ONLINE APPENDIX: A Few Bad Apples? Racial Bias in Policing

Felipe Goncalves*　　Steven Mello†

*University of California, Los Angeles; fgoncalves@econ.ucla.edu
†Dartmouth College and NBER; steven.m.mello@dartmouth.edu

Contents

A Data Appendix

A.1 Citations Data

Our data cover the universe of citations written by the Florida Highway Patrol for the years 2005-2015, comprising 2,614,119 observations. We make several restrictions that reduce the number of observations, (FDOC, 2019):

1. all FHP speeding citations (3,864,034 observations)
2. no crash is involved (3,861,667 observations, 99.9% of previous sample)
3. speed is between 0 and 40 over the limit (3,836,934, 99.4%)
4. posted speed limit is between 25MPH and 75MPH (3,834,404, 99.9%)
5. citations not from an airplane (3,826,027, 99.8%)
6. unique citation IDs (3,816,404, 99.7%)
7. race/ethnicity is white, black or Hispanic (3,052,641, 80.0%)
8. not missing driver's license state, gender, or age (3,051,122, 99.9%)
9. officer is identifiable (2,335,481, 76.5%)
10. officer has at least 100 tickets, and at least 20 for minorities and 20 for whites (2,299,308, 98.5%)
11. driver has no more than 20 citations in Florida for period 2005-2015 (2,122,555, 92.3%)

Our data also include the universe of non-FHP citations issued in the state over 2005-2015 (i.e., those issued by municipal police and sheriff departments). For reference, FHP accounts for about 20% of all citations in the state and about 35% of speeding citations. Because patrol officers have access to full citation histories regardless of issuing agency when examining a driver's record, we use both FHP and non-FHP tickets when constructing our measures of a driver's ticketing history. We also use all tickets (FHP and non-FHP) to construct our recidivism measure used in our tests for statistical discrimination.

Beginning in 2013, about 40% of tickets are geocoded with the latitude and longitude of a stop (271,164 observations). We link the geocoded tickets to a Florida Department of Transportation roadmap shapefile using ArcGIS.[1] The shapefile is at the level of road "segments," which are on average 6.7 miles long and roughly correspond to entire streets within cities and uninterrupted stretches of road on interstates and highways. Tickets are linked to the nearest segment, and we remove tickets that are more than 100 meters from the nearest road (dropping 1.5% of observations). Officers in more rural areas and on interstates are given priority for vehicles with GPS, as they cannot clearly describe the location of their ticket using street intersections. 60% of officers have fewer than 5% of their stops geocoded, and there is some variation across counties in the share of tickets geocoded.

A.2 Linking Offenses to Personnel Information

Officers enter their information by hand onto each speeding ticket, leading to inconsistencies in how their names are recorded. Some names are misspelled, and sometimes officers place only their last name and first initial. The Florida Department of Law Enforcement (FDLE) maintains a record of each certified officer in the state, along with demographic information. We link these using each officer's last name and first three letters of first name (if available on ticket) using a fuzzy match algorithm in Stata (reclink). We restrict attention to officers who are unique up to last name and first three letters of first name in the FDLE data. Among tickets where only the first initial is listed, we keep matches where the last name and first initial of an officer are unique in the FDLE data. Of the 10,872,998 speeding tickets in our data, 8,148,816 match successfully to the FDLE data.

A.3 Hours and Shifts of Tickets

Officers manually enter time of day, and there are several inconsistencies in how these are recorded. Most officers use either a 12-hour time and clarify AM versus PM, and others use 24-hour military time. Some officers regularly use 12 hour time and do not record AM versus PM. We set these times to be missing.

The FHP has three shifts, 6am to 2pm, 2pm to 10pm, and 10pm to 6am. We record

[1] http://www.fdot.gov/planning/statistics/gis/road.shtm; We use the "Basemap Routes with Measures" shapefile.

these directly from the hour of the ticket if it is properly recorded above. Of the 2,122,555 observations in our sample, 781,163 have shift missing.

A.4 Traffic Court Dispositions

Traffic court dispositions from the Traffic Citation Accounting Transmission System (TCATS) database were also shared by the Florida Clerk of Courts and matched to citations using citation identifiers (FCCC, 2016). Disposition verdicts can take on the following values:

1 = *guilty*; 2 = *not guilty*; 3 = *dismissed*; 4 = *paid fine or civil penalty*; 6 = *estreated or forfeited bond*; 7 = *adjudication withheld (criminal)*; 8 = *nolle prosequi*; 9 = *adjudged delinquent (juvenile)*; A = *adjudication withheld by judge*; B = *other*; C=*adjudication withheld by clerk (school election)*; D = *adjudication withheld by clerk (plea nolo and proof of compliance)*; E = *set aside or vacated by court.*

Verdicts 1 (5.9%), 2 (0.1%), 3 (8.3%), 4 (43.7%), A (20.0%), and C (22.0%) account for over 99% of speeding tickets in our sample with disposition information (88% of citations). These shares are consistent with the fact that the remaining codes ought not apply to speeding infractions, and we focus on this subset of codes when examining disposition outcomes.

Several of the disposition codes are difficult to interpret. For example, verdicts *3* and *A* likely indicate judge involvement and suggest a possible reduction in penalty. However, a penalty reduction can take many forms, such as a small fine or point reduction, and the nature of the reduction is not observable in the data. Moreover, the FCC estimates an on-time payment rate over 90% on traffic citations, which means that many citations with verdict codes *3* and *C* are ultimately paid in-full even if there is court involvement. Hence, the verdict codes to not offer sufficient information to determine whether a contesting driver wins or loses her traffic court case.

To construct our measure of contesting, we focus on the two most easily interpretable verdict codes – straight pay (*4*) and traffic school election (*C*). Traffic school participation requires on-time fine payment. We assume this subset of drivers pay their fines and individuals with any other disposition code (*1*. *2*, *3*, *A*) contest the ticket in court. Our measure is a conservative one that likely overestimates the contest rate.

A.5 Background on Fine Payment Institutions

After receiving a citation, payment is due to the county clerk within thirty days. Drivers with outstanding fines at that point are considered delinquent and receive a driver license suspension, effective immediately. Driving with a suspended license is a misdemeanor (rather than civil) offense carrying a fine of at least \$300 and the possibility of jail time, especially for repeat offenders. Alternatively, individuals can request a court date to contest the ticket in return for a \$75 court fee, where a judge or hearing officer typically decides to either uphold the original charge or reduce the punishment.

There is scant evidence on the ability to pay traffic fines and nonpayment is very difficult to infer from our data. Our conservative estimate indicates that at least 63% of fines are paid in-full and on-time (see Appendix Section A.4), while the Florida Clerk of Courts estimate a payment rate above 90%. Proxying personal income with zip code per-capita income, Table 1 suggests that, on average, white drivers earn about \$15,100 more than Black drivers and \$8,700 more than Hispanic drivers, suggesting that ability to be pay could be correlated with race. Consistent with this hypothesis, we observe a higher "straight pay" rate for white drivers than nonwhite drivers using our conservative measure of whether a ticket was paid on-time.

B Difference-in-Differences Correction for Speed Differences

Here we provide more detail on the difference-in-differences correction for potential differences across races in underlying speed distributions, presented in Section 3.1. Using the simple model from Section 2, we can decompose our diff-diff coefficient into a weighted average of discrimination at different speeds and a term that reflects racial differences in speeding:

$$\begin{aligned}
\beta &= [Pr(X = x_d|W, \text{Lenient}) - Pr(X = x_d|W, \text{Non-Lenient})] \\
&\quad -[Pr(X = x_d|M, \text{Lenient}) - Pr(X = x_d|M, \text{Non-Lenient})] \\
&= f_w(9) + \sum_{k>9} [f_w(k) Pr(X = x_d|W, \text{Lenient}, k)] - f_w(9) \\
&\quad f_m(9) + \sum_{k>9} f_M(k) Pr(X = x_d|M, \text{Lenient}, k) - f_m(9) \\
&= \sum_{k>9} [f_w(k) Pr(X = x_d|W, \text{Lenient}, k) - f_m(k) Pr(X = x_d|M, \text{Lenient}, k)] \\
&= \sum_{k>9} \left[f_w(k) \times [Pr_w(k) - Pr_m(k)] \right] + \sum_{k>9} \left[(f_w(k) - f_m(k)) \times Pr_m(k) \right]
\end{aligned}$$

where we use $Pr_w(k)$ to denote the probability of discounting by a lenient officer. Our object of interest is the first expression, and the second expression reflects the bias in our coefficient induced by racial differences in the distribution of stopped speeds. We can use the non-lenient officers to estimate the set of both $\{f_w(k) - f_m(k)\}$ and $\{Pr_m(k)\}$.

To estimate $f_w(k) - f_m(k)$ for each speed k, we run a regression among the non-lenient officer sample where the outcome is that the driver is ticketed speed k and that includes location and time fixed effects, driver covariates, and whether driver race is white,

$$S_i^k = \alpha^k White_i + X_i \gamma + \epsilon_i$$

and take α^k as an estimate for the difference in density.

To estimate $Pr_m(k)$, we use that the probability a minority driver is ticketed at k is $f_m(k)$ for non-lenient officers and $(1 - Pr_m(k)) \cdot f_m(k)$ for lenient officers. We restrict attention to minority drivers and regress whether the driver is ticketed speed k on whether the officer is lenient, in addition to shift-time fixed effects and driver covariates,

$$S_i^k = \rho^k Lenient_i + X_i \gamma + \epsilon_i$$

and take ρ^k as an estimate for $-\Pr(k) f_m(k)$. We then divide the negative of ρ by the statewide share of minority drivers at k for non-lenient officers (our estimate of $f_m(k)$). Our estimate of the bias correction is then $-\sum_{k>10} \alpha^k \rho^k / f_m(k)$.

We estimate the degree of bias from racial differences in speeding to be 0.0065, (s.e.

0.0006). Standard errors are calculated by bootstrap, where we draw random sets of observations of the same size as our main sample from our data with replacement, calculate the statistic, and iterate 100 times. This estimate of the bias from differences in speeding is 11.1% of our baseline DD coefficient, and we can rule out at the 1% significance level that the true value for bias is more than 14% of the baseline DD coefficient.

One complication to note is that our estimation approach leads to estimates of the probability of discounting that are negative at four speeds: 14, 24, 29, and 39 MPH over. We believe this inconsistency is due to the fact that there is a small amount of bunching down to other speeds. In particular, the number of points added to a driver's record increases from 3 to 4 at 15 MPH over the speed limit. In Appendix Section I, we modify our parametric model to allow officers to discount to 14 MPH over, and the main results do not change meaningfully. Our baseline estimate of the bias correction replaces the probability of discount at these points to be the midpoints of the two adjacent probabilities.

An alternative approach to our baseline estimation strategy is to assign the probability that most increases our estimate of bias. Since empirically $f_w(14) > f_m(14)$ and $f_w(s) < f_m(s)$ for $s \in \{24, 29, 39\}$, the most conservative estimate is to allow $P_m(14) = 1$ and $P_m(s) = 0$ for $s \in \{24, 29, 39\}$. When we do so, the estimate of bias increases from 0.0065 to 0.0088 (s.e. 0.0007), which is 14.9% of our main estimate.

An alternative and extreme approach to bounding the bias in our diff-diff is to impose that $P_m(k) = 1$ whenever $f_w(k) > f_m(k)$, and $P_m(k) = 0$ whenever $f_w(k) < f_m(k)$, which generates an upper bound estimate for the impact of our observed differences in speeding on the difference-in-differences coefficient. Doing so, we get an estimate of 0.024 (s.e. 0.0015), which is 40.6% of our DD coefficient.

C Lenience Versus Discrimination

As noted in Section 3, our object of interest is the practice of discrimination rather than internal racial animus, though the two are likely related. Because officers have to be lenient to discriminate, one third of officers by construction practice no discrimination. A different question that may be of interest is how the share discriminatory would compare if all officers were forced to practice lenience, which we call the "overall share discriminatory."

We can think of this question of identifying overall discrimination as a selection problem.

Officers have to select into lenience, and the choice to practice lenience may be correlated with discrimination propensity. If officers with a higher propensity to discriminate are more likely to select into lenience, the share discriminatory among lenient officers is an upper bound for the overall share discriminatory. Conversely, if officers with high propensity to discriminate are less likely to be lenient, the share discriminatory among lenient officers is a lower bound.

To address this selection problem, we think of the troop of the officer as an instrument for selection into lenience. Five of the nine troops in our sample have around 90% of officers practicing lenience. If we suppose that officers are similar across troops, we can then estimate the share discriminatory in these troops. An officer's troop is a less-than-ideal instrument for selection. Officers may differ in unobservable ways across troops, which cover large geographic areas. However, this approach provides an approximation to the share of officers who would discriminate if all officers practiced lenience.

We re-conduct our MLE procedure from Section 4.1 separately for each troop, which we present in Figure A.4. The left panel plots estimates of the share of all officers who are discriminatory against the share of officers who are lenient, while the right panel shows the estimate of the share of lenient officers who are discriminatory. From the x-axis, we see the substantial variation in lenience across troops and the fact that five troops have 85% or more officers who are lenient. While these troops will naturally have more discriminatory officers overall, the share discriminatory *among lenient officers* is remarkably consistent across departments, with most confidence intervals covering or close to the department-wide rate of 70.0%. We conclude that our share discriminatory among lenient officers is a roughly accurate estimate of the share discriminatory if all officers practiced lenience.

D Applications of Officer Heterogeneity

Relative to the literature, our central contribution is the ability to generate officer-level estimates of discrimination, as presented in Section 4. Here, we present two analyses that can be conducted with these estimates. First, we study how discrimination varies by officer demographics. Second, we evaluate whether a measure of discrimination from an officer's early patrolling is predictive of discrimination later in their career.

D.1 Do Officer Characteristics Predict Discrimination?

In Section 4.1, we document how discrimination varies with officer race. A natural next question is how discrimination correlates with the full set of officer characteristics and behaviors, which we address here.

In addition to officer demographics, the records from the Florida Department of Law Enforcement also include every misconduct investigation made by the state against an officer, the type of alleged violation, and the ultimate verdict of the state. From the FHP, we also collected information on all use of force incidents and civilian complaints against officers for the period 2010-2015, which list the name of the officer, the date of the incident, and a description of the incident (FHP, 2015*a*,*b*).

In Table A.9, we present regressions of officer-level discrimination on officer characteristics. Here we have disaggregated officer discrimination to be calculated separately against Black drivers and Hispanic drivers, as described below in Appendix Section E. All observations are weighted by the variance of the noise in our estimate of the officer's bias.

As with the density plots, the clear takeaway from the regressions is that minority officers are more lenient toward minority drivers, as we might expect. Female officers appear less biased against Hispanic drivers and marginally less biased against Black drivers. Officers with more years experience are more discriminatory against Hispanic drivers, though the standard errors are large. There appears to be no relationship between officer discrimination and level of education, number of civilian complaints, or number of use-of-force incidents.

