Benchmarking Traffic Stop Data:

Examining Patterns in North Carolina and the City of Raleigh

Frank R. Baumgartner

Richard J. Richardson Distinguished Professor of Political Science
UNC Chapel Hill
http://fbaum.unc.edu/
Frankb@unc.edu

February 1, 2022

The question of the proper statistical "benchmark" when reviewing traffic stop has been debated for many years. When considering agency or officer traffic stop data to assess claims of disparate enforcement, what is the most appropriate point of comparison to? One logical comparison is to the local population. Another might be to the driving population. Another might be to the driving population that is actually violating a law, justifying an enforcement action. So that benchmark question is clearly an important one.

To aid in the analysis of North Carolina traffic stop data, this article explores the benchmarking question and related implications. First, what are appropriate benchmarks? Second, how do the traffic stop patterns in North Carolina compare to appropriate benchmarks? Third, does available evidence suggest that black and white drivers in North Carolina are treated differently by officers? Fourth, are there important public safety benefits that might outweigh any evidence of disparate enforcement? And finally, what was the intent of the legislature in mandating that police agencies collect traffic stop statistics, and did the legislature mandate a particular benchmark?

What are appropriate benchmarks for comparison to traffic stops?

Census records are widely available and generally correspond to other available indicators of which drivers are eligible to be pulled over by an officer. Comparisons to the census population records are the most common benchmarks used by police departments themselves in their efforts to monitor officer enforcement patterns and build community trust.

Other possible benchmarks might include such things as:

a. Census data statistically adjusted to account for different rates of driving across race and gender categories, by time of day, or with other adjustments based on transportation models.
b. Driver's license data or motor vehicle registrations.
c. Observations of roadway usage, potentially focusing on violations of the traffic code.
d. Vehicle crash or accident data.
e. Officer-to-officer comparisons.
f. The "veil of darkness" technique, comparing traffic stops before and after sundown during the period around the shift to and from daylight savings time.

There is no question that option c) would be the favored solution; a team of researchers could be stationed at carefully selected street corners across a jurisdiction at selected time periods during the day and week and record the apparent race and gender of drivers passing by. Cameras and radar guns could be used; vehicles travelling at the speed limit could hold observers noting drivers who pass them (and who therefore are speeding). Such studies indeed have occasionally been done. One well-cited study reviewed traffic on the New Jersey Turnpike, revealing that over 90 percent of the drivers were speeding (Lamberth, John. 1994. Revised Statistical Analysis of the Incidence of Police Stops and Arrests of Black Drivers/Travelers on the New Jersey Turnpike between Exits or Interchanges 1 and 3 from years 1988 through 1991. Available online at: http://www.mass.gov/eopss/docs/eops/faip/new-jersey-study-report.pdf).

It seems unnecessary to focus on the share of law-breakers since such a large majority of drivers are in violation of the law. Others have done so and noted some differences, particularly

that blacks, young people, and men are more likely to be excessive speeders in 65 mile an hour zones (Tillyer, Rob, and Robin S. Engel. 2012. Racial differences in speeding patterns: Exploring the differential offending hypothesis. *Journal of Criminal Justice* 40: 285–295.). This option, however, is prohibitively expensive in most circumstances and not a feasible option for investigations of racial profiling allegations against a particular officer or in a particular criminal or civil case. These are large research projects typically designed to assess overall patterns in a police agency. For use in a single case, one would have to re-create such a large study for each individual, clearly an unreasonable demand.

Option f), the Veil of Darkness (VoD) technique, involves a strong assumption that racial differences in traffic stops are due to an officer first peering through a windshield and determining the race of the driver before deciding to pull the car over (see Grogger, Jeffrey T., and Greg Ridgeway. 2006. Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association* 101: 878–887). This is an inappropriate assumption, given the data below showing that black drivers constitute a much higher share of drivers at night rather than in the day; the logic behind the VoD method would lead to the conclusion that officers are targeting whites during the day when it is easier to observe who is driving. Rather, the dynamic may have to do with location rather than observations of the driver's race before the traffic stop is initiated; police departments can deploy officers differentially into areas more commonly frequented by black or white drivers. My published research has shown that vehicle age, vehicle type, and a number of other visible cues are correlated with differential police behaviors, so there is ample reason to believe that officers are able to use numerous visible cues beyond the observed race of the driver in decided with whom to engage. For these reasons the VoD methodology is flawed.

The most important limitation of these benchmarks is not so much how they differ from one-another. They do not differ like night and day; they differ only marginally. Therefore very large, or stark, statistical patterns would appear to be starkly different from the baseline no matter which benchmark is chosen. Further, no matter what benchmark one picks, a critic could always say that it is not perfectly accurate. What about out-of-town or out-of-state drivers? What about people who drive much more than others? What about observations at the precise location of the stop in question, or at the exact time of day, day of week, or time of year? Clearly, people drive across municipal boundaries on a routine basis. Some small North Carolina beach towns or tourist areas in the western mountain regions of the state see more traffic stops than they have local residents (unpublished finding from the research leading to *Suspect Citizens*). There is no question that sometimes the population benchmark is inaccurate because most drivers are not local. For a city the size of Raleigh, this is likely a small concern.

The main question about benchmarking is whether different benchmarks would lead to different conclusions, and whether it is reasonable to expect that a person alleging disparate enforcement of the law should have access to various types of data to constitute a benchmark. It is an important concern to question whether a critic of any baseline comparison could always "move the goalposts" to counter-argue that no baseline is good enough. In this area, the perfect can be the enemy of the good. Census comparisons, particularly for a large jurisdiction or the state as a whole, are good enough.

