NBER WORKING PAPER SERIES

ENDOGENOUS DRIVING BEHAVIOR IN TESTS OF RACIAL PROFILING

Jesse Kalinowski
Matthew B. Ross
Stephen L. Ross

Working Paper 28789
http://www.nber.org/papers/w28789

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2021

For insightful comments, we thank Talia Bar, Patrick Bayer, Hanming Fang, Felipe Goncalves, Jeffrey Grogger, William Horrace, John MacDonald, Steve Mello, Magne Mogstad, Greg Ridgeway, Shawn Rohlin, Jesse Shapiro, Austin Smith, Jeremy D. West, John Yinger and Bo Zhao. We also thank participants at the NBER Summer Institute 2017 Law and Economics Workshop, 2017 Association for Policy Analysis and Management 2017 Research Conference, Syracuse University, University at Albany, 2016 Urban Economics Meetings, Federal Reserve Bank of Boston, Miami University of Ohio, University of Connecticut, and Ohio State University. We are also grateful for the help of Bill Dedman who provided us with the Massachusetts traffic stop data from his 2003 Boston Globe article with Francie Latour and Sharad Goel and the Stanford Open Policing Project for providing the Tennessee Highway Patrol stop data, as well as Ken Barone and James Fazzalaro for their invaluable perspective on policing. All remaining errors are our own. This research has not been directly funded by any outside agency. However, J. Kalinowski and M. Ross acknowledge prior support from the State of Connecticut and the State of Rhode Island to analyze and conduct research on racial bias in traffic stops. All data used in this study is administrative in nature and publicly available, and so not subject to IRB review. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Endogenous Driving Behavior in Tests of Racial Profiling
Jesse Kalinowski, Matthew B. Ross, and Stephen L. Ross
NBER Working Paper No. 28789
May 2021
JEL No. H11,I38,J71,J78,K14,K42,R41

ABSTRACT

African-American motorists may adjust their driving in response to increased scrutiny by police. In daylight, when their race is more easily observable, minority motorists are the only group less likely to have fatal motor vehicle accidents. In Massachusetts and Tennessee, we find that African-Americans are the only group of stopped motorists with slower speeds in daylight. Consistent with an illustrative model, these speed shifts are concentrated at higher percentiles of the distribution. Calibration of this model indicates this behavior creates substantial bias in conventional tests of discrimination that rely on changes in the odds a stopped motorist is a minority.

Jesse Kalinowski
Department of Economics
Quinnipiac University
275 Mt Carmel Ave
Hamden, CT 06518
jesse.kalinowski@uconn.edu

Matthew B. Ross
Claremont Graduate University
Claremont, CA 91711
mbross.econ@gmail.com

Stephen L. Ross
Department of Economics
University of Connecticut
341 Mansfield Road, Unit 1063
Storrs, CT 06269-1063
and NBER
stephen.l.ross@uconn.edu

An online appendix is available at: http://www.nber.org/data-appendix/w28789

1. Introduction

The possibility that police treat minority motorists differently than other groups has become a source of protest and social unrest.[1] The public's most frequent interaction with police is through motor vehicle enforcement, which can serve as the precipitating event for more serious actions like searches, arrests, or use-of-force.[2] Most studies of police vehicle and pedestrian stops use the composition of stops as their primary statistic comparing it to some benchmark such as community composition (Ramirez et al. 2000), not-at-fault accidents (Alpert et al. 2003), stop rates of other officers (Robbins et al. 2017; Ridgeway and MacDonald 2009), or the composition at night when race cannot be observed (Grogger and Ridgeway 2006). However, tests based on the racial composition of traffic stops will be biased away from finding discrimination if motorists expect and respond to police discrimination by changing their driving behavior and thus reducing their likelihood of being stopped.

Intuitively, if some motorists with weak preferences for infractions are indifferent between committing an infraction and obeying traffic laws, an increase in discrimination will cause marginal minority motorists to stop committing infractions. This reduces the minority share of motorists that are risk of being stopped for a traffic infraction. Therefore, while discrimination directly increases the fraction minority among stopped motorists, the response of motorists to discrimination moves the fraction minority in the opposite direction creating an ambiguity in the overall effect. As we show in the mathematical appendix, the effect of motorists choosing to not infract (extensive margin) can dominate the increased likelihood of being stopped (intensive margin). Similar effects would arise if minority motorists reduced their number of trips in daylight due to concern about discrimination in police stops.

[1] See Arthur et al. (2017) and Nix et al. (2017) for recent media coverage on police shootings.

[2] 23 states have mandated the collection and analysis of traffic stop data for assessing racial differences in police stops. Also see policy initiatives like Obama's Task Force on 21st Century Policing as well as funding made available via the National Highway Safety Traffic Authority (NHTSA). See NHTSA SAFETEA-LU and Fast Act S. 1906 funding for FY 2006 to 2019.

Such behavioral responses to discrimination are relevant whenever examining police actions such as a stop, arrest, or use of force where the population at risk is not observed. Traffic stop analyses are a natural example, typically documenting the share minority among stops, but not observing the fraction of white and minority drivers stopped as a share of those at risk of being stopped.[3] This concern also applies to recent studies of police shootings by Goff et al. (2015), Swaine et al. (2015), Dolven et al. (2017) and in several of the analyses in Fryer (2019) because as noted by Fryer (2019) these studies do not observe a representative sample of individuals in situations that place them at risk of a police shooting. Even examining police actions for individuals arrested for the same crime or stopped for the same offense, measures of discrimination may be biased if use of force decisions are based on subsequent, unobserved behaviors and minorities are less likely to commit such behaviors based on an expectation of police discrimination. This bias does not arise if one observes a representative sample of individuals who exhibit the behaviors that may solicit a police response.[4]

This paper provides empirical evidence that minority motorists change their behavior in response to real or perceived police discrimination. In order to provide this evidence, we identify circumstances where comparable motorist populations are exposed to different levels of discrimination. In particular, we develop a counterfactual using a popular approach developed by Grogger and Ridgeway (2006), the "Veil of Darkness" (VOD). The VOD leverages seasonal variation in daylight to compare

[3] Exceptions exist where researchers observe a representative sample of motorists (Lamberth 1994; Lange et al. 2001; McConnell and Scheidegger 2004; Montgomery County MD 2002), but such approaches are considered prohibitively expensive (Kowalski and Lundman 2007; p. 168; Fridell et al. 2001, p. 22).

[4] For example, police search decisions (Knowles et al. 2001; Dharmapala and Ross 2004; Anwar and Fang 2006; Antonovics and Knight 2009; Marx 2018; Gelbach 2018) where the population of stopped motorists represent those at risk of search. Also, in the case of police shootings, Fryer's (2019) analysis in Houston focuses on arrest codes like assault on officer, resisting arrest or interfering in arrest that likely represents the behavior that was associated with police use of force. Arnold, Dobbie and Yang (2018) study racial disparities in bail in a sample of pre-trial arraignments. However, Knox, Lowe and Mummolo (2019) raise concerns about relying on administrative data, such as the Houston data collected by Fryer (2019), when analyzing such decisions.

stops made at the same time of day and day of week where some stops are made in daylight and others in darkness. The VOD operates under the premise that motorist race is less easily observed by police after sunset, but that the distribution of motorists committing infractions at a given time of day is unaffected by lighting conditions. The analysis controls for differences in preferences between majority and minority motorists by conditioning on the minority share of stops in darkness, which in principle could not have arisen from police discrimination. With over 22 applications across the country, VOD has become the gold standard for assessing racial differences in police traffic stops.[5]

For our first analysis, we examine the racial composition of fatal traffic accidents using lower accident rates as an indicator of safer driving. This approach intuitively follows Alpert, Smith, and Dunham (2003) who use racial composition of not-at-fault accidents as a benchmark for examining racial composition of traffic stops.[6] The U.S. National Highway Traffic Safety Authority's Fatality Analysis Reporting System (FARS) contains race/ethnicity and information on the circumstances surrounding all automobile accidents that result in one or more fatalities. We estimate models that are similar to VOD tests by regressing motorist race on whether the fatal accident occurred in daylight/darkness conditional on time of day, day of week, and year by location fixed effects. Consistent with minority motorists driving more carefully during daylight because they expect to face more scrutiny by police, traffic fatalities are 1.5 percentage points less likely to involve an African-American motorist in daylight relative to a share of 13 percent in the overall sample. For comparison, DeAngelo and Hansen (2014) show that a 33 to 37 percent reduction

[5] Applications include Grogger and Ridgeway (2006) in Oakland, CA; Ridgeway (2009) Cincinnati, OH; Ritter (2017) in Minneapolis, MN; Worden et al. (2012) as well as Horace and Rohlin (2016) in Syracuse, NY; Renauer et al. (2009) in Portland, OR; Taniguchi et al. (2016a, 2016b, 2016c, 2016d) in Durham Greensboro, Raleigh, and Fayetteville, North Carolina; Masher (2016) in New Orleans, LA; Chanin et al. (2016) in San Diego, CA; Ross et al. (2019, 2017) in Connecticut and Rhode Island; Criminal Justice Policy Research Institute (2017) in Corvallis PD, OR; Milyo (2017) in Columbia, MO; Smith et al. (2017) in San Jose, CA; and Wallace et al. (2017) in Maricopa, AZ.

[6] We thank Jesse Shapiro for this suggestion.

in state police manpower in traffic enforcement lead to a 12 to 14 percent increase in highway traffic fatalities. We also find that the daylight effect is largest in states with greater racial disparities in police shootings and in those that rank highly on a Google trends racism index. Other than race, we do not observe differences in accident rates between daylight and darkness across any other motorist or vehicle attributes.

