Fair & Impartial Policing in Vermont
Traffic Stop & Race Data Collection and Data Quality
and a Study of Three Jurisdictions

Submitted to:
Department of Public Safety

Submitted by:
Robin Joy, J.D., Ph.D.
Crime Research Group
P.O. Box 1433
Montpelier, VT 05601
802-230-4768
www.crgvt.org

September 2019

This Project was funded by
BJS 2016 State Statistical Analysis Grant #BJ-CX-K016
Background

Act 193 amended 20 VSA § 2366, which went into effect June 17, 2014. Section (e)(1) states as follows (in part):

On or before September 1, 2014, every State, local, county, and municipal law enforcement agency shall collect roadside stop data consisting of the following:
   (A) the age, gender, and race of the driver;
   (B) the reason for the stop;
   (C) the type of search conducted, if any;
   (D) the evidence located, if any; and
   (E) the outcome of the stop, including whether:
      (i) a written warning was issued;
      (ii) a citation for a civil violation was issued;
      (iii) a citation or arrest for a misdemeanor or a felony occurred; or
      (iv) no subsequent action was taken.

20 VSA § 2366 (e)(1).

The statute further states that LEAs shall work to collect uniform data, adopt uniform storage methods, and ensure that the data can be analyzed. This roadside stop data, reports, and any analysis are to be made public.¹

In the 2016 legislative session, the House Committee on Judiciary introduced Act 147 as a way to continue to emphasize and require the collection and analysis of traffic stop and race data. Pursuant to 20 VSA § 2366, the Vermont Criminal Justice Training Council (VCJTC) was authorized to select a vendor to collect the traffic stop and race data. VCJTC selected Crime Research Group (CRG), which collected, reviewed, and posted the datasets to its website (www.crgvt.org/tsrd). As discovered by CRG during this process, there were many issues that negatively impacted the ability of state and local law enforcement agencies to extract the data into a format that could be published and analyzed. CRG also discovered problems with data quality and challenges with the collection of uniform data. CRG, providing services as the Statistical Analysis Center (SAC) for Vermont, through a contract with the Department of Public Safety (DPS), addressed these challenges and analyzed the data from the LEAs willing to participate in the initial analysis.

Goals and Objectives

CRG proposed the development of a consistent sustainable approach to traffic stop and race data collection and management to support research and policy through five objectives over two years. CRG worked with all law enforcement agencies in Vermont to improve data collection and conducted several different analyses of traffic stop and race data for the purpose

¹ 20 VSA § 2366(e)(1).
of developing an analysis that could be used for all Vermont law enforcement agencies in the future. The objectives for this project were:

**Objective 1:** Develop Strategies to Enhance Law Enforcement Data Collection for Traffic Stop and Race Data

**Objective 2:** Work with Law Enforcement Agencies to Facilitate the Extraction of Traffic Stop and Race Data

**Objective 3:** Assess Law Enforcement Agencies’ Data Collection and Quality Needs to Assist in Crime Analysis and Evaluation

**Objective 4:** Analyze Traffic Stop and Race Data

**Objective 5:** Disseminate Traffic Stop and Race Data on a public website

**Objective 1: Develop Strategies to Enhance Law Enforcement Data Collection for Traffic Stop and Race Data**

CRG worked with law enforcement agencies (LEAs), including municipal police departments, county sheriff’s offices, the Vermont State Police, and local constables, to understand Act 147 and identify the data required by the legislation. Several requests were sent to all the agencies as reminders to extract and send their data to the vendor. Agencies struggled with this as they did not know how to extract the data, or did not have the staff time to learn the data extraction process. Glitches were worked out regarding extraction and sending the data. Data quality issues were identified as the data was sent in, cleaned, and published. The data sets needed to be standardized by timeframe, fields, and missing data needed to be added.

The first set of data collected and published consisted of mismatched timeframes, inconsistent fields, missing data, and a variety of formats including Excel, CSV, text, pdf, and handwritten submissions. The challenge was to create standardized data fields and to collect the data for the same timeframe using a consistent format. CRG worked with LEAs to improve the traffic stop and race data, as well as collect the 2016 and 2017 traffic stop and race data for publication. CRG posted the 2016, 2017, and 2018 traffic stop and race data on the website.

**Objective 2: Facilitate the Extraction of the Traffic Stop and Race Data**

Data Extraction: Vermont LEAs use two records management systems (RMS): Spillman and Valcour. DPS, the manager of Spillman, wrote instructions for extracting traffic stop and race data from the Spillman CAD/RMS and the instructions were distributed to the LEAs using Spillman. Valcour agencies, on the other hand, were able to extract data using a customizable report. CRG worked with the 75 Vermont LEAs to collect the data required by Act 147 for each traffic stop for 2015/2016. All but three extracted traffic stop and race data and sent it for
publication. The three agencies not reporting data were very small agencies not having the capacity to extract the data. At the conclusion of the data extraction process, CRG posted the spreadsheets to its website and distributed the link to DPS and VCJTC. CRG worked with the Vermont State Police to identify specific issues with the data. Together, they developed the *Factors Impacting the Validity of the Data*, and posted these on the website to warn anyone using the 2015/2016 data of the issues with the data. It is attached here as Appendix A.

The challenge for 2017 traffic stop and race data was to create standardized data fields, a uniform timeframe, preferrably calendar year, and a consistent format.

It was the intention of DPS and CRG at the start of this project that the data extraction be automated in Year 1 using the Vermont Justice Information Sharing System (VJISS). VJISS is a statewide portal used to share justice information between and among authorized users and systems in Vermont. It was designed to promote data-driven justice decisions, as well as enhance the effectiveness of justice policies and practices throughout the State of Vermont. VJISS provided a suite of statewide data and information sharing tools to justice professionals and other stakeholders in support of more effective strategic, tactical, and operational decisions. Part of this project was to identify the data needed for robust and long-term analysis of traffic stop and race data, pull it from the RMSs into the VJISS Analytical Data Store (ADS), then push it out to a publicly accessible website in a format that allowed for data analysis and data visualization.

CRG and DPS explored the opportunities available through VJISS to explore the analytical data store’s capabilities for creating the data reports and pushing the traffic stop and race data to the website. Working with the SEARCH Open Justice Broker Consortium (OJBC), and the Valcour and VIBRS Governance Boards, which oversees the process of sharing data with outside agencies, it became apparent for a variety of reasons that uploading the data into the VJISS/ADS was not going to be a short-term solution. CRG began working with DPS to extract traffic stop and race data for the Spillman agencies into one file; and with Crosswind, the vendor for the agencies using Valcour, to extract traffic stop and race data into a second file.

The Spillman 2017 traffic stop and race data were extracted into one file with the exception of the Vermont State Police data and one police department. This meant that approximately 33 police departments had their traffic stop and race data in one file. The Valcour data was still submitted in separate files by the LEAs using Valcour. For 2017 data, three Spillman files and 35 agencies using Valcour were posted on the CRG website. Discussions continued between CRG and Crosswind for the 2018 and 2019 data.

**Objective 3: Assess Law Enforcement Agencies’ Data Collection and Quality**

CRG continued to identify and contact the LEAs having gaps in their data and needing help with data extraction. As the law enforcement data was collected, a gap analysis documented the missing data. CRG worked with Spillman and Valcour system users to gain an understanding of
the strengths and weaknesses of each CAD/RMS and the fields available to improve traffic stop and race data.

CRG completed several analyses using best practices in the field, and has determined that moving forward the following fields will be helpful for researchers and citizens choosing to analyze the data. These additional data fields will allow for a more robust and accurate analysis of traffic stop and race data.

The data fields being requested are:

- **Date, time, and location of the stop** as a way to determine and eliminate duplicates and needed to conduct the veil of darkness analysis.

- **Make, model, and year of vehicle** can explain disparities in vehicle equipment stops.

- **State of the vehicle plate** will allow researchers to remove out of state vehicles from any analysis involving Vermont census data.

- **State and town of residence of the driver** will allow researchers to remove out of jurisdiction drivers from analysis using Vermont census data.

- **More detailed reason for the stop** - the current broad categories of vehicle stops may be obfuscating differences. Knowing if a class of people is routinely pulled over for certain offenses versus others is helpful in analysis. For example, is there less tolerance for out of state drivers pulled over for speeding just a few miles per hour over the speed limit than for instate drivers not getting pulled over unless speeding at least 15 miles per hour over the limit?

- **Reason for the arrest** - analysis indicates that disparities almost disappear when the reason for arrest (e.g. driving with license suspended) is included.

- **Type of contraband found** will allow researchers and citizens to understand the relationship between types of contraband and the traffic stop.

The recommended changes to the legislation for additional data elements have been presented to the Vermont Association of Chiefs of Police, the Vermont Sheriff’s Association, the Valcour Governance Board, and the VIBRS Governance Board (for Spillman agencies). Once the additional data are approved, they will be presented to the Vermont Legislature for inclusion in the statute on traffic stop and race.
Objective 4: Analyzing the Traffic Stop and Race Data

Methodologies for Measuring Disparities – History

Benchmarking History

Interest in using traffic stop data to measure racial disparities began in the mid-1990s. The earliest studies used census data to estimate driving populations in jurisdictions. Those estimates were then used as a benchmark against which stops were measured. As the United States Court of Appeals for the 7th Circuit noted in *Chavez vs. Illinois State Police*: “Census data can tell us very little about the numbers of Hispanics and African-Americans driving on Illinois interstate highways, which is crucial to determining the population of motorists encountered by the [ISP] officers.”