Census data is often relied on as an adequate benchmark

One indication that census comparisons are good enough is that police departments, highway patrol agencies, and other law enforcement agencies use these numbers themselves in their own publications. The US Department of Justice used census numbers as a benchmark in its

investigation of the Ferguson MO Police Department (U.S. Department of Justice (USDOJ), Civil Rights Division. 2015. *Investigation of the Ferguson Police Department.* Washington: US DOJ. March 4.) as well as its investigation of the Baltimore City Police Department (U.S. Department of Justice (USDOJ), Civil Rights Division. 2016. *Investigation of the Baltimore City Police Department.* Washington: US DOJ. August 10). In the Ferguson case, it used city-wide census numbers whereas in the Baltimore case it used residential population numbers in the various patrol districts: "In each of BPD's nine police districts, African Americans accounted for a greater share of BPD's stops than the population living in the district" (DOJ Baltimore investigation, p. 7).

The State of Illinois produces extensive reports comparing each police agency's traffic stops with population "benchmarks"; these are based on census estimates of the driving-age population for each jurisdiction in conjunction with driver's license data adjusting the population data for the share of those in each age group with a driver's license; these numbers are generally zero for those under age 16, approximately 80 to 90 percent for those between ages 16 and 80; p. 37). The numbers also differ by race, with 72 percent of whites but 68 percent of blacks having a driver's license in Illinois (p. 38). The authors of the more recent (2020) report note: "While this methodology has some limitations (…), it provides a transparent, standardized approach to developing benchmarks for the nearly 1,000 law enforcement agencies in Illinois" (Illinois Department of Transportation. 2021. Illinois Traffic and Pedestrian Stop Study. 2020 ANNUAL REPORT: TRAFFIC STOP ANALYSIS. Part I Executive Summary and Appendices. https://idot.illinois.gov/transportation-system/local-transportation-partners/law-enforcement/illinois-traffic-stop-study, p. 34). This would correspond to an "adjusted census

data" methodology, since it is based on census reports, adjusted by an estimate of the likelihood that a person has a driver's license.

It is clear that cities have racially segregated neighborhoods and that officers patrolling in any given area may not encounter the same racial mix of drivers as officers in other areas. Therefore, it is useful to review data for an entire city and an entire police department. Further, generally one does not know where an officer patrols, and we cannot assume that all of an officer's traffic stops are in the same area or areas. Finally, there is a question of whether benchmark data are available for particular patrol areas, neighborhoods, or quadrants of the city. In most North Carolina cities with which I am familiar, they are not.

More important than whether census or other benchmark data are available for small geographic units (such as neighborhoods or patrol districts), the state's traffic stops database does not indicate the location of a traffic stop. The SBI-122 form (the form that officers fill out after conducting a traffic stop) does not indicate the location of the stop other than to note the city or county of the stop and the police agency for which the officer works. Therefore, the traffic stops database cannot be used to assess neighborhood-level differences in rates of stop.

Census data comparisons understate racial disparities in law enforcement

It is important to understand that census comparisons are generally biased against any finding of disparate treatment with regards to race. That is, for a black or other minority driver, use of the census generally would be an adverse benchmarking tool. (This is not in itself a good thing, to be clear, but it represents a high bar, rather than a low bar, for an allegation of disparate treatment.) It is well documented that whites typically have greater access to vehicles and tend to drive longer distances than minority drivers. This is consistent with what we wrote in *Suspect Citizens*:

One reason to proceed with a population comparison is that, on average, whites drive more than blacks and Hispanics. This is because having a driver's license, owning a car, and driving regularly are all more common among white Americans than black Americans or Hispanics [citations omitted]. If whites drive more than blacks or Hispanics, and we show with population comparisons that whites are less likely to be pulled over, then it means that whatever disparity we document using population numbers would be an under-, not an over-estimate of the true degree of disparity. (*Suspect Citizens*, p. 65).

We reviewed a number of factors in *Suspect Citizens*; compared to whites, blacks were 39% less likely to own a car; they drove 13 to 19 percent less per year; they were 13 percent less likely to be a driver; they made 18 percent fewer trips per day (see *Suspect Citizens*, Table 3.3, p. 75).

Of course, there are differences among the benchmark statistics, and some allow us to incorporate time of day into the analysis while others do not. When considering the inferences raised by the traffic stops-to-benchmark comparison, the starkness of the difference is a significant factor. Where an officer or agency's traffic stop data differs only slightly from the census population data benchmark, one may be less confident in a conclusion of disparate enforcement. Where the difference is more extreme, the benchmark may be treated as reasonably reliable. Again, whatever the degree of disparity, it should be understood as an understatement rather than an overstatement of the actual disparities in enforcement.

When considering selective enforcement of the traffic laws, how much data is sufficient to demonstrate a problem?

