Race or Place?
The Persistence of Race Effects in Police Behavior following Traffic Stops

Frank R. Baumgartner
frankb@unc.edu

Leah Christiani
christiani@unc.edu

Derek A. Epp
depp@austin.utexas.edu

Kevin Roach
kevroach@live.unc.edu

Kelsey Shoub
shoub@unc.edu

Abstract

Evidence that minority drivers are searched at higher rates than whites is widespread, but there are enduring concerns that simple measures of racial difference are being confounded by various non-racial factors that may influence an officer's decision to search a motorist. We provide evidence from over 40 million traffic stops in four states that such concerns are overstated. While race is not the only factor that determines if an officer will conduct a search, racial disparities in traffic stop outcomes are large and widespread, and they cannot be explained away by contextual factors such as the reason why the car was pulled over, the neighborhood that the stop took place, the time of day, or the age of the vehicle, or the race of the officer. Black and Hispanic drivers are consistently subjected to harsher police behaviors than whites.

May 20, 2018
University of North Carolina at Chapel Hill, Department of Political Science
http://www.unc.edu/~fbaum/traffic.htm

The United States is currently going through a period of renewed attention to questions of racial justice. The concern that police officers direct undue scrutiny to blacks and other minority groups has prompted state and local governments and media organizations to seek out and engage with data on police-citizen encounters, so that the extent of any disparities can be fully documented. However, how such disparities should be measured and what information is needed to do so is debated. These debates rest on the answer to two questions: how pervasive or localized is disparate treatment within a single agency; and what causes disparate treatment? To the first question, disparate treatment may be pervasive throughout an entire department, regardless of where the stop took place or which police officer stopped the individual. Alternatively, disparate treatment may occur in only certain areas of city or only among a small number of "bad apple" police officers. The first step to understanding racially disparate outcomes in policing is to identify whether racially disparate outcomes are widespread or confined to specific contextual circumstances.

Once we understand the scope of the disparities, the second question must be tackled: what causes these disparate outcomes? The answer may be that people of different races and ethnicities experience different treatment above and beyond the circumstances of the stop. No matter the location of the stop, the individual officer performing the stop, the reason for the stop, or any other explanatory variable—racially biased outcomes may persist. Alternatively, any differences in outcomes may be entirely explained by the circumstances and reason behind a stop, and those factors may vary systematically by racial groups. Once these contextual factors are accounted for, the explanatory power of race may disappear, though it is important to note that those contextual factors themselves may be inextricably linked to race. If minorities drive older cars, or live in certain neighborhoods where policing is more intense, perhaps these factors generate the disparate outcomes. Controlling, then, for non-racial factors may either explain away the apparent racial difference, or explain the systemic mechanisms that cause it.

In order to evaluate these questions, we have compiled publicly available records on all traffic stops in recent years in Connecticut, Illinois, Maryland, and North Carolina, for a total of over 40 million records. With such a large database, we can first assess the degree of racial disparity in outcomes and, because each state also collects numerous contextual variables about the circumstances of the stop, we can evaluate whether these apparent differences across race can be explained by non-racial factors, such as time of day, why the car was pulled over, or the neighborhood where the stop occurred.

We find that this comparison of the rates at which minority and white drivers are searched, which is calculated using only the raw number of stops and searches, is an extremely robust indicator of racial difference. While contextual factors do have some explanatory power, the role of race in the outcomes of traffic stops constantly remains a significant explanatory variable. Racial disparities, in which black and Hispanic drivers are searched at higher rates than white drivers, persist due to race itself. Further, where the simple ratio of the minority search rate to the white search rate is high, so also is the odds ratio for race in the most complete logistic regression we can run, given the data collected in each state. Thus, we document widespread, consistent, and alarmingly high racial disparities in traffic stop outcomes, and we show that these cannot be “explained away” by non-racial factors.

To deepen our analysis after our multi-state analysis, we focus on one city, Charlotte NC, which collects the most comprehensive contextual information about traffic stops that we have found. Through its open data portal, the Charlotte Police Department makes available information about every stop, including the demographic characteristics of the officer making the stop and the neighborhood in which the stop occurred. We show that even controlling for division within the city and officer characteristics, racial disparities persist and a comparison of the search rate ratios and odds ratios is still robust. We arrive at this conclusion by conducting the analysis in two ways: a series of regressions including division and officer controls and difference of proportions test on data that is exactly matched on all possible variables except for driver race and whether a search occurred.

Our methodological approach is straightforward, even if our empirical scope is large. The substantive implications of our findings are clear: racial disparities in traffic stop outcomes are accurately assessed using simple methods available to any police agency collecting even the most basic demographic information, and do not require sophisticated statistical models. This means that important conversations about race and policing can be had in a great number of jurisdictions even when more comprehensive information about the contextual factors surrounding traffic stops is unavailable. Instead, the simple search rate ratio is able to accurately reflect the extent to which racial disparities exist in policing in that particular jurisdiction—allowing for broader and more widespread conversations about racial disparities in policing. Moreover, our results demonstrate that racial disparities are large and ubiquitous across the country, so the fact that they cannot be "explained away" by other contextual factors makes them all the more troubling.

Racial Disparities in Traffic Stops

Previous work has consistently demonstrated high levels of racial disparities in traffic stop outcomes (Baumgartner, Epp, et al. 2017; Tillyer and Engel 2013; Tillyer, Klahm, and Engel 2012; Petrocelli, Piquero, and Smith 2003; Tomaskovic-Devey, Mason, and Zingraff 2004). In many states, black and Hispanic drivers are substantially more likely to be searched or arrested following a traffic stop, and they are frequently pulled over at rates that far exceed their numbers in the population (Peffley and Hurwitz 2010; Burch 2013; Lerman and Weaver 2014; Epp et al. 2014; Moore 2015; Baumgartner, Christiani, et al. 2017; Pierson et al. 2017; Baumgartner, Epp, and Shoub 2018). Altogether, the evidence for profound and widespread racial differences is indisputable, but the cause of the disparities themselves is debated. Are these differences really attributable to race? Or can they be attributed to some contextual factor that is merely correlated with race, like neighborhood or stop purpose?

A number of factors might generate racial disparities in policing outcomes. Michael Tonry (1999) lays out four possible reasons: differential criminality on the part of members of different racial groups; officer-level differences in attention to drivers of different races (whether through "old fashioned" racism or implicit bias); systemic differences or implicit biases shared by the vast majority of officers (e.g., not just a few "bad apples"); and systemic or institutional practices that yield disparate outcomes (for example, focusing on "poverty crimes" such as expired tags, or encouraging different types of police practice in different neighborhoods, themselves distinguishable by race). These categories are not mutually exclusive, and there may be overlap and connections between the distinct categories. Similarly, Jack Glaser (2015) focuses on the inaccurate use of criminal stereotypes and demographic profiles and differential policing by place.

