NBER WORKING PAPER SERIES

AN ECONOMIC ANALYSIS OF BLACK-WHITE DISPARITIES IN NYPD'S STOP AND FRISK PROGRAM

Decio Coviello
Nicola Persico

Working Paper 18803
http://www.nber.org/papers/w18803

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2013

An Economic Analysis of Black-White Disparities in NYPD's Stop and Frisk Program
Decio Coviello and Nicola Persico
NBER Working Paper No. 18803
February 2013
JEL No. K42

ABSTRACT

We analyze data on NYPD's "stop and frisk program" in an effort to identify racial bias on the part of the police officers making the stops. We find that the officers are not biased against African Americans relative to whites, because the latter are being stopped despite being a "less productive stop" for a police officer.

Decio Coviello
Institute of Applied Economics
HEC Montreal
room 4.152 (4th floor),
3000, Cote Ste-Catherine,
H3T 2A7, Montreal, QC
decio.coviello@hec.ca

Nicola Persico
Kellogg School of Business
room 542 Jacobs Center
2001 Sheridan Road
Evanston, IL 60208
and NBER
nicola@nicolapersico.com

1 Introduction

New York City's "stop and frisk program" is a police strategy whereby pedestrian are briefly stopped by police officers, and potentially searched. The stop and frisk program has given rise to widespread allegations of racial profiling.[1] The racial impact of the program has given rise to public protests[2] and to much social activism, including even a dedicated activist website.[3] The media have reported widely on this issue. Figure 1 reports the monthly number of New York Times articles which dealt with stop and frisk.

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 blacks and Latinos, as compared to whites, who become entangled in the criminal justice system."[4] The case, *Floyd et al.* v. *City of New York*, was decided on January 2013 in favor of the plaintiff, that is, against the police.

In this paper we analyze the possible racial bias of the police involved in the stop and frisk program.

[1]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 black 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:

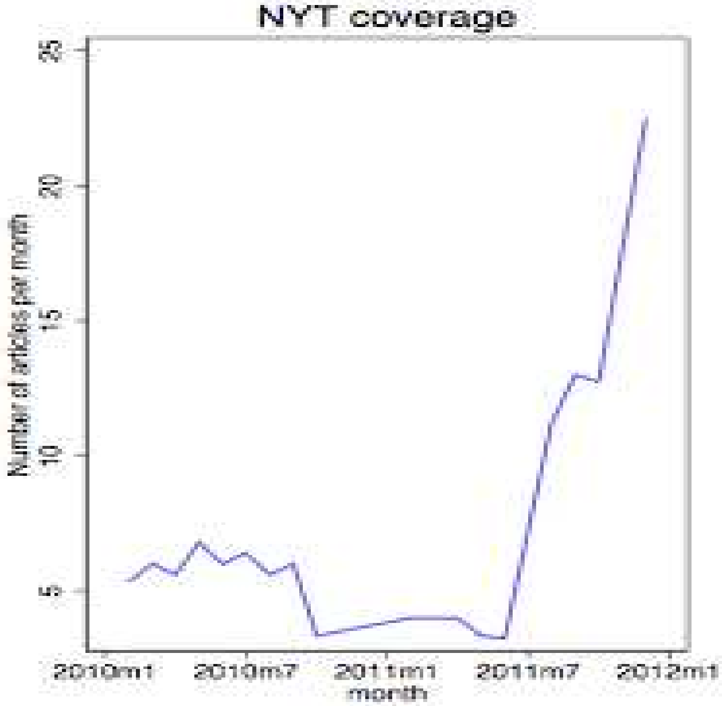
> When a majority of those targeted by police are young men of color and when the bulk of them are innocent, what else are we to conclude other than the fact that the NYPD has been implementing a policy of racial profiling and discrimination?

[2]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 black and Latino residents" (see "Thousands March Silently to Protest Stop-and-Frisk Policies," *New York Times*, June 17, 2012).