When comparing officer, agency, or statewide data to census population data, the question becomes how much data is enough to demonstrate cause for concern? We demonstrated in *Suspect Citizens* that blacks are a greater share of drivers pulled over than of the population. We

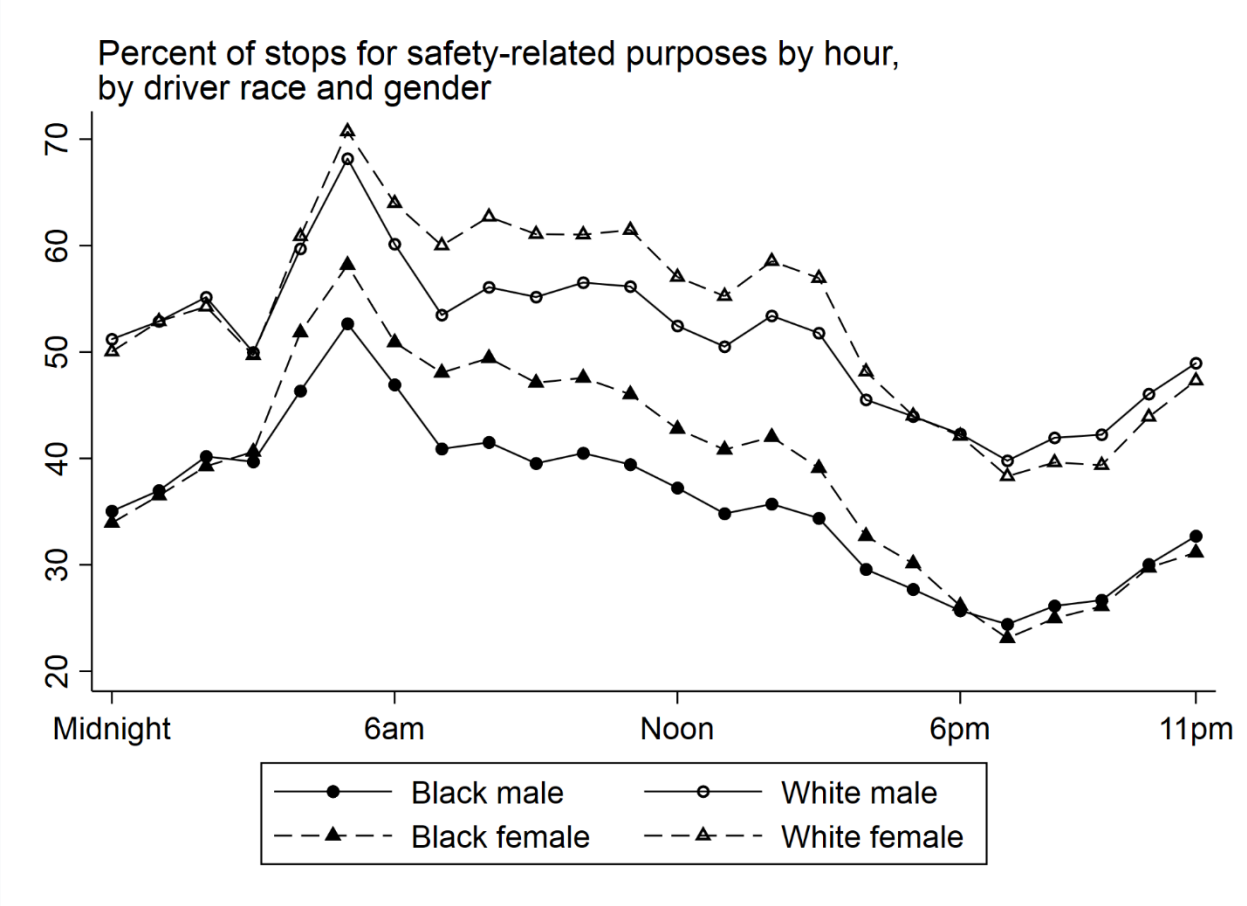
estimated (*Suspect Citizens*, table 3.1, p. 68) that, across North Carolina, black drivers are 63 percent more likely to be pulled over than whites, as a share of the population. As discussed above, we further noted that this is an under-estimate since the US Department of Transportation and others estimate that blacks drive, on average, about 15 to 20 percent less than whites (*Suspect Citizens*, table 3.3, p. 75). Taking both of these into consideration, we estimated that a black driver is approximately 96 percent more likely to be pulled over compared to a white driver (*Suspect Citizens*, p. 76). In plain English, blacks are about twice as likely to be pulled over compared to whites. (One pulled over, we showed that they are about twice as likely to be searched, as well; see *Suspect Citizens*, p. 86.)

Are black and white drivers treated differently by officers in the city of Raleigh, and state-wide?

Data collected through the SBI 122 form allow clear demonstrations that black drivers are treated differently compared to whites and that such differences reflect constitute unequal treatment based on race.

Figure 1 shows how drivers of different race and gender characteristics are more or less likely to be stopped for "investigatory" reasons rather than traffic safety reasons. The SBI-122 form lists ten stop "purposes." In this graph I classify speeding, stop sign violations, DWI, and "unsafe movement" as "safety-related purposes" and all other stop purposes (such as equipment violations, regulatory stops [e.g., expired tags] as other. (I exclude checkpoint stops throughout this report because those are recorded only when there is an adverse action such as a search or arrest of the driver, rendering them incomplete.) Safety-related stops are moving violations and may be more likely to be related to keeping the roads safe. Other traffic stops may be more related to the desire to conduct an informal criminal investigation.