In this paper, our goal is not to make a definitive argument about the precise process by which racial disparities in policing are produced. Instead, we seek to test whether race remains a significant explanatory variable even after a multitude of contextual factors are taken into consideration. We expect that it will because while some of the aforementioned causal mechanisms may only be correlates of race, many of them are race-based phenomena. Especially important is research on implicit biases, which shows that such biases are common and that they work to the determinant of blacks and Hispanics. Implicit bias is different from explicit bias (old-fashion racism) because it operates on a subconscious level and may therefore affect the decision making of people who would not think of themselves as racist or outwardly appear to harbor any racial prejudices. These biases are thought to emerge on a cultural level through repeated exposure to movies, television, news coverage, and popular culture that present blacks and other minority racial groups in an unflattering or stereotypical light (Sagar and Schofield 1980; Correll et al. 2002; Eberhardt et al. 2004; Anderson 2010). Because the stimulus is inextricably tied to the American experience, implicit biases are widespread (Gillian and Iyengar 2000; Payne 2001; Plant and Peruche 2005), and there is no reason to think that police officers would be immune to them. In fact, studies have shown that in ambiguous experimental settings officers are more likely to shoot hypothetical black suspects than identically situated white suspects (Correll et al. 2002). Such biases are magnified when split-second decision making is called for, as it often is when police officers must decide whether or not to shoot someone (or stop a vehicle) (Payne 2006). Even commonplace interactions between officers and citizens can be affects by implicit bias. Voigt et al. (2017) find that officers use less respectful language to black community members than white ones during traffic stops. Of course, implicit biases are not the only potential causes we explore; institutional practices may be involved as well: targeting certain neighborhoods for more intensive patrol, for example, could be an important cause of racial disparity. Controlling for as many of these factors as possible, we want to know whether the race effect persists.

Hypothesis

Our argument is that race is one of many factors that police officers use when deciding whether to stop and search a motorist. Of course, this is a probabilistic argument. We do not anticipate that every officer makes a race-based determination before every search. Moreover, we lack the data to determine if race played a role in any one individual stop out of the millions that are recorded in the data. Instead, the idea is that race is a common enough factor among police officers that we should be able to observe large and statistically robust race-related effects on average. Further, we argue that the average police officer is more likely, perhaps only slightly, to see black or Hispanic drivers as suspicious, and therefore more likely to search drivers from these demographic groups.

This argument has a very strong theoretical basis. Taking a historic view, race has obviously been an important explanatory variable for a huge range of criminal justice and social outcomes. So, we might wonder when exactly race stopped being so important. That is, knowing what we do about the history of racial hierarchy in America, the prior probability we would assign to race mattering in any given criminal justice outcome would be high.

We also know that implicit biases are virtually ubiquitous and operate to the detriment of blacks and Hispanics, since they are more likely to be stereotyped as violent or involved in crime. These biases are more likely to influence decision making when individuals are asked to make snap judgments, which is often the case when an officer must quickly determine whether or not to search a driver after only a short conversation. Even a subtle bias, if it is widespread, would result in substantial race-based differences in the likelihood that black and Hispanic drivers are searched. Implicit bias might also operate at the agency or even state level if leaders of a police force determine that a predominately minority neighborhood deserves some special scrutiny. Such decisions may be justified on the basis of crime rates, with police chiefs allocating more officers to high crime areas. But if we take seriously the ubiquity of implicit biases, then that implies they will affect decision making at both the street and agency level.

Hypothesis: race will continue to have an explanatory role in the reason an officer decides to search a car, even after controlling for contextual factors

Our hypothesis that race matters has very simple observational implications. First, and most obviously, we expect to find racial differences in search rates and we expect these differences to be widespread across different states and police departments. Second, we expect such differences to be statistically robust so that controlling for potentially confounding factors does not eliminate them. We test these expectations and find strong support for both of them in the pages to follow.

Data and Methods

Many law enforcement agencies across the country make some basic traffic stop data available, but four states mandate the collection and public availability of detailed contextual information about each stop from every police agency, not only the highway patrol.[1] These states make available not only annual summaries, but the full micro-level databases, with a record for each individual traffic stop. The details collected with regard to the stop vary, but generally include the race and gender of the driver stopped and the outcome of the stop, including whether or not a search occurred. We focus on searches rather than other traffic stop outcomes (such as citations or arrests) because searches are directly tied to the concept of suspicion. Concerns around racial profiling are often that police officers stereotype blacks and Hispanics, and especially young black and Hispanic men, as criminals and that this leads them to search drivers matching that

[1] A number of other states require the collection of data about the racial breakdown of who is stopped and what happens to them afterwards, but many of these do not make the data publicly available. For a summary of this see Baumgartner, Christiani, et al. 2017.

demographic profile at unduly elevated rates as compared to white drivers (see Baumgartner, Epp, and Shoub 2018 for more information about the different outcomes, such as warnings, citations, and arrest).

We propose that if race, in and of itself, is an important consideration in who is searched, then a comparison of the percentage of those searched from different race and ethnic groups will yield very similar information as a regression controlling for other factors. Contextual factors should only marginally attenuate the explanatory power from the role of race. To test this, we rely on the micro-level data made available by four states: Connecticut, Illinois, Maryland, and North Carolina. Using this data, we first calculate a simple search rate ratio comparing the percentage of black or Hispanic drivers searched to the percentage of white drivers searched, for every agency included in the dataset, over 1,800 agencies. Then, using those same data, we fit logistic regressions for each agency, controlling for every contextual factor about the stop that the dataset allows. We then compare the search rate ratio to the logistic odds ratio produced from the regression. We then test how well the search rate ratios predict the logistic regression coefficients, by regressing the search rate ratios on the logistic odds ratios for each agency. Finally, we explore one city in greater detail, exploring neighborhood effects, officer demographics, and using an exact matching methodology to test the robustness of our findings.

Table 1 summarizes the data we use. As mentioned, our data comes from the individual-level traffic stop databases collected from Connecticut, Illinois, Maryland, and North Carolina. We aggregate this data by agency and by time period. A threshold is imposed upon the agency-time period windows in order to ensure that small jurisdictions with few stops and searches do not bias our results. The total number of searches and the search rate by race are reported.

Table 1. Traffic stops, searches, and search rates by state and race

State	Years	Agency-Periods	N (stops) Logistic Regressions	N (stops) Rate-Ratios	Searches	Search Rate Total	Search Rate White	Search Rate Black	Search Rate Hispanic
CT	2013-15	15	460,343	461,517	14,796	3.2%	2.0%	7.3%	5.9%
IL	2004-14	1,222	21,464,656	21,958,971	924,674	4.2%	2.5%	7.3%	9.2%
MD	2013-16	68	2,333,523	2,548,013	75,071	3.2%	2.7%	4.1%	3.8%
NC	2002-16	552	9,517,134	19,240,543	517,621	5.4%	3.9%	7.3%	6.6%
Total		1,857	33,775,647	44,209,044	1,532,162				

Note: N's are lower for the logistic regressions because of missing data on covariates. Numbers are slightly different for Hispanic-White comparisons: 1,838 agency-periods, 33.7 million stops for the logistic regressions, 44.6 million for the rate ratios, and 1.6 million searches.

