AEA Papers and Proceedings 2019, 109: 143–147
https://doi.org/10.1257/pandp.20191026

RACE AND POLICING[‡]

Now You See Me, Now You Don't: The Geography of Police Stops[†]

By Jesse J. Kalinowski, Matthew B. Ross, and Stephen L. Ross*

The high share of minorities involved in traffic stops is frequently cited as evidence of racial profiling. On the other hand, police often attribute these disparities to racial and ethnic differences in guilt. The problem in assessing the racial share of traffic stops is that we do not know the composition of motorists at risk of being stopped.

A recent solution to this problem, the Veil of Darkness (VOD) test, was proposed by Grogger and Ridgeway (2006) and recently applied by Ridgeway (2009) and Horrace and Rohlin (2016). In VOD, the racial composition of stops after sunset is used to assess the distribution of stops in daylight (at the same time of day and day of week) since race is unlikely to be observed by police after sunset.

A key maintained hypothesis of the VOD is that neither motorists nor police change their behavior based on the amount of light. However, Kalinowski, Ross, and Ross (2017) shows that black motorists drive slower during the day when their race can be observed. Fazzalaro et al. (2018) presents maps and describes conversations with police suggesting that suburban town police change patrol locations at night, moving closer to urban centers and shifting away from speeding and toward equipment violation enforcement.

This paper uses state police stop data in Texas to assess patrol activity. We find evidence that both the types of stops and the allocation of resources change in darkness relative to daylight, and that the changes in stop type and manpower are correlated within police officers. Further, we find that counties receiving more police resources in darkness have a higher share of minority residents.

I. Texas Police Stop Data

The paper uses data from the Stanford Open Policing Project on stops made by Texas Highway Patrol officers from 2010 to 2015. These officers are assigned to 1 of 19 highway patrol districts, and each district contains between 3 and 30 counties, on average approximately 13 counties per district.[1] The data identify the location, date and time, and the reason for the stop, i.e., either the type of warning or citation issued, the race and ethnicity of the motorist, and an identifier for the police officer making the stop.[2]

[‡]*Discussants:* Jennifer Doleac, Texas A&M University; Felipe Goncalves, Princeton University.

*Kalinowski: Department of Economics, Quinnipiac University, 275 Mt. Carmel Avenue, Hamden, CT 06518 (email: jesse.kalinowski@gmail.com); Ross: Wagner School of Public Service, New York University, 295 Lafayette Street, New York, NY 10012 (email: mbr1@nyu.edu); Ross: Department of Economics, University of Connecticut, 365 Fairfield Way, U-1063, Storrs, CT (email: stephen.l.ross@uconn.edu). M. Ross and J. Kalinowski received prior support from the State of Connecticut to conduct research on racial bias in traffic stops. We thank Sharad Goel and the Stanford Open Policing Project for access to the Texas Highway Patrol stop data. We are also grateful to Ken Barone and James Fazzalaro of the Connecticut Racial Profiling Prohibition Project for their invaluable perspective on policing. All remaining errors are our own.

[†]Go to https://doi.org/10.1257/pandp.20191026 to visit the article page for additional materials and author disclosure statement(s).

[1]The data report a patrol district for each officer, but this variable has poor coverage. Therefore, we assign officers to patrol districts based on where they made the majority of their stops within a given year.

[2]Our analysis restricts the sample to stops of white motorists in the inter-twilight window by officers in the seventy-fifth percentile of total district stops. The sample includes 1,311,191 total stops and 997 officers.

We establish an inter-twilight window using data from the United States Naval Observatory (USNO) with a lower bound of the earliest time of day that sunset begins during the year in the easternmost county of Texas and with an upper bound of the latest end to the evening, civil twilight in the westernmost county.[3] We select all stops that fall within the inter-twilight window except for those stops that fall during actual twilight for the stop date, again using the earliest start and latest end of twilight in Texas.

We restrict our sample to officers issuing citations by selecting only those officers in the seventy-fifth or higher percentile on number of warnings/citations issued within their district. We also only use warnings and citations issued to non-Hispanic, white motorists.[4]

II. Type of Police Stops

We first document that the distribution of police manpower by violation type varies considerably between daylight and darkness. Figure 1 shows a substantially larger number of stops associated with warnings for equipment violations at night versus a much larger number of speeding stops and miscellaneous warnings in daylight. These findings are consistent with the anecdotal stories of police in Connecticut (see Fazzalaro et al. 2018).

To assess whether this variation is systematic, we calculate the fraction of manpower based on each officer's share of stops for each of the 20 warning/citation types separately for daylight and darkness. We then sum the absolute value of the daylight-darkness difference over types and divide by the number of types. The mean difference is 1.8 percentage points, while random allocation with 20 types implies an average citation share of 5 percent.

