RTI-STAR Report

Report generated at: 04/01/18 09:07:35
Original filename: RP 2012-17.csv
Columns used:

Date and time: Stop Date
Time: Stop Date
Target: Race
Officer ID: StatePID

Target group: B
Restrict dates to +/- 30 days from DST transition: False
Date range of cases: 01/03/12 - 12/29/17
Intertwilight period (ITP) range: 17:30:22 - 21:07:00
Original number of stop records: 16,483
Number of stop records used for analysis: 1,539
Number of daylight events in ITP: 710
Number of darkness events in ITP: 829

Based on statistical significance alone (p>=0.05), this evidence does not suggest the existence of racial disproportionality in traffic stops. However, given the typically large sample sizes available for traffic stop data, we recommend that interpretation of results be more focused on whether the size of the difference in the reported percentages (or, equivalently, their ratio) is meaningfully large for your community.

The regression model indicates that the risk of being in the reference group was 21.63% during the light portion of the intertwilight period and 24.65% during the dark portion of the intertwilight period, with a corresponding risk ratio of 1.04. The p-value for the effect of light versus dark was 0.8121.

Methodology

To study the racial distribution of traffic stops, the tool uses the Veil of Darkness (VOD) approach, which is based on the logic that police officers are less capable of determining the race of a motorist after dark than they are during daylight. Using this method, the existence of racial disproportionality in traffic stops is assessed by comparing the race distribution of stops made during daylight to the race distribution of stops made after dark, after adjusting for other factors in a regression model. The analysis is limited to stops that occur during the evening intertwilight period (roughly between 5:00 PM and 9:00 PM, depending on location) in order to reduce the variation in travel patterns that are conditional on time of day.

The VOD method was developed and first employed by Jeffery Grogger and Greg Ridgeway in an analysis of traffic stops in Oakland, California, and Cincinnati, Ohio. The method has also been used in studies focusing on the nature of traffic stops in Minneapolis, Minnesota, Syracuse, New York, and more recently in San Diego, California, and the state of Connecticut.

Our tool incorporates one enhancement from these previous studies. That is, our model accounts for within-officer correlation that is likely to occur (when this information is available in the data). By doing so, we recognize that officers may have inherent differences in the percentage of a racial or ethnic subset of the population that they are likely to encounter. These differences may be caused by factors such as geographic deployment or unit assignment.

How to interpret results

One benefit of the VOD approach is its statistical quantification of disproportionality present in a community using a reliable data source. If the reported **risk ratio** (i.e., risk of being in a traffic stop reference group during light vs. dark periods) is acceptably close to 1.0, or in other words if the **percentages** of reference group traffic stops during light vs. dark intertwilight are acceptably close to one another, it suggests that daylight was not meaningfully associated with the race or ethnicity of the driver who was stopped. Alternatively, if the risk ratio is meaningfully *greater than 1.0*, it suggests possible racial disproportionality (i.e., that reference group motorists, a racial or ethnic subset of the population of interest, are more likely to be among those stopped during times when visibility is higher compared to times when visibility is lower).

We provide the **level of statistical significance** for the effect of daylight vs. intertwilight as well. Significance levels between 0 and 0.05 (when accompanied by risk ratios *greater than 1.0*) are more indicative of a potential problem with disproportionality, and significance levels between 0.05 and 1.0 are less indicative of a potential problem. However, given the large sample sizes available for traffic stop data, interpretation of results should be more focused on the size of the difference in the percentages (or, equivalently, their ratio).

This tool provides a quantified description of traffic stop disparities in a community. The results can be useful in a number of ways: to determine whether there is meaningful evidence of racial disproportionality in traffic stops for a given reference group, to compare results across various subgroups of stops (e.g., across gender, across police units) to pinpoint subgroups that need more attention than others, and to compare results before vs. after an intervention is implemented.

This method identifies patterns of disproportionate contact with law enforcement in regard to traffic stops for selected reference groups in a defined time range. This tool does **not** alone conclusively identify agencies and/or officers that are engaging in the practice of racial bias. Finally, this method does not analyze or provide context for the reason or cause for a traffic stop or set of traffic stops to be conducted by a law enforcement officer.

This tool does not provide a customized analysis for each dataset and results should be considered preliminary. Further analyses may still be warranted to better understand the results and verify that the preliminary analysis has been conducted in a proper and legitimate manner.

How we calculated your results

Traffic stop profiling analysis was performed using Python 3.5.1 and statsmodels version 0.6.1.

For analysis with an officer id, a marginal regression model was fit using Generalized Estimating Equations (GEE). Default initialization parameters were used with the exceptions of a logit link, binomial distribution and an exchangeable

covariance structure with a robust variance estimator. The fitted model included categorical variables representing light/dark, day of week, year and a B-spline (degree = 6) of time in seconds since midnight. Officer id was specified as the group (cluster or "subject" variable).

For analysis without an officer id a binary choice logit model was fit. The form of the model was identical to the GEE model but no within-subject covariance structure was specified.