The “Gold Standard” for benchmarking is field observation study, where researchers observe the race of drivers in a jurisdiction over seasons and varying times of day. From these observations, an estimated driving population is constructed for the benchmark. These studies are often cost-prohibitive for small departments. They also need to be repeated over periods of time as demographics change.

In the early 2000s, Northeastern University’s Institute on Race and Justice (IRJ) created an estimated driving population using a very sophisticated analysis of census data. IRJ first identified communities within a 30-minute driving time radius and assumed that those communities would contribute to the driving population of the community. Then, it accounted for vehicle ownership, commute times, retail, and entertainment destinations. Using these factors, IRJ created an estimated driving population.

This methodology is a significant improvement over the use of census data. However, it, too, is often cost-prohibitive for agencies to undertake. Since Northeastern University’s advancement, two other benchmarking techniques have been developed. One uses the not-at-fault driver in two-car accidents to estimate the driving population. Because Vermont’s crash data is missing race in over one-third of all crashes, it is currently not a reliable estimate. However, if reporting rates improve, this approach holds promise.

The second innovation in determining the estimated driving population came from Connecticut’s Institute for Regional and Municipal Policy Planning at Connecticut State University. The commuting hours analysis estimates the worker population in a jurisdiction and then looks only at stops made during commuting hours. Connecticut performs an additional analysis on stops in the jurisdiction of residents only. These are the methodologies CRG used to attempt to create a benchmark driving population for jurisdictions in Vermont.

---

Prior Research on Traffic Stop and Race – Vermont

Stephanie Seguino and Nancy Brooks released “Driving While Black and Brown in Vermont,” in January, 2017. This was an analysis of 26 police departments’ traffic stop and race data. The study, however, was seriously flawed and its conclusions do not stand up to academic rigor. First, the data used included multiple ticket/warnings for events. This means that if a driver received a warning and a ticket at one stop, the study counted it as two separate stops. And if a search was conducted, two separate searches. Second, the authors do not provide clarity on resolution of data inconsistencies relating to searches and hit rates. One agency in that study tried to replicate the authors' findings with its own data and found that the authors had included inconsistencies. For example, if an officer entered “No Search” but “Contraband Found” it was counted as a “hit” and if the officer entered “Search,” but nothing or “no search conducted” in the Search Outcome field, that was counted as a fruitless search. Third, the data provided by police departments generally included data from stops/tickets that occurred in a town other than that of the police department’s jurisdiction. For example, an officer may assist in another town. These tickets issued elsewhere should be excluded from analysis. These data issues affect the analysis conducted and make the study’s conclusions questionable.

The Seguino and Brooks study also uses inappropriate methodology for benchmarking. In some jurisdictions, the researchers used crash data for the county and applied it to the city or town. This is an incorrect application of the methodology. The race of the not-at-fault drivers was missing for 35% of the vehicle crashes in one county studied. In a prior study of the Vermont State Police, the same authors found a missing rate of 21% unacceptable for analysis of that agency. In other jurisdictions, the study used the county population to create the estimated driving population (EDP) for individual towns. This methodology is flawed. First, in the population estimates the authors included children and infants. Second, the data included out-of-state residents. In the data provided for the analysis below, one jurisdiction had 3,235 stops of which, in 1,000 stops, drivers had out-of-state residences. An additional 246 drivers had residences in Vermont but outside of that county. This means that more than one-third of the drivers stopped had no relationship to the demographics of the county in which they were stopped.

Methodologies for Measuring Disparities - Vermont

There are three generally accepted ways to measure racial disparities: 1) Benchmarking stops to an Estimated Driving Population (EDP); 2) Veil of Darkness Analysis; and, 3) Disparities in post-stop outcomes to determine if minority drivers are treated differently than white drivers.

---

For this project, CRG used all three methods to test for their viability for use in Vermont and to provide a more holistic approach to understand how the law enforcement agencies interact with the motorists that are stopped.

CRG applied Connecticut’s methodology for the analysis of traffic stops and race data in Vermont. The purpose was to test the methods and make recommendations on those suited for statewide analysis and to determine a method for analyzing traffic stop and race data that would work for all LEAs in Vermont going forward. The three analyses completed used the commuting population analysis (benchmarking), resident driver analysis (benchmarking), and the Veil of Darkness in addition to analyzing disparities in post stop outcomes. For the purpose of this project we tested the methodologies in three jurisdictions that consented to participate in our research and provided data for calendar year 2016.

Jurisdiction 1 is a local police department in a city that contains a hospital and a college.
Jurisdiction 2 is a statewide agency.
Jurisdiction 3 is a local police department in a city that contains a hospital and a small college.

**Benchmarking Stops to an Estimated Driving Population (EDP)**

Connecticut pioneered the use of a database known as the LEHD Origin-Destination Employer Statistics (LODES). LEHD is an acronym for “Local Employer Household Dynamics.” This is a database of unemployment insurance data supplied by the states. Every employee who is covered by unemployment insurance is captured, along with work and home addresses. The database also contains the number of jobs by race and other demographics in a jurisdiction. The data come from a variety of sources including census data but it is also supplemented with social security records and federal tax returns.\(^8\)

CRG downloaded the Estimated Driving Population (EDP) for each Vermont town - using census data and LODES (Local Employer Household Dynamics (LEHD)\(^9\) Origin - Destination Employer Statistics). The commuting population analysis was completed for three jurisdictions that volunteered and had enough commuters to conduct this analysis. For each, the LODES data was used to identify all those employed in the town, but residing in some other location regardless of how far away they live from the target community. The numbers of all commuters from the contributing towns were totaled and represent the nonresident portion of the given town’s EDP. This was combined with the town’s resident driving population. The combined nonresident and resident numbers from the towns complete the EDP. To avoid double counting, those both living and working in the target town were counted as part of the town’s resident population and not its commuting population. The EDP is used to analyze traffic stops during commuting hours only.

---


\(^9\) LEHD is a partnership between the U.S. Census Bureau and its partner states. LODES data is available through an on-line application called *OnTheMap* operated by the Census Bureau and the American Community Survey (ACS).
The steps for conducting this analysis are:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>For each town, use LODES data to identify all those employed in the town, but residing in some other location regardless of how far away they live from the target community.</td>
</tr>
<tr>
<td>2</td>
<td>Use ACS five-year average estimated data to adjust for individuals commuting by some means other than driving, such as those using public transportation.</td>
</tr>
<tr>
<td>3</td>
<td>For all Vermont towns contributing commuters, racial and ethnic characteristics of the commuting population to be determined by using the jurisdictions’ 2010 census demographics.</td>
</tr>
<tr>
<td>4</td>
<td>For communities contributing more than 10 commuters who live outside of Vermont, racial and ethnic characteristics of the commuting population to be determined using the jurisdictions’ 2010 census demographics.</td>
</tr>
<tr>
<td>5</td>
<td>For communities contributing fewer than 10 commuters who live outside of Vermont, racial and ethnic characteristics of the commuting population to be determined using the demographic data for the county in which they live.</td>
</tr>
<tr>
<td>6</td>
<td>The numbers for all commuters from the contributing towns are totaled and represent the nonresident portion of the given town’s EDP. This will be combined with the town’s resident driving age population. The combined nonresident and resident numbers form the town’s complete EDP.</td>
</tr>
<tr>
<td>7</td>
<td>To avoid double counting, those both living and working in the target town will be counted as part of the town’s resident population and not its commuting population.</td>
</tr>
</tbody>
</table>

Examples: Jurisdictions 1 and 2

Jurisdiction 1

To construct the estimated commuting population, CRG modified Connecticut’s approach slightly. Like Connecticut, the analysis started with the number of jobs reported by the LODES data. These data are not an estimate, but are all jobs where the employee is covered by unemployment insurance. Connecticut then pulls the demographic data from the census for those 16 years old and older from the home towns that provided workers to a jurisdiction and begins to construct the population. Since Connecticut pioneered this, the number of jobs by race in a jurisdiction were added to the LODES data.

The assumption is made that residents of driving age are all equally likely to be driving during commuting hours, not just to work, but for school, errands, and daily life. To avoid double counting residents, the analysis attempts to back out local residents from the workforce. It is assumed that their demographics in the workforce are the same as the demographics of the community. It turns out this assumption is false. The work is shown here only to demonstrate why this method of benchmarking fails for these jurisdictions.
Because the LODES data uses census designated categories for race and treats Latinx as an ethnicity, this analysis is only applied to race and not ethnicity. Vermont LEAs treat Latinx origin as a race category.

Workers in Jurisdiction 1

Using the LODES data, employers in Jurisdiction 1 employed 9,621 workers in 2015, the latest year of data available. Of those workers, 3,304 resided in the jurisdiction. The remaining 6,317 workers live outside of the jurisdiction. Table 1 shows the top 10 towns outside of the jurisdiction that contribute to the workforce.

