A Few Bad Apples? Racial Bias in Policing*

Felipe Goncalves[†] Steven Mello[‡]

June 15, 2020

Abstract

We estimate the degree to which individual police officers practice racial discrimination. Using a bunching estimation design and data from the Florida Highway Patrol, we show that minorities are less likely to receive a discount on their speeding tickets than white drivers. Disaggregating this difference to the individual police officer, we find that 40% of officers explain all of the aggregate discrimination. We then apply our officer-level discrimination measures to various policy-relevant questions in the literature. In particular, reassigning officers across locations based on their lenience can effectively reduce the aggregate disparity in treatment.

JEL Classification: J15, K42
Keywords: Discrimination, Racial Bias, Police, Traffic Enforcement.

*We are grateful to Will Dobbie, Ilyana Kuziemko, and Alex Mas for guidance and support throughout this project. We benefited from helpful comments by Leah Boustan, Jessica Brown, Nicholas Buchholz, Janet Currie, Rebecca Diamond, Jennifer Doleac, Nik Engbom, Hank Farber, Sara Heller, Nathaniel Hendren, Daniel Herbst, Bo Honoré, Sierra Kuzava, Michael LaForest, Andrew Langan, Michael Mueller-Smith, Christopher Neilson, Emily Owens, Aurelie Ouss, Jakob Schlockermann, Petra Todd, and participants in various seminars. We thank Beth Allman, Jeffrey Bissainthe, Kiara Guzzo, Wilton Johnson, Timothy Kutta, Stacy Lehmann, and Brenda Paige for assistance with data from various agencies. The Princeton University Industrial Relations Section provided generous financial support. Any errors are our own.

[†]UCLA; fgoncalves@econ.ucla.edu

[‡]Dartmouth College and NBER; steven.m.mello@dartmouth.edu

1 Introduction

The disparate treatment of whites and minorities in the criminal justice system is a central policy concern in the United States. Blacks and Hispanics are more likely to be stopped by the police (Coviello and Persico, 2013), convicted of a crime (Anwar *et al.*, 2012), denied bail (Arnold *et al.*, 2018), and issued a lengthy prison sentence (Rehavi and Starr, 2014) relative to observably similar whites. In light of these disparities, a literature has developed to test whether these outcomes can be explained by discrimination on the part of police officers, judges, and other criminal justice agents (Knowles *et al.*, 2001; Anwar and Fang, 2006; Grogger and Ridgeway, 2006; Antonovics and Knight, 2009; Persico, 2009; Abrams *et al.*, 2012; Horrace and Rohlin, 2016; Fryer, 2018; Arnold *et al.*, 2018). The view that discrimination is responsible for these disparate outcomes has gained traction in recent years, particularly within minority communities, following several highly publicized police killings of minorities. A 2013 Gallup poll found that half of black adults agreed that racial differences in incarceration rates are "mostly due to discrimination," while only 19% of white respondents agreed.[1]

While current methods focus on detecting the presence of racial discrimination *on average*, an unresolved challenge is how to identify discrimination at the level of the individual criminal justice agent. Existing approaches largely do not differentiate between discrimination that is widespread versus that which is concentrated among a few agents. However, the optimal policy for mitigating the presence of discrimination depends crucially on how it varies across individuals. Without knowing which agents are discriminatory, it is not possible for institutions to target individuals for discipline or training. More generally, the optimal remedy will depend on the concentration of discrimination across agents. If misbehavior is widespread, a targeted policy of disciplining specific individuals will be ineffectual, and the appropriate response may require a department-wide solution.[2]

[1]See www.gallup.com/poll/175088/gallup-review-black-white-attitudes-toward-police.aspx.

[2]The question of whether misbehavior is systemic or the product of a few bad individuals has also garnered policy interest with regard to federal oversight of local police departments. In January 2017, Attorney General nominee Jeff Sessions stated, "I think there's concern that good police officers and good departments can be sued by the Department of Justice when you just have individuals within a department who have done wrong. These lawsuits undermine the respect for police officers and create an impression that the entire department is not doing their work consistent with fidelity to law and fairness."

In this paper, we study traffic policing by the Florida Highway Patrol and examine whether officers discriminate when enforcing punishments for speeding. We exploit a common institutional feature in traffic policing and use a bunching estimation design to identify discrimination. In many states, the punishment for speeding increases discontinuously with the speed of the driver, exhibiting "jumps" in harshness. A jump may involve not only a higher fine, but also a mandated court appearance or permanent mark on the driver's record. Officers are free to choose what speed to charge, and it is thus a common practice for officers to reduce the written speed on a driver's ticket to right below a jump in the fine schedule.[3] Our objective is to identify discrimination in discounting at the level of the individual officer, where we define discrimination as the differential treatment of drivers on the basis of their race when stopped for the same speed.

Several features of our setting are ideal for studying discrimination. When testing for discrimination in many criminal justice outcomes, a central concern is accounting for unobserved differences in criminality across individuals. In the context of speeding tickets, guilt is summarized by the driving speed, which is both one-dimensional and typically observed by the ticketing officer. Further, in many criminal justice contexts, the lenience of an agent is calculated relative to his peers' behavior. In our setting, officers make an explicit decision to reduce a driver's speed, allowing us to see each officer's absolute degree of lenience and observe officers who practice no lenience. Perhaps most importantly, we observe agents making many decisions in very similar contexts, which allows us to construct an accurate measure of discrimination for each officer by comparing his treatment of white and non-white drivers.

As shown in Figure 1, the distribution of speeds ticketed by the Florida Highway Patrol between 2005 and 2015 shows substantial excess mass at speeds just below the first fine increase, where speeds are reported relative to the speed limit. Meanwhile, a remarkably small portion of tickets are issued for speeds just above. We take this bunching as evidence that officers systematically manipulate the charged speed, commonly charging speeds just below fine increases after observing a higher speed, perhaps to avoid an onerous punishment for the driver. However, when disaggregated by driver race in Figure 2, we see that minorities are significantly less likely to be found at the bunch point.

The first task of this paper is to confirm that this disparity is evidence of officer discrim-

[3] This practice is similar to teachers' bunching up of grades on high-stakes exams (Dee *et al.*, 2016; Diamond and Persson, 2016).

ination. Our central challenge is in ruling out that racial differences in treatment are due to differences in criminality. Minorities may be driving faster than whites when stopped, leading officers to treat them less leniently. While our data record the speed that is charged on a ticket, we do not observe the true *stopped* speed of the drivers in our data. To deal with this challenge, we use the fact that one-third of officers practice no lenience. Namely, they exhibit no bunching in their distribution of ticketed speeds.[4] For these officers, we argue that their distribution of ticketed speeds reflects the true distribution of driven speeds among stopped and ticketed drivers. We show that, conditional on location and time, driver characteristics are not predictive of whether the officer he encounters is lenient. Non-lenient officers do not write fewer tickets than lenient officers, and a similar share of their tickets are for speeding offenses. These facts suggest that lenient and non-lenient officers are pulling over similar types of drivers, and thus non-lenient officers can be used to identify the "true" distribution of speeds. The speed distributions for non-lenient officers indicate that, while minorities are driving slightly faster speeds when stopped, the racial gap in speeding is minimal. We therefore argue that our estimates of differences in discounting can be interpreted as evidence of discrimination. Further, we provide a simple approach to adjusting our estimates to account for any racial differences in speeding.

Using a difference-in-differences framework, we then find that white drivers differentially benefit from being stopped by a lenient officer. White drivers stopped by lenient officers are six percentage points more likely to be discounted than minority drivers off a base of 45%. This gain stems from the fact that minorities are treated less leniently when stopped for speeds ranging from 10 to 23 MPH over the limit.

The central contribution of our paper is to further provide an estimate of the discrimination of *each individual officer*. Specifically, we compute an officer's lenience toward minorities relative to his own treatment of white drivers, differencing out the treatment of each race by non-lenient officers and adjusting for other features of the stop, and treat that difference as the officer's discrimination. Disaggregating to the officer level reveals significant heterogeneity in the degree of discrimination. An officer at the 90th percentile of discrimination is nearly twice as likely to discount a white driver as a minority driver. The modal officer practices no discrimination, and forty percent of officers explain the entirety of the aggregate

[4]The existence of non-lenient officers also leads us to conclude that the bunching of ticketed speeds is not due to drivers strategically driving below the jump in fine.

disparity in treatment. Correlating officer-level discrimination to demographics, we find that minority and female officers tend to practice less discrimination than other officers.

The remainder of the paper exploits our officer-level measures of lenience and discrimination to understand the mechanisms that lead to the disparity in treatment. To what extent are minorities being discounted less often because they are driving faster? Conversely, how much of the gap in discounting is caused by discrimination? And what policies can be used to reduce any disparity that is due to discrimination?

To answer these questions, we estimate a simple model that identifies both differences in driving speeds, by each race and county, and preferences for discounting, by each officer and race of driver. Model estimates indicate that, within location, forcing all officers to treat minority drivers the same as they treat white drivers removes 83% of the gap in discounting. Only 17% of the gap is due to minorities driving faster. Across locations, a large share of the disparity in treatment is due to the fact that minorities drive in areas where officers are less lenient to all motorists.

Performing the counterfactuals discussed above, we find that policies that target discrimination directly are only mildly effective for reducing the treatment gap. Firing the most discriminatory officers (both for and against minorities) reduces the gap, as does increasing the presence of minority or female officers, but the gains are limited. Perhaps most effective and easily implemented, reassigning officers across counties within their troops so that minorities are exposed to more lenient officers can remove essentially the entire white-minority discounting gap.

While differentiating between taste-based (Becker, 1957) and statistical (Arrow, 1973; Phelps, 1972) discrimination is not our central focus, several pieces of evidence point to a taste-based interpretation. First, our setting is less conducive to statistical discrimination than many other criminal justice interactions because officers directly observe criminality (i.e., the driving speed) rather than infer it ex ante. Second, we find that minority and female officers are less discriminatory on average, suggesting that preferences rather than statistical inference explain the observed discrimination. While we find some evidence that offers statistically discriminate on the basis of whether an individual is likely to contest a harsher ticket in court, this selection cannot explain the racial disparity in discounting. Therefore, while we use the term discrimination throughout the paper, our results are more consistent with a taste-based rather than statistical model of discrimination.

This paper contributes to a growing literature on methods for detecting discrimination in the criminal justice system and beyond, whose approaches include audit studies that vary individual race (Bertrand and Mullainathan, 2004; Edelman *et al.*, 2017; Agan and Starr, 2016), exploiting variation in the observability of race or gender (Goldin and Rouse, 2000; Grogger and Ridgeway, 2006; Donohue, 2014), and the use of rich controls for underlying behavior and context (Fryer, 2018). Another strand of research uses the "hit rate test" (Becker, 1957), where discrimination is identified by comparing the success in treatment across two groups where the treator ostensibly cares about a single objective (Knowles *et al.*, 2001; Arnold *et al.*, 2018; Marx, 2018).

Our approach falls broadly into a literature using *benchmarking* procedures, where the behavior of one agent is compared to a proposed control group, to identify discrimination.[5] In the paper most closely related to ours, Anbarci and Lee (2014) study the discounting behavior of traffic officers using a benchmarking design and find that at least one racial group of officers is biased in favor of their own race. To our knowledge, the only existing study aiming to identify discrimination by individual criminal justice agents is Ridgeway and MacDonald (2009), who compare the racial makeup of NYPD officers' stops and frisks to those of nearby officers and are able to identify a set of officers with a disproportionately high share of minority stops.

Relative to this existing literature, a strength of our approach is that non-lenient officers are, by construction, non-discriminatory. This fact allows us to avoid the possibility that the comparison group is itself discriminatory, a common problem faced in benchmarking designs.

Methodologically, our approach builds on recent research using "bunching" estimators to recover behavioral parameters (Kleven, 2016). In contrast with most applications (Chetty *et al.*, 2011; Saez, 2010), which infer a hypothetical true distribution by examining areas away from the manipulation region, our approach is similar to Best *et al.* (2015) in that we use panel data and differences across individuals in propensity to bunch to identify the true underlying distribution.

The rest of the paper is organized as follows. Section 2 provides institutional background on the Florida Highway Patrol and describes the data. Section 3 presents a conceptual framework, and Section 4 describes our empirical strategy. Section 5 presents the central findings, and Section 6 considers specification checks and alternative interpretations of our

[5] See Ridgeway and MacDonald (2010) for a review of the benchmarking literature.

results. In Section 7, we present and estimate a model of officer behavior and perform counterfactuals, and Section 8 concludes.

2 Institutional Background and Data

2.1 Institutions of the Florida Highway Patrol

State-level patrols are the primary enforcers of traffic laws on interstates and many highways. When on patrol, officers are given an assigned zone, within which they combine roving patrol and parked observation patrol. During the course of a traffic stop for speeding, officers have two primary ways to exercise discretion. They can give a written or verbal warning, which leads to no fine or points on the driver's license, or they can reduce the speed charged on the ticket. Florida Highway Patrol (FHP) officers are told explicitly in their training manuals that no enforcement actions during a traffic stop can be based on any demographic characteristics, including race and gender.

In Florida, driving 10 MPH over the limit leads to about a $75 higher fine than 9 MPH over. While drivers receive points on their license for speeding, tickets received for 9 and 10 MPH over the limit carry the same number of license points. While it is also common to find a jump in fine between 19 and 20 MPH over, the data strongly suggest that officers prefer to reduce the ticket to 9 MPH over.[6]

Officers in the FHP are divided into one of 12 troops, almost all of which patrol six to eight counties each. Officer assignments operate on eight-hour shifts and cover an assignment region that roughly corresponds to a county, though the size of a "beat" can vary based on the population density of the region. In practice, because we do not observe the exact beat policed by an officer, we will use the county of the stop as a proxy for the officer's assignment region.

Officers face no revenue incentive to collect tickets, as all fines paid by drivers are collected by the government of the county in which the fine was issued. There is also, to the best of our knowledge, no quota system for a minimum number of tickets officers must write.[7]

[6]The actual fine schedule depends on the county in Florida, though 10 MPH over always includes at least a $50 increase in the fine. The point schedule is identical statewide and does not contain a jump at 10 MPH over.

[7]We checked for a spike in the number of issued tickets at certain days of the month or days of the week, and found no evidence of an "end of the period" effect.

Officers do, however, potentially have a promotion incentive to write a certain number of tickets, as the number of tickets they write appears on their performance evaluations. We believe these set of institutional factors contribute to an environment in which officers are encouraged to write tickets but also have the freedom to write reduced charges, which is ideal for our research design.

While all speeding beyond 5 MPH over the limit commands a statutory fine, the evidence suggests that drivers are not regularly pulled over for less than 10 MPH over, and the data show very few tickets for 8 MPH over and 10 MPH over. As we will reiterate in Section 4, many officers have almost no tickets issued at 9 MPH over the limit, suggesting that the majority of the bunching of tickets is for higher speeds that have been reduced.

2.2 Data

From the Florida Court Clerks & Comptrollers, we obtained data on all traffic citations issued in Florida for the years 2005-2015. These data include all information provided on the stopped motorist's driver's license – name, address, race, gender, height, and date of birth, as well as driver's license state and number. The make and year of the stopped automobile is provided in 84% of tickets, and we link this information to an online database of vehicle price estimates.[8] While we see the speed charged on the citation, we do not see the original speed of the stop. We also do not see stops and interactions that do not result in a traffic citation.[9]

The citing officer is identified by name, rank, troop number, and badge number. To supplement the citations data, we obtained officer demographic information from the Florida Department of Law Enforcement (FDLE). These data include officer race, sex, age, education level, and the Florida law enforcement employment history of all law enforcement officers employed in the State of Florida. We focus on stops by the Florida Highway Patrol (FHP), both because officers are more consistently identified in the data in FHP stops and because traffic enforcement is the primary responsibility of FHP.

While the citations record the driver race, there appear to be inconsistencies in the

[8] See https://www.kaggle.com/jpayne/852k-used-car-listings/data.

[9] The problem of only seeing interactions that lead to enforcement is common in the discrimination literature. We discuss this point further in Section 6.1 and Appendix Section F. For a recent paper that addresses this issue directly, see West (2018).

recording of Hispanic. For example, Miami-Dade County issues fewer than 1% of their tickets to Hispanic drivers. To address this issue, we match the drivers' names to Census records, which record all names that appear more than 1,000 times and the share of white, black, Hispanic, and other that carry that name. If an individual in our data has a name that is more than 80% Hispanic, we record them as such.

Using the driver's license number, we are able to link individuals across tickets. Doing so, we construct for each ticket a measure of the number of tickets received in the previous three years, including from non-FHP stops. Using the driver's date of birth and full name, we also link each individual to prison spell records from the Florida Department of Corrections and construct an indicator for any past incarceration. At the point of a stop, the officer is able to see a driver's full criminal history, including arrest history that does not lead to prison, so we consider tickets and incarceration as the best available approximation to an individual's criminal history at the point of the stop.

We restrict the sample to speeding citations in which no accident is reported; the cited speed is between zero and 40 above the posted speed; race of the driver is reported as white, black, or Hispanic (or is imputed as such); and the gender, age, and driver's license number are not missing. To link citations and officer information, we first narrowed the list of FDLE personnel to include only officers with an employment spell as a sworn officer with the FHP covering some portion of the 2005-2015 period. We then match the list of candidate officers with the citations data using the officer name. We exclude stops that cannot be matched to an officer. Lastly, we restrict the sample to officers issuing at least 100 citations, with at least 20 given to minorities and 20 to whites.

The final sample includes 1,142,628 citations issued by 1,591 officers, from an initial sample of 2,124,692 speeding citations. The two most binding restrictions are requiring that race be specified (84% of tickets) and requiring that the officer be linkable to the FDLE (77%). In the appendix Section A we include a table that documents the sample reduction from each restriction we make. In all of our analyses, we consider speed relative to the speed limit (or posted speed) rather than absolute speed. We often refer to this quantity as *MPH Over* or simply as "the speed."

Beginning in 2013, about 40% of tickets are geocoded with the latitude and longitude of a stop (135,586 observations). Using ArcGIS, we link the geocoded tickets to 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.[10] Throughout the analysis, we also provide results for the restricted sample of geocoded tickets with corresponding fixed-effects at the road-segment level. The road-segment analysis allows us to consider a more granular comparison of drivers.

2.3 Summary Statistics

Table 1 presents summary statistics for the sample, broken out by driver race. 58% of drivers are white, 18% are black, and about 23% are Hispanic. Drivers are 35% female and about 36 years old on average, with Hispanics less likely to be female and minority drivers typically younger. In-state drivers account for 84% of tickets. The average driver has been cited about 0.34 times in the past year, though minorities have 0.13 more prior tickets. On average, minority drivers are charged with higher speeds than whites: just over 1 MPH higher for blacks and almost 3 MPH higher for Hispanics. Consistent with Figures 1 and 2, drivers of all races have a high probability of being ticketed at 9 MPH over the limit, which is just below the first jump in the fine schedule. However, minority drivers are also less likely to be charged this speed. As we show in Appendix Tables A.1 and A.2, these disparities in speed and ticketing below the jump persist after controlling for all stop characteristics and time and location fixed effects.

A notable feature of the distribution of tickets is the heaping of charged speeds at multiples of five above the bunch point. This heaping occurs because, in many instances, officers do not use a radar gun, and their recording of the speed may be approximate. In cases where a "method of arrest" is recorded, 76% of tickets indicate the officer observe the speed through "visual inspection," and the remaining 24% report the use of a radar or laser gun. We report in Appendix Figure A.1 the distribution of ticketed speeds for this latter subsample, where we find no heaping at multiples of five.[11]

The bottom panel of Table 2.3 presents summary statistics for the set of officers in our sample. We note here the large number of tickets written by the average officer (717) and the fact that the population of officers is slightly more white (63%) than our sample of drivers (58%).

[10] See Appendix A for further details.

[11] In Table 4, we also show that our main result is not changed when restricting attention to this subsample.

3 Conceptual Framework

In the previous section we documented the disparity in ticketing at 9 MPH over between whites and minorities. Here we introduce a simple framework of officer decision-making that can explain the disparity in discounting through two mechanisms – differences in speeding and discrimination – and motivates our empirical strategy in Section 4 and our modeling exercise in Section 7.

Officer j stops motorist i for speeding. His stopped speed s^* is drawn from some discrete distribution $f_r(\cdot)$, which can be a function of the driver's race r. For simplicity, we suppress here the possible dependence of the distribution on other driver characteristics. If the driver's speed is above s_d, the officer has the choice to reduce the charged speed to s_d to reduce the fine the driver will face. Otherwise, the speed is set at s^*. When deciding whether to reduce the ticket, we suppose the officer weighs a mix of personal concerns such as the inconvenience of attending traffic court; policing objectives such as the blameworthiness of the individual and the potential deterrence effect of ticketing the individual; and bias against certain groups r. Balancing these objectives, the officer has some probability $P_{jr}(s^*)$ of discounting the individual, which may be a function of the driver's race r and their stopped speed s^*.

In this framework, it is natural to define discrimination in the following way: We say that officer j is discriminatory at speed s^* if $P_{jw}(s^*) > P_{jm}(s^*)$ for a given speed s^*. While we describe the officers' preferences as potentially reflecting bias, we are not yet taking a stand on whether any disparity in treatment is taste-based versus statistical. For example, it is possible that some officers prefer whites because they believe the likelihood of having to go to court later is lower. In Section 6, we discuss how to differentiate between taste-based and statistical discrimination. Note further that we define discrimination specificallly through action. If an officer exhibits $P_j(s^*, r(i)) = 0$ for all s^* and r, we treat him as not discriminatory by definition, despite the potential for internal animus. We discuss this issue further in Section 4.

The first empirical step we take is to model the likelihood of an individual appearing at the discount point and above, given his observables. To differentiate between the stopped and charged speed, we denote the latter by S_i. The probability of being charged the discount speed is the summed likelihood of appearing at or above that speed times the likelihood of being discounted:

$$\Pr(S_i = s_d | \, i, j\,) = f_r(s_d) + \sum_{s^* > s_d} f_r(s^*) \cdot P_{jr}(s^*) \tag{1}$$

and the probability of appearing at a point above the discount point,

$$\Pr(S_i = s > s_d | \, i, j\,) = f_r(s^*) \cdot \big(1 - P_{jr}(s^*)\big) \tag{2}$$

is the likelihood of having driven that speed and then *not* being discounted.

