HCEO WORKING PAPER SERIES

Endogenous Driving Behavior in Tests of Racial Profiling in Police Traffic Stops

Jesse Kalinowski
Stephen L. Ross
Matthew B. Ross

Working Paper

2017-017

02/2017; Revised 08/2019

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Endogenous Driving Behavior in Tests of Racial Profiling in Police Traffic Stops

August 30, 2019

Jesse J. Kalinowski[1], Matthew B. Ross[2], and Stephen L. Ross[3]

Keywords: Police, Crime, Discrimination, Racial Profiling, Disparate Treatment, Traffic Stops
JEL Codes: H1, I3, J7, K14, K42

Abstract: Policy evaluations of police traffic stops have increasingly used variation in the timing of sunset, "Veil of Darkness" (VOD) tests, based on the assumption that police are better able to detect race during daylight. Here, we propose that African-American motorists adjust their driving behavior in response to increased scrutiny by police during daylight when their race is more easily observed, potentially biasing estimates of discrimination. Using nationally representative data, we demonstrate that African-Americans are less likely to have fatal motor vehicle accidents when driving in daylight. These effects are largest in states with historically high levels of structural racism and that rank highest in terms of police shootings of unarmed African-Americans. We find no such daylight effects on fatal accidents across any other observed motorist or vehicle attributes. Using police traffic stop data, we also find that the speed distribution of stopped African-American motorists in Massachusetts and Tennessee is shifted towards slower speeds in daylight with no evidence of such shifts for whites or over other observables. Finally, we develop and calibrate a model of police stop and motorist speeding behavior and use this model to demonstrate the effects of changes in driving speed on the VOD test. Theoretically, we show that the VOD test statistic can be reversed by motorist responses to the presence of police prejudice, and our calibration model demonstrates that substantial bias is introduced into the VOD test statistic by these responses.

[1] Quinnipiac University, Department of Economics, Hamden, CT.
[2] New York University, Wagner School of Public Service, New York, NY.
[3] University of Connecticut, Department of Economics and Public Policy, Storrs, CT.

Corresponding Author: M.B. Ross, Wagner School of Public Service, New York University, 370 Jay St., 12th Floor; New York, NY; MBR1@NYU.edu

Acknowledgments: For insightful comments, we thank Talia Bar, Hanming Fang, Felipe Goncalves, Jeffrey Grogger, William Horrace, John MacDonald, Steve Mello, 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 Economcs 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.

1. Introduction

The possibility that police treat minority motorists differently than other groups has recently become a source of protest and social unrest in the United States.[4] The public's most frequent interaction with police is through motor vehicle enforcement, and such enforcement can serve as the precipitating event for more serious actions like searches, arrests or use-of-force. Over the last decade, many states have mandated the collection and analysis of traffic stop data for assessing racial differences in police stops.[5] Although these data tend to show that the proportion of traffic stops made of minority motorists is high and often far exceeds their share of the local population, attribution of these differences to discrimination by police is very challenging because the motorists at risk of being stopped are virtually never observed.[6] An approach developed by Grogger and Ridgeway (2006) called the "Veil of Darkness" (VOD) test leverages seasonal variation in daylight to compare stops made at the same time where some stops were in daylight and others in darkness based on the premise that motorist race cannot be easily identified after nightfall. With over 22 applications across the country, this approach has quickly become the gold standard for assessing racial differences in police stops.[7]

In this paper, we examine one of the key maintained assumptions of the VOD test: the behavior of drivers on a given roadway at a given time of day and day of the week is unaffected by changes in the timing of sunset, an assumption that seems especially reasonable for the evening

[4]In response to recent high profile, police shootings of unarmed minority men, the "Black Lives Matter" movement has emerged to challenge traditional approaches to policing. See Arthur et al. (2017), Goff et al. (2015), and Nix et al. (2017) for recent media coverage on race and police shootings.

[5] These efforts are directly linked to national policy initiatives like Obama's Task Force on 21st Century Policing as well as funding made available to states via the National Highway Safety Traffic Authority (NHTSA). Having received NHTSA funding, 23 states have implemented programs to collect and analyze traffic stop data for evidence of discrimination. See NHTSA SAFETEA-LU and Fast Act S. 1906 funding for FY 2006 to 2019.

[6] Some exceptions exist where researchers observe a representative sample of the motorists at risk of stop (Lamberth 1994; Lange et al. 2001; McConnell and Scheidegger 2004; Montgomery County MD 2002), but applying such approaches is considered to be prohibitively expensive (Kowalski and Lundman 2007; p. 168; Fridell et al. 2001, p. 22). Another common solution is to examine a secondary outcome such as vehicle searches where the general counterfactual, individuals involved in a police encounter, is observed (Knowles et al. 2001; Dharmapala and Ross 2004; Anwar and Fang 2006; Antonovics and Knight 2009; Marx 2018; Gelbach 2018). Also see Arnold, Dobbie and Yang (2018) and Fryer (2019) who examine bail among a population of those arrested and use-of-force among a population with police interactions, respectively. However, in these cases, the question of whether discrimination arose in the initial encounter, such as stop or arrest, is unanswered.

[7] Applications of the test include Grogger and Ridgeway (2006) in Oakland, CA; Ridgeway (2009) Cincinnati, OH; Ritter and Bael (2009) and Ritter (2017) in Minneapolis, MN; Worden et al. (2010; 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. (2015; 2016; 2017a; 2017b) 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.

commute. Even if the composition of motorists is constant between daylight and darkness, as has been argued by Grogger and Ridgeway (2006) and others, the VOD test still requires that the driving behavior of motorists on the road does not change. If minority motorists believe that they face discrimination by police and that their race can be more easily observed in daylight, they might rationally respond by driving more conservatively and committing fewer moving violations during daylight periods when they face increased scrutiny. Recognizing this possibility, we empirically test for evidence of behavioral changes by minority motorists in response to visibility using both accident data as well as data on speeding stops. We then use a simple calibrated model to examine the potential impact of such behavioral changes on the VOD test statistic.

While it might seem natural to focus immediately on police stops, we begin our analysis examining national data on accidents because police stop decisions may be influenced by the race of motorists. This makes it difficult to isolate the effects of motorist behavioral changes from the direct effect of discrimination in stops on the composition of stopped motorists.[8] Therefore, we turn to data from the U.S. National Highway Traffic Safety Authority's Fatality Analysis Reporting System (FARS) that both contains information on all automobile accidents that result in one or more fatalities and identifies the race and ethnicity for each accident victim. We then estimate traditional VOD tests using the white and black subsamples of motorists who died in motor vehicle accidents and regressing the race of the motorist on whether the accident was in daylight or darkness conditional on the time of day and day of week of the accident. We find a strong correlation between race and daylight with fatal accidents being less likely to involve African-American motorists, consistent with those motorists driving more carefully in daylight relative to darkness. Daylight fatalities are one and a half percentage points less likely to involve an African-American motorist relative to a base African-American share of fatalities of 13 percent. Further, the effects are largest in states that have a higher score on an index capturing historic racism and in states that have a greater racial disparity in the incidence of police shootings. Finally, we demonstrate that unlike motorist race the sample of fatal accidents exhibits balance between daylight and darkness in terms of motorist age and gender and vehicle age and whether a vehicle is domestic or imported.

Next, we examine data on police speeding stops. We select speeding stops as our outcome because the speed the motorist was traveling provides a convenient variable for assessing the severity

[8] We thank Jesse Shapiro for pushing us to identify a sample that would not be selected based on police stop decisions. Also see Knox, Lowe and Mummolo (2019) for further discussion about concerns when researchers rely on administrative data that has been collected in response to police enforcement decisions.

of the infraction.[9] We conduct our empirical analyses of speeding stops using data from the states of Massachusetts and Tennessee. We select these two states because, to our knowledge, they are the only states with traffic stop data containing information on the speed traveled for stops in which a warning was issued.[10] This is important since we wish to observe the full speed distribution and examining only ticketed stops might result in a truncated distribution which would be a particular issue if warning rates differed by race. We first conduct standard VOD tests for discrimination in the racial composition of speeding stops. We find that, all else equal, daylight stops are more likely to be of African-American motorists than darkness stops in Massachusetts and West Tennessee with the largest differences in Massachusetts, but observe no differences in East Tennessee.[11]

We then examine changes in the relative speed of motorists stopped for speeding violations between day and night. Consistent with the large differences in the racial composition of stops in Massachusetts, we observe that the speed distribution of African-American stopped motorists is substantially slower in daylight than darkness. On average, African-American stopped motorists are traveling 5 percent slower in daylight as a fraction of the speed limit. On the other hand, the speed distribution shift is inconsistent with the VOD test in Tennessee. In particular, we find that black stopped motorists are driving somewhat slower at night in East Tennessee, on average between 1 and 1.5 percent slower in daylight, even though the VOD test reveals no evidence of racial discrimination in stops. On the other hand, in West Tennessee, the speed distribution shift is only about 0.5 percent even though the VOD test indicated racial differences in stops. While statistically weaker due to a smaller sample size, the evidence in East Tennessee is potentially consistent with a situation where the change in motorist behavior at night may have dominated the change in police stop behavior and so prevented the VOD test from detecting discrimination. Notably, we find no evidence of a change in the speed distribution for white motorists between daylight and darkness, nor any evidence that motorist gender or vehicle type leads to changes in the speed distribution of white motorists between daylight and darkness.

[9] Another reason to focus on speeding stops is due to concerns that darkness may affect other types of traffic stops, like cell phone use or equipment failures (Grogger and Ridgeway 2006; Kalinowski, Ross and Ross 2019). In principal, researchers could include a broader sample of moving violations and construct a severity measure using fines.

[10] In Tennessee, the data explicitly distinguishes between warnings and tickets. In Massachusetts, all tickets identify the fine amount, and we interpret the presence of a sizable number of speeding tickets with fines of zero dollars as evidence that the data contains something equivalent to a warning.

[11] As discussed in detail later, we divide Tennessee into two regions that are east and west of the time zone boundary and remove counties falling directly along this divide.

To better understand this process, we develop a simple model of motorist speeding and police stops in which some motorists may choose not to speed. In our model, discrimination against African-Americans may lead minority motorists to drive more conservatively in daylight when their race is more easily observable to police. We demonstrate that a lower stop-cost for police stopping African-Americans (i.e. bias against African Americans) can have an indirect effect on the VOD test statistic through fewer African-Americans speeding during daylight. This indirect effect can exceed the direct effect of higher rates of stops based on lower stop-costs, and then the VOD test would be unable to detect discrimination. We also derive the conditions under which the speed distribution of African-American motorists who commit infractions will be slower in daylight, and this analysis implies that the largest changes in the speed distribution should occur at the top of the speed distribution of stopped motorists. This implication is consistent with the empirical patterns observed in a quantile regression analysis of the Massachusetts and Tennessee speed distributions. In Massachusetts, the daylight-darkness difference in speed for African-Americans is near zero for the 10^{th} to 30^{th} deciles and rises to between 8 and 12 percent for the 70^{th} to the 90^{th} deciles, in East Tennessee the difference is about 1 percent in low deciles and rises to between 2 and 3 percent for the highest deciles and in West Tennessee differences start near zero and rise to near 1 percent.

For all three samples, we calibrate our model to the speed distribution and the share of stops made of African-Americans motorists in daylight and darkness. The calibrated models do a very good job of matching both the information from the speed distributions and the racial composition of stops. Most significantly, the calibration for East Tennessee is able to match both the shift in the speed distribution for African-American Motorists between daylight and darkness and the VOD test statistic that is near 1 in magnitude and typically interpreted as consistent with no discrimination. The calibrated daylight stop-cost for African-American motorists is substantially below the darkness stop-costs, and the implied difference in officer utility for stopping an African-American motorist in East Tennessee is approximately the same as the utility gain arising from a two standard deviation increase in the speed of stopped motorists. Consistent with earlier papers that use the Massachusetts data, the calibrated racial differences in Massachusetts are very large implying utility differences similar to a five standard deviation increase in the speed of a stopped motorist, while in West Tennessee the small shift in the speed distribution implies utility differences similar to ½ of a standard deviation change in speed.[12] We also conduct a simulation of what the VOD test statistic would have been if African-American motorists did not respond to changes in stop-cost between daylight and darkness. In Massachusetts, the VOD test statistic increases from 1.38 to 2.74 based on the very large observed shift in the speed distribution. In East and West Tennessee, the increases are more modest from 1.00 to 1.22 and from 1.09 to 1.17, respectively.

Our finding of unexplained racial differences in speeding stops against African-Americans by Massachusetts police and Tennessee state police contribute to a large literature that examines 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 2019), fines (Goncalves and Mello 2017, 2018), police 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). Further, the paper contributes to a broader literature on discrimination that considers how minorities respond to disparate treatment both in relation to theoretical models of statistical discrimination (Lundberg and Startz 1983; Lundberg 1991; Coates and Loury 1993; Moro and Norman 2003, 2004) and to empirical attempts to measure discrimination including in terms of investment in skills and education (Neal and Johnson 1996; Lang and Manove 2011; Arcidiacono, Bayer and Hizmo 2010) and the interpretation of paired testing, audit and correspondent studies (Heckman 1998; 2004, National Research Council 2004 p109-113).

2. Evidence from Accident Data

In this section, we examine a national sample of traffic accidents for evidence of whether minority motorists adjust their driving behavior in response to darkness. As noted above, unlike data on police stops, accident data should provide indirect evidence on the driving behavior of minority motorists using a sample where the composition is not directly affected by police stop behavior. Therefore, we believe that patterns uncovered within this sample can be attributed to changes made by motorists between daylight and darkness perhaps in response to actual or perceived discrimination in stops, rather than by choices made by police officers. Our analytical sample is derived from the National Highway Traffic Safety Authority's Fatality Analysis Reporting System (FARS) data, which contains

[12] The calibrated models also yield daylight police stop-costs for white motorists conditional on the common darkness police stop-cost. Given the minimal speed shifts observed for white motorists between daylight and darkness, the calibration for all three data sets yields white daylight stops costs that are close in magnitude to darkness stop-costs.

information on all automobile accidents in the United States which result in one or more fatalities. Critically, this data set documents the race and ethnicity of accident victims, and we restrict our sample to accidents where the motorist died and was either an African-American or a Non-Hispanic white. The overall sample consists of 282,924 motorist fatalities from a total of 615,826 accidents involving a fatality that occurred in the contiguous United States from 2000 to 2017.[13]

Following Grogger and Ridgeway (2006), we further limit our analytical sample to 39,076 traffic fatalities where the accident occurred within a window of time between the earliest and latest sunset of the year, i.e. the so-called inter-twilight window (ITW). Changes to the timing of sunset occur within this window due to both seasonal variation and the discrete spring/fall daylight savings time (DST) shifts. We identify accidents occurring within the ITW based on data from the United States Naval Observatory (USNO). Since the timing of sunset varies geographically, we conservatively denote the bounds of the ITW using the eastern and westernmost coordinates of each county, where the lower bound of the window is the earliest annual easternmost sunset and the upper bound is the latest westernmost end to civil twilight. Unlike many VOD studies using traffic stop data, the FARS data also contains detailed reporting on the lighting conditions when an accident occurred. In order to minimize measurement error in visibility conditions, we use this self-reported measure rather than estimates of visibility based on USNO data. For a more thorough discussion of measurement error in VOD daylight measures, see Kalinowski et al (2019).[14]

Table 1 presents descriptive statistics for the sample 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 of motorist fatalities is more male, older, drives newer vehicles, is more likely to drive imported vehicles, and is more likely to be involved in accidents that occur on weekends and in darkness.

[Insert Table 1]

We follow the standard logic of the VOD test by placing race (R_i) on the left-hand side of the equation and testing whether stops made in daylight ($\overline{v}_i$) are more likely to be of African-American

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

[14] In Appendix A Table A1-A4, we present comparable results using USNO definitions of daylight and darkness and results are robust. As is standard, we disregard stops occurring each day during actual twilight when visibility is somewhere between daylight and darkness again accounting for geographic variation in the timing of sunset at the county level. Again, we also disregard accidents occurring outside of the earliest daily sunset and latest daily end to civil twilight.

motorists using a linear probability model. We condition on day of the week (d) and hourly time of the day (t) fixed effects to assure that the effect of daylight is identified by comparing stops that were made when the composition of the drivers is expected to have been the same. The resulting estimation equation is

$$R_{idt} = \beta \overline{v}_{idt} + \delta_d + \gamma_t + \varepsilon_{idt} \tag{1}$$

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 fixed effects or state by year fixed effects. Since many models involve high dimensional fixed effects, we estimate linear probability models rather than logistic regression as used in Grogger and Ridgeway (2006), see Kalinowski et al. (2019) who demonstrate the equivalence of the linear probability and logistic regression tests in the Grogger and Ridgeway (2006) VOD framework.[15] Standard errors are clustered at the state level when the model includes state fixed effects, but at the state by year level when the model includes state by year fixed effects.

Table 2 reports the results from estimating equation (1) using our full sample of fatal accidents involving African-American or Non-Hispanic white motorists within the ITW. Column 1 presents estimates for a model containing the controls in equation (1) 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 controls for motorist age and gender and vehicle age and whether the vehicle was an import. The estimates are sizable and highly significant with the probability that a fatal accident involves an African-American increasing by 1.5 to 1.6 percentage points in daylight relative to darkness.

[Insert Table 2]

Regardless of police enforcement activity, the behavior of minority motorists is likely to be shaped by their perceptions of police behavior. Therefore, we anticipate that the daylight differences in the race of motorists in fatal accidents will depend upon the salience of race as a factor in police

[15] The applicability of the linear probability model can be seen by starting with equation (6) in Grogger and Ridgeway (2006), setting the second term to zero (one prior to taking the log) based on their maintained assumption that motorist composition does not change between daylight and darkness, replacing the conditional probabilities for a representative motorist with the predicted probabilities arising from a linear probability model. For positive β in equation (1) above holding the incidental controls fixed, their test statistic is greater than one consistent with discrimination, and the statistic increases with increases in β.

traffic stops. Thus, we collect data on two different measures that might capture African-American perceptions about the likelihood that they might be treated differently than whites by police. The first proxy is 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.[16] The second proxy is a measure of structural racism within the state developed by Mesic et al. (2018) and includes measures of residential segregation, incarceration rates, disparities in educational attainment, and disparities in employment status. Both variables are cross-sectional and so categorize states during the entire period as ones where African-American motorists may be more or less concerned about being disproportionally targeted by police. The first three columns of Table 3 interact the relative odds that a police shooting involves an African-American with the dummy variable for daylight, where the values range from a low value of 0.04 (odds of 1.04) in Connecticut to 16.76 in Rhode Island. The last three columns interact daylight with the racism index, which originally ranged from 0 to 100 but has been standardized and now range from -1.87 (index of 25.9) in Montana to 2.61 (index of 74.9) in Wisconsin. We observe a consistent pattern where the proxy for the salience of potential discrimination in police stops 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 the racial differences associated with daylight fatalities of 0.4 percentage points, while a one standard deviation increase in the racism index implies a 0.7 percentage point increase.