Table 1: Top Ten Towns Outside Jurisdiction 1 That Contribute to the Workforce

<table>
<thead>
<tr>
<th>Town, State</th>
<th>Number of Workers</th>
<th>Share of Work Force</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoosick Falls Village, NY</td>
<td>367</td>
<td>3.8%</td>
</tr>
<tr>
<td>Cambridge Village, NY</td>
<td>132</td>
<td>1.4%</td>
</tr>
<tr>
<td>South Shaftsbury, VT</td>
<td>123</td>
<td>1.3%</td>
</tr>
<tr>
<td>Rutland City, VT</td>
<td>116</td>
<td>1.2%</td>
</tr>
<tr>
<td>Pittsfield City, MA</td>
<td>112</td>
<td>1.2%</td>
</tr>
<tr>
<td>Arlington, VT</td>
<td>87</td>
<td>0.9%</td>
</tr>
<tr>
<td>Burlington City, VT</td>
<td>55</td>
<td>0.6%</td>
</tr>
<tr>
<td>Manchester Center, VT</td>
<td>49</td>
<td>0.5%</td>
</tr>
<tr>
<td>New York City, NY</td>
<td>42</td>
<td>0.4%</td>
</tr>
<tr>
<td>Troy City, NY</td>
<td>39</td>
<td>0.4%</td>
</tr>
</tbody>
</table>

Jurisdiction 1 businesses employ people who reside as far away as California and as close as Shaftsbury. Because of the geographic diversity of workers, the perils of using town or even county census data as a benchmark alone for all stops becomes clear.

Table 2 illustrates the number and percent of jobs by race.

Table 2: Number and Percent of Jobs by Race in Jurisdiction 1

<table>
<thead>
<tr>
<th>Race</th>
<th>Number of Jobs</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Alone</td>
<td>9,325</td>
<td>96.9%</td>
</tr>
<tr>
<td>Black or African American Alone</td>
<td>135</td>
<td>1.4%</td>
</tr>
<tr>
<td>Native American or Alaskan Native Alone</td>
<td>23</td>
<td>0.2%</td>
</tr>
<tr>
<td>Asian Alone</td>
<td>70</td>
<td>0.7%</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander Alone</td>
<td>6</td>
<td>0.1%</td>
</tr>
<tr>
<td>Two or More Races</td>
<td>62</td>
<td>0.6%</td>
</tr>
<tr>
<td>Total</td>
<td>9,621</td>
<td>99.5%</td>
</tr>
</tbody>
</table>

10 This jurisdiction town has a hospital and college. Hospitals often employ traveling nurses or other medical staff who may consider another state home. Likewise, the college attracts students from all over the country, and their residence on a paycheck would likely reflect their home and not their college address.
11 Numbers do not add to 100% due to rounding.
Drivers in Jurisdiction 1

The data used to construct the estimated driving population of Jurisdiction 1 residents comes from the American Community Survey 2011-2015 5-year Estimates. Only those residents who are eligible to receive a learner’s or driver’s permit are used. Therefore, only those 15 and older are counted.

Table 3: Estimated Resident Driving Population in Jurisdiction 1

<table>
<thead>
<tr>
<th>Race</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Alone</td>
<td>12,475</td>
<td>94.55%</td>
</tr>
<tr>
<td>Black Alone</td>
<td>151</td>
<td>1.14%</td>
</tr>
<tr>
<td>Native American or Alaskan Native</td>
<td>22</td>
<td>0.17%</td>
</tr>
<tr>
<td>Asian Alone</td>
<td>122</td>
<td>.92%</td>
</tr>
<tr>
<td>Hawaiian or Pacific Islander</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Some Other Race</td>
<td>70</td>
<td>.53%</td>
</tr>
<tr>
<td>Two or More Races</td>
<td>354</td>
<td>2.68%</td>
</tr>
<tr>
<td>Total</td>
<td>13,194</td>
<td>99.99%</td>
</tr>
</tbody>
</table>

Construction of the Commuting Hour Population in Jurisdiction 1

As stated above, we assumed that Jurisdiction 1 residents work in the workforce at the same racial proportion. There were 3,304 residents working within Jurisdiction 1. Accordingly, the assumed breakdown of workers is presented in Table 4.

Table 4: Assumed Breakdown Residents Who Work in Jurisdiction 1

<table>
<thead>
<tr>
<th>Race</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Alone</td>
<td>3,124</td>
<td>94.55%</td>
</tr>
<tr>
<td>Black Alone</td>
<td>38</td>
<td>1.14%</td>
</tr>
<tr>
<td>Native American or Alaskan Native</td>
<td>6</td>
<td>.17%</td>
</tr>
<tr>
<td>Asian Alone</td>
<td>30</td>
<td>.92%</td>
</tr>
<tr>
<td>Two or More/Some other Race</td>
<td>106</td>
<td>3.21%</td>
</tr>
<tr>
<td>Total</td>
<td>3,304</td>
<td>99.99%</td>
</tr>
</tbody>
</table>

Table 5 shows where the assumption that Jurisdiction 1 residents contribute to the workforce in equal proportions fails. The LODES data reports 62 jobs held by people who identify as Two or More Races, but calculations show that Jurisdiction 1 would supply 106 workers who identify as Two or More Races.

---

12 Numbers do not add to 100% due to rounding.
13 We combine the categories of “Two or More” and “Some Other Race” for comparison into the LODES data.
14 Numbers do not add up to 100% due to rounding.
Table 5: Race by Worker Residence in Jurisdiction 1

<table>
<thead>
<tr>
<th>Race</th>
<th>Total Number of Jobs</th>
<th>Jobs Held by Bennington Residents</th>
<th>Jobs Held by Non-Residents</th>
<th>% of Jobs by Race for Non-Residents</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Alone</td>
<td>9,325</td>
<td>3,124</td>
<td>6,201</td>
<td>98.16%</td>
</tr>
<tr>
<td>Black or African American Alone</td>
<td>135</td>
<td>38</td>
<td>97</td>
<td>1.54%</td>
</tr>
<tr>
<td>Native American or Alaskan Alone</td>
<td>23</td>
<td>6</td>
<td>17</td>
<td>.27%</td>
</tr>
<tr>
<td>Asian Alone</td>
<td>70</td>
<td>30</td>
<td>40</td>
<td>.63%</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>.09%</td>
</tr>
<tr>
<td>Two or More Races</td>
<td>62</td>
<td>106</td>
<td>-44</td>
<td>-.69%</td>
</tr>
<tr>
<td>Total</td>
<td>9,621</td>
<td>3,304</td>
<td>6,317</td>
<td>100%</td>
</tr>
</tbody>
</table>

Although this will not be a useful benchmarking tool for Jurisdiction 1, the discussion provides some insight on who is coming into the town for work, and where they come from. This should be kept in mind with any other benchmarking attempts to measure racial disparities in policing.

Jurisdiction 2

Unfortunately, this methodology did not work for Jurisdiction 2. This department covers a large area, and the LODES data did not have employee data for some towns, despite there being a school, post office, and other employers in the jurisdiction.

Using the LODES data for available towns, businesses in Jurisdiction 2 employed 96,705 workers in 2015, the latest year of data available. Of these workers, 78,258 are residents. The remaining 18,447 workers are non-residents. Table 6 shows the top 10 towns outside of Jurisdiction 2 that contribute to the workforce.

Table 6: Top Ten Towns Outside of Jurisdiction 2 That Contribute to the Workforce

<table>
<thead>
<tr>
<th>City/Town, State</th>
<th>Number of Workers</th>
<th>Share of work force</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Albans, Vermont</td>
<td>1308</td>
<td>1.4%</td>
</tr>
<tr>
<td>Rutland City, Vermont</td>
<td>572</td>
<td>0.6%</td>
</tr>
<tr>
<td>Montpelier, Vermont</td>
<td>544</td>
<td>0.6%</td>
</tr>
<tr>
<td>Barre City, Vermont</td>
<td>527</td>
<td>0.6%</td>
</tr>
<tr>
<td>Vergennes, Vermont</td>
<td>487</td>
<td>0.5%</td>
</tr>
<tr>
<td>Swanton, Vermont</td>
<td>355</td>
<td>0.4%</td>
</tr>
<tr>
<td>Waterbury, Vermont</td>
<td>297</td>
<td>0.3%</td>
</tr>
<tr>
<td>Middlebury, Vermont</td>
<td>238</td>
<td>0.3%</td>
</tr>
<tr>
<td>Plattsburgh, New York</td>
<td>211</td>
<td>0.1%</td>
</tr>
<tr>
<td>St. Johnsbury, Vermont</td>
<td>126</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

Employers in Jurisdiction 2 employ people who reside as far away as California.15

---

15 As in Jurisdiction 1, Jurisdiction 2 also includes a hospital which often employs nurses and medical staff from out of state. And several colleges which attract students from all over the country.
**Recommendation:** Because of the geographic diversity of workers, the perils of using county census data as the sole benchmark for all stops becomes clear therefore it is not recommended that the commuting population analysis be used in Vermont.

**Resident Driver Analysis**

Analyzing how members of their communities experience the police is useful information for police. If segments of the population are experiencing, or even perceiving, more negative contact with the police, some of the fundamentals of mutual trust begin to erode. This can lead to a more dangerous policing environment for everyone involved.