4 Empirical Strategy

From Equations 1 and 2, we see that racial differences in the likelihood of appearing at the bunch point and above can arise from either differences in speeds $f_r(s^*)$ or differences in speed-specific discounting, $P_{jr}(s^*)$. Primarily in the latter case will the disparity be of policy interest, as it would be due to discrimination rather than differences in behavior. To determine whether the observed disparity is due to differences in driving speed, we use the fact that a large share of officers in our sample practice no lenience. In other words, these officers have no bunching in their distribution of speeds.

In Figure 3, we motivate this approach by documenting the significant heterogeneity in discounting across officers. Panel A plots the officer-level distribution of lenience, defined as the share of tickets written for 9 MPH or above that are for exactly 9 MPH. A large share of officers appear to exhibit very little lenience, with 30% writing less than 1% of tickets for this bunching speed. Further, this heterogeneity across officers cannot be fully explained by the locations and times when they are patrolling. Panel B plots the share of tickets in each county and shift that is written by officers for whom fewer than 2% of their tickets are for 9 MPH over. In the 241 county-shifts in our data, 217 have a share strictly between 0 and 1, indicating significant overlap in the patrolling of lenient and non-lenient officers.[12]

The lower two panels confirm that officers are persistent in their level of lenience across time and location. In Panel C, we plot each officer's residualized lenience in his year with the

[12] Another way to see the importance of officer behavior in generating the observed bunching is to consider the explanatory power of various predictors of a discounted charge. In a regression of a discount indicator on officer fixed effects and location-time fixed effects (described below), the R^2 attributable to the officer effects is 0.312, while that attributable to the location-time effects is 0.26. This finding provides further evidence that officers exhibit significant variability in discount behavior after accounting for location and time.

second-most stops (y-axis) against his residualized lenience in his year with the most stops (x-axis), where we residualize county and month-of-stop fixed effects and driver characteristics. A strong correlation is evident: an officer who charges 9 MPH relatively more often in one year also does so in other years. In Panel D, we plot residualized lenience in the county where the officer has made the second most stops against residualized lenience in the county where he has made the most stops, confirming that officer lenience is highly correlated over space.

To identify the officers who practice no lenience, we use the Frandsen (2017) test for manipulation in bunching. In our setting, this test implies that, under the null hypothesis of no manipulation, the conditional probability of being found at the bunch point is in a range around one third, $P(X = 9|X \in [8, 10]) \in \left[\frac{1-k}{3-k}, \frac{1+k}{3+k}\right]$, where k is a restriction on the second finite difference, $\Delta^{(2)}P(S = 9) \equiv P(S = 8) - 2P(S = 9) + P(S = 10)$, such that $|\Delta^{(2)}P(S = 9)| \leq k\big(P(S = 8) - P(S = 10)\big)$. Intuitively, if the distribution of ticketed speeds is un-manipulated, the share of tickets at 9 MPH over among those charged between 8 and 10 MPH over should be approximately one-third, where the deviation k is due to curvature in the distribution of speeds. We calculate k by assuming the distribution $P(S)$ is Poisson and estimate the mean parameter λ using the empirical mean of ticketed speeds.[13] We say that an officer is non-lenient if we fail to reject that $P(S = 9|S \in [8, 10]) \leq \frac{1+k}{3+k}$ at the 99% confidence level. Out of 1591 officers, we identify 468 as non-lenient.[14]

We use this set of non-lenient officers for two purposes. First, we suppose that these officers' ticketed speeds reflect the true distribution of speeds and use them to uncover the true racial difference in speeding. Secondly, we use these officers as a control group in a difference-in-differences style framework to estimate the effect of encountering a lenient officer on the likelihood of being discounted for each racial group.

To do so, we run a linear probability model, where the outcome is an indicator S_{ij}^k of whether a driver is stopped at a given speed k, and the race of the driver is interacted with

[13] An alternative approach is to follow up with a second step where λ is calculated using the non-lenient officers, and then rerun the test of manipulation for each officer. In practice, this approach leads to no difference in which officers are labelled as lenient.

[14] All of our results are qualitatively similar if we instead identify lenient officers as those with 2% or more of tickets at 9 MPH over.

the lenience of the officer:

$$S_{ij}^{k} = \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2 \cdot \text{Lenient}_j + \beta_3 \cdot \text{White}_i \cdot \text{Lenient}_j + X_i\gamma + X_i \cdot \text{Lenient}_j\, \alpha + \epsilon_{ij} \quad (3)$$

For all regressions, the primary coefficient of interest is β_3, the interaction between white driver and lenient officer. For the bunch point of 9 MPH over the limit, β_3 reflects how much more a white driver "benefits" from encountering a lenient officer than a minority driver. For all speeds above 9 MPH, the interaction reflects how much less likely minorities are to be discounted by a lenient officer. X_i contains the set of all observable characteristics of the drivers, including gender, age, age squared, number of previous tickets, any prior incarceration, whether the driver is in-state, the log average income of the driver's home zip code, vehicle age and age squared, estimated vehicle price, and indicators for vehicle make (where missing is its own category). We also include interactions between the indicator for officer lenient and the above driver covariates (excluding vehicle make fixed effects) to isolate lenience based on race from lenience based on other characteristics that are potentially correlated with race.

We also include fixed effects interacted at the level of the stop's year, month, day of the week, shift, county, and whether it was on a highway, which we henceforth refer to as the time and location of the stop. The purpose of the fixed effects is to make the difference-in-differences comparison among drivers stopped in the same beat and shift. As mentioned earlier, county is our best available approximation to an officer's beat. To provide an even more granular comparison, we will also report results for our GPS sample, where we include fixed effects interacted at the year, month, day of the week, shift, and road segment level.

To calculate each officer's individual discrimination coefficient, we take a similar approach and use non-lenient officers as a control for the baseline frequency of tickets at 9 MPH over, but we allow the coefficients for Lenient_j and $\text{White}_i{\cdot}\text{Lenient}_j$ to vary by individual officer:

$$S_{ij}^{9} = \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2^j \cdot \text{Lenient}_j + \beta_3^j \cdot \text{White}_i \cdot \text{Lenient}_j + X_i\gamma + X_i \cdot \text{Lenient}_j\, \alpha + \epsilon_{ij} \quad (4)$$

The coefficients of interest, β_3^j, are identified from each officer's difference in discounting between whites and minorities, differencing out the disparity in ticketing for non-lenient officers. We denote β_3^j as officer j's degree of discrimination.

We note here two important points about how we are estimating and interpreting discrimination. First, officers can only be as discriminatory as they are lenient. As we document in Appendix Figure A.3, officers with very low levels of discounting have smaller disparities in discounting than officers with moderate lenience. Second, for the purpose of reporting the distribution of discrimination across officers, we treat non-lenient officers as having $\beta_3^j = 0$, since by definition they cannot be discriminatory. These two points serve to emphasize that our object of interest is not an internal measure of racial animus but rather the practice of discrimination. It may be the case that some officers are racially prejudiced but do not express it through discrimination because they are not lenient. While our baseline estimates of the distribution of discrimination take as given these caveats, in Appendix Section C we consider how to estimate the share discriminatory if all officers were to practice lenience.

The intuition for our difference-in-differences procedure is shown in Figure 4. Here we plot the histogram for non-lenient officers over the histogram for lenient officers, separately by driver race. The gap in histograms between lenient and non-lenient officers above 9 MPH over indicates the speeds at which drivers are reduced to 9 MPH over. The difference in these gaps between white and minority drivers indicate the difference in discounting between races for each speed.

For lenient officers to be a valid control group, it must be the case that, conditional on location and time of the stop, the lenience of the officer is uncorrelated with the error term, $\text{Cov}(\text{Lenient}_j, \epsilon_{ij}) = 0$. This assumption entails two presumptions about the stop. First, we require that officers in the same shift and beat are not systematically different in who they stop; second, officers do not systematically differ in the characteristics of drivers to whom they give a warning, which would lead to differential selection into our data. As mentioned above, we see no information about stops that do not result in a ticket, so one concern is that officers who differ in their lenience toward discounting may also differ in their lenience in the initial margin of whether to even write a ticket.

In Figure A.2, we evaluate how the characteristics of an officer's stops vary with whether the officer is lenient or not, where both variables have been residualized with location-time fixed effects. The first three panels show that an officer's lenience is uncorrelated with the share of tickets written to minorities, uncorrelated with the share of tickets where race is missing, and only minimally correlated with the share of tickets that are for speeding. The final panel shows that an officer's lenience is not predictive of of the number of tickets written per day. For this figure we calculate both measures at the annual level, during which officers write most of their tickets in one county, allowing us to control for county-by-year fixed effects.

To further test for selection on observables, Table 2 estimates how officer lenience varies with driver characteristics. Column 1 reports a regression of whether a driver is ticketed on location-by-time fixed effects and driver characteristics. The joint F-test of the null that all driver coefficients are zero has an F-value of 30.6 and a p-value 0.000. When the outcome is changed to the officer's lenience indicator in Column 2, the F-value declines to 0.974, and the p-value increases to 0.477, consistent with the view that officer-driver matches are conditionally random. Columns 3 and 4 replicate Column 2 for our GPS sample. Both with and without fixed effects for the road-segment of the stop, we find that our indicator for officer lenience is uncorrelated with driver characteristics.[15]

4.1 Interpreting the Diff-in-Diff Coefficient

The difference-in-differences regression coefficient β_3 in Equation 3 for 9 MPH over reflects the difference in probabilities of discounting between white and minority drivers among lenient officers relative to non-lenient officers. Using the simple model from Section 3, we can decompose this coefficient into a weighted average of discrimination at different speeds and a term that reflects racial differences in speeding:

$$\begin{aligned}
\beta_3 &= Pr(S = s_d|W, \text{Lenient}) - Pr(S = s_d|W, \text{Non-Lenient}) \\
&\quad -[Pr(S = s_d|M, \text{Lenient}) - Pr(S = s_d|M, \text{Non-Lenient})] \\
&= \sum_{s^*>9} \left[f_w(s^*) \times [Pr_w(s^*) - Pr_m(s^*)]\right] + \sum_{s^*>9} \left[(f_w(s^*) - f_m(s^*)) \times Pr_m(s^*)\right] \qquad (5)
\end{aligned}$$

The first term in the last equation is a weighted average of discrimination among lenient officers at each speed, where the weights are the speed distribution for white drivers, and is our object of interest. The second term represents a bias that is non-zero when white

[15] In Appendix Table A.7, we present a similar check for randomization, but using the officer's estimated discrimination coefficient as the outcome variable. We similarly fail to reject that driver observables are unrelated to officer discrimination conditional on location and time fixed effects.

and minority drivers have different speed distributions. In other words, our estimate of the difference-in-differences coefficient will deviate from the true discrimination coefficient we care to identify when drivers differ in underlying speed in such a way that the second summation above differs from zero.

Using the drivers stopped by non-lenient officers, we can directly measure each race's speed distribution, which we plot in Figure 5. The left panel controls for the time and location of stops, and the right panel additionally controls for all other demographic characteristics.[16] The speed distributions are remarkably similar across races, with average speed gaps of 0.53 and 0.33 MPH over with and without demographic controls, respectively. In Appendix Table A.3, we show that this small gap is consistent across various specifications for time and location controls.

This difference in speeding, while small, is statistically significant. However, we can further exploit the non-lenient officers in our sample to directly estimate the potential bias this gap may induce in our difference-in-differences coefficient. We describe our procedure in detail in Appendix Section B and describe briefly here.

To estimate $f_w(s^*) - f_m(s^*)$ for each speed s^*, we run a regression among the non-lenient officer sample where the outcome is that the driver is ticketed speed s^* and includes location and time fixed effects, driver covariates, and an indicator for driver race being white. We take the coefficient on white driver to be our estimate for the difference in density. To estimate $Pr_m(s^*)$, we use that the probability a minority driver is ticketed at k is $f_m(s^*)$ for non-lenient officers and $(1 - Pr_m(s^*)) \cdot f_m(s^*)$ for lenient officers. We restrict attention to minority drivers and regress whether the driver is ticketed speed s^* on whether the officer is lenient, in addition to shift-time fixed effects and driver covariates. Our estimate for $Pr_m(s^*)$ is then derived by dividing the negative of the coefficient on lenient officer by the statewide share of minority drivers at s^* for non-lenient officers (our estimate of $f_m(s^*)$).

Note that in Section 7, where we introduce a parametric model for conducting counterfactuals, we will jointly estimate officer preferences and driver speeds. In that setup, our estimates will directly account for differences in speeds and avoid the need for a post-estimation correction.

[16]More specifically, for each speed we regress an indicator for being ticketed at that speed on controls and an indicator for minority driver. The plotted distribution for white drivers is the empirical distribution, and the distribution for minority drivers is the white distribution plus the minority coefficient from each speed's regression.

5 Results

Table 3 reports the results of our difference-in-differences test of discrimination where the outcome of interest is whether the driver is ticketed at the discount speed. Columns 1-2 show estimates from a specification with location-time fixed effects only and our preferred specification that includes location-time fixed effects, covariates, and covariates interacted with lenience. The coefficient on the interaction between white drivers and lenient officers indicates that white drivers are 5.8 to 6.9pp more likely to receive a discount than minority drivers, relative to a mean probability of 45 percent. Columns 3-4 show that our discrimination estimates persist, but shrink slightly to 4.3-6pp, when adjusting for stretch-of-road fixed effects using our GPS sample. These coefficients indicate how much more likely a *lenient* officer is to discount a white driver. To calculate a differential probability of discount by an average officer, we use the fact that two-thirds of tickets are written by lenient officers and scale accordingly, finding that an average encounter leads to a 4pp higher discount probability for white drivers, off a base of 30%.

Worth noting is the statistical similarity between the discrimination coefficients in Columns 1 and 2 of Table 3. This finding allows us to rule out a form of "discrimination by proxy." In other words, the racial disparity in treatment we observe is not driven by lenience towards other characteristics that are correlated with race.

Implementing the correction procedure from Section 4.1 to these estimates, we find the degree of bias in our difference-in-differences coefficient induced by racial differences in speeding to be statistically insignificant. Our point estimate for the bias expression in Equation 4.1 is 0.0079, (s.e. 0.0055[17]), which is 13.5% of our baseline coefficient.[18] This measure of bias indicates to us that our difference-in-differences estimator is not meaningfully different from the true behavioral parameter we would like to identify, and the small gap is due to the fact that white and minority drivers have very similar speeds in our data. However, the above correction allows for our approach to be used in other settings where the distribution

[17] We calculate bootstrapped standard errors, where we draw random sets of observations from our data with replacement, calculate the statistic, and iterate 100 times.

[18] As noted in Appendix Section B, our approach requires imputing the probabilities of discounting at 14 MPH and 24 MPH over, where our procedure generates values outside the range of 0 and 1. When we impose the most conservative estimates for these values ($Pr_m(k) = 1$ when $f_w(k) > f_m(k)$ and $Pr_m(k) = 0$ when $f_w(k) < f_m(k)$), we estimate the bias to be 0.0099 (s.e. 0.006), which is 16.9% of the baseline coefficient and also statistically insignificant.

of driving speed (or other measure of criminality) varies substantially across races.

The graphical analogue to Table 3 is presented in Figure 6. Here we report the coefficients on interactions between white driver and lenient officer from regressions where an indicator for being charged speed s (x-axis) is the outcome of interest. The two specifications correspond to Columns 1 and 2 of Table 3. These coefficients indicate where minority drivers are disproportionately being ticketed, and thus the speeds at which white drivers are being differentially discounted. The interaction coefficient is negative and significant for almost all speeds between 10 and 23 MPH, suggesting that at these speeds minorities are less likely to receive a break. Note from Equation 5 that, under equal speed distributions across races, our diff-diff coefficient for 9 MPH over is equal to a weighted sum of coefficients for higher speeds. Because of the insignificant estimate we find for the bias component from speed differences, we think of the coefficient for 9 MPH over approximately reflecting this weighted sum of higher coefficients and conclude that the majority of the disparity in discounting is driven by discrimination at 10 to 23 MPH over.

5.1 Officer-level Heterogeneity

Officer-level results from estimating Equation 4 are reported in Figure 7. The figure displays the across-officer distribution of the interaction coefficient $\hat{\beta}_3^j$, where non-lenient officers are assigned $\hat{\beta}_3^j = 0$. The line represents a kernel density plot of our measure of discrimination against minority drivers, so that the farther right an officer is in the distribution of discrimination, the greater his level of discrimination. The unit of our measure is difference in percentage points probability of discounting. An officer whose discrimination against minorities is 0.1, for example, is 10 percentage points more likely to offer a fine reduction to a white than a minority driver. The percentiles of officer discrimination are also reported in Appendix Table A.6.

The first fact to note is the substantial heterogeneity in discrimination across officers. While the modal officer practices no discrimination, we find a large mass of officers with positive discrimination. Officers at the 10th and 90th percentiles of discrimination have a 14 percentage point difference in their racial disparity. When calculating their lenience toward minorities as a share of their lenience toward whites, officers at the 90th percentile are more than 40% less likely to discount minorities.

The second notable fact is that the median level of discrimination is quite small, three percentage points off of a base of 30% of tickets at 9 MPH over. While this disparity is comparable to the black-white wage gap (Neal and Johnson, 1996), it is possible that the officer in question is not aware of such a disparity. A large literature has explored the role of implicit bias as a source of discrimination (Greenwald and Krieger, 2006; Banks *et al.*, 2006), and in many cases the individual in question is not aware of his bias. We believe that for the median officer our results are consistent with such a theory. However, for higher percentiles of the distribution, it is hard to explain large gaps in treatment as a practice that is imperceptible to the officer. An officer at the 75th percentile has a 6.8pp difference in treatment, and this gap nearly doubles to 12.8pp at the 90th percentile.

Even under a data-generating process in which officers all have the same true discrimination, our estimates would have a distribution due to sampling error. This scenario, however, cannot explain the heterogeneity we find. The average standard error for an officer's $\hat{\beta}_3^j$ is 0.014 – less than one-fourth the standard deviation of $\hat{\beta}_3^j$ across officers, 0.068. In the scenario in which true discrimination is uniform, these numbers would be similar in magnitude. We thus conclude that the majority of the variation is due to true officer differences in discrimination rather than estimation error.[19]

Another approach to understanding the variance in discrimination across officers is to estimate what share of officers are discriminatory. We know that each officer's discrimination measure is an additive function of his true discrimination plus estimation error, $\hat{\theta}_j = \theta_j + \epsilon_j$. We can assume an officer's discrimination can take on a finite set of values on a fine grid, $\theta_j \in \{\theta^k\}$, and calculate the likelihood of observing each officer's discrimination measure $\hat{\theta}_j$ given the noise in the measure and the true distribution $f(\theta_k)$:

$$\text{Prob}(\hat{\Theta}_j = \hat{\theta}_j) = \sum_{\{\theta_k\}} f(\theta_k) \cdot \text{Prob}(\epsilon_j = \theta_k - \hat{\theta}_j)$$

We then estimate $\{f(\theta_k)\}$ by maximum likelihood, where we impose that ϵ_j is normally distributed with σ_j^2 taken from the officer-level regression. Using this approach, and calculating $1 - \sum_{\theta_k < 0} \hat{f}(k)$ as the share, we find that 41% (s.e. 1.33) of officers are discriminatory.[20]

[19]One way to calculate officer heterogeneity's accounting for noise is to do a Bayes shrinkage procedure. When we replicate the approach of Aaronson *et al.* (2007), our distribution of discrimination looks nearly identical to the unshrunk version.

[20]This approach is a discretized version of a deconvolution procedure (Delaigle *et al.*, 2008).

In contrast, we find that only 7% (0.79) of officers have $\theta_j < 0$, i.e., practice reverse discrimination. We can also apply our correction from Section 4.1 here and subtract from each officer's discrimination coefficient our estimate of statistical bias induced by racial differences in speeding. Doing so, we estimate that 32% (1.36) of officers are discriminatory. When removing officers who are non-lenient and thus cannot practice discriminate, our estimate of share discriminatory increases to 52% (2.14).

The left panel of Figure 8 shows how our measure of discrimination varies by officer race. Perhaps consistent with intuition, white officers are much more likely to be discriminatory against minority drivers, with a greater rightward skewness in their distribution. However, minority officers are still, on average, discriminatory against minority drivers. Among black officers, a very small percentage are discriminatory in favor of minority drivers. Some of the disparity in discrimination across officer race is driven by minority officers being less likely to be lenient overall. This fact is due in part to minority officers working in troops in which all officers are less lenient. In the right panel of Figure 8, we show the distribution of discrimination only for lenient officers. The white officers' distribution continues to be shifted farther to the right.

The ability to identify discrimination separately by officer race is a notable advance beyond the previous literature. Several benchmarking papers detect bias using comparisons across officer race (Anwar and Fang, 2006; Antonovics and Knight, 2009; Price and Wolfers, 2010; Anbarci and Lee, 2014). With such an approach, we can know that some race of officers is acting in a discriminatory manner, but not which group. With our method, we can see the magnitude of discrimination separately for each officer race.[21]

In Appendix Section D, we conduct a further analysis of how our full set of officer demographics predict discrimination and lenience. While we find that officer experience and education are unrelated to discrimination, female officers officers are on average less discriminatory. Appendix Section D also probes the stability of our officer-level estimates

Doing the continuous deconvolution leads to an identical estimate for the share of officers who are discriminatory. The grid is 99 points spanning the 1st to 99th percentiles of the empirical distribution of $\hat{\theta}_j$. Confidence intervals are calculated through bootstrapping by performing 100 draws of the set $\{\hat{\theta}_j\}$ and performing MLE on each draw.

[21]Note that the "hit-rate" approach to testing for discrimination, most recently used in Marx (2018) and Arnold *et al.* (2018), also provides an "absolute" measure of discrimination that does not use a comparison group and could in principle be used to estimate discrimination separately by race of the criminal justice agent.

across an officer's career and documents a significant positive relationship between estimates drawn from an officer's early and late samples.