[Insert Table 3]

As discussed above, the VOD approach has been applied in many states. Here, 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. Thus, we test whether we observe evidence of behavioral adjustments of minority motorists in the specific states where the test failed to find discrimination.[17] We then repeat the analyses of Table 2 using this sample, and the results are presented in Table 4. We find even larger racial differences in fatalities in this subsample. Specifically,

[16] Our data on police shootings come from Mesic et al. (2018). However, our findings are robust to constructing a shootings ratio from the Fatal Encounters (https://fatalencounters.org/) or Mapping Violence (https://mappingpoliceviolence.org/) databases.

[17] 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 as the VOD traffic stop studies referenced above.

we find that daylight motorist fatalities are over 3 percentage points more likely to involve African-American motorists than equivalent fatalities that occurred in darkness. While the fatality differences do not imply discrimination in police stops in these states, the data is suggestive that minority motorists are concerned about such stops and government attempts to detect discrimination in stops may be affected by the associated behavioral adjustments. All standard errors are clustered at the state by year level.[18]

[Insert Table 4]

Lastly, we address the potential concern that the overall composition of motorists might be changing in response to daylight within the ITW. Formal tests of balance are wholly absent in the existing applications of the VOD test because it is impossible to use traffic stop data alone to disentangle changes in enforcement activity that affect the composition of stops from compositional changes in traffic patterns. However, in the case of our accident data, we can reasonably expect that police traffic stop behavior did not directly affect the composition of motorists and vehicle attributes associated with traffic fatalities, at least for those fatalities involving white motorists. We examine the composition of white non_hispanic motorists involved in fatal accidents in Table 4. 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 in the FARS data. 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 any of these variables. None of the t-statistics across motorist or vehicle characteristics are significant, and in the full model in column 5 the F-statistic associated with the four estimates parameters is 1.37 (p=0.24). Motorist race appears to be the only motorist or vehicle characteristic available in the FARS data for which fatality rate differences correlate with daylight conditional on time of day.

[18] We cluster standard errors by state by year in the state fixed effects model in column 1 due to the small number of states. This decision is conservative empirically in that clustering at the state level yields smaller standard errors than arise with state by year clustering.

[Insert Table 4]

While the use of traffic fatality data avoids mingling changes in police treatment of minority motorist and minority motorist behavior, one might be concerned that motorists differ in their selection into the sample of motorist fatalities. Recall that we have detailed data on all motorists involved in a fatal traffic accident, but only race and ethnicity for the fatalities themselves. Therefore, we also estimate models where observations are weighted based on the likelihood that an individual is involved in a fatal traffic accident using vehicle attributes and information on restraints, i.e. airbags and seatbelt usage. Thus, we conduct an inverse probability weighting analysis by reweighting the sample of fatal accidents to resemble the overall sample of both fatalities and non-fatalities for those accidents involving at least one fatality. While this only controls for selection on the observed variables, the results presented above are robust to the use of this alternative model and can be found in Appendix Tables A5-A8. 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.

In this section., we have presented evidence that minority motorists are involved in accidents at a lower rate nationally during periods of daylight relative to equivalent periods of darkness. We provide strong evidence that changes in minority accident rates in response to visibility are larger in states with more police shootings and with a history of racism. In fact, 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 when they are driving during daylight and so during daylight hours choose to drive more conservatively and more carefully. 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, as opposed to changes that arise across many motorist attributes.

3. 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 violations for headlights, seatbelt, and cellphones are likely correlated with visibility for reasons other than race. Our focus on speeding stops also has the added advantage of having a clear measure of infraction severity that we can use to assess changes in motorist driving behavior, e.g. 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.[19] As noted above, we selected these two states because the stop records for these states contain information on the speed traveled for stops in which a warning was issued.[20] In the Massachusetts data, we observe both stops by local police and state police. In order to focus on stops containing a reasonable number of African-American motorists, we restrict our analysis to state police stops and stops made by town police departments of the 10 largest towns.[21] In the Tennessee data, we make a distinction between patrol districts lying on the Eastern and Western side of the time zone border that bisects the state.[22] 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.[23]

Table 6 presents descriptive statistics for the samples of speeding stops within the ITW, but not in actual twilight. The Massachusetts sample numbered 10,203 speeding stops, while the samples in East and West Tennessee, respectively, contain 23,515 and 102,054 speeding stops.[24] 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

[19] The Massachusetts data was collected by Bill Dedman for the Boston Globe study and used by Antonovics and Knight (2009) to study discrimination in police searches. Antonovics and Knight (2009) only use stops by the Boston Police Department due to limitations in the availability of data on police officers. The Tennessee data was obtained through the Stanford Open Policing Project.

[20] 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.

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

[22] We also exclude three rural patrol districts (of eight total) that lie directly adjacent or on top of the border between the Eastern and Central Standard Time zone boundary. A significant portion of those traffic stops occur on opposing sides of the time zone from a patrol district's headquarters. Thus, it is unclear whether the timestamp is denoted in the time zone of the headquarters or the location where the stop is made potentially creating measurement error in our main variable of interest, i.e. daylight. We find that estimates using the overall sample are less precise but qualitatively and quantitatively similar to our preferred specification which excludes these patrols.

[23] 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 and had no information related to the minute. As a result, only traffic stops that occurred during the ITW in an hour of complete daylight or darkness were included. Although this additional restriction reduces the overall sample size, it occurs randomly and poses no threat to our identification strategy. In Massachusetts, the earliest Easternmost sunset was recorded in Orleans, MA while the latest Westernmost end to civil twilight occurred in Mount Washington, MA. In Tennessee, the ITW was bounded by recordings of the earliest Easternmost sunset and latest Westernmost end to civil twilight occurring in Erwin, TN and Memphis, TN respectively.

[24] The Massachusetts sample has a lower share of stops that fall within the ITW because we only have the hour that a stop was made (not the minute) and must disregard stops made in hours that were not fully darkness/daylights. As we noted before, we do not believe this poses any threat to identification other than a reduction in power.

to be of African-Americans, but stops of males were less likely to be of African-Americans in east Tennessee and more likely to be of African-Americans in west Tennessee.

[Insert Table 6]

Table 7 presents the VOD estimator for all three geographic samples of speeding violations. The model follows equation (1) from the traffic fatality data almost exactly except that the sample is of speeding stops, rather than traffic fatalities, and the geographic fixed effects are within state. In terms of controlling for geography, we use town and state police barracks fixed effects because counties are quite large in Massachusetts. On the other hand, in Tennessee, counties are small in size relative to state police patrol districts so those models are estimated with county fixed effects. Standard errors are clustered at the city or state police barracks level for Massachusetts, and at the county by year level for Tennessee.[25] Columns 1, 3 and 6 present estimates for models that include time of day, day of week, geographic and in Tennessee year fixed effects. Columns 2, 4 and 7 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. Finally, as an additional robustness test, in columns 5 and 8, we present results with county by year fixed effects for the east and west Tennessee analyses since the Tennessee sample covers 10 years.

In two of the three jurisdictions, Massachusetts and West Tennessee, we find evidence suggesting that the odds that a minority motorist is stopped increases in daylight relative to darkness. Our linear probability models suggest that a daylight stop in Massachusetts during this time period 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 involve African-Americans. The magnitude 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 discrimination in East Tennessee using the traditional VOD estimator. The classic interpretation of these results would be that Massachusetts and West Tennessee show evidence of discriminatory policing towards

[25] In principle, we would 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 and so has a very small number of clusters. We have confirmed that the standard errors based on clustering at the county level are smaller than the standard errors clustered at the county by year level and so to be conservative we cluster at the county by year level. Standard errors in West Tennessee, which contains more counties, are very similar when comparing clustering at the county and at the county by year level.

minority motorists, but that East Tennessee does not. Appendix Table A9 presents comparable estimates following Grogger and Ridgeway and using the logistic regression instead of the linear probability model. The estimates are qualitatively very similar.

[Insert Table 7]

Next, we explore our motivating hypothesis that the speed of stopped minority motorists changes across daylight and darkness in response to real or perceived discrimination. Specifically, we estimate a model using Ordinary Least Squares of the form

$$S_{idt} = \beta_0 + \beta_1 R_{idt} + \beta_2 \overline{v}_{idt} + \beta_3 (R_{idt} \cdot \overline{v}_{idt}) + \delta_d + \gamma_t + \varepsilon_{idt} \tag{2}$$

where S_{idt} is recorded speed relative to the speed limit, e.g. $speed/speed\ limit$. We calculate a relative speed based on both our intuition that the same absolute speed limit violation will be more concerning to police 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. As above, 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 β_3 is the coefficient on the interaction of these two variables, which captures racial heterogeneity in speed distribution shift with daylight. δ_d and γ_t are day of week and time of day fixed effects, respectively. As above, we add geographic fixed effects, city/police state barracks for Massachusetts and county for Tennessee, and for the Tennessee samples where we have a longer time horizon we also include year fixed-effects and estimate robustness tests using county by year fixed effects.

Table 8 contains the results from applying equation (2) to our speeding data from Massachusetts and Tennessee. For the estimates using the Massachusetts sample, we find evidence of a large and statically significant decrease in the speed of stopped minority motorists in daylight. The estimates indicate that the speed of stopped motorists decreases in daylight by over 5 percentage points relative to the speed limit. The next largest speed shift arises in East Tennessee where daylight implies that African-Americans travel 1.5 percentage points slower, although the estimates are only significant at the 10 percent level possibly due to the smaller sample size. The smallest increase is in West Tennessee where speeds decrease by only 0.5 percentage points. The larger speed distribution shift in East Tennessee relative to West Tennessee is suggestive that racial differences in police stops may exist that were not detected by the VOD test in Table 7. The magnitudes of the estimated effects are quite stable to controls for motorist and vehicle attributes. However, the estimated effect size in East Tennessee is eroded by approximately 20 percent with the addition of county by year fixed effects, and the t-statistic falls to 1.5, below standard thresholds for significance. Nonetheless, the estimated effect in East Tennessee is still substantially larger than the estimated effect size in West Tennessee. Notably, the estimates on daylight are always near zero and statistically insignificant suggesting that the speed distribution of stopped white motorists does not change between daylight and darkness.

[Insert Table 8]

Next, we estimate unconditional quantile regressions following Firpo, Fortin, and Lemieux (2009) using a software package described in Borgen (2016). By following these methods, we obtain marginal effect estimates at different quantiles of the data based on the level of the dependent variable so that the estimates for the 10th percentile represent the effect for the subsample whose level of speeding, again relative to the speed limit, places them at the 10th percentile of speeders. 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 speed distribution, and (3) use RIF as the outcome variable in a linear model, so-called RIF-OLS or unconditional quantile regression (Firpo, Fortin, and Lemieux 2009). We first create a kernel density distribution by smoothing speeds such that we can observe an estimated density for any discrete point in the speed distribution

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

where S_i is again simply recorded speed relative to the speed limit. The bandwidth parameter h is selected following a standard procedure in kernel density estimation that would minimize the mean integrated squared error if the data were Gaussian and a Gaussian kernel were used.[26] The kernel

[26] 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.

function K is robust to a variety of alternative functional forms but is specified as Epanechnikov in our estimates. We obtain a smoothed speed and density at each numeric decile τ of the distribution since we now have a continuous representation of the distribution. We then calculate the Recentered Influence Function (RIF) for each decile in the kernel smoothed speeding data within the inter-twilight sample as follows

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

where q_τ and f_{spd} are the estimated speed and density at decile τ based on the kernel smoothing estimate of the speed distribution, and $\mathbb{I}$ is an indicator function. Using the decile RIF's for each i observation, we estimate changes in the speeding distribution by estimating linear models for the RIF at each decile. Specifically, we estimate unconditional quantile regressions of the form

$$RIF_{\tau,idt} = \beta_{\tau,0} + \beta_{\tau,1} R_{idt} + \beta_2 \overline{v}_{idt} + \beta_{\tau,3}(R_{idt} \cdot \overline{v}_{idt}) + \delta_{\tau,d} + \gamma_{\tau,t} + \varepsilon_{\tau,idt} \tag{3}$$

where the variables follow the convention discussed with respect to (2). The parameter of interest $\beta_{\tau,3}$ is again the coefficient estimate on the interaction between daylight and minority status, which here captures racial heterogeneity in terms of motorist speed at a specific τ quantile.

Table 9 presents the results from applying equation (3) plus geographic and in the case of Tennessee year fixed effects to the same sample of speeding stops used for the estimates in Table 8. We continue to find evidence of slower speeds in daylight for African-American motorists, but in all cases the speed distribution shift tends to be larger for the higher percentiles. In Massachusetts, the shift is quite large starting near zero at the 10th percentile and rising over 10 percentage points at the 80th and 90th percentiles. Again, 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. The shift in West Tennessee is smaller starting at zero and reaching a maximum near 1 percentage point at the higher percentiles. In the next section, we will demonstrate that this phenomenon arises naturally in a sample of motorists who commit speeding violations when some individuals choose not to speed due to the risk of police stops.

[Insert Table 9]

The quantile regressions yield multiple estimates for each sample and raise concerns about conducting inference in the context of multiple hypothesis testing. We follow Bifulco et al. (2008) and conduct a simulation exercise to assess the likelihood of the pattern of results for each subsample having arisen by chance.[27] Bifulco et al. (2008) suggest exploiting the logic of a traditional permutation test in a bootstrap framework by 1. ordering the t-statistics arising from the coefficients for each quantile by magnitude, 2. drawing bootstrap samples of the same size as the original sample with replacement under the null of no conditional correlation between speed and 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 set of t-statistics arising from the data. While the t-tests indicated in Table 9 are two-sided, this permutation test is conducted in a one-sided manner where the signed and ordered bootstrap t-statistics dominate the sample signed and ordered t-statistics if all T-statistics for a bootstrap sample have a lower value than the corresponding sample T-statistics. We can strongly reject the null hypothesis of no negative shift in the speed distribution for all three subsamples. In Massachusetts, the likelihood of this pattern arising by chances is 0.013 percent. In East and West Tennessee, the likelihoods are 0.005 and 0.001 percent, respectively. We also re-estimate these models adding the available motorist and vehicle controls (Appendix Table A.10), and in the Tennessee subsamples adding the county by year fixed effects (Appendix Table A.11). The addition of motorist and vehicle controls has no impact on the quantile estimates. As in Table 8, the addition of county by year fixed effects erodes the speed shift in East Tennessee somewhat with the point estimates of the relative speed effects in the upper percentiles being between 15 and 20 percent smaller, but entire pattern of the speed distribution shift 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 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 9 replacing race in (2) with either motorist male or whether domestic vehicle. To avoid confounding the effects of race and the motorist and vehicle attributes, we restrict this analysis to the

[27] This test examines whether the entire set of estimated parameters could have arisen by chance or the likelihood of incorrectly rejecting a composite null hypothesis of no change over all percentiles, in contrast to much more conservative Bonferroni-style corrections which are designed to allow us to conclude whether the null hypothesis is rejected for a specific parameter.

sample of non-Hispanic white motorists only. These results are shown in Table 10 for the analysis over driver gender and in Table 11 for the analysis over vehicle type. While we see some evidence in Massachusetts that stopped male motorists drive faster and some evidence in West Tennessee that stopped motorists who drive domestic vehicle travel at higher speeds, we observe minimal evidence in any of the subsamples of an interaction between either the motorist or vehicle attribute and daylight, unlike the strong pattern of results on motorist race. 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.[28] For Massachusetts, we also can conduct these analyses for whether the driver is young (below the age of 30), the vehicle is an older model (older than 5 years) and whether the vehicle is red. As above, we find no evidence of a systematic change in speeds with daylight associated with these motorist and vehicle attributes. Finally, for Tennessee, we can repeat the analyses above on male motorists and domestic vehicles including county by year fixed effects, and the lack of significant findings is robust. These additional analyses are shown in Appendix Tables A12-A16.

[Insert Tables 10 and 11]

In this section, we have presented evidence on the speed distribution of stopped motorists based on traffic stop data documenting speed distribution shifts where relative to white motorists African-American motorists travel more slowly in daylight, when presumably their race can be observed. Not unexpectedly, the largest differences in the speed distribution arise in the Massachusetts sample where we also observed the largest racial differences between daylight and darkness stops consistent with a high level of discrimination causing a large decline in the speed of African-American motorists. At the same time, we observed that in Tennessee 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. To some extent, this evidence from East Tennessee is similar to the evidence from the fatality data where we found that the representation of African-Americans among fatal accidents was lower in daylight in the set of nine states where no or mixed evidence was found of discrimination when applying a VOD test. A key limitation of our analysis of motorist speed is that we only observe the speed distribution of stopped motorists. In the next two sections, we will develop

[28] We follow the same permutation strategy for testing the null hypothesis except that the test is two-sided using the absolute value of the T-statistics because we have no initial hypothesis concerning how these attributes might shift the speed distribution.

a simple model of police stop and motorist speeding behavior and calibrate that model to the data on stopped motorists from Massachusetts and East and West Tennessee so that we can more explicitly account for selection into being stopped by police when considering the observed speed distributions.

4. Simple Model of Police-Motorist Interaction

In this section, we develop a simple economic framework for presenting the basic intuition of the VOD test and the concerns that arise when drivers can change their behavior in response to discrimination. We begin by imposing assumptions consistent with the basic requirements for the VOD test to be valid. Specifically, we assume that the distribution of motorists *f* over motorist type $d \in \{w, m\}$ does not depend upon visibility (v).

Assumption 1. $f(i, d, v) \equiv f(i, d)$.