A few caveats about this estimate. First, it assumes that all residents 15 or older have a learner’s permit or a license, which is likely untrue, but to what extent is unknown. Second, this estimate is based on the ACS 2011-2015 5-year survey, which does have high margins of error for the non-white populations. In Jurisdiction 1, the ACS estimates the total Black Alone population as 154 with a margin of error of +/- 57 meaning that the true population could be anywhere from 97 to 211. In Jurisdiction 2, one town had an ACS estimate of five Black Alone residents, with a margin of error of +/-10, meaning that the true population is anywhere from 0 to 15 residents. In Jurisdiction 3, the Black Alone population is estimated at 206, with a margin of error of +/- 121, placing the true population between 48 and 327. Finally, the census categories for race and ethnicity do not correspond with Vermont’s traffic stop data collection. Latinx is considered an ethnicity in census data and race in traffic data. There are no multi-race categories in Vermont traffic data but there are in census data.

Despite the caveats, this type of analysis can be useful for local jurisdictions, provided the limitations are understood. The resident driver analysis for Jurisdiction 1 is presented in Table 7 for illustration:

<table>
<thead>
<tr>
<th>Race of Operator (Residents)</th>
<th>Reason for Stop</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DUI</td>
<td>Equipment Violation</td>
</tr>
<tr>
<td>Latinx</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Missing</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Non-White Not-Latinx</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>White</td>
<td>1</td>
<td>417</td>
</tr>
<tr>
<td>Grand Total</td>
<td>1</td>
<td>448</td>
</tr>
</tbody>
</table>

In Table 7, the percent of Non-White Not-Latinx resident drivers who are stopped in Jurisdiction 1 is 3.59%. Below in Table 8 is the estimated driving population for Jurisdiction 1.
Table 8: Estimated Driving Population for Jurisdiction 1

<table>
<thead>
<tr>
<th>Race</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Alone</td>
<td>12,475</td>
<td>94.55%</td>
</tr>
<tr>
<td>Black Alone</td>
<td>151</td>
<td>1.14%</td>
</tr>
<tr>
<td>Native American or Alaskan Native</td>
<td>22</td>
<td>0.17%</td>
</tr>
<tr>
<td>Asian Alone</td>
<td>122</td>
<td>0.92%</td>
</tr>
<tr>
<td>Hawaiian or Pacific Islander</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Some Other Race</td>
<td>70</td>
<td>0.53%</td>
</tr>
<tr>
<td>Two or More Races</td>
<td>354</td>
<td>2.68%</td>
</tr>
<tr>
<td>Total</td>
<td>13,194</td>
<td>99.99%</td>
</tr>
</tbody>
</table>

The estimated driving population for Non-Whites is 5.44%, but account for 3.59% of the stopped resident drivers. Whites are estimated to make up for 94.55% of the driving population and account for 93.53% of the stopped resident drivers. From this benchmark, it does not appear that Non-Whites are stopped at a disproportionate rate to their estimated driving population. However, the driving population estimate is just that, an estimate. Some of the assumptions made in creating it may not be true. This benchmark is illustrative, but not dispositive.

**Recommendation:** It is useful for law enforcement agencies in Vermont to conduct the resident driver analysis for jurisdictions that serve one town or city. To accomplish this, police departments would have to report out the town/state of the driver.

**Veil of Darkness Analysis**

The Veil of Darkness analysis was developed in 2002 by researchers in Oakland, California and it does not attempt to benchmark a driving population. The analysis is conducted on a subset of stops. The Veil of Darkness method looks at stops before and after the sun rises or sets on a given day during the inter-twilight of dawn and dusk. It assumes that the driving population at 5 p.m. in January is the same population driving at 5 p.m. in June. Therefore, if there is racial bias by a police department, whether explicit or implicit, one would expect more minorities to be stopped during daylight hours (in June), when officers can see into the vehicle than in the dark during January when officers may not be able to perceive the race of the driver. The steps taken to conduct this analysis were:

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Download the US Naval Observatory Data for Vermont and construct variables for dawn and dusk inter twilight periods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2</td>
<td>Conduct appropriate analysis for the inter twilight periods</td>
</tr>
</tbody>
</table>

The analysis the focused on the 30 days before and after a time switch. This helps eliminate some of the differences that may be observed because of seasonal driving differences. For example, an area may see more traffic during the winter months than in the summer. Limiting

---

16 Numbers do not add to 100% due to rounding.
the analysis to a sixty-day period helps strengthen the analysis. Below, in Table 9, is the Fall time switch analysis, for the afternoon hours, for Jurisdiction 1.

Table 9: Fall Switch to Standard Time for Jurisdiction 1

<table>
<thead>
<tr>
<th>Race of Operator</th>
<th>Missing</th>
<th>Asian</th>
<th>Black</th>
<th>Unknown</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Dark</td>
<td>13</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>128</td>
</tr>
</tbody>
</table>

As you can see from Table 9, more people were pulled over in the dark hours than daylight hours. If racial bias was influencing the decision to pull over, the theory argues that more minorities would be pulled over during daylight hours.

Testing for racial bias is done through regression modeling. Regression models control for variables that may affect the stop, such as gender, vehicle make and model, state of the plate, and race. Jurisdiction 1 did not have the age of the operator in its data, so a regression analysis was not possible. Jurisdictions 2 and 3 did have age in their data and below are findings from the regression analysis. The full regression models are attached in Appendix B.

Regression modeling assigns a probability that something happened more than by chance. The process of conducting a regression analysis allows a determination of which factors matter most, which factors can be ignored, and how these factors influence each other. Regression also “controls” for other factors, for example, regression can tell if females were more likely to be pulled over, controlling for age. Although every jurisdiction had enough traffic stops during the evening hours of the 60-day period in the fall and spring, the regression modeling did not produce any relevant factors. That is, none of the variables tested predicted whether one would be stopped during the day or night. The variables tested were: state of the vehicle, race, gender, and age.

There are several possible reasons why the regression models did not produce any relevant factors. First, it could be that the variables do not, in fact, predict whether one is stopped during the day or night. However, the models returned some numbers that indicate that the data subset is too small, with little or no variance, to obtain an accurate picture. For example, one jurisdiction pulled over Latinx men and no Latinx women during the time period. This means that no test could be completed to see if gender made a different in treatment by police because there were no women in the sample. Every jurisdiction tested had similar issues - too few people with disparate traits in the sample.

**Recommendation:** The Veil of Darkness analysis is useful, even if on these jurisdictions the regression modeling failed to produce relevant factors. The analysis can be performed on any jurisdiction and the raw numbers might warrant a closer look at a particular LEA. However, it should be noted that it is essentially looking at one shift of an LEA, the nighttime commuting
hours shift. In some departments that might be one officer. The caveat with this method is that it is not a measure of the actions of all officers in the department.

**Post-Stop Outcomes**

Post-Stop Outcomes are an indication of the difference in treatment by police after the stop has occurred for minorities who are stopped versus non-minorities who are stopped. Post-Stop Outcomes include issuing tickets and/or warnings, arrests, and searches. These measures do not rely on benchmarking to driving populations. The race of the driver is perceived after the stop when the decision to issue a ticket or the decision to search is made. A weakness in looking at Post-Stop Outcomes is that this analysis does not account for the full range of variables that an officer uses when exercising discretion. Several analyses are presented here for descriptive purposes, with suggestions to improve the measures.

**Arrests**

Table 10 below shows the post-stop outcomes for Jurisdiction 2. There were 20 discretionary arrests in the Jurisdiction, of which two were Non-White operators, or 10% of the arrests. Jurisdiction 2 also arrested two people for DUI in a traffic stop. In working with Jurisdiction 2, CRG learned that officers are not issuing tickets when they are arresting people for DUI or other serious motor vehicle offenses.

By obtaining data from DPS on DUI arrests for Jurisdiction 2, the number of DUI arrests during the study period was 94.\(^{17}\) There were 89 arrests for White drivers, yet only 18 White driver arrests appear in the traffic stop data. There were additionally three arrests for Black drivers and two for Hispanic drivers. Omitting stop and arrest data for DUI and GNO traffic stops creates data issues for the analysis of post-stop outcomes.

Considering the additional data on arrests in the DUI traffic stops resulted in arrests for Black drivers dropping from 5.0% to 3.5%, arrests for Latinx/Hispanic drivers dropping from 5.0% to 2.6% and for White drivers increasing from 90% to 93.86%. This illustrates the point that small numbers make a big difference when moving from raw numbers to percentages. And to emphasize that any discussion of traffic stop data on post-stop outcomes in Vermont should include a discussion of the omission of DUI stops and arrests from the traffic stop data.

Jurisdiction 2 is working with officers to fill out the ticket for DUI stops even when the driver is arrested, this agency may have more officers than other departments which gives them the ability to change their practices.