As is clear, searches are relatively rare. On average, the total search rate in a state is less than five percent, though this varies by racial group. In order to produce robust estimates, because searches are rare, and because many police agencies are so small, we set a multi-condition threshold for agencies and eliminate those agencies that fall below the threshold. First, agencies are included only if they have a minimum of 10,000 traffic stops. Second, this must include at least 100 stops of white drivers and 100 stops of black drivers. (For our analyses of Hispanic-white disparities, the agency must make 100 stops of white and 100 stops of Hispanic drivers in a given time period.) Typically, these thresholds are met by using a single year of data. However, if they are not met, we add the following year of data for the same police agency, and we continue this process until we meet the threshold. For example, if a police agency made 11,000 total stops in 2009, including 4,000 stops of black drivers, and in 2010 made 12,000 total stops and 3,000 stops of blacks, both years of data would meet the threshold independently and be included in the analysis. (In Table 1, we would count 2009 and 2010 as two agency-period observations.) But if an agency made only 7,000 stops in 2009 and only 4,000 in 2010, then we would combine these two years of data to form one agency-period observation that meets the 10,000 stop threshold. Some agencies never meet these thresholds even after combining every year for which data is available and, as a result, are dropped from the analysis. The net result of these procedures is to drop very small police agencies from our analysis and to eliminate those with very low levels of racial diversity. These thresholds are not particularly restrictive, but given that our dependent variable is relatively rare, they are necessary to obtain accurate, robust estimates.[2]

Next, we calculate the search rate ratio. This ratio is simply the percentage of black drivers searched divided by the percentage of white drivers searched. For Hispanics, the statistic is the same with the black search rate replaced with the Hispanic search rate. The formula for this simple statistic is represented below.

$$SRR = \frac{Searches\ of\ Minority\ Drivers / Stops\ of\ Minority\ Drivers}{Searches\ of\ White\ Drivers / Stops\ of\ White\ Drivers}$$

Values below 1 indicate that white drivers are searched at higher rates than minority drivers, while values above 1 indicate that minority drivers are searched at higher rates. If the minority search rate is 6 percent and the white rate is 3 percent, the SRR is 2.0.

Of course, the simple SSR could potentially obscure non-racial reasons for any disparities found. Therefore, we also conduct the most conservative analysis we can, based on the data made available in each state and compare the results of these logistic regressions with the simple comparison of search rates. In each state, we estimate a logistic regression predicting whether a given traffic stop will lead to a search, controlling for other factors. We strive to estimate the most conservative model for each state and, since each state collects different contextual factors

[2] We eliminated approximately 10.43 percent of all traffic stops through these procedures, all from small jurisdictions.

about the traffic stop, we estimate a slightly different model in each state. The independent variables included in each regression are listed in Table 2.[3]

Table 2. Summary of variables available by state

Variable	CT	IL	MD	NC
Race	X	X	X	X
Gender	X	X	X	X
Driver Age	X	X	X	X
Stop Purpose	X	X	X	X
Hour of Day	X	X	X	X
Day of Week	X	X	X	X
Out of State	X	X	X	
High Disparity Officer	X		X	X
Vehicle Age		X		

Note: X indicates the variable was included. A blank indicates the variable was not available.

Variables for race, gender, stop purpose, hour of day, day of week, out-of-state plates, and high disparity officer are included as dichotomous variables (that is, each stop purpose, hour and day of the week is identified separately in the model). Driver and vehicle age are included as continuous variables. The independent variable for race is a categorical variable with values for white, black, Hispanic, and "other." White is the baseline racial category and as a result, the odds ratios produced are the odds of a search for minority drivers, as compared to white drivers.[4]

Three states make available a variable identifying (usually anonymously) the officer who made the traffic stop. We generate a "high disparity officer" variable coded as 1 if the officer has: a) at least 50 stops of white drivers; b) at least 50 stops of black (or Hispanic, depending on the analysis) drivers; c) an overall search rate higher than the average for their agency; and d) a rate of search for black (Hispanic) drivers at least twice that of white drivers. This allows a

[3] For logistic regressions, we also imposed a threshold that the agency had to conduct 10 searches of the racial group in question. Our analyses showed that odds-ratios became unstable with very low numbers of searches.

[4] Agency-level fixed effects, applied for individual states, showed results virtually identical to those reported here, but required further restrictions eliminating small agencies, so we have reported results here without agency fixed effects. Findings are robust to this specification issue.

conservative test of the “bad apple” hypothesis (see Baumgartner, Epp, and Shoub 2018 for more about the construction and interpretation of this variable). Most importantly, inclusion of the variable means that any effect for the race variables is over and above the effect driven by “a few bad apples” on the force.[5]

Illinois includes a variable for the age of the vehicle (or rather, model year, which we recode to be the age of the vehicle). Since wealthier people may replace their cars more often, we include vehicle age as a proxy for economic status, which of course is likely correlated with race. Therefore, if a race effect persists after controlling for vehicle age, it relates to the effect of race above and beyond that of economic status. Similarly, our controls for stop purpose include variables indicating that the driver was stopped for expired tags, cracked brake lights, and other poverty-related factors. In sum, if poverty leads to traffic stops because low-income individuals are more likely to drive older cars or those with equipment or regulatory issues, by controlling for these factors we can get a “clean” estimate of the race effect. If drivers are more likely to be searched when they are driving late at night, on the weekends, or for specific traffic violations, these effects will be captured with the control variables for day or week and time of day. Overall, these data represent the most complete models that we can estimate, given the data collected in each state, and they allow us to estimate the effect of being black or Hispanic, net of all these other factors.

