***
					(0.069)	(0.048)					(0.065)	(0.045)
Above 18-yrs							0.776***	0.473***			0.783***	0.518***
							(0.180)	(0.154)			(0.174)	(0.150)
Stop between 7pm-6am									-0.969***	-0.933***	-0.984***	-0.942***
									(0.153)	(0.159)	(0.151)	(0.158)
Constant	5.862***	6.750***	5.323***	6.856***	5.385***	6.938***	4.846***	6.579***	5.985***	7.417***	5.273***	6.270***
	(0.453)	(0.303)	(0.524)	(0.275)	(0.541)	(0.278)	(0.567)	(0.322)	(0.562)	(0.290)	(0.469)	(0.334)
P-value of $H_0 : u_i = 0$		0.001		0.001		0.001		0.001		0.001		0.001
Prob. of Arrest	5.843											
Fraction of African American	0.840											
Fraction of Female			0.0742									
Fraction of Pedestrians above 6 feet					0.613							
Fraction of non-juvenile stops							0.847					
Fraction of night stops									0.459			
Number of precincts		75		75		75		75		75		75
Observations	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774	2,892,774

Notes. Estimates are on 76 precincts. The dependent variable is the probability of being arrested conditional on being stopped in New York City (in %). *African American* is an indicator variable coding the pedestrian's race; *Female* is an indicator variable coding the pedestrian's gender; *Above 6 feet* is an indicator variable for pedestrians above 6 feet height; *Above 18-yrs* is an indicator variable coding whether the pedestrian is 18 years old or more race; *Stop between 7pm-6am* is an indicator variable for night stops. To control for possible time trend in the dependent variable and precincts specific characteristics regressions additionally include year fixed effects (9 dummies); even columns include precincts fixed effects (75 dummies). *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Cluster adjusted at precinct level standard error are reported in parenthesis. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.8: Summons Issued

Model	OLS	OLS	OLS	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
African American	0.070*	0.095**	0.095	-1.753***	-1.736***	-1.736***	-1.721***
	(0.038)	(0.038)	(0.360)	(0.047)	(0.047)	(0.295)	(0.276)
Constant	6.122***						
	(0.035)						
Mean outcome				6.18%			
Fraction of African American				84%			
P-value of $H_0 : u_i = 0$				0.001	0.001	0.001	0.001
Number of precincts				76	76	76	76
Observations	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865
Cluster SE	no	no	yes	no	no	yes	yes
Time FE	no	yes	yes	no	yes	yes	yes
Precincts FE	no	no	no	yes	yes	yes	yes
Time FE · Precincts FE	no	no	no	no	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of a summons being issued conditional on being stopped in New York City (in %). *African American* is an indicator variable coding the pedestrian's race. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Columns 3, 5-7, shows show clustered standard errors at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.9: Summons Issued, Controlling for the Type of Crime

Model	OLS (1)	FE (2)	FE (3)
African American	-0.254	-1.464***	-1.487***
	(0.311)	(0.293)	(0.276)
Mean outcome		6.13%	
Fraction of African American pedestrians		84%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	2,496,267	2,496,267	2,496,267
Cluster SE	yes	yes	yes
Time FE	yes	yes	yes
Precincts FE	no	yes	yes
Crime FE	yes	yes	yes
Time FE · Precincts FE	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of a summons being issued conditional on being stopped in New York City (in %). *African American* is an indicator variable coding the pedestrian's race. All the regressions include 13 indicators of the type of crime representing 95% of the crimes (as recorded by the officer on Form UF-250). To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Standard errors are clustered at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).

Table B.10: Arrest Made and Summons Issued (Weighted)

Model	OLS	OLS	OLS	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
African American	-0.322***	-0.331***	-0.331	-0.048	-0.063*	-0.063	-0.072
	(0.030)	(0.030)	(0.379)	(0.037)	(0.037)	(0.177)	(0.171)
Constant	6.136***						
	(0.028)						
Mean outcome				5.86%			
Fraction of African American				84%			
P-value of $H_0 : u_i = 0$				0.001	0.001	0.001	0.001
Number of precincts				76	76	76	76
Observations	2,947,867	2,947,867	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865
Cluster SE	no	no	yes	no	no	yes	yes
Time FE	no	yes	yes	no	yes	yes	yes
Precincts FE	no	no	no	yes	yes	yes	yes
Time FE · Precincts FE	no	no	no	no	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is ($\pi(\alpha) = \alpha \cdot \mathrm{I}_{\mathrm{arrest}} + (1-\alpha) \cdot \mathrm{I}_{\mathrm{summons}}$) the weighted sum of the probability of being arrested and the probability of a summons being issued conditional on being stopped in New York City (in %). The weights $(\alpha, 1-\alpha)$ are .8 and .2. *African American* is an indicator variable coding the pedestrian's race. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Columns 3, 5-7, shows show clustered standard errors at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.11: Arrest Made and Summons Issued, Controlling for the Type of Crime (Weighted)