While some officer demographics are predictive of discrimination, we are also interested in the usability of our measures of discrimination to predict other officer behavior. A growing literature is interested in identifying the factors that can predict officer misconduct (Chalfin et al., 2016). Here we ask whether our measures of lenience and discrimination can be used to predict an officer's propensity to receive a civilian complaint or use force on the job. To make the analysis at the officer-level – but still account for differences in years and locations worked – we run regressions of the following form:

$$Y_i = \alpha_0 + \alpha_1 \cdot \text{Lenience}_i + \alpha_2 \cdot \text{Bias}_i + X_i \cdot \beta + \sum_k \gamma_k \text{Troop}_i^k + \sum_k \theta_k \text{Year}_i^k + \epsilon_i$$

where Y_i is an outcome of either receiving a civilian complaint or using force. Troop_i^k is an indicator for troop k being an officer's modal ticket location, and Year_i^k indicates whether an officer appears in our traffic data in year k. X_i are other officer-level characteristics.

The results, reported in Table A.10, indicate that lenience is statistically predictive of both civilian complaints and use of force. An increase of one standard deviation in lenience (25pp change in discounting) correlates to 0.15 fewer civilian complaints and a 4pp decreased likelihood of receiving any complaints. Similarly, a one SD increase in lenience is associated with 0.04 fewer incidents of force and 1.3pp lower likelihood of any force, though these estimates are insignificant. Black officers are less likely to engage in force, as are older officers. Female officers are less likely to receive complaints but are similarly likely as male officers to use force. Discrimination against minorities seems to be positively related to force and negatively related to complaints, though the standard errors are too large to say conclusively.[2]

Our finding that officer lenience is negatively associated with civilian complaints and use of force is noteworthy, though answering whether this relationship is causal is outside the scope of this paper. Further research that identifies the causal impact of lenience on other officer outcomes and considers the tradeoff between these different dimensions of behavior would be an important advancement.

D.2 Early Versus Late Discrimination

We argue in this study that the central value of estimating the distribution of discrimination is its use for conducting policy. Knowing who is discriminatory is crucial for identifying who to train or discipline. Given this motivation, a natural question is whether the measure we have constructed for each officer can be quickly and stably estimated early in their career. Specifically, we ask whether discrimination measured at the start of an officer's sample– where we use multiple measures for the start– is correlated with discrimination measured later in their sample.

[2]Our estimates of lenience and discrimination are both measured with error, leading to attenuation in the relationship between these measures and misconduct. To attempt to account for this error, we also do a split-sample instrumental variables procedure. We divide each officer's data randomly in half and estimate their bias and lenience for each sample. We then use one estimate as an instrument for the other. Doing so, we find the coefficients on discrimination increase overall in magnitude, though the standard errors remain too large to definitively say whether there is a true relationship.

To calculate the early measure of discrimination, we first predict whether a ticket is going to be at the discount point using only our sample of non-lenient officers, fitting $E(S^9_{ij}|X_{ij}) = X_{ij}\beta$. We then calculate $\epsilon_{ij} = S^9_{ij} - X_{ij}\hat{\beta}$ for each ticket, including those by lenient officers. Then, we take each officer's first 100 tickets and calculate discrimination as the difference in residuals across his white and minority drivers, $D^{\text{early}}_j = \overline{\epsilon^{\text{white}}_{ij}} - \overline{\epsilon^{\text{min}}_{ij}}$. We construct an analogous measure of late discrimination, D^{late}_j, using only stops after an officer's first 100. We also construct these discrimination measures where the cutoff for early is the first 200 tickets or the first 12 months of patrolling. The latter constitutes the "probationary period" for FHP officers and during which they can be terminated for any reason.

We report in Figure A.5 the results of conducting local linear regressions of various measures of late discrimination on early discrimination. We restrict attention to officers who are found to be lenient in the early period. The top left panel conducts regressions of an officer's percentile of late discrimination on their percentile of early discrimination. We find a strong relationship between early and late behavior. An officer at the 80th percentile of early discrimination is 14 percentiles higher in the distribution of late discrimination than an officer at the 20th percentile (57 v. 43).

The top right panel plots local linear regressions of the average measure of late discrimination against percentile of early discrimination. Here we find that an increase in early discrimination leads to a higher prediction for the magnitude of an officer's discrimination in the later period. The bottom left panel plots regressions of whether an officer's late discrimination is found to be significant (at the 5% level) against early discrimination percentile, where we also find a significant relationship between early and late discrimination.

There are two potential explanations for the fact that our early measures of discrimination are not perfectly predictive of late discrimination. The first is that officer underlying degree of discrimination is constant, but estimation error leads to a less than perfect prediction of late discrimination from early discrimination. The second explanation is that an officer's underlying degree of discrimination changes of their career. To probe this question further, we compare the regression plots for the 100-ticket measures and the 200-ticket measures. If estimation error is driving less-than-perfect relationship, we should expect the relationship to be stronger using the 200-ticket measures, where early discrimination is measured with better precision. Because the average officer writes over 700 tickets, the reduction in late-period tickets for the 200-ticket approach should not significantly reduce precision. However, the regression plots look nearly identical between the two cutoffs for early v. late. We therefore conclude that our estimates are more consistent with officers changing their underlying degree of discrimination somewhat over their careers.

Taken together, these calculations suggest that our early measure can be useful for identifying officers for training as part of an early-warning system (Walker, Alpert and Kenney, 2000). However, we caution against disciplining or removing officers on the basis of our early measures, as officers do change their underlying behavior somewhat as their career progresses.

E Black and Hispanic Measures of Discrimination

While our main analysis focuses on the degree of discrimination against Black and Hispanic drivers as a composite, an important follow-up question is how our measures of discrimination change when estimated separately for Black and Hispanic drivers. To answer this question, we first run a modified version of the baseline Diff-Diff regression, Equation 3, where officer lenience is interacted with indicators for Black and Hispanic driver:

$$\begin{aligned} S_{ij}^k &= \beta_0 + \beta_1 \cdot \text{BlackDriver}_i + \beta_2 \cdot \text{HispanicDriver}_i + \beta_3 \cdot \text{Lenient} \\ &\quad + \beta_4 \cdot \text{BlackDriver}_i \cdot \text{Lenient}_j + \beta_5 \cdot \text{HispanicDriver}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij} \end{aligned} \tag{1}$$

We present the results of these regressions in Appendix Table A.12. Analogous to Table 3, the first column presents the baseline regression for the full sample with our standard fixed effects but without any covariates. The second column presents the same regression but with all driver covariates and interactions between officer lenience and driver demographics. We find that both Black and Hispanic drivers are less likely to be discounted by lenient officers than white drivers. However, the coefficient for discrimination against Hispanic drivers is more than twice the magnitude of the gap for Black drivers. While the coefficients are slightly smaller in magnitude in Column 2 relative to Column 1, these differences are not statistically significant. This similarity in coefficients indicates that the disparity we observe is not driven by "discrimination by proxy," as ruled out for our main results.

In Columns 3 and 4, we perform the same regressions but on the sample of GPS tickets. The coefficients we find are similar in magnitude and also indicate that Hispanic drivers face more discrimination than Black drivers.

We next run an alternate version of the officer-level regression, where each lenient officer has an interaction with indicators for Black driver and Hispanic driver:

$$
\begin{aligned}
S_{ij}^{k} &= \beta_0 + \beta_1 \cdot \text{BlackDriver}_i + \beta_2 \cdot \text{HispanicDriver}_i + \beta_3^{j} \cdot \text{Lenient}_j \\
&\quad + \beta_4^{j} \cdot \text{BlackDriver}_i \cdot \text{Lenient}_j + \beta_5^{j} \cdot \text{HispanicDriver}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}
\end{aligned} \tag{2}
$$

Where the objects of interest are β_4 and β_5, whose negatives are our measures of officer discrimination against Black drivers and Hispanic drivers, respectively. As in the primary specification, we say that officers who are non-lenient have values of $\beta_4 = 0$ and $\beta_5 = 0$.

In Figure A.6 below, we produce the distribution of officer-level discrimination against Blacks and Hispanic drivers. The left panel plots the distribution for all officers, and the right panel restricts attention to lenient officers. The distributions look very similar, though discrimination against Hispanic drivers appears to be somewhat higher.

Applying our method from Section 5.1, we find that $36.4\% \pm 1.3$ of officers exhibit discrimination against Black drivers and $39.8\% \pm 1.4$ against Hispanic drivers. Among lenient officers, these shares are $53.7\% \pm 1.8$ and $58.4\% \pm 1.7$, respectively.

The correlation in officer-level discrimination against the two groups is 0.449 among lenient officers. The standard deviation across discrimination estimates are 0.063 and 0.078 for Black and Hispanic discrimination, respectively. The average discrimination coefficient has a standard error of 0.010 and 0.011, respectively. Using these values to account for the fact that the correlation is across noisy coefficients, we estimate that the true correlation is 0.462.

F Accounting for Stopping Margin Selection

As discussed in Section 5, one concern we face is that we do not observe interactions that do not result in a ticket. Therefore, officer differences in lenience and discrimination on whether to give a ticket may bias our estimates of discrimination on whether to give a discount. Here we write down a simple selection model to discuss the potential bias from selection into the data and present a procedure to correct our estimates for officer-by-race differences in ticketing.

Consider a model of ticketing where there is a first margin of whether or not a driver is ticketed at all:

$$
\begin{aligned}
D^*_{ij} &= \theta^W_j + \theta^M_j \cdot M_i + \epsilon_{ij} \\
Z_{ij} &= \alpha^W_j + \alpha^M_j \cdot M_i + \eta_{ij}
\end{aligned}
$$

D^*_{ij} is a latent variable for whether the driver receives a discount, and Z_{ij} is a latent variable for whether the officer tickets the driver at all, where we assume $\eta_{it} \sim N(0,1)$. We observe D_{ij} if Z_{ij} crosses zero and the officer chooses to ticket the driver:

$$
D_{ij} = \begin{cases} \mathbb{I}(D^*_{ij} \geq 0) & \text{if } Z_{ij} \geq 0 \\ \text{missing} & \text{otherwise} \end{cases}
$$

Therefore, the comparison we make to determine the degree of discrimination is based on the difference in discounting among observed drivers[3]:

$$
\begin{aligned}
\hat{\theta}^M_j &= E[D^*_{ij}|M_i = 1, Z_{ij} > 0] - E[D^*_{ij}|M_i = 0, Z_{ij} > 0] \\
&= \theta^M_j + E[\epsilon_{ij}|\eta_{ij} > -\alpha^W_j - \alpha^M_j] - E[\epsilon_{ij}|\eta_{ij} > -\alpha^M_j]
\end{aligned}
$$

If there's a difference in treatment in the first margin ($\alpha^M_j \neq 0$) and $corr(\epsilon_{ij}, \eta_{ij}) \neq 0$, then our estimate of θ^M_j will be inconsistent. In particular, if $\alpha^M_j > 0$ (discrimination in ticketing) and $\text{corr}(\epsilon_{ij}, \eta_{ij}) < 0$ (drivers more likely to be ticketed are less likely to be discounted), then the error term above will be positive, suggesting that our measure of discrimination will be biased toward zero.

To deal with the issue of potential correlation between ticketing on the first margin and discounting, we will use an approach similar to the Heckman (1979) correction. Imagine that all officers working in the same county and year face the same number of drivers of a certain race on a given day of work, N_r. Officers choose whether or not to write a ticket for the driver, Z_{ij}, and thus the daily rate of tickets for that officer for that race-county-year is $N_{rj} = N_r \cdot P(Z_{ij} = 1)$.

Under the presumption that all officers in the same county-year face the same quantity of drivers who could potentially be ticketed for speeding, we can compare officers to calculate their propensity to give a ticket. Within each county-year-race, we calculate the average

[3] We abstract here from the lenient v. non-lenient approach from the main text as well as including observable characteristics. However, when implementing the correction procedure we return to both.

daily number of tickets given by each officer. To account for large right-tail values, we allow the 95th percentile across officers of N_{rj} for each county-year-race to be our value for N_r. Then for each officer-race-county-year, $P(Z_{ij} = 1) = \frac{N_{rj}}{N_r}$, which we call P_{ij}. Using this value, we can identify the expectation for the error term η_{ij} in the ticketing equation for each driver:

$$\begin{aligned} P_{ij} &= Pr(\alpha_j^W + \alpha_j^M \cdot M_i + \eta_{ij} \geq 0) \\ &= \Phi(\alpha_j^W + \alpha_j^M \cdot M_i) \\ \implies E(\eta_{ij}|Z_{ij} = 1) &= \frac{\phi(\alpha_j^W + \alpha_j^M \cdot M_i)}{\Phi(\alpha_j^W + \alpha_j^M \cdot M_i)} \\ &= \frac{\phi(\Phi^{-1}(P_{ij}))}{P_{ij}} \end{aligned}$$

Note that in the uncorrected approach, the conditional expectation of the error term is potentially nonzero because of a correlation with the ticketing error term:

$$\begin{aligned} E(\epsilon_{ij}|\eta_{ij} > -\alpha_j^W - \alpha_j^M \cdot M_i) &= \rho \cdot E(\eta|\eta_{ij} > -\alpha_j^W - \alpha_j^M \cdot M_i) \\ &= \rho E(\eta_{ij}|Z_{ij} = 1) \end{aligned}$$

Therefore, we can address the potential selection into the data using our officer-county-year-race-specific expected value for the ticketing error term, which we call the Heckman Correction term, and re-run the main regression with this addition. The results of this procedure are presented in Table 4. Column 1 presents again the baseline regression, and Column 2 presents the same regression with the additional Heckman Correction term. The addition of the correction does not change the value of the interaction term on Driver White and Officer Lenient or any other coefficients, suggesting that our result is not due to any issues with sample selection. This finding should not be surprising, as we found in the bottom right panel of Figure A.2 that officer lenience is uncorrelated with ticketing frequency.

F.1 Alternate Approach to Evaluating Selection Bias

As noted in Section 6, officers appear to discount some drivers to 14 MPH over and 29 MPH over. Because we can observe drivers who are discounted to 9 MPH over, one approach to evaluating the impact of selection bias induced by lenience is to examine whether discounting to 9 MPH over impacts estimates of discounting to 14 MPH over and 29 MPH over.

To do so, we first use the Frandsen (2017) approach to identify whether officers are lenient at each speed threshold of 9, 14, and 29 MPH over. Namely, we identify an officer as lenient at a threshold if a statistically higher share than 1/3 of tickets around that speed are at the discount point and construct variables Lenient_j^9, Lenient_j^{14}, and Lenient_j^{29}. Because of the substantial reduction in sample size around 29 MPH over, it is not clear we can precisely identify officer lenience at this margin. However, we move forward with the presumption that we can estimate whether an officer is lenient at 29 MPH over.

We then run regression of whether an individual is ticketed at 9 MPH over, using our baseline Diff-Diff regression,

$$\begin{aligned} S_{ij}^9 &= \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2 \cdot \text{Lenient}_j^9 \\ &\quad + \beta_3 \cdot \text{White}_i \cdot \text{Lenient}_j^9 + X_{ij}\gamma + \epsilon_{ij}^9 \end{aligned}$$

and store our residual estimates, $\hat{\epsilon}_{ij}^9$. We then run a similar regression for 14 MPH over, where we restrict attention to drivers who have been ticketed 10 MPH over or higher (i.e. did not get discounted),

$$\begin{aligned} S_{ij}^{14} &= \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2 \cdot \text{Lenient}_j^{14} \\ &\quad + \beta_3 \cdot \text{White}_i \cdot \text{Lenient}_j^{14} + \lambda^9 \hat{\epsilon}_{ij}^9 + X_{ij}\gamma + \epsilon_{ij}^{14} \end{aligned} \tag{3}$$

We include the estimated residual $\hat{\epsilon}_{ij}^9$ from the regression on the discount to 9 MPH over. Analogously, we store our residual estimate, $\hat{\epsilon}_{ij}^{14}$, and then run a similar regression for 29 MPH over, where we restrict attention to drivers who have been ticketed 15 MPH over or higher,

$$\begin{aligned} S_{ij}^{29} &= \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2 \cdot \text{Lenient}_j^{29} \\ &\quad + \beta_3 \cdot \text{White}_i \cdot \text{Lenient}_j^{29} + \pi^9 \hat{\epsilon}_{ij}^9 + \pi^{14} \hat{\epsilon}_{ij}^{14} + X_{ij}\gamma + \epsilon_{ij}^{29} \end{aligned} \tag{4}$$

The coefficients of interest are λ^9, π^9, and π^{14}, which indicate the degree of correlation in discount propensity across the three margins. Note from Appendix Section F that a key determinant of bias in our discrimination estimates is the correlation in treatment propensity across treatment margins. Note also that we are only able to identify the impact of the residuals on different treatment margins because officers have different degrees of lenience at each margin, allowing the residuals to be linearly independent of the officer lenience measures.