Results

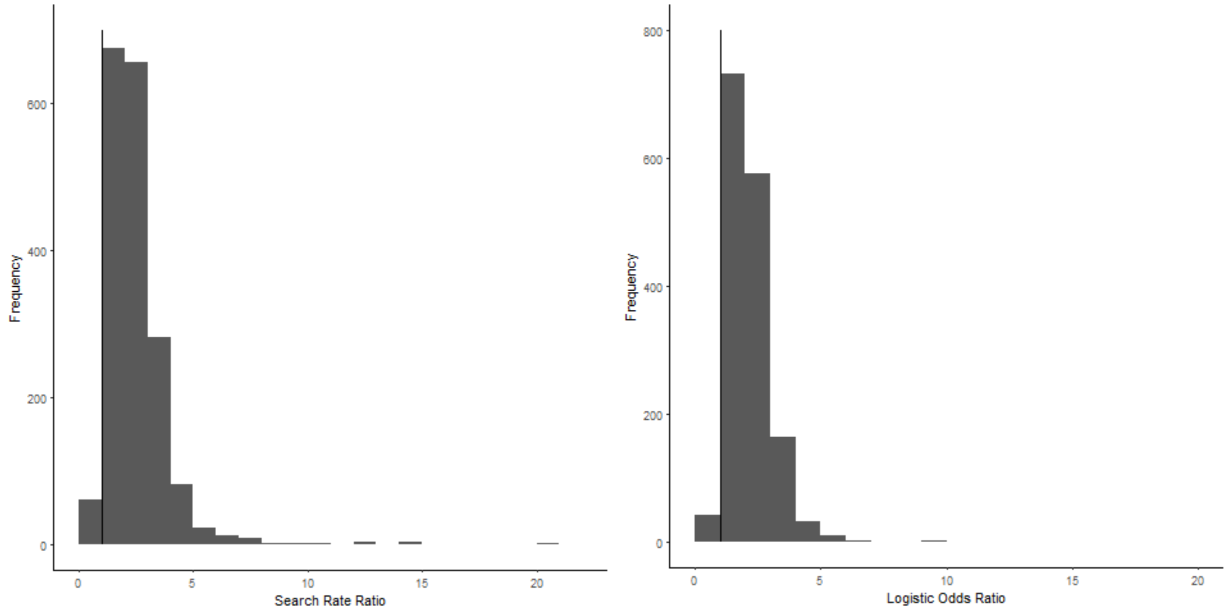
We estimate a logistic regression for every agency in our dataset, and compare it to the search rate ratio. Figure 1 shows the distribution of the search rate ratio and the odds ratio produced

[5] In North Carolina, approximately one-third of all officers were identified as “bad apples” using this analysis.

from the logistic regression for black-white (A) and Hispanic-white (B) comparisons. There is a vertical line at 1, indicating the point of equality.

Figure 1. Search Rate Ratios Distributions

A. Black-White Comparisons



B. Hispanic-White Comparisons

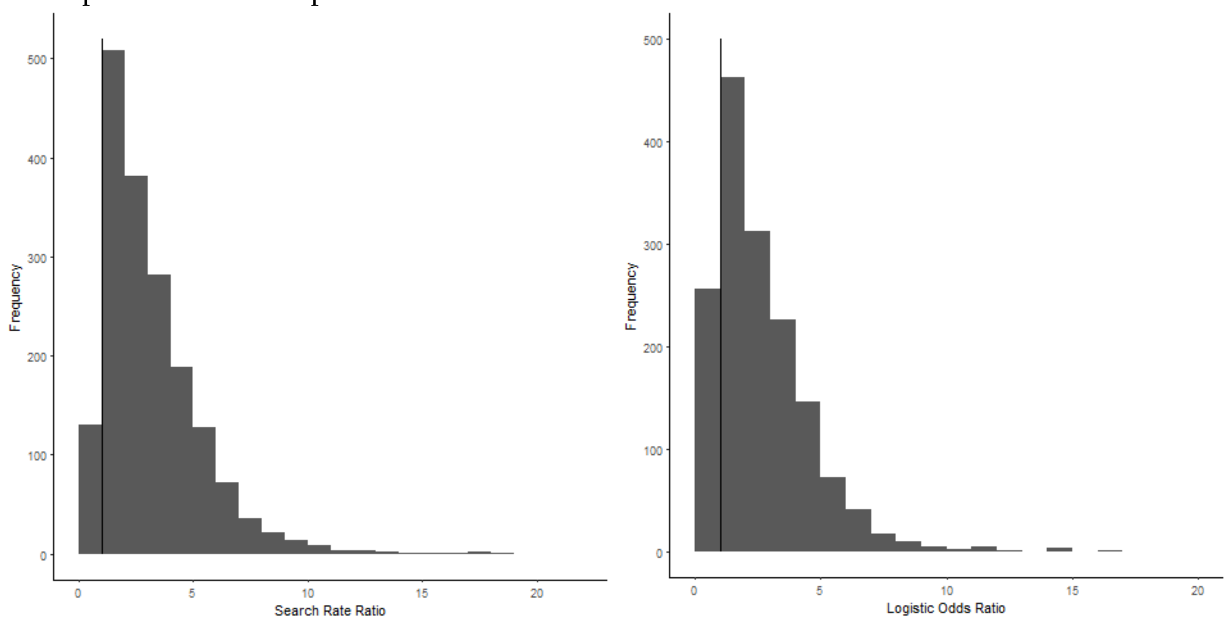


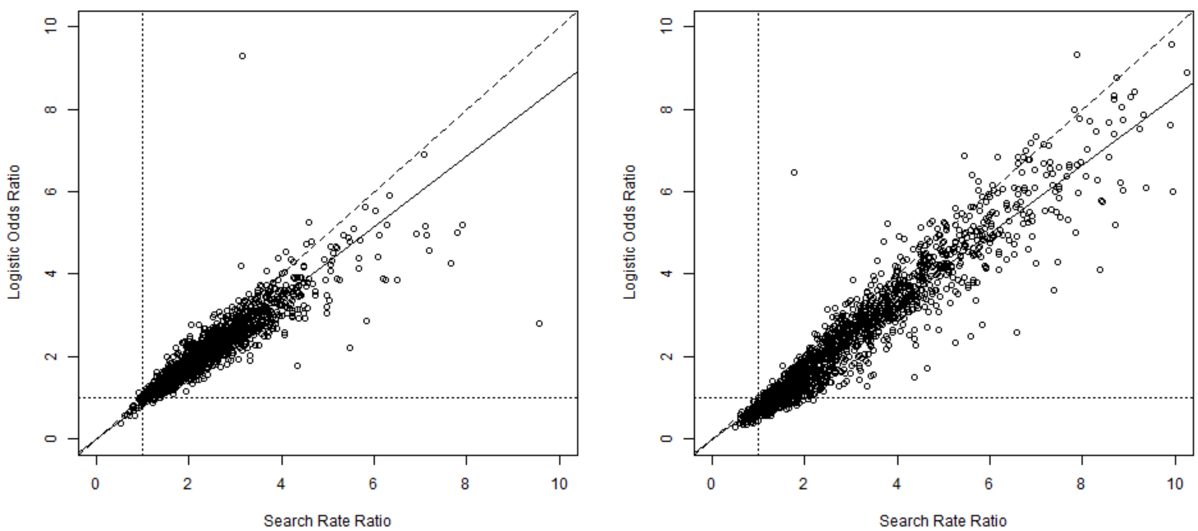
Table 3. Descriptive Statistics

	Min	1st Q.	Median	Mean	3rd Q.	Max
Black Logistic Odds Ratio	0.37	1.54	2.01	2.14	2.56	9.91
Hispanic Logistic Odds Ratio	0.28	1.22	2.21	2.64	3.52	16.83
Black Search Rate Ratio	0.00	1.64	2.20	2.42	2.92	20.66
Hispanic Search Rate Ratio	0.00	1.65	2.61	3.22	4.16	72.85

The majority of our data falls to the right of the vertical line, with search rate ratios and logistic odds ratios that exceed 1. This means that, on average, black and Hispanic drivers are more likely to be searched than white drivers, even after contextual factors are taken into account. These racial disparities are large, widespread, and race appears to be a statistically robust factor in a police officer's decision to search a car. The median value of every value falls between 2.01 and 2.61, indicating that minority drivers are more than twice as likely to be searched. The logistic odds ratios do tend to be slightly lower than the raw search rate ratios, indicating that some proportion of the disparities observed is indeed explained by non-racial factors. But the high values for the logistic odds ratios on average mean that race remains a significant factor. Note, however, that the number of observations below the 1.0 equality indicator is considerably higher for the logistics odds ratios, as are the number of cases only slightly to the right of the line. Clearly, some part of the raw disparities are indeed explained away. But not most.