In our second empirical analysis, we examine data on police speeding stops. We focus on speeding stops because the motorist's speed provides a convenient variable for assessing infraction severity under the premise that minority motorists might respond to discrimination by reducing the severity of infractions committed.[7] However, motorist behavioral responses will have an ambiguous influence on the distribution of infractions over stopped motorists. Specifically, if an increase in discrimination/daylight causes minority offenders who commit minor infractions to obey the speed limit, these "nighttime only" offenders will no longer be part of the distribution of motorists stopped for speeding shifting the speed distribution up, in the opposite direction of the direct response of motorists driving more slowly in daylight to reduce their likelihood of stop. Therefore, any change in speed represents a rejection of the null of no behavioral response. Further, we additionally focus our test on changes in speed at the top of the speeding distribution following the intuition that dropping motorists from the bottom of the distribution has minimal effect on the percentile ranking of motorists near the very top of the distribution.[8]

We conduct our analysis using data from Massachusetts and Tennessee. To our knowledge, these samples are the only statewide data available with information on the speed of traffic stops resulting in a warning, rather than only for tickets/fines, which is important because police will observe race after the stop and prior to issuing

[7] Darkness may also affect traffic stops for non-moving violations, like cell phone use or equipment failures (Grogger and Ridgeway 2006; Kalinowski, Ross and Ross 2019a). Researchers might use statutorily established fines to measure severity for a broader set of moving violations.

[8] In the mathematical appendix, we demonstrate that discrimination will increase infraction severity of minority motorists near the top of the distribution of stopped motorists, as long as the tails of the distribution of motorist preferences are not too heavy.

either a ticket or warning whether or not the stop is in daylight or darkness.[9] We first conduct traditional VOD tests using the racial composition of speeding stops. We find that daylight stops are more likely to be of African-American motorists than darkness stops in Massachusetts (4.5 percentage points) and West Tennessee (1 percentage point). We observe no differences in East Tennessee.[10]

We then calculate the speed of motorists relative to the speed limit and measure the speed distribution among stopped motorists in both daylight and darkness. First, we find no effect of daylight on the speed distribution of white motorists. Consistent with the intuition above, we find no effect of daylight on stopped minority motorist speeds at the 10^{th} and 20^{th} percentiles for Massachusetts and West Tennessee, and only a 1 to 1.5 percentage point decline in speeds for East Tennessee in daylight. However, the negative shift in the minority motorist speed distribution increases in magnitude at higher percentiles. Massachusetts has a decrease of speed in daylight of 11 to 12 percentage points at the 80^{th} and 90^{th} percentile and East Tennessee has a decrease of 3 percentage points at the 70^{th} percentile. These results are consistent with motorists changing their driving behavior in response to daylight when race is visible. In West Tennessee, however, the maximum shift in the speed distribution is less than one percentage point.

The much larger shift in Massachusetts appears reasonable given the much higher rates of minority motorist stops in the Massachusetts data, four and half times those in West Tennessee. The next largest shift in speed occurs in East Tennessee. Notably, this finding occurs even though the VOD test revealed no evidence of racial discrimination in stops for East Tennessee. East Tennessee is consistent with the change in motorist behavior in daylight having dominated racial differences in police stop behavior, preventing the VOD test from detecting discrimination. Further, we find no evidence of speed distribution shifts for white motorists between daylight and

[9] In Tennessee, the data explicitly identifies warnings and tickets. In Massachusetts, many speeding tickets have zero fine, which we interpret as somewhat equivalent to a warning.

[10] Tennessee is divided at the time zone boundary removing counties on the boundary.

darkness or over available motorist and vehicle attributes. Only African-American motorists appear to change their driving behavior in response to changes in visibility.

Finally, in order to quantify the magnitude of these effects, we calibrate our model of motorist infractions and police stops to the speed distribution of stopped motorists and the African-American share of stops in daylight and darkness.[11] The calibrated models match the empirical moments well. Most significantly, the calibration for East Tennessee is able to match both the shift in the speed distribution of stopped African-American motorists between daylight and darkness and the VOD test statistic that is near one, which is typically interpreted as evidence of equal treatment. In East Tennessee, the daylight police stop cost for African-American motorists is substantially below the darkness stop costs suggesting a preference for stopping African-American motorists (discrimination), and the daylight decrease in stop cost is similar to the increase in officer pay-off arising from a two standard deviation increase in relative speed. These results arise even though the VOD test statistic appeared consistent with equal treatment. The calibrated racial differences in stop costs for Massachusetts are very large implying pay-off differences similar to a five standard deviation increase in the speed. In West Tennessee, the small shift in the speed distribution implies much smaller racial differences in stop costs equivalent to only one-half of a standard deviation change in speed. The large stop costs differences in Massachusetts imply a 50 percentage point increase in share of minority motorists not infracting between daylight and darkness, while the more modest differences in East Tennessee imply a smaller 15 percentage point increase in the share.[12]

Minority responses to adverse treatment are documented in several contexts including labor markets, health, and the arts (see Arcidiacono, Bayer and Hizmo 2010;

[11] We calibrate to aggregate moments, more common in macroeconomics, rather than estimating the structural model using micro data due to the large computational requirements.

[12] We also simulate the model while forcing motorist behavior to remain unchanged in daylight, which implies an increase in the VOD test statistic from 1.00 to 1.22 in East Tennessee, a noticeably smaller increase of 1.09 to 1.17 in West Tennessee, and a very large increase of 1.38 to 2.74 in Massachusetts.

Institute of Medicine 2003; and Goldin and Rouse 2000). However, work on behavioral responses to police discrimination is rare. The literature has documented decreasing criminal behavior as police enforcement rises, and to the extent that discrimination can be interpreted as increased scrutiny by police, our paper also contributes to that literature. Finally, the empirical evidence in this study contributes to the literature examining racial differences in the legal system including police stops (Grogger and Ridgeway 2006; Ridgeway 2009, Horrace and Rohlin 2016, Ritter 2017, and Kalinowski, Ross and Ross 2019b), fines (Goncalves and Mello 2017, 2018), searches (Knowles, Persico, and Todd 2001; Dharmapala and Ross 2004; Anwar and Fang 2006; Antonovics and Knight 2009; Marx 2018), use-of-force (Fryer 2019; Knox, Lowe and Mummolo 2019), bail (Ayres and Waldfogel 1994; Arnold, Dobbie and Yang 2018) and jury trials (Anwar, Bayer, and Hjalmarsson 2012; Flanagan 2018).

2. Implications of Motorists Behavioral Adjustments

In this section, we present police and motorist objective functions and describe some key results from our theoretical appendix. First, the motorist infractions and police stop problems below imply that some motorists with weak preferences for committing infractions choose to obey the traffic laws. As a result, rationale motorists will respond to a higher likelihood of police stops by exiting the population of motorists at risk of stop due to a traffic violation. In the appendix, we demonstrate that when motorists behave in this way the share of stopped motorists who are minority can actually increase as discrimination against those minority motorists increases. Similarly, while infraction severity, e.g. speeding, may appear to be a natural strategy for detecting the responses of minority motorists to discrimination in police stops, data on infractions is almost always restricted to individuals who were stopped for the infraction. At low infraction levels, minority motorists may stop committing infractions when discrimination increases, and the distribution of infraction severity among stopped motorists could then increase with discrimination even if minority motorists as a whole are driving more safely or slowly.

The model proposed here differs from traditional models of police search, such as Knowles et al. (1999), because officers observe the violation in most motorist stops, while searches are made of stopped motorists based on the likelihood that an individual is carrying contraband. As Knowles et al. (1999) note, motorists will have an incentive to always break the law if they face no chance of being stopped or punished, but will never break the law if they know that they will be punished with certainty. To create uncertainty in stops when the infraction is observable, we extend the standard police cost function to include a stochastic term that represents factors around the context or circumstance of the infraction. These factors may include whether a police officer observed the infraction, differences in officer stop thresholds, impact of pending stop quotas, rainy or cold weather, the officer's current mood, and whether the officer is busy with other activities.

Specifically, the officer's decision to stop a motorist $\gamma(i, d, \phi)$ is made after observing a non-negative infraction severity i (e.g. speed above the limit) that yields a pay-off from stop of $u(i)$, motorist type/demography d, and circumstances surrounding the stop ϕ. The officer's utility maximization problem takes the form

$$\max_{\gamma(i,s_d,\phi)} [u(i) - h(\phi) - s_d]\gamma(i, s_d, \phi) \quad (1)$$

where we define s_d as a fixed component of stop costs associated with a motorist type while $h(\phi)$ represents circumstantial costs.[13]

As a result, the solution to the officer's problem implies an optimal infraction threshold above which the officer makes a stop with certainty and below which the officer does not make a stop. Given the officer's net utility of $u(i) - h(\phi) - s_d \ \forall \ i$, the solution to her utility maximization problem is simply

$$\gamma(i, s_d, \phi) = \begin{cases} 1, & \text{if } u(i) > h(\phi) + s_d \\ 0, & \text{otherwise.} \end{cases}$$

[13] The model appears to assume uniform stop costs across officers, but if motorists receive a random draw of officer when observed committing an infraction then officer specific heterogeneity in stop costs can be included in the circumstances term ϕ and all results will continue to hold.

Solving for the infraction level with zero net pay-off implies a threshold severity of

$$i^*(\phi, s_d) = u^{-1}(h(\phi) + s_d) \tag{2}$$

Conditional on infraction severity and stop costs, we can solve Equation (2) for the circumstances $\phi^*(i, s_d)$ when net pay-off is zero by exploiting the monotonicity of $h(\phi)$.

$$\phi^*(i, s_d) = h^{-1}(u(i) - s_d) \tag{3}$$

Further, we can assume that ϕ is distributed uniform without loss of generality if minimal restrictions are placed on h, and then Equation (3) represents the unconditional (i.e. circumstances not observed) probability that an officer stops a motorist with infraction level i.

The motorist problem can be characterized as a trade-off between the benefit of committing an infraction $b(i, c)$, which depends on motorist preferences c, e.g. recklessness, criminality, stress, timing of trip, sleep deprivation, etc. and the expected cost of being stopped, or

$$\max_{i(c, s_d)} b(i, c) - \tau(i)\phi^*(i, s_d) \tag{4}$$

where the cost of being stopped for committing an infraction is $\tau(i)$ and the probability of being stopped is $\phi^*(i, s_d)$.