\(^{17}\) These arrests that are not in the traffic stop data are noted in red font in Table 10.
Table 10: Post-stop Outcome by Race for Jurisdiction 2

<table>
<thead>
<tr>
<th>Race of Operator</th>
<th>Reason for Stop</th>
<th>Outcome of Stop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Missing</td>
<td>Arrest</td>
</tr>
<tr>
<td>Missing</td>
<td>Missing</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>Equipment Violation</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Moving Violation</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>DUI</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Equipment Violation</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Investigatory</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Moving Violation</td>
<td>1</td>
</tr>
<tr>
<td>Latinx</td>
<td>DUI</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Equipment Violation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Investigatory</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Moving Violation</td>
<td>1</td>
</tr>
<tr>
<td>Native American</td>
<td>Moving Violation</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>Moving Violation</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Missing</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>DUI</td>
<td>2+89</td>
</tr>
<tr>
<td></td>
<td>Equipment Violation</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Investigatory</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Moving Violation</td>
<td>10</td>
</tr>
<tr>
<td>Grand Total</td>
<td>11</td>
<td>20/114</td>
</tr>
</tbody>
</table>

Search and Hit Rates

Researchers have tried to use search and hit rates to measure bias. There are two main methods of doing so, both with flaws. The first is called the KPT Hit Rate. Developed in a series of papers by Knowles, Pearson, and Todd, this test looks at the success rates of searches of White drivers and compares them to success rates of non-White drivers. The second method applies the Veil of Darkness analysis to post-stop behavior. Both methods are presented here for descriptive analysis only.

KPT Hit Rate Analysis

This model is based on the economic Game Theory. Game theory posits that we all act to maximize our desired outcomes. In the case of police officers, they would act to successfully discover contraband. In the case of criminals, they would act to minimize the risk of being detected. The KPT hit rate argues that if an officer wants to find illicit drugs, and the officer is

---

intentionally biased against Blacks, then he will search Black drivers more frequently, but find more contraband on White drivers. Eventually, the theory argues, there will be equilibrium because Black drivers will begin to carry less contraband and the officer - still wanting to maximize the outcome - will search White drivers more frequently.

The theory is not without its critics. First, it assumes rationality on everyone’s part. Given the amount of crime driven by mental illness and addiction, rationality of the defendants may not be the best assumption. Second, it assumes that the types of crimes for which people will be searched and contraband will be found is equal among all crime categories and that all races participate in all crimes equally. It is important to understand the assumptions in the model and know that the data does not allow us to test for these assumptions.

Some searches were eliminated due to inconsistencies in the Search Reason and Search Outcome fields. In some cases, an officer indicated a search in the type of search field, but No Search Conducted in the Search Outcome Field. In other cases, the Search Outcome field indicated a search was conducted, but the Search field indicated No Search.

Search Outcomes in Jurisdiction 2

In Jurisdiction 2, of the 4,935 stops, a total of 27 searches were conducted without a warrant. One Latinx driver was searched and contraband was found. White drivers were searched in 26 stops, 23 yielded contraband and three did not.

**Table 11: Search Outcomes for Jurisdiction 2**

<table>
<thead>
<tr>
<th>Race of Operator</th>
<th>Search Outcome</th>
<th>Missing</th>
<th>No Search</th>
<th>Search Reasonable Suspicion</th>
<th>Search Warrant</th>
<th>Search Probable Cause</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>Missing</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No Search Conducted</td>
<td>112</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>112</td>
</tr>
<tr>
<td>Black</td>
<td>Contraband Found</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No Search Conducted</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>140</td>
</tr>
<tr>
<td>Latinx</td>
<td>Contraband Found</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No Search Conducted</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>Native American</td>
<td>No Search Conducted</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Unknown</td>
<td>No Search Conducted</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>White</td>
<td>Missing</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Contraband Found</td>
<td></td>
<td>2</td>
<td></td>
<td>21</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>No Contraband Found</td>
<td></td>
<td>1</td>
<td></td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>No Search Conducted</td>
<td>4575</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4575</td>
</tr>
</tbody>
</table>
Solar Powered Searches

In 2013, Ritter\textsuperscript{19} argued that applying the Veil of Darkness analysis to searches would eliminate some of the problems with the assumptions in the KPT Hit Rate Analysis. He calls his theory Solar Powered Searches. The theory argues that if there are fewer searches of minority drivers in darkness, shown by using regression analysis, then officers may be pulling over minority drivers when race is visible with the intention of searching them. The analysis is conducted the same as the Veil of Darkness analysis for stops. Only those stops occurring between the inter-twilight period are considered. In Jurisdiction 1 there were no searches conducted during this time period. Jurisdiction 2, presented below, had six searches during this time period, all of white drivers.

\textbf{Table 12: Jurisdiction 2 – Solar Powered Searches}

<table>
<thead>
<tr>
<th>Dark</th>
<th>Race of Operator</th>
<th>Search</th>
<th>Search Reasonable Suspicion</th>
<th>Search Probable Cause</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Missing</td>
<td>No Search</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light</td>
<td>Asian</td>
<td>23</td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>31</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Latinx</td>
<td>25</td>
<td></td>
<td></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Native American</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>1</td>
<td>1057</td>
<td>3</td>
<td>1061</td>
</tr>
<tr>
<td>Dark</td>
<td>Asian</td>
<td>20</td>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>18</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Latinx</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>2</td>
<td>676</td>
<td>1</td>
<td>681</td>
</tr>
<tr>
<td>Grand Total</td>
<td>3</td>
<td>1861</td>
<td>1</td>
<td>5</td>
<td>1870</td>
</tr>
</tbody>
</table>

\textbf{Recommendations:} Police should continue to report on post-stop outcomes, but for testing of racial disparities the data and current methods of assessing disparity are not particularly helpful. First, DUI is one of the more common crimes in the state and the traffic stop data are not reflecting DUIs or GNOs. One cannot draw conclusions about what officers are doing post-stop without that data. There are methodological concerns about the KPT Hit-Rate and Solar Powered Searches that add little to the discussion. It is beneficial for citizens to be able to see how the police are working, but no conclusions can be drawn from it.

Objective 5: Website for the Dissemination of Traffic Stops and Race Data

CRG upgraded its website (www.crgvt.org) to facilitate access to the traffic stop and race data and to facilitate downloading capabilities and analysis. CRG posted the original data sets from all LEAs on the CRG website (www.crgvt.org/tsrd) for 2015/16, 2017, and 2018. The data has improved over time. CRG is in the process of working with Crosswind, the Valcour vendor, to extract one traffic stop and race file for all agencies using Valcour, similar to the process DPS is using for the Spillman agencies. CRG anticipates that the 2019 data will be in four or less files. CRG will then work with the State of Vermont website on which all state data is posted and the tools for visualizing and analyzing the data will be added.

Conclusion

CRG standardized the timeframe for collection of traffic stop and race data and facilitated a process for extracting the data from the two CAD/RMSs being used by Vermont LEAs. The data has improved over time and will continue to improve as challenges with the data are identified. The Vermont Legislature requires the collection and publishing of traffic stop and race data annually. CRG will continue to assist with this process to enable researchers and the public to have quality traffic stop and race data published annually in an analyzable format on the Vermont data website.

The purpose of the study was to test different methods of assessing racial disparities in traffic stops for their applicability for Vermont law enforcement agencies. The Commuting Hour analysis pioneered by Connecticut fails when applied to Vermont agencies. Resident Driver analysis is useful for understanding how residents of a town may be treated different than non-residents. It should be included in future analysis of the individual LEAs. The Veil of Darkness analysis is the easiest to perform consistently over time, and we recommend that this be the primary analysis going forward. For the year 2016, this analysis did not find evidence of disparities. It is important to acknowledge that statistical tests check for disparities for which there may be many reasons including racial bias. The fact that there are disparities is not proof of racial bias.

“It is important to understand that a key component of external procedural justice—the practice of fair and impartial policing—is built on understanding and acknowledging human biases, both explicit and implicit. These principles form the foundation of fair and impartial policing. This is not a short term fix but a long range vision.”\textsuperscript{20} In alignment with the national movement on transparency and accountability, the State of Vermont is embracing this move toward recognizing implicit bias and exploring the issues of race in the criminal justice system. The passing of Act 147 is an example of this. It is, however, very important to ensure that the anecdotal evidence is supported by quality data and valid analysis. This project has been an opportunity to continue to delve deeper into these issues.

\textsuperscript{20} 21\textsuperscript{st} Century Policing Report
Appendix A

Factors affecting the validity of the traffic stop and race data (for 2015/2016):

Number of people stopped vs number of tickets: The traffic stop counts in the data as submitted by departments are for tickets, not people. A police officer may give out more than one ticket per stop which can be miscalculated as number of people stopped. State law requires that data associated with an officer’s interactions with a vehicle operator be collected at all roadside stops.

Non-discretionary stops: Tickets and warnings issued as the result of crashes, marijuana tickets, under-age drinking tickets, and externally generated stops may be included in the data sets as if they were discretionary (e.g., a crash may have been coded as a moving violation).

Numbers too small for analysis: In many of the spreadsheets, the numbers of stops for operators of color are too small for valid analysis of the data.

Non-standardization of data entry and coding: Data entry and coding in the law enforcement data collection systems have not been standardized, and officer training on data entry has not been available. Therefore, some percentage of the data may be inaccurate.

Timeframe: Not all data was extracted for the same timeframe.

Missing data: Data was frequently missing from tickets (e.g., race was not coded), or more than one box was checked in a category. Tickets or warnings issued to a company (i.e., commercial motor vehicle stops) were missing data, including age and gender of the driver.

Non-audited data: Data include tickets improperly coded by police officers and have not been checked for accuracy.

Data missing from some police departments: Some police departments were unable to extract the data from their records management systems, so that data is not available.

