**Appendix 1. Methods**

To construct Figure 1, we estimate models for each individual year in the data as well as the combined 2014-22 data. In particular, we estimate a linear probability model of the form

|  |  |
| --- | --- |
|  | (1) |

where the dependent variable is an indicator which is equal to one if traffic stop *i* was made of a person of color and zero otherwise. The primary independent variable is an indicator which is equal to one if a traffic stop was made in daylight and zero otherwise. Additional control variables include six indicators for the day of the week (), 5 indicators for each hour between 4 p.m. and 9 p.m. (), and 22,178 indicators for each unique officer in each agency . For the model where we combine all years in the data, we interact all of the control variables with indicators for each year from 2014 to 2022. We note that stops occurring during twilight (neither daylight or darkness) are dropped from the sample and that each underrepresented group (i.e. Black/African American, Hispanic/Latinx, and an aggregate grouping of any race/ethnicity) is only compared against a sample of Caucasian (i.e. non-Hispanic/Latinx) motorists. Across all of the estimates, we two-way cluster standard errors by officer and agency.

To construct Appendix Figure A.1, we estimate a linear probability model of the form

|  |  |
| --- | --- |
|  | (2) |

where the dependent variable is an indicator which is equal to one if traffic stop *i* was made of a motorist of color and zero otherwise. The primary independent variable is an indicator which is equal to one if a traffic stop was made in the period after the spring and before the fall daylight savings time shift. The variable is a running variable equal to as little as -21 before the shift to a period of more light and +21 afterwards. We interact this variable with an indicator for fall vs. spring both alone and in combination with . As before, we include additional control variables including six indicators for the day of the week (), 5 indicators for each hour between 4 p.m. and 9 p.m. (), and indicators for each unique officer in each season . For the model where we combine all years in the data, we interact all of the control variables with indicators for each year from 2014 to 2022.

To construct Appendix Figure A.2, we estimate the model from Equation (1) on a restricted sample of only speeding violations.

To construct Appendix Figure A.3, we estimate the model from Equation (1) but include additional controls for whether the motorist was female or young (age 25 or less) as well as whether their vehicle was registered in Massachusetts, new (vehicle age 6 years or less), or a bright color (red, yellow, or orange).

**Appendix 2. Robustness Tests**

**Figure A.1: Daylight Savings Time Robustness Check**

A graph with purple dots

Description automatically generatedA graph with red dots

Description automatically generated

(a) Black/African American (b) Hispanic/Latinx

A graph of a graph with numbers and dots

Description automatically generated with medium confidence

(c) All individuals from underrepresented groups

Notes: The results for panels (a), (b), and (c) are estimated on a sample traffic stops made of White non-Hispanic and Black/African American, Hispanic/Latinx, or any motorist of color during the inter-twilight for each 303 policing agencies. The figure plots the change in the likelihood a motorist of color is stopped in daylight relative to darkness. The estimates were obtained from a linear probability model regressing an indicator for race on an indicator for the daylight side of DST, a running variable, and an interaction between the two. We also include controls for season interacted with time of day, day of week, and badge. In the aggregate sample, the estimated point estimate was -0.53 (-4%, p<0.15) for the Black/African American comparison, 0.66pp (5%, p<0.14) for the Hispanic/Latinx comparison, and 0.19pp (1%, p<0.32) for all underrepresented groups comparison.

**Figure A.2: Speeding Violation Robustness Check**

A graph with purple dots and numbers

Description automatically generatedA graph with numbers and lines

Description automatically generated

(a) Black/African American (b) Hispanic/Latinx

A graph with numbers and lines

Description automatically generated

(c) All individuals from underrepresented groups

Notes: The results for panels (a), (b), and (c) are estimated on a sample traffic stops made of White non-Hispanic and Black/African American, Hispanic/Latinx, or any motorist of color stopped for speeding during the inter-twilight. The figure plots the change in the likelihood a motorist of color is stopped in daylight relative to darkness as well as the associated 95 confidence interval. The estimates were obtained from a linear probability model regressing an indicator for race on an indicator for daylight as well as controls for time of day, day of week, and badge. The standard errors, used to construct the confidence intervals and to conduct the hypothesis test denoted by the p-value, were two-way clustered at the badge and agency. In the aggregate sample, the estimated point estimate was 0.12 (1%, p<0.14) for the Black/African American comparison, 0.69pp (6%, p<0.12) for the Hispanic/Latinx comparison, and 0.74pp (2%, p<0.3) for all underrepresented groups comparison.

**Figure A.3: Balance Failure Robustness Check**

A graph with purple dots and numbers

Description automatically generatedA graph with numbers and lines

Description automatically generated

(a) Black/African American (b) Hispanic/Latinx

A graph with numbers and lines

Description automatically generated

(c) All individuals from underrepresented groups

Notes: The results for panels (a), (b), and (c) are estimated on a sample traffic stops made of White non-Hispanic and Black/African American, Hispanic/Latinx, or a combined sample of all motorists of color during the inter-twilight. The figure plots the change in the likelihood a motorist of color is stopped in daylight relative to darkness as well as the associated 95 confidence interval. The estimates were obtained from a linear probability model regressing an indicator for race on an indicator for daylight as well as controls for time of day, day of week, badge, motorist age, motorist gender, vehicle age, vehicle color, and state residence. The standard errors, used to construct the confidence intervals and to conduct the hypothesis test denoted by the p-value, were two-way clustered at the badge and agency. In the aggregate sample, the estimated point estimate was 0.3 (2%, p<0.15) for the Black/African American comparison, 0.85pp (6%, p<0.14) for the Hispanic/Latinx comparison, and 0.93pp (3%, p<0.32) for all underrepresented groups comparison. .

**Figure A.4: Individual Agency Veil of Darkness Estimates, 2022**

A graph with purple dots

Description automatically generatedA graph of a line with orange dots

Description automatically generated with medium confidence

(a) Black/African American (b) Hispanic/Latinx

A graph of a graph showing the number of dots

Description automatically generated with medium confidence

(c) All individuals from underrepresented groups

Notes: The results for panels (a), (b), and (c) are estimated on a sample traffic stops made of White non-Hispanic and Black/African American, Hispanic/Latinx, or a combined sample of all motorists of color during the inter-twilight for each 303 policing agencies. The figure plots the change in the likelihood a motorist of color is stopped in daylight relative to darkness. The estimates were obtained from a linear probability model regressing an indicator for race on an indicator for daylight as well as controls for time of day, day of week, and badge.

**Figure A.5: Individual Agency Veil of Darkness Estimates, 2014-22**

A purple dot on a white background

Description automatically generatedA graph with red dots

Description automatically generated

(a) Black/African American (b) Hispanic/Latinx

A graph of a graph showing the number of motorists

Description automatically generated with medium confidence

(c) All individuals from underrepresented groups

Notes: The results for panels (a), (b), and (c) are estimated on a sample traffic stops made of White non-Hispanic and Black/African American, Hispanic/Latinx, or a combined sample of all motorists of color during the inter-twilight for each 303 policing agencies. The figure plots the change in the likelihood a motorist of color is stopped in daylight relative to darkness. The estimates were obtained from a linear probability model regressing an indicator for race on an indicator for daylight as well as controls for time of day, day of week, and badge.