Figure 1 makes clear that black and Hispanic drivers are more likely to be searched in almost every police jurisdiction where we have data. We want to know, however, whether for any given agency the simple ratio accurately predicts the logistic odds-ratio. To do this, we first calculate the correlation between the two sets of values. Second, we regress the odds ratios from the logistic regressions on the search rate ratios. To find evidence for our claim that a simple search rate ratio captures much of the same information as a logistic, the two sets of measures should be highly correlated. Figure 2 plots the search rate ratio against the odds ratio, for black-white and Hispanic-white comparisons.

Figure 2. Search Rate Ratios Compared to Odds Ratios from Logistic Regressions
A. Black-White Comparisons B. Hispanic-White Comparisons



Each point is a particular agency-period observation. The graph is scaled from 0 to 10 on each axis, and each graph includes a dotted line representing a perfect 1:1 relationship between the search rate ratio and the logistic odds ratio. The regression line, which is depicted as a solid black line, goes directly through the data points and we can compare this to the dashed line, which depicts equality. For both the A and B panel of Figure 2, the black line hews slightly to the right of the dashed line, revealing that the odds-ratio does tend to be slightly lower than the search rate ratio. This reflects the fact that there are indeed some systematic differences between minority and white drivers. The relationship between the simple search rate ratio and the logistic odds ratio is not exactly 1:1, so the covariates included in the logistic regression do have some explanatory power. Table 4 gives the slope of these lines: 0.86 for Blacks, and 0.83 for Hispanics. In other words, about 15 percent of the observed difference in search rates can be “explained away” by other factors such as why the car was pulled over. So, there is some value in that extra information, but the vast bulk of the observed racial difference remains even after controls. Table 4 reports regression results corresponding to Figure 2, overall and separately for each state.

Table 4. Comparing Search-Rate Ratios and Odds Ratios from Logistic Regressions

Driver Race Comparison	State	Slope	N	Adjusted R^2
Black – White	All States	0.86 (0.00)	1,557	0.97
	IL	0.85 (0.00)	1,002	0.97
	MD	0.85 (0.01)	59	0.96
	NC	0.89 (0.01)	482	0.99
Hispanic – White	All States	0.83 (0.00)	1,558	0.97
	IL	0.85 (0.00)	1,003	0.98
	MD	0.79 (0.01)	59	0.98
	NC	0.64 (0.01)	482	0.96

Note: All slope estimates are statistically significant at the 0.05 level. Entries are the results from a regression predicting the logistic odds-ratio using the search rate ratio as the predictor variable, with no constant term (standard errors are in parentheses). Separate results are not presented for Connecticut because there are only five observations in that state.

Turning to the substance of the results produced, note that the absence of racial disparity in traffic stop outcomes would result in a clustering of all the data points at 1.0 on either scale. Any number above 1.0 on the search rate ratio (the x-axis) indicates that minorities are searched at a higher rate than whites. Any number above 1.0 for the logistic odds ratio (the y axis) indicates that minorities have a higher chance of being searched following a traffic stop than the baseline racial category, whites. Figure 2 demonstrates that these values both consistently exceed 1, indicating that there is a higher likelihood of searches, following a traffic stop, for minority drivers than for white drivers. For the black-white comparison, the bulk of the data clusters between 1 and 4, indicating that black drivers are searched up to a rate of four times that of whites, on average. For the odds ratios, this disparity is as great as 9 and for the search rate ratios, almost 10. For the Hispanic-white comparison, the bulk of the data clusters between 1 and 6 for the search rate ratio and between 1 and 5 for the odds ratio. There is a lot of variance, especially with the Hispanic-white comparisons, but in some cases Hispanics are searched up to 10 times that of whites.

If these differences were to be explained away by the statistical covariates, then the odds ratio (y-axis) would be at or below 1.0, and the slope of the regression line would approach zero. While we still might except to see high variance along the x-axis, for the search rate ratio, the y axis should be clustered around 1.0 because the effect of race would be zero—white and minority drivers would be searched at a 1:1 rate and thus, the odds ratio would equal 1. Such a result would suggest that the search rate ratio, which appears to point to large racial differences, is being confounded by covariates. Clearly, however, this is not the case. Where the search rate ratio is high, so is the odds-ratio. The two correlate almost perfectly, as indicated by the high R^2 values.

Figure 2 and Table 4 make clear that the search rate ratio is a robust indicator of racial disparity in the likelihood of search. Where the search rate ratio is extremely low, so is the odds-ratio produced from a logistic regression. This means that the easily computed and highly prevalent search rate ratio is actually a robust measure of racial disparity. A logistic regression, even when it controls for a variety of contextual factors, conveys much of the same information as the simple search rate ratio. Crucially, contextual factors do not "explain away" the racially disparate results we find in almost every jurisdiction studied.

A Deeper Dive into Officer and Neighborhood Effects in One City

Even though our regression models use the most complete data available, there may still be factors that are unaccounted for because they are not collected by the state or agency. The neighborhood or precinct in which the stop occurred, or the officer involved in the stop, may play a role in determining whether or not the driver is searched. On the one hand, responsible policing may require that police chiefs allocate more officers to patrol high-crime neighborhoods, which would naturally result in more stops and, likely, more searches simply based on neighborhood. On the other hand, if officers use a lower threshold in their decision to search a driver simply due to the location of the stop, this process may be inherently tied to and informed by race. Perhaps neighborhood effects generate the racial disparities, but the data do not allow us to test for this. Most states and agencies do not collect this level of detailed data, but we do have such data for the city of Charlotte, North Carolina.