The resulting first order condition is

$$\mathrm{FOC} \equiv \frac{\partial \mathrm{b}(\mathrm{i}, \mathrm{c})}{\partial \mathrm{i}} - \frac{\mathrm{d}\tau(\mathrm{i})}{\mathrm{di}} \phi^*(\mathrm{i}, \mathrm{s_d}) - \tau(\mathrm{i}) \frac{\partial \phi^*(\mathrm{i}, \mathrm{s_d})}{\partial \mathrm{i}} = 0 \tag{5}$$

Because both costs and the probability of stop depend upon infraction severity, we impose a number of technical assumptions that are detailed in the appendix to assure that conditional on c motorists select a unique, optimal finite infraction severity i^*>0.

In this model, discrimination arises if police officers have lower demographic cost of stopping a minority *(m)* relative to the majority *(w)*, $s_m < s_w$. A standard statistic for evaluating racial discrimination in stops is the relative share of stops involving minority motorists. However, when motorist infraction severity is a choice, the standard statistic for racial discrimination in police stops depends upon the distribution of motorists over preferences $g(c, d)$.

Definition 1. $K_g \equiv \frac{p[m|stopped, s_m, g(c,m)]}{p[w|stopped, s_w, g(c,w)]} = \frac{\int_{c^*(s_m)}^{\infty} g(c,m)\phi^*(i'(c,s_d),s_d)\, di}{\int_{c^*(s_w)}^{\infty} g(c,w)\phi^*(i'(c,s_d),s_d)\, di}$

A decrease in stop cost operates through two effects on K_g: 1. a change in the probability of stop for motorist's who were infracting and 2. an increase in the threshold at which motorists begin to commit infraction.

The purpose of this model is to allow us to examine whether the behavioral adjustments of motorists can reverse the standard assumption in tests for discrimination in police stops that decreases in minority motorist stop costs lead to a higher share of minorities among stopped motorists. To capture the effect of increasing discrimination, we examine the impact of decreasing $\mathrm{s_m}$ on $\mathrm{K_g}$.

$$\frac{dK_g}{ds_m} = \frac{1}{p[w|stopped, s_w, g(c,w)]}\left(-\left(\frac{dc^*}{ds_m}\right)_{c=c^*} g(c^*, m)\tilde{\phi}(c^*, s_m) + \int_{c^*}^{\infty} g(c,m)\frac{d\tilde{\phi}}{ds_m}\, dc\right)$$

where $\tilde{\phi}(c, s_m) = \phi^*(i'(c, s_m), s_m)$. The first term in parentheses captures the effect of marginal minority motorists choosing not to commit infractions when discrimination increases and these exits from the population at risk drive down the fraction of minority stopped. The first term can dominate the second if either the density of motorists at c^* or the change in c^* with stop cost is large enough to counteract changes in stop probabilities, and the proof in the appendix creates a counterexample. Notably, this counterexample is constructed without imposing any restrictions on the preferences of majority motorists, and while holding the mean of preferences for minority motorists fixed.[14]

The critical feature of our model that leads to a reversal of the standard assumption is that some motorists choose not to commit infractions if their

[14] While the model implies that the marginal offenders are those who commit the least serious infractions, the ambiguity in the VOD test will arise if any minority motorists exit the population of those committing infractions in daylight, even if for example they exit this population by not driving.

preferences for committing an infraction are sufficiently weak. Therefore, we must acknowledge several consequential assumptions on the police and motorists problems that were required to obtain this result. Specifically, we assume that police receive a pay-off and motorists face a cost for any positive infraction level and that the pay-off and cost are both bounded away from zero as the infraction level approaches zero. Then, the stochastic component of stop costs assures that motorists always face a positive expected cost of committing an infraction even for minimal infraction levels. We also assume that benefits from committing an infraction increase smoothly with the infraction level as motorists shift from zero infraction to positive levels.

Our second key result in the theoretical appendix relates to how the distribution of infractions of stopped minority motorists changes as discrimination increases. As with the statistic in Definition 1, the distribution of infractions will experience two conflicting effects: 1. Holding preferences fixed minority motorists will commit less severe infractions as discrimination increases, and 2. Low preference minority motorists will change from committing relatively non-severe infractions to committing no infractions shifting the distribution of motorists at risk of stop towards higher levels of infractions.

When some low infraction level motorists, say x percent of those committing infractions, are dropped from the bottom of the distribution, the motorists just above these dropped motorists in terms of preferences have their percentile in the distribution moved from just above x to close to the bottom of the distribution, or a movement in percentile ranking of approximately x. On the other hand, the motorist with the strongest preferences for committing infraction at the top of the infraction severity distribution among motorists committing infractions has their percentile of 100% unaffected by this change. Therefore, at the very top of the distribution of motorists committing infractions, the only effect on speed is the direct effect of the reduction in infraction level due to higher discrimination, and so the effect on infraction severity at the very top is unambiguous.

To account for a continuous distribution of c and focus on stopped motorists, we characterize the observed infraction severity distribution at a specific percentile x in the speed distribution of stopped motorists. Conditional on s_m and motorist preference $c \geq c^*(s_m)$, we write a stopped motorist percentile by integrating over the product of the pdf of c and the equilibrium probability of stop, or

$$x(c, s_m) = \frac{\int_{c^*(s_m)}^{c} g\,(c')\tilde{\phi}(c, s_m)\,dc'}{\int_{c^*(s_m)}^{\infty} g\,(c')\tilde{\phi}(c, s_m)\,dc'}$$

where the numerator captures the mass of stopped motorists below c and the denominator captures all stopped motorists. Similarly, we can pick a percentile x and write the preference parameter of that motorist as an implicit function $c_x(x, s_m)$. Finally, we define the equilibrium infraction level of stopped motorists at each percentile by substituting c_x into i'.

Definition 3. $i_x(x, s_m) \equiv i'(c_x(x, s_m), s_m)$

Differentiation of $\mathrm{i_x(x, s_m)}$ in Definition 3 yields

$$\frac{\mathrm{di_x}}{\mathrm{ds_m}} = \frac{\mathrm{di'}}{\mathrm{ds_m}} + \frac{\mathrm{di'}}{\mathrm{dc}}\frac{\mathrm{dc_x}}{\mathrm{ds_m}}$$

which again illustrates direct effect on infraction severity fixing c^* and the indirect effect because the preference parameter at a given percentile changes with stop costs. Signing this relationship as x limits one or at the top of the distribution requires one additional assumption. Namely that the right tail of the minority preference distribution is not too heavy. Notably, this result will hold distributions that have substantially heavier tails than the normal distribution.

3. Evidence from Accident Data

Exploiting the logic of the Veil of Darkness (VOD) logic, we examine a national sample of traffic accidents for evidence of whether minority motorists adjust their driving behavior in response to lighting conditions, possibly driving more conservatively and safely in daylight when race can be observed. Unlike the data on police stops, accident data provides evidence on the driving behavior of minority motorists where the racial composition is not directly affected by the composition of police stops. Therefore, we believe that the patterns uncovered in the accident data can be attributed to changes in motorist driving behavior, presumably in response to actual or perceived discrimination.

Our sample is drawn from the National Highway Traffic Safety Authority's Fatality Analysis Reporting System (FARS) data, which documents all automobile accidents in the United States involving one or more fatalities. This dataset documents the race and ethnicity of fatalities, and we restrict our sample to accidents where the motorist died and were either an African-American or a Non-Hispanic white motorist. The overall sample consists of 282,924 motorists from a total of 615,826 accidents involving a fatality that occurred in the contiguous United States from 2000 to 2017.[15] We do not exploit information on whether a motorist was designated as at-fault because police at-fault designation may be influenced by motorist race (West 2018). To our knowledge, there is no comparable national dataset containing non-fatal accidents.

Following Grogger and Ridgeway (2006), we further limit our sample to 39,076 fatalities where the accident occurred within a window of time between the earliest and latest sunset of the year based on the location of the fatality. This time period is referred to as the inter-twilight window (ITW) and is selected so that stops during a particular time of day are only included in the sample if that time of day has variation in daylight across the seasons of the year. Changes in the daylight variable occur due to both seasonal variation in sunset and the discrete spring/fall daylight savings time (DST) changes. We identify accidents occurring within the ITW based on data from the United States Naval Observatory (USNO) denoting the bounds of the

[15] Observations are weighted by the inverse number of fatalities per accident. For instance, when both drivers from a two-car accident die, we give each of those fatalities a weight of one-half.

ITW using the eastern and westernmost coordinates of each county where the accident occurred. The lower bound of the county-specific window is the earliest annual easternmost sunset and the upper bound is the latest westernmost end to civil twilight. Unlike many VOD studies of traffic stops, the FARS data also contains detailed reporting on the lighting conditions when an accident occurred. We use this self-reported measure rather than estimates of daylight based on USNO data to minimize measurement error in visibility. In Appendix Table B2, we present comparable results using USNO definitions and results are robust.[16] For a more thorough discussion of measurement error in VOD daylight measures, see Kalinowski et al (2019).

Appendix Table B1 presents descriptive statistics with column 1 showing the means for the entire ITW sample, column 2 for the sample of accidents involving fatalities of African-American motorists and column 3 for the sample of white motorist fatalities. The African-American population is more male, older, drives newer vehicles, more likely to drive imported vehicles, and more likely to be involved in accidents that occur on weekends and in darkness.