Analysts should be cautious when comparing results from different jurisdictions because of lack of data, missing data, differing timeframes for data extractions, differences in coding, etc.
Appendix B

R code

######## Libraries
library("compiler", lib.loc="C:/Program Files/R/R-3.4.2/library")
library("datasets", lib.loc="C:/Program Files/R/R-3.4.2/library")
library("graphics", lib.loc="C:/Program Files/R/R-3.4.2/library")
library("utils", lib.loc="C:/Program Files/R/R-3.4.2/library")

###creates new variable
DST.2.Jurisdiction10$dark2 <- NA
DST.2.Jurisdiction10$dark2[DST.2.Jurisdiction10$dark == "Dark"] <- 1
DST.2.Jurisdiction10$dark2[DST.2.Jurisdiction10$dark == "Light"] <- 0
print(DST.2.Jurisdiction10$dark2)

###creates new variable
DST.2.Jurisdiction10$gender2 <- NA
DST.2.Jurisdiction10$gender2[DST.2.Jurisdiction10$Gender == "Female"] <- 0
DST.2.Jurisdiction10$gender2[DST.2.Jurisdiction10$Gender == "Male"] <- 1
print(DST.2.Jurisdiction10$gender2)

### regression model
summary(model9)

Results

Call:

Residuals:
   Min     1Q Median     3Q    Max
-1.6628 -0.9280 -0.9231  1.3585  1.8179

Coefficients: (3 not defined because of singularities)
               Estimate Std. Error z value
(Intercept) -1.857e+01  6.523e+03  -0.003
DST.2.Jurisdiction3$Race.of.OperatorBlack-9.148e+01  8.505e+03  -0.011
DST.2.Jurisdiction3$Race.of.OperatorLatinx  3.171e+00  6.159e+00   0.515
DST.2.Jurisdiction3$Race.of.OperatorWhite  1.845e+01  6.523e+03  -0.003
DST.2.Jurisdiction3$Age1.201e-04  6.370e-02   0.002
DST.2.Jurisdiction3$gender21.700e+01  6.523e+03   0.003
DST.2.Jurisdiction3$Race.of.OperatorBlack:DST.2.Jurisdiction3$Age5.946e+00  2.901e+02   0.020
DST.2.Jurisdiction3$Race.of.OperatorBlack:DST.2.Jurisdiction3$gender27.449e+01  1.243e+04   0.006
DST.2.Jurisdiction3$Race.of.OperatorWhite:DST.2.Jurisdiction3$gender2-1.749e+01  6.523e+03  -0.003
DST.2.Jurisdiction3$Age:DST.2.Jurisdiction3$gender26.175e-03  1.499e-02   0.413

Pr(>|z|)

(Intercept)  0.998
2016/2017 BJS SJS Grant – Fair & Impartial Policing ~ Traffic Stop & Race

Coefficients:

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorBlack</td>
<td>0.991</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorLatinx</td>
<td>0.607</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorWhite</td>
<td>0.998</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Age</td>
<td>0.998</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$gender2</td>
<td>0.998</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorBlack:DST.2Jurisdiction3$Age</td>
<td>0.984</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorLatinx:DST.2Jurisdiction3$Age</td>
<td>0.907</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorWhite:DST.2Jurisdiction3$Age</td>
<td>0.915</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorBlack:DST.2Jurisdiction3$gender2</td>
<td>0.995</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorLatinx:DST.2Jurisdiction3$gender2</td>
<td>NA</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorWhite:DST.2Jurisdiction3$gender2</td>
<td>0.998</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Age:DST.2Jurisdiction3$gender2</td>
<td>0.680</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorBlack:DST.2Jurisdiction3$Age:DST.2Jurisdiction3$gender2</td>
<td>0.988</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorLatinx:DST.2Jurisdiction3$Age:DST.2Jurisdiction3$gender2</td>
<td>NA</td>
</tr>
<tr>
<td>DST.2Jurisdiction3$Race.of.OperatorWhite:DST.2Jurisdiction3$Age:DST.2Jurisdiction3$gender2</td>
<td>NA</td>
</tr>
</tbody>
</table>

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 502.16 on 380 degrees of freedom
Residual deviance: 483.36 on 368 degrees of freedom
AIC: 509.36

Number of Fisher Scoring iterations: 17

Deviance Residuals:
- Min 1Q Median 3Q Max
-1.5395 -1.1257 -0.8035 1.2042 1.8915

Coefficients: (2 not defined because of singularities)

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>362.11</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.OperatorBlack</td>
<td>-361.33</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Latinx</td>
<td>582.61</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1White</td>
<td>-361.93</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Age</td>
<td>-15.08</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$gender2</td>
<td>-364.76</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Black:DST.2Jurisdiction1$Age</td>
<td>15.01</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Latinx:DST.2Jurisdiction1$Age</td>
<td>-16.70</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1White:DST.2Jurisdiction1$Age</td>
<td>15.06</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Black:DST.2Jurisdiction1$gender2</td>
<td>364.83</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Latinx:DST.2Jurisdiction1$gender2</td>
<td>NA</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1White:DST.2Jurisdiction1$gender2</td>
<td>NA</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Age:DST.2Jurisdiction1$gender2</td>
<td>364.67</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Black:DST.2Jurisdiction1$Age:DST.2Jurisdiction1$gender2</td>
<td>15.21</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Latinx:DST.2Jurisdiction1$Age:DST.2Jurisdiction1$gender2</td>
<td>NA</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1White:DST.2Jurisdiction1$gender2</td>
<td>NA</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Age:DST.2Jurisdiction1$gender2</td>
<td>15.15</td>
</tr>
<tr>
<td>DST.2Jurisdiction1$Race.of.Operator1Black:DST.2Jurisdiction1$Age:DST.2Jurisdiction1$gender2</td>
<td>15.15</td>
</tr>
</tbody>
</table>

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 496.48 on 361 degrees of freedom
Residual deviance: 468.81 on 348 degrees of freedom
AIC: 496.81

Number of Fisher Scoring iterations: 18

Deviance Residuals:
- Min 1Q Median 3Q Max
-1.3010 -1.0393 -0.9805 1.3181 1.4179

Coefficients: (8 not defined because of singularities)
2016/2017 BJS SJS Grant – Fair & Impartial Policing ~ Traffic Stop & Race

Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.707e+01 2.400e+03 -0.007 0.994
DST.2Jurisdiction2$Race.of.OperatorBlack 1.658e+01 6.174e+03 0.000 1.000
DST.2Jurisdiction2$Race.of.OperatorNative American 2.025e+02 3.393e+03 0.000 1.000
DST.2Jurisdiction2$Race.of.OperatorWhite 1.622e+01 2.400e+03 0.007 0.995
DST.2Jurisdiction2$Age 1.671e-02 1.785e-02 0.936 0.349
DST.2Jurisdiction2$gender2 5.534e-01 8.364e-01 0.662 0.508
DST.2Jurisdiction2$Race.of.OperatorBlack:DST.2Jurisdiction2$Age NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorNative American:DST.2Jurisdiction2$Age NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorWhite:DST.2Jurisdiction2$Age NA NA NA NA
DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorNative American:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorWhite:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorBlack:DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorNative American:DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorWhite:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorBlack:DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorNative American:DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Race.of.OperatorWhite:DST.2Jurisdiction2$gender2 NA NA NA NA
DST.2Jurisdiction2$Age:DST.2Jurisdiction2$gender2 NA NA NA NA

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 237.08 on 174 degrees of freedom
Residual deviance: 231.86 on 167 degrees of freedom
AIC: 247.86

Number of Fisher Scoring iterations: 15

Call:
glm(formula = DST.2Jurisdiction4$dark2 ~ DST.2Jurisdiction4$Race.of.Operator * DST.2Jurisdiction4$Age * DST.2Jurisdiction4$gender2, family = binomial(link = "logit"), data = DST.2Jurisdiction4)