We also assume that the behavior of police officers in terms of motorist stop likelihood, as captured by stop-costs *s*, only depends upon visibility through the inability of officers to observe motorist type. Therefore, we assume that low visibility ($\underline{v}$) stop-costs are equal across type, high visibility ($\overline{v}$) stops are higher for whites (w) than minorities (m) i.e. no reverse discrimination, and low visibility stop costs are bounded by the high visibility ($\overline{v}$) stop-costs associated with the majority and minority groups.[29]

Assumption 2. $s_{\underline{v}} = s_{\underline{v},w} = s_{\underline{v},m}$, $s_{\overline{v},w} \leq s_{\overline{v},m}$ *and* $s_{\underline{v}} \in [s_{\overline{v},w}, s_{\overline{v},m}]$

The following assumptions, 3 through 5, are imposed in order to create a model where some motorists choose not to commit an infraction when they face low police stop costs, but may commit infractions if stop costs increase. A similar model can be derived from police and motorist utility maximization problems, and using this model assumptions 3 through 5 can be derived while only imposing assumptions on primitive model parameters. This more extensive model and the derivation of results are shown in Appendix B. First, we intuitively assume that stop probability of a motorist

[29] Validity of the VOD test requires that officer returns and costs of stops are unaffected by visibility except that group membership is unobserved in darkness. If there is a favored and disfavored group, rational officers would not assign higher stop costs for the disfavored group (or lower costs for the favored group) relative to an individual whose group membership is unknown.

who commits an infraction ϕ^* is increasing in motorist speed, decreasing in police stop-costs, and bounded away from zero for finite stop-costs.[30]

Assumption 3. $\frac{\partial \phi^*}{\partial i} > 0$, $\frac{\partial \phi^*}{\partial s_d} < 0$, *and finite.* $\lim_{i \to 0} \phi^*(i, s_d) > 0$ *for all* s_d.

Given that ϕ^* increases with i, the assumption on the limit of ϕ^* implies that the probability of stop is always positive.

Based on equation (4) in Grogger and Ridgeway (2006), our simple test statistic *K* for the VOD test is the ratio of minority to majority stops when race is visible scaled by the ratio of minority to majority stops when race is not visible, the counter-factual, or

Definition 1. $K_{VOD} \equiv \frac{p[m|stopped,\overline{v}]}{p[w|stopped,\overline{v}]} \frac{p[w|stopped,\underline{v}]}{p[m|stopped,\underline{v}]} = \frac{\int_0^\infty f(i,m)\phi^*(i,s_{\overline{v},m})di}{\int_0^\infty f(i,w)\phi^*(i,s_{\overline{v},w})di} \frac{\int_0^\infty f(i,w)\phi^*(i,s_{\underline{v},w})di}{\int_0^\infty f(i,m)\phi^*(i,s_{\underline{v},m})di}$

Proposition 1. *Under the null hypothesis of equal stop-costs/treatment in high visibility and assumptions 1-3,* $s_{\underline{v}} = s_{\overline{v},w} = s_{\overline{v},m}$ *implies that* $K_{VOD} = 1$. *If* $s_{\overline{v},m} < s_{\overline{v},w}$, $K_{VOD} > 1$.

Proof. *The first part of the proposition holds by inspection. Under the null, the allowed interval for* $s_{\underline{v}}$ *in assumption 2 is degenerate to a single point with equal stop-costs. Once stop-costs are the same over visibility, the integrals over the speed distribution for a type d are identical between high and low visibility and those integrals cancel out of the expression in Definition 1.*

The second part of the theorem is established by examining the impact of increasing $s_{\overline{v},w}$ *and/or decreasing* $s_{\overline{v},m}$. $\frac{\partial}{\partial s_{\overline{v},d}} \int_0^\infty f(i,d)\phi^*(i, s_{\overline{v},d})\, di = \int_0^\infty f(i,d) \frac{\partial \phi^*}{\partial s_{\overline{v},d}} di < 0$. *The derivative is negative based on assumption 3. Therefore, any decrease in* $s_{\overline{v},m}$ *increases the numerator of* K_{VOD}, *and any increase in* $s_{\overline{v},w}$ *decreases the denominator of* K_{VOD}. *Starting at* $K_{VOD} = 1$ *for equal treatment with any arbitrary* $s_{\underline{v}} = s_{\overline{v},w} = s_{\overline{v},m}$, *assumption 2 implies that we can only*

[30] Although stop probability can limit to zero as stop costs approach infinity.

deviate from equal treatment by decreasing $s_{\overline{\upsilon},m}$ or increasing $s_{\overline{\upsilon},w}$ and both of these changes increase K_{VOD} above 1. QED

Next, we describe a set of functions that will characterize motorist behavior in response to stop-costs and the resulting likelihood of being stopped, effectively relaxing assumption 1. Motorists are distributed following the probability distribution *g* over motorist type and a parameter c that determines their benefits arising from *i*. We assume that the optimal infraction level (i'), given that the motorist commits an infraction, is always positive, and assume that optimal motorist infraction severity increases with officer stop-costs and with the motorist's benefit parameter c, or

Assumption 4. $i'(c, s_d) > 0$ for all c and s_d, $\frac{\partial i'}{\partial s_d} > 0$, $\frac{\partial i'}{\partial c} > 0$ *and finite.*

For convenience, we will refer below to a higher level of infraction severity as traveling faster as in the context of a speeding violation.

Next, we assume that some individuals with low benefits from infraction at their optimal infraction level i' will choose not to speed and so be at a corner solution with an infraction level of zero. This corner solution may arise for example because there is always some positive risk of being stopped for speeding no matter how small the infraction level, see Assumption 3. We define c^* as the preference parameter value for which a motorist is indifferent between committing the infraction and not, and i^{**} as the actual infraction level selected. We also assume that the number of motorists choosing to infract (based on changes in c^*) rises as stop-costs rise.

Assumption 5. *There exists a finite* c^* *such that* $i^{**}(c, s_d) = \begin{cases} i'(c, s_d) > 0 & if\ c > c^*(s_d) \\ 0 & otherwise \end{cases}$ *for all* $s_d > 0$, $\frac{\partial c^*}{\partial s_d} < 0$ *and finite.*

In principle, we also might reasonably assume that the likelihood of being stopped falls with stop-costs even though motorist travel faster when stop-costs rise. We specify a new function for the probability that a motorist with a given benefit parameter is stopped in equilibrium as $\tilde{\phi}(c, s_d) \equiv \phi^*(i'(c, s_d), s_d)$, and impose

Assumption 6. $\frac{\partial\tilde{\phi}}{\partial s_{\overline{v},d}} = \frac{\partial\phi^*}{\partial s_d} + \frac{\partial\phi^*}{\partial i}\frac{\partial i'}{\partial s_d} < 0$

We do not rely on assumption 6 below. While Assumption 6 is derived in Appendix B, the derivation of assumption 6 requires imposing an assumption upon an equilibrium function arising from the solution of the motorist's problem, rather than on the primitives of the problem. Intuitively, however, it seems reasonable that motorists who travel faster in response to a decreased likelihood of stop do not travel so much faster that the likelihood of stop increases in equilibrium, and we provide the condition above for illustrative purposes.

Finally, we impose a somewhat restrictive assumption on the density of motorists over c requiring an upper bound of the density distribution with non-zero density at that upper bound. We also assume that the distribution is non-zero, continuous and differentiable over its relevant domain.

Assumption 7. *There exists a* c_h *such that* $g(c,d) = 0$ *for all* $c > c_h$ *and* $g(c,d) > 0$ *for all* $c \leq c_h$, *and* $g(c,d)$ *is continuous and differentiable over* $(-\infty, c_h)$.

The model in Appendix B also imposes this assumption.

We now define a new formulation of our test statistic that allows for the behavioral response of motorists so that the speed distribution changes in response to police stop-costs.

Definition 2. $K_{ERS} \equiv \frac{\int_{c^*(s_{\overline{v},m})}^{c_h} g(c,m)\tilde{\phi}(c,s_{\overline{v},m})di}{\int_{c^*(s_{\overline{v},w})}^{c_h} g(c,w)\tilde{\phi}(c,s_{\overline{v},w})di}\frac{\int_{c^*(s_{\underline{v},w})}^{c_h} g(c,w)\tilde{\phi}(c,s_{\underline{v},w})di}{\int_{c^*(s_{\underline{v},m})}^{c_h} g(c,m)\tilde{\phi}(c,s_{\underline{v},m})di}$

Proposition 2. *Under the null hypothesis of equal stop-costs/treatment in high visibility (*$s_{\underline{v}} = s_{\overline{v},w} = s_{\overline{v},m}$*) and assumptions 2-5 and 7,* $K_{ERS} = 1$. *If* $s_{\overline{v},m} < s_{\overline{v},w}$, *parameters exist for any set of functions satisfying assumptions 3-5 and 7 such that* $K_{ERS} < 1$.

Proof. *The first part of proposition 2 is established again by inspection. Once stop-costs are the same over visibility, the integrals over the speed distribution for a type d are identical between high and low visibility and those integrals cancel out of expression in Definition 2.*

The second part of the theorem is again established by examining the impact of increasing $s_{\overline{v},w}$ *and/or decreasing* $s_{\overline{v},m}$. $\frac{\partial}{\partial s_{\overline{v},d}} \int_{c^*}^{c_h} g(c,d)\tilde{\phi}(c, s_{\overline{v},d})\, di = -\frac{\partial c^*}{\partial s_{\overline{v},d}} g(c^*,d)\tilde{\phi}(c^*, s_{\overline{v},d}) + \int_{c^*}^{c_h} g(c,d)\frac{\partial \tilde{\phi}}{\partial s_{\overline{v},d}} di$

The first term is positive by Assumption 5 as stop-costs rise new motorists with lower values of c *begin to speed raising the group's share in the population. The second term is ambiguous. If assumption 6 is violated, the second term is non-negative, and the entire derivative is positive. Therefore, if we start at equal stop-costs and* $K_{ERS} = 1$*, decreases in the stop-costs for minorities in high visibility will lower the numerator and increases in the stop-costs for whites will raise the denominator, both lowering the value of* K_{ERS} *below 1. In this case, higher relative stop-costs for whites tend to lead to higher relative speeds for white and so raises the representation of whites among stopped motorists.*

However, if assumption 6 is satisfied as expected intuitively, the second term is negative leading to an ambiguity in the derivative. The rest of the proof will proceed by constructing a general method for selecting parameters that yield $K_{ERS} < 1$. *Select an* $\underline{s}_d > 0$. *By assumption 5, there exists a value* c_h *such that* $c_h = c^*(\ \underline{s}_d)$. *Now set the maximum* c *for both the white and minority distributions over* c *to that value* c_h. *Let* $s_{\underline{v}} = s_{\overline{v},w} = s_{\overline{v},m} = \underline{s}_d - \varepsilon$ *where* ε *is a very small number so that almost no motorists speed and* $K_{ERS} = 1$ *since stops costs are equal. By the assumptions of continuity for all functions over their relevant domains, the second term of the derivative above can be brought arbitrarily close to zero by setting* ε *arbitrarily close to zero. Very few people are speeding so an increase in stop-costs has a minimal marginal effect on the speed of all motorists. Further, the first term of the derivative is positive and bounded away from zero based on the non-zero density for values of* c *at or below* c_h *imposed by assumption 7, the non-zero derivative implied by assumption 5 and the non-zero stop probability from assumption 3. Once we have found an* ε *small enough to yield a positive derivative, an increase in* $s_{\overline{v},w}$ *or a decrease in* $s_{\overline{v},m}$ *implies prejudice against minorities, but will increase the magnitude of the denominator or decrease the magnitude of the numerator, respectively, based on changes in* c^* *reducing* K_{ERS} *below 1. QED*

While the proof focuses on the validity of the VOD test under the null hypothesis and the potential for a test statistic below one when minority motorists face discrimination, the proposition clearly illustrates the potential for substantial bias towards zero in the VOD test. Even though the stop probability of minority motorists should fall at night, the increase in the incidence of minority motorists committing infractions implies that the lower stop probability is drawn from a larger share of minority motorists at night reducing the VOD test statistic towards one (and perhaps beyond one in extreme cases).

We can also use this framework to examine the speed distribution of motorists who commit infractions. Assumption 4 implies that optimal motorist speed increases as stop-costs rise, but the changes in the actual distribution of speed is ambiguous. The speed distribution of motorists who commit infractions is affected in two ways by higher stop-costs: motorists who would have committed an infraction anyway driving faster, and additional motorists who have weaker preferences for the infraction level and so tend to drive slower now choose to speed. While we only observe the speed distribution of stopped motorist in our empirical work above, we focus on the speed distribution of motorists infracting for simplicity of exposition, and we derive similar results to proposition 3 based on the same intuition for the speed distribution of stopped motorists. These derivations are shown in the Proposition B4 in Appendix B.

We can characterize the speed distribution of speeding motorists by examining the effect of a change in stop-costs on the travel speed i_x of motorists at a specific percentile x of motorists who speed. If *G* is the CDF of *g*, the percentile for a given c can be written as

$x(c, s_d) = \frac{G(c) - G(c^*(s_d))}{1 - G(c^*(s_d))}$ or equivalently $c_x(x, s_d) = G^{-1}\big((1 - G(c^*))x + G(c^*)\big)$

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

Proposition 3. *For all* s_d *there exists* $\tilde{x}$ *such that* $\frac{\partial i_x}{\partial s_d} > 0$ *for all* $x > \tilde{x}$

Proof. *Substituting the expression for* $c_x(x, s_d)$ *into the definition of* $i_x(x, s_d)$ *and differentiating yields*

$$\frac{\partial i_x}{\partial s_d} = \frac{\partial i'}{\partial s_d} + \frac{\partial i'}{\partial c}\frac{dG^{-1}}{d\{\}}(1 - x)\frac{dG}{dc}\frac{dc^*}{ds_d}$$

By assumption 4, the first term of the derivative is positive and bounded away from zero, by assumptions 4, 5 and 7 the derivatives in the second term are finite, and by assumption 7 $x = 1$ *when* $c^* \geq c_h$. *Therefore, the derivative of* i_x *is positive when* $x = 1$ *and by continuity there are values of* x *just below one where the derivative is also positive. QED*

Proposition 3 provides additional insights into the effects of motorist behavioral changes on the speed distributions of motorists who choose to speed or stopped motorists. The bias in both the VOD test statistic and the speed distribution shift is driven by changes in c^*. Therefore, places with the largest bias towards or below one in the test statistic are also likely to have the largest difference between the speed distribution shift for all motorists and distribution shift for speeding motorists, based on the second term in the derivative above. Under the alternative hypothesis of discrimination, places that have an inconclusive VOD test due to this bias are also likely to be places where the shift in the speed distribution is hardest to detect when only observed for stopped motorists. Perhaps, this feature of our model helps explain why it was difficult to statistically detect the shift in the speed distribution in East Tennessee.

However, this phenomenon grows weaker as we move further out the speed distribution because the additional speeders added at the bottom of the distribution result in only a fraction of those speeders being shifted across any percentile, see the scale factor of $(1 - x)$ in the second term of the derivative. As the percentile approaches one (the top of the speed distribution), the share shifted approaches zero, while the first term (the partial derivative of i') remains bounded away from zero. Therefore, evidence of a significant increase in the speed of minority motorists near the top of the speed distribution of stopped motorists, as found for all three samples considered above, is suggestive that the VOD test is biased away from finding discrimination.

5. Calibration and Simulation

In order to simulate the motorist speed distribution and calculate the VOD test statistic for this distribution, we need to parameterize the key functional forms for the model above. First, we allow motorist preferences c to follow a skew-normal distribution with separate parameters for whites and blacks. The skew-normal probability density function with skewness a is defined by

$$f(t) = 2\phi(t)\Phi(at)$$

Where ϕ and Φ are the normal PDF and CDF respectively, and $t = (x - e)/w$ with location e and scale w. Note that we do not attempt to impose the restriction associated with Assumption 7 on the density function.

Next, we need to parameterize the probability of being stopped ϕ^* as a function of speed/infraction severity, visibility, and group-based police stop-costs. We begin by specifying a simple additive function combining a monotonic function of motorist speed with police stop-costs. Specifically,

$$u(i, s_d) = i^\eta + u_0 - s_d \quad \text{for } i > 1 \text{ and } \; u(0, s_d) = 0$$

where $\eta > 1$ to allow the return to a stop to increase non-linearly with the severity of the infraction and $u_0 > \max(s_{v,d})$ for all $\{v, d\}$ as a baseline return to police stops to assure that $u > 0$ for all positive infraction levels. One might interpret $u(i, s_d)$ as a net return to police from stops.

Next, we need a monotonic mapping from $u(i, s_d)$ over $R^+$to stop probability ϕ^* where the probability limits to a positive value as $i \to 0$, as noted above in assumption 3. To achieve this, we specify a function $\phi(h)$ that is positive for all positive values of h and where $\phi(0) = 0$. In this way, the requirement that $u_0 > \max(s_{v,d})$ assures that the argument of ϕ is always greater than zero so that that the probability of a stop is always positive for a non-zero infraction level. Specifically,

$$\phi^*(i, s_d) \equiv \phi(u(i, s_d)) \;\text{ where }\; \phi(h) = \begin{cases} \frac{h^a}{h^a + K} & h > 0 \\ 0 & otherwise \end{cases}, a > 0 \text{ and } \; K > 0$$

The function also limits to a probability of 1 as $h \to \infty$. If $a < 1$, the function has a negative second derivative for $h > 0$. Otherwise, the second derivative of the function can change sign with *i*. Additionally, as $h \to \infty$ the second derivative must be negative.

The motorist speed function should naturally be a function of the stop probability and motorist preferences c. To provide simple mico-foundations to this relationship, we assume a penalty upon being stopped that depends on i and a benefit function from speeding or committing the infraction that depends on both i and c. We assume that the penalty for being stopped is

$$\tau(i) = i^\mu + \tau_0$$

where $\mu > 1$ and $\tau_0 > 0$ so that costs are convex in the infraction level and that the benefit from speeding is

$$b(i,c) = b_0 i^{\alpha_1} e^{\alpha_2 c}$$

where $b_0 > 0$, $0 < \alpha_1 < 1$ so that there are diminishing marginal returns in benefits from the infraction level, and $\alpha_2 > 0$. A rational motorist would solve the problem

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

to find the optimal speed $i'(c,s_d)$. The resulting functions satisfy Assumption 4 under reasonable assumptions on the primitive parameters.

While a closed-form solution does not exist for $i'(c,s_d)$, we can impose the implicit function theorem exploiting the monotonic nature of $i'(c,s_d)$ and define $c'(i,s_d) = i'^{-1}(i,s_d)$. The first order equation from the motorist's optimization problem then yields a closed-form solution for

$$c'(i,s_d) \equiv \frac{1}{\alpha_2} ln\left(\mu\phi^*(i,s_d) i^{\mu-\alpha_1} + \frac{\partial\phi^*}{\partial i}\frac{(i^\mu + \tau_0)}{i^{\alpha_1 - 1}}\right) - \frac{\ln(\alpha_1 b_0)}{\alpha_2}$$

Rather than solving numerically for $i'(c,s_d)$ over c, which would be computationally burdensome, we calculate $c'(i,s_d)$ over a fine grid of values of i and create a piece-wise approximation of $i'(c,s_d)$ for any value of c by linearly interpolating between the two nearest points in the grid for c'.

