

Suspect Citizens

What 20 Million Traffic Stops Tells Us about Policing and Race

Frank R. Baumgartner, Derek A. Epp, and Kelsey Shoub
Cambridge University Press, 2018

Appendix E

Shortcomings in the Official Data and Recommendations

Of course, our study can only be as good as the data on which it is based, so we take seriously questions of data quality and highlight here some flaws that we know permeate the official records that we use. Here we review various elements in the state data collection process and provide our feedback with the hope of improving practice.

Officer ID Numbers

By law, each agency is to assign an anonymous ID number to each officer, and that ID number is to be associated with each traffic stop. The correlation between the officer ID and the officer's identity is not a public record. For the most part, this system works quite well. We quickly noted a number of anomalies in the officer ID assignment, however, and review those briefly here.

First, different agencies assign the same numbers. We have resolved that by creating our own ID number, combining a code for the agency with that of the officer within the agency. The state should consider assigning a unique officer ID that would stay with the officer when the officer moves from agency to agency. Now, if an officer were to transfer from one agency to another, those numbers would be un-related.

Second, large numbers of typographical errors plague the dataset. We found a few Officer ID fields left blank, including 2,487 for the Highway Patrol and small numbers of a few

other agencies. Third, space characters, apostrophes, and miscellaneous other additions to the typical numeric ID numbers affect the data. If an officer's ID is, say, 650, but sometimes it appears as _650 or as 650' or as 650~ or as /650 or as " 650" then when we link all the observations by officer ID these will not be seen as the same officer. It is a simple software matter to either pre-code the officer ID into the software, or to cause the software to reject inputs such as space characters which do not conform to the expected format. The Brunswick County Sheriff's Department has an officer ID of "BIG DAWG" with 84 stops, but unfortunately this officer apparently spelled it "BIGDAWG" 31 times as well. That same department had ID's of "BLULITEMU" and "BLULITEMUP". Henderson PD has an officer listed as "COCAINE" and another as "COCAIN" and one as "Cocaine". Several have officers listed as "Fatfingers" and of course that is misspelled in numerous different ways. What all this means is that very large numbers of officer ID's are used only once or twice, since they are typos, not real numbers. In fact, the median number of observations in our entire dataset for each unique officer ID is just 3. By contrast, 112 IDs appear at least 10,000 times, 4,356 appear 1,000 times or more, and 20,895 appear at least 100 times. On the other hand, 58,997 Officer ID numbers appear fewer than five times, and 40,133 appear just once. While we have no way of knowing for sure, it seems likely that most of those represent typos. These problems could be fixed with better software and training.

Contraband Types

Another problem with the SBI-122 data collection is that when officers find contraband they are asked to record it as either drugs, alcohol, money, weapons, or other. These broad categories make it impossible for analysts to make crucial distinctions between "hard drugs" (such as cocaine or heroin) and marijuana. Likewise, the form asks officers to record only the number of

weapons, but we do know if the officers recovered pocket knives or assault rifles. For that matter, it is somewhat unclear under what context weapons are considered contraband as owning a firearm is legal in North Carolina. One possibility is that officers list firearms as contraband when a motorist is unable to produce a gun license. In any event, analysis in Chapter 5 showed that when officers find contraband they arrest the motorist less than half the time, so whatever contraband they are finding it would often appear to be relatively trivial in nature.

One possibility to improve the data collection is to list the most common types of drugs on the form and have the officer record the amount of each.

Contraband Amounts when Less than 0.5

A much more serious problem with the contraband data collection, and one that could be more easily solved with software updates, is that very low contraband amounts appear to be rounded down to zero either by the local agency computers or by the NCDOJ computers when the data is compiled. We believe this is what is happening when a stop is associated with a contraband identification number, which happens only when contraband is found, but no contraband type or amount is recorded. As Chapter 5 documented, no amount is recorded for 25 percent of all contraband hits, so this happens quite frequently (with the implication that when the police find contraband it is often in very low quantities). Allowing the database files to accept more digits after the decimal point would fix this problem and make the data collection much more accurate.

Contraband Associated with the Driver, Passenger, etc.