Deviance Residuals:
Min 1Q Median 3Q Max
-1.2396 -0.9709 -0.7778 1.3571 1.7478

Coefficients: (8 not defined because of singularities)
Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.514e+01 1.455e+03 -0.010 0.992
DST.2Jurisdiction4$Race.of.OperatorBlack -5.461e+01 4.842e+03 -0.111 0.991
DST.2Jurisdiction4$Race.of.OperatorLatinx 4.931e+01 1.456e+03 0.000 1.000
DST.2Jurisdiction4$Race.of.OperatorMissing -1.662e+01 2.058e+03 0.000 1.000
DST.2Jurisdiction4$Race.of.OperatorUnknown -4.284e+01 6.427e+03 0.000 1.000
DST.2Jurisdiction4$Race.of.OperatorWhite 1.426e+01 1.455e+03 0.010 0.992
DST.2Jurisdiction4$Age 6.290e-03 1.402e-02 -0.443 0.658
DST.2Jurisdiction4$gender2 6.280e-01 6.532e-01 0.961 0.336
DST.2Jurisdiction4$Race.of.OperatorBlack:DST.2Jurisdiction4$Age 1.990e+00 1.286e+02 0.015 0.998
DST.2Jurisdiction4$Race.of.OperatorLatinx:DST.2Jurisdiction4$Age 2.579e+01 4.156e+01 0.621 0.535
DST.2Jurisdiction4$Race.of.OperatorMissing:DST.2Jurisdiction4$Age -4.717e+03 6.259e-02 -0.075 0.940
DST.2Jurisdiction4$Race.of.OperatorUnknown:DST.2Jurisdiction4$Age 6.209e+03 1.286e+02 0.000 1.000
DST.2Jurisdiction4$Race.of.OperatorWhite:DST.2Jurisdiction4$Age NA NA NA NA
DST.2Jurisdiction4$Gender2 NA NA NA NA
DST.2Jurisdiction4$Race.of.OperatorBlack:DST.2Jurisdiction4$gender2 NA NA NA NA
DST.2Jurisdiction4$Race.of.OperatorLatinx:DST.2Jurisdiction4$gender2 NA NA NA NA
DST.2Jurisdiction4$Race.of.OperatorMissing:DST.2Jurisdiction4$gender2 1.513e+01 1.455e+03 0.010 0.992
DST.2Jurisdiction4$Race.of.OperatorUnknown:DST.2Jurisdiction4$gender2 1.063e+01 6.260e+03 0.002 0.999
DST.2Jurisdiction4$Race.of.OperatorWhite:DST.2Jurisdiction4$gender2 NA NA NA NA
DST.2Jurisdiction4$Age:DST.2Jurisdiction4$gender2 -5.230e-04 1.657e-02 -0.032 0.975
DST.2Jurisdiction4$Race.of.OperatorBlack:DST.2Jurisdiction4$Age:DST.2Jurisdiction4$gender2 NA NA NA NA
DST.2Jurisdiction4$Race.of.OperatorLatinx:DST.2Jurisdiction4$Age:DST.2Jurisdiction4$gender2 NA NA NA NA
DST.2Jurisdiction4$Race.of.OperatorMissing:DST.2Jurisdiction4$Age:DST.2Jurisdiction4$gender2 NA NA NA NA
DST.2Jurisdiction4$Race.of.OperatorUnknown:DST.2Jurisdiction4$Age:DST.2Jurisdiction4$gender2 6.463e-02 1.286e+02 0.001 1.000
DST.2Jurisdiction4$Race.of.OperatorWhite:DST.2Jurisdiction4$Age:DST.2Jurisdiction4$gender2 NA NA NA NA

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 498.67 on 390 degrees of freedom
Residual deviance: 482.85 on 375 degrees of freedom
(3 observations deleted due to missingness)
AIC: 514.85

Number of Fisher Scoring iterations: 14
Call: glm(formula = DST.2Jurisdiction5$dark2 ~ DST.2Jurisdiction5$Race.of.Operator * DST.2Jurisdiction5$Age * DST.2Jurisdiction5$gender2, family = binomial(link = "logit"), data = DST.2Jurisdiction5)

Deviance Residuals:
  Min 1Q Median 3Q Max
-1.4938 -0.9882 -0.9081 1.3735 1.7417

Coefficients: (3 not defined because of singularities)
  Estimate Std. Error z value Pr(>|z|)
  (Intercept) 5.5394 10.1052 0.548 0.584
  DST.2Jurisdiction5$Race.of.OperatorBlack -25.1054 18037.1947 -0.001 0.999
  DST.2Jurisdiction5$Race.of.OperatorLatinx -229.9960 26761.0225 -0.009 0.993
  DST.2Jurisdiction5$Race.of.OperatorNative American -61.8561 13072.5917 -0.005 0.996
  DST.2Jurisdiction5$Race.of.OperatorWhite -5.5349 10.1141 -0.547 0.584
  DST.2Jurisdiction5$Age 0.2116 0.3646 -0.580 0.562
  DST.2Jurisdiction5$gender2 116.3679 19596.2630 0.006 0.995
  DST.2Jurisdiction5$Race.of.OperatorBlack:DST.2Jurisdiction5$Age 2.2116 612.1048 0.000 1.000
  DST.2Jurisdiction5$Race.of.OperatorLatinx:DST.2Jurisdiction5$Age 5.0436 583.7947 0.009 0.993
  DST.2Jurisdiction5$Race.of.OperatorNative American:DST.2Jurisdiction5$Age NA NA NA NA
  DST.2Jurisdiction5$Race.of.OperatorWhite:DST.2Jurisdiction5$Age 0.1955 0.3648 0.536 0.592
  DST.2Jurisdiction5$Race.of.OperatorBlack:DST.2Jurisdiction5$gender2 -95.7628 26633.6971 -0.004 0.997
  DST.2Jurisdiction5$Race.of.OperatorLatinx:DST.2Jurisdiction5$gender2 1048.2024 43458.4128 0.000 1.000
  DST.2Jurisdiction5$Race.of.OperatorNative American:DST.2Jurisdiction5$gender2 NA NA NA NA
  DST.2Jurisdiction5$Race.of.OperatorWhite:DST.2Jurisdiction5$gender2 116.7848 19596.2630 -0.006 0.995
  DST.2Jurisdiction5$Age:DST.2Jurisdiction5$gender2 3.7703 625.4604 0.000 0.995

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 738.74 on 560 degrees of freedom
Residual deviance: 713.50 on 544 degrees of freedom
(1 observation deleted due to missingness)
AIC: 747.5

Number of Fisher Scoring iterations: 18

Call: glm(formula = DST.2Jurisdiction6$dark2 ~ DST.2Jurisdiction6$Race.of.Operator * DST.2Jurisdiction6$Age * DST.2Jurisdiction6$gender2, family = binomial(link = "logit"), data = DST.2Jurisdiction6)

Deviance Residuals:
  Min 1Q Median 3Q Max
-1.0173 -0.8212 -0.7218 1.3557 1.9008

Coefficients: (5 not defined because of singularities)
  Estimate Std. Error z value Pr(>|z|)
  (Intercept) -1.723e+01 1.546e+04 -0.001 0.999
  DST.2Jurisdiction6$Race.of.OperatorUnknown 1.739e-08 2.827e+03 0.000 1.000
  DST.2Jurisdiction6$Race.of.OperatorWhite 1.653e+01 1.546e+04 0.001 0.999
  DST.2Jurisdiction6$Age 8.020e-03 4.198e+02 0.000 1.000
  DST.2Jurisdiction6$gender2 6.615e-01 9.375e-01 0.706 0.480
  DST.2Jurisdiction6$Race.of.OperatorUnknown:DST.2Jurisdiction6$Age NA NA NA NA
  DST.2Jurisdiction6$Race.of.OperatorWhite:DST.2Jurisdiction6$Age -2.099e-02 4.198e+02 0.000 1.000
  DST.2Jurisdiction6$Race.of.OperatorUnknown:DST.2Jurisdiction6$gender2 NA NA NA NA
  DST.2Jurisdiction6$Race.of.OperatorWhite:DST.2Jurisdiction6$gender2 NA NA NA NA
  DST.2Jurisdiction6$Age:DST.2Jurisdiction6$gender2 -8.020e-03 2.412e-02 -0.332 0.740
  DST.2Jurisdiction6$Race.of.OperatorUnknown:DST.2Jurisdiction6$Age:DST.2Jurisdiction6$gender2 NA NA NA NA
  DST.2Jurisdiction6$Race.of.OperatorWhite:DST.2Jurisdiction6$Age:DST.2Jurisdiction6$gender2 NA NA NA NA
  DST.2Jurisdiction6$Race.of.OperatorUnknown:DST.2Jurisdiction6$Age:DST.2Jurisdiction6$gender2 NA NA NA NA
  DST.2Jurisdiction6$Race.of.OperatorWhite:DST.2Jurisdiction6$Age:DST.2Jurisdiction6$gender2 NA NA NA NA

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 234.62 on 198 degrees of freedom
Residual deviance: 227.64 on 192 degrees of freedom
AIC: 241.64
Number of Fisher Scoring iterations: 15

Call:
```r
glm(formula = DST.2Jurisdiction7$dark2 ~ DST.2Jurisdiction7$Race.of.Operator * (Intercept
DST.2Jurisdiction7$Age * DST.2Jurisdiction7$gender2, family = binomial(link = "logit"),
data = DST.2Jurisdiction7)
```