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. We conduct these analyses in Appendix Section E. We estimate a significant degree of discrimination against both minority groups, though our estimates are somewhat larger for Hispanic drivers.

6 Robustness Checks and Alternative Explanations

In this section we report various specification and robustness checks to evaluate the strength of our findings. In particular, we consider various explanations of our findings that are not officer racial discrimination.

In Table 4 we report the primary difference-in-differences results with various changes in the regression specification, with Column 1 re-reporting the baseline specification. In Column 2, we conduct a split-sample analysis where we calculate whether an officer is lenient using a randomly-selected 20% of officers' tickets, which we exclude from the regression. In Column 3, lenience is calculated separately for each officer's year of ticketing, allowing for changes in officer behavior over a career. In Column 4, we re-weight the set of observations so that the "share" minority in each county is the same. This approach is borrowed from Anwar and Fang (2006) and accounts for the possibility that officers differ across counties in their lenience, which could be correlated with minority status.

One concern with our baseline estimates is the sometimes-significant coefficient on white driver. In principle, this coefficient characterizes speed differences between white and non-white drivers among non-lenient officers and should be small based on our analysis in Figure 5. The significance could be due to a model misspecification which attributes heterogeneity in the lenience coefficients to the location and time fixed effects. This issue arises in difference-in-differences models where parts of the treated group (here, the lenient officers) operate as a control for other treated observations by helping to estimate the fixed effects (Borusyak and Jaravel, 2017; Goodman-Bacon, 2018). In Column 5, we re-run the baseline regression where the non-lenient officer observations are given weight of 1000 and lenient officer observations are given weight of 1. By doing so, the lenient officers are only used to identify the lenient officer variable coefficients. The resulting coefficient on white driver is insignificant and greatly reduced in magnitude.

One feature of the data discussed earlier is that the histogram of ticketed speeds exhibits jumps at multiples of five, and we find that this heaping only occurs among "visual" stops. In Column 6 of Table 4, we find that our baseline regression is essentially unchanged when restricting to the sample of tickets from a radar or laser gun.

In all these specifications, the interaction coefficient between officer lenient and driver race is significant and quantitatively similar to the baseline specification.[22] The largest disparity is evident in the re-weighted specification, where the coefficient reduces from 5.8pp to 3.7pp. This difference suggests that some of the gap in treatment between whites and minorities is due to minorities disproportionately driving in counties where officers are less lenient overall. These differences across counties could be due to differences in how much drivers exceed the speed limit. In our model in Section 7, we explicitly account for the possibility that counties and races differ in speeds and continue to find a disparity in discounting between races.

6.1 Selection into the Data

As we state in Section 2, our data are constrained by the fact that we do not observe interactions that do not result in a ticket. One concern is that differences on the margin of whether to give a ticket vary across officers and that this difference may make our estimates of officer-level discrimination inconsistent.

As documented in Section 4, officer lenience is not correlated with any characteristics of the stopped driver or the daily frequency of tickets, consistent with the view that officer-driver matches are random. In Appendix Table A.8, we additionally document that, while minorities are slightly over-represented in our data relative to the Florida population, the racial shares in our ticket sample matches closely the racial shares among individuals involved in a car accident in Florida. These data likely correspond more closely to the demographic composition of speeders than the general population. Overall, we do not have the impression that minorities are severely overrepresented or underrepresented in the tickets data relative

[22] In Appendix Table A.5, we show that our discrimination estimates are also unchanged when restricting to subsamples where speeding is either the only offense or one of at least two offenses, alleviating concerns that minorities are treated more harshly because they are also facing charges for other offenses.

to the general population or the population of speeding drivers in Florida.

We further believe that any discrimination on the stopping margin would likely bias our results toward finding less discrimination in discounting. To see this argument, imagine a minority driver who is on the margin of being ticketed, such that if he were white he would have been let off with a warning. This driver appears in our data only because he is a minority. Because he is at this margin, it is very likely the officer will give him a discount. Therefore, discrimination on the ticketing margin places too many minority drivers in our sample who are disproportionately at the discount point. Thus, the disparity in discounting would be even greater without a hypothetical disparity in ticketing.[23]

In Appendix Section F, we formalize this logic with a simple selection model that allows for officer differences in propensity to let drivers off with a warning. Using this model, we implement a sample selection correction, as in Heckman (1979), that accounts for officer-by-race differences in propensity to appear in the data. We report the results of this regression in Column 7 of Table 4, and all of our primary coefficients look identical to our baseline specification.[24]

6.2 Racial Difference in Requesting a Break

Consistent with the existing literature, our study documents racial differences in the quality of police-civilian interactions (Najdowski, 2011; Najdowski *et al.*, 2015; Trinkner and Goff, 2016; Voigt *et al.*, 2017). However, differences in the quality of the interaction leave open the possibility that white drivers are actually more likely to request a break than minorities. If officers are open to requests for a discount, this difference in solicitations could generate a disparity in lenience.

We argue that racial differences in propensity to request a break cannot explain our findings. Relative to existing studies in the discrimination literature, one strength of our data is that individuals can be linked across tickets, allowing us to evaluate whether there

[23]As pointed out in Brock *et al.* (2012), it is not necessarily the case that an individual at the margin of appearing in the data is guaranteed a certain treatment once in the data. In light of their argument, our selection correction procedure allows for an arbitrary relationship between an individual's propensity to be ticketed and propensity to be discounted.

[24]In Appendix Table A.11, we document how our results vary by daytime versus nighttime, when officers are less able to discern race (Grogger and Ridgeway, 2006). In the potentially less selected sample of nighttime stops, we find a slightly larger coefficient, consistent with the argument that discrimination in the ticketing margin biases our estimates downwards.

are individual-level differences in propensity to receive a discount. To probe this question, we restrict attention to individuals with at least two tickets, comprising 172,810 observations. Running a regression of discounting on individual characteristics has an R^2 of 0.318, and the addition of officer fixed effects leads to an increase to 0.527. In contrast, the further addition of individual-fixed effects only increases the R^2 to 0.542. This small increase shows that, beyond individual covariates and the stopping officer, the specific individual has little explanatory power for whether a discount is given, indicating that individual differences in propensity to request a break is likely not a substantial factor in the disparity in discounting.

6.3 Statistical Discrimination v. Taste-Based Discrimination

We have so far argued that our findings are consistent with officers discriminating against minority drivers, though we have remained agnostic on whether the discrimination is taste-based or statistical. While we feel that statistical discrimination is unlikely in our setting because criminality is observed directly by officers, race could be correlated with unobserved characteristics of interest related to other policing objectives. In this case, disparities in treatment could arise in the absence of internal animus. For example, officers may be choosing whom to discount on the basis of how individuals respond after the stop. Some drivers may be more deterrable and speed less after a harsh ticket, and others may respond by contesting the ticket in court. 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, these estimates do not rule out racial differences in the *responsiveness* to the ticket.

In Appendix Section G, we present a simple test for whether officers are attempting to minimize court contesting or maximize deterrence, which we report in Table 5.[25] To evaluate the impact of a discounted ticket, we instrument for receiving a discount using the stopping officer's persistent (leave-out) level of lenience.[26] Our test then follows the logic of Heckman *et al.* (2010) and claims that non-linearities in the relationship between the outcome and the propensity score reflect sorting of individuals on the basis of their responsiveness.

[25]See Appendix Section A.4 for more information on our measure of court contesting and Appendix Section A.5 for additional background on the institution details regarding fine repayment.

[26]This procedure is very commonly used in the criminal justice literature when judges differ in their punitiveness (Kling, 2006; Dobbie and Song, 2015) We use this approach to evaluate how individuals respond to their ticket in a follow-up paper, Goncalves and Mello (2017).

We find no evidence that officers choose who to discount on the basis of deterrability: the impact of a discount on future speeding is positive but constant across levels of officer lenience. However, we do find that officers choose who to discount based on whether they will contest their ticket in court: among officers who are not very lenient, the marginal impact of giving a driver a discount is a large reduction in likelihood of contesting the ticket. In contrast, more lenient officers have a marginal impact of a discount on court contestation that is significantly smaller, suggesting that more responsive drivers are discounted first. We then perform in Column 5 a hit-rate test similar to Arnold *et al.* (2018) and find that officers' statistical discrimination on court contestation cannot explain the racial disparity in discounting.

Note however that our tests for statistical discrimination only rule out that the pursuit of deterrence or fewer court contestations cannot explain our findings. It may certainly be the case that officers are statistically discriminating on the basis of another objective. In particular, we do not see whether drivers are deterred from non-traffic offending. It may also be the case that officers are "inaccurately" statistically discriminating, in that they are incorrect in their belief about the correlation between race and the objective they care to maximize (Bohren *et al.*, 2019). The test we present is not able to identify this form of statistical discrimination.

7 Model and Counterfactuals

One of the central motivations of our paper is the need to understand how various personnel policies affect the aggregate disparity in treatment between whites and minorities. We have argued that the key input into the outcome of these policies is the distribution of discrimination across officers. To perform counterfactual analyses, however, we need to know both how driver speeds are generated and how officers then choose to discount these speeds. To do so, we present a simple model that allows us to simultaneously estimate officers' taste parameters for each racial group and speed parameters for each race-by-county. This model allows us to perform counterfactuals that change the distribution of discrimination across officers and the population of drivers each officer faces.

Individual i drives at a speed s^* that is drawn from a Poisson distribution $P_{\lambda_i}(s^*)$, where $\lambda_i = \lambda_{rc} + \gamma Z^{(1)}$ is a function of the county-by-race of the driver and other demographics

$Z^{(1)}$. We include in $Z^{(1)}$ the driver's gender, age, and number of tickets in the previous three years.

Within a county, officers and drivers match randomly with each other. If the driver is stopped for a speed s^* at or below the discount point s_d, the officer charges s^*. If $s^* > s_d$, the officer has the choice to discount the driver to s_d. He makes this decision by weighing a cost to discounting, which we impose to have the form $c(s^*) = b \cdot s$, against the "value" of discounting, $t_{ij} = t_{rj} + \alpha Z_i^{(2)} + \epsilon_{ij}$, where t_{rj} depends on the officer identity and driver race, $Z^{(2)}$ are driver demographics, and ϵ_{ij} is a standard normal random variable reflecting differences in preference not captured by driver demographics. Thus, the driver has her speed reduced to s_d if

$$t_{rj} + \alpha Z_i^{(2)} + \epsilon_{ij} > a + b \cdot s_i^*$$

In addition to the $Z^{(1)}$ demographics, $Z^{(2)}$ includes the share of drivers in a county who are minorities. We include this share to account for the possibility that officers change their behavior depending on the racial mix of the county's drivers.

Two simplifications of the model should be discussed here. First, we do not allow the driver's distribution of speeds to respond to the lenience of the officers in their county. We are comfortable in making this restriction because we find that there is no cross-sectional relationship between the county lenience rate and the speeds charged.[27]

Second, we provide no micro-foundation for an officers' decision to discount a driver. In Appendix Section G, we provide a series of tests for identifying what the officer is maximizing. However, for the purposes of conducting the counterfactuals, it suffices to identify differences across officers in their propensity to discount.

7.1 Identification

In principle, our model can be identified using only aggregate information, as if all data came from one officer and one county. Intuitively, the tickets provide 40 moments (for each

[27] In Goncalves and Mello (2017), we find that drivers do respond ex-post to receiving a harsh ticket by speeding less. This should lead to a steady-state relationship between lenience and the frequency of traffic tickets. However, the magnitude of the deterrence effect is small enough that the racial gaps in the counterfactuals would not be meaningfully impacted. For example, in the 11 years of our sample, if all minority drivers were treated as white drivers, there would only be about 70 more car accidents and fewer than one more death in expectation.

potential speed) to estimate three parameters (discount slope, preference for discounting, and true speed). Such an estimation approach relies heavily on the functional form assumptions of a Poisson speed distribution.

In practice, our estimation is similar to our difference-in-differences regressions, in that it relies heavily on the heterogeneity across officers in discount lenience. While all officers' data enter the maximum likelihood equations, the speed parameters are primarily identified using officers who exhibit no lenience, from which we get an estimate of the true distribution of speeds. To do so, we strongly rely on the assumption that officers and drivers are randomly sorting within a county, allowing us to suppose that the underlying distribution of speeds are the same for non-lenient and lenient officers.

Our estimation also depends heavily on the smoothness and parameterization of the underlying speed distribution. Any excess mass at the bunch point is taken to be lenience on the part of the officer. As argued earlier, we believe this assumption is valid, and drivers are not systematically choosing to bunch below the fine increase.

We estimate the model via maximum likelihood. The model parameters to be identified are the 67×2 county-race speeds λ_{rc}; 3 demographic speed parameters γ; 1592×2 officer average racial preferences, t_{rj}; 4 demographic preference parameters α; and the slope of the cost function b, totaling 3,326 parameters. Details of how the estimation is carried out in practice are provided in Appendix Section H.

7.2 Model Estimates

The results of the model estimation are reported in Appendix Table A.14. Because the estimates are closely aligned to the findings from our difference-in-differences approach, we leave our full discussion of these estimates to Appendix Section H.1. In short, we find that the average officer practices substantial lenience, with a significant variance across officers. Off a baseline of 35.7% likelihood of discounting a driver from 10 MPH to 9 MPH, the average officer is 2pp less likely to discount minority drivers. We find that minorities drive significantly faster than white drivers, as do males, younger drivers, and drivers with previous tickets. The average officer is also more generous to female drivers, old drivers, and drivers with fewer previous tickets. They are also less lenient to all drivers when ticketing in a county with more minorities.

7.2.1 Decomposing the Gap in Discounting

A first-order question in the study of discrimination is the extent to which an aggregate racial disparity can be explained by the measured amount of discrimination. Table 6 seeks to answer this question by decomposing the measured racial discounting disparity into discrimination by officers, sorting of officers across counties, and differential speeding by racial groups. We do so by simulating the model with different restrictions on the behavior and location of the officers. In each simulation, drivers are randomly re-assigned a new officer from their county and drawn a new speed s from their individual specific distribution $P_{\lambda_i}(s)$. If the driver's speed is above the discount point, the officer draws a preference shock ϵ and gives the driver a discount to 9 MPH over if $t_{ij} + \epsilon > b \cdot x$. Standard errors are calculated by iterating the simulation 100 times, as explained in Appendix Section H.2.

The "Baseline" row of Table 6 shows how the charged speeds of drivers appear in a simulation of the model that does not change any of the parameters of the model. All of the decompositions are benchmarked to this baseline. In the "No Discrimination" row, we remove discrimination by making each officer treat minority drivers like they treat white drivers. This restriction reduces the gap in discounting by 25%. In the "No Sorting" row, drivers and officers match randomly from throughout the entire state rather than the initial county. Here we find that 28% of the gap in discounting is removed, consistent with the earlier finding that officers tend to be more lenient overall in neighborhoods with fewer minorities. Removing both sorting and discrimination, the gap in speeding is reduced by 45%. The remaining gap is due exclusively to the fact that minorities are driving faster speeds. In the second panel of Table 6 we report the same decompositions, where the gap is conditional on the county of the stop. Removing the sorting of officers no longer has any effect, since that only leads to differences *across* counties. Further, notice that over 80% of the within-county disparity can be explained by discrimination, leaving only about 17% of the disparity to be explained by differences in speeding across races. In Appendix Table A.15, we perform these same calculations, where the outcome of interest is the average speed rather than share discounted, and find similar results.

7.3 Policy Counterfactuals

Reported in Table 7, we now use the estimates to conduct a series of policy counterfactuals to explore how best to curb discrimination in speeding tickets. The results of these counterfactuals are compared relative to a baseline simulation, reported in the first row, that retains the empirical pool of officers and their distribution across counties. As with Table 6, the calculation of standard errors is discussed in Appendix Section H.

7.3.1 Firing and Hiring

We first consider the most direct policy for mitigating the disparity in treatment: removing the most discriminatory officers. We take officers in the 95th percentile and above of discrimination and remove them from the pool of officers. This cutoff removes officers with a difference in discounting of 16 percentage points or greater between whites and minorities. For symmetry, we also remove officers who reverse discriminate by that amount (comprising only 0.4% of officers).

The statewide disparity in treatment barely changes in response to removing these officers, falling by less than 4%. The lack of effectiveness from this policy partly stems from the fact that discriminatory officers are on average very lenient. When they are removed, drivers are left to be stopped by officers who, while less discriminatory, are also less lenient overall. This fact can be seen by noting that the average discount rate goes down for both white and minority drivers.

The next counterfactuals we consider are increased hiring of minority and female officers. Given our earlier finding that minority and female officers exhibit lower levels of discrimination, we should expect that increasing their presence might lead to lower levels of aggregate bias. We calculate this counterfactual by re-simulating which officer each driver draws, taken from within his county, where the probability of drawing a minority or female officer is exogenously changed. Consistent with our intuition, the gap in probability of discount declines, though very modestly. Increasing the share of female officers from 8% to 18% of the force leads to a 7.5% reduction in the discount gap. An increase in minority officers from the empirical share of 35% to 45% reduces the gap by 13.5%.

Demographic policies have been suggested in the past as a possibility for systemically changing police behavior, particularly toward poor and minority communities. Donohue III and Levitt (2001) find that an increase in minority officers leads to an increase in arrests of white offenders, no effect on non-white offenders, and vice versa for an increase in white officers. Our results, though only counterfactuals, are qualitatively consistent with their findings.

7.3.2 Resorting Officers

The final counterfactuals we consider are to reassign officers to specific areas based on their behavior and the share of minorities in each county. Officers are assigned to troops, which patrol 6-10 counties. Within the troops, officers regularly vary in which locations they patrol. It may be potentially feasible for a senior officer to, for example, change the assignment of officers such that minorities face less biased officers. The bottom two rows of Table 7 present the results of such a policy. Column 1 sorts officers within a troop such that the least biased officers are in counties with the most minorities. Column 2 sorts officers within a troop such that the most lenient officers are in counties with the most minorities.

Surprisingly, sorting officers to expose minorities to the least discriminatory has a very small effect on the treatment gap. The least biased officers are also not very lenient on average, dampening the impact of their equal discounting across races and reducing the gap in discounting by only 11%. Much more effective in reducing the gap in treatment is assigning the most *lenient* officers to minority counties. This policy reduces the treatment gap by 86%.

In short, the counterfactual analyses highlight the importance of absolute lenience as a consideration separate from discrimination. The policy aimed at exposing minorities to lenience is much more effective than removing overall bias through firing biased officers or hiring minority and female officers.

7.4 Caveats

Our simplified modeling framework and counterfactuals are meant to be suggestive of how the racial treatment gap might change when various personnel policies are considered. That being said, many caveats must be recognized. We are not taking a strong normative stance on the social welfare function, and the only outcome we consider is the statewide disparity in discounting. Other outcomes could be relevant to the policy makers's problem that we do not consider here.

For example, increasing lenience uniformly may lead to increased speeding, which we show to be the case in a separate study (Goncalves and Mello, 2017). Changing leniency standards may also lead officers to give drivers verbal warnings rather than a reduced charge. A full consideration of the welfare impact of the ensuing policies would likely consider additional outcomes, such as the speeding response to changes in enforcement (Gehrsitz, 2017; Goncalves and Mello, 2017; Chalfin and McCrary, 2017) and the tradeoff between the level and inequality in lenience.

One additional concern is that officers will change their lenience behavior in response to being reassigned counties. We address this concern in part by allowing officer behavior to vary by the share of drivers who are minorities, though it is important to note that officers may respond in other ways.

Finally, we have abstracted from some aspects of the data in order to focus on the central features of the setting. In particular, we do not model the fact that officers sometimes reduce their driver's speed to a multiple of five or that they sometimes appear to discount the speed to 14 MPH over instead of 9. In Appendix Section I, we present an extended model that allows for these additional features, and we find that our baseline estimates of officer discrimination are unchanged.

8 Conclusion

The large racial disparities in the criminal justice system have led many to claim discrimination as the root cause. In this paper, we argue that identifying discrimination at the level of the individual criminal justice agent is crucial for understanding the best policy for mitigating disparities in outcomes.

We study speeding tickets and the choice of officers to discount drivers to a speed just below an onerous punishment. Using a bunching estimator in a difference-in-differences framework, we document that minority drivers are significantly less likely to be given a discounted speed on their ticket. A key advantage of our approach is the ability to explore the entire distribution of both lenience and discrimination on the part of officers. Our estimates reveal significant heterogeneity in behavior across officers, with 40% of the force explaining the entire aggregate disparity. Estimates from our parametric model of driver speeding and officer decision-making confirm that, while minorities drive slightly faster on average, our officer-level estimates of discrimination, which leverage non-lenient officers as a control group, are not confounded by differences in speeding across racial groups. On net, we attribute 83% of the racial gap in discounting to discrimination by patrol officers.

Using our model estimates to explore various counterfactuals, we find that policies targeting discrimination directly have only a modest effect on the aggregate racial gap in treatment. The limited effectiveness of such policies is due to the fact that minorities tend to reside in regions where officers are less lenient towards all drivers. On the other hand, policies targeting officer lenience, such as reassigning lenient officers to minority neighborhoods, are much more effective at reducing the aggregate disparity. These counterfactuals highlight our central argument that the impacts of various policy reforms depend crucially on the *distribution across officers* in their degree of discrimination.

Our paper raises several interesting questions for future research. While we find that the relationship between average lenience and neighborhood racial composition has critical implications for policy effectiveness, diagnosing the root cause of this empirical fact is beyond the scope of our paper. Understanding the extent to which institutional inputs such as troop leadership or peer effects can explain the distribution of officer behavior could raise other potential policy solutions. Finally, evidence on how officers respond in practice to policies targeting discriminatory policing will be critical for policy design in the future.

References

Aaronson, D., Barrow, L. and Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics*, **25** (1), 95–135.

Abrams, D. S., Bertrand, M. and Mullainathan, S. (2012). Do judges vary in their treatment of race? *The Journal of Legal Studies*, **41** (2), 347–383.

Agan, A. Y. and Starr, S. B. (2016). Ban the box, criminal records, and statistical discrimination: A field experiment.