Yet another issue with the recording of contraband is that the NCDOJ data files link contraband hits with stops, rather than with the individual motorists who were involved with a stop. This is problematic when there is a passenger involved in the stop as it is impossible to tell if the contraband was found on the driver or the passenger. We drop passengers from our analysis

because their presence is not systematically recorded by the police, but it remains possible that we are erroneously associating contraband found on passengers with drivers. Contraband is found in cars with passengers only 13,899 times in the entire 20-million stops dataset (and passengers and drivers are almost always of the same race), so we are confident that this is not a source of bias in our analysis. Still, it is an issue with the data collection that could be resolved with only moderate adjustments to the SBI-122 form.

Multiple Search Types

Only a single search type is recorded, though officers may record more than one. From discussions with officers and chiefs, it appears that the software simply records the first search type in the numeric list of search types used.

Time and Minute

Perusing the data, it is clear that an inordinate number of stops occur at exactly midnight. The time and date of the stop is recorded in a format including year-month-day hour : minute : second . thousands-of-a-second. Five illustrative data observations are presented in Table E-1.

Table E-1. Selected Dates and Times from Individual Traffic Stops

2002-01-01 00:25:00.000
2002-01-01 04:20:00.000
2002-01-01 00:10:00.000
2002-01-01 03:56:00.000
2002-01-01 00:00:00.000
2002-01-01 00:00:00.000
2002-01-01 00:00:00.000

The seconds and fractions of seconds are always missing; these data are never reported.

But note that while all these stops occurred on the first day of our data collection, January 1, 2002, with the first one occurring 25 minutes after midnight, the second at 4:20am, and so on, the last three cases were reported as having occurred exactly at midnight. In fact, of all 20 million traffic stops from 2002 through 2016, 51 percent were recorded as having occurred at

midnight (hour 00). Of those 10,319,248 midnight stops, 9,847,518, or 95 percent were recorded as occurring at minute 00 as well. The next highest observations for minute among stops having occurred just after midnight was 16,697 stops occurring at 12:01am, one minute later. These represented 0.16 percent of all stops in that hour. Clearly, the hour-minute combination of 00:00 is a default observation equivalent to missing data. We treat it as such in our analysis.

Some agencies essentially never report the time of their stops, in particular the State Highway Patrol, or if they do these do not find their way into the database released by the SBI and on which we have relied. Table E-2 shows the percent of cases listed as exactly 00:00 for hour and minute by agency. As a point of reference, among those observations occurring at noon, 5 percent occurred exactly at noon, as compared to during another minute of the hour. No hour of the day shows a value as high as 7 percent for stops exactly on the hour, not one minute after, except for midnight, which shows a value of 95 percent.

Table E-2. Missing Hour and Minute Observations, by Agency

Agency Name	Total Stops	Exactly Midnight (00:00)	Percent
NC State Highway Patrol	8,850,616	8,838,621	99.86
Gastonia Police Department	92,514	83,294	90.03
Gaston County Police Department	97,798	70,958	72.56
Goldsboro Police Department	100,574	56,968	56.64
Guilford County Sheriff's Office	94,958	24,756	26.07
Winston-Salem Police Department	466,248	97,016	20.81
Asheville Police Department	120,364	24,495	20.35
Kannapolis Police Department	115,047	20,209	17.57
Mooresville Police Department	101,682	16,452	16.18
Durham Police Department	290,115	44,838	15.46
Greenville Police Department	138,784	18,299	13.19
Burlington Police Department	114,569	4,786	4.18
High Point Police Department	306,816	10,265	3.35
Fayetteville Police Department	551,276	15,887	2.88
Wake County Sheriff's Office	147,421	2,762	1.87
Concord Police Department	202,592	3,210	1.58
Forsyth County Sheriff's Office	122,413	1,509	1.23
Wilmington Police Department	198,776	2,130	1.07

Jacksonville Police Department	216,133	2,294	1.06
Greensboro Police Department	578,974	5,469	0.94
Hickory Police Department	110,346	1,001	0.91
Apex Police Department	101,779	361	0.35
Cary Police Department	251,873	309	0.12
Charlotte-Mecklenburg Police Department	1,704,765	1,784	0.10
Raleigh Police Department	899,923	864	0.10
All Others	4,259,390	498,981	11.71
Total	20,235,746	9,847,518	48.66

Note: Time value is considered missing if the stop was listed as occurring at 00:00:00.000, exactly at midnight. Note that for agencies reporting very small numbers of such cases (e.g., below 3 percent), these may be valid observations. Where the number is extremely high, however, it most likely represents a failure to enter the time, or a loss of data.