We report the results of estimating Equations 3 and 4 in Table A.13. We find an insignificant impact of $\hat{\epsilon}_{ij}^{9}$ on whether an individual is ticketed to 14 MPH over. This insignificant coefficient suggests that individuals with a higher propensity to be discounted to 9 MPH are no more or less likely to be discounted to 14 MPH over conditional on no discount to 9 MPH, which suggests that discrimination in discounting to the lowest margin should not bias our estimates of discrimination in the next lowest margin.

Our estimates of the residual coefficients from Equation 4, presented in Column 2, are negative and statistically significant. These coefficients are of opposite sign of what we would expect based on our intuition and suggest that individuals who are pre-disposed to be discounted to 9 MPH over are *less* likely to be discounted to 29 MPH over. The magnitude of the coefficients is somewhat small; a two standard deviation change in each residual leads to a 0.005 and 0.009 increase in likelihood of a discount to 29 MPH over, respectively.

One possibility for why our estimates in the second regression are in the opposite direction from what we expect is that our measures of the residuals are identified from variation in the lenience of stopping officers, which may be noisily measured. For a fixed set of characteristics and conditional on not receiving a discount to 9 or 14 MPH over, a driver's residual will be lower if their officer is lenient. If an officer's lenience at 9 or 14 is positively correlated with lenience at 29, then some of the negative relationship may be driven by the fact that we have measurement error in officer lenience. Measurement error in officer lenience is likely to be quite high for discounting to 29 MPH over.

Overall, our conclusion from these estimates is that discriminatory discounting from one margin of officer lenience should not substantially bias our estimates of discrimination at higher margins.

G Testing for Statistical Discrimination

Our paper argues that racial disparities in officer lenience reflect bias. However, a compelling alternative explanation is that officers are using race as a signal for an unobserved driver type. Our baseline regressions show that officers differentiate between white and minority drivers after controlling for previous tickets, suggesting that the observed disparity does not reflect statistical discrimination on the level of criminality. However, officers may be sorting individuals on how they *respond* to a discount. For example, officers may be trying to identify drivers who will react to a harsh ticket by speeding less in the future. Alternatively, they may choose to discount a particular driver because they are likely to respond by not contesting the ticket. To formalize these stories, imagine that drivers who are stopped for speeding have some outcome after the ticket, Y_i, that depends on whether a discount D_i is given:

$$Y_i = X_i\beta + \alpha_i D_i + \epsilon_i$$

Whether or not they speed, or contest the ticket, is potentially a function of the treatment given to them by the current stopping officer. As throughout the paper, the officer chooses whether to give a discount, and he does so on the basis of demographics, but also potentially other unobservables:

$$D_i = \mathbb{1}\big(Z_{ij}\theta - v_i \geq 0\big)$$

where Z_{ij} is written to encapsulate both the individual covariates X_i and an instrument for discounting based on the officer identity, which we discuss below. The story we are interested in testing is whether officers choose who to discount on the basis of how Y_i responds. In other words, do we have $\alpha_i \perp\!\!\!\perp D_i | X_i$ or not? Heckman, Schmierer and Urzua (2010) provide a number of tests for whether there is such a correlation, from which we borrow directly below. In particular, they show that a lack of correlation between discounting and treatment effect implies a linear relationship between the outcome and propensity score for treatment. To see this, we first reformulate the discount equation:

$$\mathbb{1}(D = 1) = \mathbb{1}(v \leq Z_{ij}\theta) = \mathbb{1}(F_v(v) \leq F_v(Z_{ij}\theta)) = \mathbb{1}(U_d \leq P(Z_{ij}))$$

where U_d is a uniform random variable and $P(z_{ij}) = Pr(D = 1|Z_{ij} = z_{ij})$ is the propensity score. The marginal treatment effect is defined as the treatment effect for an individual at a given propensity to be treated (Björklund and Moffitt, 1987):

$$MTE(x, u_d) = E(\alpha_i | X = x, U_d = u_d)$$

The conditional expectation of Y_i as a function of X_i and Z_i can then be written as a function of the marginal treatment effects:

$$\begin{aligned}
E(Y|Z=z) &= X_i\beta + E(\alpha_i D_i|z) \\
&= X_i\beta + E(\alpha_i D_i|P(z)) \\
&= X_i\beta + E(\alpha_i|D=1, P(z)) \cdot p \\
&= X_i\beta + \int_0^p E(\alpha_i|U_d = u_d)du_d
\end{aligned}$$

Under no correlation between α_i and D_i, then $E(\alpha_i|U_d = u_d) = E(\alpha_i)$. Therefore, the conditional expectation of Y_i should be linear in $P(z)$:

$$\begin{aligned}
\frac{\partial E(Y|Z=z)}{\partial P(z)} &= E(\alpha_i|U_d = P(z)) \\
&= E(\alpha_i) \quad \text{under} \quad \alpha_i \perp\!\!\!\perp D_i|X_i
\end{aligned}$$

Therefore, a test for the linearity of Y_i in $P(Z)$ tells us whether officers are sorting individuals on the basis of their treatment effect of D_i on Y_i. Under linearity, the marginal treatment effects of individuals with different propensities to be treated (in our case, stopped by different officers) will be the same.

The instrument Z_{ij} we use for whether an individual receives a discount is based on the identify of the officer and is a leave-out measure of the officer's propensity to give a discount:

$$Z_{ij} = \frac{1}{N_j - 1} \sum_{k \in \mathcal{J} \setminus i} D_k$$

where N_j is the number of individuals stopped by officer j. This average-lenience-of-treater instrumenting design has been used in various settings to study the effect of criminal sentence length (Kling, 2006; Mueller-Smith, 2014), bankruptcy protection (Dobbie and Song, 2015), foster care (Doyle, 2007; Doyle Jr, 2008), and juvenile incarceration (Aizer and Doyle Jr, 2015).

We then calculate an individual's propensity to receive a discount based on their stopping officer and demographic characteristics. Because an officer's lenience can vary with the race of the driver, we interact the instrument with driver race:

$$P(Z, X) = X_i\gamma + \theta^0 Z_{ij} + \theta^M \text{DriverMinority}_i Z_{ij}$$

We then run regressions of Y_i on specifications that are linear and quadratic in $P(z, x)$, where the outcomes we consider are whether a driver receives another ticket in the year following the FHP stop[4] and whether the driver contests the ticket.

The results of this analysis are presented in Table 5. We restrict attention to in-state drivers with a ticket at 9 MPH or over for whom we have a court record of whether the driver contested. These restrictions leave us with 1,513,913 tickets. The first two columns treat an individual's recidivism as the outcome. In Column 1 we see that an increase in the probability of receiving a discount increases an individual's likelihood of recidivating.[5] In the second column, we see that a quadratic term for the propensity score is also significant. However, the effect is negative, suggesting that the impact of a discount on recidivism is largest for the drivers with a low propensity score, which are those who are stopped by a harsh officer. This finding runs counter to selection on gains. The "first" drivers to receive a discount (those who are at the margin of receiving a discount by harsh officers) are those with the largest recidivism response to receiving a discount. If officers were discounting on the basis of deterrability, the first drivers to be discounted would have the smallest recidivism effect.

Columns 3 through 5 use as an outcome whether the driver contests the ticket in court. As with recidivism, we find an effect of receiving a discount: drivers stopped by officers who are more likely to give discounts are less likely to contest their ticket. When we add a quadratic term in Column 4, we also find a non-linear relationship, with the quadratic having a significant positive coefficient. Drivers stopped by very harsh officers have a larger marginal response to discounting than drivers stopped by less harsh officers. This finding is consistent with officers picking who to discount in order to avoid court contestations.

The intuition for this result is the following: imagine an officer who is very lenient toward his drivers. If he is going to be harsh to one driver, he will pick someone who is not very responsive to a harsh ticket and will not contest. We will thus see that officer have a small effect of discounting on contesting. In contrast, imagine an officer who is harsh toward

[4]The recidivism of the driver is calculated as an indicator for whether they receive any traffic ticket in the state of Florida in the following year. We link drivers by driver's license number. More information is available in Goncalves and Mello (2017).

[5]Though we do not report it here, the first-stage coefficient on the instrument is close to 1 and slightly smaller for minority drivers. The first-stage relationship is essentially linear, indicating that any non-linearity in the reduced form regressions presented here are not due to differences in the strength of the instrument at different levels.

nearly all drivers. If he is going to give a break to someone, that discount should give him a large return in reduced court time. We should thus expect a large contest response among that officer's drivers. Our findings are thus consistent with the story that officers do try to identify driver's propensity to not contest their ticket.

While we therefore do find evidence of statistical discrimination on court contest response, our primary objective is to determine whether any form of statistical discrimination can explain the disparity we observe between whites and minorities. To do so, we implement a test based on Arnold, Dobbie and Yang (2018) and Marx (2018). They implement the logic of the Becker (1957) hit-rate test in the random-judge design and show that, under no discrimination, the impact of a treatment should be the same at the margin across racial groups. To conduct this test, we interact the propensity score with the race of driver in Column 5. Doing so, we find that the marginal effect of a discount on contesting is statistically larger for minorities than for white drivers, indicating that the discrimination we observe cannot be explained by sorting on contest response.

H Notes on Model Estimation

While the setup of the model is simple, the non-parametric identification of the distribution of officer bias and the distribution of county-by-race speeds leads to a significant number of parameters to be identified. We estimate the model through maximum likelihood, programmed in Matlab. We provide the program with the gradient vector and utilize "fminunc" with a quasi-newton search algorithm option. The variance matrix of the parameters is calculated as the inverse of the information matrix, which we calculate as the variance of the score functions.

One issue to note is that the log likelihood function is essentially flat for certain regions of the parameter space for some officer preference parameter values. This flatness occurs because some officers have no (all) drivers at the bunch point, consistent with an infinitely negative (positive) "t." The optimization algorithm reaches values that are large in magnitude. However, because the score function is essentially flat at these large values, the parameters's standard errors are extremely large.

To deal with this issue, we treat these parameters (specifically, the t estimates for officers with $P(Discount \mid X = 10) < 0.02$ or $P(Discount \mid X = 10) > 0.98$) as known and set their variances to be zero.

H.1 Model Estimates Discussion

Table A.14 presents estimates of the model parameters. Columns in the top panel present the mean and variance of each class of parameters, broken down by race, and the final column compares differences across racial groups in the mean parameter estimates. The slope parameter is positive and significant at 0.0425. Consistent with our intuition, officers face an upward-sloping cost with respect to speed, meaning that tickets are less likely to be discounted the higher the observed speed. The parameter t represents an officer's mean valuation of a racial group. We find both significant heterogeneity and a significant disparity across whites and minorities in how officers value discounting drivers, with officers' mean valuation for whites being 0.0185 higher than for minorities.

While the values of t are by themselves hard to interpret, the racial differences in treatment are more easily understood in terms of the probability of discount (i.e., fine reduction). $Pr(Discount^9|S = 9, E(Z))$ represents the likelihood of receiving a reduced ticket if the driver is at the speed right above the bunching speed, where, besides race, the driver has the average demographics Z. Consistent with the reduced-form evidence, the average officer is substantially lenient, with a large variance across officers. Officers are 2 percentage points less likely to discount minorities than whites, off a baseline of 33.9% likelihood of discount. Figure A.7 further shows this disparity, highlighting how racial bias results in a decreased mass of officers with very high lenience and an increase in mass of officers with very low lenience. Figure A.8 shows how the disparity only arises among officers with some degree of lenience.

The λ estimates tell us how races-by-counties differ in their underlying speeds prior to officers' choice of lenience. As we found in Section 4 when restricting our attention to non-lenient officers, model estimates suggest that minorities on average drive significantly faster than whites, on the order of 0.43 MPH. Figure A.9 presents this gap by county, showing that minority speeds stochastically dominate white speeds. These results are in line with previous studies of highway patrol ticketing, which argue that much of the gap in ticketing between whites and minorities can be explained by higher speeds by minorities (Smith et al., 2004; Lange, Johnson and Voas, 2005). However, these previous studies and the news coverage that followed implicitly argued that the racial difference in speeds rules out the presence of bias by officers. Our study highlights how this thinking is incorrect by showing that disparities in driving and racial bias coexist in our setting. As shown in Figure A.10, the distribution of bias across officers looks very similar to the distribution found in our reduced form estimates from Section 4.

The bottom panel of Table A.14 presents the demographic-specific speed and preference parameters. Female drivers, older drivers, and those with fewer tickets all drive slower speeds on average and are more likely to be discounted. The effect of county minority share indicates that officers are less likely to discount everybody in a more minority neighborhood, regardless of the race of the stopped individual.

We report in Figure A.11 various estimates of model fit to the data. For each panel, we construct the model statistics by simulating 100 times and averaging across iterations. The top left panel compares the aggregate histograms of speeds. The top right panel compares the average ticketed speeds by race-county. The bottom left panel compares the share of tickets at 9 MPH over by officer-race. The bottom right panel compares the racial disparity in bunching at 9 MPH over by officer. In all cases, the model estimates match very closely with the true data.

H.2 Counterfactuals

Here we provide information on how the counterfactuals and their standard errors are calculated. There are several sources of uncertainty in the estimation that lead to standard errors on our calculations: 1) uncertainty of our parameter estimates, 2) randomness of the matching between officers and drivers, 3) randomness in the speed draws for the drivers, and 4) randomness in the officers' decisions to discount. We therefore calculate standard errors through a sampling procedure as follows:

- Draw a sample of parameters $\theta^{(1)} \sim N(\hat{\theta}, \hat{\Sigma})$, where $\hat{\theta}$ and $\hat{\Sigma}$ are our parameter point estimates and variance matrix, respectively.

- Within each county, randomly match officers and drivers. In the baseline estimation, the probability of encountering an officer is the share of tickets in the data which that officer gave. All the counterfactuals consist of changing the distribution of officers being matched.

- Drivers draw a speed from their Poisson distribution, $s \sim P_{\lambda_i}$.

- We draw a set of $\epsilon_{ij} \sim N(0,1)$ for all stops, and an officer discounts her driver if $t_{rj} + \alpha Z^{(2)} + \epsilon_{ij} > b \cdot s$.

- Iterate 100 times.

Then, our estimates and standard errors for the racial gaps in each counterfactual are the average and standard deviation across all iterations.

Here we describe explicitly how each counterfactual is performed:

- Decomposition with no sorting: Rather than matching drivers and officers randomly within a county, they are matched randomly across the entire state.