Deviance Residuals:
Min 1Q Median 3Q Max
-1.2396 -0.9709 -0.7778 1.3571 1.7478

Coefficients: (8 not defined because of singularities)
Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.514e+01 1.455e+03 -0.010 0.992
DST.2Jurisdiction7$Race.of.OperatorBlack -5.461e+01 4.842e+03 -0.111 0.991
DST.2Jurisdiction7$Race.of.OperatorLatinx 4.931e-01 1.456e+03 0.000 1.000
DST.2Jurisdiction7$Race.of.OperatorMissing -1.662e-01 2.058e+03 0.000 1.000
DST.2Jurisdiction7$Race.of.OperatorUnknown -4.284e-01 6.427e+03 0.000 1.000
DST.2Jurisdiction7$Race.of.OperatorWhite 1.426e+01 1.455e+03 0.010 0.992
DST.2Jurisdiction7$Age 6.209e-03 1.402e-02 -0.443 0.658
DST.2Jurisdiction7$gender2 6.280e-01 6.532e-01 0.961 0.336
DST.2Jurisdiction7$Race.of.OperatorUnknown:DST.2Jurisdiction7$Age 1.990e+00 1.286e+02 0.151 0.888
DST.2Jurisdiction7$Race.of.OperatorLatinx:DST.2Jurisdiction7$Age 2.579e-01 4.155e-01 0.621 0.535
DST.2Jurisdiction7$Race.of.OperatorMissing:DST.2Jurisdiction7$Age -4.717e-03 6.259e-02 -0.075 0.940
DST.2Jurisdiction7$Race.of.OperatorUnknown:DST.2Jurisdiction7$gender2 6.209e-03 1.286e+02 0.000 1.000
DST.2Jurisdiction7$Race.of.OperatorWhite:DST.2Jurisdiction7$Age NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorBlack:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorLatinx:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorMissing:DST.2Jurisdiction7$gender2 1.513e+01 1.455e+03 0.010 0.992
DST.2Jurisdiction7$Race.of.OperatorUnknown:DST.2Jurisdiction7$gender2 1.063e+01 6.260e+03 0.002 0.999
DST.2Jurisdiction7$Race.of.OperatorWhite:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Age:DST.2Jurisdiction7$gender2 -5.230e-04 1.657e-02 -0.032 0.975
DST.2Jurisdiction7$Race.of.OperatorBlack:DST.2Jurisdiction7$Age:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorLatinx:DST.2Jurisdiction7$Age:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorMissing:DST.2Jurisdiction7$Age:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorUnknown:DST.2Jurisdiction7$Age:DST.2Jurisdiction7$gender2 NA NA NA NA
DST.2Jurisdiction7$Race.of.OperatorWhite:DST.2Jurisdiction7$Age:DST.2Jurisdiction7$gender2 6.463e-02 1.286e+02 0.001 1.000

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 498.67 on 390 degrees of freedom
Residual deviance: 482.85 on 375 degrees of freedom
(3 observations deleted due to missingness)
AIC: 514.85

Number of Fisher Scoring iterations: 14

Call:
```r
glm(formula = DST.2Jurisdiction8$dark2 ~ DST.2Jurisdiction8$Race.of.Operator * (Intercept
DST.2Jurisdiction8$Age * DST.2Jurisdiction8$gender2, family = binomial(link = "logit"),
data = DST.2Jurisdiction8)
```

Deviance Residuals:
Min 1Q Median 3Q Max
-1.0664 -0.9382 -0.9076 1.3938 1.6014

Coefficients: (13 not defined because of singularities)
Estimate Std. Error z value Pr(>|z|)
(Intercept) 5.471e-02 2.976e+00
DST.2Jurisdiction8$Race.of.OperatorBlack 8.916e+01 9.295e+03
DST.2Jurisdiction8$Race.of.OperatorLatinx 6.479e+01 1.320e+04
DST.2Jurisdiction8$Race.of.OperatorMissing 2.000e+01 6.523e+03
DST.2Jurisdiction8$Race.of.OperatorNative American 1.962e+01 6.523e+03
DST.2Jurisdiction8$Race.of.OperatorUnknown -1.883e+01 2.352e+04
DST.2Jurisdiction8$Race.of.OperatorWhite -0.255e-01 2.929e+00
DST.2Jurisdiction8$Age -4.887e-02 1.171e-01
DST.2Jurisdiction8$gender2 3.371e-02 6.813e-01

26
Coefficients:

```
Call:
  data = DST.2.Jurisdiction)

Deviance Residuals:
  Min       1Q   Median       3Q      Max
-1.4225  -0.8143  -0.6899  1.2294  2.0896

Coefficients:

|                      | Estimate | Std. Error | z value | Pr(>|z|) |
|----------------------|----------|------------|---------|----------|
| (Intercept)          | 0.049    | 0.961      | -0.011  | 0.996    |
| DST.2.Jurisdiction8$Age:DST.2.Jurisdiction8$gender2 | -0.417   | 0.676      | -0.632  | 0.527    |
| DST.2.Jurisdiction8$Race.of.OperatorBlack:DST.2.Jurisdiction8$Age | -0.016   | 0.987      | -0.016  | 0.996    |
| DST.2.Jurisdiction8$Race.of.OperatorLatinx:DST.2.Jurisdiction8$Age | -0.005   | 0.996      | -0.005  | 0.996    |
| DST.2.Jurisdiction8$Race.of.OperatorMissing:DST.2.Jurisdiction8$Age | 0.003    | 0.998      | 0.003   | 0.998    |
| DST.2.Jurisdiction8$Race.of.OperatorNative American:DST.2.Jurisdiction8$Age | 0.003    | 0.998      | 0.003   | 0.998    |
| DST.2.Jurisdiction8$Race.of.OperatorUnkown:DST.2.Jurisdiction8$Age | 0.000    | 1.000      | 0.000   | 1.000    |
| DST.2.Jurisdiction8$Race.of.OperatorWhite:DST.2.Jurisdiction8$Age | -0.316   | 0.752      | -0.417  | 0.676    |
| DST.2.Jurisdiction8$Race.of.OperatorBlack | -0.010   | 0.992      | -0.010  | 0.992    |
| DST.2.Jurisdiction8$Race.of.OperatorLatinx | 0.005    | 0.996      | 0.005   | 0.996    |
| DST.2.Jurisdiction8$Race.of.OperatorMissing | 0.003    | 0.998      | 0.003   | 0.998    |
| DST.2.Jurisdiction8$Race.of.OperatorNative American | 0.003    | 0.998      | 0.003   | 0.998    |
| DST.2.Jurisdiction8$Race.of.OperatorUnkown | -0.001   | 0.999      | -0.001  | 0.999    |
| DST.2.Jurisdiction8$Race.of.OperatorWhite | -0.316   | 0.752      | -0.316  | 0.752    |
```

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 395.88 on 302 degrees of freedom
Residual deviance: 378.11 on 288 degrees of freedom
(1 observation deleted due to missingness)
AIC: 408.11

Number of Fisher Scoring iterations: 17
Deviance glm(formula = DST.2$sta$Race.of.Operator ~ DST.2$sta$Age:DST.2$sta$gender2, family = binomial(link = "logit"), data = DST.2$sta)

Deviance Residuals:

Min 1Q Median 3Q Max
-1.3072 -0.0788 -0.0435 1.2181 1.4790

Coefficients: (6 not defined because of singularities)

Estimate Std. Error z value Pr(>|z|)

(Intercept) -1.944e+01 1.262e+04
DST.2$sta$Race.of.OperatorBlack 4.79e-01 2.648e+04
DST.2$sta$Race.of.OperatorLatinx 1.878e+00 1.501e+04
DST.2$sta$Race.of.OperatorNative American 3.513e+01 4.845e+03
DST.2$sta$Race.of.OperatorOperatorUnknown 2.870e+01 1.262e+04
DST.2$sta$Race.of.OperatorWhite 1.811e+01 1.262e+04

DST.2$sta$Race.of.OperatorWhite:DST.2$sta$Age:DST.2$sta$gender2 3.410e-02 2.797e+02
DST.2$sta$Race.of.OperatorLatinx:DST.2$sta$Age:DST.2$sta$gender2 1.878e+00 1.111e+00
DST.2$sta$Race.of.OperatorNative American:DST.2$sta$gender2 NA NA
DST.2$sta$Race.of.OperatorUnknown:DST.2$sta$Age:DST.2$sta$gender2 -1.888e+01 2.797e+02
DST.2$sta$Race.of.OperatorUnknown:DST.2$sta$gender2 NA NA
DST.2$sta$Race.of.OperatorUnknown:DST.2$sta$Age:DST.2$sta$gender2 -3.349e+02 2.797e+02
DST.2$sta$Race.of.OperatorUnknown:DST.2$sta$gender2 NA NA
DST.2$sta$Race.of.OperatorUnknown:DST.2$sta$Age:DST.2$sta$gender2 NA NA
DST.2$sta$Race.of.OperatorUnknown:DST.2$sta$gender2 NA NA

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 175.79 on 141 degrees of freedom
Residual deviance: 156.61 on 127 degrees of freedom
AIC: 186.61

Number of Fisher Scoring iterations: 16

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Intercept)  -1.557e+01 1.455e+03  -0.011  0.991
DST.2sta$Race.of.OperatorBlack  3.113e+01 2.058e+03  0.015  0.988
DST.2sta$Race.of.OperatorWhite  1.499e+01 1.455e+03  0.010  0.992
DST.2sta$Age  5.291e-05 1.940e-02  0.003  0.998
DST.2sta$gender2  1.093e+00 8.962e-01 1.219  0.223
DST.2sta$Race.of.OperatorBlack:DST.2sta$Age NA NA NA NA
DST.2sta$Race.of.OperatorWhite:DST.2sta$Age NA NA NA NA
DST.2sta$Race.of.OperatorBlack:DST.2sta$gender2 NA NA NA NA
DST.2sta$Race.of.OperatorWhite:DST.2sta$gender2 NA NA NA NA
DST.2sta$Age:DST.2sta$gender2 -1.768e-02 2.364e-02  -0.748  0.455
DST.2sta$Race.of.OperatorBlack:DST.2sta$Age:DST.2sta$gender2 NA NA NA NA
DST.2sta$Race.of.OperatorWhite:DST.2sta$Age:DST.2sta$gender2 NA NA NA NA

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 227.75 on 166 degrees of freedom
Residual deviance: 221.38 on 161 degrees of freedom
(1 observation deleted due to missingness)
AIC: 233.38

Number of Fisher Scoring iterations: 14