For the top 25 agencies in the state, the table shows that a few have extremely high numbers of suspicious times, in particular the largest agency, the Highway Patrol. This is a shame, since we know that the time of the traffic stop is an important predictor of the outcome. Eleven of the 25 agencies listed show inordinately high numbers of stops exactly at midnight. Among smaller agencies (not individually listed here, but available on our web site), the problems is common as well.

Missing Stops

While the law clearly requires every agency covered by the law to report each traffic stop, it is clear that many stops are missing. Stops may be omitted from the database for a number of reasons, including when officers fail to file a report following a stop; through clerical error; through computer error at the reporting agency or at the SBI; or for other reasons.

We started the book with the story of the traffic stop by undercover officer Kelly Stewart in Durham. In conjunction with the criminal case against Carlos Riley, expert witness and attorney Ian Mance provided testimony that officer Stewart never filed an SBI-122 form associated with Mr. Riley’s “unsafe driving” stop. While that may have been understandable

since this particular stop led to a shooting, Mance's review of all the SBI-122 forms filed by that particular officer showed that while he had "reported 1,188 stops, he reported only 3 stops between December 2013 and September 2014, all of them on the same day in April. Over the remainder of 2014, he reported only 14 stops." (Mance, expert report, *North Carolina v. Riley*, 4).

In response to an informal review by Mance suggesting large numbers of missing SBI 122 reports, an Asheville Police Department report to the city council noted that the city had reported 691 stops to the SBI in February 2015 and 622 in August of that year, but just 9 and 2 stops were listed in the SBI database for those months (City of Asheville Police Department 2017, 4). To paraphrase an old adage goes, to err is human, but to err on a very large scale requires a computer. Rather than just missing a stop here and there as an officer forgets to fill out the form, or as a clerical worker misses a document in a stack of paperwork, here we are looking at very large numbers, in this instance over 1,300 traffic stop records disappeared in two months.

We have compiled monthly totals of traffic stops for every agency in the state database, and the patterns there clearly suggest massive failure to report, or reports lost through computer or other types of error. From 2002 through 2016, large agencies (which meet the threshold for reporting throughout the period) should have filed 180 monthly reports. In Table E-1 we summarize the number of reports filed for the top 26 agencies by number of stops. For each agency, the table shows the total number of traffic stops, the average per month, the number of months missing from the SBI database, and the number of months in which the agency reported extremely low numbers of traffic stops. These unusually low numbers of reports per month could, of course, be accurate indicators of the level of traffic stop activity, but they may suggest days or weeks of lost or misplaced data. When one officer fails to report, he or she could be on

desk duty or could have failed to report stops that were made. When an entire agency moves in a single month from thousands of stops to just a dozen or so, it suggests the data were misplaced, lost, or never sent to the SBI. It is also possible that the data were sent to the SBI and were lost, misplaced, or deleted; we have no way of knowing where any data may have been lost, but Table E-3 certainly suggests that tens of thousands of traffic stops have been omitted from the official record.

Table E-3. Missing Data Problems

Agency	Total	Monthly Average	Number of Monthly Observations			
			Missing	<200	<100	<10
NC State Highway Patrol	8,745,844	48,588	0	0	0	0
Charlotte-Mecklenburg Police Department	1,704,765	9,471	0	2	0	0
Raleigh Police Department	899,923	5,590	19	0	0	0
Greensboro Police Department	578,974	3,234	1	1	1	1
Fayetteville Police Department	551,276	3,063	0	0	0	0
Winston-Salem Police Department	466,248	2,711	8	2	2	2
High Point Police Department	306,816	1,733	3	6	5	3
Durham Police Department	290,115	1,707	10	0	0	0
Cary Police Department	251,873	1,399	0	3	3	3
Jacksonville Police Department	216,133	1,201	0	1	1	1
Concord Police Department	202,592	1,282	22	2	2	1
Wilmington Police Department	198,776	1,235	19	4	4	1
Wake County Sheriff's Office	147,421	824	1	5	4	2
Greenville Police Department	138,784	780	2	15	11	1
Forsyth County Sheriff's Office	122,413	684	1	0	0	0
Asheville Police Department	120,364	688	5	3	3	2
Kannapolis Police Department	115,047	639	0	0	0	0
Burlington Police Department	114,569	636	0	4	2	1
Hickory Police Department	110,346	685	19	9	8	5
SHP-Motor Carrier Enforcement Section	104,772	2,015	128	3	3	3
Apex Police Department	101,779	565	0	0	0	0
Mooresville Police Department	101,682	624	17	5	3	2
Goldsboro Police Department	100,574	562	1	7	0	0
Gaston County Police Department	97,798	565	7	2	0	0
Guilford County Sheriff's Office	94,958	549	7	11	2	1