- Decomposition with no bias: Identical to the baseline, officers and drivers are matched randomly within a county. Officers preferences for minority drivers is set to be their white preference, t_{wj}.

- Firing 5% most discriminatory officers: Calculate $P_j^{\text{bias}} \equiv Pr_j(Discount \mid X = 10, E(Z), r = w) - Pr_j(Discount \mid X = 10, E(Z), r = m)$, and find the 5th percentile for the entire state and "remove" all officers below this threshold. We also remove officers that cross the same threshold of discrimination against *white* drivers. The probability of an individual encountering a specific officer is that officer's share of tickets among the remaining officers.

- Hiring more minority officers: We increase the share from 35% to 45%. We do so by proportionately increasing the number of minority officers in each county. e.g. a county that previously was 10% minority officers is now 16% minority. The distribution of officer tastes t_{rj} is the same as the existing distribution *within* officer race. The procedure is identical for increasing the share of female officers.

- Re-assigning officers based on discrimination: Within a troop, officers are ranked based on their discrimination. In the county of that troop with the most minorities, the lowest-ranked officers are assigned. The second-most minority county receives the next-least discriminatory officers, and so on. Officers write as many tickets as in the true data, so some officers may write tickets in two counties that are adjacent in their share minority. The procedure is identical when assigning officers based on their lenience, where the most lenient officers are assigned to the most minority neighborhoods.

I Model Extension with 14-MPH Bunching and Multiple-of-5 Heaping

In the primary model presented in Section 8, we assume that officers face a single choice of whether to reduce the driver's speed to 9 MPH over in cases where the observed speed is higher. Here we allow for two more margins of choices for the officer. In addition to discounting to 9MPH over, he may also discount the driver to 14 MPH over, or he may reduce the speed to the closest multiple of five. We will call this latter practice "heaping." To add these two additional margins of decision-making, we will simplify the degree of officer heterogeneity by allowing officers in each county to be one of only two types, as described below.

We have the following model of the officer's decision problem. The officer, who we denote by j, observes a driver, denoted by i, with speed s^* above the limit, drawn from a poisson distribution with mean value λ_i. The officer has a mean valuation towards the driver's demographic group, $t_{X(i)j}$, and an idiosyncratic preference for the particular driver, $\epsilon_i \sim N(0, 1)$.

If the observed speed is at or below 9, the officer reports the true speed. If the speed is above 9, the officer discounts the driver to 9 if $t_{X(i)j} + \epsilon_i > b \cdot s^*$. If the observed speed is above 14 and the officer's valuation of the driver is not sufficiently high to discount the driver to 9, he will discount the driver to 14 if $t_{X(i)j} + \epsilon_i > b \cdot s^* - C_{14}$. If the driver is stopped at a speed above 15 and has not been discounted to either 9 or 14, the officer will write down the closest multiple of five below s^* with probability h. This "heaping probability" is exogenous and independent of s^*, $t_{X(i)j}$, and ϵ_i. In other words, the ticketed speed S will be a function of s^*, $t_{X(i)j}$, and ϵ_i:

$$S = \begin{cases} s^* & \text{if} \quad s^* \leq 9 \\ 9 & \text{if} \quad s^* \geq 10 \quad \& \quad t_{X(i)j} + \epsilon_i > b \cdot s^* \\ 14 & \text{if} \quad s^* \geq 15 \quad \& \quad t_{X(i)j} + \epsilon_i \in (b \cdot s^* - C_{14}, b \cdot s^*] \\ s^* & \text{if} \quad s^* \in [10, 14] \quad \& \quad t_{X(i)j} + \epsilon_i \leq b \cdot s^* \\ s^* \quad \text{w/ prob } (1-h) & \text{if} \quad s^* \geq 15 \quad \& \quad t_{X(i)j} + \epsilon_i \leq b \cdot s^* - C_{14} \\ 5 \times \text{floor}(s^*/5) \quad \text{w/ prob } h & \text{if} \quad s^* \geq 15 \quad \& \quad t_{X(i)j} + \epsilon_i \leq b \cdot s^* - C_{14} \end{cases}$$

We will conduct all estimation at the county level, so all parameters will be estimated separately in each county. We will allow the driving speed poisson parameter to be a function of the driver's race, gender, age, and any previous tickets, $\lambda_i = \lambda_r + X\beta$. For the officer's preference parameters, we will make some simplifications relative to the baseline model in Section 6. We will assume that the officer is drawn from one of two types, $\pi \in \{NL, L\}$, which we call the non-lenient type and the lenient type. When presenting the mean valuation of an officer, type will be denoted by a superscript. For $\pi = NL$, the officer has mean valuation of $t^{NL} = -5$ for all drivers, which in practice corresponds to no drivers receiving a break from the officer. If the officer is of type $\pi = L$, he has mean valuations that may vary by race of driver, t^L_W, t^L_M.

We will estimate the model using maximum likelihood. As before, the log likelihood function includes summations over the possible true values s^*. However, this fact makes the model more challenging to estimate due to the greater number of ticketed speeds S for which drivers may have been drawn from multiple speeds. Further, the fact that we are drawing officers from two types rather than using fixed effects requires a summation over officer type.

To address these issues, we will use the Expectation-Maximization Algorithm. Intuitively, this approach treats the true speed s^* and officer type π as missing data, expands the data to have an observation for every true observation in our data and a potential s^* and potential π, and weights these observations by the probability of that s^* and π being the true values, which we denote by $\gamma_{s^*\pi i}$. After estimating the model parameters using these given weights, we then update the weights using the new values for the model parameters. We do this process until the values of the model parameters and probabilities of the unobserved data converge. Formally, we do the following:

1. Pick starting guesses for all parameters, θ^0. Use these values to construct the probability of any given s^* for each individual, $\gamma_{s^*i} = Poiss_{\lambda_i^0}(s^*)$, and set the initial probability of each individual's officer being lenient, $\gamma_{\pi j(i)}$, to be $\gamma_\pi^0 = 0.5$. We then set $\gamma_{s^*\pi i} = \gamma_{s^*i} \times 0.5$.

2. Construct the log likelihood conditional on potential speed and officer type,

$$l_\theta(x) = \sum_i \sum_{s^*\pi} l_\theta(x|s,\pi) \cdot \gamma^0_{s^*\pi i}$$

   and solve for model parameters $\theta^1 = \{\lambda_{rc}, \beta, t^L_{cW}, t^L_{cM}\}$ that maximize the log likelihood.

3. Update the probability of each unobserved type for each observation, where we condition on all realized ticketed speeds S for all individuals:

$$\begin{aligned}
\gamma^1_{s^*\pi i} &= Pr(s_i^*, \pi_{j(i)}|S) \\
&= Pr(s_i^*|\pi_{j(i)}, S) \cdot Pr(\pi_{j(i)}|S) \\
Pr(s_i^*|\pi_{j(i)}, S) &= \frac{Pr(S_i|\pi_{j(i)}, s_i^*) \cdot Poiss_{\lambda^1}(s_i^*)}{\sum_{s'} Pr(S_i|\pi_{j(i)}, s') \cdot Poiss_{\lambda^1}(s')} \\
Pr(\pi_{j(i)}|S, i) &= \frac{Pr(S_i|\pi) \cdot \gamma^0_\pi}{\sum_{\pi'} Pr(S_i|\pi') \cdot \gamma^0_{\pi'}} \\
\text{where} \quad Pr(S_i|\pi) &= \prod_i \prod_{s^*} \gamma^0_{s^*\pi i} \cdot Pr(S_i|\pi, s^*)
\end{aligned}$$

4. Return to 1. with new observation weights $\gamma^1_{s^*\pi i}$.

Because we are conducting the analysis at the county level, we need each officer to have sufficient observations in each county. We therefore restrict attention to tickets that are issued by officers with at least ten tickets in the ticket's county.

Model Estimates

The parameter estimates are presented in Table A.16. The top panel presents all the parameters of officer preferences, while the bottom panel presents the parameters of driver speeds. Consistent with the previous model, we find that officers have a discounting slope b of 0.0262 (v. 0.0425). In the average county, drivers who are stopped at 15 MPH over have a 38% chance of being discounted to 9 MPH over and a 3.4% chance of being discounted to 14 MPH over.

Because our estimates are at the county level, they are not directly comparable to the baseline model, which estimates discrimination at the officer level. In Figure A.12, we average over all individuals in the state and present the estimated probabilities of discounting from each stopped speed from each model specification. Despite the fact that the extended model allows for a probability of discounting to 14 MPH over, we find that the difference in probability of discounting to 9 MPH over is less than one percentage point between the two model specifications.

Figure A.13 plots the empirical distribution of ticketed speeds separately for white and minority drivers, where we overlay the distribution of ticketed speeds generated by the extended model. We find that, in addition to matching the frequency of tickets at the discount point, our model matches well the share of tickets at multiples of five. Because we use one parameter per county to match these frequencies, we do not perfectly match each multiple of five but are able to match the average frequency at multiples of five.

To further evaluate the fit of the model across counties, Figure A.14 plots the empirical and simulated share of tickets at different points in the speeding distribution. We find that the model is able to fit the variation across counties in the share of tickets at 9 MPH over, 14 MPH over, and at multiples of five.

In light of these estimates, we conclude that an extended version of our model can fit discounting to 14 MPH over and heaping of tickets to multiples of five. Further, this extension leads us to the same conclusion about discounting to 9 MPH over, lending additional confidence to our baseline model and the abstractions we make from these additional model features.

Table A.1: Racial Disparity in Speeding

	Full Sample			GPS Sample	
	(1) MPH Over	(2) MPH Over	(3) MPH Over	(4) MPH Over	(5) MPH Over
Driver Black	1.031	0.708	0.598	0.739	0.681
	(0.0712)	(0.0318)	(0.0285)	(0.0481)	(0.0475)
Driver Hispanic	2.715	0.794	0.637	0.885	0.667
	(0.122)	(0.0358)	(0.0327)	(0.0649)	(0.0543)
Female			-0.526	-0.371	-0.341
			(0.0186)	(0.0334)	(0.0349)
FL License			-0.310	-0.481	-0.380
			(0.0353)	(0.0601)	(0.0565)
Age/100			-4.095	-3.914	-3.576
			(0.111)	(0.131)	(0.136)
1 Prior Ticket			0.192	0.154	0.156
			(0.0134)	(0.0342)	(0.0358)
2+ Prior Ticket			0.643	0.583	0.541
			(0.0167)	(0.0359)	(0.0367)
Any Past Prison Spell			0.714	0.724	0.680
			(0.0543)	(0.124)	(0.139)
Log Zip Income			0.0232	-0.0515	-0.0823
			(0.0212)	(0.0319)	(0.0292)
Zip Income Missing			0.946	0.951	0.482
			(0.104)	(0.171)	(0.142)
Mean	16.210	16.229	16.229	15.769	15.851
Vehicle FE			X	X	X
Location + Time FE		X	X	X	
GPS FE					X
Observations	2122555	2078315	2078315	262369	225355

Notes: Table reports regressions where the outcome is the speed for which the individual is ticketed. "Location FE" are fixed effects at the county by posted speed limit. "Location + Time FE" are fixed effects at the county by posted speed limit by year by month by day of week by hour fixed effects. "GPS FE" are fixed effects at the road segment by posted speed limit by year by month by day of week by hour fixed effects. GPS sample are tickets with the GPS location available. Standard errors are clustered at the county level.

Table A.2: Racial Disparity in Discounting

	Full Sample			GPS Sample	
	(1) MPH Over	(2) MPH Over	(3) MPH Over	(4) MPH Over	(5) MPH Over
Driver Black	-0.0269	-0.0229	-0.0213	-0.0296	-0.0282
	(0.00567)	(0.00261)	(0.00233)	(0.00401)	(0.00378)
Driver Hispanic	-0.140	-0.0387	-0.0346	-0.0480	-0.0338
	(0.00841)	(0.00288)	(0.00259)	(0.00510)	(0.00462)
Female			0.0251	0.0183	0.0182
			(0.00159)	(0.00273)	(0.00280)
FL License			0.0157	0.0252	0.0190
			(0.00326)	(0.00471)	(0.00458)
Age/100			0.123	0.128	0.110
			(0.00979)	(0.0120)	(0.0116)
1 Prior Ticket			-0.00813	-0.00802	-0.00965
			(0.000998)	(0.00264)	(0.00265)
2+ Prior Ticket			-0.0252	-0.0225	-0.0220
			(0.00157)	(0.00294)	(0.00290)
Any Past Prison Spell			-0.0256	-0.0287	-0.0256
			(0.00359)	(0.00782)	(0.00889)
Log Zip Income			-0.00608	-0.00106	0.000341
			(0.00165)	(0.00235)	(0.00208)
Zip Income Missing			-0.0594	-0.0574	-0.0239
			(0.00787)	(0.0129)	(0.0108)
Mean	16.210	16.229	16.229	15.769	15.851
Vehicle FE			X	X	X
Location + Time FE		X	X	X	
GPS FE					X
Observations	2122555	2078315	2078315	262369	225355

Notes: Table reports regressions where the outcome is an indicator for the individual being ticketed at 9MPH over the limit. "Location FE" are fixed effects at the county by posted speed limit. "Location + Time FE" are fixed effects at the county by posted speed limit by year by month by day of week by hour fixed effects. "GPS FE" are fixed effects at the road segment by year by month by day of week by hour fixed effects. GPS sample are tickets with the GPS location available. Standard errors are clustered at the county level.

Table A.3: Racial Disparity in Speeding, Non-lenient Officers

	Full Sample			GPS Sample	
	(1) MPH Over	(2) MPH Over	(3) MPH Over	(4) MPH Over	(5) MPH Over
Driver Black	1.376	0.624	0.488	0.693	0.577
	(0.110)	(0.0514)	(0.0481)	(0.0858)	(0.0912)
Driver Hispanic	1.800	0.429	0.272	0.382	0.344
	(0.144)	(0.0341)	(0.0337)	(0.0596)	(0.0598)
Female			-0.465	-0.366	-0.295
			(0.0303)	(0.0514)	(0.0540)
FL License			-0.309	-0.330	-0.215
			(0.0611)	(0.0742)	(0.0899)
Age/100			-4.013	-3.035	-2.926
			(0.121)	(0.176)	(0.192)
1 Prior Ticket			0.208	0.132	0.0739
			(0.0228)	(0.0601)	(0.0567)
2+ Prior Ticket			0.584	0.488	0.438
			(0.0280)	(0.0565)	(0.0659)
Log Zip Income			-0.0673	-0.0386	-0.0651
			(0.0256)	(0.0555)	(0.0533)
Mean	19.714	19.746	19.746	19.116	19.192
Vehicle FE			X	X	X
Location + Time FE		X	X	X	
GPS FE					X
Observations	526478	506236	506236	49846	41562

Notes: Table reports regressions where the outcome is the speed for which the individual is ticketed, restricting attention only to non-lenient officers. ”Location FE” are fixed effects at the county by posted speed limit. ”Location + Time FE” are fixed effects at the county by posted speed limit by year by month by day of week by hour fixed effects. ”GPS FE” are fixed effects at the road segment by year by month by day of week by hour fixed effects. Standard errors are clustered at the county level.