[3]http://stopandfrisk.org/

[4]Cited from "Court Strikes Challenge to Stop-and-Frisk Trial," Courthouse News Service, November 14, 2008. The case is Floyd v. City of New York.

Figure 1: **Stop and Frisk Media Coverage**



Notes. The figure reports the number of articles talking about the New York City's "stop and frisk program" on the New York Times searched using the NYT search engine.
Source. Articles appeared between 2010-2012.

2 The Data

We use data collected by the NYPD on individual stops, questionings and frisks in the City of New York between 2003-2011.[5] The database contains information on whether the person was frisked, issued a summons or arrested, the type of crime which is ascertained, the race of the pedestrian, the timing and location of the stop. We restrict the sample to black and white pedestrians, setting Hispanics aside, because the charge of racial bias seems to have special force with reference to the African American population. In this restricted sample

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

of 2,600,929 stops, approximately 6 percent of the stopped pedestrinas were arrested and 84 percent of the stops are of black pedestrians, the rest of whites. Most of of the crimes fall into one of these categories: Possession of a Weapon (27%); Robbery (16%); Rent Gouging (12%); Grand Larceny Auto (9.3%); Burglary (8.7%).[6] 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 (UF-250 form) 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).[7] Stops which occurred but did not give rise to one of these outcomes need not be recorded. Therefore, recorded stops (the ones in our database) may be a selected sample of all stops.

Nevertheless, we think there is reason to trust the sample somewhat as being representative of all stops. This is because 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.[8] This suggest that officers may have an incentive to record stops, perhaps as a way of demonstrating productivity 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.

It is tempting to attempt to address the selective recording concern by restricting the sample

[6] Other crimes are: Grand Larceny (4.1%); Illegal Possession of Substances (3.8%); Marihuana (3.8%); Assault (3.2%); Illegal Sales of Substances (3.1%); Petit Larceny (2.5%); Mischief (1.2%); Graffiti (1.1%).

[7] See Chapter 5 of the US Commission for Civili Rights Report (USCCR 2010).

[8] The outcome "refused to identify" is not recorded in the data. We proxy for it using the field "evasive response to questioning."

Table 1: Descriptive Statistics

	Mean	sd	n
Outcomes			
Arrest made	5.8	23	2,600,927
Race of the pedestrian			
Black	84	37	2,600,927
Crime details			
Possession of a Weapon	27	44	2,149,330
Robbery	16	36	2,149,330
Rent Gouging	12	33	2,149,330
Grand Larceny Auto	9.3	29	2,149,330
Burglary	8.7	28	2,149,330
Grand Larceny	4.1	2	2,149,330
Illegal Possession of Substances	3.8	19	2,149,330
Marihuana	3.8	19	2,149,330
Assault	3.2	18	2,149,330
Illegal Sales of Substances	3.1	1.7	2,149,330
Petit Larceny	2.5	1.6	2,149,330
Mischief	1.2	11	2,149,330
Graffiti	1.1	11	2,149,330
Other Crimes	4.4	21	2,149,330

Notes. Variables expressed in percent. *Black* 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*.

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. Therefore, as we will discuss below, our analysis cannot meaningfully be applied to this subsample. For this reason we will utilize the full sample of all stops in the database.

3 Disparate Impact vs. Intent to Discriminate

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 blacks, 33.7 percent were of Latinos, while whites accounted for only 9.3 percent of the stops.[9] This disparate impact is unfortunate, but should not be surprising if we believe that crime and therefore policing are disproportionally concentrated in minority-rich neighbourhoods.

However, mere disparate impact is not the same as impermissible behavior. Discrimination law in the United States generally does not prohibit disparate impact, as long as it does not reflect an intent to discriminate.[10] Therefore, if one is interested in impermissible behavior, it is helpful to have an empirical strategy which goes beyond merely documenting disparate impact, and can detect racial animus on the part of the police.

4 Detecting Intentional Discrimination: Hit Rates Analysis

Gelman et al. (2007) have analyzed similar data from the 1988-89 years. Most of their analysis focuses on documenting disparities in impact, but they address racial animus in

[9]See NYCLU (2011), pg.5.