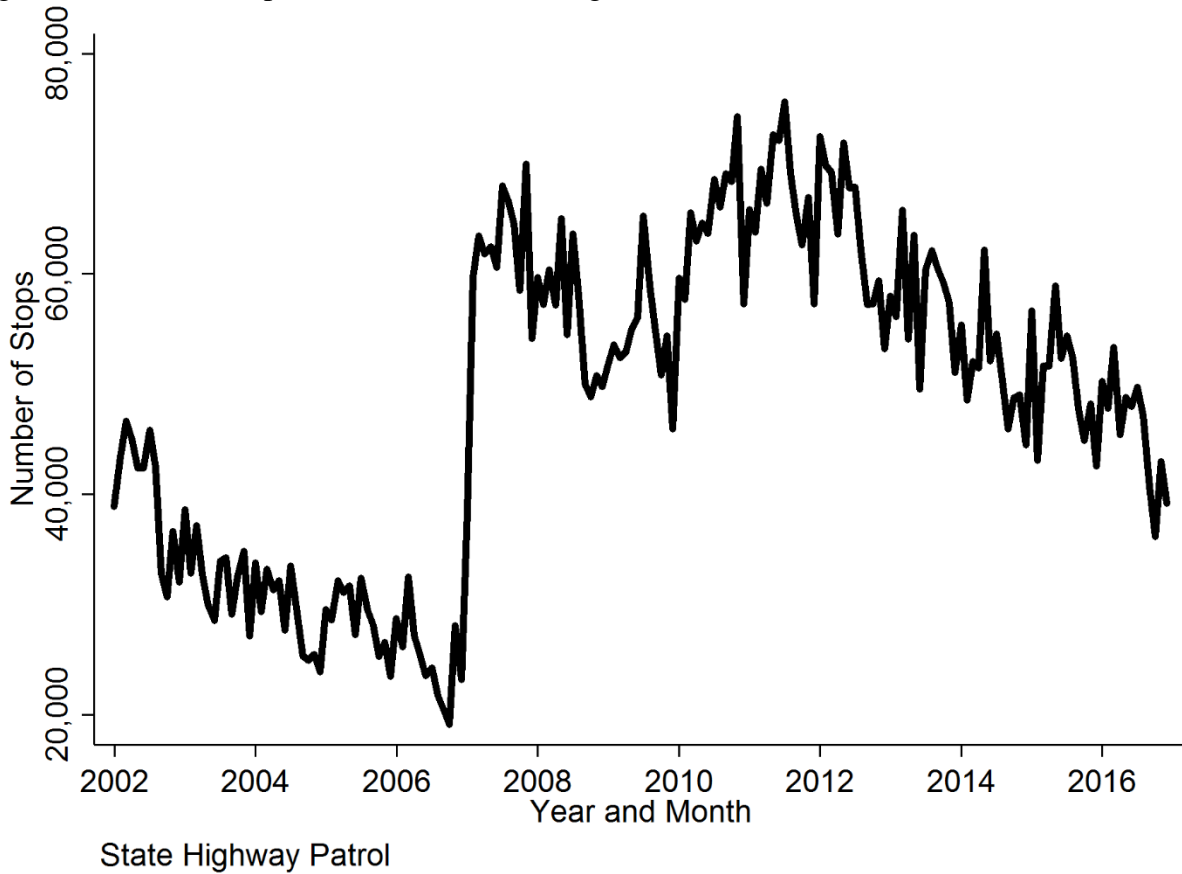
Gastonia Police Department	92,514	535	7	0	0	0
----------------------------	--------	-----	---	---	---	---

Note: The table shows the total number of traffic stops from January 1, 2002 through December 31, 2016, by month, for the top 26 reporting agencies in the state. The SHP-Motor Carrier Enforcement Division reports were included in the SHP totals during some months but reported separately from January 2003 through January 2007, with very small numbers reported in three later months. Observations listed here as missing were completely absent for the entire month. We also note cases with extremely low values. For example, Charlotte reported these values for September through December 2008: 10,119; 112; 10,103; 108. The two extremely low values suggest non-reporting.