Aizer, A. and Doyle Jr, J. J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics*, **130** (2), 759–803.

Anbarci, N. and Lee, J. (2014). Detecting Racial Bias in Speed Discounting: Evidence from Speeding Tickets in Boston. *International Review of Law and Economics*, **38**, 11–24.

Antonovics, K. and Knight, B. (2009). A New Look at Racial Profiling: Evidence from the Boston Police Department. *Review of Economics and Statistics*, **91** (1), 163–177.

Anwar, S., Bayer, P. and Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *The Quarterly Journal of Economics*, **127** (2), 1017–1055.

— and Fang, H. (2006). An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *The American Economic Review*, **96** (1), 127–151.

Arnold, D., Dobbie, W. and Yang, C. S. (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics*, **133** (4), 1885–1932.

Arrow, K. (1973). The theory of discrimination. *Discrimination in labor markets*, **3** (10), 3–33.

Banks, R. R., Eberhardt, J. L. and Ross, L. (2006). Discrimination and implicit bias in a racially unequal society. *California Law Review*, **94** (4), 1169–1190.

Becker, G. S. (1957). *The Economics of Discrimination: an Economic View of Racial Discrimination*. University of Chicago.

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *The American Economic Review*, **94** (4), 991–1013.

Best, M. C., Cloyne, J., Ilzetzki, E. and Kleven, H. J. (2015). Interest rates, debt and intertemporal allocation: Evidence from notched mortgage contracts in the united kingdom.

Björklund, A. and Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics*, pp. 42–49.

Bohren, J. A., Haggag, K., Imas, A. and Pope, D. G. (2019). *Inaccurate Statistical Discrimination*. Tech. rep., National Bureau of Economic Research.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs. *Available at SSRN 2826228*.

Brock, W. A., Cooley, J., Durlauf, S. N. and Navarro, S. (2012). On the observational implications of taste-based discrimination in racial profiling. *Journal of Econometrics*, **166** (1), 66–78.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J. and Mullainathan, S. (2016). Productivity and selection of human capital with machine learning. *The American Economic Review*, **106** (5), 124–127.

— and McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, **55** (1), 5–48.

Chetty, R., Friedman, J. N., Olsen, T. and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics*, **126** (2), 749–804.

Coviello, D. and Persico, N. (2013). An Economic Analysis of Black-White Disparities in NYPD's Stop and Frisk Program. *NBER Working Paper*, pp. 1–28.

Dee, T., Dobbie, W., Jacob, B. and Rockoff, J. (2016). The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examination. *NBER Working Paper*, pp. 1–67.

Delaigle, A., Hall, P. and Meister, A. (2008). On deconvolution with repeated measurements. *The Annals of Statistics*, pp. 665–685.

Diamond, R. and Persson, P. (2016). The Long-Term Consequences of Teacher Discretion in Grading of High-Stakes Tests. *SIEPR Discussion Paper*, **16-003**, 1–70.

Dobbie, W. and Song, J. (2015). Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. *The American Economic Review*, **105** (3), 1272–1311.

Donohue, J. (2014). An Empirical Evaluation of the Connecticut Death Penalty System Since 1973: Are There Unlawful Racial, Gender, and Geographic Disparities? *Journal of Empirical Legal Studies*, **11** (4), 637–696.

Donohue III, J. J. and Levitt, S. D. (2001). The impact of race on policing and arrests. *The Journal of Law and Economics*, **44** (2), 367–394.

Doyle, J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *The American Economic Review*, **97** (5), 1583–1610.

Doyle Jr, J. 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.

Edelman, B., Luca, M. and Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics*, **9** (2), 1–22.

Frandsen, B. 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*, Emerald Publishing Limited, pp. 281–315.

Fryer, R. G. (2018). An Empirical Analysis of Racial Differences in Police Use of Force. *NBER Working Paper*, **22399**, 1–63.

Gehrsitz, M. (2017). Speeding, punishment, and recidivism: Evidence from a regression discontinuity design. *The Journal of Law and Economics*, **60** (3), 497–528.

Goldin, C. and Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review*, **90** (4), 715–741.

Goncalves, F. and Mello, S. (2017). Does the punishment fit the crime? speeding fines and recidivism. *Unpublished Manuscript*, pp. 1–48.

Goodman-Bacon, A. (2018). *Difference-in-differences with variation in treatment timing*. Tech. rep., National Bureau of Economic Research.

Greenwald, A. G. and Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review*, **94** (4), 945–967.

GROGGER, J. and RIDGEWAY, G. (2006). Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness. *Journal of the American Statistical Association*, **101** (475), 878–887.

HECKMAN, J. J. (1979). Sample selection as a specification error. *Econometrica*, **47** (1), 153–161.

—, SCHMIERER, D. and URZUA, S. (2010). Testing the correlated random coefficient model. *Journal of Econometrics*, **158** (2), 177–203.

HORRACE, W. C. and ROHLIN, S. M. (2016). How Dark Is Dark? Bright Lights, Big City, Racial Profiling. *Review of Economics and Statistics*, **98** (2), 226–232.

KLEVEN, H. J. (2016). Bunching. *Annual Review of Economics*, **8**, 435–464.

KLING, J. R. (2006). Incarceration length, employment, and earnings. *The American economic review*, **96** (3), 863–876.

KNOWLES, J., PERSICO, N. and TODD, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy*, **109** (1), 203–229.

LANGE, J. E., JOHNSON, M. B. and VOAS, R. B. (2005). Testing the racial profiling hypothesis for seemingly disparate traffic stops on the new jersey turnpike. *Justice Quarterly*, **22** (2), 193–223.

MARX, P. (2018). An absolute test of racial prejudice. *Unpublished Manuscript*, pp. 1–22.

MUELLER-SMITH, M. (2014). The criminal and labor market impacts of incarceration. *Unpublished Working Paper*.

NAJDOWSKI, C. J. (2011). Stereotype threat in criminal interrogations: Why innocent black suspects are at risk for confessing falsely. *Psychology, Public Policy, and Law*, **17** (4), 562.

—, BOTTOMS, B. L. and GOFF, P. A. (2015). Stereotype threat and racial differences in citizens' experiences of police encounters. *Law and Human Behavior*, **39** (5), 463.

NEAL, D. A. and JOHNSON, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy*, **104** (5), 869–895.

PERSICO, N. (2009). Racial Profiling? Detecting Bias Using Statistical Evidence. *Annual Review of Economics*, **1** (1), 229–254.

PHELPS, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, **62** (4), 659–661.

PRICE, J. and WOLFERS, J. (2010). Racial Discrimination Among NBA Referees. *Quarterly Journal of Economics*, **125** (4), 1859–1887.

REHAVI, M. M. and STARR, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, **122** (6), 1320–1354.

RIDGEWAY, G. and MACDONALD, J. (2010). Methods for Assessing Racially Biased Policing. In S. Rice and M. White (eds.), *Race, Ethnicity, and Policing New and Essential Readings*, pp. 180–204.

— and MacDonald, J. M. (2009). Doubly robust internal benchmarking and false discovery rates for detecting racial bias in police stops. *Journal of the American Statistical Association*, **104** (486), 661–668.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, **2** (3), 180–212.

Smith, W., Tomaskovic-Devey, D., Zingraff, M., Mason, H. M., Warren, P. and Wright, C. (2004). The North Carolina Highway Traffic Study. *NCJRS Grant Report*, pp. 1–407.

Trinkner, R. and Goff, P. A. (2016). The color of safety: The psychology of race & policing. *The SAGE Handbook of Global Policing*, pp. 61–81.

Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., Jurgens, D., Jurafsky, D. and Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences*, p. 201702413.

Walker, S., Alpert, G. P. and Kenney, D. J. (2000). Early warning systems for police: Concept, history, and issues. *Police Quarterly*, **3** (2), 132–152.

West, J. (2018). Racial Bias in Police Investigations. *Working Paper*, pp. 1–36.

Table 1: Summary Statistics

A. Tickets	(1) White Driver	(2) Black	(3) Hispanic	(4) All
Driver Female	0.362	0.402	0.305	0.356
	(0.481)	(0.490)	(0.461)	(0.479)
Age	0.037	0.034	0.034	0.036
	(0.015)	(0.012)	(0.012)	(0.014)
Florida License	0.825	0.851	0.896	0.846
	(0.380)	(0.356)	(0.306)	(0.361)
Zip Code Income	52.716	37.620	44.001	47.922
	(51.933)	(29.938)	(41.765)	(46.708)
Vehicle Price / 1000	18.856	17.964	19.046	18.728
	(9.759)	(8.542)	(9.792)	(9.557)
Citations in Past Year	0.288	0.427	0.408	0.341
	(0.721)	(0.909)	(0.877)	(0.799)
MPH Over	15.560	16.658	18.334	16.404
	(6.524)	(7.033)	(6.988)	(6.825)
Discount	0.343	0.314	0.204	0.306
	(0.475)	(0.464)	(0.403)	(0.461)
Fine Amount	182.060	187.999	197.436	186.636
	(76.130)	(80.366)	(80.401)	(78.154)
Share	0.584	0.184	0.231	1
N	667086	210272	264270	1141628

B. Officers	(1) White Officer	(2) Black	(3) Hispanic	(4) All
Female	0.070	0.140	0.075	0.081
	(0.256)	(0.348)	(0.265)	(0.273)
Bachelor's Degree	0.516	0.598	0.364	0.507
	(0.500)	(0.493)	(0.484)	(0.500)
Total Tickets	716.045	846.396	594.443	717.554
	(707.639)	(899.260)	(1047.451)	(852.605)
Yearly Tickets	101.714	118.349	98.389	104.428
	(84.953)	(110.845)	(112.529)	(98.665)
Tenure	11.894	12.294	7.427	10.947
	(9.966)	(9.626)	(7.070)	(9.544)
Share	0.627	0.157	0.200	1
N	998	250	318	1591

Notes: Standard deviations in parentheses. Zip code income is missing for 3.0% of White stops, 2.7% of Black stops, 4.1% of Hispanic stops. Vehicle information is missing for 26.1% of White stops, 26.4% of Black stops, 34.2% of Hispanic stops. To account for the fact that a large share of fine amounts are missing or zero in our data, we impute the fine amount with the modal nonzero fine for each county × speed over the limit cell. In all tables, Discount is an indicator for a charged speed of 9 MPH over the posted limit.

Table 2: Officer Lenience Randomization Check

	Full Sample		GPS Sample	
	(1) Discount	(2) Lenient	(3) Lenient	(4) Lenient
Driver Black	-0.0209	-0.000130	-0.000618	0.00369
	(0.00254)	(0.00253)	(0.00397)	(0.00360)
Driver Hispanic	-0.0334	0.000360	-0.00661	0.00439
	(0.00277)	(0.00306)	(0.00608)	(0.00361)
Driver Female	0.0260	0.00259	0.00223	0.000720
	(0.00172)	(0.00173)	(0.00291)	(0.00275)
Florida License	0.0117	0.00707	0.000434	-0.00412
	(0.00358)	(0.00332)	(0.00410)	(0.00280)
Age	1.311	-0.0272	0.0539	-0.0430
	(0.108)	(0.112)	(0.0997)	(0.0890)
1 Prior Ticket	-0.0114	-0.000963	-0.00236	-0.00105
	(0.00118)	(0.00124)	(0.00264)	(0.00330)
2+ Prior Ticket	-0.0278	-0.000121	-0.00177	-0.00462
	(0.00191)	(0.00212)	(0.00371)	(0.00429)
Any Past Prison	-0.0278	-0.00396	-0.0151	-0.0112
	(0.00465)	(0.00384)	(0.00881)	(0.0106)
Log Zip Code Income	-0.00559	0.000183	0.000515	-0.00180
	(0.00169)	(0.00204)	(0.00296)	(0.00290)
Vehicle Price / 1000	-0.000462	-0.0000644	-0.000207	0.000100
	(0.000114)	(0.000148)	(0.000277)	(0.000322)
Vehicle Age	0.00104	-0.0000153	-0.000368	0.000125
	(0.000205)	(0.000235)	(0.000448)	(0.000559)
F-val	30.6	.974	.959	.663
F-test	0	.477	.495	.812
Mean	.302	.74	.803	.789
Location + Time FE	X	X	X	
GPS FE				X
Observations	1096077	1096077	126841	70330

Notes: All regressions includes vehicle make fixed effects and fixed effects at the interacted level of county, road type (highway v. not), year, month, day of the week, and shift. The F-test reports the joint hypothesis test that variables Driver Black through vehicle age are zero, in addition to indicators for whether income, vehicle price, and vehicle age are missing. Standard errors are clustered at the county level. "Location + Time FE" indicate fixed effects at the interacted level of county, road type (highway v. not), year, month, day of the week, and shift. "GPS FE" includes road segment by year by month by day of the week by shift fixed effects.

Table 3: Difference-in-Differences Results

	Full Sample		GPS Sample	
	(1) Discount	(2) Discount	(3) Discount	(4) Discount
Driver White	-0.0198	-0.0152	-0.00424	-0.00107
	(0.00500)	(0.00509)	(0.00501)	(0.00306)
Officer Lenient	0.217	0.168	0.123	0.0637
	(0.0181)	(0.0305)	(0.0298)	(0.0287)
Driver White	0.0689	0.0583	0.0603	0.0431
× Officer Lenient	(0.00651)	(0.00671)	(0.00781)	(0.00606)
Mean	.302	.302	.314	.299
Covariates		X	X	X
Location + Time FE	X	X	X	
GPS FE				X
Observations	1096077	1096077	126841	102172

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 4: Alternative Difference-in-Differences Specifications

	(1) Discount	(2) Discount	(3) Discount	(4) Discount	(5) Discount	(6) Discount	(7) Discount
Driver White	-0.0142	-0.0154	-0.0109	0.00147	0.00124	0.0120	-0.0156
	(0.00513)	(0.00478)	(0.00347)	(0.00180)	(0.000516)	(0.00649)	(0.00478)
Lenient	0.165	0.169	0.268	0.185	0.210	0.175	0.163
	(0.0312)	(0.0264)	(0.0121)	(0.0361)	(0.0297)	(0.0333)	(0.0311)
Driver White	0.0584	0.0617	0.0600	0.0373	0.0563	0.0450	0.0577
× Officer Lenient	(0.00677)	(0.00649)	(0.00524)	(0.00342)	(0.00713)	(0.00863)	(0.00625)
Selection Correction							0.0453
							(0.0177)
Specification	Baseline	Split-Sample	Lenience by Year	Race Re-weighted	Lenience Re-weighted	Radar Gun Sample	
Difference		.003	.001	-.022	-.003	-.014	-.001
		(.009)	(.008)	(.007)	(.009)	(.01)	(.009)
Mean	.306	.305	.306	.307	.306	.371	.306
R2	.37	.378	.403	.377	.303	.353	.371
N	1078973	853357	1078973	1073770	1078973	109478	1078973

Notes: All regressions include vehicle type fixed effects and fixed effects for county-year-month. Standard errors are clustered at the county level. The baseline specification is the same regression as Column 3 from Table 3. Column 2 reports a regression where a random sample of 20% of the data is used to estimate whether an officer is lenient, and the remaining 80% is used in the regression. Column 3 allows officer lenient/non-lenient to vary by year. Column 4 re-weights the observations so that the relative weight given to minority drivers is equalized across county-year-month. Column 5 interacts officer lenient/non-lenient with all observable demographics of drivers. Column 6 restricts attention to a sample of tickets where the officer reports that he/she used a radar gun to identify the driver's speed. Column 7 reports the baseline regression with the addition of the Heckman Correction term, as explained in Section 6.1 and Appendix Section F.

Table 5: Testing for Statistical Discrimination

	(1) Recidivism	(2) Recidivism	(3) Court	(4) Court	(5) Court
P(Discount)	0.00751	0.0181	-0.133	-0.173	-0.122
	(0.00266)	(0.00767)	(0.00741)	(0.0187)	(0.00748)
Driver Minority	0.00633	0.00627	0.0470	0.0473	0.0560
	(0.000922)	(0.000922)	(0.00166)	(0.00166)	(0.00244)
P(Discount)^2		-0.0125		0.0464	
		(0.00854)		(0.0204)	
P(discount) × Driver Minority					-0.0304
					(0.00489)
Location + Time FE	X	X	X	X	X
Mean	.105	.105	.389	.389	.389
R2	.019	.019	.37	.37	.37
N	802507	802507	802507	802507	802507

Notes: Columns 1-2 use as outcome whether an individual receives another speeding ticket in Florida in the following year. Column 1 regresses recidivism on driver demographics and the propensity score for receiving a discount. The propensity score uses driver demographics and an instrument for officer lenience interacted with driver race, as explained in Appendix Section G. Column 2 additionally includes a quadratic term for the propensity score. Columns 3 and 4 are analogous to Columns 1 and 2, where the outcome is whether the driver contests the ticket in court. Column 5 regresses court contestation on propensity score, where propensity score is also interacted with driver race. For all regressions, 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.

Table 6: Discounting Gap Decomposition

	State-Wide Disparity			
	(1) White Mean (MPH)	(2) Minority Mean	(3) Difference	(4) Percent
Baseline	0.347	0.266	-0.081	100
	(0.001)	(0.001)	(0.001)	
No Discrimination	0.347	0.286	-0.061	75.553
	(0.001)	(0.001)	(0.001)	(0.010)
No Sorting	0.327	0.269	-0.059	72.045
	(0.001)	(0.001)	(0.001)	(0.014)
Neither	0.327	0.291	-0.037	45.016
	(0.001)	(0.001)	(0.001)	(0.012)

	County-Level Disparity			
	(1) White Mean (MPH)	(2) Minority Mean	(3) Difference	(4) Percent
Baseline	0.347	0.321	-0.027	100
	(0.001)	(0.001)	(0.001)	
No Discrimination	0.347	0.343	-0.005	17.903
	(0.001)	(0.001)	(0.001)	(0.033)
No Sorting	0.327	0.300	-0.027	100.888
	(0.001)	(0.001)	(0.001)	(0.043)
Neither	0.327	0.322	-0.005	17.656
	(0.001)	(0.001)	(0.001)	(0.035)

Notes: Table presents how the racial gap in discounting and changes without bias and sorting of officers across counties. The probability gap is the probability of being discounted if you are at the speed right above the jump in fine. Both gaps are 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. The county-level disparities re-weight the minority observations so that the "share" minority is identical across counties.

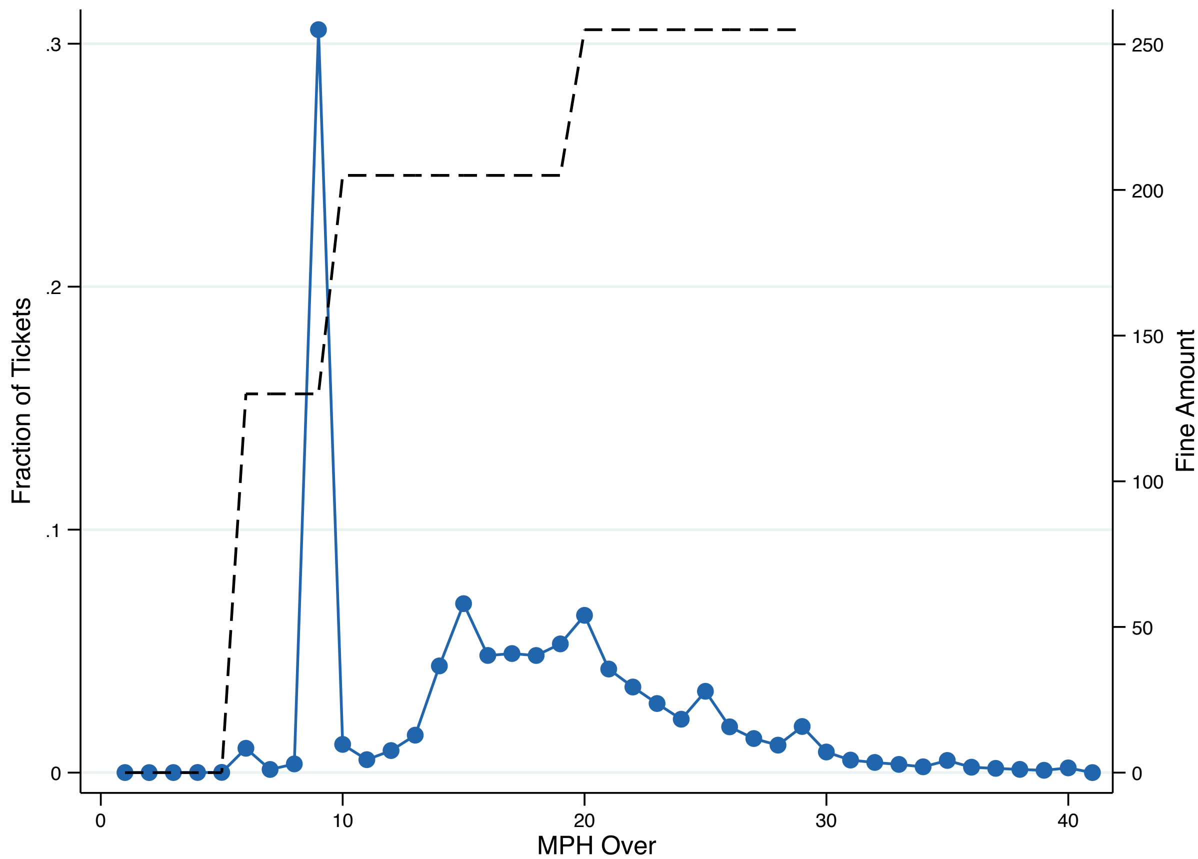
Table 7: Model Counterfactuals

	Hiring & Firing			
	(1) White Mean	(2) Minority Mean	(3) Difference	(4) Percent
Baseline	0.3473	0.2661	-0.0813	100
	(0.0007)	(0.0007)	(0.0010)	
Fire 5% Most Biased Officers	0.3440	0.2655	-0.0785	96.5733
	(0.0013)	(0.0012)	(0.0012)	(0.0154)
Increase Female Share 10pp	0.3423	0.2671	-0.0752	92.4890
(Base of 8%)	(0.0007)	(0.0008)	(0.0011)	(0.0131)
Increase Minority Share 10pp	0.3057	0.2354	-0.0703	86.5128
(Base of 35%)	(0.0016)	(0.0018)	(0.0024)	(0.0293)