We then resample police officers retaining the county and time of each officer's stop and selecting a date (seasonal variation) randomly, simulating the daylight/darkness treatment for each stop under the null hypothesis of no changes between daylight and darkness. The average statistic under the null is 1.12 with a standard deviation of 0.01, and the fraction of simulations exceeding 1.8 points is far below 1 in a 1,000. Similar results arise restricting our sample to stops made within 28 days on either side of daylight saving time (DST).[5]

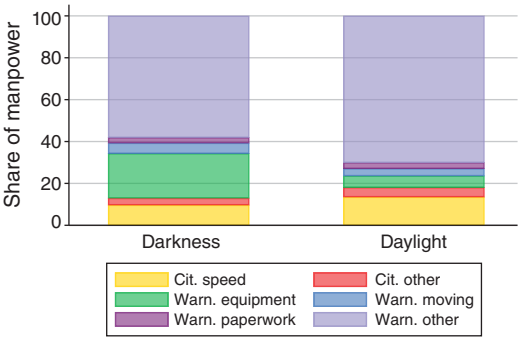

Figure 1. Change in Violation Distribution

Notes: Each column shows the fraction of police manpower associated with each type of citation/warning. Stops with multiple citations or warnings have their contribution divided evenly between each category.

III. Location of Stops

Next, we describe changes in the allocation of police manpower across counties between daylight and darkness. For each officer, we allocate their time hour by hour each day based on the county in which they issued one or more warning/citation. If warnings/citations are issued in more than one county in an hour, their time is evenly divided between the counties. This approach captures the location of officers at a given time without using information on the number of stops, which might be correlated with darkness or the share of minority motorists.

We then calculate the fraction of police resources allocated to each county in Texas by daylight and darkness plus the daylight-darkness difference. The seventy-fifth percentile county has a 0.044 percentage point increase in the share of resources from daylight to darkness, while the twenty-fifth percentile has a 0.036 point decrease. Equal allocation across all counties would imply 0.39 percent of resources allocated per county so that the twenty-fifth and

[3] Civil twilight is defined as the time period beginning at the end of sunset and ending when the sun is six degrees below the horizon.

[4] Results are robust to officers above the median or above the ninetieth percentile.

[5] See the online Appendix DST sample results.

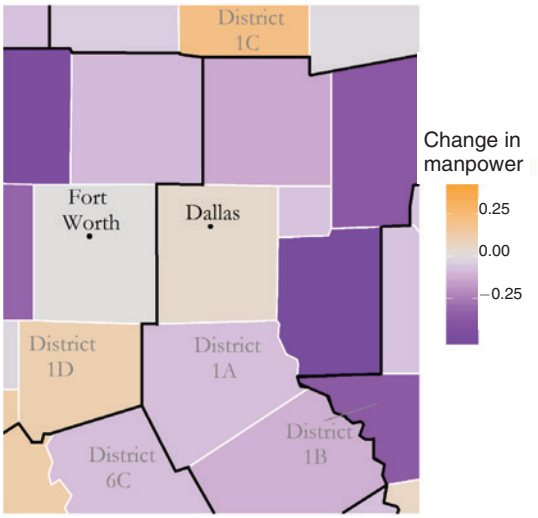

FIGURE 2. CHANGE IN MANPOWER

Notes: The daylight to darkness change in share of state police resources is based on non-Hispanic, white stops. The orange shading represents a larger fraction of resources allocated in daylight with purple representing more resources in darkness.

seventy-fifth percentile counties both represent a 10 percent shift relative to the average allocation.[6]

Figure 2 shows an example of this exercise for state police districts located in and around Dallas-Ft. Worth. The figure shades counties based on differences in allocation of police resources between daylight and darkness.

The orange shading shows a shift in resources toward Dallas in daylight and to suburban counties surrounding Dallas and Ft. Worth at night. The orange county at the top in district 1C contains Sherman, the primary urban area in 1C. Unlike in Connecticut, state police resources in these districts are shifted toward suburban counties after sunset.

More generally, we summarize the extent of change across counties within districts by averaging the absolute value of this difference over all counties. The resulting statistic summed across counties is 0.07 percentage points, which is approximately 18 percent of the mean allocation. Then, we resample police officers from the data retaining the county and the time of stop and selecting a date randomly again simulating the daylight/darkness treatment under the null hypothesis. The average statistic under the null is 0.059 with a standard deviation of 0.001, and the fraction of simulations that exceed 0.07 percentage points is far below 1 in 1,000. Results are robust to the use of speeding stops or of stops made within 28 days of daylight saving time (DST).[7]

IV. Stops and Racial Composition

Finally, we compare the geographic manpower shifts to changes in types of stops, racial composition of counties and size of the counties. We create a sample of police officer by county observations calculating the daylight to darkness shift in the officer's resource allocation to that county using either all stops or only speeding stops. For each officer, we calculate the change in the violation type as the average over all types of the absolute value of the daylight to darkness change in percent of stops of each type. For each county, we use 2017 census data to calculate the fraction of residents who are black, the fraction non-Hispanic white, and identify the highest population county in each district.