Our parameterization also assures that Assumption 5 is satisfied under reasonable parameter values. For a given set of parameters, we can calculate the motorist's optimal speed for each value of c, and then solve for the value $c^*(s_d)$ where net benefits at the optimal speed are exactly equal to zero. Based on the monotonic nature of the net benefits function, motorists net benefits are negative for all values of c below $c^*(s_d)$, and those motorists select an infraction level of zero. With $\phi^*(i,s_d)$, $i'(c,s_d)$ and $c^*(s_d)$ in hand for a given set of parameters, we can solve for the equilibrium speed distribution and the speed distribution of stopped motorists by drawing a large sample of motorists from the distribution of c. By assigning separate skewness, location and scale parameters for white and minority motorists, a common police stop-cost $s_{\underline{v}}$ in darkness and separate daylight police stop-costs for white and minority motorists, $s_{\overline{v},w} > s_{\underline{v}}$ and $s_{\overline{v},m} < s_{\underline{v}}$, respectively; we can use the same sample over c to simulate white and minority speed distributions in daylight and in darkness. Then, we can vary the representation of minority motorists by applying a weight to the minority distribution and based on that representation calculate the share of stops in daylight and darkness that involve minority motorists.

In order to calibrate the model to the data, 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} percentiles to help capture the skewed nature of the speed distribution. Further, to better fit key features of the model, we calibrate using 12 moments associated with the speed distribution of white and minority motorists in daylight, 12 moments associated with the difference between the daylight and darkness speed at each percentile in the white and minority speed distributions. Similarly, we calibrate to one moment for the percentage (fraction times 100) of motorists stopped during the darkness who are minority and one moment for the VOD test statistic in Definition 1 again times 100, which is completely determined by the daylight and darkness fraction of stops that are minorities. In order 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, add the sample means back to the residuals so that each group of stopped motorists are compared for common observables. Finally, we convert these relative speeds back to miles per hour using the mode speed limit in each sample.

Theory does not provide guidance for establishing the weights on the moments. Our simulation is attempting to match 6 statistics: black and white daylight speed distribution, black and white daylight to darkness shift in the speed distribution, fraction stopped motorists minority in darkness and VOD test statistic. Equal weights with 6 statistics would imply a weight of 16.7 percent for each statistic. However, one might want to place more weight on the speed distribution statistics since they represent the sum of 6 individual moment squared deviations. On the other hand, we do not want to put too much weight on these moments since the number of moments is arbitrary based on the number of speed percentiles considered. For our baseline models, we place three times the weight on the speed distribution statistics so that the weight on those four are 21.5 percent each, and the weight on the fraction stopped motorists minority and VOD test statistic are 7 percent each. We also run robustness tests where we use an equal weight of 16.7 percent, and where we place six times the weight based on the 6 moments of the speed distribution for a weight of 23.25 percent for the four speed distribution statistics and 3.5 percent for both percent minority stopped and the VOD test statistic.

The functional forms above contain ten parameters shared by both white and minority simulated motorists. Mean, variance, and skewness of our preference distribution, and daylight stop-costs, must also be determined separately for white and minority motorists. Finally, we add a parameter for the fraction of motorists who are minority. We initialize the darkness stop-cost $s_{\underline{v}}$ to 44 allowing both the daylight stop-cost of both groups and the minimum return to a stop u_0 to vary relative to this fixed value. Finally, we must calibrate the fraction of minority motorists for the simulated population. Therefore, in total 18 free parameters are calibrated for each area. We minimize a mean squared error (MSE) optimization function of the weighted moments for all calibrations. Because the surface of this function is highly non-linear and appears to contain multiple local minima and inflections points, we first use a derivative-free Simplex-based optimization algorithm, Subplex (Rowan, 1990), to identify local minima. Once we have identified a local minimum, we use a second optimization routine based on quadratic approximations to the surface, BOBYQA (Powell 2009), to precisely locate that minimum and verify that the gradients over all parameters are approximately zero in this location. Finally, after identifying a specific local minimum that fits the data well, we will identify a global minimum using a modified evolutionary-based optimization routine, ESCH as described in da Silva Santos (2010) and accessed via an open source library for non-linear optimization (NLopt). The nature of evolutionary algorithms used for global optimization requires that limits be placed on the range of each parameter, and we use the information generated from the various local optimizations to place these limits. The specific limits for each parameter are shown in Appendix Table B1.[31]

We calibrate the parameters separately for the moments from Massachusetts, East Tennessee, and West Tennessee samples in a series of stages using the results of each stage as initial values in the next stage.

1. First, we focus on matching the minority daylight speed distribution while calibrating just distributional parameters, i.e. mean, variance, and skewness, while targeting an additional moment based on a specific positive fraction of minorities

[31] This second routine also requires that the analyst place limits on the parameter space, but this is a relatively non-restrictive process since we are simply refining an already identified local minimum. In practice, the search for the local minimum never crosses the bounds that we set on the parameters.

not infracting in daylight using the Simplex-based algorithm. Holding other parameters fixed at values that were found based on experimentation.

2. We next match both the daylight speed distribution and the difference between the daylight and darkness distributions for minorities additionally calibrating all motorist parameters that are common between groups plus the minority daylight stop-cost, i.e. $\alpha_1, \alpha_2, b_0, \mu,$ and τ_0. At this stage, we also drop the target on the fraction of minorities not speeding, which was simply used to anchor the initial calibration.
3. We then target all 26 moments and calibrate all 18 parameters. We first identify the local minimum using the simplex-based algorithm, but as mentioned above, we locate the local minimum precisely using quadratic approximations to the surface.
4. We repeat the process outlined in steps 1-3 for initial fractions minority not infraction in daylight between 0.05 and 0.40 in increments of 0.05 typically identifying different local minima for each percent not infracting value (even though that moment restriction is removed starting in step 2). We then identify the local minimum arising from an initial fraction not infracting moment restriction in step 1 that results in the lowest overall Mean Squared Error in step 3. We also verify that this minimum is internal to the range of fractions considered.
5. Finally, we use an evolutionary-based optimization routine using the best local optimum identified in step 4 and imposing parameter limits that were developed by observing the optimization over many possible local minima. Again, the quadratic approximation technique is used to precisely locate the minimum once the entropy-based routine has identified the minimum.

Note that the optimization also includes a penalty function starting below 2 percent of minority motorists not infracting in daylight in order to rule out corner solution equilibria where all motorists commit infractions. The final local and global optimums always imply a percent minority motorists not infracting above 2 percent so that the penalty function has no direct impact on the final optimum identified.

Table 12 presents the results of the calibration for minority motorists and the minority share variables with the first two columns presenting the empirical and the simulated moments for

Massachusetts and the next four columns presenting the same results for East and West Tennessee (majority motorist moments are shown in Appendix Table B2). At the bottom, the table also presents the fraction of motorists not infracting for minorities and majority in daylight and in darkness. The model does a very good job of matching both the daylight speed distribution and the change in the speed distribution between daylight and darkness. The model also closely matches both the fraction of stops in darkness that involve minority motorists and the VOD test statistic. The results for East Tennessee are notable in that the model fits both the empirical VOD test statistic that is just below one, and the shift in the speed distribution with stopped minority motorists at the upper speed percentiles driving substantially faster in darkness than the same percentile in daylight. The East Tennessee calibration quantitatively illustrates the potential for the VOD test to fail to identify racial differences in motorist stops as suggested earlier in Proposition 2. The calibrated parameters are shown in Appendix Table B3.

[Insert Table 12]

Table 13 summarizes relevant information from the calibration for considering the impact of race on police stop behavior. The first row presents the minority stop-cost in daylight, which is 0.006 in Massachusetts, 30.113 in East Tennessee, and 37.753 in West Tennessee all in comparison to a nighttime stop-cost of 44.0. Consistent with previous studies 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 of 44, during the relevant time period. Interestingly, we observe higher levels of prejudice in the calibration for East Tennessee as compared to West Tennessee, based on the strong shift in the speed distribution in East Tennessee even though the VOD test statistic for East Tennessee was near 1.0.

[Insert Table 13]

Further, if we interpret $u(i, s_d)$ as the return to a police stop, we can use the calibrated parameters for both police stop-costs and η, which captures non-linearities in return to police stops associated with the severity of the infraction, in order to compare the lower minority stop-costs in daylight to the utility gains 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 simulated mean level of infractions among stopped motorists. Specifically, we find the mean μ and standard deviation σ of the level of infraction, 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. Daylight when race is observable raises the returns of stop for minority motorists in Massachusetts by more than the effect of raising speed infraction by five standard deviations above the mean. In East Tennessee, daylight raises the return to stop for minority motorists by an amount comparable to a 2 standard deviation increase in infraction level, but in West Tennessee where we observe a smaller speed distribution shift daylight raises the return by about ½ a standard deviation.

The second panel of Table 13 presents the VOD test statistic from the calibration in the first row and a counterfactual or adjusted VOD test statistic that would have arisen in the calibration if minority motorists in daylight did not change their infraction behavior, i.e. behaved as if they faced the police costs for stops in darkness during daylight when they actually face higher stop-costs.

Definition 4. $K_{ADJ} \equiv \frac{\int_{c^*(s_{\underline{v}})}^{c_h} g(c,m)\phi^*(i'\left(c,s_{\underline{v}}\right),s_{\overline{v},m})di}{\int_{c^*(s_{\underline{v}})}^{c_h} g(c,w)\phi^*(i'\left(c,s_{\underline{v}}\right),s_{\overline{v},w})di} \frac{\int_{c^*(s_{\underline{v}})}^{c_h} g(c,w)\phi^*(i'\left(c,s_{\underline{v}}\right),s_{\underline{v}})di}{\int_{c^*(s_{\underline{v}})}^{c_h} g(c,m)\phi^*(i'\left(c,s_{\underline{v}}\right),s_{\underline{v}})di}$

where K_{ADJ} takes the same form as K_{ERS} except that c^* and i' in daylight depend upon the darkness police stop-cost. As noted earlier, the shift in the speed distribution increases with the calibrated value of the VOD test statistic so that the counterfactual VOD test statistic increases the most in Massachusetts from 1.38 to 2.74, the next most in East Tennessee increasing from 1.00 to 1.22, and exhibits the smallest increase in West Tennessee with a value of 1.17 relative to the calibrated value of 1.09. Again, the East Tennessee results are notable in that the counterfactual VOD test statistic is substantially larger than 1 even though the traditional VOD test found no evidence of police prejudice in East Tennessee.

Finally, we conduct a robustness test by modifying the weights above. The first panel of Table 14 presents the results from Table 13. The second panel applies an equal weight of 16.7 percent to the four speed distribution components and two moments based on the percent minority stopped. Recognizing that the speed distribution components contain six moments each, the third panel assigns approximately six times the weight to the four speed distribution components so that each of those receive a weight of 23.25 percent and the percent minority stopped based moments receive a weight of 3.5 percent each.[32] The basic results are relatively robust with similar daylight minority stop-costs across the three calibrations, and substantially larger VOD test statistics after adjusting for minority

[32] The calibrated parameters for these alternative weights are shown in Appendix Tables B4, B5 and B6 for the three sites.

driver changes in behavior. The magnitude of the adjusted VOD test statistics is notably sensitive to the weights only for West Tennessee. The largest adjusted VOD test statistic arises for the third panel where a larger weight is placed on matching the speed distribution shift, which makes sense since the baseline calibration understated the speed distribution shift in West Tennessee. Surprisingly, placing lower weight on the speed distribution contributions also increases the adjusted VOD test statistic. A better match to the VOD test statistic, which now has higher weight, requires lower police stop-costs for minorities in daylight, which may have increased the shift in the speed distribution even as the total fit of the speed distribution moments eroded due to having lower weight.

6. Conclusion

The Veil of Darkness test uses seasonal variation to compare the racial composition of police stops in daylight and darkness at the same time of day and has quickly become a gold standard for evaluating administrative data on police stops. This paper observes that even if the composition of motorists is the same between daylight and darkness the behavior of motorists may change under the alternative hypothesis of unequal treatment in traffic stops. If race is observable in daylight, but not in darkness, minority motorists might rationally choose to drive more conservatively and commit less infractions or less severe infractions in daylight if they anticipate being stopped for infractions at higher rates when race can be observed.

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 white motorist decreases by about one and a half percentage points in daylight relative to darkness. In the traffic stop data, we find a large shift in the speed distribution of African-Americans between daylight and darkness for Massachusetts with motorists stopped for speeding traveling about 5 percent slower in daylight relative to the speed limit. We find smaller differences of about 1.5 percent on average in East Tennessee with the smallest differences in West Tennessee of only 0.5 percent, although the results in East Tennessee are less precisely estimated due to the smaller sample. 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.

We develop and calibrate a simple theoretical model of police stop and motorist infraction behavior where in equilibrium some motorists may choose not to speed. The model matches the empirical moments in the data relatively well including capturing the empirical 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 of increasing the motorist speed above the speed limit by 5 standard deviations relative to the mean. These larger differences are consistent with both the high VOD test statistic and the large minority speed distribution change observed in Massachusetts. 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 the substantial shift in the minority speed distribution between daylight and darkness. While the statistical evidence of a speed distribution shift is weaker in East Tennessee due to the smaller sample size, the calibration results are strongly suggestive that VOD tests can fail to find evidence of discrimination in police stops due to motorist changes in behavior. In West Tennessee where the speed distribution shift is smaller, the VOD test statistic suggests more modest racial differences in treatment by police, and the calibrated differences in police preferences between daylight and darkness are relatively small.

In summary, the VOD test remains one of the best techniques available for providing convincing evidence of discrimination in police stops. However, this paper has documented substantial empirical evidence that minorities likely adjust their behavior in daylight to reflect actual or perceived police discrimination in stops when race can be observed. Our model calibrations suggest that the bias in the VOD test arising from changes in minority motorist behavior can be large, and in the case of East Tennessee this bias appears to have completely eliminated any observable evidence of discrimination in police stops. Going forward, states that collect data on traffic stops should attempt to collect objective information on the severity of the infraction where possible including when officers choose to issue warnings. Further, researchers that apply the VOD test should also use these techniques to test for differences in infraction severity that could be consistent with motorist adjustments that might bias the VOD test away from finding discrimination.

References

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2017). *When Should You Adjust Standard Errors for Clustering?* NBER working Paper (No. 24003). https://doi.org/10.3386/w24003

Anbarci, N., & Lee, J. (2014). Detecting racial bias in speed discounting: Evidence from speeding tickets in Boston. *International Review of Law and Economics*. https://doi.org/10.1016/j.irle.2014.02.002

Antonovics, K., & Knight, B. G. (2009). A new Look at racial profiling: Evidence from the Boston Police Department. *Review of Economics and Statistics*. https://doi.org/10.1162/rest.91.1.163

Anwar, S., Bayer, P., & Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *Quarterly Journal of Economics*. https://doi.org/10.1093/qje/qjs014

Anwar, S., & Fang, H. (2006). An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review*. https://doi.org/10.1257/000282806776157579

Arcidiacono, P., Bayer, P., & Hizmo, A. (2010). Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.2.4.76

Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*. https://doi.org/10.1093/qje/qjy012

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" A Market Test for Race Discrimination in Bail Setting. In *Stanford Law Review*. Retrieved from http://digitalcommons.law.yale.edu/fss_papers/1527

Bell, G. C., Hopson, M. C., Craig, R., & Robinson, N. W. (2014). Exploring Black and White Accounts of 21st-Century Racial Profiling: Riding and Driving While Black. *Qualitative Research Reports in Communication*. https://doi.org/10.1080/17459435.2014.955590

Bifulco, R., Fletcher, J. M., & Ross, S. L. (2011). The Effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1259570

Borgen, N. T. (2018). Fixed Effects in Unconditional Quantile Regression. *The Stata Journal*. https://doi.org/10.1177/1536867x1601600208

Brent, R. P. (1973). An Algorithm with Guaranteed Convergence for Finding a Zero of a Function. In *Algorithms for Minimization without Derivatives*. Englewood Cliffs, NJ: Prentice-Hall.

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*.

Cheng, C., Long, W., & Long,) Wei. (2018). *The Spillover Effects of Highly Publicized Police-Related Deaths on Policing and Crime: Evidence from Large U.S. Cities*.

Coate, S., & Loury, G. (1993). Will Affirmative-Action Policies Eliminate Negative Stereotypes? *American Economic Review*, *83*(5), 1220–1240.

Coviello, D., & Persico, N. (2016). An Economic Analysis of Black-White Disparities in the New York Police Department's Stop-and-Frisk Program. *The Journal of Legal Studies*. https://doi.org/10.1086/684292

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

Dedman, B., & Latour, F. (2003). Speed Trap: Who Gets a Ticket and Who Gets a Break? *The Boston Globe*.

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.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). UNCONDITIONAL QUANTILE REGRESSIONS. *Econometrica*, *77*(3), 953–973. https://doi.org/10.3982/ECTA6822

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*. https://doi.org/10.1086/701423

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

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*. Retrieved from www.washingtonpost. com/sf/investigative/2015/12/26/a-yearof-reckoning-police-fatally-shoot-nearly1000/

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

Goncalves, F., & Mello, S. (2017). *Does the Punishment Fit the Crime? Speeding Fines and Recidivism*. Retrieved from 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*. https://doi.org/10.1198/016214506000000168

Heckman, J. J. (1998). Detecting Discrimination. *Journal of Economic Perspectives*. https://doi.org/10.1257/jep.12.2.101

Horrace, W. C., & Rohlin, S. M. (2016). How dark is dark?: Bright lights, big city, racial profiling. *Review of Economics and Statistics*. https://doi.org/10.1162/REST_a_00543

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*. https://doi.org/10.1257/pandp.20191026

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. https://doi.org/10.1257/aer.101.4.1467

Knowles, J., & Persico, N. (2001). Racial Bias in Motor Vehicles: Theory and Evidence. *Journal of Politcal Economy*. https://doi.org/10.1086/318603

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*. https://doi.org/10.1016/j.jcrimjus.2007.01.004

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,.

Lundberg, S. J. (2006). The Enforcement of Equal Opportunity Laws Under Imperfect Information: Affirmative Action and Alternatives. *The Quarterly Journal of Economics*. https://doi.org/10.2307/2937919

Lundberg, S. J., & Start, R. (1983). Private Discrimination and Social Intervention in Competitive Labor Markets. *American Economic Review*. https://doi.org/10.2307/1808117

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

Masher, J. (2016). What The "Veil of Darkness" Says About New Orleans Traffic Stops. Retrieved February 22, 2017, from 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*.

McCullagh, P. (Peter), & Nelder, J. A. (n.d.). *Generalized linear models*.

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*. https://doi.org/10.1016/j.jnma.2017.12.002

Milyo, J. (n.d.). *DWB in COMO: Understanding Race Disparities in Vehicle Stops and Searches*. Retrieved from 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*.