We follow the logic of the VOD test placing race (R_i) on the left-hand side of the equation and testing whether accidents occurring in daylight ($\overline{v}_i$) are more likely to involve African-American motorists using a linear probability model, but similar results arise regressing daylight fatality on race. We condition on day of the week (d) and hourly time of the day (t) fixed effects to assure that the model is identified by comparing stops where the composition of the drivers is expected to be the same. The resulting estimation equation is

$$R_{idt} = \beta\overline{v}_{idt} + \delta_d + \gamma_t + \varepsilon_{idt} \tag{6}$$

where δ_d is the vector of day of the week fixed effects and γ_t contains the time of the day fixed effects. We also add state and year or state by year fixed effects. Since many models involve high dimensional fixed effects, we estimate linear probability models

[16] As is standard, we disregard stops occurring each day during actual twilight when visibility is somewhere between daylight and darkness.

rather than logistic regression as used in Grogger and Ridgeway (2006). Kalinowski et al. (2019) demonstrate the equivalence of the linear probability and logistic regression tests in Grogger and Ridgeway (2006).[17] Standard errors are clustered at the state level in columns 1 and 2, but at the state by year level when the model includes state by year fixed effects.

Panel 1 of Table 1 reports the results from estimating Equation (6) using our sample of fatal accidents. Column 1 presents estimates for a model containing the controls in Equation (6) plus state and year fixed effects, while column 2 presents estimates for models that contain state by year fixed effects. Column 3 presents estimates after adding controls for motorist and vehicle attributes including motorist age and gender and vehicle age and whether the vehicle was an import. The estimates imply that the likelihood of a fatal accident involving an African-American decreases by 1.5 to 1.6 percentage points in daylight, relative to a mean of 12.8%. Lower fatality rates of African-Americans in daylight are consistent with African-American motorist driving more conservatively in daylight when race can be observed. The magnitude of these effects appears comparable to a moderate increase in traffic enforcement. For comparison, DeAngelo and Hansen (2014) examines the effect of a large budget cut to Oregon state police. The authors show that the cuts lead to a 33 to 37 percent reduction in the number of state police allocated to traffic enforcement and estimates that, as a result, highway traffic fatalities increased by between 12 and 14 percent.

The behavior of minority motorists is also likely to be shaped by their perceptions of police behavior. Panels 2 and 3 present estimates based on interacting daylight with one of two different measures that might capture African-American perceptions about police treatment of minority motorists. The first proxy is the odds

[17] Starting with Equation (6) in Grogger and Ridgeway (2006), they set the second term to zero (in the equation prior to taking the log) based on the assumption that motorist composition does not change between daylight and darkness. Then, one can replace the conditional probabilities for a representative motorist with the predicted probabilities arising from a linear probability model. For positive β in Equation (10) above, the test statistic is greater than one consistent with discrimination, and the statistic increases with increases in β.

that an unarmed individual involved in a police shooting in a given state is African-American divided by the fraction of state residents who are African-American, where the values range from 0.04 (odds of 1.04) in Connecticut to 16.76 in Rhode Island..[18] The second proxy is a measure of real and perceived racism constructed using Google Trends data in a similar manner as Stephens-Davidowitz (2014).[19] The index that google trends produces is between 0 and 100, but has been standardized and so ranges from –2.16 (index of 48.6) in Montana to 2.38 (index of 89) in Maryland. Both variables are cross-sectional characterizing states over the period from 2004 to 2020. The proxy for the perception of discrimination is positively associated with the reduction in the share of fatal accidents involving African-Americans in daylight relative to darkness. A doubling of the black-white odds of police shooting from even odds to odds of 2 to 1 implies an increase in racial differences associated with daylight fatalities of 0.4 percentage points, while a one standard deviation increase in the racism index implies a 1.2 percentage point increase in differences.

Next, in Panel 4, we restrict our FARS sample to the 9 states where the VOD test has been conducted on police traffic stops and either failed to find or found mixed evidence of discrimination.[20] We find even larger racial differences in this subsample. Daylight motorist fatalities are over 3 percentage points more likely to involve African-American motorists relative to a dependent mean of 13.2%, as compared to 1.5 percentage points relative to a mean of 12.8 for the entire sample. While these fatality differences do not imply discrimination in police stops, the data is suggestive that

[18] Police shootings data comes from Mesic et al. (2018). However, findings are robust to shootings ratios from Fatal Encounters (https://fatalencounters.org/) or Mapping Violence (https://mappingpoliceviolence.org/).

[19] Stephens-Davidowitz (2014) uses the frequency of searches for racial slurs to capture the sentiment of whites about minorities. In our case, we are interested in. the sentiment of minorities in terms of real or perceived discrimination, particularly by police. Thus, we construct an index using Google Trends from 2004-20 searching for the following words: police shooting, discrimination, racial profiling, prejudice, racism, and police complaint. Similar results arise using an index by Mesic et al. (2018) based on residential segregation, incarceration rates, and disparities in education and employment status.

[20] The states are Arizona, California, Connecticut, Louisiana, Missouri, North Carolina, Ohio, Oregon, and Rhode Island. For convenience and to maintain a reasonably sized sample, we do not restrict our accident sample to the exact same time periods of VOD traffic stop studies in these states.

minority motorists are concerned about such stops, potentially affecting previous tests for discrimination.[21]

Lastly, we address the concern that the overall composition of motorists might change in response to daylight. Formal tests of balance are wholly absent in existing applications of the VOD test because traffic stop data alone cannot be used to disentangle changes in enforcement activity from compositional changes in traffic patterns. In our accident data, however, we can reasonably expect that police traffic stop differences between daylight and darkness did not directly affect the composition of white motorists experiencing traffic fatalities or their vehicle attributes. We examine the composition of white non-Hispanic motorists involved in fatal accidents in Table 2. Columns 1-4 present models where daylight is regressed on whether the vehicle is domestic rather than import, the age of the vehicle in years, whether the motorist was male and whether the motorist was under the age of 30. Column 5 presents a model that includes all four of the motorist and vehicle attributes available. All models included hour of day, day of week and state by year fixed effects. The composition of fatal accidents for non-Hispanic white motorists does not vary between daylight and darkness for these variables. No t-statistics are significant, and in the full model the F-statistic associated with the four estimates is 1.37 ($p=0.24$). Thus, motorist race appears to be the only motorist or vehicle characteristic available for which differences in fatality rates correlate with daylight.[22]

In this section, we present evidence that minority motorists are involved in accidents at a lower rate during periods of daylight relative to equivalent periods of

[21] We cluster standard errors by state by year due to the small number of states. This decision is conservative empirically in that clustering at the state level yields smaller standard errors..

[22] Motorists might differ in their selection into the sample of fatalities. We have detailed data on all accidents involving a fatality, but only race and ethnicity for the fatalities themselves. Therefore, we also estimate inverse probability weighted models based on the likelihood that the motorist dies during a fatal traffic accident using vehicle attributes and information on restraints, i.e. airbags and seatbelt usage. The results presented above are robust to selection on these observables (Appendix Table B3). We do not include controls for airbags and seatbelt use in the models above because those controls may be endogenous to motorist risk-taking behavior.

darkness. These changes in minority accident rates are larger in states with more police shootings and where there is a higher perception of racism. Further, these responses are especially large in states where VOD analyses of traffic stops have failed to find evidence of discrimination. This evidence is supportive of a view that African-American motorists realize that their race can be identified by police in daylight, and so choose to drive more conservatively and carefully during daylight hours. We also found that the accidents rates of non-Hispanic white motorists are invariant to changes in visibility across several motorist and vehicle characteristics, suggesting that this responsiveness to daylight is a phenomenon that is primarily about race.

4. Evidence from Traffic Stop Data

In this section, we present the results from an analysis of police traffic stops. Following previous studies, we focus on a subsample of stops made for moving violations, in our case speeding, since other violations (e.g. headlights, seatbelt, and cellphones) are possibly correlated with both visibility and race. Our focus on speeding stops also has the added advantage of providing a clear measure of infraction severity that we can use to assess changes in motorist driving behavior, i.e. speed relative to the speed limit. We analyze speeding stops in Massachusetts from April 2001 to January 2003 made by either the State Police or large municipal police departments in Massachusetts and by Tennessee State Police from 2006 to 2015.[23] As noted above, we selected these two states because the stop records contain information on the speed traveled for stops in which a warning was issued.[24] In Massachusetts, we observe stops by local and state police. To focus on stop populations containing a reasonable number of African-Americans, we restrict our analysis to state police stops and stops made by town police departments of the 10 largest towns.[25] In Tennessee, we make a distinction between

---

[23] Massachusetts data was collected by Bill Dedman for the Boston Globe and used by Antonovics and Knight (2009) to study police searches. Tennessee data is from the Stanford Open Policing Project.

[24] In Tennessee, warnings are explicitly included in the data. In Massachusetts, there are a large number of traffic stops with zero-dollar fines listed which we believe represent warnings.

[25] These towns include Boston, Worcester, Springfield, Lowell, Cambridge, Brockton, New Bedford, Quincy, Lynn, and Newton of which Newton is the smallest with a population of under 90,000.

patrol districts lying on the Eastern and Western side of the time zone border that bisects the state.[26] As before, we select only traffic stops that occur within the Inter-Twilight window (ITW) which we bound between the earliest recorded Easternmost sunset and latest Westernmost end to civil twilight in each county.[27]

Appendix Table B4 presents descriptive statistics for the ITW speeding stop samples, excluding actual twilight. The Massachusetts sample numbered 10,203 speeding stops, while samples in East and West Tennessee, respectively, contain 23,515 and 102,054 stops. In Massachusetts, speeding stops were more likely to involve African-American motorists in daylight, for female drivers, for imported vehicles, and on Saturdays. In Tennessee, weekend stops were more likely to be African-Americans, but stops of males were less likely to be African-Americans in east Tennessee and more likely to be African-Americans in west Tennessee. Stopped African-American motorists tend to travel at lower speeds in terms of miles per hour over the speed limit, but in Massachusetts African-Americans are stopped for more severe infractions as measured by percent over the speed limit likely due to their stops being concentrated in more urban locations with lower speed limits.