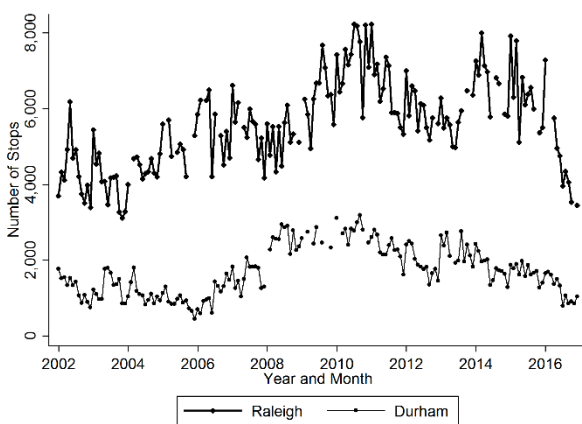
Table E-3 shows no missing data issues for the largest agency, the Highway Patrol.

Similarly, the Charlotte PD has no months with missing data. However, in October and December 2008, the agency reported just 112 and 108 traffic stops. In both surrounding months there were over 10,000 stops. Raleigh shows no traffic stops at all in 19 months; Greensboro has one month missing entirely and one with fewer than 10 traffic stops, compared to an average per month of over 3,000. Figure E-1 provides some graphical illustrations of the number of stops over time for selected large agencies.

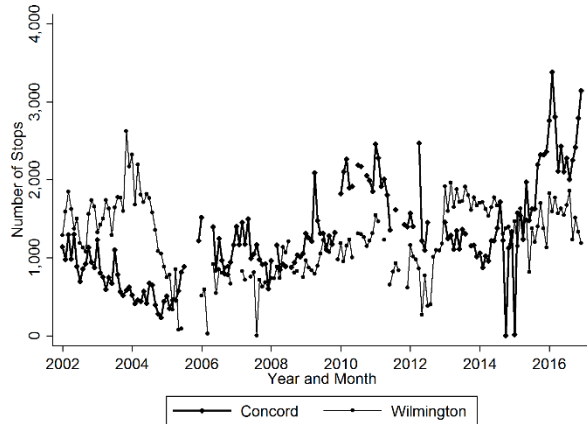
Figure E-1. Traffic Stops over Time, Selected Agencies



a. State Highway Patrol



b. Raleigh and Durham



c. Concord and Wilmington

Figure E-1a shows monthly values for the State Highway Patrol. While Table E-1 showed there were no missing values, the agency obviously made a large transformation either in its volume of traffic stops or in its reporting to the SBI in early 2007. It went from reporting

38,396 stops in January of that year to 59,754 in February and reported numbers in the 60,000 range and higher throughout most of the remainder of the year; it had not reported more than 40,000 since August of 2002. Raleigh and Durham show a different issue: entirely missing months, where no data was reported; these are visible as breaks in the lines in Figure E-1b. Raleigh has 19 missing months, and Durham has 10, as shown in the table. Concord and Wilmington have both large numbers of missing months (22 and 19, respectively), but also a number of extremely low values, as shown by sharp and incongruous dips in the lines to values close to zero. Our on-line appendix includes a spreadsheet with monthly counts of traffic stops as well as the numbers presented in Table E-3 for all 313 police agencies listed in the SBI database. Some smaller agencies are missing for particular months or entire years because they did not reach the statutory threshold for reporting to the SBI. Many others, however, failed to report at all, or appear to have failed to report all of the stops their officers did make.