[10]The expression "intentional discrimination" has a specific legal meaning: it is taken to mean that the treator engaged in disparate treatment "because of," not merely "in spite of," its adverse effects upon an identifiable group. A mere awareness of the consequences of an otherwise neutral policy does not suffice. See Persico (2009).

their Section 5.3. They tentatively conclude that police were indeed racially biased against blacks. This tentative conclusion is based on the statistical fact that blacks were less likely than whites to be arrested conditional on being stopped. This fact is informative about the bias of the police officer choosing whom to stop. Indeed, an officer who was not biased against blacks, and was motivated by the prospect of making an arrest, should cut down on less-productive stops of blacks and increase the more-productive stops of whites. A perfectly unbiased police force would generate equal arrest rates ("hit rates") between the stops of white and black pedestrians. Conversely, a disparity in hit rates across races is indicative of bias against the race that is least likely to produce a "hit" for the officer.

This argument based on the productivity of stops, and the associated statistical test, are referred to as "hit rates analysis" and were introduced in the policing context by Knowles et al. (2001).[11,12]

A previously unexplored issue in this literature is whether the hit rates analysis can be carried out on a subsample of all stops, where the subsample is selected based on the information acquired *after the pedestrian has been stopped*. We argue that it cannot. This is because the information that is used to exclude some of the stops is not available to the officer at the time of the stop—the officer has no way to distinguish in- and out of- sample stops. Conditioning our analysis on such ex-post information would mean conditioning on information

[11]See Ayres (2002), Persico and Castleman (2005), Todd (2006), Whitney (2008), and Persico (2009) for reviews of this strategy. Ayres and Waldfogel (1994) earlier applied 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 Hernandez-Murillo and Knowles (2004), Persico and Todd (2006, 2008), Sanga (2009), Childers (2012). Also, Anwar and Fang (2006) importantly extend the hit rate analysis.

[12]The hit rates analysis is also robust to the presence of information which is used by the police officer to identify whom to stop, but which is unobserved by the econometrician. See Persico (2009) for a discussion of this point.

not possessed by the officer at the time of the stop. Put differently, the outcomes contained this restricted data set would not correspond to the outcomes generated by an officer's decision problem at the time of the stop. Thus the hit rate analysis cannot properly be applied to such a subsample. Consequently the subsample of stops that are required by law to be recorded, which was mentioned in Section 2, cannot be the object of hit rate analysis.

5 Importance of Conditioning on Precincts in Hit Rates Analysis

We begin by replicating Gelman *et al*'s (2007) "hit rates" test on the more recent data made available by the NYPD. We set Hispanics aside and focus exclusively on the relative treatment of black and white pedestrians, because the charge of racial bias seems to have special force with reference to the African American population. As in Gelman et al. (2007), we focus on arrests as the measure of the officer's productivity.

Our analysis confirms two patterns which were also found by Gelman et al. (2007) in their sample.[13] First, few (only 6%) of the persons stopped are actually arrested. Second, there is a racial disparity in arrest rates: 1 in 15 whites stopped were arrested, compared with approximately 1 in 17 blacks. To gauge the statistical significance of this disparity, in Table 2 we regress the percent probability of being arrested on an indicator variable coding the pedestrian's race. Depending on the specification, black pedestrians who are stopped are between 0.338% and 0.356% less likely to be arrested compared to whites. (Whites have

[13] Details in Section 2

about a 6% probability of being arrested). 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. Thus like in Gelman et al. (2007), in our sample, too, black pedestrians can be significantly (in a statistical sense) less likely to be arrested then whites conditional on being stopped.