Table A.4: Diff-Diff Regressions by Whether Any Additional Citations Issued

	(1) Discount	(2) Discount	(3) Discount
Driver White	-0.0154	-0.0144	-0.0181
	(0.00478)	(0.00480)	(0.00540)
Officer Lenient	0.211	0.232	0.0866
	(0.0323)	(0.0328)	(0.0454)
Driver White	0.0592	0.0557	0.0625
× Officer Lenient	(0.00626)	(0.00626)	(0.00814)
Specification	Baseline	No Other Citation	Any Other Citation
Location + Time FE	X	X	X
Mean	0.313	0.321	0.251
R2	0.363	0.369	0.346
Observations	2078315	1773541	264540

Notes: Table presents estimates from regressions of Equation 3. Column 2 restricts attention to stops where the driver is not issued any citations other than speeding, and Column 3 restricts attention to all other stops, i.e. cases where at least one other citation is issued.

Table A.5: Difference-in-Differences Officer-Level Results

A. Lenient + Non-Lenient	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%	(6) N
All Officers	-0.017	0.000	0.003	0.053	0.110	1851
White Officers	-0.013	0.000	0.017	0.068	0.122	1168
Black Officers	-0.037	-0.002	0.000	0.014	0.056	292
Hispanic Officers	-0.014	0.000	0.000	0.030	0.080	364

B. Lenient Only	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%	(6) N
All Officers	-0.029	-0.000	0.033	0.075	0.125	1289
White Officers	-0.023	0.006	0.043	0.089	0.136	877
Black Officers	-0.070	-0.021	0.001	0.040	0.071	174
Hispanic Officers	-0.032	-0.004	0.021	0.050	0.111	214

Notes: Table reports percentiles of the distribution of officer-level discrimination, as calculated from Equation 4.

Table A.6: Share Discriminatory by Officer Race

	A. *Lenient + Non-Lenient*		
	(1) Share	(2) Share "Reverse"	(3) N
All Officers	0.420 (0.012)	0.084 (0.008)	1851
White Officers	0.503 (0.017)	0.062 (0.009)	1168
Black Officers	0.229 (0.028)	0.174 (0.025)	292
Hispanic Officers	0.297 (0.027)	0.079 (0.018)	364
Other Officers	0.588 (0.124)	0.134 (0.089)	27

	B. *Lenient Only*		
	(1) Share	(2) Share "Reverse"	(3) N
All Officers	0.603 (0.017)	0.121 (0.011)	1289
White Officers	0.670 (0.023)	0.082 (0.012)	877
Black Officers	0.384 (0.047)	0.292 (0.043)	174
Hispanic Officers	0.505 (0.046)	0.134 (0.030)	214
Other Officers	0.662 (0.139)	0.151 (0.100)	24

Notes: Table reports share of officers who are discriminatory and reverse discriminatory, as described in Section 4.1. Estimates in Panel B are constructed by dividing the share discriminatory in Panel A by the share lenient officers for each race of officer.

Table A.7: Officer Discrimination Randomization Check

	Full Sample	GPS Sample	
	(1) Discrimination	(2) Discrimination	(3) Discrimination
Driver Black	0.000918	0.00181	0.000660
	(0.000305)	(0.000743)	(0.000340)
Driver Hispanic	-0.000525	-0.000858	0.0000813
	(0.000334)	(0.000930)	(0.000348)
Female	-0.000157	0.0000932	0.000126
	(0.000155)	(0.000263)	(0.000196)
FL License	-0.000293	0.00000740	0.00000868
	(0.000318)	(0.000799)	(0.000282)
Age/100	-0.000502	-0.000596	-0.000428
	(0.000876)	(0.00151)	(0.000818)
1 Prior Ticket	0.000110	0.000296	0.000258
	(0.0000954)	(0.000268)	(0.000284)
2+ Prior Ticket	0.000229	0.000536	0.0000240
	(0.000139)	(0.000345)	(0.000229)
Any Past Prison	0.000130	0.0000118	-0.000347
	(0.000398)	(0.000949)	(0.000829)
Log Zip Income	0.000273	0.000376	0.0000120
	(0.000256)	(0.000495)	(0.000202)
Zip Income Missing	0.000269	-0.00437	-0.00218
	(0.00136)	(0.00256)	(0.00146)
Vehicle Price / 1000	0.0000139	0.00000451	0.0000172
	(0.0000132)	(0.0000328)	(0.0000205)
Vehicle Age	0.0000403	0.0000184	0.00000854
	(0.0000240)	(0.0000431)	(0.0000323)
F-val	2.568	1.857	1.135
F-test	0.001	0.027	0.322
Mean	0.031	0.037	0.037
Location + Time FE	X	X	
GPS FE			X
Observations	2078315	262369	225355

Notes: All regressions includes vehicle type fixed effects and county fixed effects. The F-test reports the joint hypothesis test that variables Driver Black through Log Zip Code Income are zero. Standard errors are clustered at the county level. "Location FE" includes county by highway fixed effects. "Location + Time FE" includes county by highway by year by month by day of the week by shift fixed effects. "GPS FE" includes road segment by county by highway by year by month by day of the week by shift fixed effects.

Table A.8: Characteristics of Cited Drivers Relative to Other Data Sources

	(1) Citations	(2) ACS - Any	(3) ACS - Drivers	(4) Crash - Any	(5) Crash - Injury
Female	0.371	0.514	0.478	0.436	0.453
Age	0.354	47.938	42.110	40.507	40.529
White	0.598	0.635	0.616	0.548	0.573
Black	0.170	0.146	0.144	0.195	0.203
Hispanic	0.232	0.219	0.240	0.257	0.223

Notes: ACS data are from 2006-2010 include individuals aged 16 or older and sampling weights are used. We obtained these data from Integrated Public Use Microdata Series (Ruggles et al., 2020). So that the samples are parallel, we use only citations and accidents from 2006-2010 and keep only white, Black, or Hispanic individuals aged 16 or over in the ACS. We use sampling weights when computing the shares from the ACS data. To match to the shares in our data, we restrict attention to ACS respondents who report their race as white, Black, or Hispanic. Crash records acquired from FDOC (2010).

Table A.9: Predicting Officer Discrimination

	(1) White Lenience	(2) Black Discrimination	(3) Hispanic Discrimination
Black Officer	-0.029	-0.015	-0.012
	(0.014)	(0.002)	(0.002)
Hispanic Officer	-0.032	-0.009	-0.011
	(0.014)	(0.002)	(0.002)
Other Race	0.001	0.010	-0.002
	(0.044)	(0.010)	(0.007)
Female Officer	-0.043	-0.003	-0.006
	(0.016)	(0.002)	(0.002)
Age / 10	0.005	0.001	0.001
	(0.009)	(0.001)	(0.001)
Experience / 10	0.065	0.001	0.003
	(0.034)	(0.004)	(0.005)
Failed Entrance Exam	0.012	-0.001	-0.002
	(0.018)	(0.002)	(0.002)
Any College	-0.012	-0.001	-0.001
	(0.012)	(0.002)	(0.002)
Number of Civilian	-0.005	-0.000	-0.000
Complaints	(0.002)	(0.000)	(0.000)
Use of Force Incidents	-0.007	-0.000	-0.000
	(0.004)	(0.001)	(0.001)
Mean	0.214	0.022	0.035
Observations	1,630	1,630	1,630
R2	0.213	0.097	0.112

Notes: Robust standard errors in parentheses. Outcomes are derived from the regression $S^9_{ij} = \beta_0 + \beta_1 \cdot \text{Black}_i + \beta_2 \cdot \text{Hispanic}_i + \beta^j_3 \cdot \text{Lenient}_j + \beta^j_B \cdot \text{Black}_i \cdot \text{Lenient}_j + \beta^j_H \cdot \text{Hispanic}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$. White Lenience is calculated as $\beta_0 + \beta^j_3 \text{Lenient}_j$. Black Bias and Hispanic Bias are calculated as $\beta^j_B \cdot \text{Lenient}_j$ and $\beta^j_H \cdot \text{Lenient}_j$, respectively. The sample of officers is reduced from 1591 to 1411 because of the restriction that each officer stop both Black and Hispanic drivers.

Table A.10: Predicting Officer Complaints/Force

	(1) # Complaints	(2) Any Complaints	(3) # Use of Force	(4) Any Use of Force
Lenience	-0.618	-0.156	-0.146	-0.0504
	(0.180)	(0.0523)	(0.126)	(0.0451)
Discrimination	-0.306	-0.0595	0.198	0.0349
	(0.636)	(0.207)	(0.490)	(0.174)
Black Officer	0.102	0.00211	-0.164	-0.0928
	(0.153)	(0.0362)	(0.0753)	(0.0291)
Hispanic Officer	-0.00261	-0.00649	0.0220	-0.0164
	(0.129)	(0.0346)	(0.0823)	(0.0325)
Other Race	0.278	0.0792	-0.165	-0.0665
	(0.352)	(0.0934)	(0.156)	(0.0861)
Female Officer	-0.222	-0.0869	-0.0674	-0.0174
	(0.131)	(0.0429)	(0.0826)	(0.0360)
Age / 10	-0.154	0.127	-0.524	-0.147
	(0.274)	(0.0812)	(0.163)	(0.0680)
Age Squared	0.0207	-0.0182	0.0357	0.00803
	(0.0396)	(0.0117)	(0.0204)	(0.00927)
Experience / 10	-0.288	-0.198	-0.592	-0.0238
	(0.368)	(0.117)	(0.290)	(0.101)
Exp Squared	-0.0220	0.0286	0.0284	0.00802
	(0.0671)	(0.0226)	(0.0434)	(0.0172)
Failed Entrance Exam	0.241	0.0329	-0.0673	0.0140
	(0.185)	(0.0449)	(0.104)	(0.0416)
Any College	-0.194	-0.0231	0.0982	0.00820
	(0.0921)	(0.0274)	(0.0836)	(0.0244)
Sought Promotion	-0.0965	-0.0515	-0.0732	-0.00469
	(0.108)	(0.0272)	(0.0880)	(0.0247)
Mean	1.186	0.528	0.518	0.278
Observations	1630	1630	1630	1630
Regression	OLS	OLS	OLS	OLS

Notes: Heteroskedasticity-robust standard errors in parentheses. Column title indicates the dependent variable. Data is at the officer level. Regressions have indicator variables for years when and districts where the officer worked.

Table A.11: Diff-Diff Regressions by Additional Subsamples

	(1) Discount	(2) Discount	(3) Discount	(4) Discount	(5) Discount
Driver White	-0.0154	-0.0168	-0.0146	-0.0158	0.000952
	(0.00478)	(0.00551)	(0.00473)	(0.00671)	(0.00450)
Officer Lenient	0.211	0.0145	0.217	0.239	0.206
	(0.0323)	(0.0603)	(0.0322)	(0.0227)	(0.0402)
Driver White	0.0592	0.0594	0.0581	0.0654	0.0458
× Officer Lenient	(0.00626)	(0.00736)	(0.00624)	(0.00784)	(0.00688)
Specification	Baseline	Young Men	All Others	Nighttime	Daytime
Location + Time FE	X	X	X	X	X
Mean	0.313	0.272	0.322	0.347	0.310
R2	0.363	0.320	0.374	0.354	0.321
Observations	2078315	445501	1586615	926253	289942

Notes: Table presents estimates from regressions of Equation 3. Column 2 restricts the sample to young male drivers, and Column 3 restricts the sample to all other observations. Column 4 restricts the sample to stops made at night (7pm-5am), and Column 5 restricts the sample to stops made during the day (7am-5pm).

Table A.12: Difference-in-Differences Results by Black and Hispanic Drivers

	Full Sample		GPS Sample	
	(1) Discount	(2) Discount	(3) Discount	(4) Discount
Driver Black	0.00730	0.00477	0.00651	0.00153
	(0.00333)	(0.00311)	(0.00520)	(0.00337)
Driver Hispanic	0.0230	0.0200	0.0117	0.00712
	(0.00682)	(0.00679)	(0.00576)	(0.00323)
Officer Lenient	0.282	0.263	0.197	0.123
	(0.0186)	(0.0336)	(0.0311)	(0.0236)
Driver Black	-0.0389	-0.0342	-0.0447	-0.0371
× Officer Lenient	(0.00501)	(0.00468)	(0.00741)	(0.00566)
Driver Hispanic	-0.0864	-0.0776	-0.0791	-0.0558
× Officer Lenient	(0.00929)	(0.00917)	(0.00864)	(0.00656)
Mean	0.313	0.313	0.325	0.316
Covariates		X	X	X
Location + Time FE	X	X	X	
GPS FE				X
Observations	2078315	2078315	262369	225355

Notes: Table reports linear probability estimates where the outcome variable is whether an individual is ticketed for 9 MPH over the limit, as in Equation 3. Standard errors are clustered at the county level. "GPS FE" includes road segment by year by month by day of the week by shift fixed effects.

Table A.13: Difference-in-Differences at Higher Margins

	(1) 14 MPH Discount	(2) 29 MPH Discount
Driver White	0.000836	0.0000409
	(0.00120)	(0.000488)
Officer Lenient at 14	0.0934	
	(0.0211)	
Driver White × Officer Lenient at 14	0.0410	
	(0.00649)	
$\hat{e}^9\_i$, Discount to 9 Residual	0.000718	-0.0201
	(0.0148)	(0.00731)
Officer Lenient at 29		0.0785
		(0.00628)
Driver White × Officer Lenient at 29		-0.0107
		(0.00122)
$\hat{e}^{14}\_i$, Discount to 14 Residual		-0.0774
		(0.0188)
Mean	0.066	0.031
Sample	$\geq$ 10 MPH	$\geq$ 15 MPH
Observations	1378600	1196503

Notes: County-level clustered standard errors in parentheses. Dependent variables are indicators for whether the Officer is Black or Hispanic. Troops consist of 6-10 counties, so "County FE" subsumes "Troop FE."

Table A.14: Model Parameter Estimates

	White			Minority			
	(1) μ	(2) σ^2	(3) # Param	(4) μ	(5) σ^2	(6) # Param	(7) Mean Diff
b, slope	0.0425	—	1	—	—	—	—
	(0.0004)						
t, officer valuations	-0.1823	4.2201	1851	-0.2007	3.7423	1851	0.0185
	(0.0026)	(0.1388)		(0.0029)	(0.1230)		(0.0039)
λ, speeds	20.3561	2.6103	67	20.7879	1.8637	67	-0.4318
	(0.0274)	(0.4544)		(0.0269)	(0.3244)		(0.0384)
Pr($Discount^9$ | S = 15, E(Z))	0.3390	0.1194	1851	0.3193	0.1120	1851	0.0197
	(0.0006)	(0.0000)		(0.0006)	0.0000		(0.0009)

	Speed Parameters γ		Preference Parameters α	
	(1)	(2)	(3)	(4)
Female	-0.4491	(0.0064)	0.1341	(0.0026)
Age	-0.0434	(0.0002)	0.0054	(0.0001)
Previous Tickets	0.1157	(0.0015)	-0.0226	(0.0007)
County Minority Share			-1.8760	(0.0195)

Notes: This table presents estimates of the model introduced in section 6. b is the slope parameter for how officers weight the speed of drivers in choosing to discount, t is each officer's mean valuation of a racial group in choosing to discount, and λ is the poisson speed parameter for each race by county. $Pr(Discount^9|S = 15, E(Z)) = \Phi(t_{rj} + E(Z)\alpha - 15b)$, i.e. the probability of being discounted when driving 15 MPH over for an average driver. The variances are empirical variances of the estimates, not adjusted for sampling error.