Charlotte is the largest city in North Carolina and it has sufficiently large populations of white, black, and Hispanic residents for an analysis that compares each group. Not only is it a racially diverse city, but it is an economically diverse city with distinct neighborhoods that the police department has divided into 13 different divisions. Further, the city of Charlotte makes additional traffic stop data available through its open data portal, for the years of 2016 and 2017. This includes all the traffic stop information required by the state (e.g., summarized in Table 2 above) as well as information on the division within the department that made the stop and officer characteristics. This makes the Charlotte data set unique among cities; very few publicly available data sets provide officer characteristics or match it with the specific division within the department that makes the stop. We will examine this in two ways: first, by running regressions similar to those in the previous section with additional controls for division and officer characteristics; second, by matching stops based on driver, officer, and contextual information and then performing a difference of proportions test between different racial groups.

Before turning to this analysis, we introduce the data in two tables. Table 5 presents the search rates and search rate ratios by division. Table 6 presents the search rates and search rate ratios by officer race. Each presents the overall number of stops and overall search percent, then breaks out the number of stops and search percent by race for white drivers, black drivers, and Hispanic drivers. Other drivers (Asian drivers, Native American drivers, and Pacific Islander drivers) are excluded from the table. In the last two columns are the search rate ratios—black-white and Hispanic-white.

In Table 5, the raw search rates of both black and Hispanic drivers are uniformly higher than white drivers, and this is true in each division of the city. However, the extent to which these search rates differ seems to vary by division. The black-white and Hispanic-white search rate ratios always exceed 1, indicating that black and Hispanic drivers are always searched at higher rates than white drivers. The black-white search rate ratios vary widely, from 1.61 to 7.87. The disparities are consistent, but their magnitude is variable. For Hispanics, the variation is more muted. Hispanic drivers are always searched as higher rates than whites, as the Hispanic-white search rate ratio always exceeds 1, but it ranges from just 1.04 to 3.80.

Looking across officer racial groups, Table 6 demonstrates that again, black and Hispanic drivers are always searched at higher rates than white drivers, regardless of the race of the officer. The black-white search rate ratios always exceed 1 and, in Charlotte, black officers have the highest racial disparities in searches, at 4.01. White officers have a search rate ratio of 3.66 and Hispanic officers have a rate ratio of 1.90. For the Hispanic-white search rate ratio, all officers search Hispanic drivers almost twice as often as white drivers, with black officers at the highest level again (2.98). Note, however, that white officers have consistently higher search rates than other types of officers, in particular black officers. Black officers search just 2.5 percent of all drivers, where white officers search almost twice as many. Both sets of officers, however, maintain a high search rate ratio, targeting their searches particularly at minority drivers.

Table 5. Search Percentages and Minority: White Rate Ratios by Division, Charlotte, NC

			White Drivers		Black Drivers		Hispanic Drivers		SRR	
Patrol Division	Stops	Search %	Stops	Search %	Stops	Search %	Stops	Search %	B:W	H:W
Central	7,987	4.24	2,566	1.13	4,616	6.37	401	2.74	5.64	2.43
Eastway	17,560	3.56	4,236	1.58	10,014	4.60	2,510	3.63	2.91	2.29
Freedom	7,749	6.56	1,625	4.00	5,467	7.68	385	5.19	1.92	1.30
Hickory Grove	13,816	5.70	2,406	3.99	8,884	6.66	1,983	4.14	1.67	1.04
Independence	10,085	4.78	3,005	3.49	5,537	5.63	1,091	5.13	1.61	1.47
Metro	7,723	11.32	862	6.03	6,453	12.29	202	8.91	2.04	1.48
North	13,910	2.24	3,510	1.08	9,278	2.83	596	1.34	2.62	1.24
North Tryon	13,260	7.40	1,305	4.29	10,094	8.44	1,444	4.50	1.97	1.05
Providence	19,570	1.82	10,549	0.51	6,897	4.03	926	1.94	7.87	3.80
South	17,557	1.14	11,035	0.90	3,786	2.06	1,257	1.43	2.30	1.60
Steele Creek	11,526	4.29	3,491	2.15	5,583	5.34	1,930	5.70	2.48	2.65
University City	10,387	3.01	2,431	1.93	6,605	3.56	589	3.57	1.84	1.84
Westover	12,869	6.74	3,332	2.58	7,984	9.07	1,039	4.91	3.51	1.90
Total	163,999	4.34	50,353	1.73	91,198	6.12	14,353	3.97	3.54	2.29

Table 6. Search Percentages and Search Rate Ratios by Driver and Officer Race, Charlotte, NC

			White Drivers		Black Drivers		Hispanic Drivers		SRR	
Officer Race	Stops	Search %	Stops	Search %	Stops	Search %	Stops	Search %	B:W	H:W
White	118,210	4.91	36,774	1.91	65,663	6.99	10,237	4.19	3.66	2.20
Black	28,241	2.50	8,625	0.88	15,992	3.53	1,980	2.63	4.01	2.98
Hispanic	7,245	4.47	1,797	2.34	4,178	5.62	984	4.67	2.41	2.00
Other	10,636	2.55	3,534	1.61	5,287	3.06	1,217	3.70	1.90	2.29
Total	164,332	4.34	50,730	1.73	91,120	6.12	14,418	3.97	3.54	2.29

Beyond simply looking at the search percentages and search rate ratios, we can test whether the driver race effect disappears when we include controls for officer characteristics and the division within which the stop took place. The first way we do so is by replicating the logistic regression analyses that were estimated in the previous section. We integrate the new Charlotte-specific variables in a series of three models. The first model includes the variables included in the previous regressions. These are: dummy variables for driver race (white, black, Hispanic, or other), a dummy variable for gender (female or male), driver age, and stop purpose. (Unfortunately, the dataset from the open data portal does not include either time of day or day of week.) In the second regression, controls for officer characteristics are included. These are: dummy variables for officer race (white, black, Hispanic, or other), a dummy variable for gender (female or male), and officer years of service. In both the first and second regression, white drivers and officers are the reference categories for each variable respectively, and female drivers and officers are the reference categories for each variable respectively. Finally, fixed effects for division are included in the third model. Results are presented in Table 7.

Table 7. Explaining Who Gets Searched Following a Traffic Stop, Charlotte, NC

	Model 1	Model 2	Model 3
Driver Race: Black	2.96**	2.70**	2.20**
	(0.11)	(0.10)	(0.09)
Driver Race: Hispanic	1.38**	1.33**	1.12*
	(0.08)	(0.08)	(0.07)
Driver Race: Other	0.57**	0.54**	0.51**
	(0.06)	(0.06)	(0.06)
Driver Gender: Male	3.57**	3.42**	3.37**
	(0.12)	(0.11)	(0.11)
Driver Age	0.95**	0.95**	0.95**
	(0.00)	(0.00)	(0.00)
Officer Race: Black		0.50**	0.49**
		(0.02)	(0.02)
Officer Race: Hispanic		0.70**	0.71**
		(0.04)	(0.04)
Officer Race: Other		0.60**	0.69**
		(0.04)	(0.05)
Officer Gender: Male		1.80**	1.80**
		(0.09)	(0.09)
Officer Years of Service		0.94**	0.94**
		(0.00)	(0.00)
Stop Purpose	Included	Included	Included
City Division Control			Included
Intercept	0.05*	0.06 *	0.06*
	(0.01)	(0.01)	(0.01)
AIC	51,226	49,042	47,518
BIC	51,377	49,242	47,838
Log Likelihood	-25,598	-24,501	-23,727
Num. obs.	166,328	164,332	162,027

Note: ** $p < 0.05$; * $p < 0.10$. Entries are odds ratios with standard errors in parentheses.