Table 2: Arrest Made

Model	OLS (1)	OLS (2)	OLS (3)	FE (4)	FE (5)	FE (6)	FE (7)
Black	-0.338***	-0.356***	-0.356	0.414***	0.388***	0.388*	0.388*
	(0.039)	(0.039)	(0.466)	(0.048)	(0.048)	(0.200)	(0.199)
Constant	6.069***						
	(0.036)						
Mean outcome				5.79%			
Fraction of black				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	2,600,929	2,600,929	2,600,927	2,600,927	2,600,927	2,600,927	2,600,927
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

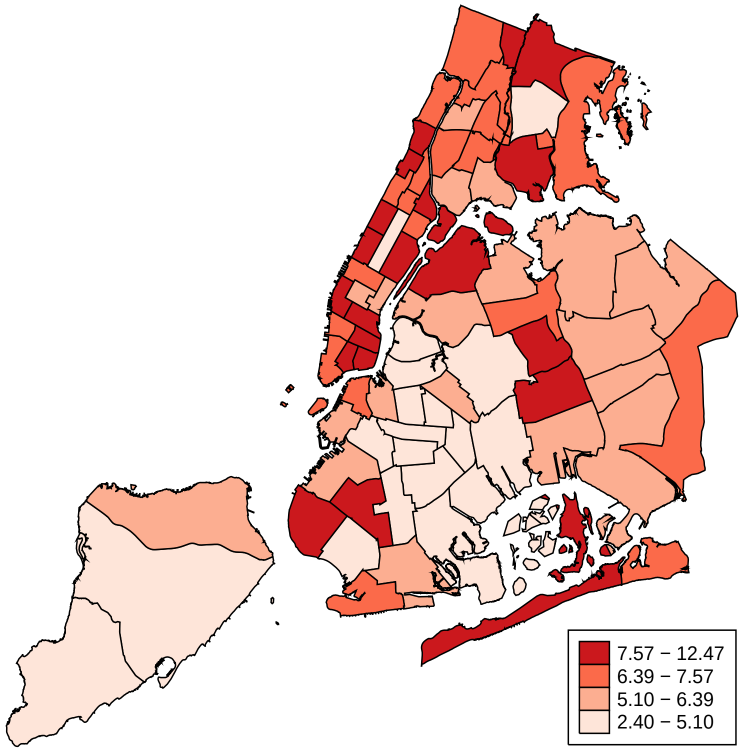
Notes. Estimates are on 76 precincts. The dependent variable is the probability of being arrested conditional on being stopped in New York City (in %). *Black* is an indicator variable coding the pedestrian's race. To control for possible time trend in the dependent variable and precincts specific characteristics, when denoted with "yes", regressions additionally include year fixed effects (8 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (8 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% (***).

However, we need to be careful in interpreting the outcome of this test as evidence of police bias against blacks. This test masks heterogeneity in arrest rates across precincts. Figure 2 shows that precincts vary considerably in the likelihood that a stop translates into an arrest. This heterogeneity might lead to a fallacy in aggregation.

To see this fallacy, consider precincts as separate jurisdictions, so that an officer from one precinct cannot stop pedestrians in another precinct.[14] If the police officers in each precinct

[14] 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". The NYPD is organized in 76 precincts, each of which is responsible for a specific geographic area.

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



Notes. The figure reports Probability of Being Arrested Conditional on Being Stopped in New York City (in %).
Source. Statistics for the City of New York, Years 2003-2011.

were unbiased, then *within each precinct* the arrest rates of black and white pedestrians stopped should be the same. However, the levels of these arrest rates need not all be the same *across precincts*.[15] For example, suppose hypothetically that of all the blacks and whites stopped in the Bronx 3% were arrested, and 6% of the blacks 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 blacks, because in the aggregate sample most blacks are searched in the Bronx and have a 3% arrest

[15]Figure 2 which reports the arrest rates by precinct, shows that indeed precincts have very different "baseline" arrest rates. This suggests that the "separate jurisdictions" story is valid and that it is necessary to control for precinct fixed effects.

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 "black " 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 black searchees in many or all precincts, and this black-white difference in arrest rates would be picked up by the coefficient on "black,"after controlling for precinct-level fixed effects. Therefore, controlling for precinct fixed effects is necessary for the hit rates test to function properly.