Moro, A., & Norman, P. (1974). Note on an extension of the Ramsey inverse elasticity of demand pricing or taxation formula. *Journal of Public Economics*. https://doi.org/10.1016/0047-2727(74)90034-6

National Research Council. (2004). Fairness and Effectiveness in Policing: The Evidence. In *Fairness and Effectiveness in Policing*. https://doi.org/10.17226/10419

Neal, D. A., & Johnson, W. R. (1996). The Role of Premarket Factors in Black-White Wage Differences. In *Source: Journal of Political Economy* (Vol. 104).

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*. https://doi.org/10.1111/1745-9133.12269

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

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

Ridgeway, G., Schell, T., Riley, K. J., Turner, S., & Dixon, T. L. (2006). *Police-community Relations in Cincinnati: Year Two Evaluation Report*.

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

Ritter, J. A., & Bael, D. (2009). *Detecting Racial Profiling in Minneapolis Traffic Stops: A New Approach. Center for Urban and Regional Affairs*.

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*. https://doi.org/10.1016/j.jebo.2017.02.005

Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2015). *State of Connecticut Traffic Stop Data Analysis and Findings, 2013-14*.

Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2016). *State of Connecticut Traffic Stop Data Analysis and Findings, 2014-15*.

Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2019). *State of Rhode Island Traffic Stop Data Analysis and Findings, 2017*.

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, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2018). *State of Connecticut Traffic Stop Data Analysis and Findings, 2016-17*.

Ross, M. B., Fazzalaro, J., Barone, K., & Kalinowski, J. (2017). *State of Connecticut Traffic Stop Data Analysis and Findings, 2015-16*.

Ross, S. L., & Yinger, J. (1999). The Default Approach to Studying Mortgage Discrimination: A Rebuttal. In *Mortgage lending discrimination: A Review of Existing Evidence*. Washington, DC: Urban Institute.

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

Schell, T., Ridgeway, G., Dixon, T. L., Turner, S., & Riley, K. J. (2007). *Police-community Relations in Cincinnati: Year Three Evaluation Report*.

Smith, A. C. (2016). Spring forward at your own risk: Daylight saving time and fatal vehicle crashes. *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20140100

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

Swaine, J., Laughland, O., Lartey, J., & Ciara, M. (2015, December 31). Young black men killed by US police at highest rate in year of 1,134 deaths. *The Guardian*. Retrieved from www.theguardian.com/usnews/2015/dec/31/the-counted-policekillings-2015-young-Black-men

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*.

Worden, R. E., McLean, S. J., & Wheeler, A. P. (2010). *Stops by Syracuse Police, 2006-2009*.

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

Tables and Figures

Table 1: Descriptive Statistics for the FARS Accident Data

Total Accidents		615,826		
Fatal Accidents		282,924		
Inter-Twilight		39,076		
Sample		All	Black	White
Daylight		53.44%	49.93%	53.95%
Motorist	Black	12.83%	100.00%	0.00%
	Male	67.67%	72.22%	66.99%
	Young	42.74%	38.92%	43.31%
Auto.	Domestic	66.36%	62.25%	66.97%
	Old	22.05%	19.10%	22.48%
Day of Week	Sunday	14.03%	15.19%	13.86%
	Monday	13.49%	12.96%	13.57%
	Tuesday	12.91%	11.49%	13.11%
	Wednesday	13.50%	12.90%	13.59%
	Thursday	14.01%	13.52%	14.08%
	Friday	16.52%	16.15%	16.58%
	Saturday	15.54%	17.79%	15.21%
Hour of Day	4:00 PM	5.70%	3.23%	6.07%
	5:00 PM	22.97%	21.79%	23.14%
	6:00 PM	24.83%	24.83%	24.83%
	7:00 PM	21.53%	23.83%	21.19%
	8:00 PM	18.06%	19.64%	17.82%
	9:00 PM	4.87%	3.93%	5.01%
States + DC		49	49	49

Note: The overall sample includes only traffic stops involving black or Non-Hispanic white motorists.

Table 2: Estimated Change in the Accidents Rate for Minority Motorists in Daylight

LHS: Black		(1)	(2)	(3)	(4)
Daylight		-0.01752*** (0.00412)	-0.01663*** (0.00392)	-0.01566*** (0.00399)	-0.01525*** (0.00398)
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
Observations		39063	39063	39063	39063

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 black or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017. The sample is further restricted to include only accidents 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.

Table 3: Heterogenous Changes to Accident Rates

LHS: Black		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Daylight		-0.01554***	-0.00911*	-0.00691	-0.00620	-0.01798***	-0.01606***	-0.01508***	-0.01463***
		(0.00571)	(0.00557)	(0.00581)	(0.00578)	(0.00416)	(0.00395)	(0.00404)	(0.00402)
Black/White Odds Police Shooting		-0.01628***							
		(0.00209)							
Daylight x Black/White Odds Police Shooting		-0.00193	-0.00356**	-0.00415***	-0.00429***				
		(0.00150)	(0.00150)	(0.00159)	(0.00158)				
Racism Index						-0.01287***			
						(0.00369)			
Daylight x Racism Index						-0.00393	-0.00684**	-0.00690**	-0.00729***
						(0.00291)	(0.00281)	(0.00283)	(0.00282)
Controls	Hour of Day	X	X	X		X	X	X	X
	Day of Week	X	X	X		X	X	X	X
	Year	X	X			X	X		
	State		X				X		
	State x Year			X	X			X	X
	Motorist/Vehicle				X				X
Observations		39076	39076	39076	39076	39076	39076	39076	39076

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 black or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017. The sample is further restricted to include only accidents 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.

Table 4: Change to Accident Rates in States that Failed the VOD

LHS: Black		(1)	(2)	(3)	(4)
Daylight		-0.04642*** (0.01217)	-0.03559*** (0.01085)	-0.03324*** (0.01080)	-0.03381*** (0.01071)
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
Observations		6587	6587	6587	6587

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 black or Non-Hispanic white motorists which occurred within the ITW from 2000 to 2017 in a subsample of U.S. states with a jurisdiction that conducted a VOD study that failed to reject the null. These states include Arizona, California, Connecticut, Louisiana, Missouri, North Carolina, Ohio, Oregon, and Rhode Island. The sample is further restricted to include only accidents 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.

Table 5: 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. The sample is further restricted to include only accidents 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 confidence level of 77.82 percent.

Table 6: Descriptive Statistics for Massachusetts and Tennessee Traffic Stop Data

		MA			East TN			West TN		
Total Stops		401,408			489,313			1,658,611		
Speeding Stops		80,471			143,014			541,667		
Inter-Twilight		10,203			23,515			102,054		
Sample		All	Black	White	All	Black	White	All	Black	White
Daylight		66.75%	71.05%	65.78%	64.94%	67.59%	68.63%	68.44%	63.15%	65.03%
Motorist	Black	18.32%	100.00%	0.00%	4.64%	100.00%	0.00%	18.60%	100.00%	0.00%
	Male	72.76%	69.82%	73.42%	65.59%	58.54%	62.07%	61.41%	73.33%	65.21%
	Young	51.97%	50.62%	52.27%	-	-	-	-	-	-
Auto.	Domestic	32.81%	28.62%	33.75%	31.80%	34.12%	36.77%	36.28%	30.61%	31.85%
	Old	49.74%	50.62%	49.54%	-	-	-	-	-	-
	Red	10.35%	11.88%	10.01%	-	-	-	-	-	
Day of Week	Sunday	14.71%	13.48%	14.99%	11.98%	16.04%	12.98%	13.55%	14.67%	11.85%
	Monday	13.84%	12.04%	14.24%	14.26%	12.96%	13.30%	13.23%	16.04%	14.17%
	Tuesday	15.02%	15.78%	14.84%	12.59%	11.50%	13.03%	12.75%	10.54%	12.69%
	Wednesday	13.50%	13.43%	13.51%	13.27%	11.48%	13.54%	13.16%	11.82%	13.34%
	Thursday	14.09%	15.84%	13.70%	13.62%	12.99%	14.10%	13.89%	11.73%	13.71%
	Friday	14.30%	12.68%	14.66%	20.46%	18.92%	19.58%	19.46%	19.62%	20.50%
	Saturday	14.54%	16.75%	14.05%	13.81%	16.11%	13.48%	13.97%	15.58%	13.73%
Hour of Day	5:00 PM	36.84%	33.87%	37.51%	26.37%	22.69%	23.97%	23.73%	24.84%	26.44%
	6:00 PM	34.00%	38.26%	33.05%	23.53%	27.29%	28.94%	28.63%	21.91%	23.61%
	7:00 PM	16.29%	17.50%	16.02%	20.61%	22.57%	21.90%	22.02%	20.81%	20.60%
	8:00 PM	12.87%	10.38%	13.43%	17.38%	15.65%	14.53%	14.74%	19.62%	17.27%
	9:00 PM				12.12%	11.79%	10.67%	10.87%	12.83%	12.08%
Counties/Towns		18			13			44		

Note: The overall sample includes only traffic stops involving black or Non-Hispanic white motorists. MA is used in this and the following tables as an abbreviation for Massachusetts and TN is used in the following tables as an abbreviation for Tennessee.

Table 7: Canonical Veil of Darkness Estimates

LHS: Black		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		MA		East TN			West TN		
Daylight		0.0458**	0.0441**	-0.00116	-0.000921	-0.000287	0.0105***	0.00972**	0.0119***
		(0.0185)	(0.0193)	(0.00397)	(0.00395)	(0.00396)	(0.00384)	(0.00382)	(0.00372)
Controls	Day of Week	X	X	X	X	X	X	X	X
	Time of Day	X	X	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		X	X
	County x Year					X			X
Observations		10203	10203	23515	23515	23515	102054	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 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 involving black 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 8: Estimated Change in Mean Speed for Stopped Minority Motorists in Daylight

LHS: Rel. Speed		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		MA		East TN			West TN		
Daylight		0.591	0.716	-0.0716	-0.0680	0.120	-0.0606	-0.0429	0.0452
		(1.428)	(1.472)	(0.433)	(0.433)	(0.413)	(0.176)	(0.176)	(0.173)
Black		1.887*	1.771	-1.646**	-1.649**	-1.605**	0.351*	0.387*	0.427**
		(1.049)	(1.057)	(0.711)	(0.714)	(0.707)	(0.210)	(0.209)	(0.203)
Daylight*Black		-5.416**	-5.340**	-1.477*	-1.477*	-1.192	-0.498**	-0.505**	-0.502**
		(1.889)	(1.857)	(0.816)	(0.816)	(0.803)	(0.220)	(0.221)	(0.219)
Controls	Day of Week	X	X	X	X	X	X	X	X
	Time of Day	X	X	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		X	X
	County x Year					X			X
Observations		10203	10203	23515	23515	23515	102054	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 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 involving black 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. Relative speed is calculated as speed relative to the limit and multiplied by one hundred.

Table 9: 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)
	Black	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*Black	-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)
	Black	-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*Black	-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)
	Black	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*Black	-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) but robust in Tennessee to clustering on county and year separately and robust in Massachusetts to clustering by town. Bootstrapping one-thousand random samples, we find that the p-value for a one-sided permutation test of joint significance on all nine quantiles is equal to 1.3 percent for Massachusetts, 0.5 percent for East Tennessee, and 0.1 percent for West Tennessee. The sample includes only traffic stops involving black 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 10: Falsification Test over Gender with White Motorists

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-SP	Daylight	1.175	0.555	1.663	2.073	-0.137	-0.229	-1.326	-1.182	-2.187
		(1.089)	(1.516)	(1.340)	(1.566)	(2.013)	(2.616)	(2.793)	(3.466)	(3.997)
	Male	1.202	1.340	1.379**	1.474*	0.980	1.242	1.605	3.377*	6.752*
		(0.799)	(0.901)	(0.618)	(0.776)	(1.050)	(1.386)	(1.594)	(1.873)	(3.705)
	Daylight*Male	-1.265	-0.516	-0.267	-0.752	-0.226	0.0118	0.850	0.179	0.348
		(0.847)	(0.789)	(0.801)	(1.061)	(1.293)	(1.628)	(1.842)	(2.076)	(4.474)
	Obs.	8334	8334	8334	8334	8334	8334	8334	8334	8334
East TN-HP	Daylight	0.106	0.157	-0.344	-0.353	-0.283	-0.00306	0.388	0.540	0.644
		(1.288)	(1.005)	(0.539)	(0.344)	(0.346)	(0.451)	(0.681)	(1.049)	(1.455)
	Male	-0.396	-0.823**	-0.546*	-0.483*	-0.314	-0.0823	0.0747	0.121	2.262*
		(0.602)	(0.359)	(0.254)	(0.227)	(0.189)	(0.267)	(0.284)	(0.646)	(1.202)
	Daylight*Male	-0.471	-0.186	0.260	0.314*	0.110	-0.0347	-0.265	-0.306	-1.004
		(0.493)	(0.353)	(0.274)	(0.158)	(0.336)	(0.431)	(0.513)	(0.841)	(1.314)
	Obs.	22424	22424	22424	22424	22424	22424	22424	22424	22424
West TN-HP	Daylight	0.0645	0.140	-0.0757	-0.0724	-0.205	0.135	0.168	0.421	0.102
		(0.164)	(0.252)	(0.130)	(0.160)	(0.134)	(0.219)	(0.304)	(0.408)	(0.543)
	Male	-0.348***	-0.543***	-0.189	0.0792	0.0998	0.403**	0.633***	1.206***	1.544***
		(0.0957)	(0.162)	(0.137)	(0.145)	(0.125)	(0.172)	(0.171)	(0.254)	(0.319)
	Daylight*Male	-0.00342	0.0480	-0.00976	-0.0928	-0.0347	-0.307	-0.285	-0.779**	-0.285
		(0.151)	(0.197)	(0.193)	(0.193)	(0.160)	(0.187)	(0.219)	(0.333)	(0.452)
	Obs.	83076	83076	83076	83076	83076	83076	83076	83076	83076

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) but robust in Tennessee to clustering on county and year separately and robust in Massachusetts to clustering by town. Bootstrapping one-thousand random samples, we find that the p-value for a one-sided permutation test of joint significance on all nine quantiles is equal to 89.2 percent for Massachusetts, 58.7 percent for East Tennessee, and 36.8 percent for West Tennessee. The sample includes only traffic stops involving 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 11: Falsification Test over Vehicle Type with White Motorists

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-SP	Daylight	0.308	0.349	1.778	1.433	-0.534	-0.438	-0.828	-1.697	-1.679
		(1.008)	(1.194)	(1.069)	(1.199)	(1.452)	(1.587)	(2.018)	(3.276)	(3.805)
	Domestic	0.404	0.231	0.212	-0.924	-1.433*	-1.344	-1.548	-3.115	-0.370
		(0.935)	(0.695)	(0.570)	(0.627)	(0.684)	(0.775)	(1.085)	(1.991)	(2.245)
	Daylight*Domestic	-0.147	-0.554	-0.910	0.118	0.654	0.576	0.280	1.457	-1.389
		(1.158)	(1.234)	(0.914)	(1.098)	(1.007)	(0.942)	(1.302)	(2.182)	(2.221)
	Obs.	8334	8334	8334	8334	8334	8334	8334	8334	8334
East TN-HP	Daylight	-0.231	0.253	-0.0583	-0.141	-0.145	0.0218	0.356	0.513	0.447
		(1.012)	(0.942)	(0.526)	(0.417)	(0.453)	(0.485)	(0.718)	(0.974)	(1.375)
	Domestic	0.492	0.866	0.425*	0.259	0.358	0.476	0.598	0.619	1.410
		(0.411)	(0.490)	(0.238)	(0.245)	(0.251)	(0.316)	(0.538)	(0.865)	(1.489)
	Daylight*Domestic	0.104	-0.655	-0.338	0.00464	-0.192	-0.146	-0.451	-0.556	-1.559
		(0.649)	(0.718)	(0.302)	(0.323)	(0.303)	(0.372)	(0.532)	(0.866)	(1.711)
	Obs.	22424	22424	22424	22424	22424	22424	22424	22424	22424
West TN-HP	Daylight	0.0314	0.172	-0.0461	-0.0806	-0.206*	-0.0790	0.0216	-0.109	-0.190
		(0.126)	(0.227)	(0.113)	(0.139)	(0.112)	(0.176)	(0.232)	(0.330)	(0.451)
	Domestic	0.341***	0.613***	0.500***	0.657***	0.541***	0.702***	0.829***	0.936***	0.682*
		(0.118)	(0.177)	(0.125)	(0.179)	(0.134)	(0.198)	(0.283)	(0.325)	(0.381)
	Daylight*Domestic	0.131	0.0655	-0.0621	-0.119	-0.0466	0.0600	-0.105	0.0670	0.208
		(0.127)	(0.182)	(0.144)	(0.176)	(0.131)	(0.221)	(0.288)	(0.311)	(0.411)
	Obs.	83076	83076	83076	83076	83076	83076	83076	83076	83076

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) but robust in Tennessee to clustering on county and year separately and robust in Massachusetts to clustering by town. Bootstrapping one-thousand random samples, we find that the p-value for a one-sided permutation test of joint significance on all nine quantiles is equal to 85.7 percent for Massachusetts, 72.3 percent for East Tennessee, and 91.1 percent for West Tennessee. The sample includes only traffic stops involving 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 12: Calibration Results

	Massachusetts		East Tennessee		West Tennessee	
	Data	Simulation	Data	Simulation	Data	Simulation
Black Speed Distribution Daylight						
20th Percentile	13.3835	13.3344	12.1763	12.8364	11.5419	11.1629
40th Percentile	14.8568	15.5537	15.2878	14.9840	13.5632	13.4064
60th Percentile	18.3094	18.5776	17.8090	17.5080	15.7624	15.8768
80th Percentile	23.9418	23.6617	20.7542	21.5682	19.3593	19.5268
90th Percentile	28.4176	28.5276	25.3446	25.0375	22.8966	23.1396
95th Percentile	33.3009	33.1861	28.6582	28.4432	26.8071	26.5863
Difference Daylight and Darkness						
20th Percentile	-0.0899	0.2458	0.5273	-0.2437	0.0830	0.3135
40th Percentile	2.2735	2.3309	0.3049	0.3068	0.2845	0.4916
60th Percentile	3.8130	3.5311	0.7094	0.6952	0.4227	0.5726
80th Percentile	4.7673	4.6565	1.6052	0.9309	0.4644	0.6492
90th Percentile	5.6541	5.1206	1.2012	1.3029	0.7023	0.6702
95th Percentile	5.1922	5.3577	1.3253	1.4398	1.0512	0.7007
Minority Share of Stops						
Minority Share of Stops Darkness	0.1664	0.1665	0.0466	0.0466	0.1771	0.1771
VOD Test Statistic	1.3769	1.3793	0.9924	0.9973	1.0908	1.0899
Percent Minority						
Motorists	NA	0.1638	NA	0.0552	NA	0.1771
Not Infracting in Daylight	NA	0.4959	NA	0.3197	NA	0.0773
Not Infracting in Darkness	NA	0.0056	NA	0.1673	NA	0.0063

Notes: Empirical speed distribution in miles per hour based on regressing relative speed on day of week, time of day, geographic and for Tennessee year controls, calculating the residual, adding the means of the controls back to the sample and then calculating the miles per hour based on the mode speed limit of traffic stops for each site. The simulated moments arise from the global optimum identified by applying an evolutionary based optimization routine called ESCH and precisely located by applying second optimization routine based on quadratic approximations to the surface BOBYQA. The calibrated parameters used to calculate these moments are shown in Appendix Table B2.