Across all three models the coefficient associated with black drivers is greater than 1 and statistically significant at the 0.05 level. This indicates that black drivers are more likely to be searched than white drivers controlling for the context of the stop, officer characteristics, and geographic division where the stop took place. However, across the three models the coefficient does modestly change and decrease: a part, but by no means all, of the effect of race is correlated with officer characteristics and stop location. Black drivers, even once all contextual controls are used (model 3), are still over twice more likely to be searched following a traffic stop than white drivers, simply due to race.

A more complex story emerges for a comparison of Hispanic drivers to white drivers. While the coefficient is consistently above one, the statistical significance of the associated coefficient diminishes with additional controls. In the first and second models, those without division fixed effects, the coefficient is statistically significant at the 0.05 level. In the third model, that with division fixed effects, the coefficient is statistically significant only at the 0.10 level.

In addition to the race coefficients, we can also evaluate how different officer coefficients influence the probability of a driver being searched. As was seen in Table 6, non-white officers are less likely to search drivers, of any race, than their white counterparts. This is also apparent in models 2 and 3. Additionally, male officers are more likely to conduct searches than female officers. Finally, officers with more experience are less likely to perform a search than those with less experience. These officer characteristic variables are of great interest, and help explain stop outcomes. However, including them does not "explain away" the race effects we have consistently seen.

The Charlotte data allows an even more conservative analysis of possible differences in how often different groups are searched, exact matching. That is, we compare the search rates of black (or Hispanic) drivers who are otherwise identical to white drivers. To be considered "otherwise identical" (or a match), the observations must be identical along each of the following characteristics: Stop purpose (10 categories), year, month, patrol division (13 categories, see Table 5), gender and age (in years) of driver, and the race, gender, and years of service on the force of the officer who conducted the stop. Literally, the comparison is, say, a 27 year-old black male pulled over for speeding in a certain month in a certain neighborhood by a white male officer with six years of experience on the force, compared to an identically situated white driver, stopped by a similarly situated officer (or the same officer). With this methodology, we take advantage of every element of the dataset to compile a list of drivers who differ only by their race, and we look at the rates of search. We identify 9,220 matches for the black-white comparison, and 1,916 matches for the Hispanic-white comparison. Table 8 presents the results.

Table 8. Paired T-Test Using Matched Data for Search Rate Comparisons between Groups

		Percentage Searched		
Comparison	N	Minority Drivers	White Drivers	Search Rate Ratio
Black - White	9,220	2.10	1.10	1.90**
Hispanic - White	1,916	1.10	0.50	1.60 *

** $p < 0.05$; * $p < 0.10$. Note: dataset is limited to exact matches.

As can be seen, the percentage of drivers searched in this analysis is smaller than the overall search rate for Charlotte. This is due to the highly restrictive matching process; our N has gone from over 160,000 in Tables 5–7 to lower than 10,000 here. But the ratio of search rates makes a familiar pattern clear: otherwise identical black drivers are 90 percent more likely to be searched than white drivers; for Hispanics, the increased odds of search is 60 percent (both differences are statistically significant). Black and Hispanic drivers are more likely to be searched than their white counterparts in the most restrictive test that we can conduct.

Conclusion

Looking at more than 40 million traffic stops in the four states that provide the most extensive data, we have asked a simple question: Are minority drivers searched at higher rates than white drivers, and do these disparities remain after we control for potentially spurious or legally relevant factors that might explain them? The answer is that disparities are large and robust. They are partially reduced when we control for such factors as why the car was pulled over, the time of day, or the neighborhood where the stop occurred. However, significant disparities remain, even after controlling statistically for every variable made available. In our most restrictive and conservative test, there remains a 90 percent increased likelihood of stop for black drivers compared to whites, even for otherwise identical traffic stops. Few agencies are racially neutral in the odds of searching minority and white drivers after a routine traffic stop, and their greater rates of searching black or Hispanic drivers cannot be explained by "extraneous" factors, at least not any factors which are systematically collected by law enforcement officers. Note that the relevant laws mandating the collection of traffic stops statistics were uniformly motivated by concerns about the possibility of racial profiling (see Baumgartner, Epp, and Shoub 2018, chapter 2). Thus, the states mandated that such variables as the stop purpose be collected. Drivers pulled over for speeding do indeed have lower rates of search, and lower racial disparities in the search rates resulting from those stops compared to those drivers pulled over for poverty-related factors such as equipment problems or expired registration tags. Those factors do indeed explain a certain percentage of the racial disparities we observe. But the racial disparities remain after controlling for them. These simple search rate ratios should be considered valuable and meaningful indicators of racial disparities in policing.

"Driving while black" surged to the national consciousness and debate in the late-1990s. North Carolina was the first state in the nation to mandate the collection of demographic information on routine traffic stops. It is worth remembering the premise and the supposed promise of this legislation. In an editorial praising the bill, the Raleigh *News and Observer* wrote:

> The numbers ... should settle this issue of equitable treatment once and for all…. If the patrol is, as many blacks believe, unfairly targeting them, it must be stopped immediately. If not, the patrol deserves to be exonerated (Editorial Board 1999).

The data are in. Disparities are large, ubiquitous, troubling, and unexplained. But police agencies have changed from suggesting that disparities are unacceptable indicators of bias and must be eliminated to suggesting that unobserved factors explain the persistent differenced uncovered in virtually every police agency where they have been investigated.

Substantively, our conclusions are very troubling. In their review of the Ferguson, Missouri, Police Department, the US Department of Justice discovered that black drivers were 75 percent more likely to be searched than white drivers, after a traffic stop. In this analysis covering hundreds of agencies in four states, the average disparity is much higher. Across hundreds of police departments in four states, black and Hispanic drivers are searched, on average, at more than double the rate of white drivers. This disparity is not explained by different driving behavior, age, or other factors of the stop measured in the official data collection. Many possible causes may generate the disparities we have documented here. Individual, possibly implicit, bias on the part of officers is a strong contender given the widespread nature of the disparities. But institutional practices (e.g., promoting more "aggressive" policing behaviors in different parts of a city) may also be part of the story. We have given no insights here into what is generating these disparities, except to document that nothing in the current data collection protocols used by most police agencies explains them away. That is, the simple search-rate ratio is not a spurious finding. It is robust, and it points to very large differences in the way white and minority drivers experience policing on the roadway. If there are other factors absent from our datasets that might explain away these racial differences, then police departments should start collecting that data immediately. Existing datasets, which are extensive, point to large, widespread, and statistically robust, race-based differences. If they cannot be explained away by contextual factors, then they must be confronted and accepted for what they are.