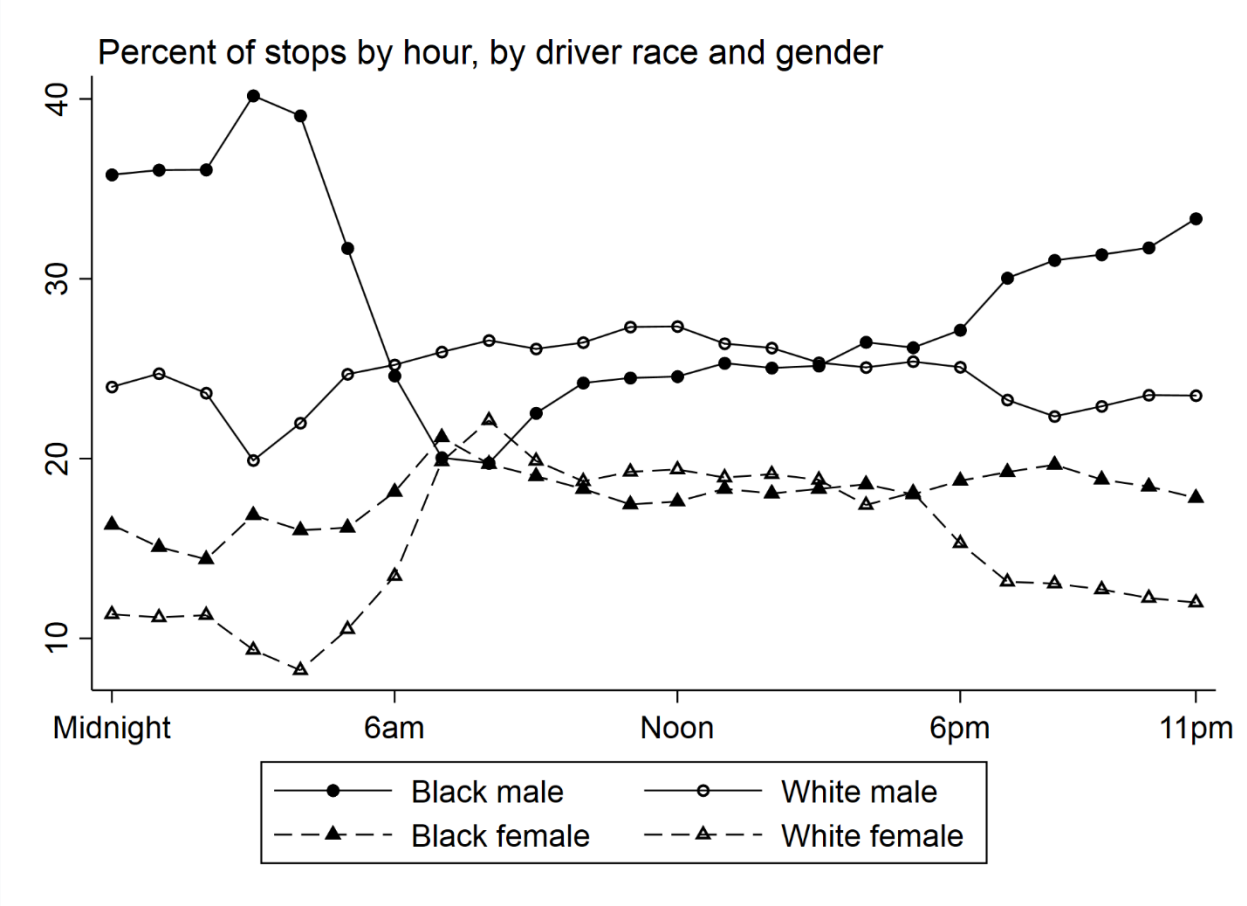
Figure 1. Percent of stops for safety-related purposes, by race, gender, and hour of day, Raleigh PD, 2002-2020.



Black drivers, whether male or female, are much more likely to be stopped for reasons other than moving violations, particularly in the evening and at night. White drivers, particularly white women, are much more likely to be stopped for moving violations. The numbers are stark: at any given time of the day (the data combines all stops from 2002 through 2020 based on the time of day), white drivers may be 20 percent higher in their share of safety-related stops than black drivers. At noon, for example, black male drivers show a value of 40 percent safety-related stops but white female drivers show a value around 60 percent.

Figure 2 shows how drivers of different race and gender characteristics represent a different share of overall traffic stops by time of the day.

Figure 2. Percent of traffic stops by race, gender, and hour of day, Raleigh PD, 2002-2020.



Black males are more than 35 percent of those pulled over during the period following midnight, but only 20 percent of drivers stopped during the morning daylight hours. White men, by contrast, show a relatively consistent pattern throughout the day. Why would this be? It may be because the traffic stop is a convenient tool for a police officer to conduct an informal criminal investigation. Officers can find a reason to pull over virtually any car, and without unduly prolonging the traffic stop, they may ask questions about the driver's identity and plans. This may involve a request to search the car, or a decision that probable cause of some kind of contraband or criminal activity is apparent.

Figure 3 shows the percent of drivers subjected to a search, and Figure 4 shows the percent of drivers arrested. With the exception of driving while impaired, most traffic or vehicle infractions do not involve an arrest. Rather, arrests come from criminal investigations relating to

the presence of contraband. When traffic stops divert from the purpose of keeping the roads safe from poor driving behaviors and into informal criminal investigations, as they often do, then black drivers, particularly males, are subjected to significantly more scrutiny.

Figure 3. Percent of drivers searched by race, gender, and hour of day, Raleigh PD, 2002-2020.