Table 13: 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 nighttime stop-cost that was fixed at 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 if the statistic implied by the simulated speed distributions in each site 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.

Table 14: Calibration Results Related to Racial Differences in Police Stop Behavior

	Massachusetts	East Tennessee	West Tennessee
	Original Weights		
Minority Stop Cost Diff	43.994	13.887	6.247
Simulated VOD Test	1.379	0.997	1.090
Adjusted VOD Test	2.736	1.223	1.173
	Equal Weights		
Minority Stop Cost Diff	43.994	13.9996	10.125
Simulated VOD Test	1.38	0.994	1.091
Adjusted VOD Test	2.736	1.226	1.271
	Speed Moments Times Six		
Minority Stop Cost Diff	43.979	12.348	12.38
Simulated VOD Test	1.38	1	1.09
Adjusted VOD Test	2.736	1.195	1.338

Notes. The first panel repeats the results from Table 12 using the original weights. The second panel presents results where the four speed distributions receive the same weight as each of the moments associated with percent minority stopped. The third panel presents results where the speed component contribution receives six times the weight because those components contain the mean squared error for six distinct moments.

Appendix A. Empirical Appendix

Table A.1: Estimated Change in the Accidents Rate for Minority Motorists in Daylight, USNO Daylight Definition

LHS: Black		(1)	(2)	(3)	(4)
Daylight (USNO)		-0.01107*** (0.00413)	-0.01019*** (0.00389)	-0.00986*** (0.00392)	-0.00960*** (0.00391)
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
Observations		39063	39063	39063	39063

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 black or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017. The sample is further restricted to include only accidents involving at least one or more non-commercial automobiles (no motorcycle or pedestrian). Observations are weighted by the inverse likelihood of experiencing a fatality, estimated using detailed vehicular characteristics and restraint use.

Table A.2: Heterogenous Changes to Accident Rates, USNO Daylight Definition

LHS: Black	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Daylight (USNO)	-0.00814	-0.00177	-0.00066	-0.00011	-0.01191***	-0.00978***	-0.00946**	-0.00916**
	(0.00576)	(0.00550)	(0.00565)	(0.00563)	(0.00418)	(0.00393)	(0.00396)	(0.00395)
Black/White Odds Police Shooting	-0.01575***							
	(0.00213)							
Daylight (USNO) x Black/White Odds Police Shooting	-0.00268*	-0.00399***	-0.00437***	-0.00451***				
	(0.00152)	(0.00146)	(0.00152)	(0.00151)				
Racism Index					-0.01319***			
					(0.00372)			
Daylight (USNO) x Racism Index					-0.00287	-0.00526*	-0.00519*	-0.00564**
					(0.00286)	(0.00275)	(0.00277)	(0.00276)
Controls: Hour of Day	X	X	X		X	X	X	X
Controls: Day of Week	X	X	X		X	X	X	X
Controls: Year	X	X			X	X		
Controls: State		X				X		
Controls: State x Year			X	X			X	X
Controls: Motorist/Vehicle				X				X
Observations	39076	39076	39076	39076	39076	39076	39076	39076

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 black or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017. The sample is further restricted to include only accidents involving at least one or more non-commercial automobiles (no motorcycle or pedestrian). Observations are weighted by the inverse likelihood of experiencing a fatality, estimated using detailed vehicular characteristics and restraint use.

Table A.3: Change to Accident Rates in States that Failed to Reject the Null, USNO Daylight Definition

LHS: Black		(1)	(2)	(3)	(4)
Daylight (USNO)		-0.04334*** (0.01162)	-0.03245*** (0.01052)	-0.03235*** (0.01037)	-0.03277*** (0.01031)
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
Observations		6587	6587	6587	6587

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 black or Non-Hispanic white motorists which occurred within the ITW from 2000 to 2017 in a subsample of U.S. states with a jurisdiction that conducted a VOD study that failed to reject the null. These states include Arizona, California, Connecticut, Louisiana, Missouri, North Carolina, Ohio, Oregon, and Rhode Island. The sample is further restricted to include only accidents 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.

Table A.4: Balancing Test of Accidents for White Motorists within the ITW, USNO Daylight Definition

LHS: Daylight (USNO)		(1)	(2)	(3)	(4)	(5)
Domestic Vehicle		0.00155				0.00210
		(0.00491)				(0.00493)
Vehicle Age			-0.00747			-0.00767
			(0.00555)			(0.00555)
Male Motorist				-0.00987**		-0.01014**
				(0.00482)		(0.00484)
Young Motorist					-0.00052	-0.00068
					(0.00463)	(0.00463)
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. The sample is further restricted to include only accidents 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 1.61 and a confidence level of 83.98 percent.

Table A.5: Estimated Change in the Accidents Rate for Minority Motorists in Daylight, Fatality Risk Weighted

LHS: Black		(1)	(2)	(3)	(4)
Daylight		-0.00879 (0.00690)	-0.01296** (0.00634)	-0.01328** (0.00648)	-0.01353** (0.00650)
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
Observations		39063	39063	39063	39063

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 black or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017. The sample is further restricted to include only accidents 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.

Table A.6: Heterogenous Changes to Accident Rates, Fatality Risk Weighted

LHS: Black		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Daylight		-0.01547***	-0.01110*	-0.00917	-0.00861	-0.01623***	-0.01564***	-0.01522***	-0.01485***
		(0.00590)	(0.00584)	(0.00604)	(0.00602)	(0.00437)	(0.00421)	(0.00428)	(0.00428)
Black/White Odds Police Shooting		-0.01603***							
		(0.00199)							
Daylight x Black/White Odds Police Shooting		-0.00107	-0.00240	-0.00312*	-0.00323**				
		(0.00156)	(0.00157)	(0.00164)	(0.00163)				
Racism Index						-0.01532***			
						(0.00344)			
Daylight x Racism Index						-0.00261	-0.00509*	-0.00496*	-0.00551*
						(0.00300)	(0.00289)	(0.00293)	(0.00291)
Controls	Hour of Day	X	X	X		X	X	X	X
	Day of Week	X	X	X		X	X	X	X
	Year	X	X			X	X		
	State		X				X		
	State x Year			X	X			X	X
	Motorist/Vehicle				X				X
Observations		39076	39076	39076	39076	39076	39076	39076	39076

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 black or Non-Hispanic white motorists which occurred within the ITW in the contiguous U.S. from 2000 to 2017. The sample is further restricted to include only accidents 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.

Table A.7: Change to Accident Rates in States that Failed to Reject the Null, Fatality Risk Weighted

LHS: Black		(1)	(2)	(3)	(4)
Daylight		-0.04623*** (0.01242)	-0.03683*** (0.01082)	-0.03601*** (0.01077)	-0.03640*** (0.01072)
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
Observations		6587	6587	6587	6587

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 black or Non-Hispanic white motorists which occurred within the ITW from 2000 to 2017 in a subsample of U.S. states with a jurisdiction that conducted a VOD study that failed to reject the null. These states include Arizona, California, Connecticut, Louisiana, Missouri, North Carolina, Ohio, Oregon, and Rhode Island. The sample is further restricted to include only accidents 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.

Table A.8: Balancing Test of Accidents for White Motorists within the ITW, Fatality Risk Weighted

LHS: Daylight		(1)	(2)	(3)	(4)	(5)
Domestic Vehicle		0.00430 (0.00572)				0.00465 (0.00575)
Vehicle Age			-0.00767 (0.00625)			-0.00749 (0.00625)
Male Motorist				-0.00399 (0.00526)		-0.00443 (0.00528)
Young Motorist					0.00431 (0.00518)	0.00420 (0.00518)
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. The sample is further restricted to include only accidents involving at least one or more non-commercial automobiles (no motorcycle or pedestrian). Observations are weighted by the inverse 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 is0.93 and a confidence level of 55.65 percent.

Table A.9: Canonical Veil of Darkness Estimates, Logit

LHS: Black		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		MA		East TN			West TN		
Daylight		0.409***	0.416***	0.0104	-0.0150	0.00300	0.0706**	0.0637**	0.0817***
		(0.0703)	(0.0989)	(0.0958)	(0.0943)	(0.0981)	(0.0289)	(0.0288)	(0.0286)
Controls	Day of Week	X	X	X	X	X	X	X	X
	Time of Day	X	X	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		X	X
	County x Year					X			X
Observations		10203	10203	23515	23515	23515	102054	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 (TN) and patrol districts (MA) but robust to clustering on county and year separately (TN), patrol district (TN), or town (MA). The sample includes only traffic stops involving black or Non-Hispanic white motorists. The two Tennessee samples also include controls for year in the first two specifications of each panel.

Table A.10: Estimated Change in Speed Distribution for Stopped Minority Motorists in Daylight, Demographic Controls

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	0.0811	0.196	1.938	0.727	-0.318	-0.0289	0.482	-0.560	-1.260
		(1.117)	(1.174)	(1.260)	(0.931)	(1.128)	(1.251)	(1.389)	(2.252)	(2.918)
	Black	0.712	0.555	2.479*	2.164***	1.467**	1.601*	1.262	5.639***	6.154**
		(1.054)	(1.053)	(1.243)	(0.738)	(0.687)	(0.860)	(1.669)	(1.801)	(2.891)
	Daylight*Black	-0.324	-0.221	-1.683	-2.230**	-5.000**	-6.748**	-7.616**	-10.79***	-11.93**
		(1.308)	(1.311)	(1.392)	(1.011)	(1.974)	(2.624)	(2.680)	(2.739)	(4.133)
	Obs.	10203	10203	10203	10203	10203	10203	10203	10203	10203
East TN	Daylight	-0.206	-0.00313	-0.190	-0.126	-0.100	-0.0471	0.181	0.211	-0.0905
		(1.084)	(0.800)	(0.469)	(0.351)	(0.337)	(0.413)	(0.566)	(0.836)	(1.118)
	Black	-2.032**	-1.751*	-1.206	-0.872	-0.767	-1.024	-0.163	-1.014	-1.887
		(0.805)	(0.930)	(0.694)	(0.803)	(0.814)	(1.062)	(1.395)	(1.833)	(2.142)
	Daylight*Black	-1.196	-1.416**	-1.081**	-0.977**	-1.431***	-1.573**	-3.068**	-2.231	-2.084
		(0.845)	(0.635)	(0.468)	(0.363)	(0.435)	(0.588)	(1.085)	(1.547)	(2.265)
	Obs.	23515	23515	23515	23515	23515	23515	23515	23515	23515
West TN	Daylight	0.0812	0.163	-0.0766	-0.132	-0.198*	-0.0350	0.0227	-0.139	-0.126
		(0.121)	(0.213)	(0.109)	(0.138)	(0.108)	(0.181)	(0.249)	(0.340)	(0.471)
	Black	0.188	0.614**	0.686**	0.691*	0.316	0.696	0.711	0.600	0.244
		(0.246)	(0.270)	(0.257)	(0.377)	(0.343)	(0.426)	(0.514)	(0.608)	(0.728)
	Daylight*Black	-0.111	-0.191	-0.554***	-0.876***	-0.543***	-0.853***	-1.005***	-0.812	-0.951
		(0.151)	(0.200)	(0.171)	(0.257)	(0.187)	(0.244)	(0.329)	(0.511)	(0.669)
	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) but robust in Tennessee to clustering on county and year separately and robust in Massachusetts to clustering by town. Bootstrapping one-thousand random samples, we find that the p-value for a one-sided permutation test of joint significance on all nine quantiles is equal to 1.4 percent for Massachusetts, 0.4 percent for East Tennessee, and 0.1 percent for West Tennessee. The sample includes only traffic stops involving black or Non-Hispanic white motorists. Controls include observed motorist and vehicle attributes, 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 A.11: Estimated Change in Speed Distribution for Stopped Minority Motorists in Daylight, County by Year and Demographic Controls

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
East TN	Daylight	0.372	0.327	0.0145	0.00964	0.0284	0.0854	0.374	0.459	-0.100
		(0.695)	(0.454)	(0.274)	(0.273)	(0.318)	(0.354)	(0.461)	(0.699)	(0.991)
	Black	-2.008	-1.807**	-1.190**	-0.877	-0.780	-0.989	-0.180	-1.187	-1.958
		(1.236)	(0.868)	(0.551)	(0.535)	(0.614)	(0.711)	(0.892)	(1.322)	(1.682)
	Daylight*Black	-0.822	-1.023	-0.812	-0.700	-1.113	-1.324*	-2.757***	-1.697	-1.684
		(1.334)	(0.980)	(0.629)	(0.574)	(0.681)	(0.738)	(0.989)	(1.486)	(2.028)
	Obs.	23515	23515	23515	23515	23515	23515	23515	23515	23515
West TN	Daylight	0.153	0.277*	0.0104	-0.0326	-0.0999	0.0761	0.138	-0.0119	-0.0466
		(0.108)	(0.156)	(0.108)	(0.146)	(0.121)	(0.171)	(0.235)	(0.296)	(0.397)
	Black	0.193	0.600***	0.665***	0.685***	0.331*	0.738***	0.754**	0.637*	0.319
		(0.131)	(0.164)	(0.135)	(0.200)	(0.180)	(0.240)	(0.309)	(0.347)	(0.502)
	Daylight*Black	-0.115	-0.214	-0.538***	-0.865***	-0.543***	-0.854***	-1.015***	-0.781**	-0.903
		(0.146)	(0.193)	(0.153)	(0.218)	(0.184)	(0.265)	(0.356)	(0.395)	(0.572)
	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) but robust to clustering on county and year separately. Bootstrapping one-thousand random samples, we find that the p-value for a one-sided permutation test of joint significance on all nine quantiles is equal to 4 percent for East Tennessee and 0.001 percent for West Tennessee. The sample includes only traffic stops involving black or Non-Hispanic white motorists. Controls include observed motorist and vehicle attributes, time of day, day of week, and county by year fixed-effects. Relative speed is calculated as speed relative to the speed limit and multiplied by one hundred.

Table A.12: Falsification Test over Gender with White Motorists, County by Year Controls

LHS: Rel. Speed		(1) 10 pct	(2) 20 pct	(3) 30 pct	(4) 40 pct	(5) 50 pct	(6) 60 pct	(7) 70 pct	(8) 80 pct	(9) 90 pct
East TN	Daylight	0.573	0.411	-0.214	-0.304	-0.250	0.0333	0.483	0.686	0.527
		(0.706)	(0.531)	(0.307)	(0.332)	(0.379)	(0.424)	(0.586)	(0.820)	(1.311)
	Male	-0.426	-0.842**	-0.581**	-0.525*	-0.378	-0.138	0.0186	0.107	2.162*
		(0.513)	(0.343)	(0.222)	(0.269)	(0.295)	(0.292)	(0.354)	(0.541)	(1.118)
	Daylight*Male	-0.304	-0.0267	0.382	0.444	0.266	0.105	-0.120	-0.234	-0.843
		(0.613)	(0.422)	(0.274)	(0.286)	(0.373)	(0.381)	(0.446)	(0.688)	(1.293)
	Obs.	22424	22424	22424	22424	22424	22424	22424	22424	22424
West TN	Daylight	0.144	0.266	0.0142	0.0264	-0.109	0.216	0.248	0.513	0.183
		(0.124)	(0.191)	(0.139)	(0.172)	(0.154)	(0.207)	(0.282)	(0.349)	(0.476)
	Male	-0.311***	-0.502***	-0.176	0.0954	0.122	0.398**	0.652***	1.223***	1.513***
		(0.0924)	(0.148)	(0.118)	(0.142)	(0.120)	(0.167)	(0.216)	(0.262)	(0.360)
	Daylight*Male	0.00403	0.0808	0.0195	-0.0758	-0.0272	-0.273	-0.258	-0.746**	-0.238
		(0.120)	(0.172)	(0.143)	(0.171)	(0.151)	(0.197)	(0.249)	(0.308)	(0.429)
	Obs.	83076	83076	83076	83076	83076	83076	83076	83076	83076

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) but robust to clustering on county and year separately. Bootstrapping one-thousand random samples, we find that the p-value for a one-sided permutation test of joint significance on all nine quantiles is equal to 64.3 percent for East Tennessee and 42.7 percent for West Tennessee. The sample includes only traffic stops involving Non-Hispanic white motorists. Controls include time of day, day of week, and county by year fixed-effects. Relative speed is calculated as speed relative to the speed limit and multiplied by one hundred.

Table A.13: Falsification Test over Age with White Motorists

LHS: Rel. Speed		(1) 10 pct	(2) 20 pct	(3) 30 pct	(4) 40 pct	(5) 50 pct	(6) 60 pct	(7) 70 pct	(8) 80 pct	(9) 90 pct
MA	Daylight	0.0415	0.166	1.072	1.162	0.0214	-0.510	-1.676	-1.542	-3.136
		(1.234)	(1.078)	(0.966)	(1.159)	(1.350)	(1.330)	(1.532)	(2.679)	(4.114)
	Young Motorist	0.644	1.485**	0.947	1.277**	2.101**	1.212	1.497	3.078*	1.907
		(0.420)	(0.564)	(0.548)	(0.603)	(0.733)	(0.986)	(1.367)	(1.631)	(1.688)
	Daylight*Young Motorist	0.302	-0.113	0.626	0.539	-0.735	0.429	1.692	0.576	1.739
		(0.482)	(0.611)	(0.713)	(0.843)	(0.947)	(1.354)	(1.467)	(1.391)	(1.853)
	Obs.	8334	8334	8334	8334	8334	8334	8334	8334	8334

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 counties (TN) and patrol districts (MA) but robust to clustering on towns (MA) and county by patrol area (TN). Bootstrapping one-thousand random samples, we find that the p-value for a two-sided permutation test of joint significance on all nine quantiles is equal to 74.8 percent for MA-SP. Results estimated using absolute rather than relative results are generally robust and qualitatively similar to our primary estimates. The sample includes only traffic stops involving Non-Hispanic white motorists. Controls include time of day, day of week, and patrol location fixed-effects. The Tennessee sample also includes year indicators. Relative speed is calculated as speed above the limit relative to the limit and multiplied by one hundred.