	Resorting Officers			
	(1) White Mean	(2) Minority Mean	(3) Difference	(4) Percent
Exposing Minorities	0.3327	0.2602	-0.0725	89.1587
To *Least Biased*	(0.0022)	(0.0018)	(0.0017)	(0.0212)
Exposing Minorities	0.2989	0.2879	-0.0110	13.5403
To *Most Lenient*	(0.0008)	(0.0010)	(0.0011)	(0.0129)

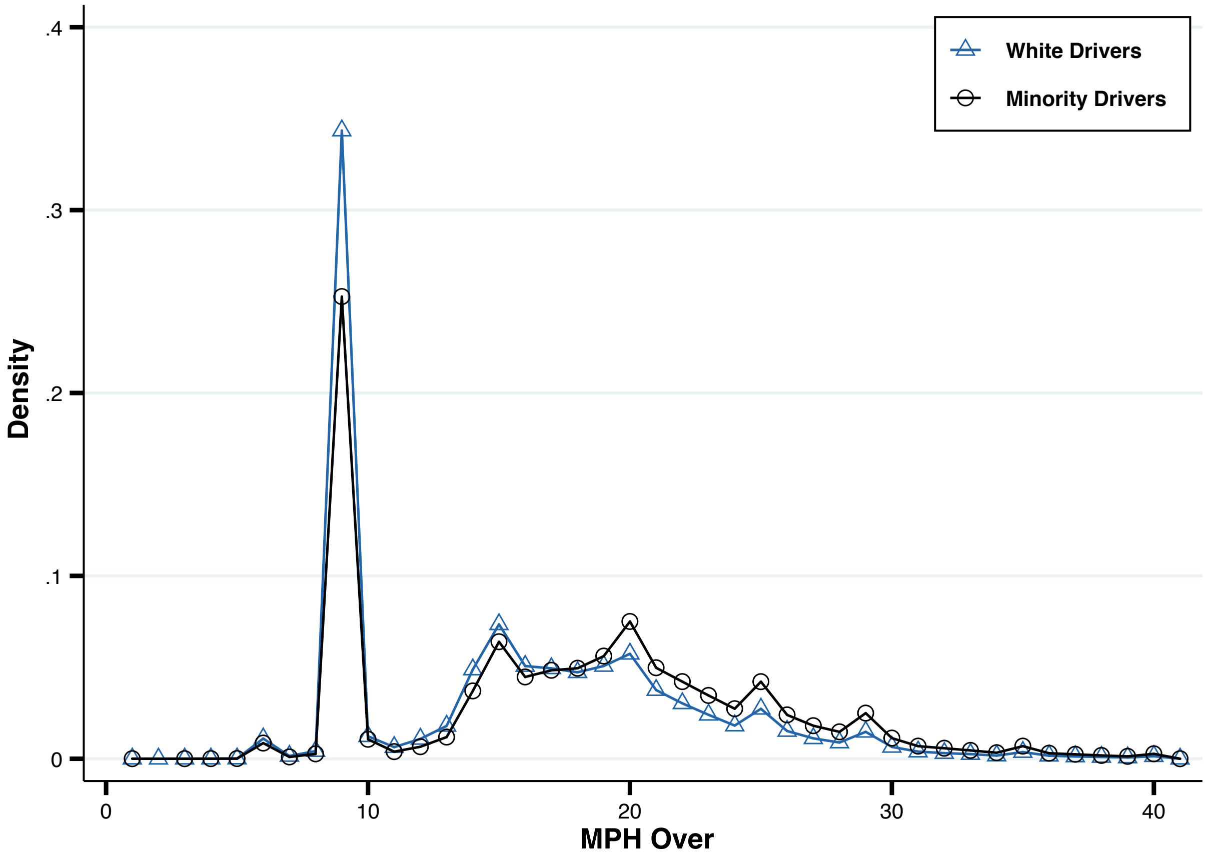
Notes: Results are reporting the probability of being ticketed 9MPH over, where the averages are statewide. In the bottom panel of counterfactuals, officers are resorted *within* troops.

Figure 1: Distribution of Charged Speeds and Fine Schedule



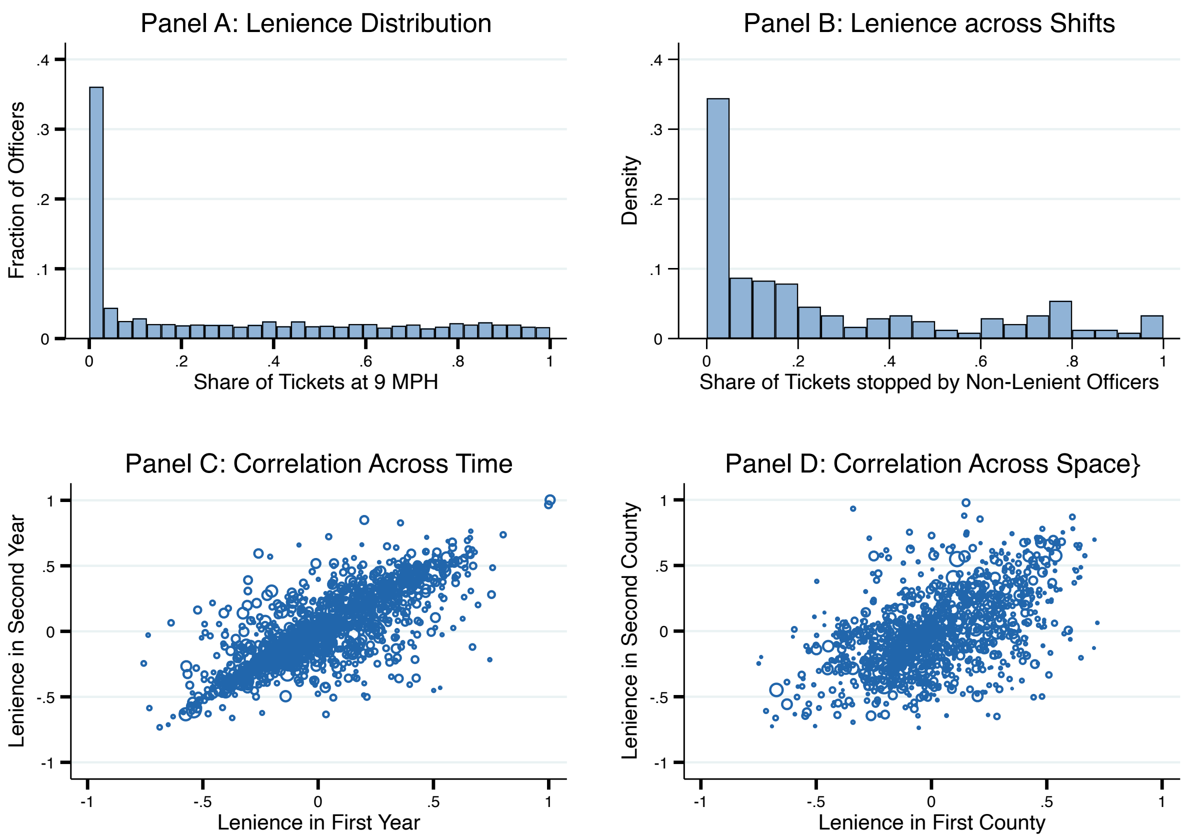
Notes: Connected line shows histogram of tickets. Dashed line plots fine schedule for Broward County. 30 MPH over is felony speeding and carries a fine to be determined following a court appearance.

Figure 2: Charged Speed Distributions by Driver Race



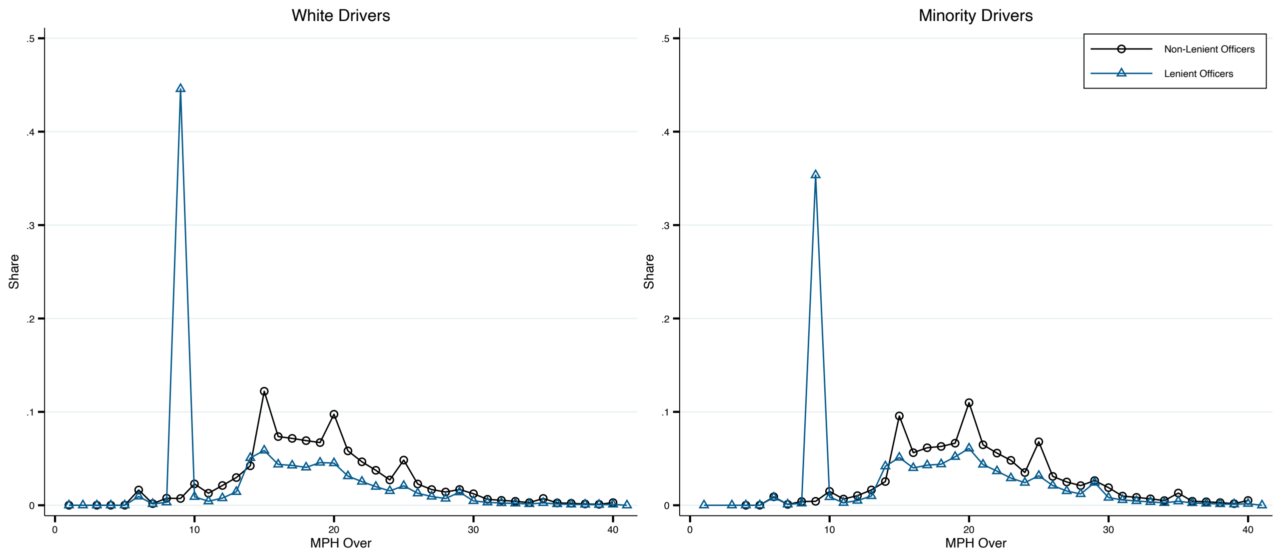
Notes: Connected line shows histogram of ticketed speeds, separately by driver race. 34.3% of tickets to white drivers are given at 9 MPH over compared to 25.2% of tickets for minority drivers.

Figure 3: Evidence of Officer Lenience



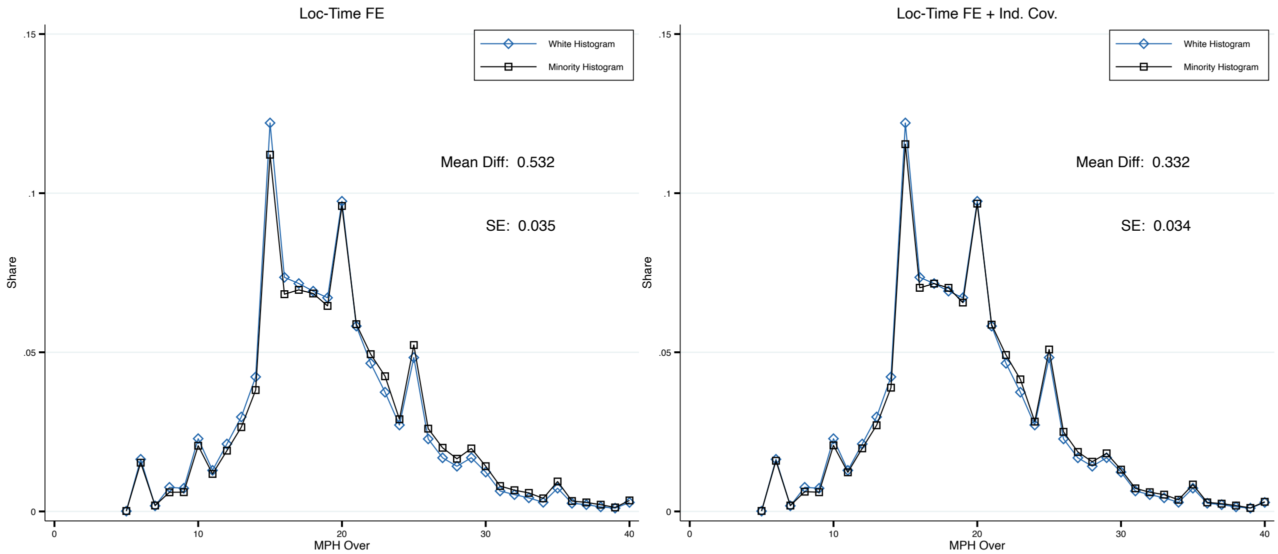
Notes: Panel A plots the across-officer distribution of lenience, calculated as the share of tickets given for 9 MPH over the limit. Panel B plots the share of tickets in each county and shift that is written by officers for whom fewer than 2% of their tickets are for 9 MPH over. Panel C plots officers' residualized lenience in the years with the most and second most citations. Panel D plots the residualized lenience in the county with the most and second most citations. Estimates residualized by conditioning on county fixed effects, speed zone fixed effects, year and month fixed effects, and day of week fixed effects.

Figure 4: Difference-in-Differences Raw Data Plot



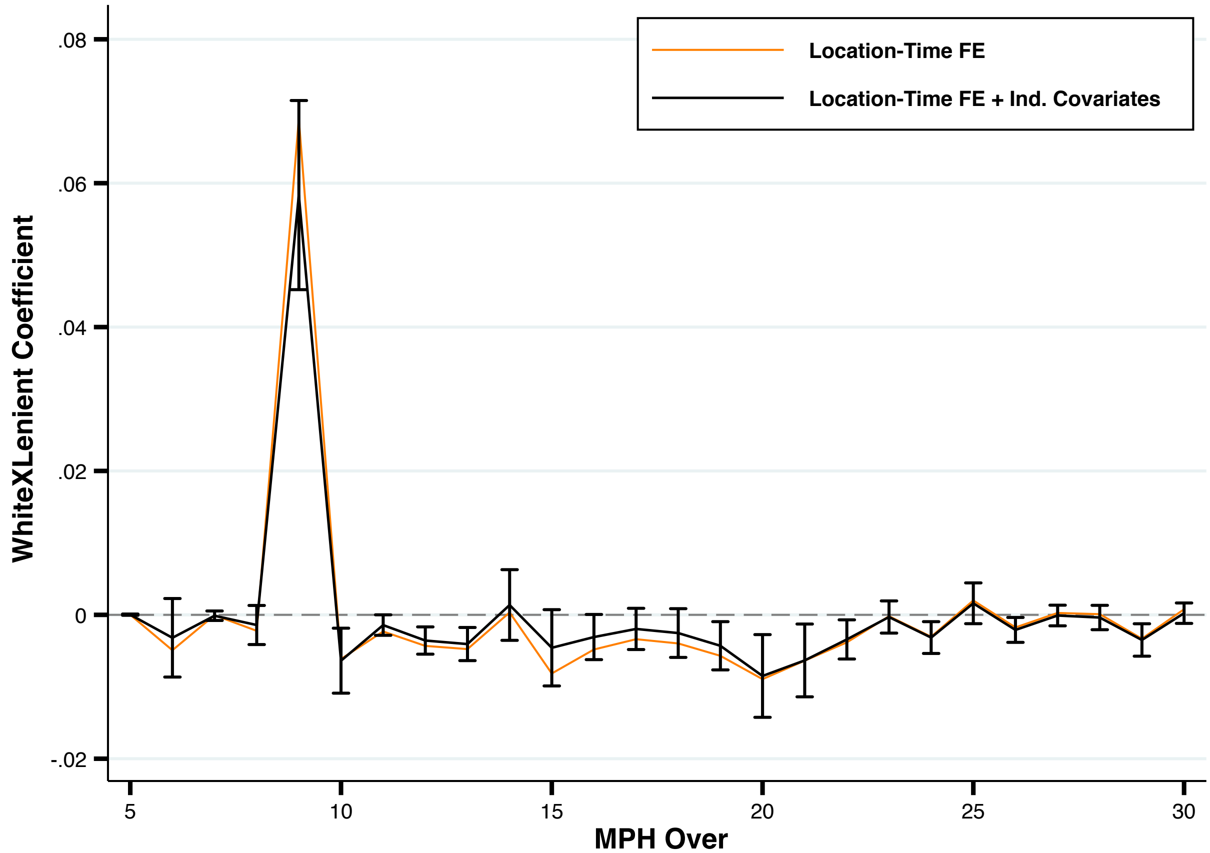
Notes: The left figure plots the histograms of speeds for white drivers, separately for stops made by lenient and non-lenient officers. The right figure plots the same histograms of speeds for minority drivers, separately by officer lenience.

Figure 5: Difference-in-Difference Raw Data Plot



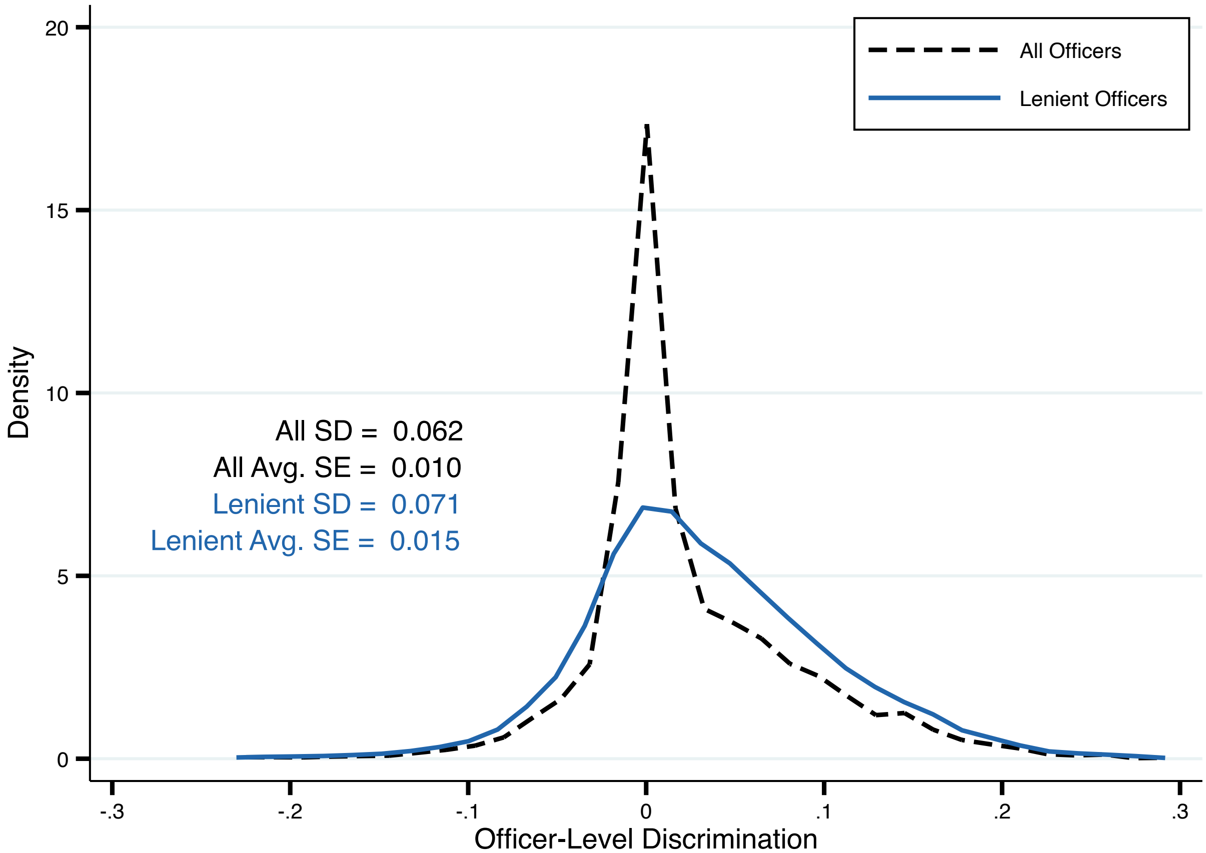
Notes: The left figure plots the histograms of speeds ticketed by non-lenient officers separately by race, where we have controlled for location and time fixed effects. Specifically, for each speed, we regress whether an individual is ticketed at that speed, controlling for minority driver and location and time fixed effects. The white histogram is the statewide distribution, and the minority histogram is the white histogram with the addition of minority regression coefficient for each speed. The right panel adds other driver demographic controls.

Figure 6: Difference-in-Difference Results



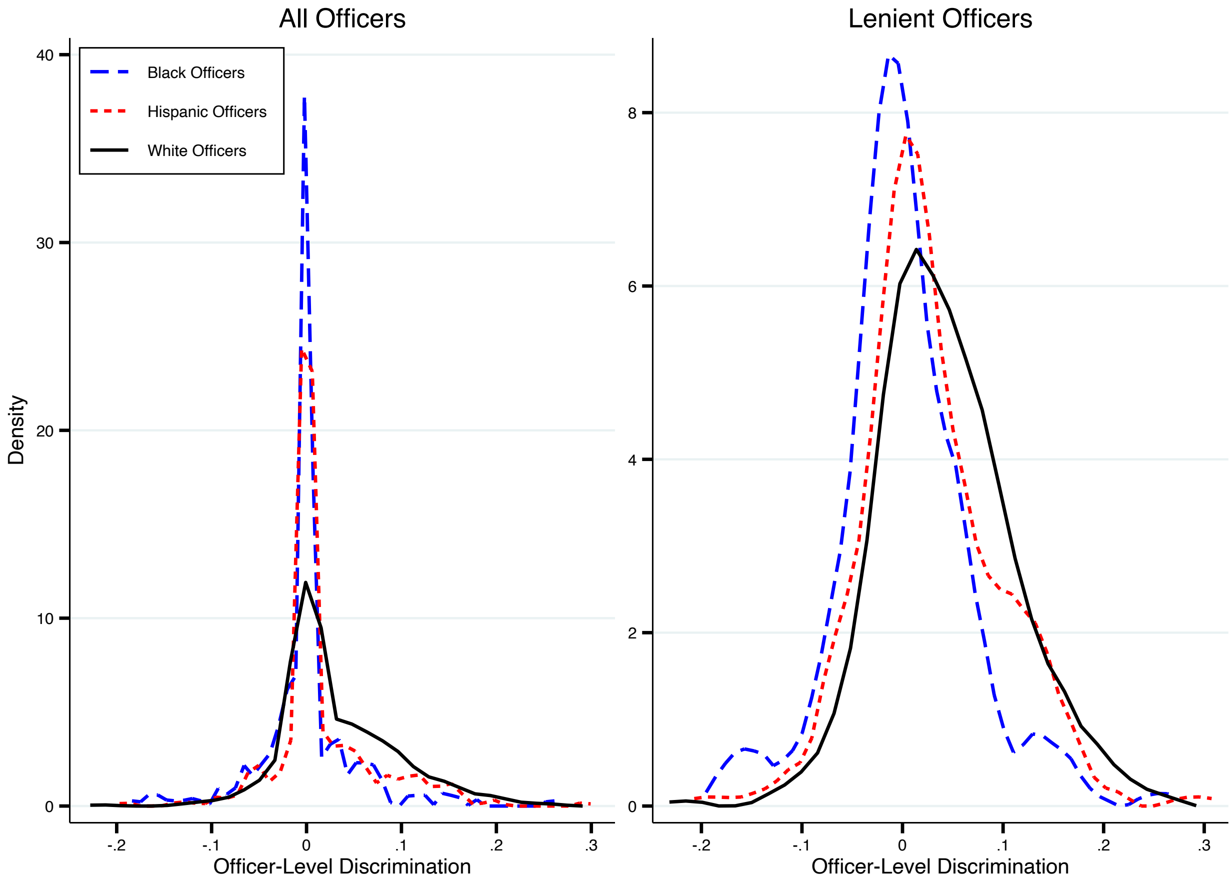
Notes: Figure plots the difference-in-difference regression results for each speed from Equation 3. The y-axis plots the interaction between driver being white and the officer being lenient. Standard errors are at the 5% level. The legend indicates the controls in each set of regressions.