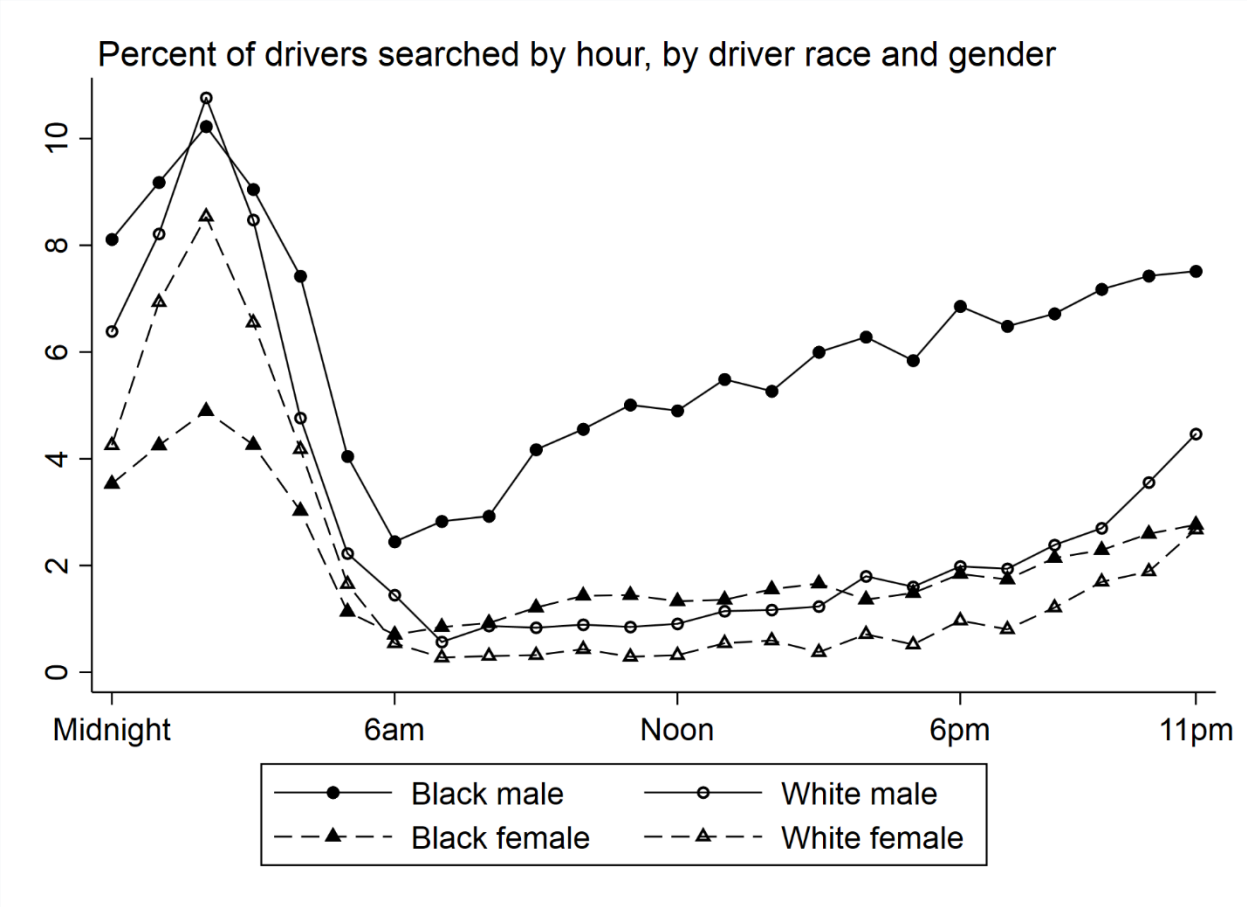
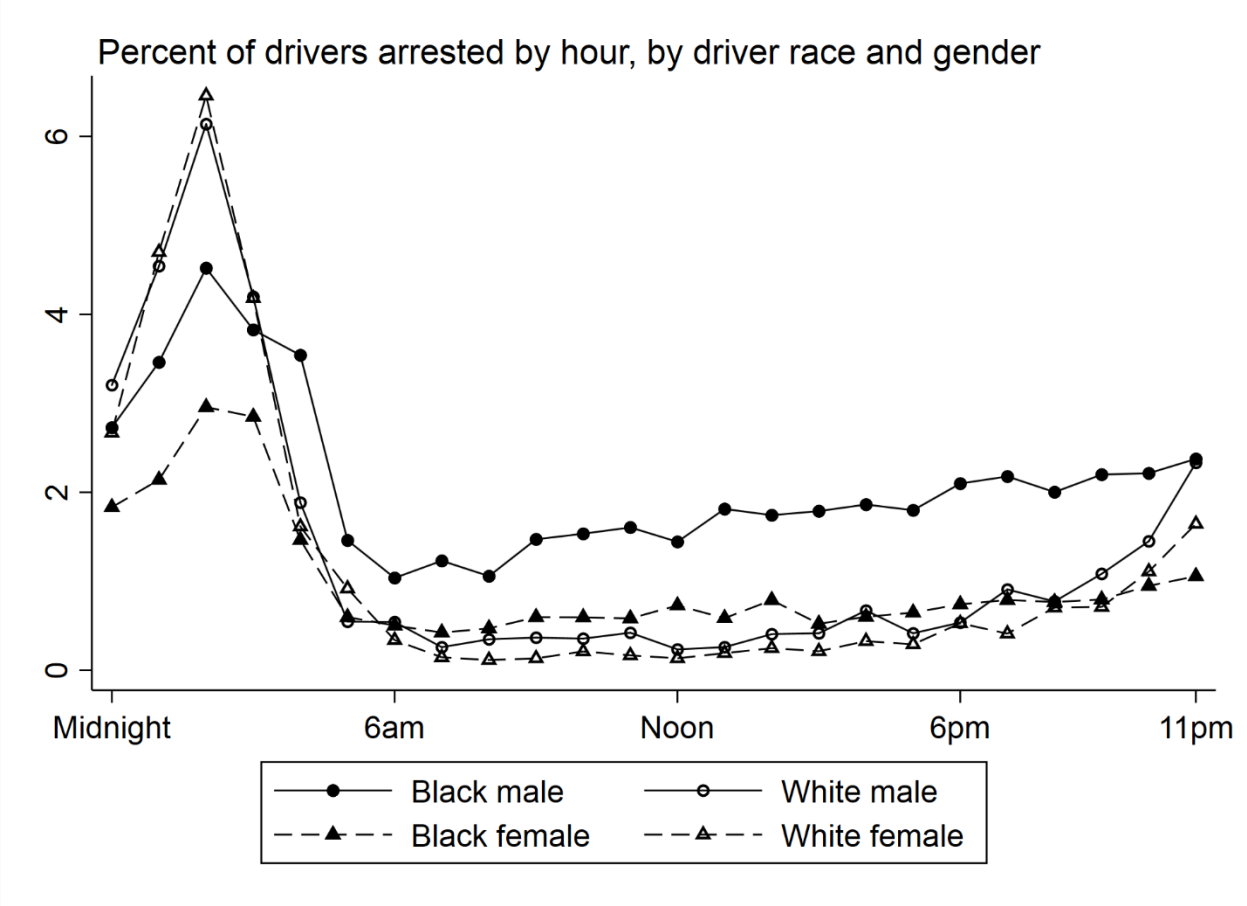