In columns 4-7 of Table 2 we run the same OLS regression, this time introducing 76 precincts-level fixed effects. Notably, the coefficient on "black" changes sign. Now, stopping a black pedestrian results in a probability of arrest which is larger by 0.388% to 0.414% compared to a white pedestrian. Our estimates also suggest that the precincts fixed effects are jointly explaining, in statistical terms, the arrest rate (i.e., the p-value for the joint test of all the precincts fixed effects equal to zero reported in Table 2 less than 5%).

That is, after accounting for the fact that different precincts have different "baseline" rates of arrest conditional on search, blacks are no longer less likely to be arrested conditional on being searched. Therefore, introducing precinct-fixed effects overturns the result: the hit rates analysis provides no evidence that the individual police officers who make the decision to stop this or that pedestrian, are biased against blacks. This is the main message of this analysis.[16]

6 Discussion Concerning our Choice of Outcome

6.1 Validity of Arrest as an Outcome

Gelman et al. (2007) raise the concern that arrests might not be a suitable outcome for hit rates analysis. The ideal outcome is a measure of productivity which the officer legitimately maximizes, and which is itself "objective," that is, cannot be affected 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 the police may be more likely, all else equal, to arrest a black than a white pedestrian after having stopped either. This is a valid concern which might invalidate the hit rates test.

We address this concern by looking at the officer's behavior *after the stop* has been made. We check whether, given the type of crime ascertained by the police officer after the stop, the officer is more likely to arrest a black 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

[16] For completeness, we ran the various specifications in Table 2 on the subsample of stops that are required by law to be reported. In this subsample the coefficient on "black" 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.

to see blacks being arrested more often than whites for the same type of crime.[17]

To implement this test, we check whether the race of the person stopped predicts the probability of arrest, after conditioning on the type of crime (as recorded by the officer on Form UF-250). Table 3 presents the results.

Table 3: Arrest Made, Controlling for the Type of Crime

Model	OLS (1)	FE (2)	FE (3)
Black	-0.495	0.228	0.234
	(0.412)	(0.213)	(0.213)
Mean outcome		5.75%	
Fraction of black pedestrians		84%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	2,149,330	2,149,330	2,149,330
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 %). *Black* 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 (8 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (8 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% (***).

After controlling for the type of crime committed, the estimates suggests that there is no evidence of a significant race effect on arrest. We interpret this result as evidence that, given a certain crime committed by 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.

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

Therefore, we conclude the outcome "arrest" is an outcome that can properly be used to carry out the hit rates test.[18]

6.2 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, Table A.2 we perform the hit rate test on the outcome "summons issued." The results are the opposite of Table 2: after controlling for precincts, the sign on "black" switches and becomes *negative*.[19] The interpretation, according to the hit rate analysis, would be that officers are biased against blacks in their decision to stop *if officers only cared about issuing a summons.* But, clearly officers care about both issuing a summons and making an arrest.[20] 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. We do not know the actual value of α. If α is close to zero then

[18] 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 and is available upon request.

[19] Table A.3 in the appendix corresponds to Table 3 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, blacks appear to be issues summons less often than whites. Therefore, we see no evidence that any police discreation which might affect the outcome "summons" is biased against blacks.

[20]

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 the variable "arrests."

What happens if we run the hit rate test on the outcome variable $\pi(\alpha)$? Will the police look biased against blacks or against whites? This depends on the value of α. If α is close to zero then we know from Table A.2 in the appendix that the police will look biased against blacks. If α is close to one then Table 2 says that the police are biased against whites. We have performed a search for the threshold value α such that the police looks unbiased. This threshold value is $\alpha = 0.82$. At this value, the police values each arrest equal to about four summons. For any value of α larger than 0.82, that is, if the police officers value arrest more than four summons, then the hit rate test would conclude that the police is not biased against blacks. For any value smaller than 0.82, then the hit rate test would conclude that the police is not biased against whites. We regard a "value ratio" of 4 or more summons per arrest as a plausible one.