Table A.15: Speed Gap Decomposition

	State-Wide Disparity			
	(1) White Mean (MPH)	(2) Minority Mean	(3) Difference	(4) Percent
Baseline	15.4 (0.006)	17.1 (0.008)	1.7 (0.011)	100
No Discrimination	15.4 (0.006)	16.9 (0.008)	1.5 (0.009)	88 (0.011)
No Sorting	15.5 (0.007)	17.0 (0.009)	1.5 (0.011)	86 (0.011)
Neither	15.5 (0.006)	16.7 (0.009)	1.3 (0.011)	15 0.010

	County-Level Disparity			
	(1) White Mean (MPH)	(2) Minority Mean	(3) Difference	(4) Percent
Baseline	15.4 (0.006)	16.0 (0.009)	0.7 (0.011)	100
No Discrimination	15.4 (0.006)	15.8 (0.008)	0.4 (0.009)	67 (0.016)
No Sorting	15.5 (0.007)	16.2 (0.009)	0.7 (0.011)	104 (0.021)
Neither	15.5 (0.006)	15.9 (0.009)	0.5 (0.011)	70 (0.018)

Notes: Table presents how the racial gap in speeds changes without bias and sorting of officers across counties. The gap is the minority drivers' outcome minus white drivers' outcome. No bias is calculated by assigning each officer's preferences toward minorities to be the same as his preference to whites. No sorting is calculated by simulating a new draw of officers for each driver, where the draw is done at the state level.

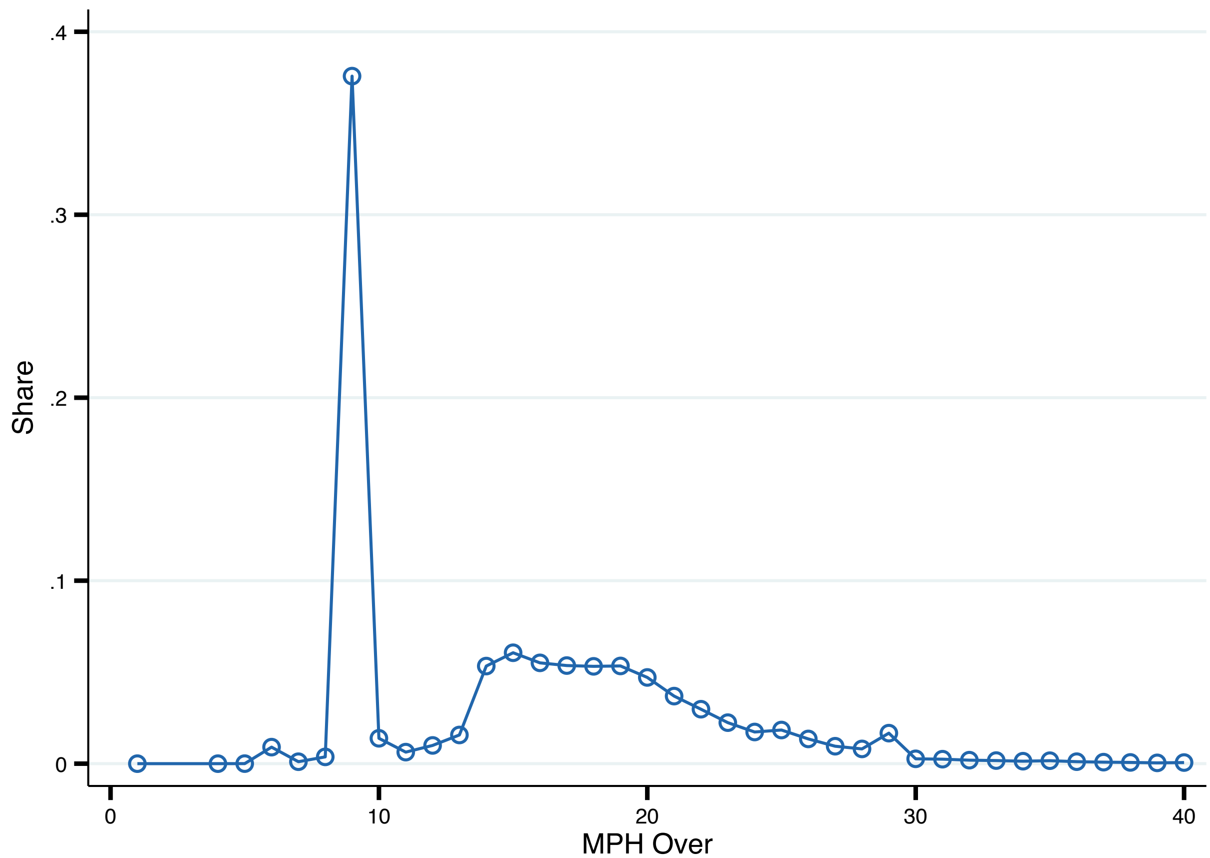
Table A.16: Alternative Model Parameter Estimates

	(1) μ	(2) σ^2	(3) # Param	(4) μ	(5) σ^2	(6) # Param	(7) Mean Diff
b, Slope	0.0262	0.0030	65	—	—	—	—
Heap Probability	0.0884	0.0028	65	—	—	—	—
c_{14}, 14 MPH Discount Cutoff	-0.2231	0.0711	65	—	—	—	—
Share Officers Lenient	0.6914	0.0696	65	—	—	—	—
	White Drivers			Minority Drivers			
t, Officer Valuations	0.1048	2.5560	65	-0.0219	2.4687	65	0.1268
λ, Speeds	20.4758	3.1672	65	21.0488	2.8431	65	-0.5731
Pr($Discount^9$ | S = 15, E(Z))	0.3930	0.0685	65	0.3642	0.0638	65	0.0288
Pr($Discount^{14}$ | S = 15, E(Z))	0.0338	0.0010	65	0.0344	0.0011	65	-0.0006

	Speed Parameters γ			Preference Parameters α			
	(1) μ	(2) σ^2	# Param	(3) μ	(4) σ^2	# Param	
Female	-0.3620	0.1136	65	0.1015	0.0064	65	
Age	-0.0439	0.0002	65	0.0039	0.0000	65	
Previous Tickets	0.1072	0.0032	65	-0.0286	0.0005	65	

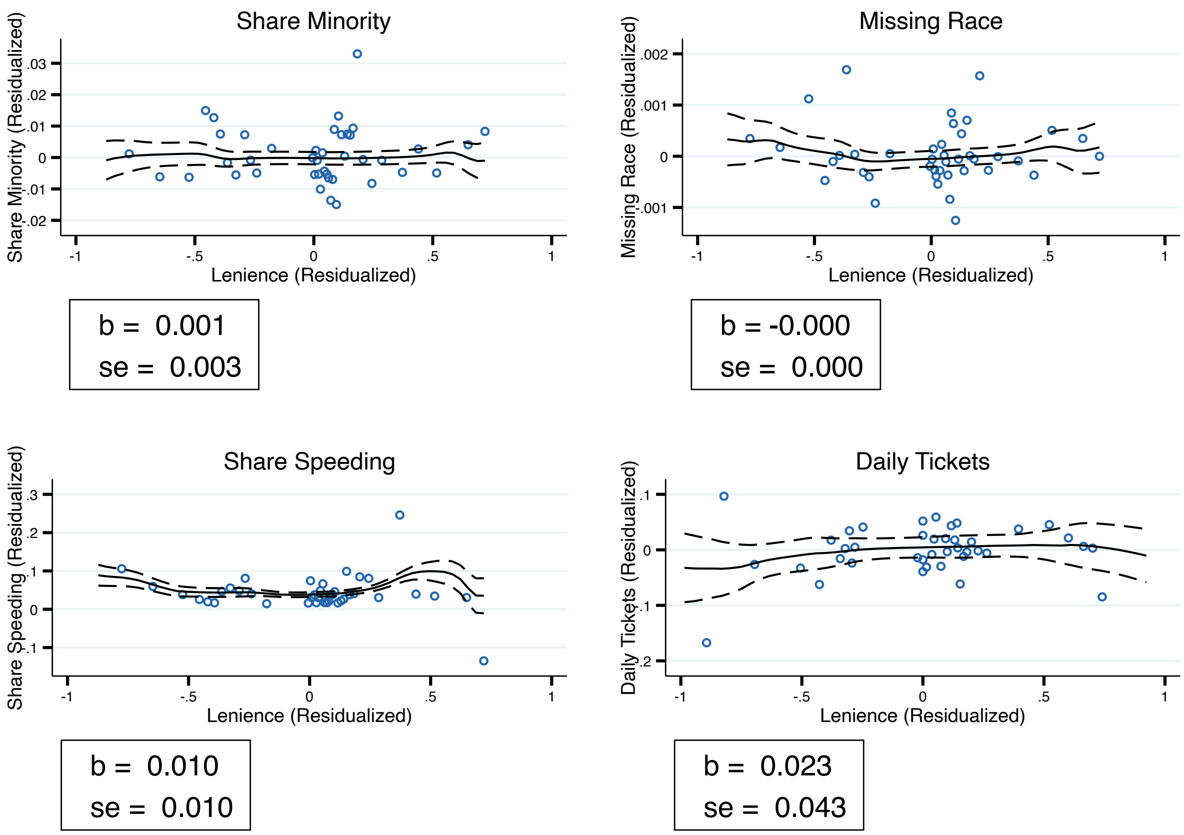
Notes: Table presents parameter estimates from the alternative model specification presented in Appendix Section I.

Figure A.1: Distribution of Charged Speeds for Radar Gun Sample



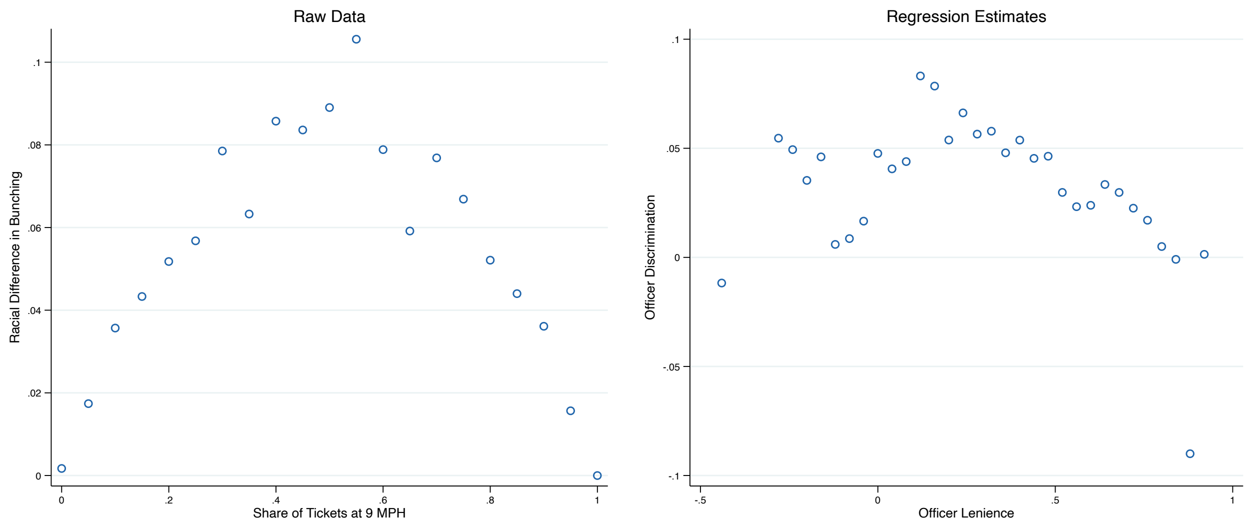
Notes: Line shows histogram of ticketed speeds for observations where the officer records that the speed is detected from a radar or laser gun (N = 126,121).

Figure A.2: Officer Lenience and Stop Characteristics



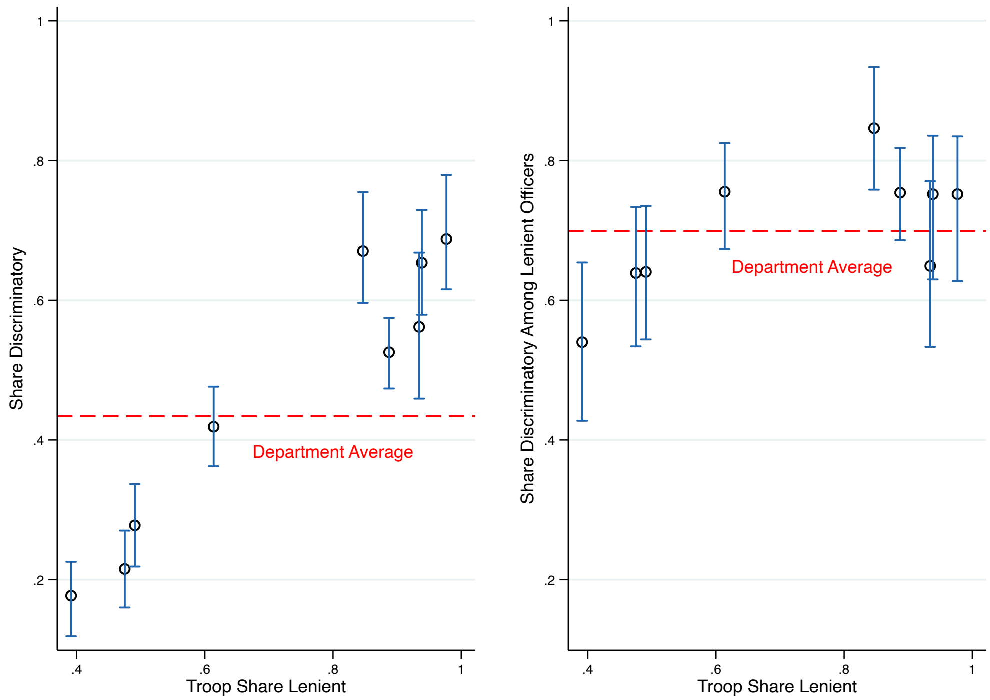
Notes: Figure plots the relationship between officer lenience and various characteristics of the officers' stops, where both officer's lenience and the stop characteristic have been residualized to remove location-time fixed effects. By officer lenience here we mean the indicator for whether an officer has more than 2% of tickets charge at 9mph over. The top left panel plots officer lenience against his share of tickets given to minority drivers, the top right the share of tickets with race missing, and the bottom left the share of tickets that are for speeding. For the bottom right figure, we calculate the number of daily tickets for each officer-by-year, and similarly calculate whether an officer is lenient in each year. We residualize both with county-by-year fixed effects. The scatterplots group the data into 40 bins by residualized lenience, where each bin contains 2.5% of observations.