References

Anderson, Kristin J. 2010. *Benign Bigotry: The Psychology of Subtle Prejudice*. New York: Cambridge University Press.

Baumgartner, Frank R., Derek A. Epp, Kelsey Shoub, and Bayard Love. 2017. Targeting Young Men of Color for Search and Arrest during Traffic Stops: Evidence from North Carolina, 2002–2013. *Politics, Groups, and Identities* 5, 1: 107–31.

Baumgartner, Frank R., Leah Christiani, Derek A. Epp, Kevin Roach, and Kelsey Shoub. 2017. Racial Disparities in Traffic Stop Outcomes. *Duke Forum for Law and Social Change* 9: 21-53.

Baumgartner, Frank R., Derek A. Epp, and Kelsey Shoub. 2018. *Suspect Citizens: What 20 Million Traffic Stops Tell Us about Policing and Race.* New York: Cambridge University Press.

Burch, Traci. 2013. *Trading Democracy for Justice.* Chicago: University of Chicago Press.

Correll, J., B. Park, C. M. Judd, and B. Wittenbrink. 2002. The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals. *Journal of Personality and Social Psychology* 83, 6: 1314–1329.

Eberhardt, Jennifer L., Phillip Atiba Goff, Valerie J. Purdie, and Paul G. Davies. 2004. Seeing Black: Race, Crime, and Visual Processing. *Journal of Personality and Social Psychology* 87, 6: 876–893.

Editorial Board. 1999. Who's being stopped? *Raleigh News and Observer*. February 19.

Epp, Charles R., Steven Maynard-Moody, and Donald Haider-Markel. 2014. *Pulled Over: How Police Stops Define Race and Citizenship.* Chicago: University of Chicago Press.

Geller, Amanda, and Jeffrey Fagan. 2010. Pot as Pretext: Marijuana, Race and the New Disorder in New York City Street Policing. *Journal of Empirical Legal Studies* 7, 4: 591–633.

Geller, Amanda, Jeffrey Fagan, R.R. Tyler, and B. Link. 2014. Aggressive Policing and the Mental Health of Young Urban Men. *American Journal of Public Health* 104, 12: 2321–27.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. 2007. An Analysis of The New York City Police Department's Stop-and-Frisk Police in the Context of Claims of Racial Bias. *Journal of the American Statistical Association* 102: 813–823.

Gilliam, Franklin D. Jr. and Shanto Iyengar. 2000. Prime Suspects: The Influence of Local Television News on the Viewing Public. *American Journal of Political Science* 44, 3: 560–573.

Glaser, Jack. 2006. The Efficacy and Effect of Racial Profiling: A Mathematical Simulation Approach. *Journal of Policy Analysis and Management* 25, 2: 395–416.

Glaser, Jack. 2015. *Suspect Race: Causes and Consequences of Racial Profiling*. New York: Oxford University Press.

Harris, David A. 1997. Driving while Black and all other Traffic Offenses: The Supreme Court and Pretextual Traffic Stops. *Journal of Criminal Law and Criminology* 87, 2: 544–582.

Harris, David A. 1999. The stories, the statistics, and the law: Why "driving while black" matters. *Minnesota Law Review* 84: 265–326.

Harris, David A. 2002a. *Profiles in injustice: Why racial profiling cannot work*. New York: New Press.

Harris, David A. 2002b. Racial profiling revisited: "Just common sense" in the fight against terror? *Criminal Justice* 17: 36–41, 59.

Lerman, Amy E., and Vesla M. Weaver. 2014. *Arresting Citizenship*. Chicago: University of Chicago Press.

Moore, Nina M. 2015. *The Political Roots of Racial Tracking in American Criminal Justice*. New York: Cambridge University Press.

Payne, B. Keith. 2001. Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon. *Journal of Personality and Social Psychology* 81, 2: 181-192.

Payne, B. Keith. 2006. Split-Second Decisions and Unintended Stereotyping. *Current Directions in Psychological Science* 15, 6: 287-291.

Peffley, Mark, and Jon Hurwitz. 2010. *Justice in America: The Separate Realities of Blacks and Whites*. New York: Cambridge University Press.

Petrocelli, Piquero, and Smith 2003, xxx missing cite

Pierson, Emma, Camelia Simoiu, Jan Overgoor, Sam Corbett-Davies, Vignesh Ramachandran, Cheryl Phillips, and Sharad Goel. 2017. A large-scale analysis of racial disparities in police stops across the United States. Working paper, available online: https://5harad.com/papers/traffic-stops.pdf.

Plant, E. A., and Peruche, B. M. 2005. The Consequences of Race for Police Officer's Responses to Criminal Suspects. *Psychological Science* 16, 3: 180–183.

Sagar, H. A., and J. W. Schofield. 1980. Racial and Behavioral Cues in Black and White Children's Perceptions of Ambiguously Aggressive Acts. *Journal of Personality and Social Psychology* 39, 4: 590–598.

Tillyer, Rob and Robin S. Engel. 2013. The Impact of Drivers' Race, Gender, and Age During Traffic Stops: Assessing Interaction Terms and the Social Conditioning Model. *Crime & Delinquency* 59 (3): 369-395.

Tillyer, Rob, Charles F. Klahm IV, and Robin S. Engel. 2012. The Discretion to Search: A Multilevel Examination of Driver Demographics and Officer Characteristics. *Journal of Contemporary Criminal Justice* 28 (2): 184-205.

Tomaskovic-Devey, Donald, Marcinda Mason and Matthew Zingraff, 2004. Looking for the Driving While Black Phenomena: Conceptualizing Racial Bias Processes and their Associated Distributions. *Police Quarterly*. 7:3–29.

Tonry, Michael. 1995. *Malign Neglect: Race, Crime, and Punishment in America*. New York: Oxford University Press.

Tyler, Tom R., Jeffrey Fagan, and Amanda Geller. 2014. Street Stops and Police Legitimacy: Teachable Moments in Young Urban Men's Legal Socialization. *Journal of Empirical Legal Studies* 11, 4: 751–785.

Tyler, Tom R., Jonathan Jackson, and Avital Mentovich. 2015. The Consequences of Being an Object of Suspicion: Potential Pitfalls of Proactive Police Contact. *Journal of Empirical Legal Studies* 12, 4: 602–636.

Voigt, Rob, Nicholas P. Camp, Vinodkumar Prabhakaran, William L. Hamilton, Rebecca C. Hetey, Camilla M. Griffiths, David Jurgens, Dan Jurafsky, and Jennifer L. Eberhardt. 2017. Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Science* (PNAS). www.pnas.org/ cgi/ doi/ 10.1073/ pnas.1702413114.