Table 1, column 1 contains estimates from regressing officer average absolute change in violation type on the absolute value of the resource-allocation change. Columns 2, 3, and 4 contain estimates from regressing the county racial composition or size indicator on the signed value of the resource-allocation change. All models include district fixed effects and cluster standards errors by district. Panel A presents the annual sample and panel B the DST sample.

With one exception, we find a strong relationship between violation type and resource allocation. Officers who reallocate time across counties between daylight and darkness also change the type of stops they make. The standard deviation of the geographic change is 3.6 percentage points between day and night, relative to an average of about 10 percent (the average police officer makes stops in ten counties). Focusing on the three significant estimates, such

[6] Note that we are using 252 rather than 254 counties because we drop the two counties on the mountain time zone border.

[7] See the online Appendix for statewide maps for annual and DST samples.

Table 1—Policing Changes

Outcome	Violation type	Black	Non-Hispanic white	Largest county
Panel A. Annual sample				
All stops	0.035	0.050	−0.055	−0.001
	(0.008)	(0.010)	(0.031)	(0.001)
Speeding stops only	0.000	0.043	−0.082	0.006
	(0.003)	(0.014)	(0.015)	(0.002)
Panel B. Daylight saving time sample				
All stops	0.076	0.007	0.018	−0.001
	(0.01)	(0.02)	(0.045)	(0.001)
Speeding stops only	0.046	0.003	0.008	−0.000
	(0.011)	(0.012)	(0.023)	(0.001)

Notes: The independent variable is the percentage change in county police resources between daylight and darkness from 0 to 100: absolute value for type and signed changes for county demographics. Dependent variables are listed in the top row. Violation-type changes are standardized, and percent black and non-Hispanic white range from 0 to 100. Models include district fixed effects with standard errors clustered by district.

a geographic change is associated with between 0.13 and 0.27 standard deviation changes in stop-type composition.

In the annual sample, we find that counties containing more minority residents receive less resources in daylight. However, the effect is small with a one standard deviation change in geography associated with changes of between 0.2 and 0.3 percentage points. For speeding stops only, we find that more resources are allocated to the largest population counties at night. A one standard deviation geographic change implies a 2 percentage point change in the likelihood of a stop's county being the largest population county in the district, relative to a 1 in 13 chance on average.

Neither the demographic nor county size results are robust to the DST sample, but our ability to identify robust demographic patterns may be limited by our use of county, as opposed to neighborhoods or cities and towns.

V. Implications for Discrimination Tests

In this paper, we demonstrate that when darkness falls, regardless of the time of day, police officers change their patrolling behavior. The types of stops change substantially with a decrease in the number of speeding stops and an increase in number of equipment warnings in darkness. We also observe a geographic shift between daylight and darkness in where stops are being made, and this geographic shift appears linked to the changes in stop type. We also find evidence of more stops in darkness in minority counties and in the largest counties. These findings mirror evidence in Connecticut.

Grogger and Ridgeway (2006) focuses on a sample of speeding stops based on concerns that discrimination might be masked because police focus on different types of stops at night. The findings support their decision to separate different types of stops when conducting Veil of Darkness (VOD) tests. However, the nature of the speeding stops could still change between daylight and darkness. If police are not focused on moving violations at night, they may only pull over the most extreme violations, and the resulting sample may not be comparable to the sample in daylight when police focus on routine speeding stops.

Grogger and Ridgeway (2006) notes that seasonal changes can confound identification in the annual VOD sample, but suggest that estimation within a window surrounding the DST change or in urban areas would likely mitigate this problem. Our findings suggest that changes in the geographic distribution of stops occur in urban areas and persist through more rigorous sample restrictions.

In practice, researchers need to understand how the geographic distribution of stops changes between daylight and darkness and then include controls for geography at the appropriate geographic level. Otherwise, VOD tests may fail to find discrimination if police shift enforcement activities in darkness toward more urban areas and/or toward places with a greater minority-residential share. Simple controls for state police barracks or town/city police departments are insufficient to address this concern because police are reallocating activity within their own jurisdiction.

REFERENCES

Fazzalaro, James, Ken Barone, Matthew B. Ross, and Jesse Kalinowski. 2018. *Traffic Stop Data Analysis and Findings, 2015–16*. New Britain: Connecticut Racial Profiling Prohibition Project.

Grogger, Jeffrey, and Greg Ridgeway. 2006. "Testing for Racial Profiling in Traffic Stops from behind a Veil of Darkness." *Journal of the*

American Statistical Association 101 (475): 878–87.

Horrace, William C., and Shawn M. Rohlin. 2016. "How Dark Is Dark? Bright Lights, Big City, Racial Profiling." *Review of Economics and Statistics* 98 (2): 226–32.

Kalinowski, Jesse, Stephen L. Ross, and Matthew B. Ross. 2017. "Endogenous Driving Behavior in Veil of Darkness Tests for Racial Profiling." Human Capital and Economic Opportunity Working Paper 2017–17.

Ridgeway, Greg. 2009. *Cincinnati Police Department Traffic Stops: Applying RAND's framework to Analyze Racial Disparities*. Santa Monica: RAND Corporation.