7 Bias in Manpower Allocation Across Precincts?

Our tests in Section 5 (columns 4-7 in Table 2) have shown that individual officers are not biased against blacks in selecting which pedestrians to stop. This suggests that no bias is present *at the within-precinct level.* Is there any impermissible behavior going on at a different level, i.e., *in the cross-precinct* resource allocation? In this section we examine this question.

Let's start with disparate impact. In our data we find evidence of disparate impact in the cross-district resource allocation. Precincts with a higher fraction of black residents experience a higher police pressure per capita (correlation coefficient 0.39).[21] Is this disparity evidence of bias, or impermissible behavior?

Any discussion of bias has to start with identifying a legally "permissible" objective or motive for the police, which defines what is unbiased behavior. Deviations from this behavior can be classified as biased. 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.[22] If we make this assumption, then we might conceptualize the "legally permissible version" of the police chief's problem as follows:

$$\max_{p_i} \sum_i C_i\left(p_i\right) \quad \text{s.t.} \quad \sum_i p_i \leq P,$$

where p_i represents the number of police officers assigned to precinct i, P represents the total amount of police officers available to the police chief, and the function $C_i\left(\cdot\right)$ represents the crime rate (per capita) in precinct i. According to this formulation, the legally permissible objective of the police chief is to minimize the sum of (per capita) crime rates across all precincts.

If this is a legally permissible objective, what would be the corresponding impermissible, or biased, version? Perhaps one in which we allow the police chief to "prioritize" different precincts. This can be conceptualized by assigning different weights to the crime rates of

[21] In Appendix, Table A.1 we report the correlation matrix.

[22] Of note, this objective is 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.

different precincts. Then, a potentially biased police chief would solve the followin problem:

$$\max_{p_i} \sum_i \beta_i C_i\left(p_i\right) \quad \text{s.t.} \quad \sum_i p_i \leq P, \tag{1}$$

where β_i is the weight assigned to precinct i.

The difficulty with this conceptualization is that one might be ambivalent about whether the configuration $\beta_i > \beta_j$ represents bias in favor or against precinct i. On the one hand, $\beta_i > \beta_j$ 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, $\beta_i > \beta_j$ means that precinct i will be assigned more police officers, so more stops and more frisks, which some civil liberty advocates object to. So, it is conceptually/normatively ambiguous whether $\beta_i > \beta_j$ 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:[23]

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

where p_i^* and p_j^* represent the optimal allocation. This equation shows that in order to

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

recover the ratios β_i/β_j we need to observe the ratio of elasticities of crime to policing, $C'_j\left(p^*_j\right)/C'_i\left(p^*_i\right)$. The difficulty is that whereas it might be possible to observe the *equilibrium levels* of crime rates $C_i\left(p^*_i\right)$, it is much harder to get persuasive estimates of their *elasticities* $C'_i\left(p^*_i\right)$. An elasticity 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.* In other words, estimating elasticities requires observing more than simply equilibrium levels of crime. This is an empirical challenge.[24]

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 β_i in problem (1).

8 Conclusions

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.[25]" 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. We have analyzed data on NYPD's "stop and frisk program" in an effort to identify racial bias on the part of the police officers making the stops. Once we control for precincts, white pedestrians are slightly less likely than african-american pedestrians to be arrested conditional on being stopped. We interpret this fact as evidence

[24]See Persico (2009) for a discussion of the kind of statistical variation that would permit the identification of the weights β_i.

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

that the officers are not biased against African Americans relative to whites, because the latter are being stopped despite being a "less productive stop" for a police officer. According to this interpretation, the small differential in probability of arrest represents a slight bias against white pedestrians.

In our view, the observed differential across races in the productivity (probability of arrest) of a stop should not be overinterpreted. Indeed its size is small. According to our estimate, after controlling for precinct, stopping a white motorist is only marginally less likely to result in arrest (4/10 of a percentage point). Such a small difference might be difficult for a police officer to detect based on his own personal experience. If that is the case, and police officers cannot distinguish the average productivities of black v. white stops, then it is proper to interpret the evidence as consistent with the officers being unbiased. Moreover, if we take into account that issuing a summons is also a productive outcome, albeit a less productive one than an arrest, then it is even easier to rationalize police behavior as unbiased.