Checkpoints and Passengers

As described in the text, we typically exclude passengers and checkpoint stops. This is because such stops are recorded only when there is an adverse outcome: passengers in a car in which the driver gets a speeding ticket or drivers going through a checkpoint stop with no action are not recorded. Thus, it is impossible to calculate any rates for such things as search or arrest, since we do not know the baseline of how many passengers or checkpoint stops were present. More generally, the law requires these stops to be recorded only when there are adverse outcomes. However, the law is quite ambiguous. Passenger information should be recorded, according to the law as reproduced in Appendix B, when there is a search, physical resistance, or force engaged. For checkpoints, information is to be recorded when there is an “adverse outcome” including warning, search, seizure, or arrest, or if there was physical resistance, force, or injury

(note the law does not mention citation, but does mention warnings, a lower sanction than a citation). Perhaps because of the ambiguity in the law, it does not appear to be followed in a reliable manner. For example, we found 355,540 passengers in the data set from 2002 through 2016. Of these, we found 34,460 cases of a search, 44,150 additional arrests, 290 additional passenger-only searches, 3,231 additional cases where force was encountered, and 349 additional cases where force was engaged. These numbers sum to 82,480, or 23.2 percent of all passengers. We are not sure why 273,060 passengers are in the dataset, since they do not appear to correspond to the statutory requirement.

Similarly when we look at checkpoint stops, we find 276,803 observations. Of these, 219,651 received a citation, or 79.35 percent of all checkpoint stop outcomes. Of these, just 3,925 individuals were searched, leaving as a mystery why 219,651 checkpoint stops leading to a ticket but no search were recorded. An additional 19,067 individuals were subject to no action and no search but were recorded following a checkpoint stop nonetheless.

Stop Location and Stop City

The law requires the location of the stop to be recorded, and notes that for the SHP that means the highway patrol district, and for all other agencies, it means the city or county of the stop. The definition of SHP highway patrol districts, however, is not easily available to the general public, nor is there readily available census or other indicators of the local population. Typically, Stop Location typically consists of a numeric code for the county, except for the SHP which reports a numerical patrol area. Stop City, on the other hand, is an alphabetic string variable corresponding in general to the jurisdiction of the agency, and is blank for all SHP traffic stops. Since most agencies correspond to a single geographic location except for the SHP, these data provide little

value, and the SBI should provide a list of SHP patrol districts and coordinates corresponding to the data reported.

Because Stop City is a string variable with no controls for data entry, typographical errors and inconsistencies abound. The city of Winston-Salem appears with more than 30 different spellings. Kings Mountain appears with bewildering combinations of close and distant spellings and capitalization patterns. Stop location and city data seem to be of little value. SHP patrol districts are of value but should be documented so that any user of the SBI database also has the list of SHP districts.

Typos and Miscellaneous Errors

Naturally in a database of this size there are many small errors, as for example we saw in our section on the Officer ID numbers above. A number of drivers are listed as younger than 15 years old; 5,636 in fact. Drivers under the age of 11 number 3,607, and indeed the database includes 845 drivers aged zero, and 167 apparent one-year old drivers. At the other end of the age distribution, 617 drivers are listed as over 100 years old, of whom 544 are exactly 110 years of age. The oldest drivers (18 observations) were listed as 119 years of age. We should note that for the most part, the driver age variable is relatively clean, with 99.97 percent of the observations falling within the range of 15 to 100 years old.

Our data show 442,918 drivers were arrested following their traffic stop, but of these, only 288,888 were searched, leaving 162,030 arrestees not searched. While this may be an accurate reflection of an error in procedure (individuals taken into custody after an arrest should be searched “incident to arrest” if they were not previously searched, it may be an error in paperwork. We suspect large numbers are errors in paperwork because when we look at the

211,363 individuals searched “incident to arrest,” only 182,177 were in fact arrested. In fact, 503 individuals searched “incident to arrest” were let go with no action at all.

One option that would resolve a number of these errors is to include more drop-down menus in the software where the data is recorded. Then, rather than typing out an officer ID number every single time that officer records a stop, the ID could simply be selected from a list of pre-loaded ID options. Typos appear to be a larger problem for the small agencies than the large ones (Charlotte’s data is relatively typo-free) and it is possible that the large agencies already have a drop-down menu system in place. Regardless, the NCDOJ should standardize the software, so that police agencies across the state are using the same data-input system.

When Does a Traffic Stop End and a Voluntary Encounter Begin?