Table A.14: Falsification Test over Vehicle Type with White Motorists, County by Year Controls

LHS: Rel. Speed		(1) 10 pct	(2) 20 pct	(3) 30 pct	(4) 40 pct	(5) 50 pct	(6) 60 pct	(7) 70 pct	(8) 80 pct	(9) 90 pct
East TN-HP	Daylight	0.370	0.616	0.157	0.00323	-0.000831	0.152	0.552	0.675	0.387
		(0.718)	(0.501)	(0.286)	(0.287)	(0.348)	(0.369)	(0.508)	(0.673)	(1.053)
	Domestic	0.614	0.820**	0.403*	0.254	0.327	0.450	0.562	0.500	1.352
		(0.551)	(0.409)	(0.218)	(0.249)	(0.299)	(0.321)	(0.441)	(0.566)	(1.078)
	Daylight*Domestic	0.0373	-0.655	-0.348	-0.0196	-0.213	-0.150	-0.464	-0.451	-1.425
		(0.584)	(0.483)	(0.282)	(0.312)	(0.369)	(0.394)	(0.527)	(0.668)	(1.308)
	Obs.	22424	22424	22424	22424	22424	22424	22424	22424	22424
West TN-HP	Daylight	0.151	0.361**	0.0871	0.0553	-0.0845	0.0481	0.147	0.0199	-0.0724
		(0.120)	(0.178)	(0.128)	(0.161)	(0.139)	(0.188)	(0.251)	(0.313)	(0.413)
	Domestic	0.335***	0.594***	0.490***	0.664***	0.547***	0.706***	0.847***	0.955***	0.669
		(0.0949)	(0.150)	(0.114)	(0.142)	(0.129)	(0.185)	(0.252)	(0.288)	(0.429)
	Daylight*Domestic	0.0265	-0.0603	-0.136	-0.199	-0.112	-0.0165	-0.197	0.00194	0.171
		(0.106)	(0.173)	(0.139)	(0.169)	(0.149)	(0.213)	(0.286)	(0.330)	(0.489)
	Obs.	83076	83076	83076	83076	83076	83076	83076	83076	83076

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 counties (TN) and patrol districts (MA) but robust to clustering on towns (MA) and county by patrol area (TN). Bootstrapping one-thousand random samples, we find that the p-value for a two-sided permutation test of joint significance on all nine quantiles is equal 82.4 percent for East TN-SP and 87.3 percent for West TN-SP. Results estimated using absolute rather than relative results are generally robust and qualitatively similar to our primary estimates. The sample includes only traffic stops involving Non-Hispanic white motorists. Controls include time of day, day of week, and patrol location fixed-effects. The Tennessee sample also includes year indicators. Relative speed is calculated as speed above the limit relative to the limit and multiplied by one hundred.

Table A.15: Falsification Test over Vehicle Age with White Motorists

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-SP	Daylight	-0.529	0.209	1.438	0.766	-0.547	-0.476	-0.578	-0.775	-2.066
		(1.091)	(1.227)	(1.108)	(1.349)	(1.608)	(1.742)	(2.008)	(2.745)	(3.677)
	Old Vehicle	-0.475	0.388	0.439	-0.812	0.147	-0.111	0.829	1.033	-0.727
		(0.669)	(0.765)	(0.902)	(0.650)	(0.571)	(0.670)	(0.738)	(0.934)	(1.404)
	Daylight*Old Vehicle	1.467*	-0.115	0.0108	1.323*	0.454	0.423	-0.229	-0.666	-0.336
		(0.771)	(0.742)	(0.782)	(0.643)	(0.654)	(0.792)	(1.079)	(1.656)	(2.345)
	Obs.	8334	8334	8334	8334	8334	8334	8334	8334	8334

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 counties (TN) and patrol districts (MA) but robust to clustering on towns (MA) and county by patrol area (TN). Bootstrapping one-thousand random samples, we find that the p-value for a two-sided permutation test of joint significance on all nine quantiles is equal to 60.7 percent for MA-SP. Results estimated using absolute rather than relative results are generally robust and qualitatively similar to our primary estimates. The sample includes only traffic stops involving Non-Hispanic white motorists. Controls include time of day, day of week, and patrol location fixed-effects. The Tennessee sample also includes year indicators. Relative speed is calculated as speed above the limit relative to the limit and multiplied by one hundred.

Table A.16: Falsification Test over Vehicle Color with White Motorists

LHS: Rel. Speed		(1) 10 pct	(2) 20 pct	(3) 30 pct	(4) 40 pct	(5) 50 pct	(6) 60 pct	(7) 70 pct	(8) 80 pct	(9) 90 pct
MA-SP	Daylight	-0.0280	-0.0749	1.356	1.315	-0.416	-0.414	-1.058	-1.397	-2.362
		(1.156)	(1.112)	(1.002)	(1.161)	(1.419)	(1.478)	(1.760)	(2.702)	(3.816)
	Red Vehicle	-2.131	-2.231	-1.137	-0.995	-0.442	-0.0736	-0.939	-1.777	0.0628
		(1.491)	(1.743)	(1.923)	(1.318)	(1.537)	(1.271)	(1.669)	(2.105)	(3.040)
	Daylight*Red Vehicle	2.563	2.179	0.942	1.616	1.209	1.836	3.364	2.690	1.561
		(1.942)	(2.257)	(2.145)	(1.740)	(1.796)	(1.959)	(2.421)	(3.013)	(4.418)
	Obs.	8334	8334	8334	8334	8334	8334	8334	8334	8334

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 counties (TN) and patrol districts (MA) but robust to clustering on towns (MA) and county by patrol area (TN). Bootstrapping one-thousand random samples, we find that the p-value for a two-sided permutation test of joint significance on all nine quantiles is equal to 46.7 percent for MA-SP. Results estimated using absolute rather than relative results are generally robust and qualitatively similar to our primary estimates. The sample includes only traffic stops involving Non-Hispanic white motorists. Controls include time of day, day of week, and patrol location fixed-effects. The Tennessee sample also includes year indicators. Relative speed is calculated as speed above the limit relative to the limit and multiplied by one hundred

B. Model and Calibration Appendix

We develop the simplest model possible that captures several key aspects of expected driver behavior in terms of speeding:

1. While police are more likely to stop drivers traveling at higher speeds, situations exist where some drivers are stopped at speeds only modestly above the speed limit and at other times drivers are not stopped for significantly more severe speeding infractions,

2. Some drivers at sometimes choose to obey the speed limit,

3. A driver who is on the margin between obeying the speed limit and speeding will if they decide to speed choose a level that is discretely over the speed limit (the choice to speed represents traveling at the speed limit versus 3 or 4 miles per hour above, rather than at the speed limit versus 1/5 or 1/2 a miles per hour above), and

4. When police stop-costs rise, motorists must in equilibrium have both a reduced likelihood of being stopped (a feature critical for the validity of the Veil of Darkness test) and an increased level of speeding (the key intuition behind our paper).

The resulting model will then allow us to derive the key assumptions made in the paper itself primarily from assumptions on primitive parameters specified with

B.1. The Police Officer's Problem

We begin by structuring the police officer's decision as the choice of selecting probability $\gamma(i, d, \phi)$ of making a stop. The officer's choice or decision is made after observing a non-negative infraction severity, i, e.g. miles per hour over the speed limit; a motorist's demography, d; and the circumstances surrounding the stop, ϕ, which might include both environmental factors and factors related to officers' idiosyncratic preferences and current circumstances. We assume for simplicity that $d \in \{m, w\}$ is a dichotomous random variable that indicates whether the motorists is a racial or ethnic minority, and s_d is the positive police stop-cost associated with motorist race and ethnicity.

The officer's maximization problem involves trading-off the stop pay-off, u, and stop-costs, which includes both race specific costs and a circumstance cost defined by the function $h(\phi)$. The maximization problem takes the following form:

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

where $\gamma \in [0,1]$. We make the following assumptions about police pay-offs and costs:

Assumption 1.1 u is a twice differentiable, non-negative function, $\frac{du(i)}{di} > 0$ and $\frac{d^2u(i)}{di^2} > 0\ \forall i > 0$, $\lim_{i \to 0} u(i) = u_0 > 0$ and $u(0) = 0$;

Assumption 1.2 $\phi \sim Uniform(0,1)$;

Assumption 1.3 h is a twice differentiable, non-negative function, $\frac{dh(\phi)}{d\phi} > 0\ \forall\ 0 \le \phi \le 1$, $h(0) = 0$ and $\lim_{\phi \to 1} h(\phi) = \infty$;

Assumption 1.4 $u_0 - s_d > 0\ \forall d$

We assume u is discontinuous at zero so that the officer receives no pay-off for stopping a motorist who has a zero level of infraction but has a pay-off bounded away from zero for any positive infraction level. This assumption is consistent with the current penalty structures in many states. We also assume that u has increasing total and marginal pay-off with respect to the severity of the infraction. The officer faces two costs for stopping a motorist: s_d, a race specific cost for stopping a motorist (henceforth *stop-cost*), and an additional circumstance specific cost, $h(\phi)$, resulting from factors like the officer's idiosyncratic preferences, geographic location, discretion, and contemporaneous opportunity cost (henceforth, circumstance). The introduction of circumstances allows for heterogeneity in whether individuals are stopped at a specific infraction level. The circumstances are drawn from a uniform (0,1) distribution without loss of generality because the monotonic function $h(\phi)$ captures possible non-linearities in the mapping between circumstances ϕ and an officer's net pay-off. We do not impose a sign restriction on the second derivative of h to allow for generality over circumstance costs. For example, if circumstance costs were distributed unimodally such as a chi-square distribution, the curvature of h must change sign over the range of ϕ. Finally, assumption 1.4 requires a positive net pay-off for a stop under some circumstances ϕ, even for an infinitesimally small positive level of infraction i. This effectively insures that the probability of stop is bounded away from zero for any motorist with a non-zero infraction level and so allows for a situation where motorists might choose not to commit an infraction.

Conditional on circumstances ϕ, demography d and the level of infraction i, the solution to the officer's problem requires an optimal infraction threshold, above which the probability an officer makes a stop, given full information, is equal to unity and otherwise the probability is zero due to the monotonic relationship between pay-off and the severity of motorist violation. Specifically, given the officer's net utility of $u(i) - h(\phi) - s_d\ \forall\ i$, the solution to her utility maximization problem is

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

Further, solving for zero net pay-off implies that an officer will stop all motorists at any infraction level above some threshold level following a specific stop-threshold function of

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

where u^{-1} maps from (u_0, ∞) to $(0, \infty)$ and $h(\phi) + s_d$ is always greater than u_0. Finally, conditional on infraction severity and exploiting the monotonicity of $h(\cdot)$, we can solve equation (2) for the circumstances when the net pay-off of a stop is zero, and officers will stop individuals with infraction level i whenever circumstances are more favorable than $\phi^*(i, s_d)$, i.e. $\phi < \phi^*(i, s_d)$. The resulting expression for the stop threshold over circumstances is

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

where h^{-1} maps from $(0, \infty)$ to $(0,1)$ and $u(i) - s_d$ is always greater than zero. Recall that ϕ is distributed uniform (0,1); thus equation (3) also represents the unconditional (i.e. circumstances have not been observed) probability that an officer stops a motorist with infraction level i. Using equation (3), we can then derive the conclusions that were imposed in Assumption 3 from the paper.

Proposition B1. *(i) The infraction level representing the optimal stop-threshold,* $i^* = u^{-1}[h(\phi) + s_d]$*, is increasing in officer circumstances and demographic based stop-cost. (ii) The probability of an officer making a stop,*

$\phi^*(i, s_d) = h^{-1}[u(i) - s_d]$*, is decreasing in stop-cost and increasing in the level of infraction.*

Proof. *Assumption 1.1 and the Implicit Function Theorem imply that* $u^{-1}(\cdot) > 0$ *over its domain* (u_0, ∞). *Then by inspection it is clear that the derivative of equation (2) implies* $\frac{\partial i^*}{\partial \phi} > 0$, $\frac{\partial i^*}{\partial s_d} > 0$. *Assumption 1.3 and the Implicit Function Theorem imply that* $h^{-1\prime}(\cdot) > 0$*, and by inspection it is clear that the derivative of equation (3) implies* $\frac{\partial \phi *}{\partial s_d} < 0$*, and* $\frac{\partial \phi^*}{\partial i} > 0$. ***QED***

B.2. The Motorist's Problem

The motorist's utility maximization problem, over infraction level, takes the following form:

$$\max_{i(c, s_d)} b(i, c) - \tau(i) \int_0^{\phi^*(i, s_{v,d})} \Gamma(\phi) d\phi \tag{4}$$

where $b(i,c)$ is the motorist pay-off for committing an infraction of a given level i, c is a motorist preference parameter, τ is motorist costs associated with being stopped when committing an infraction, and $\Gamma(\phi)$ is the probability density function of ϕ distributed as uniform (0,1). We make the following assumptions about motorist behavior:

Assumption 2.1 b is a twice differentiable, non-negative function, $\frac{\partial b}{\partial i} > 0$, and $\frac{\partial^2 b}{\partial i^2} < 0\ \forall\ c$ and $i \geq 0$, and $b(0,c) = 0$ and $\lim_{i\to\infty} \frac{\partial b}{\partial i} = 0\ \forall\ c$;

Assumption 2.2 $\frac{\partial b}{\partial c} > 0$ and $\frac{\partial^2 b}{\partial c \partial i} > 0\ \forall\ c$ and for $i > 0$;

Assumption 2.3 τ is a twice differentiable, positive function, $\frac{d\tau}{di} > 0$ and $\frac{d^2\tau}{di^2} > 0$ for $i \geq 0$, and $\tau(0) > 0$;

Assumption 2.4 $\frac{\partial b}{\partial i}|_{i=0} \geq \frac{d\tau}{di}|_{i=0} h^{-1}[u_0 - s_{v,d}] + \tau(0) h^{-1\prime}[u_0 - s_{v,d}]\ \forall\ c$;

Assumption 2.5 $\frac{\frac{d^2u}{di^2}}{\frac{du}{di}} \geq \frac{-h^{-1\prime\prime}(\cdot)}{h^{-1\prime}(\cdot)} \frac{\partial u}{\partial i}, \frac{\frac{\partial \tau}{\partial i}}{\tau(i)} > \frac{-h^{-1\prime\prime}(\cdot)}{h^{-1}(\cdot)} \frac{\partial u}{\partial i}\quad \forall i \geq 0.$

Assumption 2.6 $c \sim g(c,d)$ where there exists a $c_{h,d}$ such that $g(c,d) = 0\ \forall c > c_{h,d}$ and $g(c_{h,d}) > 0$.

The motorist maximizes the expected utility function in equation (4) with respect to infraction severity. She takes the probability of being stopped, $\gamma(i, s_{v,d}, \phi)$, from the officer's problem as given and integrates over the distribution of possible circumstances, ϕ. As such, the motorist compares the expected marginal benefits and costs when choosing an optimal i'. The term, c, captures motorist heterogeneity through context, e.g. recklessness, timing, sleep deprivation, etc. We assume that the motorist benefit or pay-off is an increasing function of infraction severity and that marginal benefit is diminishing. Additionally, both the benefit and the marginal benefit of infracting rise with recklessness, c. This assumption simply initializes the direction of effect of this parameter on motorist benefit. We assume that the motorist's cost and marginal cost are increasing in infraction severity, and motorist's cost is bounded away from zero for infinitesimally small infraction levels, which is required to assure that some motorists choose not to commit an infraction. The cost function is assumed to be invariant to recklessness. We also assume, at low levels of infraction, the marginal benefit of increasing infraction level is higher than the marginal cost of increasing infraction level in order to assure an interior solution for motorists who choose to commit an infraction. Finally, we impose two technical assumptions that the relative curvature (relative to the slope) of the officer's utility function and the relative slope of the cost function both exceed in magnitude the relative curvature of ϕ^* arising from the officers problem. The first restriction allows us to sign the second order condition of the motorist's problem. The second restriction assures that infraction severity responds to stop-costs in the expected manner increasing when police find it more costly to stop a motorist.

Based on these assumptions, we next derive the conclusions that were imposed in Assumption 4 from the paper in Proposition A2 and in Assumption 5 in Proposition A3 below.

Proposition B2. *There exists a unique optimal infraction level* i' *on* R^+ *for a motorist of type* $\{c, d\}$. *The optimal infraction level is increasing in criminality,* c, *and declining in stop-costs,* s_d.

Proof. *To show that the optimal infraction level is a function of stop-cost, we first rewrite equation (4) using Assumption 1.2, that* ϕ *follows a uniform distribution, and* $\int_0^{\phi^*(i\prime,s_d)} \Gamma\,(\phi)\,d\phi = \phi^*(i'(c,s_d),s_d)$ *as:*

$$\max_{i\prime(c,s_d)} b(i,c) - \tau(i)\phi^*(i,s_d) \tag{5}$$

Thus, the motorist will solve the maximization problem in equation (5) by choosing an optimal infraction level that satisfies the following first-order condition:

$$FOC \equiv \frac{\partial b(i,c)}{\partial i} - \frac{d\tau(i)}{di}\phi^*(i,s_{v,d}) - \tau(i)\frac{\partial \phi^*(i,s_{v,d})}{\partial i} = 0 \tag{6}$$

By Assumption 2.1, the first term in equation (6) is positive, and by Assumption 2.3 and Proposition 1 the second and third terms are negative when including the subtraction signs. Assumption 2.4 implies that the left-hand side of equation (6) is positive at $i = 0$. *Assumption 2.1 requires that the first term go to zero as* i *limits to infinity, and Assumption 2.3 implies that the second term is non-zero. Therefore, by continuity of all functions over* R^+, *a positive solution to equation (6) exists.*

The second-order condition of the motorist's problem excludes the possibility of multiple equilibria and can be written formally as:

$$SOC \equiv \frac{\partial^2 b(i,c)}{\partial i^2} - \frac{d^2\tau(i)}{i^2}\phi^*(i,s_{v,d}) - 2\frac{d\tau(i)}{di}\frac{\partial \phi^*(i,s_{v,d})}{\partial i} - \tau(i)\frac{\partial^2 \phi^*(i,s_{v,d})}{\partial i^2} < 0 \tag{7}$$

The first term in equation (7) is negative based on Assumption 2.1, the second term is negative based on Assumption 2.3, and the third term is negative based on Assumption 2.3 and Proposition 1. The final term is negative as well assuring uniqueness. In order to show why the final term is negative, we draw on the solution of the officer's problem and the monotonicity of $h(\cdot)$. *Recall that* $\phi^*(i,s_{v,d}) = h^{-1}[u(i) - s_{v,d}]$; *we use this expression to expand the second derivative of* ϕ^* *from equation (3):*

$$\frac{\partial^2 \phi^*(i,s_{v,d})}{\partial i^2} = \left(\frac{du(i)}{di}\right)^2 h^{-1\prime\prime}(u(i) - s_{v,d}) + \frac{d^2u(i)}{di^2} h^{-1\prime}(u(i) - s_{v,d}).$$

The first term is ambiguous, but the second term is positive and dominates the first term in the equation based on Assumption 2.5. Therefore, there exists a unique positive value of i' *that maximizes motorist payoff over* R^+ *conditional on* c *and* $s_{v,d}$.