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. Our interpretation of the results is simply that, whether or not the officers behavior violates the law, the behavior does not display a racial bias, conscious or not.

A different question is whether bias may be found in the allocation of officers across precincts. In Section 7 we have suggested a framework for thinking through what bias might mean in the context of such an allocation problem. Our view is that defining such bias presents conceptual difficulties; moreover, identifying it empirically presents additional difficulties.

References

[1] 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.

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

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

[4] 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.

[5] 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.

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

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

[8] 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.

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

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

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

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

[13] 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.

[14] 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.

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

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

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

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

[19] 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

Appendix A

Table A.1: Cross-correlation table

Variables	Arrest made over stops	Black over population	Stops over population
Arrest made over stops	1.00		
Black over population	-0.29	1.00	
	(0.01)		
Stops over population	-0.26	0.39	1.00
	(0.03)	(0.00)	

Notes. Correlations (and *p-values* for statistical significance in parenthesis) for 75 precincts in 2011. The precinct of central park is excluded since there is no population living in this precinct. *Arrest made over stops* is the number of arrest made relative to the total stops in the precinct. *Black over population* is the number of black residents relative to the total resident population. *Stops over population* is the number of stop-and-frisk relative to the resident population.
Source.

Table A.2: Summons Issued

Model	OLS	OLS	OLS	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Black	0.118***	0.133***	0.133	-1.804***	-1.791***	-1.791***	-1.758***
	(0.041)	(0.041)	(0.380)	(0.051)	(0.051)	(0.317)	(0.297)
Constant	6.253***						
	(0.038)						
Mean outcome				6.35%			
Fraction of black				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	2,600,927	2,600,927	2,600,927	2,600,927	2,600,927	2,600,927	2,600,927
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 %). *Black* 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 (8 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (8 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% (***).

Table A.3: Summons Issued, Controlling for the Type of Crime

Model	OLS	FE	FE
	(1)	(2)	(3)
Black	-0.254	-1.533***	-1.537***
	(0.338)	(0.321)	(0.302)
Mean outcome		6.33%	
Fraction of black pedestrians		84%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	2,149,330	2,149,330	2,149,330
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 %). *Black* 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 (8 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (8 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 A.4: Arrest Made and Summons Issued (Weighted)

Model	OLS	OLS	OLS	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Black	-0.256***	-0.268***	-0.268	0.015	-0.004	-0.004	0.002
	(0.033)	(0.033)	(0.379)	(0.041)	(0.041)	(0.172)	(0.167)
Constant	6.102***						
	(0.030)						
Mean outcome				5.89%			
Fraction of black				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	2,600,929	2,600,929	2,600,927	2,600,927	2,600,927	2,600,927	2,600,927
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 .82 and .18. *Black* 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 (8 dummies) and precincts fixed effects (75 dummies). In Column 7, we include interactions between year fixed effects (8 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% (***).

Table A.5: Arrest Made and Summons Issued, Controlling for the Type of Crime (Weighted)

Model	OLS	FE	FE
	(1)	(2)	(3)
Black	-0.452	-0.089	-0.085
	(0.333)	(0.166)	(0.164)
Mean outcome		5.85%	
Fraction of black pedestrians		84%	
P-value of $H_0 : u_i = 0$		0.001	0.001
Number of precincts		76	76
Observations	2,149,330	2,149,330	2,149,330
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 ($\pi(\alpha) = \alpha \cdot \mathrm{I}_{\mathrm{arrest}} + (1-\alpha) \cdot \mathrm{I}_{\mathrm{summons}}$) 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 .82 and .18. *Black* 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 (8 dummies) and precincts fixed effects (75 dummies). In Column 3, we include interactions between year fixed effects (8 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% (***).