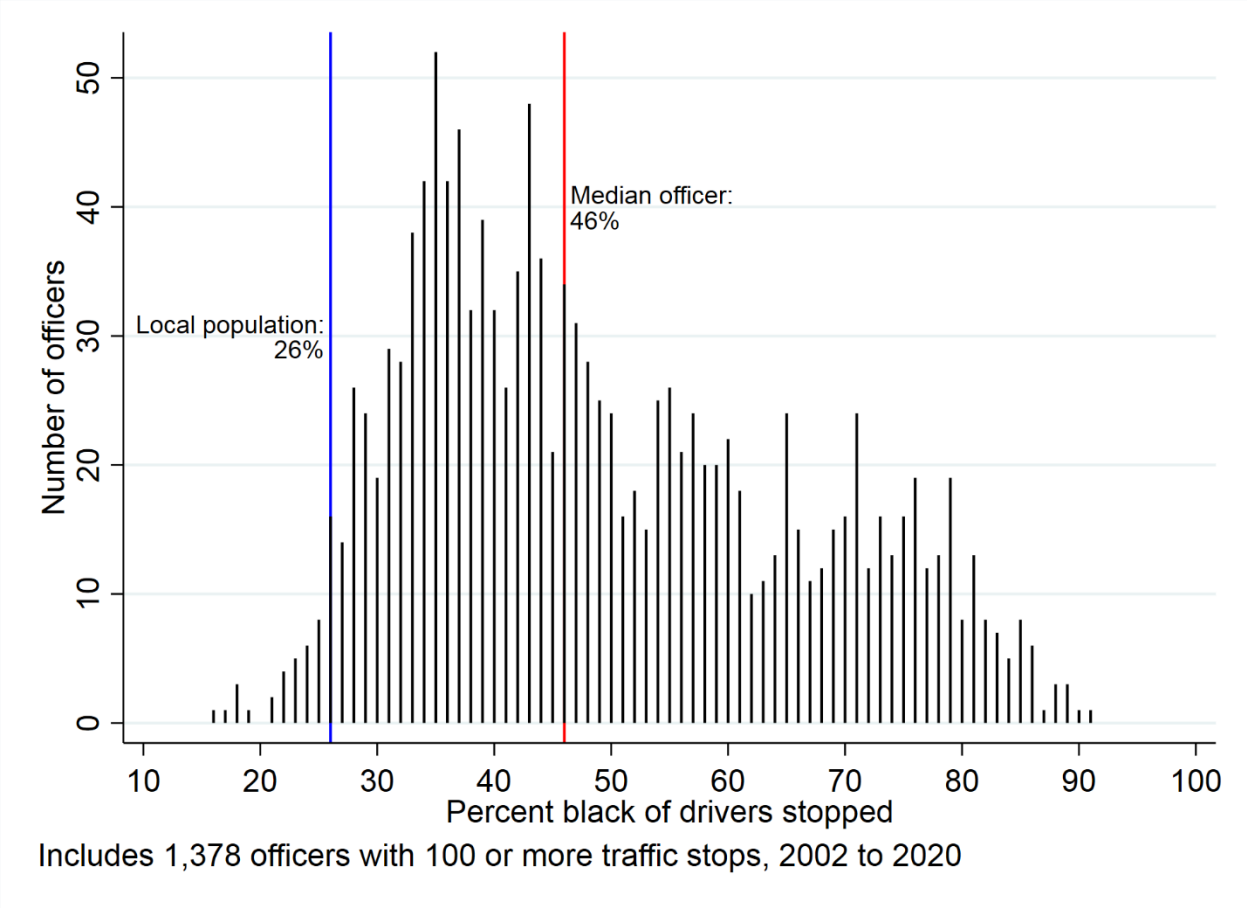
Figure 4. Percent of drivers arrested by race, gender, and hour of day, Raleigh PD, 2002-2020.