By total differentiation of the first order condition in equation (6), it is easy to show that the optimal infraction level i^{**} is increasing in criminality. Therefore,

$$\frac{di'}{dc} = -\frac{\partial(FOC)}{\partial c}\frac{1}{SOC} = -\frac{\frac{\partial^2 b}{\partial c \partial i}}{SOC} > 0\ \forall c\ and\ s_{v,d}\ such\ that\ i' > 0,$$

where SOC is the expression for the second order condition in equation (7) and is negative as shown above. Therefore, the sign of the numerator is established by Assumption 2.2.

The same exercise signs the derivative with respect to stop-costs $s_{v,d}$ where the derivative of the FOC or the numerator is

$$\frac{\partial(FOC)}{\partial s_{v,d}} = -\frac{\partial \tau}{\partial i}\frac{\partial \phi^*}{\partial s_{v,d}} - \tau(i)\frac{\partial^2 \phi^*}{\partial i \partial s_{v,d}} > 0$$

The first term is positive by assumption 2.3 and Proposition 1. Using the functional form of ϕ^* established in equation (3) and assumption 2.5, it can be readily demonstrated that the first term in the expression dominates the second term. ***QED***

Next, we define i^{**} as the actual infraction level of the motorist. If the pay-off from the optimal infraction level is positive, then $i^{**} = i'$ otherwise $i^{**} = 0$.

Proposition B3. *There exist parameter values such that a threshold c^* exists, above which motorists commit a traffic infraction at the optimal level i', and below which motorists do not commit an infraction in equilibrium for any set of functions satisfying assumptions 1.1-1.4 and 2.1-2.6. For such parameter values, $lim_{c \to c^*} i^{**} > 0$ for c above c^*, and c^* is decreasing in $s_{v,d}$.*

Proof. *The proof proceeds by construction. Assume a benefit function $b(i,c)$. Based on Assumption 2.2, this benefit function approaches a finite maximum value $\overline{b}(c)$ for any c as i increases. Now pick an arbitrary value of c. Assumption 1.4 assures that $lim_{i \to 0} \phi^*(i, s_{v,d}) = \underline{\phi^*}(s_{v,d}) > 0$. Therefore, we can set the officer and motorist cost parameters so that $\tau(0)\underline{\phi^*}(s_{v,d}) > \overline{b}(c)$. For this cost function, a motorist of type c never speeds and $\tau(i)\phi^*(i, s_{v,d})$ always lies above $b(i,c)$. Now, because τ and b are differentiable and the second order condition in the motorist problem is always negative, we can slowly and continuously reduce the function τ by multiplying by a decreasing positive scaler less than 1 (where the scaler effectively acts as a parameter of the cost function) until $\tau(i)\phi^*(i, s_{v,d})$ just touches the function $b(i,c)$ at one point. Given that the slope of the benefit function over i at $i = 0$ is steeper than the slope of $\tau(i)\phi^*(i, s_d)$, the two curves will touch, yielding zero net benefits, at a positive value of i. For the selected parameters, the arbitrarily chosen c equals c^*. The benefit curves for all values of c below c^* lie below the benefits curve for c^* (and similarly, all curves lie above for c above c^*). Therefore, for c below c^*, net benefits are always negative, and for c above c^*, net benefits are positive over some range of i.*

In the proof of Proposition 2, we show that the optimal infraction level i' *is positive for all* c *and that the function* i' *is monotonically increasing in* c. *Therefore, for values of* c *below* c^*, i' *is positive and must lie below the optimal infraction level for any* c *greater than* c^*. *This implies that the optimal infraction level for any* c *greater than* c^* *is bounded away from zero, or* $lim_{c \to c^*} i^{**} > 0$.

An increase in $s_{v,d}$ *unambiguously lowers* $\phi^*(i, s_d)$ *holding* τ *fixed. At the* c^* *above, the cost curve shifts down and net benefits are positive. Therefore, a new tangency between the two curves holding* $b(i, c)$ *and* $\tau(i)$ *fixed requires a decrease in* c *in order to lower the benefits curve down to just touch the now lower* $\tau\phi^*(i, s_{v,d})$. ***QED***

B.3. Equilibrium Motorist Stop Probability and Speed Distribution

Interestingly, the assumptions above do not assure that the equilibrium probability of a motorist being stopped falls as police stop-costs rise. Specifically, when stop-costs rise, motorists drive faster because of lower stop probabilities. The total derivative of ϕ^* is

$$\frac{d\phi^*}{ds_{v,d}} = \frac{\partial \phi^*}{\partial s_{v,d}} + \frac{\partial \phi^*}{\partial i}\frac{di'}{ds_{v,d}}$$

Notice that the first term implies a direct lower probability of being stopped from higher stop-costs, but the second implies an increase in stop probability as optimal speed increases. Typically, we would not expect motorists to drive so much faster that the higher speeds actually more than undue the original decline in stop probabilities that was the reason behind the faster speeds in the first place.

However, we cannot easily rule out this possibility by placing assumptions on the primitive parameters of the problem. In order to see this, we use the solution for $\phi^*(i, s_d) = h^{-1}\big(u(i) - s_{v,d}\big)$ in equation (3) and replace i with $i'(c, s_d)$ to show that

$$\frac{d\phi^*}{ds_{v,d}} = \frac{\partial \phi^*}{\partial s_{v,d}} + \frac{\partial \phi^*}{\partial i}\frac{di'}{ds_{v,d}} = -h^{-1'}\big(u(i) - s_{v,d}\big)\left(1 - \frac{\partial u}{\partial i}\frac{di'}{ds_{v,d}}\right) <> 0 \qquad (8)$$

Both the derivative of police utility u and the derivative of motorist optimal speed i' are positive based on assumption 1.1 and Proposition 2 and i' is an equilibrium function creating ambiguity that cannot be easily resolved by imposing restrictions on primitive parameters or function.

If we are willing to impose a stronger assumption,

Assumption 3.1 $\frac{\partial u}{\partial i}\frac{di'}{ds_{v,d}} < 1$

Then equation (8) implies that stop probability will fall as stop-costs rise consistent with Assumption 6 in the paper. This assumption requires that the utility from police stops rise slowly enough as speed increases so that the effect of stop-cost on speed does not reverse the direct effect of stop-costs on the likelihood of a stop. We recognize that Assumption 3.1 is not ideal because it is based on an equilibrium function i'. However, the imposition of an assumption that individuals do not completely undo exogenous changes in incentives through their behavioral adjustments is relatively standard.

In the paper, we theoretically examine the distribution of the infraction severity or speed of motorists who commit speeding infractions recognizing that some fraction of motorists may choose not to speed. However, we do not observe the speed distribution of motorists who are speeding, but rather only observe the distribution of stopped motorists, which will be selected because the likelihood of stop increases with motorist speed. Therefore, in this appendix, we revisit the results in proposition 3 of the paper for the speed distribution of stopped motorists.

Specifically, we can characterize the speed distribution of stopped motorists by examining the effect of a change in stop-costs on the travel speed i_x of motorists at a specific percentile x of motorists who were stopped. Conditional on c and s_d, we can write a stopped motorist percentile as

$$x(c, s_d) = \frac{\int_{c^*(s_d)}^{c} g\,(c')\phi^*(i'(c', s_d), s_d)\,dc'}{\int_{c^*(s_d)}^{c_h} g\,(c')\phi^*(i'(c', s_d), s_d)\,dc'}$$

or equivalently

$$\int_{c^*(s_d)}^{c_x(x,s_d)} g\,(c')\phi^*(i'(c', s_d), s_d)\,dc' = x\int_{c^*(s_d)}^{c_h} g\,(c')\phi^*(i'(c', s_d), s_d)\,dc' \tag{9}$$

Definition B1. $i_x(x, s_d) \equiv i'(c_x(x, s_d), s_d)$

Proposition B4. *For all s_d there exists $\tilde{x}$ such that $\frac{\partial i_x}{\partial s_d} > 0$ for all $x > \tilde{x}$.*

Proof. *Differentiation of $i_x(x, s_d)$ in Definition A1. yields*

$$\frac{\partial i_x}{\partial s_d} = \frac{\partial i'}{\partial s_d} + \frac{\partial i'}{\partial c}\frac{dc_x}{ds_d}$$

By Proposition A1, the first term of the derivative is positive and bounded away from zero. The second term contains the effect of an increase in c on speed or infraction level, which is positive also based on Proposition A1 and the derivative of $c_x(x, s_d)$, which can be found by differentiating equation (9).

$$\frac{dc_x}{ds_d} g(c_x)\phi^*(i'(c_x, s_d), s_d) = (1 - x)\frac{dc^*}{ds_d} g(c^*)\phi^*(i'(c^*, s_d), s_d) + x\int_{c^*(s_d)}^{c_h} g\,(c')\frac{d\phi^*}{ds_{v,d}}\,dc'$$
$$- \int_{c^*(s_d)}^{c_x} g\,(c')\frac{d\phi^*}{ds_{v,d}}\,dc'$$

The first term above is negative based on Proposition A3 leading to an ambiguous derivative of i_x. *This first term represents the same source of ambiguity discussed in Proposition 3 in the paper. However, as* x *approaches 1, the first goes to zero. Further, as* x *approaches 1,* c_x *approaches* c_h *so that the second and third term exactly cancel out when* $x = 1$. *Therefore, when* $x = 1$, *the derivative of* c_x *is zero leading to a positive derivative of* i_x. *By continuity, there exist lower values* x *just below 1 where the derivative is also positive*

By assumptions 4, 5 and 7 the derivatives in the second term are finite, and by assumption 7 $x = 1$ *when* $c = c_h$. *Therefore, the derivative of* i_x *is positive when* $x = 1$ *and by continuity there are values of* x *just below one where the derivative is also positive.* ***QED***

Table B.1: Minimum and Maximum Values for Parameters

Parameters	Min	Max
α_1	0	1
M	1	4
δ_0	0	50
A	0.8	1.5
α_2	0	3
K	100	500
H	1	1.5
b_0	0	200
τ_0	50	800
σ_m	0	3
σ_w	0	3
mean_m	-4	2
mean_w	-4	2
skew_m	-50	100
s_v	44	44
MA		
skew_w	-50	100
s_vm	0	15
s_vw	44	60
E TN		
skew_w	-50	600
s_vm	30	44
s_vw	44	50
W TN		
skew_w	-50	100
s_vm	30	44
s_vw	44	47

Notes. Table presents the bounds on parameter values used for the envolutionary based optimization selected based on the local optima identified during the initial stages of optimization. Most parameter limits are the same by site with the exception of the minority and white daylight stop-costs which are influenced heavily by the empirical racial composition of stops, and for the white skewness where we observed unusually high levels of skewness in the white population in some of the calibrations for East Tennessee.

Table B.2: Calibration Results

	Massachusetts		East Tennessee		West Tennessee	
	Data	Simulation	Data	Simulation	Data	Simulation
White Speed Distribution Daylight						
20th Percentile	12.58119726	12.8990268	13.31786436	13.3819359	11.28808471	11.0761
40th Percentile	16.42617039	16.2500592	16.22345474	15.8291645	13.56950803	13.4798
60th Percentile	20.63785929	20.3829396	18.95081273	18.8376353	15.9148927	16.184
80th Percentile	26.69713667	26.8608619	23.30887568	23.5328263	19.74395994	20.0753
90th Percentile	33.27770052	33.3347123	27.7676207	27.9992266	23.69906791	23.7206
95th Percentile	41.03735374	40.9859405	33.42494972	33.2383482	28.02024315	27.7701
Difference Daylight and Darkness						
20th Percentile	-0.49062111	-0.2660177	0.00422909	-0.0000143	-0.11309278	0
40th Percentile	-0.38312866	-0.2971691	-0.11176114	0.0008881	-0.09190034	-0.0001
60th Percentile	0.2679906	-0.3117633	-0.03097062	0.0001257	-0.02815239	-0.0001
80th Percentile	-0.57652247	-0.3265926	-0.03718293	-0.0000601	0.03660167	0
90th Percentile	0.16650752	-0.34577	-0.07980587	0.0004908	0.26197307	-0.0001
95th Percentile	-1.70325014	-0.5461246	-0.2129078	0.0002214	0.46968762	0

Notes: Empirical speed distribution in miles per hour based on regressing relative speed on day of week, time of day, geographic and for Tennessee year controls, calculating the residual, adding the means of the controls back to the sample and then calculating the miles per hour based on the mode speed limit of traffic stops for each site. The simulated moments arise from the global optimum identified by applying an evolutionary based optimization routine called ESCH and precisely located by applying a second optimization routine based on quadratic approximations to the surface BOBYQA. The calibrated parameters used to calculate these moments are shown in Appendix Table 18.

Table B.3: Calibrated Parameters

Parameters	Sites		
	MA	E TN	W TN
α_1	0.522029	0.509337	0.999519
M	1.55008	1.52118	2.18566
δ_0	5.14442	35.5046	3.0628
A	1.23914	1.1562	0.987308
α_2	0.509207	0.421155	1.4297
K	320.493	331.992	235.093
H	1.00387	1.00018	1.24943
b_0	16.7978	17.7386	128.356
τ_0	139.826	122.94	495.065
σ_m	1.23344	1.16092	0.513587
σ_w	1.66537	1.47121	0.53856
mean_m	-0.157625	-0.601036	-2.04262
mean_w	-1.24202	-1.02003	-2.08983
skew_m	0.269799	0.286773	2.5971
skew_w	11.551	3.47682	9.46006
s_vm	0.0057178	30.1313	37.7525
s_vw	44.9736	44.0005	44.0004
s_v	44	44	44
MSE	0.7483	0.638	0.2591

Notes. Each column of this table contains the calibrated parameters for one of the three sites for our baseline set of weights where the speed distribution components each have a weight of 21.5 and the share stops minority in darkness and the VOT test statistics (times 100) each have a weight of 3.5%. The parameters for the Massachusetts sample are in column 1 labelled MA. Column 2 contains parameters for East Tennessee labelled E TN, and column 3 is West Tennessee labeled W TN. The last row shows the mean squared error of the moments for each site.

Table B.4: Calibrated Parameters for Massachusetts with Alternative Weights

Parameters	Weights		
	Original	Equal	Speed*Six
α_1	0.522029	0.522111	0.521889
M	1.55008	1.55032	1.55012
δ_0	5.14442	5.15151	5.14515
A	1.23914	1.23914	1.23901
α_2	0.509207	0.509362	0.509277
K	320.493	320.355	320.48
H	1.00387	1.00383	1.00385
b_0	16.7978	16.7934	16.7978
τ_0	139.826	139.812	139.807
σ_m	1.23344	1.23477	1.23406
σ_w	1.66537	1.66587	1.66551
mean_m	-0.157625	-0.160836	-0.157642
mean_w	-1.24202	-1.24327	-1.24219
skew_m	0.269799	0.268804	0.26982
skew_w	11.551	11.5521	11.5507
s_vm	0.0057178	0.104016	0.0209493
s_vw	44.9736	44.9798	44.9742
s_v	44	44	44
MSE	0.7483	0.5781	0.8044

Notes. This table presents the calibrated parameters for different weights for the State of Massachusetts sample. The first column presents parameters for the baseline weights. The second column presents parameters for equal weights of 16.7% for the four speed components and the two components based on share minority stopped (share in darkness and VOD test), and the third column presents parameters for weights where the speed distribution components that contain 6 moments each have approximately 6 times the weight or 23.25% as the weight of 3.5% for the share stops minority in darkness and the VOD test statistic.

Table B.5: Calibrated Parameters for East Tennessee with Alternative Weights

Parameters	Weights		
	Original	Equal	Speed*Six
α_1	0.509337	0.509832	0.509053
M	1.52118	1.52139	1.52119
δ_0	35.5046	35.512	35.5046
A	1.1562	1.1561	0.995303
α_2	0.421155	0.423291	0.421159
K	331.992	330.47	331.94
H	1.00018	1.00002	1.00073
b_0	17.7386	17.7361	10.4695
τ_0	122.94	122.763	122.94
σ_m	1.16092	1.16421	1.16237
σ_w	1.47121	1.46317	1.47036
mean_m	-0.601036	-0.61939	-0.601093
mean_w	-1.02003	-1.0126	-1.0212
skew_m	0.286773	0.293491	0.285781
skew_w	3.47682	3.4744	3.4276
s_vm	30.1313	30.0004	31.6522
s_vw	44.0005	44.0514	44.0331
s_v	44	44	44
MSE	0.638	0.4829	0.6164

Notes. This table presents the calibrated parameters for different weights for the subsample from the portion of the State of Tennessee that is east of the time zone boundary. The first column presents parameters for the baseline weights. The second column presents parameters for equal weights of 16.7%, and the third column presents parameters for weights where the speed distribution components that contain 6 moments each have 6 times the weight or 23.25% as opposed to a weight of 3.5% for the share stops minority in darkness and the VOD test statistic.

Table B.6: Calibrated Parameters for West Tennessee with Alternative Weights

Parameters	Weights		
	Original	Equal	Speed*Six
α_1	0.999519	0.999859	0.999519
M	2.18566	2.18541	2.18367
δ_0	3.0628	3.48442	3.28581
A	0.987308	0.987312	0.987276
α_2	1.4297	1.42969	1.42973
K	235.093	235.036	233.854
H	1.24943	1.25016	1.2497
b_0	128.356	129.092	128.877
τ_0	495.065	506.544	506.865
σ_m	0.513587	0.516049	0.516706
σ_w	0.53856	0.53923	0.538645
mean_m	-2.04262	-2.04261	-2.04212
mean_w	-2.08983	-2.09014	-2.08981
skew_m	2.5971	2.59473	2.59168
skew_w	9.46006	9.43347	9.47268
s_vm	37.7525	33.8754	31.6203
s_vw	44.0004	44.0022	44.2422
s_v	44	44	44
MSE	0.2591	0.168	0.2335

Notes. This table presents the calibrated parameters for different weights for the subsample from the portion of the State of Tennessee that is west of the time zone boundary. The first column presents parameters for the baseline weights. The second column presents parameters for equal weights of 16.7%, and the third column presents parameters for weights where the speed distribution components that contain 6 moments each have 6 times the weight or 23.25% as opposed to a weight of 3.5% for the share stops minority in darkness and the VOD test statistic.