Model	OLS (1)	FE (2)	FE (3)
African American	-0.492	-0.107	-0.118
	(0.333)	(0.170)	(0.166)
Mean outcome		5.86%	
Fraction of African American pedestrians		84%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	2,496,267	2,496,267	2,496,267
Cluster SE	yes	yes	yes
Time FE	yes	yes	yes
Precincts FE	no	yes	yes
Crime FE	yes	yes	yes
Time FE · Precincts FE	no	no	yes

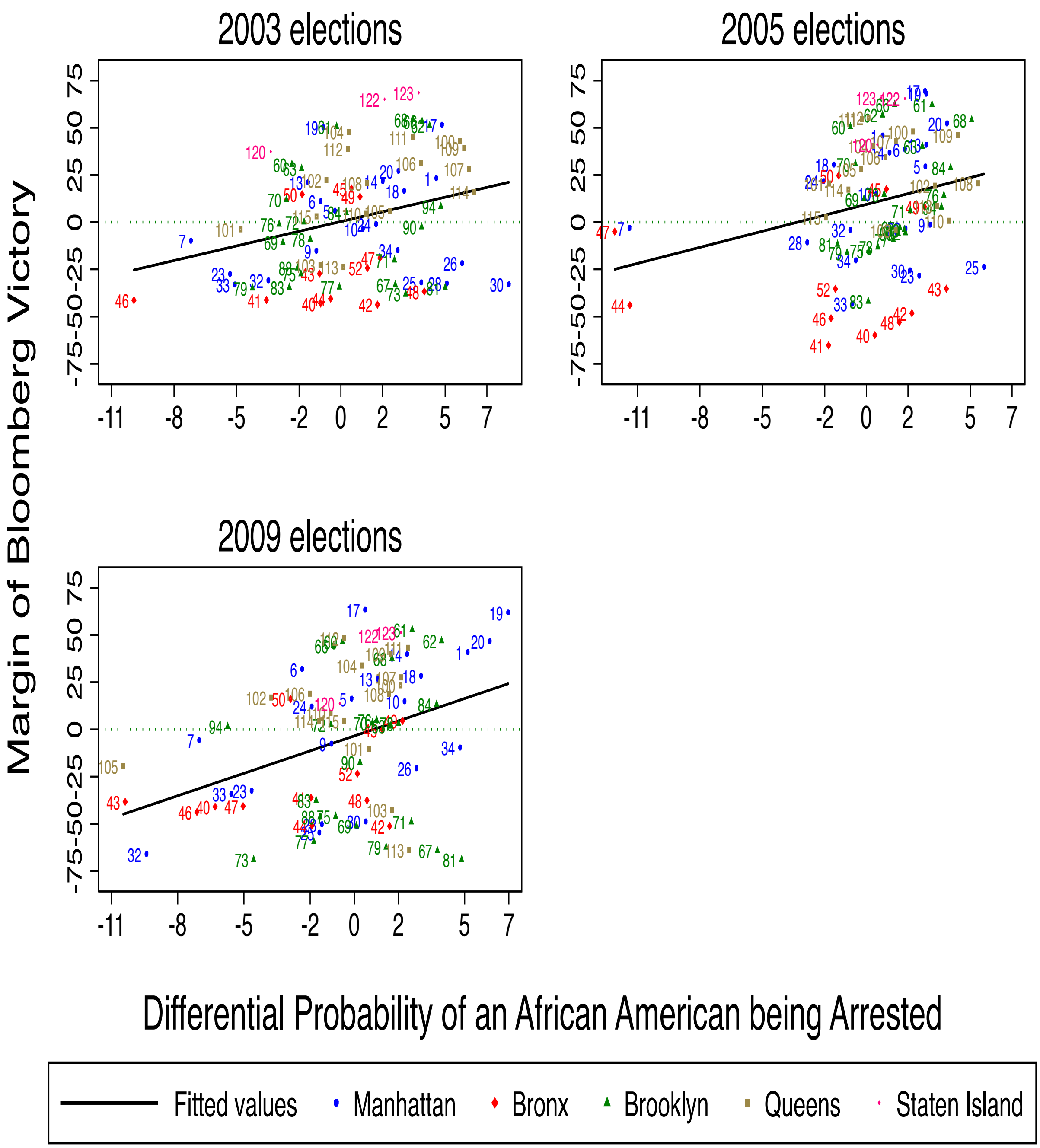
Notes. Estimates are on 76 precincts. The dependent variable is ($\pi(\alpha) = \alpha \cdot \mathrm{I}_{\text{arrest}} + (1-\alpha) \cdot \mathrm{I}_{\text{summons}}$) the weighted sum of the probability of being arrested and the probability of a summons being issued conditional on being stopped in New York City (in %). The weights $(\alpha, 1-\alpha)$ are .8 and .2. *African American* is an indicator variable coding the pedestrian's race. All the regressions include 13 indicators of the type of crime representing 95% of the crimes (as recorded by the officer on Form UF-250). To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Standard errors are clustered at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***).
Source. Statistics for the City of New York, Years 2003-2012.