Just as it may sometimes be unclear what is a “routine traffic stop” and what is a criminal investigation where the suspect happens to be driving a car, so it can sometimes be unclear when a traffic stop has come to an end. Typically, a traffic stop ends when the officer indicates that it is over. In most cases that ends when the officer thanks the driver for his or her time and returns to the patrol vehicle. In some cases, however, the officer may continue to engage in a “voluntary encounter” with the driver or passengers after the driver is technically free to go on his or her way. That is, if the officer engages in a conversation such as “Do you mind if I ask you a question?” and the driver agrees, this may not be considered to be part of the traffic stop, and the resulting outcome may not be recorded. These data should be recorded if they follow from the same encounter that was initiated by the traffic stop, and a revised law should specify this.

Previous Assessments

Deborah Weisel conducted an in-depth analysis of the database we use in a report sponsored and published by the NC Association of Chiefs of Police and NC Sheriffs’ Association, using data

from 2000 through 2011. This report notes a number of shortcomings of the data collection, as follows:

Weisel (n.d., 86 ff.) discusses problems with four stop purposes in particular, noting that they are “less clearly defined and thus subject to specification or measurement error. These types include DUI, Checkpoint, Investigation, and Other Violation”. We paraphrase her findings:

1. DUI. These stops are typically clearly defined *after* the stop has been initiated, but are not as clear before the stop is initiated. Therefore, it seems possible or even likely that a stop initiated for one purpose, but which leads to a DUI discovery might be coded as a DUI where that was not in fact the purpose of the stop, but rather the outcome. The purpose, for example, may have been failing to signal a lane change or unsafe movement.
2. Checkpoints. Since these stops should be recorded only when action is taken, two problems ensue. One is that many agencies record them when they need not be recorded, apparently based on misunderstanding the requirements. And a second is rates cannot be calculated since all stops are not recorded.
3. Investigation. The law requires “routine” traffic stops to be recorded, but Weisel quotes a 2009 SBI report saying that no report is required for “a BOLO [be on the lookout] or Amber alert, a radio broadcast “attempt to locate,” a warrant or stolen notice, or similar criminal investigation”. She writes that this is either mis-understood, inconsistent across agencies, or inappropriate: “Further, it seems likely that law enforcement officers may make numerous investigation stops that are not documented but perhaps should be reported” (87).
4. In a footnote describing the stop purposes laid out in the form, she notes: “Officers are instructed not to document traffic stops related to criminal investigations such as arising

from a warrant, notice of a stolen vehicle, look-outs (BOLO) or similar criminal investigations” (Weisel, 35, fn. 17). This raises the question of what the “investigation” stop purpose indicates.

5. Other vehicle violations. This is a large category but unclear what is included so should be clarified.

Weisel (n.d, 96) ends with these recommendations:

1. Requirements for initial stop purpose should be reviewed – DUI should be dropped, investigations clarified and other examined. Checkpoints should either include all drivers or be excluded from stops.
2. The severity of the offense for which the driver was initially stopped should be documented; for example, the officer should document both the recorded speed and the posted speed.
3. Document location of stop. Traffic stop reports should include a place variable – roadway, intersection, or other marker; when occurring within a municipality, this data should include the 100-block or nearest intersection.
4. Revisit documentation of race/ethnicity classification. The vast majority of data collection systems in the U.S. require officers to determine either the race or ethnicity of the driver – but not both. We recommend combining race and ethnicity into a single racial/ethnicity indicator for drivers – as recommended by Northeastern University (Ramirez, McDevitt, & Farrell, 2000; Fridell, 2005).
5. Document home jurisdiction (city and county) of all drivers stopped.
6. Eliminate documentation of sex, age, race and ethnicity of searched passengers.

7. Specify the arrest charge (highest charge if multiple) arising from stop and factors that lead to the arrest such as outstanding warrant.
8. Document requests for consent searches and record if the request was declined.
9. Eliminate documentation of basis for searches, particularly consent searches.
10. Citations, arrests and searches arising at checkpoints should not be documented.
11. Audit law enforcement agencies reporting data to ensure it is submitted and audit data to ensure accuracy.
12. Audit Traffic Stop Reports for accuracy and analyze data to shed light on variations between and within jurisdictions.

We agree with many of these recommendations. Given modern technology easily available on any cell phone or computer, it seems that much more precise geographic location data would be entirely feasible to collect, not only the block or street address of the stop, but the geographic coordinates, which would allow much more sophisticated spatial analysis.