Several things are noteworthy with regards to Figures 3 and 4. First, rates are very low; the vast majority of traffic stops involve neither a search nor an arrest, including for black drivers. Second, they vary dramatically by time of day. It appears that the very purpose and function of a traffic stop differs between daylight, evening, and the after-midnight period. And finally, there are strong differences in odds of search and arrest by driver race, with black male drivers being much more likely to see these outcomes.

Finally, we can look at the racial mix of drivers pulled over by each individual police officer on a given force. Figure 5 shows such results for over 1,300 officers in the Raleigh Police Department, all those who reported at least 100 traffic stops during the period of 2002 to 2020. For each officer it reports the percentage of drivers pulled over who were black.

Figure 5. Percent black of drivers pulled over by Raleigh PD officers, 2002-2020.



As indicated by the blue line toward the left of the Figure, the population of Raleigh is approximately 26 percent black. However, just 41 officers, or less than 3 percent of those shown, have a lower share of black drivers. If traffic stops were racially neutral with regards to the local population, half of the officers would be below the population average and half would be above it. Instead, we see that the average officer pulls over 46 black drivers out of 100 traffic stops, nearly twice the population average, as indicated by the red line. A significant share of officers pull over 75 or more black drivers for every 100 traffic stops they make. For any given officer, the further to the right on this graph, the greater the disparity of their stops compared to the population average. Of course, some of this could be explained by the population characteristics of their patrol district, which might not reflect the city-wide average. Still, if the policing practices of the Raleigh PD were neutral with regard to race, we would see a very different

distribution. Recall that virtually every driver routinely breaks a traffic law, generally on each trip. Either the Department systematically assigns more officers in areas of Raleigh where more black drivers are on the street, or individual officers are more likely to pull over black drivers than whites for other reasons. In any case, the odds of a black driver being pulled over are much higher than those of a white driver. This is consistent with what we showed state-wide in *Suspect Citizens*.

Are there important public safety benefits that might outweigh any evidence of disparate enforcement?

The policies and practices that lead to patterns such as those laid out in the previous section, which in my experience are typical of police traffic stops statistics across the state and nationwide, are not very effective. From the beginnings of the development of the police strategy of using the traffic and vehicle codes to fight the war on drugs and the war on crime, there has been a clear understanding that it amounts to a "needle in the haystack" strategy. Thousands of motorists would have to be momentarily detained and evaluated, and perhaps searched, in order to find a few who might be involved in transporting narcotics or other contraband, or otherwise involved in criminal behaviors. We wrote this in *Suspect Citizens*:

> "When asked about the logic behind the Drug Enforcement Agency's (DEA) Operation Pipeline, a California Highway Patrol sergeant famously explained that, 'It's sheer numbers. Our guys make a lot of stops. You've got to kiss a lot of frogs before you find a prince'" (*Suspect Citizens*, p. 98).

One important question, then, is whether the "numbers game" that the police engage in is a valuable tool to fight crime or, rather, a waste of resources. It is clear from the statistics above that the burden of these policies falls disproportionately on certain segments of the population. Indeed it could be that most middle-class white Americans are unaware of these policies, since such individuals are rarely exposed to them. In minority communities, however, this burden can

weigh heavily, explaining very different attitudes about, and trust in, the police. Table 1 lays out some basic statistics regarding the use of traffic stops as tools in fighting the war on crime. The numbers suggest that traffic stops are a very inefficient mechanism for finding contraband. Therefore, it is a reasonable conclusion that the public-safety or crime-fighting value of these policies in no way outweighs the cost of them in terms of equal enforcement of the law.

Table 1. Summary of outcomes of statewide traffic stops, 2002 to 2020.

Category	State-wide	Raleigh PD
Stops	24,980,776	1,045,289
Traffic Safety Stops	13,365,910	451,793
Searches	763,343	37,836
Hits	280,152	7,545
Arrests	500,040	15,467
Arrest From Hit From Search	105,784	2,628
Consent Searches	346,475	9,808
Arrest From Hit From Consent Search	20,759	248
Probable Cause Searches	264,963	15,176
Arrest From Hit From Probable Cause Search	54,326	1,770
Percent Traffic Safety Stops	53.50%	43.22%
Search Rate Per Stop	3.06%	3.62%
Hit Rate Per Search	33.48%	19.13%
Arrest Rate Per Hit	38.85%	35.03%
Hit-and-Arrest Rate Per Search	13.86%	6.95%
Hit-and-Arrest Rate Per Probable Cause Search	20.50%	11.66%
Hit-and-Arrest Rate Per Consent Search	5.99%	2.53%
Arrest Rate Per Stop	2.00%	1.48%
Hit-and-Arrest Rate Per Stop	0.42%	0.25%

Source: Summary of Traffic Stops Statistics: Raleigh Police Department. November 6, 2021. http://fbaum.unc.edu/TrafficStops/Reports2021/Baumgartner-TrafficStopsSummary-Raleigh.pdf