Figure 7: Difference-in-Differences Officer-Level Results



Notes: Figure plots each officer's β_3^j from Equation 4. Officers who are non-lenient are assigned $\beta_3^j = 0$. SD reports the standard deviation across β_3^j, and Avg SE. reports the average standard error for each individual β_3^j.

Figure 8: Difference-in-Differences Officer-Level Results, by Officer Race



Notes: Left figure plots the discrimination coefficient β_3^j for all officers. Right figure plots the discrimination coefficient for all lenient officers.

Appendix: For Online Publication Only

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:

1. speeding is the primary citation (2,124,692 observations)
2. no crash is involved (2,123,311 observations, 99.9% of previous sample)
3. speed is between 0 and 40 over the limit (2,109,258, 99.3%)
4. posted speed limit is between 25MPH and 75MPH (2,107,933, 99.9%)
5. citations not from an airplane (2,103,923, 99.8%)
6. race/ethnicity is not missing (1,759,257, 83.6%)
7. race/ethnicity is white, black or Hispanic (1,671,089, 95.0%)
8. not missing driver's license state, gender, or age (1,667,558, 99.8%)
9. officer is identifiable (1,215,588, 72.9%)
10. officer has at least 100 tickets, and at least 20 for minorities and 20 for whites (1,174,284, 96.6%)
11. driver has no more than 20 citations in Florida for period 2005-2015 (1,141,628, 97.2%)

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 22% 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 (135,586 observations). We link the geocoded tickets to a Florida Department of Transportation roadmap shapefile using ArcGIS.[28] 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. 40% 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 2,124,692 speeding tickets in our data, 504,644 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

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

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 these directly from the hour of the ticket if it is properly recorded above. If there is no correct hour of day, we take a two-week moving average of the officer's modal shift for his citations and impute the shift. For the remaining tickets we leave shift as missing. Of the 1.6 million initial speeding citations, 692,416 have shift missing, and 413,560 remain missing after the imputation procedure.

A.4 Traffic Court Dispositions

Traffic court dispositions from the *TCATS* database were also shared by the Florida Clerk of Courts and matched to citations using citation identifiers. 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 (9.3%), 4 (38.2%), A (22.4%), and C (24.1%) account for over 99 percent of speeding tickets in our sample with disposition information (89% of citations). This makes sense given that the rest 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 percent 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, we feel 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 percent 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 percent. 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 4.1. Using the simple model from Section 3, 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 0.0077, (s.e.

0.0055). 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 13.2% of our baseline DD coefficient and statistically indistinguishable from zero.

One complication to note is that the lenient officers actually have a higher frequency of drivers ticketed at 14 MPH over the limit, inconsistent with the presumption that their probability of ticketing at that speed reflects the non-lenient distribution times the probability of not being discounted. We believe this inconsistency is due to the fact that there is a small amount of bunching down to 14 MPH over for lenient officers, since 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. We also find that the estimate for the discount probability at 24 is a small and statistically insignificant negative value.

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 is to assign the probability that most increases our estimate of bias. Since empirically $f_w(14) > f_m(14)$ and $f_w(24) < f_m(24)$, the most conservative estimate is to allow $P_m(14) = 1$ and $P_m(24) = 0$. When we do so, the estimate of bias increases from 0.0077 to 0.0096 (s.e. 0.0061), which is 16.5% of our main estimate and still statistically insignificant.

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 0.022, which is 38.1% of our DD coefficient.

C Lenience Versus Discrimination

As noted in Section 4, 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 5.1 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 51.6%. 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 5. 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 5.1, 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.

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.[29] 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 black drivers and marginally less biased against Hispanic 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

[29] Specifically, we run $S^9_{ij} = \beta_0 + \beta_1 \cdot \text{Black}_i + \beta_2 \cdot \text{Hispanic}_i + \beta^j_3 \cdot \text{Lenient}_j + \beta^j_B \cdot \text{Black}_i \cdot \text{Lenient}_j + \beta^j_H \cdot \text{Hispanic}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$. We take -$\beta^j_B$ and -β^j_H to be our measures of discrimination against black and hispanic drivers, respectively.

worked – we run regressions of the following form:

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

where Y_i is an outcome of either receiving a civilian complaint or using force. District_i^k is an indicator for an officer ever working in District k in the years 2011-2016, 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 (25% change in discounting) correlates to 0.19 fewer civilian complaints and a 5.5% decreased likelihood of receiving any complaints. Similarly, a one SD increase in lenience is associated with 0.06 fewer incidents of force and 3% lower likelihood of any force. Black officers are less likely to engage in force, as are older officers. Female officers are less likely to receive complaints but just as likely as male officers to use force. Discrimination against minorities seems to be positively related to force and complaints, though the standard errors are too large to say conclusively.[30]

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

[30] 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.

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.

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 *et al.*, 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:

$$\begin{aligned} S^k_{ij} &= \beta_0 + \beta_1 \cdot \text{BlackDriver}_i + \beta_2 \cdot \text{HispanicDriver}_i + \beta_3 \cdot \text{Lenient} \\ &\quad + \beta_4 \cdot \text{BlackDriver}_i \cdot \text{Lenient}_j + \beta_5 \cdot \text{HispanicDriver}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij} \end{aligned} \tag{6}$$

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:

$$\begin{aligned} S_{ij}^k &= \beta_0 + \beta_1 \cdot \text{BlackDriver}_i + \beta_2 \cdot \text{HispanicDriver}_i + \beta_3^j \cdot \text{Lenient}_j \\ &\quad + \beta_4^j \cdot \text{BlackDriver}_i \cdot \text{Lenient}_j + \beta_5^j \cdot \text{HispanicDriver}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij} \end{aligned} \tag{7}$$

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 $32.6\% \pm 2.2$ of officers exhibit discrimination against black drivers and $36.4\% \pm 2.6$ against Hispanic drivers. Among lenient officers, these shares are $48.5\% \pm 3.2$ and $53.9\% \pm 3.5$, respectively.

The correlation in officer-level discrimination against the two groups is 0.3668 among lenient officers. The standard deviation across discrimination estimates are 0.081 and 0.091 for black and Hispanic discrimination, respectively. The average discrimination coefficient has a standard error of 0.018 and 0.020, respectively. Using these values to account for the fact that the correlation is across noisy coefficients, we estimate that the true correlation is 0.3855.

F Accounting for Stopping Margin Selection

As discussed in Section 6, 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:

$$\begin{aligned} D_{ij}^* &= \theta_j^W + \theta_j^M \cdot M_i + \epsilon_{ij} \\ Z_{ij} &= \alpha_j^W + \alpha_j^M \cdot M_i + \eta_{ij} \end{aligned}$$

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[31]:

$$\begin{aligned} \hat{\theta}_j^M &= E[D_{ij}^* | M_i = 1, Z_{ij} > 0] - E[D_{ij}^* | M_i = 0, Z_{ij} > 0] \\ &= \theta_j^M + E[\epsilon_{ij} | \eta_{ij} > -\alpha_j^W - \alpha_j^M] - E[\epsilon_{ij} | \eta_{ij} > -\alpha_j^M] \end{aligned}$$

If there's a difference in treatment in the first margin ($\alpha_j^M \neq 0$) and $corr(\epsilon_{ij}, \eta_{ij}) \neq 0$, then our estimate of θ_j^M will be inconsistent. In particular, if $\alpha_j^M > 0$ (discrimination in ticketing) and 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 $N_{rj} = N_r \cdot P(Z_{ij} = 1)$.

[31] 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.

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 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 7, 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),

$$\begin{aligned} S_{ij}^{14} &= \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2 \cdot \text{Lenient}_j^{14} \\ &\quad + \beta_3 \cdot \text{White}_i \cdot \text{Lenient}_j^{14} + \lambda^9 \hat{\epsilon}_{ij}^9 + X_{ij}\gamma + \epsilon_{ij}^{14} \end{aligned} \tag{8}$$

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,

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

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 6 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 8 and 9 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 9, 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 0.0038 and 0.0089 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 *et al.* (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[32] 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 844,422 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.[33] However, the quadratic in the second column is insignificant. Though not shown, a specification that includes a cubic in the propensity score also has insignificant higher terms.

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. However, when we add a quadratic term in Column 4, we 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.

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 nearly all drivers. If he is going to give a break to someone, that discount should give him a

[32]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).

[33]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.

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 *et al.* (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 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 $P(Discount \mid X = 10) < .02$ or $P(Discount \mid X = 10) > .98$) 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 0.0395. 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 0.0275 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). $Pr(Discount|E(Z), j, X = 10)$ 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 3.3 percentage points less likely to discount minorities than whites, off a baseline of 35.7% 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 5 when restricting our attention to non-lenient officers, model estimates suggest that minorities on average drive significantly faster than whites, on the order of 0.5 to 0.7 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 *et al.*, 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 5.

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 500 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:

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

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 7. 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 $\gamma^0_\pi = 0.5$. 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 0.0421 (v. 0.0395). In the average county, drivers who are stopped at 15 MPH over have a 37% chance of being discounted to 9 MPH over and a 3.5% 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	1.099	0.736	0.554	0.789	0.699
	(0.0738)	(0.0362)	(0.0335)	(0.0596)	(0.0666)
Driver Hispanic	2.775	0.804	0.604	0.953	0.709
	(0.126)	(0.0398)	(0.0364)	(0.0765)	(0.0714)
Driver Female			-0.575	-0.397	-0.345
			(0.0215)	(0.0431)	(0.0467)
Florida License			-0.271	-0.505	-0.377
			(0.0406)	(0.0740)	(0.0772)
Age			-43.57	-41.14	-36.87
			(1.246)	(1.650)	(1.790)
1 Prior Ticket			0.282	0.274	0.280
			(0.0167)	(0.0459)	(0.0534)
2+ Prior Ticket			0.789	0.668	0.721
			(0.0221)	(0.0601)	(0.0666)
Past Prison Spell			0.693	0.992	0.914
			(0.0741)	(0.181)	(0.228)
Log Zip Code Income			-0.135	-0.233	-0.182
			(0.0164)	(0.0289)	(0.0265)
Mean	16.404	16.453	16.453	15.934	16.098
Vehicle FE			X	X	X
Location + Time FE		X	X	X	
GPS FE					X
Observations	1141628	1096077	1096077	126841	102172

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	-0.0297	-0.0232	-0.0181	-0.0322	-0.0265
	(0.00572)	(0.00283)	(0.00260)	(0.00464)	(0.00473)
Driver Hispanic	-0.139	-0.0379	-0.0317	-0.0505	-0.0349
	(0.00861)	(0.00306)	(0.00275)	(0.00571)	(0.00545)
Driver Female			0.0262	0.0189	0.0183
			(0.00172)	(0.00330)	(0.00363)
Florida License			0.0143	0.0251	0.0163
			(0.00365)	(0.00555)	(0.00616)
Age			1.291	1.328	1.102
			(0.107)	(0.137)	(0.141)
1 Prior Ticket			-0.0113	-0.00875	-0.0110
			(0.00118)	(0.00320)	(0.00358)
2+ Prior Ticket			-0.0274	-0.0197	-0.0243
			(0.00191)	(0.00407)	(0.00422)
Past Prison Spell			-0.0269	-0.0375	-0.0386
			(0.00466)	(0.0120)	(0.0149)
Log Zip Code Income			0.00362	0.0106	0.00585
			(0.00122)	(0.00236)	(0.00222)
Mean	16.404	16.453	16.453	15.934	16.098
Vehicle FE			X	X	X
Location + Time FE		X	X	X	
GPS FE					X
Observations	1141628	1096077	1096077	126841	102172

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	1.402	0.642	0.460	0.702	0.544
	(0.115)	(0.0560)	(0.0524)	(0.125)	(0.128)
Driver Hispanic	1.890	0.459	0.276	0.348	0.260
	(0.161)	(0.0438)	(0.0425)	(0.0891)	(0.0940)
Driver Female			-0.505	-0.465	-0.391
			(0.0365)	(0.0740)	(0.0842)
Florida License			-0.281	-0.301	-0.206
			(0.0664)	(0.122)	(0.129)
Age			-44.48	-32.65	-33.67
			(1.497)	(2.452)	(2.840)
1 Prior Ticket			0.268	0.250	0.131
			(0.0315)	(0.0912)	(0.106)
2+ Prior Ticket			0.725	0.646	0.636
			(0.0398)	(0.112)	(0.113)
Log Zip Code Income			-0.102	-0.0991	-0.158
			(0.0194)	(0.0413)	(0.0479)
Mean	19.853	19.938	19.938	19.371	19.505
Vehicle FE			X	X	X
Location + Time FE		X	X	X	
GPS FE					X
Observations	292529	273321	273321	23281	18211

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: Officer Lenience Randomization Check

	Full Sample			GPS Sample	
	(1) Lenience	(2) Lenience	(3) Lenience	(4) Lenience	(5) Lenience
Driver Black	0.000635	0.00166	-0.000493	-0.00704	0.000622
	(0.0166)	(0.00292)	(0.00337)	(0.00549)	(0.00550)
Driver Hispanic	-0.0900	-0.00594	-0.00666	-0.0224	-0.00255
	(0.0287)	(0.00497)	(0.00462)	(0.0130)	(0.00338)
Driver Female	0.0188	0.00423	0.00251	0.00148	0.00163
	(0.00428)	(0.00217)	(0.00181)	(0.00132)	(0.00201)
Florida License	-0.0774	0.000511	0.000401	0.00769	-0.00196
	(0.0252)	(0.00346)	(0.00317)	(0.00806)	(0.00422)
Driver Age	-0.214	0.216	0.0744	0.134	0.0371
	(0.162)	(0.129)	(0.119)	(0.103)	(0.0741)
1 Prior Ticket	-0.0113	-0.000750	-0.000483	0.00244	0.0000311
	(0.00645)	(0.000896)	(0.000948)	(0.00144)	(0.00223)
2+ Prior Tickets	-0.0235	-0.00132	-0.0000563	0.00443	0.00212
	(0.0125)	(0.00145)	(0.00149)	(0.00225)	(0.00371)
Log Zip Code Income	-0.00361	0.00276	-0.00415	-0.000165	-0.000452
	(0.00855)	(0.00295)	(0.00195)	(0.00312)	(0.00211)
F-test	0	.616	.039	.419	.946
Mean	.31	.31	.309	.326	.33
Location FE		X			
Location + Time FE			X	X	
GPS FE					X
Observations	1139734	1139734	1077412	124916	135427

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 year by month by day of the week by shift fixed effects.

Table A.5: Diff-Diff Regressions by Whether Any Additional Citations Issued

	(1) Discount	(2) Discount	(3) Discount
Driver White	-0.0152	-0.0142	-0.0201
	(0.00509)	(0.00507)	(0.00650)
Officer Lenient	0.168	0.174	0.112
	(0.0305)	(0.0305)	(0.0504)
Driver White	0.0583	0.0559	0.0600
× Officer Lenient	(0.00671)	(0.00669)	(0.0105)
Specification	Baseline	No Other	Any Other
Spec2		Citation	Citation
Location + Time FE	X	X	X
Mean	.302	.308	.223
R2	.361	.364	.334
Observations	1096077	991225	79279

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.6: Difference-in-Differences Officer-Level Results

	Discrimination Percentile					
	(1) 10 %	(2) 25%	(3) 50%	(4) 75%	(5) 90%	(6) N
All Officers	-0.0113	0.0000	0.0053	0.0681	0.1275	1591
White Officers	-0.0076	0.0000	0.0231	0.0835	0.1386	1591
Black Officers	-0.0339	0.0000	0.0000	0.0228	0.0637	1591
Hispanic Officers	-0.0112	0.0000	0.0000	0.0404	0.1199	1591

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

Table A.7: Officer Discrimination Randomization Check

	Full Sample			GPS Sample	
	(1) Discrimination	(2) Disc	(3) Disc	(4) Disc	(5) Disc
Driver Black	0.000888	0.00152	0.000661	0.00233	0.00000407
	(0.00190)	(0.000604)	(0.000553)	(0.00182)	(0.000880)
Driver Hispanic	-0.00617	0.000164	-0.000751	-0.000986	-0.000444
	(0.00319)	(0.000849)	(0.000668)	(0.00258)	(0.000699)
Driver Female	0.00124	-0.000176	-0.000154	0.000703	0.000367
	(0.000615)	(0.000181)	(0.000202)	(0.000523)	(0.000431)
Florida License	-0.0111	-0.000238	-0.000563	-0.0000257	0.0000737
	(0.00304)	(0.000715)	(0.000657)	(0.00227)	(0.000606)
Driver Age	-0.0139	-0.00259	-0.000300	-0.0122	-0.00549
	(0.0274)	(0.0169)	(0.0156)	(0.0253)	(0.0153)
1 Prior Ticket	-0.000769	0.0000154	0.0000520	0.000813	0.000519
	(0.000693)	(0.000174)	(0.000178)	(0.000773)	(0.000515)
2+ Prior	-0.00198	0.0000737	0.000178	0.00114	0.000326
Tickets	(0.00144)	(0.000223)	(0.000215)	(0.000886)	(0.000511)
Log Zip Code	0.00115	0.00150	0.0000372	0.000848	0.0000294
Income	(0.00128)	(0.000721)	(0.000377)	(0.00110)	(0.000641)
F-test	0	.175	.148	.221	.687
Mean	.305	.305	.304	.323	.323
Location FE		X			
Location + Time FE			X	X	
GPS FE					X
Observations	1141628	1141628	1079250	125040	125040

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	0.359	0.514	0.474	0.424	0.441
Age	34.979	47.667	41.755	39.692	39.812
White	0.588	0.649	0.632	0.569	0.588
Black	0.177	0.143	0.142	0.193	0.197
Hispanic	0.234	0.208	0.226	0.238	0.216

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 (IPUMS). 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.