Table 3 presents the VOD model estimates for all three samples of speeding violations. The model follows Equation (6) from the traffic fatality data except that the geographic fixed effects are within state. To control for geography, we use town and state police barracks fixed effects because counties are quite large relative to the size of Massachusetts. In Tennessee, models include county fixed effects because

Restrictions based on omitting towns with African-American shares below the state average yields a similar sample of towns and similar results. Smaller towns in Massachusetts tend to be more rural and have very few African-American residents.

[26] We exclude three rural patrol districts (of eight) lying adjacent to or on top of the time zone boundary. A significant portion of those traffic stops occur on opposing sides of the time zone from the patrol district's headquarters creating ambiguity about the time of the stop. Estimates using the entire sample are less precise, but quantitatively similar to, estimates when these patrol districts are excluded.

[27] The ITW occurred in Massachusetts between 4:09 PM and 9:08 PM while in Tennessee it falls within 5:15 PM and 9:48 PM. The Massachusetts traffic stop data only contains the hour of the day that the stop was made. Any hour where a portion occurs within the ITW is included, but only when those stops that occurred in an hour that completely excluded twilight, complete daylight or darkness were included. The twilight sample excluded is similar on observables to the ITW sample.

counties are small in size relative to state police patrol districts. Standard errors are clustered at the town/state police barracks level for Massachusetts, and at the county by year level for Tennessee.[28] Columns 1, 3 and 5 present estimates for models that include time of day, day of week, geographic and in Tennessee year fixed effects. Columns 2, 4 and 6 present estimates adding the available motorist and vehicle controls that include whether the motorist is male or female and whether the vehicle is domestic or import for both states plus whether the driver is under the age of 30, whether the vehicle is older than 5 years and whether the vehicle is red for Massachusetts. Estimates for east and west Tennessee are very similar including county by year fixed effects, and all results are similar if we regress daylight on race instead.

In Massachusetts and West Tennessee, we find evidence suggesting that the odds that stops involves a minority motorist increases in daylight relative to darkness. A daylight stop in Massachusetts is approximately 4.5 percentage points more likely to involve an African-American motorist, while in west Tennessee daylight stops are 1 percentage point more likely to involve African-Americans. The magnitudes of these estimates are stable as we add controls for motorist and vehicle attributes and as we add county by year fixed effects for Tennessee. However, we find no evidence of differences in East Tennessee. The classic interpretation of these results is that Massachusetts and West Tennessee show evidence of discriminatory policing, but that East Tennessee does not. Appendix Table B5 presents similar estimates using the logistic regression as in Grogger and Ridgeway.

Next, we explore our motivating hypothesis that the speed of stopped minority motorists decreases in daylight in response to real or perceived discrimination at higher percentiles of the speed distribution. We calculate a relative speed based on both our intuition that the same absolute speed limit violation will be more concerning to police

[28] We could cluster by county for the Tennessee models in columns 3, 4, 6 and 7 where we do not include county by year FE's. However, East Tennessee only contains 13 counties. We have confirmed that the standard errors based on clustering at the county level are smaller for East Tennessee than those clustered at the county by year level and very similar in magnitude in West Tennessee.

when speed limits are low and the empirical fact that fine schedules in both states apply more severe penalties for the same absolute speed violation at lower speed limits. Specifically, we define S_{idt} as $speed/speed\ limit$. We then estimate marginal effects at each decile using unconditional quantile regressions following Firpo, Fortin, and Lemieux (2009) and using a software package described in Borgen (2016). For our problem, conditional quantile regression would be inappropriate because our goal is to detect changes at the top of the actual speed distribution, as opposed to estimating effects for individuals whose unobservables fall at different points in the distribution.

The estimation follows a three-step procedure where we (1) construct a transformed speed variable using kernel density estimation, (2) define the re-centered influence function (RIF) variable for each quantile in the transformed distribution, and (3) use RIF as the outcome in a linear model to obtain the quantile estimates (Firpo, Fortin, and Lemieux 2009). We kernel smooth speeds to obtain an estimated density at discrete points in the distribution.

$$\widehat{f_K}(S_i) = \sum_{j=1}^{n} K\left(\frac{S_i - S_j}{h}\right)$$

The bandwidth parameter h is selected following a standard procedure that minimizes the mean integrated squared error under a Guassian Kernal if the data is Gaussian.[29] The results are robust to a variety of alternative functional forms for K, but is specified as Epanechnikov in our estimates. We estimate the relative speed and density at each numeric decile τ of the distribution, and then calculate the Recentered Influence Function (RIF) for each decile in the kernel smoothed speeding data within the inter-twilight sample as follows

$$RIF\left(S_i\colon q_\tau, F_{\widehat{spd}}\right) = q_\tau + \frac{\tau - \mathbb{I}\{S_i \le q_\tau\}}{f_{spd}(q_\tau)}$$

[29] The precise calculation is $h = (9m/10n)^{1/5}$ where $m = min\left(\sqrt{var(S)}, IQR(S)/1.349\right)$ and IQR is the interquartile range.

where q_τ and f_{spd} are the estimated speed and density at decile τ, and $\mathbb{I}$ is an indicator function. Using the decile RIF's for each i observation, we estimate changes in the speeding distribution using linear models for the RIF at each decile.

$$RIF_{\tau,idt} = \beta_{\tau,0} + \beta_{\tau,1} R_{idt} + \beta_{\tau,2} \overline{v}_{idt} + \beta_{\tau,3} (R_{idt} \cdot \overline{v}_{idt}) + \delta_{\tau,d} + \gamma_{\tau,t} + \varepsilon_{\tau,idt} \quad (7)$$

where the variable R_{idt} is a dichotomous indicator variable equal to unity when the motorist was of African American descent and $\overline{v}_{idt}$ is a binary variable indicating the presence of the daylight during the traffic stop. The parameter of interest $\beta_{\tau,3}$ is the coefficient on the interaction of these two variables, which captures racial heterogeneity in speed distribution shift. As above, we add geographic fixed effects, and for the Tennessee samples we also include year or county by year fixed-effects.

Table 4 presents the results from applying Equation (7) to the same sample of speeding stops used for the VOD estimates in Table 3. We find evidence of slower speeds in daylight for African American motorists, but as suggested by our model the speed distribution shift in all three sites arises primarily for the higher percentiles. In Massachusetts, the shift is quite large starting near zero at the 10th percentile and rising to over 10 percentage points at the 80th and 90th percentiles, relative to the sample average 41 percent over the speed limit. The next largest speed distribution shift is in East Tennessee starting around 1 percentage point at the 10th percentile and reaching a maximum of 3 percentage points at the 70th percentile, relative to an average of 35 percent over the limit. The shift in West Tennessee is smaller starting at zero and reaching a maximum below 1 percentage point at the higher percentiles. Notably, the coefficients on daylight are always insignificant consistent with no shift in the speed distribution of non-Hispanic white motorists.

The quantile regressions yield multiple estimates raising concerns about multiple hypothesis testing. We follow Bifulco et al. (2008) conducting a simulation exercise to assess the likelihood that the pattern of results arose by chance. Bifulco et al. (2008) exploit the logic of a Fisher's exact permutation test in a resampling framework 1) ordering the t-statistics arising from the coefficients for each quantile by magnitude, 2) drawing 10,000 bootstrap samples of the same size as the original sample with replacement under the null of no correlation between speed and daylight (randomizing daylight), 3) re-estimating the quantile model for and ordering the t-statistics from each bootstrap sample, and 4) calculating the fraction of bootstrap samples where the set of ordered t-statistics dominate the actual set of t-statistics. While the t-tests above are two-sided, this permutation test is one-sided where a vector of signed and ordered bootstrap t-statistics lies below the actual signed and ordered t-statistics if all elements of the bootstrap vector have a lower value than the corresponding elements of the actual vector. We strongly reject the null hypothesis of no negative shift for all three sites. In Massachusetts, the likelihood of this pattern arising by chance is 0.013 percent. In East and West Tennessee, the likelihoods are 0.005 and 0.001, respectively. We also re-estimate these models adding the motorist and vehicle controls, and in Tennessee adding county by year fixed effects (Appendix Table B6). The addition of motorist and vehicle controls has no impact. The county by year fixed effects erode the speed shift in East Tennessee somewhat by between 15 and 20 percent smaller, but remains significant with a 0.04 likelihood of a type 1 error.

Next, as we did for the fatality analysis, we examine the speed distribution for non-Hispanic White motorists over other factors. In both Massachusetts and Tennessee, we observe whether the motorist is male and whether the vehicle is either a domestic or imported vehicle. We re-estimate the models in Table 4 replacing race in Equation (11) with either motorist male or whether domestic vehicle. Repeating our bootstrap analysis, we find that the likelihood that these results could have arisen by chance was 0.89 for Massachusetts, 0.59 for East Tennessee and 0.37 for West Tennessee for gender; and 85.7 percent for Massachusetts, 72.3 percent for East Tennessee, and 91.1 percent for West Tennessee for vehicle type, see Appendix Table B7.[30] For Massachusetts, we also conduct these analyses for whether driver is younger

[30] We follow the same permutation strategy except that the test is two-sided using the absolute value of the t-statistics because we have no priors concerning how these attributes might shift the speed distribution. For Tennessee, negative findings are robust to including county by year fixed effects

than 30, vehicle is older than 5 years and whether vehicle is red. As above, we find no evidence of a change in speeds with daylight, see Appendix Table B8.

In this section, we present evidence on the speed distribution of stopped motorists. African-American motorists in the upper half of the speed distribution travel more slowly in daylight, when presumably race is observed. The largest differences in the speed distribution arise in the Massachusetts sample where we also observed the largest composition differences between daylight and darkness stops. In Tennessee, we observed that the largest shift in the speed distribution of stopped African-American motorists arose in East Tennessee where the VOD tests did not identify any evidence of discrimination, consistent with behavioral changes potentially confounding the VOD test. Further, we find no evidence of speed distribution shifts for whites or shifts over other motorist or vehicle attributes.