During the period of 2002 through 2020, the Raleigh PD has made over 1 million traffic stops. Of these, 451,000 have been safety-related stops and the rest have been non-moving violations such as equipment violations, expired tags, and other stop purposes likely to be associated with informal investigations rather than keeping the roads safe. Searches have followed the traffic stop 3.62 percent of the time. Contraband has been discovered in 19 percent

of those searches (so over 80 percent of the searches led to no contraband, a very high rate). The driver has been arrested in 35 percent of the occasions when the search led to contraband. This means that, even when an officer discovers contraband following a stop, 65 percent of those drivers are not arrested. This could be because the contraband was very small in amount, because the contraband was not necessarily illegal in itself (e.g., a spoon, a baggie, or something else that could be considered "drug paraphernalia" but which is not otherwise illegal to possess), or it could be for other reasons. In any case, just 35 percent of those found with contraband were arrested, according to the Raleigh PD statistics; this equates to 7 percent of those searched, and 0.25 percent of those pulled over. (These are not dissimilar to the state-wide averages also shown the table, though the Raleigh PD numbers are even lower.)

When we combine all of these statistics into final outcomes, and whether we look at consent searches, probable cause searches, or all searches and contraband hits combined, we see that vast numbers of Raleigh and North Carolina drivers, are subjected to searches that do not lead to arrest. For the Raleigh PD, just 0.25 percent of traffic stops lead to an arrest following the discovery of contraband in a search. Overall, following more than 1 million traffic stops and almost 38,000 searches, we see a net total of just 2,628 arrests following the discovery of contraband. That means that more than 35,000 individuals were subjected to "fruitless search." Fewer than seven percent of searches conducted by the Raleigh PD, and less than 14 percent state-wide, led to arrest-worthy contraband hits. In other words, 93 percent of searches were fruitless.

If we look at the numbers of individuals subjected to search following a traffic stop, state-wide or by the Raleigh PD, we find that the vast majority of the searches were "fruitless" in the sense that they did not lead to the arrest of the driver. Statewide, the summary table above shows

that over 760,000 searches were conducted, but 105,784 arrests were made following the discovery of contraband in such a search. Subtracting one number from the other leads to the conclusion that more than 600,000 people have been subjected to search following a traffic stop, state-wide, that proved to be exonerating rather than incriminating, in the eyes of the officer who conducted the search. That is, if the officer had determined there was cause for arrest following the search, then an arrest would have followed. The fact that 600,000 individuals have been subjected to such fruitless searches should be cause for concern. In Raleigh alone, the number is 37,836 searches minus 2,628 that led to contraband and arrest, leaving 35,208 individuals searched fruitlessly. Of course, legal cases and court appearances would rarely follow these fruitless searches since the individual would not have been arrested. But our courts should bear in mind that for every one person arrested following a contraband hit in a search in the context of a traffic stop, many more individuals are subjected to a fruitless search.

Traffic stops are frustrating and can be dangerous. Searches following from traffic stops can be humiliating, time-consuming, and they can have collateral consequences. Certainly, the experience of having one's car searched indicates to the driver that they are the object of suspicion from the officer. The fact that more than 90 percent of the searches conducted by the Raleigh PD, and 85 percent of the searches conducted state-wide, do not lead to the arrest of the driver following the discovery of contraband is a clear indication of a wasteful practice. Nationwide, wide-spread police practices involve taking advantage of the fact that virtually every driver is breaking some element of the extensive traffic or vehicle codes in order to conduct an informal criminal investigation. Because these practices are rarely targeted on white middle class individuals, many may be unaware that they are even happening. Because judges typically review only the cases, that lead to the arrest of the individual in question, they may be

collectively unaware of the thousands of individuals annually are subjected to demonstrations that the police view them with suspicion, only to be released with no action. These police practices are less effective than many would like to assume, and in fact the statistics I have reviewed here show that they are very inefficient. They create a great burden on communities of color and they generate very little public safety value. For every individual who is arrested following a contraband hit in a traffic stop, there are approximately nine individuals who were subjected to a fruitless, but perhaps humiliating and costly, search, and perhaps 50 to 100 who were subjected to a traffic stop.

What was the intent of the legislature in mandating that police agencies collect traffic stop statistics, and did the legislature mandate a particular benchmark?

We reviewed the passage of the law mandating collection of traffic stops statistics in *Suspect Citizens* (chapter 2). Very briefly, the legislature was concerned about the possibility that allegations of "driving while black" were indeed true. The motivation was to gather the data that would allow this allegation to be refuted, or indeed validated. If validated, then the expectation was that police agencies would take actions to correct whatever problems might be revealed by the data. The legislature did not mandate the collection of any "benchmarking" statistics, and it did not mandate that the precise location of the stop, or even the street address, be recorded. It seems reasonable to assume that the drafters of the legislation, the first of its kind to apply statewide in the country, expected that the patterns revealed by the data collection would suffice to assess whether there were patterns of unequal enforcement of the traffic laws.

Conclusion

The traffic stop data in North Carolina speak very loudly. Black drivers are about twice as likely as whites to be pulled over, and once pulled over about twice as likely to be subjected to search.

However, these searches and this use of the traffic code to do informal investigations, yields many more innocent people than guilty ones. Disparate enforcement of the law appears wide-spread, and there seems to be little public safety benefit of these wasteful and potentially discriminatory practices. To further public discourse and the development of data-driven solutions to patterns of disparate enforcement, the use of readily available census data could be used as a general benchmark in considering the problem of racial profiling in North Carolina.