Table A.9: Predicting Officer Discrimination

	(1) White Lenience	(2) Black Discrimination	(3) Hispanic Discrimination
Black Officer	-0.016	-0.015	-0.011
	(0.018)	(0.002)	(0.002)
Hispanic Officer	-0.026	-0.007	-0.009
	(0.017)	(0.002)	(0.002)
Other Race	-0.019	0.017	0.002
	(0.049)	(0.011)	(0.008)
Female Officer	-0.038	-0.004	-0.005
	(0.022)	(0.002)	(0.002)
Age (/10)	0.010	0.001	0.000
	(0.010)	(0.001)	(0.001)
Experience (/10)	0.105	0.001	-0.002
	(0.042)	(0.005)	(0.005)
Failed Entrance Exam	0.020	-0.002	-0.004
	(0.022)	(0.002)	(0.002)
Any College	-0.013	-0.001	-0.002
	(0.014)	(0.002)	(0.002)
Number of Complaints	-0.010	-0.000	-0.000
	(0.003)	(0.000)	(0.000)
Use of Force Incidents	-0.003	-0.000	-0.001
	(0.005)	(0.001)	(0.001)
Mean	.175	.02	.032
Observations	1,402	1,402	1,402
R2	.235	.085	.098

Notes: Robust standard errors in parentheses. Outcomes are derived from the regression $S_{ij}^{9} = \beta_0 + \beta_1 \cdot \text{Black}_i + \beta_2 \cdot \text{Hispanic}_i + \beta_3^j \cdot \text{Lenient}_j + \beta_B^j \cdot \text{Black}_i \cdot \text{Lenient}_j + \beta_H^j \cdot \text{Hispanic}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$. White Lenience is calculated as $\beta_0 + \beta_3^j \text{Lenient}_j$. Black Bias and Hispanic Bias are calculated as $\beta_B^j \cdot \text{Lenient}_j$ and $\beta_H^j \cdot \text{Lenient}_j$, respectively. The sample of officers is reduced from 1591 to 1402 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	-0.622	-0.184	-0.184	-0.108
	(0.206)	(0.0546)	(0.152)	(0.0494)
Discrimination	-0.247	0.134	0.0642	-0.0698
	(0.574)	(0.192)	(0.433)	(0.166)
Black	0.111	0.00131	-0.196	-0.0863
	(0.176)	(0.0401)	(0.0908)	(0.0335)
Hispanic	-0.00981	0.0165	0.0309	0.00569
	(0.144)	(0.0372)	(0.0993)	(0.0373)
Other	0.178	0.0242	-0.234	-0.0727
	(0.380)	(0.0996)	(0.181)	(0.0938)
Female	-0.295	-0.109	-0.0149	0.0125
	(0.158)	(0.0496)	(0.104)	(0.0443)
Age	-0.120	0.163	-0.736	-0.194
	(0.332)	(0.0900)	(0.212)	(0.0793)
Age Squared	0.0167	-0.0249	0.0611	0.0138
	(0.0478)	(0.0131)	(0.0266)	(0.0108)
Experience	-0.0633	-0.0800	-0.554	-0.00694
	(0.414)	(0.130)	(0.331)	(0.117)
Exp Squared	-0.0209	0.0350	-0.00580	-0.00173
	(0.0766)	(0.0249)	(0.0457)	(0.0195)
Failed Entrance Exam	0.259	0.0432	-0.106	-0.00330
	(0.205)	(0.0483)	(0.110)	(0.0458)
Any College	-0.183	-0.0250	0.103	0.0134
	(0.104)	(0.0293)	(0.0946)	(0.0264)
Sought Promotion	-0.194	-0.0664	-0.0405	0.0239
	(0.113)	(0.0294)	(0.0884)	(0.0277)
Mean	1.26	.551	.559	.294
Observations	1402	1402	1402	1402
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	-0.0152	-0.0167	-0.0140	-0.0135	-0.000662
	(0.00509)	(0.00547)	(0.00518)	(0.00727)	(0.00514)
Officer Lenient	0.168	-0.0339	0.179	0.174	0.181
	(0.0305)	(0.0691)	(0.0296)	(0.0237)	(0.0503)
Driver White	0.0583	0.0589	0.0574	0.0616	0.0477
× Officer Lenient	(0.00671)	(0.00761)	(0.00681)	(0.00858)	(0.00809)
Specification	Baseline	Young	All Others	Nighttime	Daytime
Spec2		Men			
Location + Time FE	X	X	X	X	X
Mean	.302	.256	.313	.332	.301
R2	.361	.317	.372	.352	.307
Observations	1096077	256605	799016	473915	152398

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	0.00727	0.00263	-0.00128	-0.00442
	(0.00358)	(0.00334)	(0.00505)	(0.00343)
Driver Hispanic	0.0262	0.0217	0.00687	0.00383
	(0.00731)	(0.00725)	(0.00548)	(0.00337)
Officer Lenient	0.288	0.224	0.179	0.105
	(0.0195)	(0.0334)	(0.0316)	(0.0287)
Driver Black	-0.0399	-0.0303	-0.0415	-0.0298
× Officer Lenient	(0.00536)	(0.00499)	(0.00789)	(0.00673)
Driver Hispanic	-0.0921	-0.0801	-0.0770	-0.0540
× Officer Lenient	(0.00976)	(0.00969)	(0.00944)	(0.00790)
Mean	.302	.302	.314	.299
Covariates		X	X	X
Location + Time FE	X	X	X	
GPS FE				X
Observations	1096077	1096077	126841	102172

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	0.00142	-0.000417
	(0.00118)	(0.000581)
Officer Lenient at 14	0.0606	
	(0.0262)	
Driver White × Officer Lenient at 14	0.0411	
	(0.00783)	
$\hat{e}^9\_i$, Discount to 9 Residual	0.00256	-0.0150
	(0.0147)	(0.00760)
Officer Lenient at 29		0.0947
		(0.00761)
Driver White × Officer Lenient at 29		-0.00918
		(0.00168)
$\hat{e}^{14}\_i$, Discount to 14 Residual		-0.0783
		(0.0216)
Mean	0.063	0.031
Sample	$\geq$ 10 MPH	$\geq$ 15 MPH
Observations	734250	638073

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	0.0395	—	1	—	—	—	—
	(0.0006)						
t, officer valuations	-0.2824	4.5876	1591	-0.3099	4.2300	1591	0.0275
	(0.0031)	(0.1627)		(0.0035)	(0.1500)		(0.0046)
λ, speeds	20.5058	2.7202	67	20.9833	2.1300	67	-0.4775
	(0.0517)	(0.4735)		(0.0407)	(0.3708)		(0.0658)
Pr($Discount^9$ | S = 15, E(Z))	0.3383	0.1168	1591	0.3180	0.1087	1591	0.0204
	(0.0007)	(0.0000)		(0.0008)	0.0000		(0.0010)

	Speed Parameters γ		Preference Parameters α	
	(1)	(2)	(3)	(4)
Female	-0.4813	(0.0087)	0.1353	(0.0036)
Age	-0.0453	(0.0003)	0.0057	(0.0001)
Previous Tickets	0.1868	(0.0027)	-0.0388	(0.0013)
County Minority Share			-1.8714	(0.0270)

Notes: This table presents estimates of the model introduced in section 7. 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. $Pr(Discount \mid E(Z), j) = \Phi(t_{rj} + E(Z)\alpha - 10b)$, i.e. the probability of being discounted when driving right above the bunch point 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	15.531	17.296	1.764	100
	(0.009)	(0.011)	(0.014)	
No Discrimination	15.530	17.087	1.557	88.244
	(0.008)	(0.012)	(0.013)	(0.014)
No Sorting	15.645	17.166	1.521	86.193
	(0.009)	(0.012)	(0.015)	(0.015)
Neither	15.644	16.927	1.283	15.644
	(0.009)	(0.012)	(0.014)	0.013

	County-Level Disparity			
	(1) White Mean (MPH)	(2) Minority Mean	(3) Difference	(4) Percent
Baseline	15.531	16.194	0.662	100
	(0.009)	(0.011)	(0.014)	(NaN)
No Discrimination	15.530	15.967	0.436	65.868
	(0.008)	(0.012)	(0.013)	(0.022)
No Sorting	15.645	16.341	0.695	104.980
	(0.009)	(0.012)	(0.015)	(0.027)
Neither	15.644	16.106	0.462	69.714
	(0.009)	(0.012)	(0.014)	(0.024)

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.

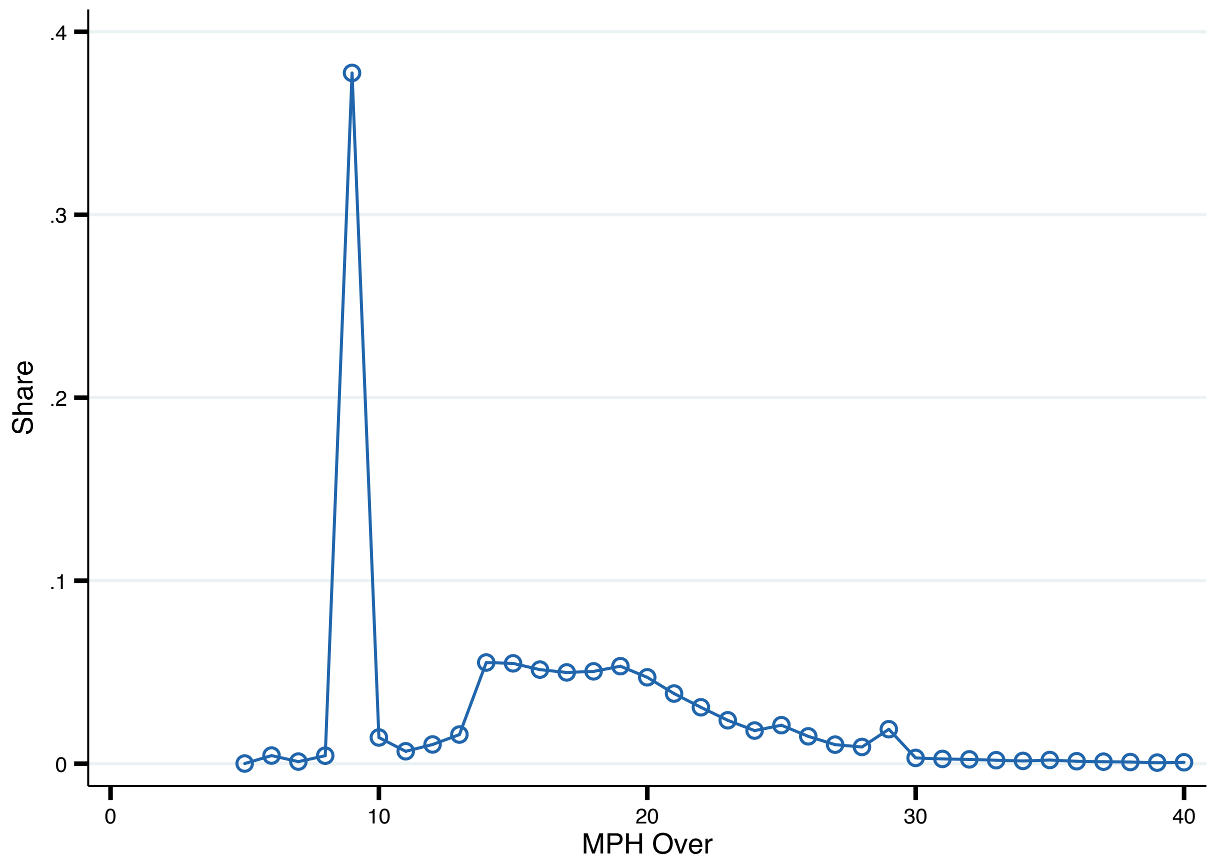
Table A.16: Alternative Model Parameter Estimates

	(1) μ	(2) σ^2	(3) # Param	(4) μ	(5) σ^2	(6) # Param	(7) Mean Diff
b, Slope	0.0421	0.0206	65	—	—	—	—
Heap Probability	0.0855	0.0027	65	—	—	—	—
c_{14}, 14 MPH Discount Cutoff	-0.2225	0.0767	65	—	—	—	—
Share Officers Lenient	0.6830	0.0715	65	—	—	—	—
	White Drivers			Minority Drivers			
t, Officer Valuations	0.2561	5.5875	65	0.1199	5.3932	65	0.1361
λ, Speeds	20.4923	3.3512	65	21.1354	3.0184	65	-0.6431
Pr($Discount^9$ | S = 15, E(Z))	0.3747	0.0687	65	0.3468	0.0627	65	0.0279
Pr($Discount^{14}$ | S = 15, E(Z))	0.0352	0.0014	65	0.0356	0.0015	65	-0.0004

	Speed Parameters γ			Preference Parameters α			
	(1) μ	(2) σ^2	# Param	(3) μ	(4) σ^2	# Param	
Female	-0.3609	0.1587	65	0.1324	0.0206	65	
Age	-0.0439	0.0002	65	0.0040	0.0000	65	
Previous Tickets	0.1636	0.0134	65	-0.0593	0.0184	65	

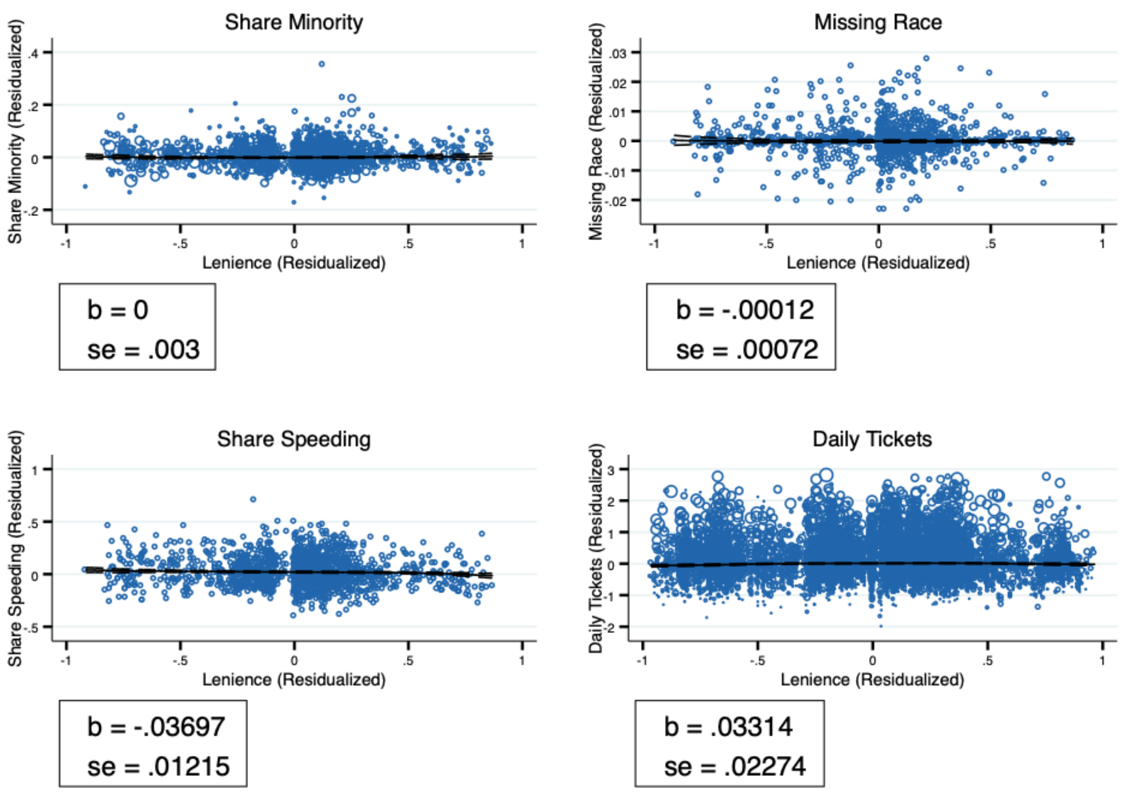
Notes: This table presents estimates of the model introduced in section 7. b is the slope parameter for how officers weight the speed of drivers in choosing to discount, t is each officer's mean valuation of a racial group in choosing to discount, and λ is the poisson speed parameter for each race by county. $Pr(Discount \mid E(Z), j) = \Phi(t_{rj} + E(Z)\alpha - 10b)$, i.e. the probability of being discounted when driving right above the bunch point for an average driver. The variances are empirical variances of the estimates, not adjusted for sampling error.

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



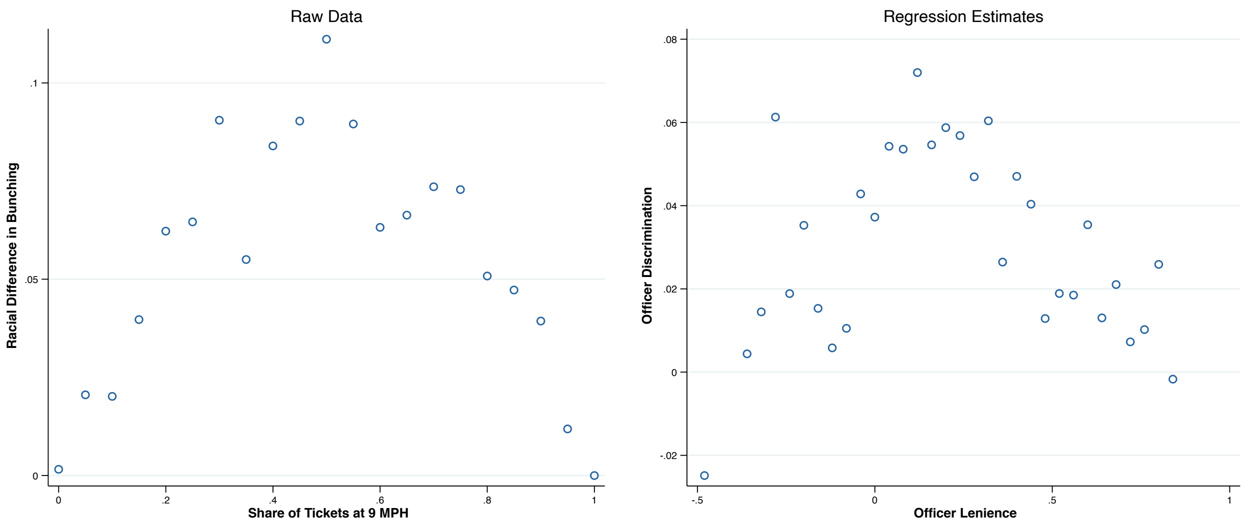
Notes: Line shows histogram of ticketed speeds for observations where the officer records that the speed is detected from a radar gun (N = 101,716).

Figure A.2: Officer Lenience and Stop Characteristics



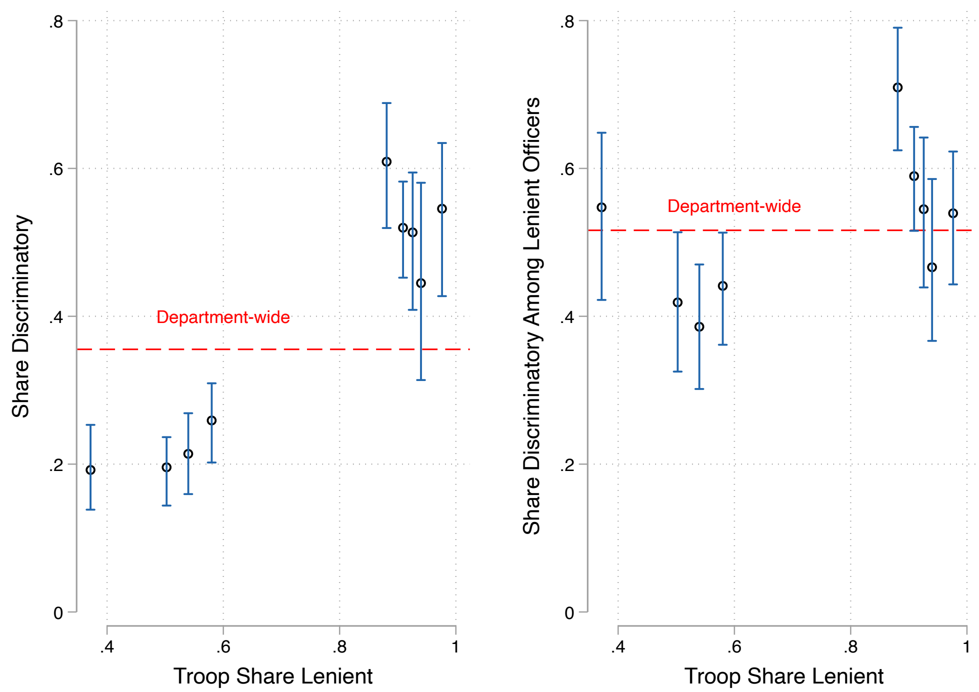
Notes: Figure plots the relationship between officer lenience and various characteristics of the officers' stops, where both officer's lenience and the stop characteristic have been residualized to remove location-time fixed effects. By officer lenience here we mean the indicator for whether an officer 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.

Figure A.3: Officer-Level Discrimination Against Lenience



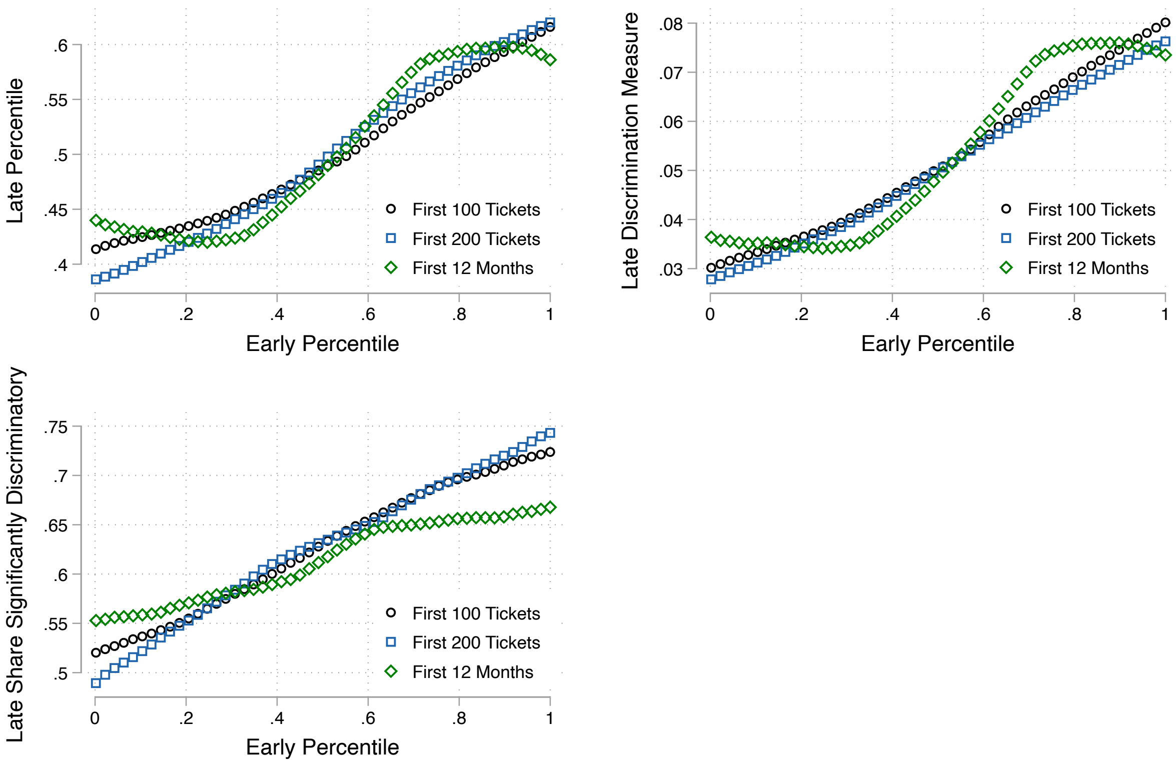
Notes: The left panel plots the officer-level racial differences in share of tickets issued at 9 MPH over (white minus minority) against the share of tickets issued at 9 MPH over for all individuals, which are binned to multiples of 0.05. The right panel plots the officer-level estimates of discrimination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