5. Calibration and Simulation

To help quantify the magnitude of these effects, we calibrate our model to the data on stopped motorists from Massachusetts and East and West Tennessee to calculate racial differences in police stop costs in daylight and darkness. We also use the darkness police stop costs to calculate counterfactual VOD test statistics that would have arisen if African-American motorists did not respond to increased scrutiny by police in daylight by driving more slowly. The details on the parameterization of the model and the optimization procedure please see Appendix C.[31]

To calibrate the model, we calculate six speed percentiles (20^{th}, 40^{th}, 60^{th}, 80^{th}, 90^{th}, and 95^{th}) in miles per hour over the speed limit for each combination of daylight/darkness and minority/non-minority, the fraction of motorists stopped during daylight who are minority, and the fraction of motorists stopped in darkness who are minority. Beyond the quintiles, we add moments for the 90^{th} and 95^{th}

[31] We conduct a macro-style calibration using aggregate moments, rather than structural estimation using micro data. This decision is based on computational demands given that each calibration takes several weeks to run. Therefore, we rely on the quantile regressions above for inference.

percentiles to help capture the skewed nature of the speed distribution. To assure that the speed moments are comparable to the estimations above, we remove the time of day, day of week and geographic fixed effects in our relative speed model and add the sample means back to the residuals yielding motorists with effectively common observables. Finally, to assure common circumstances, we convert these relative speeds back to miles per hour using the mode speed limit in each sample.[32]

Table 5 summarizes the impact of race on police stop behavior in the calibration. The first row presents the minority stop cost in daylight, which is 0.006 in Massachusetts, 30.1 in East Tennessee, and 37.8 in West Tennessee all in comparison to a darkness stop cost of 44.0. White stop costs in daylight are all near the darkness stop cost, consistent with the quantile regression estimates that showed no change in the speed distribution in daylight for white motorists. Consistent with previous studies of the Massachusetts data and the large shift in the speed distribution, we find evidence of high levels of police prejudice in Massachusetts, i.e. a daylight stop cost far below the darkness stop cost. We observe higher levels of prejudice (lower stop costs) for East Tennessee than West Tennessee, based on the shift in the speed distribution in East Tennessee, even though the VOD test statistic for East Tennessee was near 1.0.

Further, we can use the calibrated parameters for police stop costs and $u(i)$ to compare the lower minority stop costs in daylight to the police pay-offs that arise from stopping a motorist whose speeding infraction is more severe. The next three rows show the change in return to a police stop if the speed of the motorist increases by ½, 2 or 5 standard deviations relative to the mean level of infractions among stopped motorists. Specifically, we find the mean μ and standard deviation σ of the number of miles per hour over the speed limit within the simulation for motorists committing infractions, and calculate $(\mu + \alpha\sigma)^{\eta} - (\mu)^{\eta}$ where α takes on the values of ½, 2 and 5 and η is the exponent parameter in $u(i)$. Daylight raises the effective

[32] The results and parameters of the calibration are shown in Appendix Tables C2 to C4.

net returns to stopping minority motorists in Massachusetts by more than the effect of raising speed by five standard deviations above the mean. In East Tennessee, daylight raises the return to stopping minority motorists by an amount comparable to a 2 standard deviation increase in speed, but in West Tennessee where the speed distribution shift is smaller daylight raises the return by ½ a standard deviation.

The second panel of Table 5 presents the VOD test statistic from the calibration and a counterfactual VOD statistic that would arise if minority motorists did not change their infraction behavior in daylight, i.e. behaved in daylight as if they faced the police costs for stops in darkness. Following Grogger and Ridgeway (2006), the VOD test statistic is

Definition 4. $K_{VOD} \equiv \frac{p[m|stopped,\overline{\upsilon}]}{p[w|stopped,\overline{\upsilon}]}\frac{p[w|stopped,\underline{\upsilon}]}{p[m|stopped,\underline{\upsilon}]}$

We can calculate the alternative statistic K_{ADJ} by calculating the above probabilities in K_{VOD} except c^* and i' in daylight $\overline{\upsilon}$ are assumed to depend on the darkness $\underline{\upsilon}$ police stop cost.

Definition 5. $K_{ADJ} \equiv \frac{\int_{c^*(s_{\underline{\upsilon}})}^{c_h} g(c,m)\phi^*\left(i'\left(c,s_{\underline{\upsilon}}\right),s_{\overline{\upsilon},m}\right)di}{\int_{c^*(s_{\underline{\upsilon}})}^{c_h} g(c,w)\phi^*\left(i'\left(c,s_{\underline{\upsilon}}\right),s_{\overline{\upsilon},w}\right)di}\frac{\int_{c^*(s_{\underline{\upsilon}})}^{c_h} g(c,w)\phi^*\left(i'\left(c,s_{\underline{\upsilon}}\right),s_{\underline{\upsilon}}\right)di}{\int_{c^*(s_{\underline{\upsilon}})}^{c_h} g(c,m)\phi^*\left(i'\left(c,s_{\underline{\upsilon}}\right),s_{\underline{\upsilon}}\right)di}$

The counterfactual VOD test statistic increases most in Massachusetts from 1.38 to 2.74, next most in East Tennessee from 1.00 to 1.22, and by the smallest amount in West Tennessee to 1.17 from the calibrated value of 1.09. The results in Table 5 are repeated for alternative weights in Appendix Table C5.

6. Conclusion

Utilizing the Veil of Darkness approach developed by Grogger and Ridgeway (2006), we document empirical evidence of behavioral changes using both national data on traffic fatalities and data on traffic stops from the states of Massachusetts and Tennessee. Using the national accident fatality data, we find that the likelihood of a motorist fatality being an African-American as opposed to a white motorist decreases by about 1.5 percentage points in daylight. In the traffic stop data, we find a large shift in the speed distribution of African-Americans between daylight and darkness near the top of the distribution for Massachusetts, 7 to 12 percent slower in daylight relative to the speed limit. We find a smaller, but sizable, shift for East Tennessee, 1.5 to 3 percent slower, but very little shift in West Tennessee, one percent slower or less. We do not observe similar changes in fatalities or speeding over any observable motorist or vehicle characteristics, nor do we observe such changes in speeding for white motorists. Our empirical findings from both the accident and traffic stop samples provide strong evidence suggesting that minority motorists are adjusting their driving behavior during daylight, when their race is observable to police, in response to real or perceived police discrimination.

We calibrate our theoretical model of police stop and motorist infraction behavior to the samples of traffic stop data. The model matches the empirical moments well including capturing the fact that the observed decreases in the infraction level of African-Americans in daylight is largest at the highest percentiles of the speed distribution. The calibrated differences in police stop costs for minority motorists between daylight and darkness is very large in Massachusetts, equivalent to the return to police from increasing motorist speed above the speed limit by 5 standard deviations relative to the mean. These large differences are consistent with both the high VOD test statistic and the large speed distribution shift. On the other hand, the VOD test statistic in East Tennessee is near one and yet we observe substantially lower calibrated police stop costs for minority motorists in daylight, equivalent to an increase in motorist speed of 2 standard deviations. The failure of the VOD test statistic to detect discrimination in East Tennessee appears attributable to a substantial shift in the minority speed distribution between daylight and darkness.

In summary, we have established that the commonly used statistic, the fraction of stopped motorists who are minorities, can yield biased results due to the behavioral responses of minority motorists. We have also provided empirical evidence of such responses based on the daylight-darkness difference in the minority share of fatal traffic accidents and shifts between daylight and darkness in the speed distribution of stopped minority motorists. Our findings have direct implications for tests of police discrimination in traffic stops relying on exogenous shocks to visibility (Grogger and Rdigeway 2006), as well as benchmarking approaches based on Census data (Ramirez et al. 2000), not-at-fault accidents (Alpert et al. 2003), or synthetic controls using stops by other officers (Robbins et al. 2017; Ridgeway and MacDonald 2009). More generally, our model predictions are relevant whenever police engage in some action where the specific population at risk is not observed and the police observe the offense or illegal behavior that could result in a police action, e.g. stop, arrest, or use-of-force.

References

Alpert, G. P., Smith, M. R., & Dunham, R. G. (2003). Toward a Better Benchmark: Assessing the Utility of Not-at-Fault Traffic Crash Data in Racial Profiling Research. *Confronting Racial Profiling in the 21st Century: Implications for Racial Justice, Boston.*

Antonovics, K., & Knight, B. G. (2009). A new Look at racial profiling: Evidence from the Boston Police Department. *Review of Economics and Statistics*.

Anwar, S., Bayer, P., & Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *Quarterly Journal of Economics*.

Anwar, S., & Fang, H. (2006). An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review*.

Arcidiacono, P., Bayer, P., & Hizmo, A. (2010). Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics*.

Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*.

Arthur, R., Dolven, T., Hamilton, K., McCann, A., & Sherman, C. (2017, December 11). Shot by cops and forgotten. *VICE News*. Retrieved from news.vice.com/en us/article/xwvv3a/shot-by-cops

Ayres, I., & Waldfogel, J. (1994). A Market Test for Race Discrimination in Bail Setting Recommended Citation "A Market Test for Race Discrimination in Bail Setting" *Stanford Law Review*.

Bjerk, D. (2007). Racial Profiling, Statistical Discrimination and the Effect of a Color Blind Policy on the Crime Rate. *Journal of Public Economic Theory* 9(3), 521-545.

Bifulco, R., Fletcher, J. M., & Ross, S. L. (2011). The Effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health. *American Economic Journal: Economics Policy*.

Borgen, N. T. (2018). Fixed Effects in Unconditional Quantile Regression. *The Stata Journal*.

Chalfin, A. & McCrary, J. (2017). Criminal Deterrence: A Review of the Literature. *Journal of Economic Literature*. Vol. 55 (1), 5–48.

Chanin, J., Welsh, M., Nurge, D., & Henry, S. (2016). *Traffic enforcement in San Diego, California An analysis of SDPD vehicle stops in 2014 and 2015*.