Table B.12: Suspect Frisked

Model	OLS	OLS	OLS	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
African American	14.433***	14.176***	14.176***	10.761***	10.499***	10.499***	10.236***
	(0.079)	(0.079)	(1.595)	(0.096)	(0.095)	(1.166)	(1.141)
Constant	39.258***						
	(0.072)						
Mean outcome				51.38%			
Fraction of African American				84%			
P-value of $H_0 : u_i = 0$				0.001	0.001	0.001	0.001
Number of precincts				76	76	76	76
Observations	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865	2,947,865
Cluster SE	no	no	yes	no	no	yes	yes
Time FE	no	yes	yes	no	yes	yes	yes
Precincts FE	no	no	no	yes	yes	yes	yes
Time FE · Precincts FE	no	no	no	no	no	no	yes

Notes. Estimates are on 76 precincts. The dependent variable is the probability of being frisked conditional on being stopped in New York City (in %). *African American* is an indicator variable coding the pedestrian's race. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (9 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (9 dummies) and precincts fixed effects (75 dummies). Columns 3, 5-7, shows show clustered standard errors at the precinct level. *P-value of* $H_0 : u_i = 0$ is the p-value for the joint test of all the precincts fixed effects equal to zero. Significance at the 10% (*), at the 5% (**), and at the 1% (***). Source. Statistics for the City of New York, Years 2003-2012.

Figure B.1: **Mayor Bloomberg margin of victory and Differential Probability of an African American being Arrested**



Notes. The figure plots the Mayor Bloomberg margin of victory against the differential probability of an African American being arrested in the year of the elections for each of the 75 precincts. The latter is the estimated coefficient in the univariate regression of arrests on an indicator variable for african-american pedestrians. The line represents OLS fitted values; circles (blue) Manhattan, diamonds (red) Bronx; triangles (green) Brooklyn; squares (brown) Queens and dots (red) Staten Island.
Source. Statistics for the City of New York, Years 2003-2012.

C Variables, Descriptions, and Sources

Variable	Description	Source
African American	Indicator variable coding whether the pedestrian is African American	NYPD, Stop-and-Frisk Database
Hispanic	Indicator variable coding whether the pedestrian is Hispanic	NYPD, Stop-and-Frisk Database
Relative police pressure	$\frac{\frac{Arrests of African Americans}{African American population}}{\frac{Arrests of Whites}{White population}}$ in each of the New York City police precinct	NYPD Stop-and-Frisk Database
Margin of Bloomberg victory	Difference between M. Bloomberg and the first running opponent M. Green, or F. Ferrer, or B. Thompson vote share in the 2001, 2005, 2009 elections, respectively. Missing years are computed using moving averages. Original data: Electoral district data (2001, 2005, 2005 districts) matched with police precincts using the Stata routine gpsmap.ado	NYC Board of Elections
Fraction of African American	is the percentage of the population that is African American in the precinct in 2010. Original data: ZIP Code level data matched with police precincts using the Stata routine gpsmap.ado	U.S. Census Bureau, 2011
Income	Inflation adjusted median income, 2010 precinct average. Original data: ZIP Code level data matched with police precincts using the Stata routine gpsmap.ado	American Community Survey (5-year), U.S. Census Bureau
Age	Median age, 2010 precinct average. Original data: ZIP Code level data matched with police precincts using the Stata routine gpsmap.ado	American Community Survey (5-year), U.S. Census Bureau
Fraction of female	is the precinct average percentage of the population that is female in 2010. Original data: ZIP Code level data matched with police precincts using the Stata routine gpsmap.ado	American Community Survey (5-year), U.S. Census Bureau
Fraction of college degree	is the precinct average percentage of the population aged 15-24 with a college degree in 2010. Original data: ZIP Code level data matched with police precincts using the Stata routine gpsmap.ado	American Community Survey (5-year), U.S. Census Bureau
Serious crime	Number of annual crimes (murders, rapes, robberies, fel. assaults, burglaries, grand larcenies, grand larceny autos) in each precinct divided by the precinct population in 2010 (in 1,000 habitants) for the years 1998, 2001, 2012. Missing years are computed using moving averages. Original data: Data matched with police precincts by the NYPD	NYPD, crime statistics
Graffiti	Number of annual graffiti in each precinct in 2011 divided by the precinct population in 2010 (in 1,000 habitants). Original data: Data matched with police precincts by the NY Police Department	Graffiti Locations, NYC Open Data
Social capital	Number of annual civic initiatives (education, emergency preparedness, environment, helping neighbors in need, strengthening communities) in each precinct, in 2011 divided by the precinct population in 2010 (in 1,000 habitants). Original data: Latitude-Longitude geo-coded data matched with police precincts using the Stata routine gpsmap.ado	NYC Service Volunteer Opportunities, NYC Open Data
African American commander	Indicator for African American commanding officers. Original data: NY police borough matched with police precincts	NYC Journal articles