Figure A.3: Officer-Level Discrimination Against Lenience



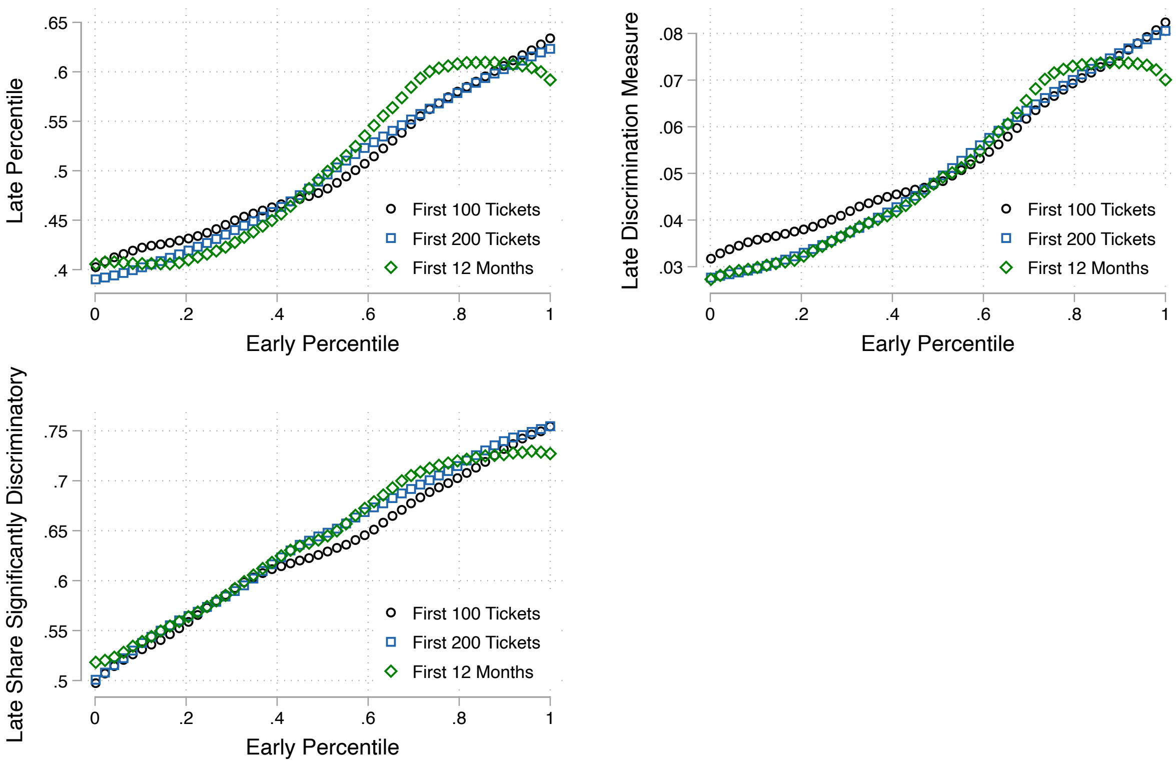
Notes: The left panel plots the officer-level racial differences in share of tickets issued at 9 MPH over (white minus minority) against the share of tickets issued at 9 MPH over for all individuals, which are binned to multiples of 0.05. The right panel plots the officer-level estimates of discrimination against officer-level estimates of lenience from Equation 4, which are separated into 25 bins of equal size.

Figure A.4: Share Discriminatory by Troop-Level Lenience



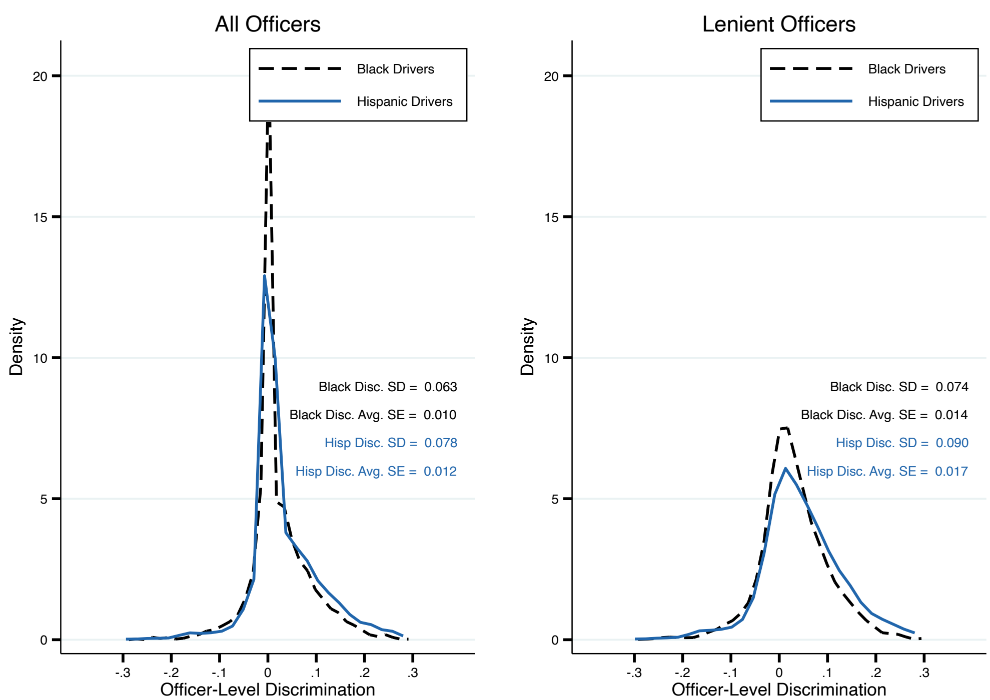
Notes: This figure plots troop-level estimates of the share of officers who are discriminatory against the share of each troop's officers who are lenient. The left panel calculates the share discriminatory among all officers, while the right panel calculates the share discriminatory among only lenient officers.

Figure A.5: Early and Late Discrimination



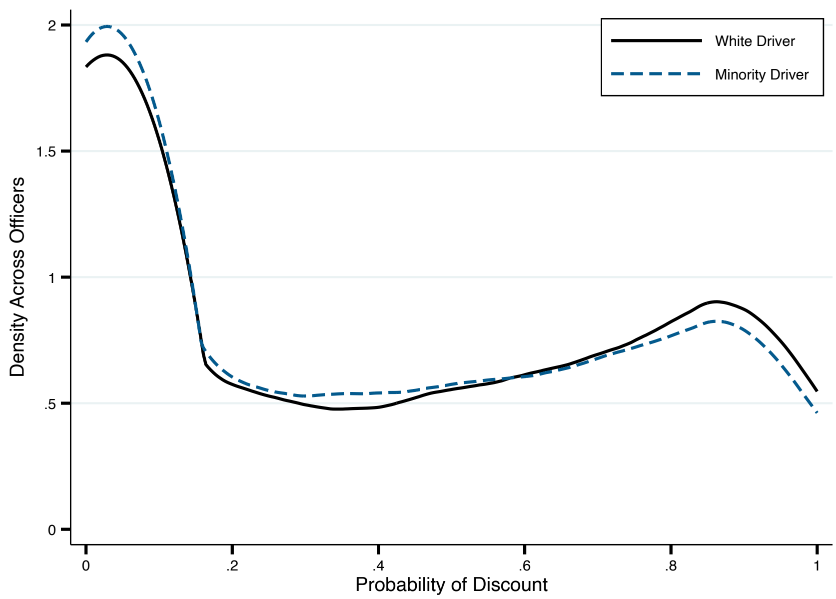
Notes: Figure corresponds to analysis in Appendix Section D.2. The top left panel regresses an officer's late sample discrimination percentile on early sample discrimination percentile. The top right panel regresses an officer's late sample measure of discrimination on their early sample discrimination percentile, and the bottom left panel regresses an officer's late sample indicator for a significant discrimination coefficient on their early sample discrimination percentile. The legends indicate the rule used to designate an officer's early sample.

Figure A.6: Distribution of Officer-Level Discrimination, Separately by Race



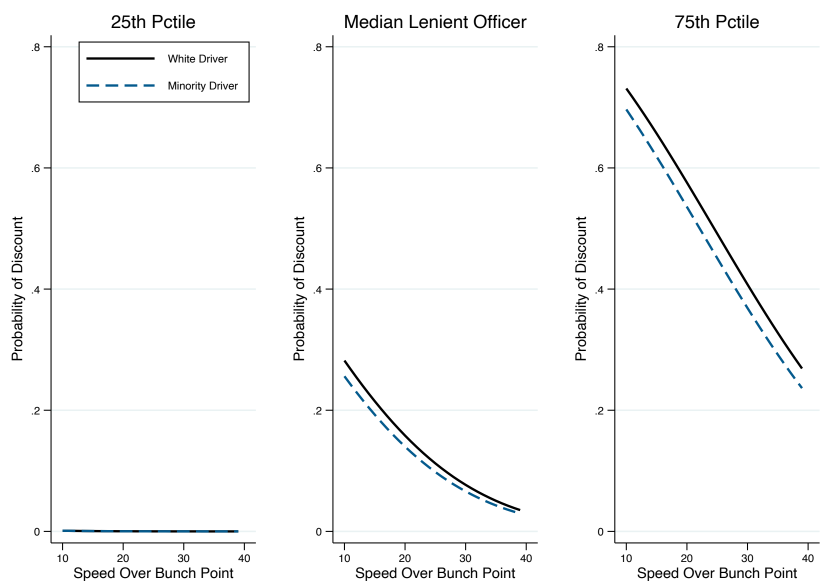
Notes: Figure plots each officer's $-\beta_4^j$ and $-\beta_5^j$ from the regression 2. Officers who are non-lenient are assigned $\beta_3^j = 0$ and are excluded from the right panel. SD reports the standard deviation across $-\beta_4^j$ and $-\beta_5^j$, and Avg SE. reports the average standard error for each individual $-\beta_4^j$ and $-\beta_5^j$.

Figure A.7: Model Estimates: Officer Lenience by Race



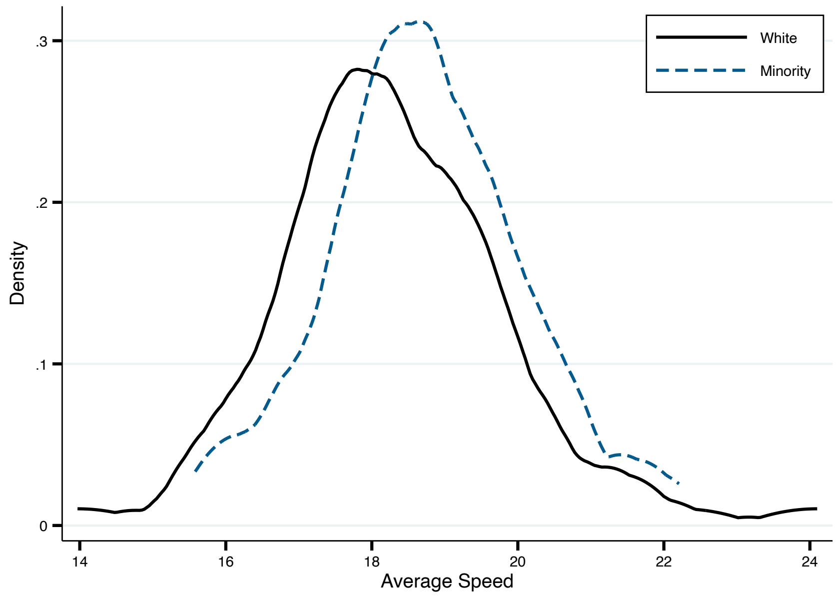
Notes: $P_{rj} \equiv P_j(\text{Discount}|X = 10, \text{ Driver Race} = r, Z = E(Z))$

Figure A.8: Model Estimates: Percentiles of Officer Lenience



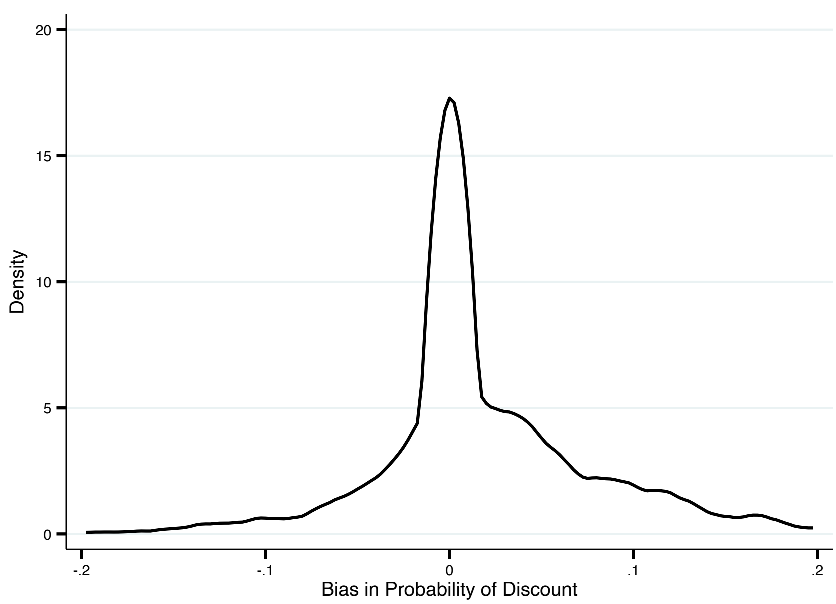
Notes: $P_{rj} \equiv P_j(\text{Discount}|X = 10, \text{ Driver Race} = r, Z = E(Z))$

Figure A.9: Model Estimates: Speed Distribution



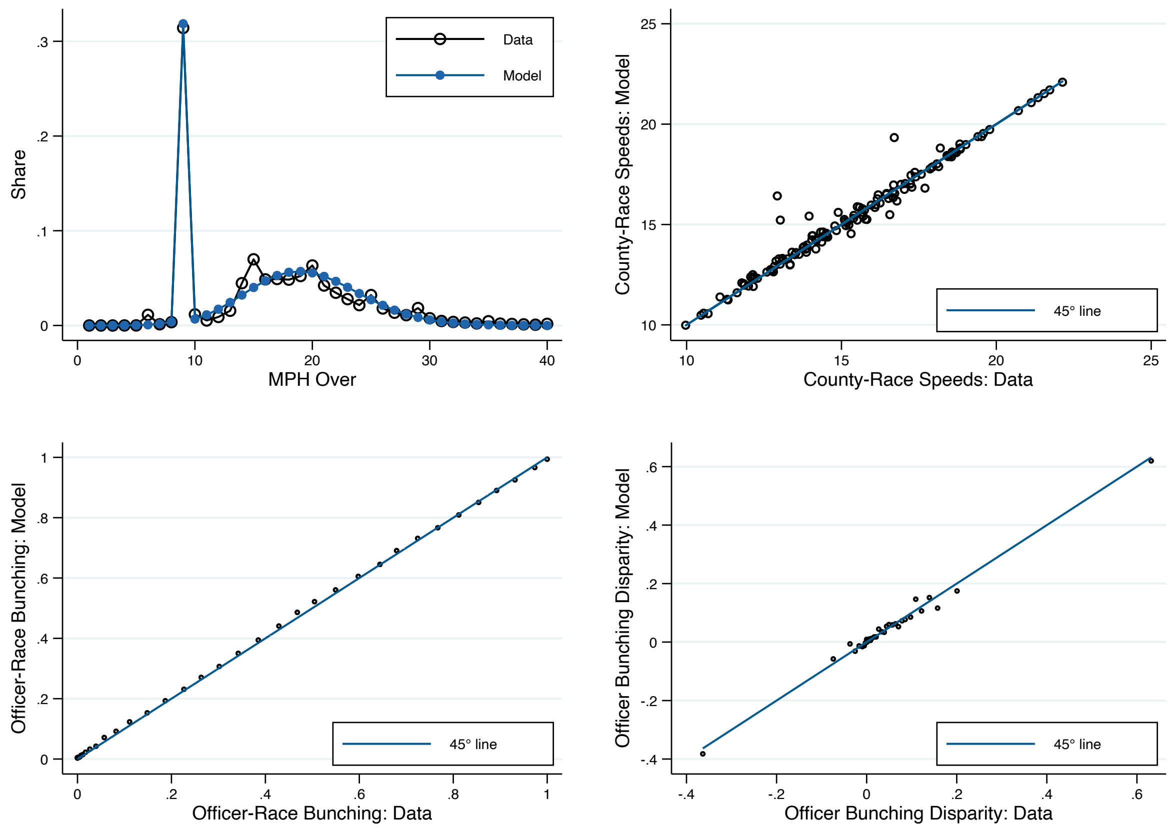
Notes: Figure plots the distribution of speed parameters λ across counties, separately by race of the driver, where individual covariates are set to the average value. In other words, we plot $\lambda = \lambda_{cr} + \gamma E(Z)$