Criminal Justice Policy Research Institute. (2017). *Law Enforcement Contacts Policy and Data Review Committee: 2016 Annual Report*. https://www.pdx.edu/cjpri/sites/www.pdx.edu.cjpri/files/LECC_2016_Annual_Report_FINAL 011917_0.pdf

C. H. da Silva Santos, M. S. Gonçalves, and H. E. Hernandez-Figueroa, "Designing Novel Photonic Devices by Bio-Inspired Computing," *IEEE Photonics Technology Letters* **22** (15), pp. 1177–1179 (2010).

DeAngelo, G., & Hansen, B. (2014). Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities. American Economic Journal: Economic Policy 6(2): 231–257.

Dharmapala, D., Ross, S., Dharmapala, D., & Ross, S. (2004). Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence. *The B.E. Journal of Economic Analysis & Policy*, *3*(1), 1–23.

Evans, W. N. & Owens, E. G. (2007). COPS and Crime. Journal of Public Economics. Vol. 91 (1), 181–201.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). UNCONDITIONAL QUANTILE REGRESSIONS. *Econometrica*, *77*(3), 953–973.

Flanagan, F. X. (n.d.). *Race, Gender, and Juries: Evidence from North Carolina **.

Fridell, L., Robert, L., Diamond, D., & Kubu, B. (2001). *Racially Biased Policing: A Principled Response*. Washington, DC.

Fryer, R. G. (2019). An Empirical Analysis of Racial Differences in Police Use-of-force. *Journal of Political Economy*.

Gelbach, J. B. (2018). *Search Distaste at the Highway Stop Level* (Vol. 23). Retrieved from https://openpolicing.stanford.edu/data/.

Gehrsitz, M. (2017). Speeding, Punishment, and Recidivism: Evidence from a Regression Discontinuity Design. *Journal of Law and Economics* 60(3), 497-528.

Goff, P. A., Kindy, K., Fisher, M., Tate, J., & Jenkins, J. (2015, December 26). A Year of Reckoning: Police Fatally Shoot Nearly 1,000. *The Washington Post.*
www.washingtonpost. com/sf/investigative/2015/12/26/a-yearof-reckoning-police-fatally-shoot-nearly1000/

Goldin, C. & Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review*. American Economic Association, vol. 90(4), pages 715-741, September.

Goncalves, F., & Mello, S. (2018). *A Few Bad Apples? Racial Bias in Policing*. http://www.princeton.edu/~fmg/JMP

Goncalves, F., & Mello, S. (2017). *Does the Punishment Fit the Crime? Speeding Fines and Recidivism*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064406

Grogger, J., & Ridgeway, G. (2006). Testing for racial profiling in traffic stops from behind a veil of darkness. *Journal of the American Statistical Association*.

Hansen, B. (2017). Punishment and Deterrence: Evidence from Drunk Driving. *American Economic Review* 105(4), 1581–1617.

Horrace, W. C., & Rohlin, S. M. (2016). How dark is dark?: Bright lights, big city, racial profiling. *Review of Economics and Statistics*.

Institute of Medicine (2003). Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: *The National Academies Press*.

Kalinowski, J. J., Ross, M. B., & Ross, S. L. (2019). Now You See Me, Now You Don't: The Geography of Police Stops. *AEA Papers and Proceedings*.

Kalinowski, J., Ross, M., & Ross, S. L. (2019). Addressing Seasonality in Veil of Darkness Tests for Discrimination: An Instrumental Variables Approach. *Working Papers*.

Kevin Lang, B., & Manove, M. (2011). Education and Labor Market Discrimination. *American Economic Review*, *101*, 1467–1496.

Knowles, J., Persico, N., & Todd, P. (2001). Racial Bias in Motor Vehicles: Theory and Evidence. *Journal of Politcal Economy*.

Knox, D., Lowe, W., & Mummolo, J. (2019). The Bias Is Built In: How Administrative Records Mask Racially Biased Policing. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3336338

Kowalski, B. R., & Lundman, R. J. (2007). Vehicle stops by police for driving while Black: Common problems and some tentative solutions. *Journal of Criminal Justice*.

Lamberth, J. (1994). *Revised Statistical Analysis of the Incidence of Police Stops and Arrests of Black Drivers/Travelers on the New Jersey Turnpike Between 2 Exits or Interchanges 1 and 3 from the Years 1988 through 1991*.

Lange, J. E., Blackman, K. O., & Johnson, M. B. (2001). *Speed violation survey of the New Jersey Turnpike: Final Report*. Calverton, Md. : Public Services Research Institute.

Levitt, S. (1997). Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime. *American Economic Review*. Vol. 87, 270–290.

Marx, P. (2018). *An Absolute Test of Racial Prejudice*. Working Paper.

Masher, J. (2016). What The "Veil of Darkness" Says About New Orleans Traffic Stops. https://nolacrimenews.com/2016/09/08/what-the-veil-of-darkness-says-about-new-orleans-traffic-stops.

McConnell, E. H., & Scheidegger, A. R. (2004). Race and Speeding Citations: Comparing Speeding Citations Issued by Air Traffic Officers With Those Issued by Ground Traffic Officers. *Paper Presented at the Annual Meeting of the Academy of Criminal Justice Sciences April, 4–8*.

Mello, S. (2019). More COPS, Less Crime. *Journal of Public Economics*. Vol. 172, 174-200.

Mesic, A., Franklin, L., Cansever, A., Potter, F., Sharma, A., Knopov, A., & Siegel, M. (2018). The Relationship Between Structural Racism and Black-White Disparities in Fatal Police Shootings at the State Level. *Journal of the National Medical Association*.

Milyo, J. (n.d.). *DWB in COMO: Understanding Race Disparities in Vehicle Stops and Searches*. https://www.columbiamissourian.com/news/local/traffic-stop-report-shows-racial-disparities-as-officials-

Montgomery County, M. (2002). *Traffic Stop Data Collection Analysis: Third Report*.

Nix, J., Campbell, B. A., Byers, E. H., & Alpert, G. P. (2017). A Bird's Eye View of Civilians Killed by Police in 2015: Further Evidence of Implicit Bias. *Criminology and Public Policy*.

Persico, N. (2009). Racial Profiling? Detecting Bias Using Statistical Evidence. *Annual Review of Economics*, 1, pp. 229-54.

Persico, N., & Todd, P. (2006). *Generalizing the Hit Rates Test for Racial Bias in Law Enforcement, With an Application to Vehicle Searches in Wichita*.

Powell, M. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives*. https://doi.org/Technical Report NA2009/06

Ramirez, D., J. McDevitt, and A. Farrell. (2000). A Resource Guide on Racial Profiling Data Collection Systems: Promising Practices and Lessons Learned. Washington, D.C.: U.S. Department of Justice.

Renauer, B. C., Henning, K., & Covelli, E. (2009). *Portland Police Bureau. Report*.

Ridgeway, G. (2009). *Cincinnati Police Department Traffic Stops: Applying RAND's framework to Analyze Racial Disparities*.

Ridgeway, Greg and John MacDonald. 2009. "Doubly Robust Internal Benchmarking and False Discovery Rates for Detecting Racial Bias in Police Stops." Journal of the American Statistical Association, Vol. 104, No. 486

Ritter, J. A. (2017). How do police use race in traffic stops and searches? Tests based on observability of race. *Journal of Economic Behavior and Organization*.

Robbins, M., Saunders, J., Kilmer, B. A Framework for Synthetic Control Methods With High Dimensional, Micro-Level Data: Evaluating a Neighborhood-Specific Crime Intervention. Pages. Journal of the American Statistical Association. May 2017

Ross, S. L. (1997). Mortgage lending discrimination and racial differences in loan default: A simulation approach. *Journal of Housing Research* 8, 277-298.

Ross, S. L. (1996). Mortgage lending discrimination and racial differences in loan default. *Journal of Housing Research* 7, 117-126.

Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2019). *State of Connecticut Traffic Stop Data Analysis and Findings, 2018*.

Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2017). *State of Rhode Island Traffic Stop Data Analysis and Findings, 2016*.

Ross, S. L. and Yinger, J. (2002). *Color of Credit: Mortgage Discrimination, Research Methods, and Fair Lending Enforcement*. MIT Press: Cambridge, MA.

Rowan, T. (1990). *Functional Stability Analysis of Numerical Algorithms*. University of Texas at Austin.

Smith, A. C. (2016). Spring forward at your own risk: Daylight saving time and fatal vehicle crashes. *American Economic Journal: Applied Economics*.

Smith, M. R., Rojek, J., Tillyer, R., & Lloyd, C. (2007). *San Jose Police Department Traffic and Pedestrian Stop Study*.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics*, Elsevier, vol. 118(C), pages 26-40.

Swaine, Jon, Oliver Laughland, Jamiles Larteyand, and Ciara McCarthy. 2015. "Young Black Men killed by US Police at Highest Rate in Year of 1,134 Deaths." The Guardian, December 31.

Taniguchi, T., Hendrix, J., Aagaard, B., Strom, K., Levin-Rector, A., & Zimmer, S. (2016). *A test of racial disproportionality in traffic stops conducted by the Fayetteville Police Department.*

Taniguchi, T., Hendrix, J., Aagaard, B., Strom, K., Levin-Rector, A., & Zimmer, S. (2016). *Department, A test of racial disproportionality in traffic stops conducted by the Greensboro Police.*

Taniguchi, T., Hendrix, J., Aagaard, B., Strom, K., Levin-Rector, A., & Zimmer, S. (2016). *A test of racial disproportionality in traffic stops conducted by the Raleigh Police Department.*

Taniguchi, T., Hendrix, J., Aagaard, B., Strom, K., Levin-Rector, A., & Zimmer, S. (2016). *Exploring racial disproportionality in traffic stops conducted by the Durham Police Department.*

Wallace, D., Louton, B., Nelson, J., Gricius, M., Tyler, D. H., Coble, S., Kramer, K. (2017). *Yearly Report for the Maricopa County Sheriff's Office: Years 2015-16.*

West, J. (2018). Racial Bias in Police Investigations. University of California, Santa Cruz Working Paper.