A second assessment of the North Carolina experience with data collection was conducted by Raul Pinto of the NC ACLU (see American Civil Liberties Union of North Carolina 2014). His recommendations, in order:

1. First, “adhere to the old real estate adage: location, location, location” (p. 16). The more precise the location information, the better. He notes that police often use location as an explanation of the racially disparate nature of their traffic stops or outcomes, but the datasets do not allow independent assessment of this.
2. Second, facilitate the use of officer-identifiers by police departments to assess the “bad apple” problem (p. 16). Solve the Officer ID problem we discussed above, and promote the use of these data within agencies.

3. Third, Pinto urges officer training in racial profiling but also in the merits and values of the data collection required by law (p. 17).

Missing Elements in the Current Form

Other states include a number of useful items on various forms and if the General Assembly were to consider revisions it should take careful look at what is done in other states. For example, Texas data includes the name and address of the driver who was stopped. Such data would allow an assessment of repeated interactions. Other useful items include the make and model of the driver's car, the posted speed limit and the observed speed, if the car was a rental or had out-of-state license plates, and the race, gender, and years of service of the officer who made the stop.

Particularly critical in any assessment of policing decisions is consideration of the location of the activity. Therefore, the address or the GIS coordinates of where the traffic stop occurred should be included in the form.

For traffic stops leading to citation or arrest, there is currently no linkage between two state databases: The NC Administrative Office of the Courts (AOC) maintains records, including the name and address of the individual charged with any infraction or crime (including traffic infractions, but ranging up to homicide), and the result of that arrest or charge. For cases leading to an entry in the NC AOC database, a linkage should be possible, allowing researchers and police administrators better to understand the eventual outcomes of the traffic stops documented here.

Finally, we should add that the general idea should be to include any information that a police department might use to explain any observed disparities. That is, we know now that racial disparities are widespread. We do not know their causes. If officials have ideas about the

causes of the disparities we have documented in this book, the way to test those ideas is to gather the data. We have controlled for every variable collected in the current version of the form, and unexplained racial disparities remain strong. If these can be explained by other factors, then the way to find out is to add those variables to the form so that the data can be analyzed.

Promote Comparison with Other States and Jurisdictions with Open Data

It is also possible that the NCDOJ could facilitate collaborative efforts with other states, pooling traffic stops data in a central location. This would allow analysts and policymakers to put the NC numbers into a wider national context. Other options would be to work directly with academic organizations such as Stanford's project (<https://openpolicing.stanford.edu>) or opendatapolicing.org. We believe that publicizing the data is of crucial importance. The NCDOJ has been collecting traffic stops data for almost 20 years, but for much of that period the significance of the database was underappreciated. Working with outside groups is one way to ensure that the data gets the attention it deserves. It also facilitates another one of our key points, which is that analysis of racial effects in policing should be conducted by disinterested outsiders who do not have a personal stake in the findings.

A Final Assessment

Looking at the intent of the NC General Assembly in 1999 when the initial data collection law was passed, it is clear that the lawmakers believed that the data would either disprove any assertion that drivers from different racial groups experienced differences in treatment by the police, or that it would lead to immediate reform if such differences were found. It is clear, however, that dramatic disparities exist, not only in North Carolina, but throughout the country. When they are found, one side may assert that they show meaningful differences and that reforms are needed. Another group may rightly argue, however, that drivers of different racial

groups may behave differently, may have different levels of poverty, education, or employment that cause race to be confused with poverty or some other factor that may be the root cause of the issue. Our firm belief is that the differences we have uncovered in this book are too consistent, too strong, and too meaningful to be explained away by such arguments. But there is merit in those arguments. Some differences, such as poverty, are correlated with race and may well be the root cause of some of the disparities we have documented. The challenge is simply to collect the relevant data. So we would end this discussion about what data should be collected with the simple idea that any information which would seem to be relevant to why drivers experience diverse outcomes following a traffic stop should be collected. The legislature did the state a great service in mandating that, if we are concerned about something we all agree is unacceptable, then we should find out if it is happening and take steps to remedy it. Going forward, it seems clear that more information is needed. The disparities are real, strong, ubiquitous, and consistently present over time. Their explanations, however, can only be as good as the additional data we collect to understand them.