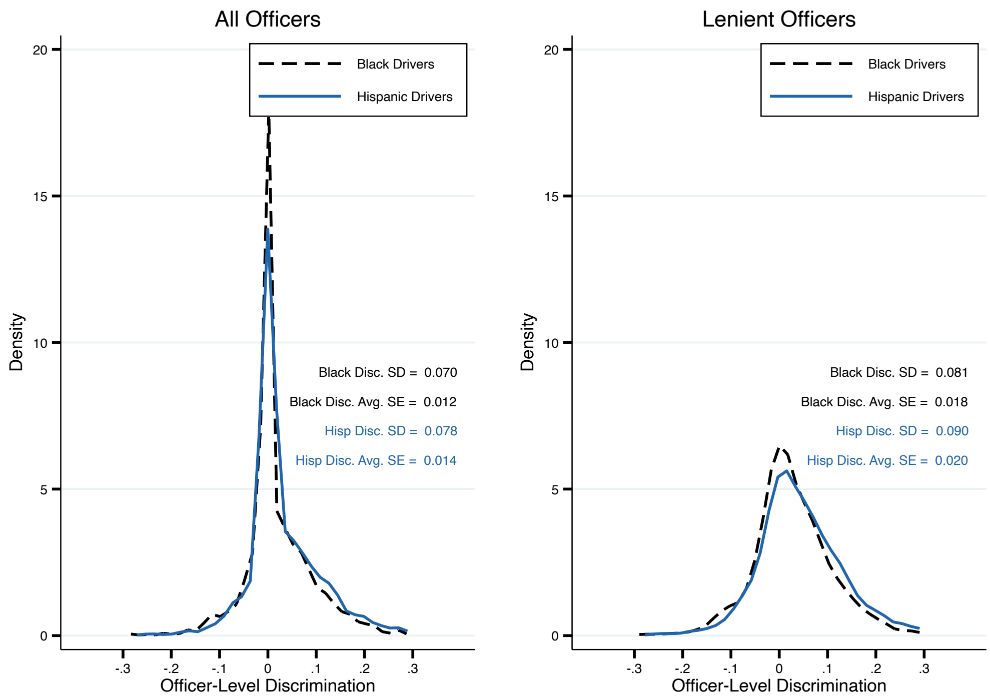
Notes: This figure plots troop-level estimates of the share of officers who are discriminatory against the share of each troop's officers who are lenient. The left panel calculates the share discriminatory among all officers, while the right panel calculates the share discriminatory among only lenient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

Figure A.5: Early and Late Discrimination



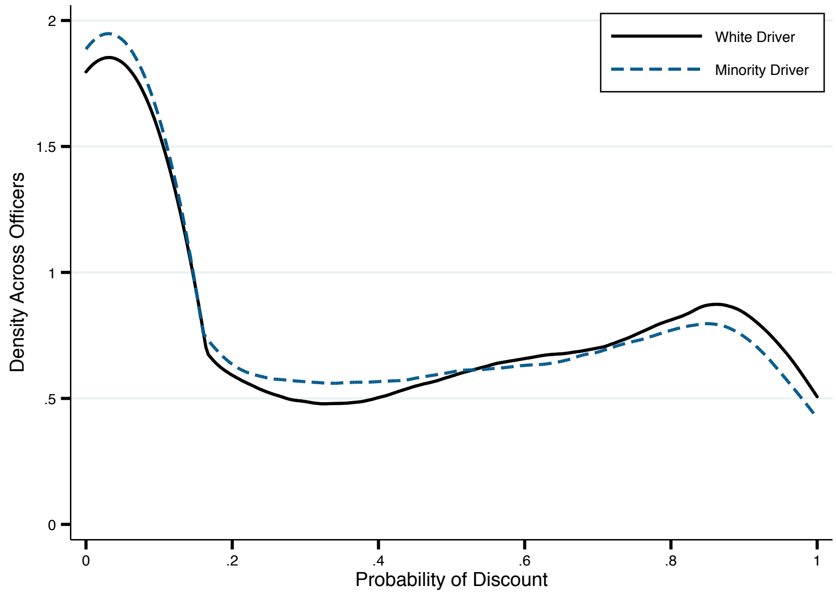
Notes: Figure corresponds to analysis in Appendix Section D.2. The top left panel regresses an officer's late sample discrimination percentile on early sample discrimination percentile. The top right panel regresses an officer's late sample measure of discrimination on their early sample discrimination 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



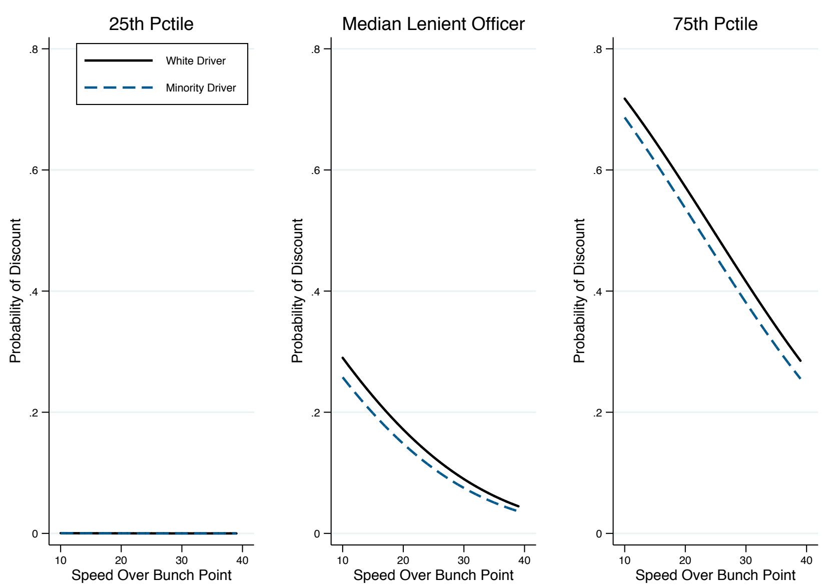
Notes: Figure plots each officer's $-\beta_4^j$ and $-\beta_5^j$ from the regression 7. Officers who are non-lenient are assigned $\beta_3^j = 0$ and are excluded from the right panel. SD reports the standard deviation across $-\beta_4^j$ and $-\beta_5^j$, and Avg SE. reports the average standard error for each individual $-\beta_4^j$ and $-\beta_5^j$.

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



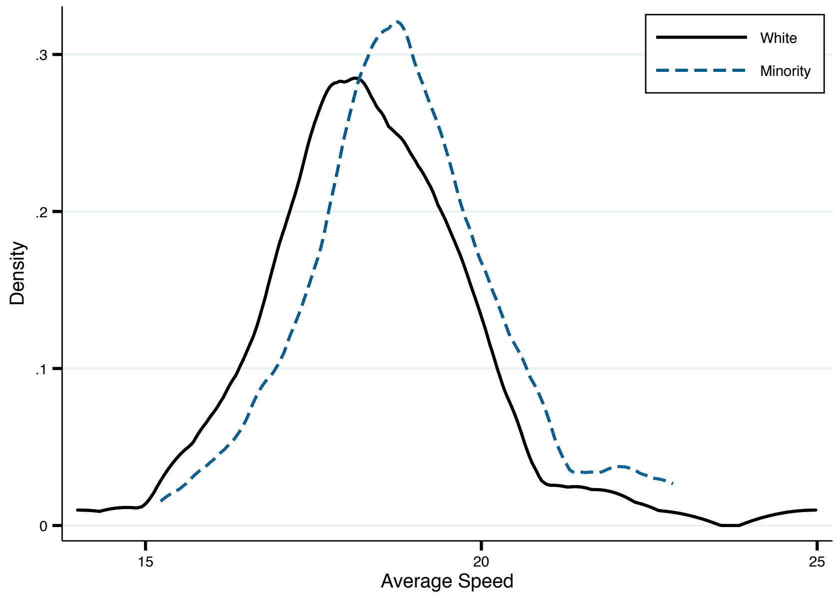
Notes: $P_{rj} \equiv P_j(\text{Discount}|X = 10, \text{ Driver Race} = r, Z = E(Z))$

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



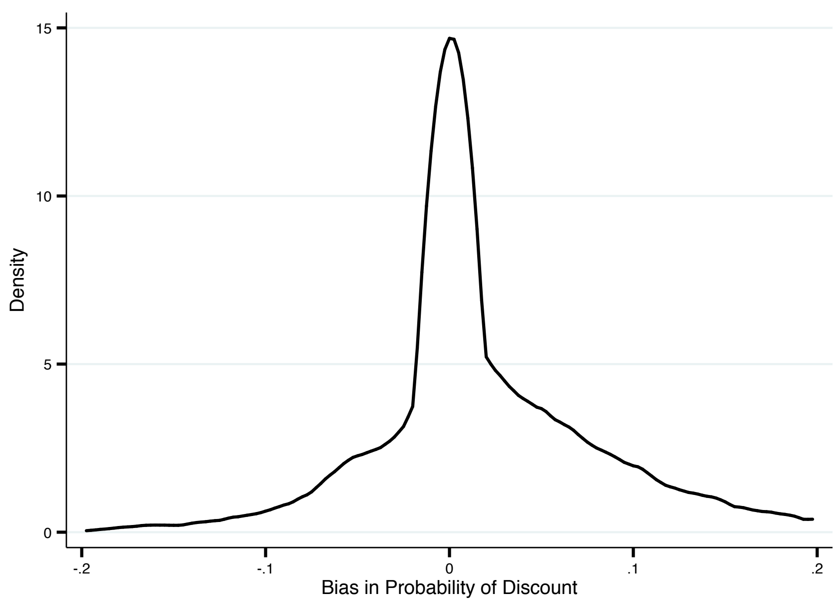
Notes: $P_{rj} \equiv P_j(\text{Discount}|X = 10, \text{ Driver Race} = r, Z = E(Z))$

Figure A.9: Model Estimates: Speed Distribution



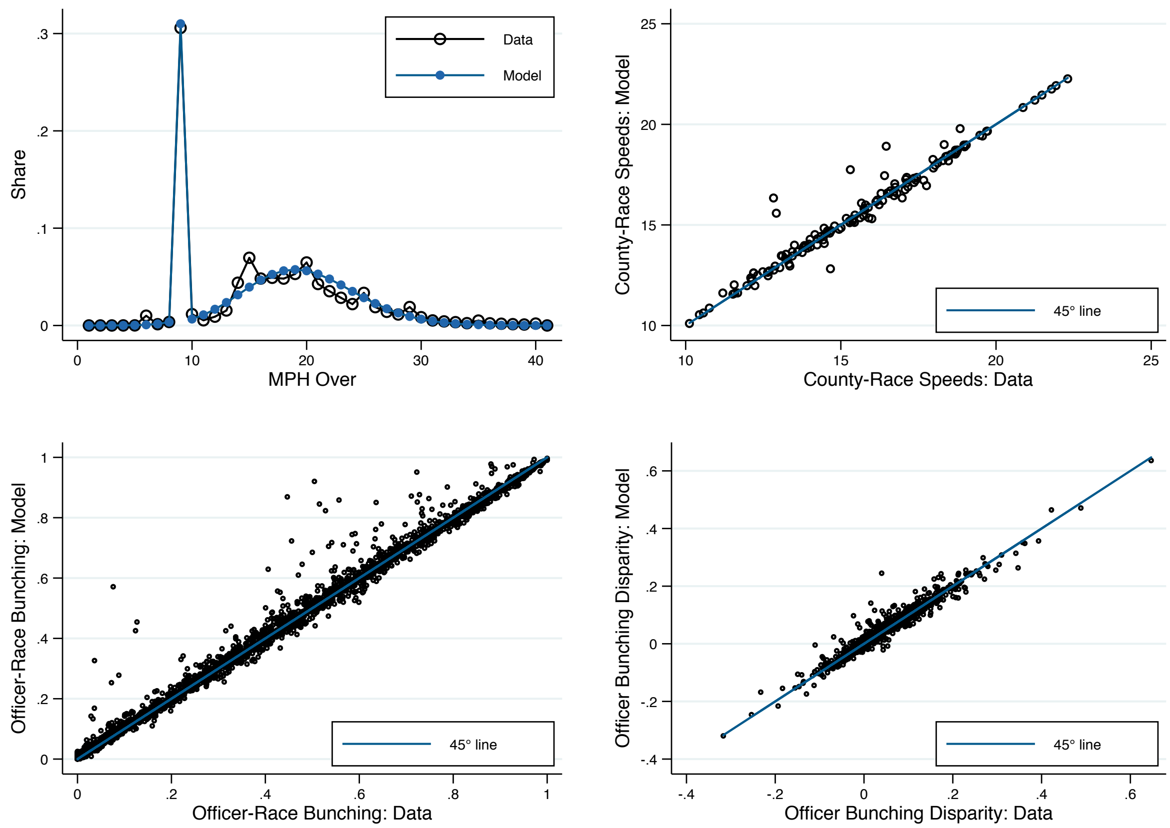
Notes: Figure plots the distribution of speed parameters λ across counties, separately by race of the driver, where individual covariates are set to the average value. In other words, we plot $\lambda = \lambda_{cr} + \gamma E(Z)$

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



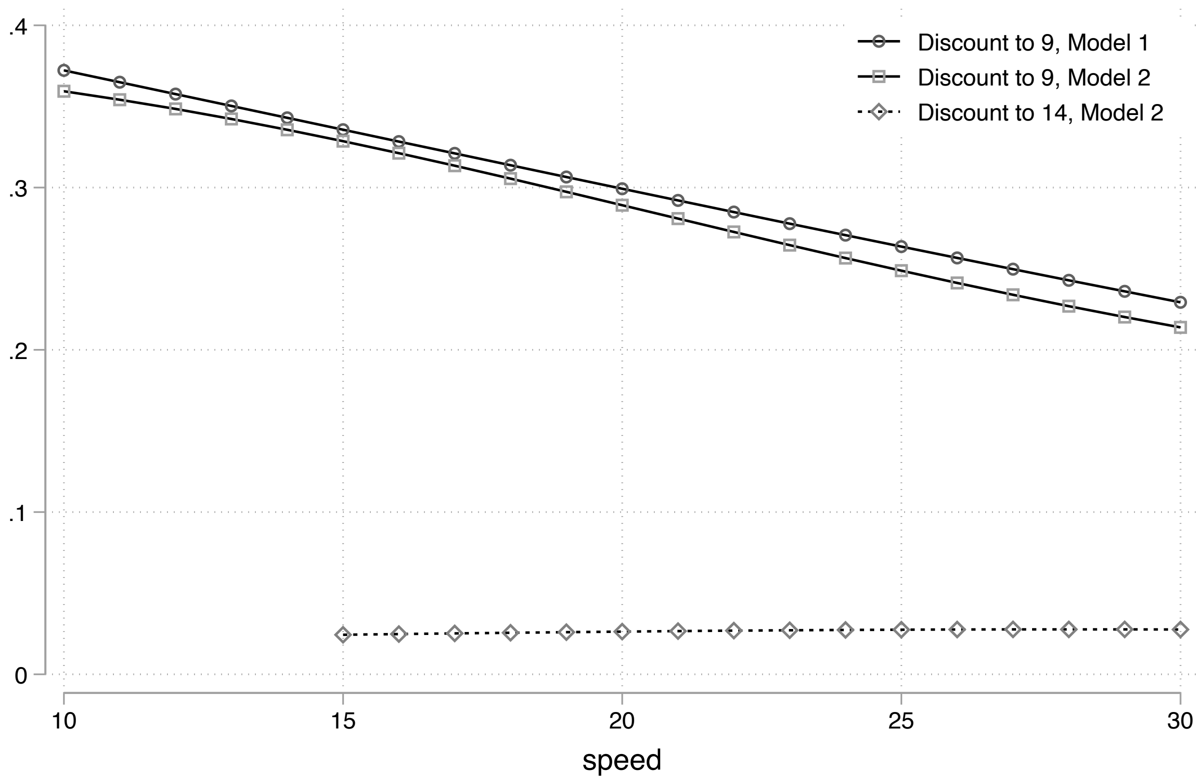
Notes: $P_j(\text{Discount}|X = 10, \text{Driver Race} = \text{White}, Z = E(Z)) - P_j(\text{Discount}|X = 10, \text{Driver Race} = \text{Minority}, Z = E(Z))$

Figure A.11: Model Diagnostic Figures



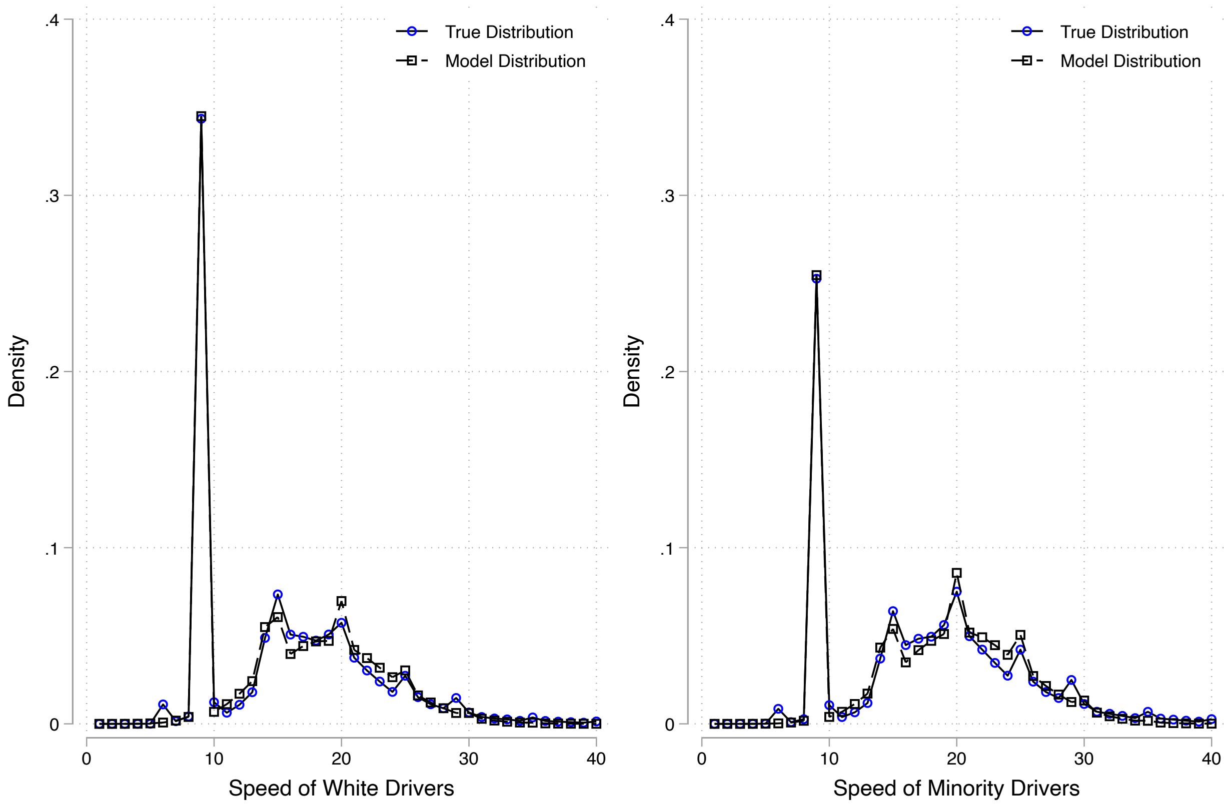
Notes: Figures compare various model estimates with their counterparts in the true data. Model estimates are found by simulating 100 iterations of the model and calculating averages across iterations. The top left panel compares the aggregate histograms of speeds. The top right panel compares the average 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.

Figure A.12: Comparison of Models



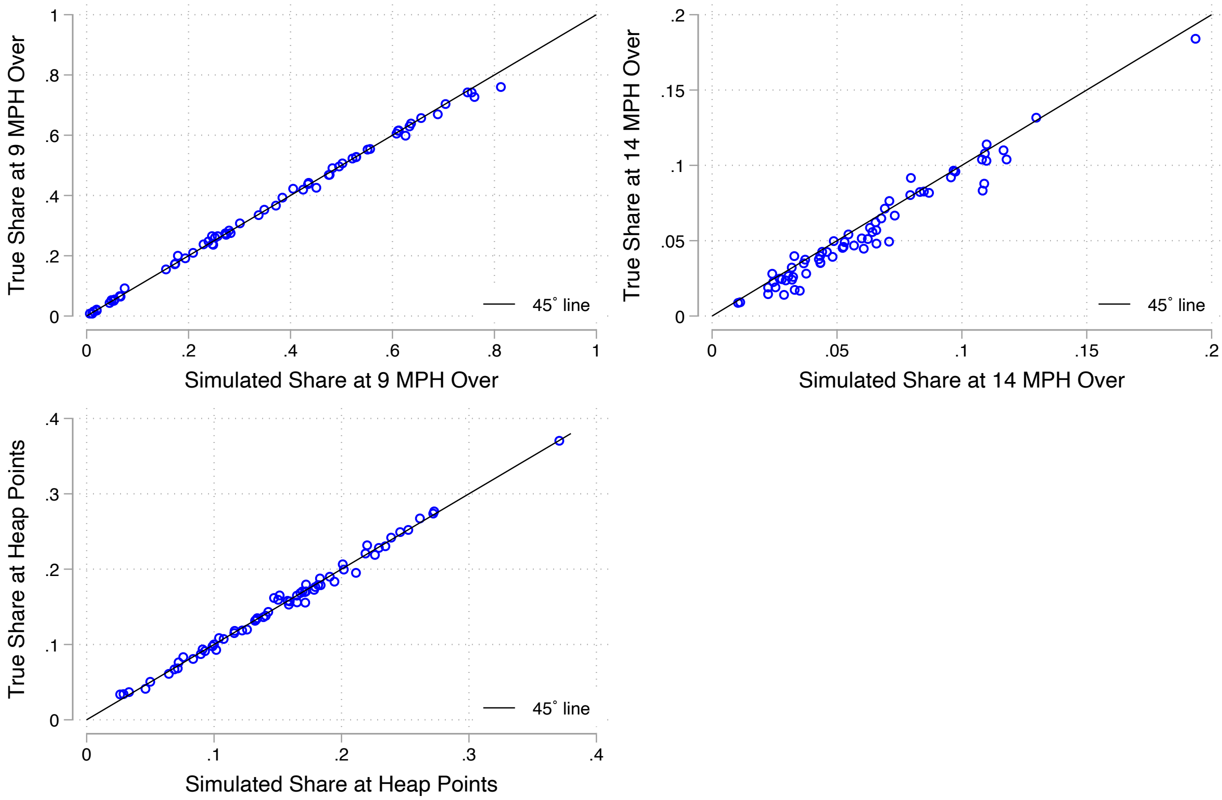
Notes: Figure plots estimates of the probability of discounting to 9 MPH over and 14 MPH over from every potential stopped speed, derived from the models in Section 7 and Appendix Section I. The estimates are calculated for the average driver across the full sample.

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.