Worden, R. E., McLean, S. J., & Wheeler, A. P. (2012). Testing for Racial Profiling with the Veil-of-Darkness Method. *Police Quarterly.*

Table 1: Estimated Change in the Accidents Rate for Minority Motorists in Daylight

LHS: African-American		(1)	(2)	(3)	(4)
Baseline					
Daylight		-0.01752*** (0.00412)	-0.01663*** (0.00392)	-0.01566*** (0.00399)	-0.01525*** (0.00398)
Observations		39076	39076	39076	39076
Interaction – Black-White Police Shootings Odds Ratio					
Daylight x Police Shootings		-0.00193 (0.00150)	-0.00356** (0.00150)	-0.00415*** (0.00159)	-0.00429*** (0.00158)
Observations		39063	39063	39063	39063
Interaction – Google Search Racism Index					
Daylight x Racism Index		-0.00886** (0.00364)	-0.01169*** (0.00345)	-0.01131*** (0.00358)	-0.01182*** (0.00355)
Observations		39063	39063	39063	39063
VOD Inconclusive States					
Daylight		-0.04642*** (0.01217)	-0.03559*** (0.01085)	-0.03324*** (0.01080)	-0.03381*** (0.01071)
Observations		6587	6587	6587	6587
Controls	Hour of Day	X	X	X	X
	Day of Week	X	X	X	X
	Year	X	X		
	State		X		
	State x Year			X	X
	Motorist/Vehicle				X

Notes: Coefficient estimates are presented where * represents a p-value .1, ** represents a p-value .05, and *** represents a p-value .01 level of significance. Standard errors are clustered at the state by year level. The sample includes only fatal accidents involving African-American or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017 involving at least one or more non-commercial automobiles (no motorcycle or pedestrian). Observations are weighted by the inverse number of observations per accident included within the sample. Panel 2 adds an interaction between daylight and the odds that an unarmed individual involved in a police shooting in a given state is African-American divided by the fraction of residents in the state who are African-American. Panel 3 adds an interaction between daylight and a statewide, standardized google trends index using the terms: "police shooting", "discrimination", "racial profiling", "prejudice", "racism", and "police complaint". Panel 4 repeats panel 1 for the subsample of states where the VOD test was conducted and results were inconclusive: Arizona, California, Connecticut, Louisiana, Missouri, North Carolina, Ohio, Oregon, and Rhode Island.

Table 2: Balancing Test of Accidents for White Motorists within the ITW

LHS: Daylight		(1)	(2)	(3)	(4)	(5)
Domestic Vehicle		0.00428 (0.00510)				0.00487 (0.00512)
Vehicle Age			-0.00589 (0.00547)			-0.00575 (0.00548)
Male Motorist				-0.00629 (0.00490)		-0.00668 (0.00491)
Young Motorist					0.00572 (0.00470)	0.00571 (0.00471)
Controls	Hour of Day	X	X	X	X	X
	Day of Week	X	X	X	X	X
	State x Year	X	X	X	X	X
R^2		0.35243	0.35243	0.35245	0.35244	0.35252
Observations		34050	34050	34050	34050	34050

Notes: Coefficient estimates are presented where * represents a p-value .1, ** represents a p-value .05, and *** represents a p-value .01 level of significance. Standard errors are clustered at the state by year level but robust to clustering on just state or year. The sample includes only fatal accidents involving Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017 involving at least one or more non-commercial automobiles (no motorcycle or pedestrian). Observations are weighted by the inverse number of observations per accident included within the sample. Results are robust to restricting the sample to not-at-fault accidents as well as weighting the fatal accidents based on the likelihood of experiencing a fatality, estimated using detailed vehicular characteristics and restraint use. The F-statistic for the main variables of interest in specification five is 1.4 and a p-value of 77.82 percent.

Table 3: Canonical Veil of Darkness Estimates

LHS: African-American		(1)	(2)	(3)	(4)	(6)	(7)
		MA		East TN			West TN
Daylight		0.0458** (0.0185)	0.0441** (0.0193)	-0.00116 (0.00397)	-0.000921 (0.00395)	0.0105*** (0.00384)	0.00972** (0.00382)
Controls	Day of Week	X	X	X	X	X	X
	Time of Day	X	X	X	X	X	X
	County (or Town)	X	X	X	X	X	X
	Year			X	X	X	X
	Motorist/Vehicle		X		X		X
Observations		10203	10203	23515	23515	102054	102054

Notes: Coefficient estimates are presented where * represents a p-value .1, ** represents a p-value .05, and *** represents a p-value .01 level of significance. Standard errors are clustered on county by year in East and West Tennessee (TN) and town or state highway patrol districts in Massachusetts (MA), but robust in Tennessee to clustering on county and year separately and robust in Massachusetts to clustering by town. The sample includes only traffic stops for speeding violations involving African-American or Non-Hispanic white motorists. The models using the Tennessee samples also include controls for year in the first two specifications of each panel and county by year fixed effects in the last.

Table 4: Estimated Change in Speed Distribution for Stopped Minority Motorists in Daylight

LHS: Rel. Speed		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		10 pct	20 pct	30 pct	40 pct	50 pct	60 pct	70 pct	80 pct	90 pct
MA	Daylight	0.00519	0.114	1.847	0.628	-0.415	-0.120	0.449	-0.749	-1.372
		(1.092)	(1.140)	(1.221)	(0.904)	(1.130)	(1.247)	(1.351)	(2.177)	(2.819)
	African-American	0.664	0.548	2.551**	2.187***	1.551**	1.728*	1.477	5.959***	6.514**
		(1.029)	(0.993)	(1.202)	(0.720)	(0.685)	(0.850)	(1.672)	(1.816)	(2.946)
	Daylight*African-American	-0.273	-0.213	-1.718	-2.228**	-5.032**	-6.839**	-7.783***	-10.99***	-12.24**
		(1.298)	(1.286)	(1.376)	(1.004)	(1.946)	(2.585)	(2.682)	(2.803)	(4.239)
	Obs.	10203	10203	10203	10203	10203	10203	10203	10203	10203
East TN	Daylight	-0.200	0.00734	-0.186	-0.123	-0.0979	-0.0470	0.181	0.210	-0.113
		(1.094)	(0.806)	(0.471)	(0.351)	(0.336)	(0.411)	(0.565)	(0.835)	(1.116)
	African-American	-2.116**	-1.861*	-1.254*	-0.909	-0.795	-1.039	-0.178	-1.029	-1.732
		(0.763)	(0.903)	(0.688)	(0.804)	(0.824)	(1.063)	(1.402)	(1.827)	(2.132)
	Daylight*African-American	-1.158	-1.384**	-1.070**	-0.965**	-1.419***	-1.560**	-3.069**	-2.232	-2.117
		(0.842)	(0.629)	(0.465)	(0.365)	(0.440)	(0.589)	(1.084)	(1.542)	(2.248)
	Obs.	23515	23515	23515	23515	23515	23515	23515	23515	23515
West TN	Daylight	0.0879	0.174	-0.0740	-0.137	-0.205*	-0.0470	0.00219	-0.168	-0.176
		(0.120)	(0.212)	(0.109)	(0.140)	(0.110)	(0.183)	(0.250)	(0.341)	(0.472)
	African-American	0.182	0.606**	0.676**	0.671*	0.296	0.664	0.671	0.546	0.170
		(0.249)	(0.272)	(0.259)	(0.378)	(0.344)	(0.428)	(0.514)	(0.607)	(0.723)
	Daylight*African-American	-0.102	-0.176	-0.545***	-0.867***	-0.536***	-0.843***	-0.996***	-0.802	-0.948
		(0.151)	(0.202)	(0.172)	(0.258)	(0.188)	(0.243)	(0.328)	(0.511)	(0.668)
	Obs.	102054	102054	102054	102054	102054	102054	102054	102054	102054

Notes: Coefficient estimates are presented such that * represents a p-value .1, ** represents a p-value .05, and *** represents a p-value .01 level of significance. Standard errors are clustered on county by year in East and West Tennessee (TN) and patrol districts in Massachusetts (MA). The sample includes only traffic stops for speeding violations involving African-American or Non-Hispanic white motorists. Controls include time of day, day of week, and geographic location fixed-effects. The two Tennessee samples also include controls for year. Relative speed is calculated as speed relative to the speed limit and multiplied by one hundred.

Table 5: Calibration Results Related to Racial Differences in Police Stop Behavior

	Massachusetts	East Tennessee	West Tennessee
Police Return and Cost of Stops			
Minority Stop Cost Diff	43.994	13.887	6.247
Return to Increase in Speed			
0.5 SD Increase			6.405
2.0 SD Increase		13.002	
5.0 SD Increase	42.940		
VOD Test Statistics			
Simulated VOD Test	1.379	0.997	1.090
Adjusted VOD Test	2.736	1.223	1.173

Notes The minority stop cost difference is calculated by subtracting the calibrated stop cost for minorities in daylight from the darkness stop cost of 44. The return to a specific number $\boldsymbol{\alpha}$ of standard deviations $\boldsymbol{\sigma}$ increase in miles per hour over the speed limit is calculated relative to the mean speeding violation $\boldsymbol{\mu}$ by $(\boldsymbol{\mu} + \boldsymbol{\alpha\sigma})^{\boldsymbol{\eta}} - (\boldsymbol{\mu_i})^{\boldsymbol{\eta}}$ using the calibrated parameters and the simulated speed distribution for each site. Finally, the simulated VOD test statistics is the statistic implied by the simulated speed distributions based on the calibrated parameters, and the adjusted VOD test statistic is calculated using the darkness minority speed distribution for daylight stops, but having police stop motorists based on their daylight stop costs.