Figure A.10: Model Estimates: Racial Discrimination by Officer



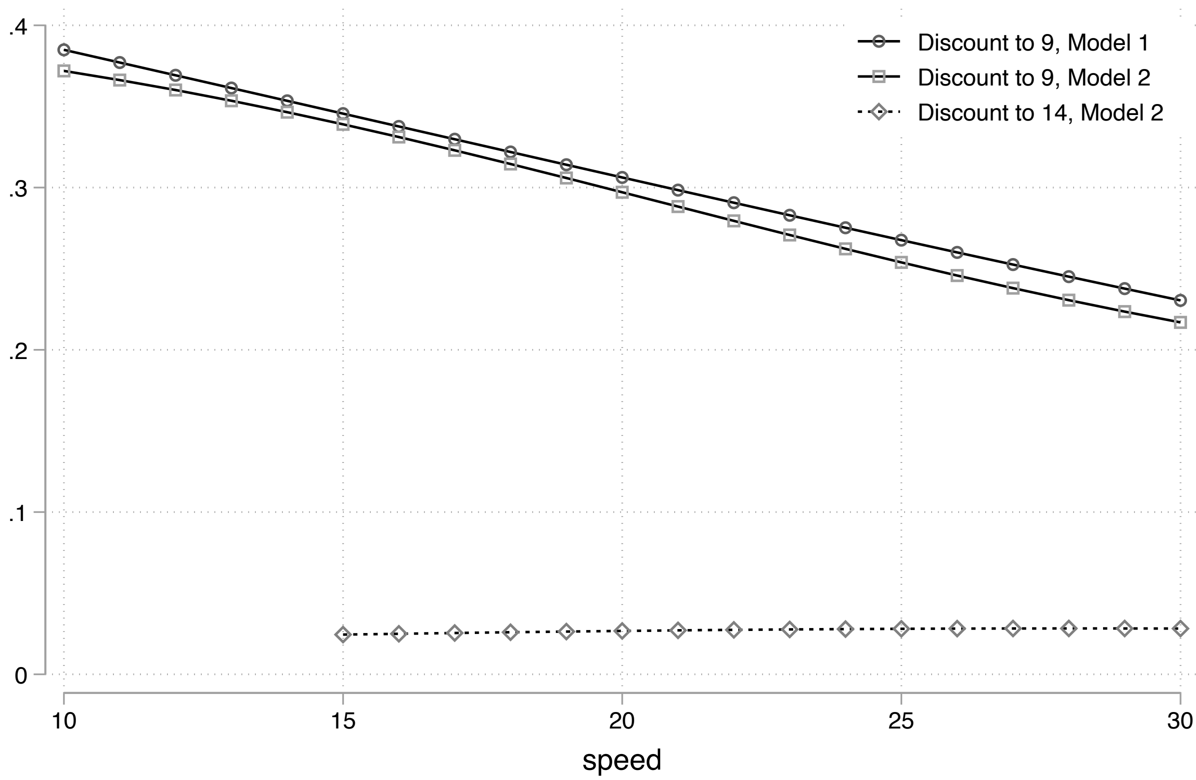
Notes: $P_j(\text{Discount}|X = 10, \text{Driver Race} = \text{White}, Z = E(Z)) - P_j(\text{Discount}|X = 10, \text{Driver Race} = \text{Minority}, Z = E(Z))$

Figure A.11: Model Diagnostic Figures



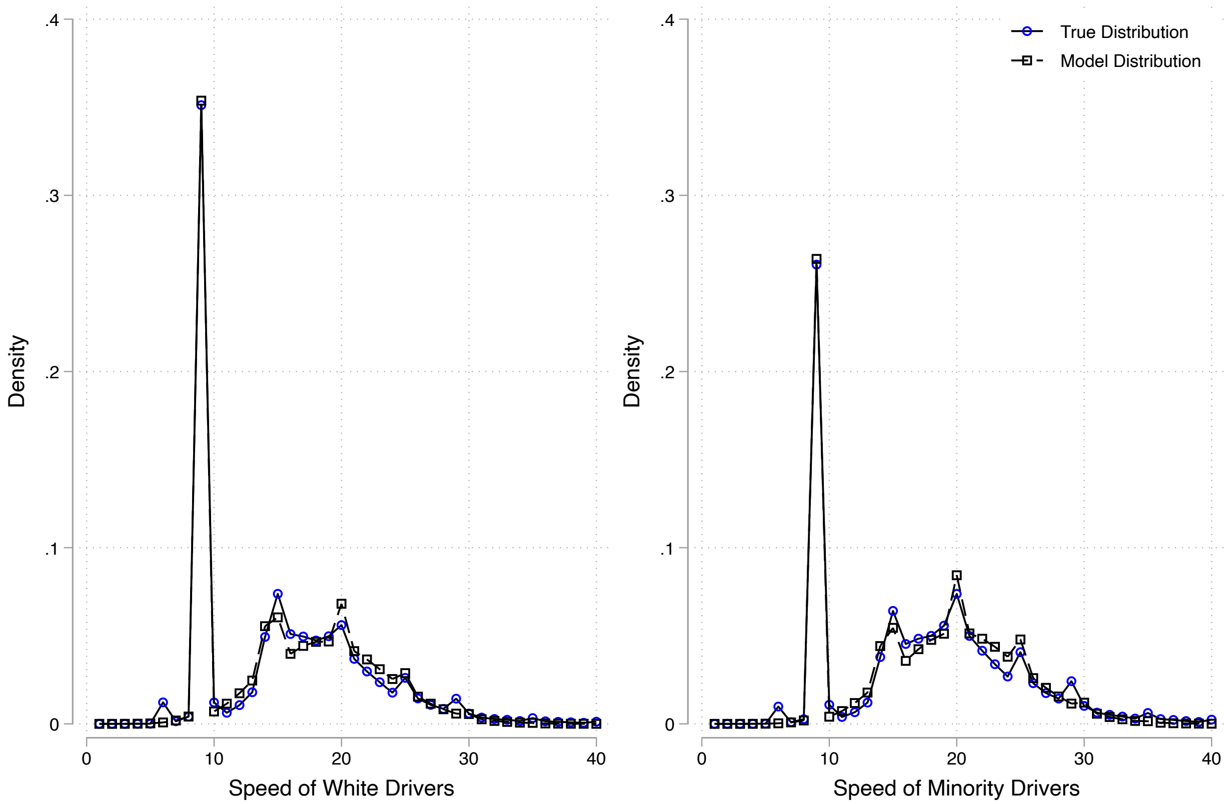
Notes: Figures compare various model estimates with their counterparts in the true data. Model estimates are found by simulating 100 iterations of the model and calculating averages across iterations. The top left panel compares the aggregate histograms of speeds. The top right panel compares the average ticketed speeds by race-county. The bottom left panel compares the share of tickets at 9 MPH over by officer-race. The bottom right panel compares the racial disparity in bunching at 9 MPH over by officer. The scatterplots of the bottom two panels group the data into 40 bins, where each bin contains 2.5% of observations.

Figure A.12: Comparison of Models



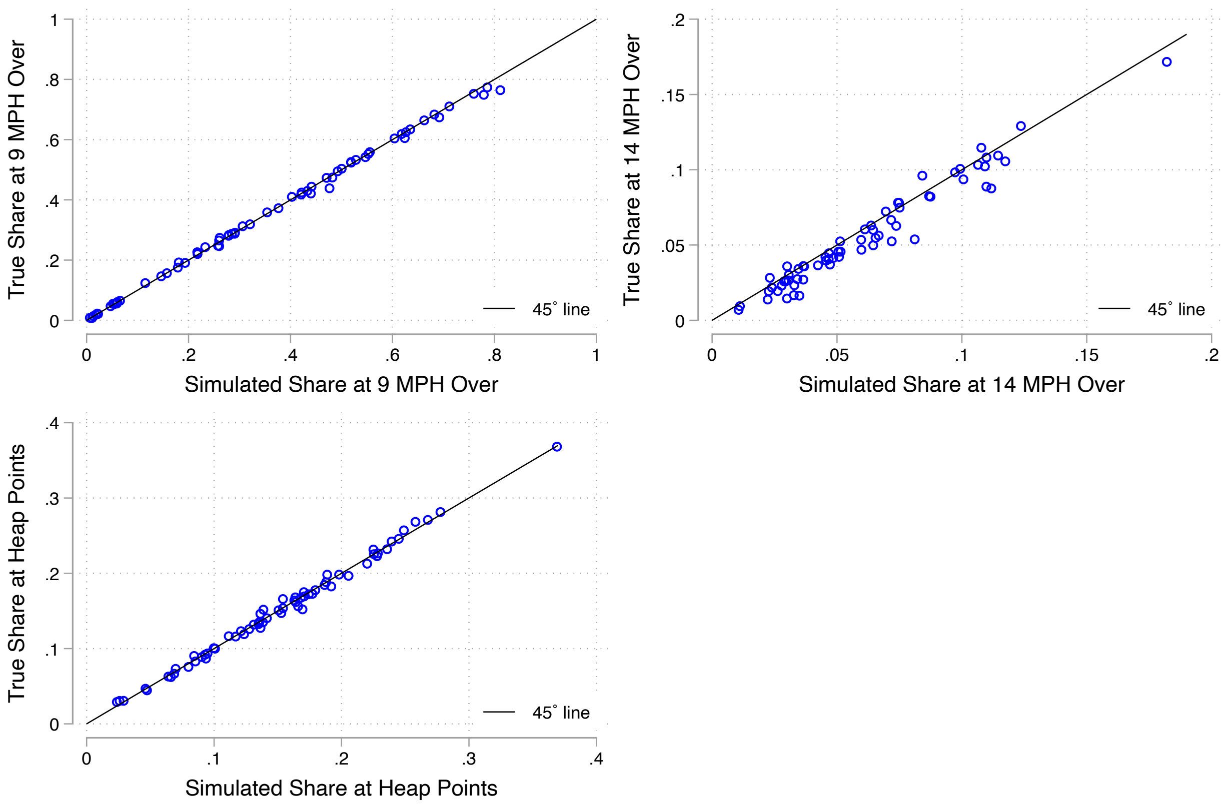
Notes: Figure plots estimates of the probability of discounting to 9 MPH over and 14 MPH over from every potential stopped speed, derived from the models in Section 6 and Appendix Section I. The estimates are calculated for the average driver across the full sample.

Figure A.13: Fit of Alternative Model



Notes: Figure plots the fit of the model in Appendix Section I to the aggregate speed distribution. The left panel plots the true and simulated speed distributions for white drivers, and the right panels plots the same distributions for minority drivers.

Figure A.14: Fit of Alternative Model Across Counties



Notes: Figure plots the fit of the model in Appendix Section I to the county-level moments of the speed distribution. The top left panel plots the true and simulated values for the share of tickets at 9 MPH over across counties. The top right panel plots the same values for the share of tickets at 14 MPH over, and the bottom right panel plots the same values for the share of tickets at multiples of 5 at and above 15 MPH over.

References

Aizer, Anna, and Joseph J Doyle Jr. 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *The Quarterly Journal of Economics*, 130(2): 759–803.

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics*, 133(4): 1885–1932.

Becker, Gary Stanley. 1957. *The Economics of Discrimination: an Economic View of Racial Discrimination.* University of Chicago.

Björklund, Anders, and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *The Review of Economics and Statistics*, 42–49.

Chalfin, Aaron, Oren Danieli, Andrew Hillis, Zubin Jelveh, Michael Luca, Jens Ludwig, and Sendhil Mullainathan. 2016. "Productivity and Selection of Human Capital with Machine Learning." *The American Economic Review*, 106(5): 124–127.

Dobbie, Will, and Jae Song. 2015. "Debt Relief and Debtor Outcomes: Measuring the Effects of Consumer Bankruptcy Protection." *The American Economic Review*, 105(3): 1272–1311.

Doyle, Joseph. 2007. "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *The American Economic Review*, 97(5): 1583–1610.

Doyle Jr, Joseph J. 2008. "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care." *Journal of Political Economy*, 116(4): 746–770.

FCCC. 2016. "Florida Court Clerks and Comptrollers Traffic Citation Accounting Transmission System [data file]." Accessed: 2016-2-8.

FDOC. 2010. "Florida Department of Transportation 2010 Crash Reports [data file]." Accessed: 2016-8-27.

FDOC. 2019. "Florida Department of Corrections OBIS Database [data file]." *http://www.dc.state.fl.us/pub/obis_request.html*, Accessed: 2019-8-29.

FHP. 2015*a*. "Florida Highway Patrol Civilian Complaints Public Records [data file]." Accessed: 2017-5-11.

FHP. 2015*b*. "Florida Highway Patrol Use of Force Public Records [data file]." Accessed: 2017-5-10.

Frandsen, Brigham R. 2017. "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete." In *Regression Discontinuity Designs: Theory and Applications.* 281–315. Emerald Publishing Limited.

Goncalves, Felipe, and Steven Mello. 2017. "Does the Punishment Fit the Crime? Speeding Fines and Recidivism." *Unpublished Manuscript*, 1–48.

Goncalves, Felipe, and Steven Mello. 2021*a*. "Replication Data for: A Few Bad Apples? Racial Bias in Policing." *American Economic Association* [publisher], Inter-university Consortium for Political and Social Research [distributor]. *https://doi.org/10.3886/E120242V1*.

Goncalves, Felipe, and Steven Mello. 2021*b*. "Supplementary Data for: A Few Bad Apples? Racial Bias in Policing." *American Economic Association* [publisher], Inter-university Consortium for Political and Social Research [distributor]. *https://doi.org/10.3886/E123921V1*.

Heckman, James J. 1979. "Sample Selection as a Specification Error." *Econometrica*, 47(1): 153–161.

Heckman, James J, Daniel Schmierer, and Sergio Urzua. 2010. "Testing the Correlated Random Coefficient Model." *Journal of Econometrics*, 158(2): 177–203.

Kling, Jeffrey R. 2006. "Incarceration Length, Employment, and Earnings." *The American economic review*, 96(3): 863–876.

Lange, James E, Mark B Johnson, and Robert B Voas. 2005. "Testing the Racial Profiling Hypothesis For Seemingly Disparate Traffic Stops on the New Jersey Turnpike." *Justice Quarterly*, 22(2): 193–223.

Marx, Philip. 2018. "An Absolute Test of Racial Prejudice." *Unpublished Manuscript*, 1–22.

Mueller-Smith, Michael. 2014. "The Criminal and Labor Market Impacts of Incarceration." *Unpublished Working Paper*.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "IPUMS USA: Version 10.0 [dataset]." *https://doi.org/10.18128/D010.V10.0*.

Smith, William, Donald Tomaskovic-Devey, Matthew Zingraff, H Marcinda Mason, Patricia Warren, and Cynthia Wright. 2004. "The North Carolina Highway Traffic Study." *NCJRS Grant Report*, 1–407.

Walker, Samuel, Geoffrey P Alpert, and Dennis J Kenney. 2000. "Early Warning Systems for Police: Concept, History, and Issues." *Police Quarterly*, 3(2): 132–152.