A Few Bad Apples? Racial Bias in Policing*

Felipe Goncalves Steven Mello †

JOB MARKET PAPER

January 9, 2018

For Latest Version, Click Here:
http://www.princeton.edu/~fmg/JMP

Abstract

We estimate the degree to which individual police officers practice racial discrimination. Traffic police regularly discount the charged speed on drivers' tickets to avoid a discrete jump in the fine schedule. This behavior leads to an excess mass in the distribution of charged speeds just below the jump. Using a bunching estimation design and data from the Florida Highway Patrol, we show that minorities are less likely to receive this break than white drivers. We disaggregate to the individual police officer level and find significant heterogeneity across officers in their degree of discrimination, with 40% of officers explaining the entirety of the aggregate discrimination. Our measure of discrimination is easy to calculate and can be used by police departments as part of an early warning system. Using a simple personnel policy that reassigns officers across locations based on their lenience, departments can effectively reduce the aggregate disparity in treatment.

JEL Classification: J71, 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 Peter Bergman, Leah Boustan, Jessica Brown, Nicholas Buchholz, Janet Currie, Rebecca Diamond, Nik Engbom, Kirill Evdokimov, Hank Farber, Jeremy Fox, Sara Heller, Nathaniel Hendren, Daniel Herbst, Bo Honoré, Sierra Kuzava, Andrew Langan, Michael Luca, Neale Mahoney, Michael Makowsky, Michael Mueller-Smith, Christopher Neilson, Emily Owens, Aurelie Ouss, Jakob Schlockermann, Petra Todd, and participants of the Hamilton-Colgate Economics Seminar, the NBER Summer Institute Crime Session, the Transatlantic Conference on the Economics of Crime, and various Princeton 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.

†Industrial Relations Section, Princeton University, Princeton, NJ 08544. Goncalves: fmg@princeton.edu. Mello: smello@princeton.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., 2017), and issued a longer 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, 2016; Arnold et al., 2017). The view that discrimination is responsible 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 individuals. 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 setting 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. Although officers typically observe a driver's speed via radar before stopping them, they are free to choose what speed to charge. 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, lenience is relative, while in our setting officers make an explicit decision to reduce a driver's speed. 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 nonwhites and whites.

This paper proceeds in three parts. First, we document that officers are on average discriminatory against minority drivers. Second, we estimate the degree of discrimination of each individual officer. Third, we apply our officer-level measures of discrimination to various policy-relevant questions.

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. 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

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

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 discrimination. 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. 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 tickets.[4] For these officers, we argue that their distribution of ticketed speeds reflects the true distribution of driven speeds among stopped drivers. We show that, conditional on location and time, driver characteristics are not predictive of the average lenience of the officer he encounters. 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.

Using a differences-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 6 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 12 to 25 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 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 disparity. Correlating

[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.

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

We then show that a police department could feasibly use our approach to identify discriminatory officers early in their careers. We construct our measure of each officer's discrimination using only his first 100 tickets and show that this early measure is closely correlated with the full-sample estimate of her discrimination. An officer in the top 2% of discrimination in the early measure is on average at the 8th percentile of discrimination in our full-sample estimate, suggesting that a department can quickly identify the worst offending 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 87% of the gap in discounting. Only 13% 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 the central focus of the paper is not to differentiate between taste-based discrimination (Becker, 1957) and statistical discrimination (Arrow, 1973; Phelps, 1972), several pieces of evidence suggest that the discrimination we observe is taste-based. First, our setting is not as conducive to statistical discrimination as other criminal justice interactions.

In a speeding stop, the officer is aware of the crime committed (i.e., the speed driven) and does not need to use race as a signal of criminality. This knowledge contrasts with cases such as vehicle searches or stop and frisk, where the officer may use demographics to infer whether an individual is carrying contraband. Further, we see a full distribution of discrimination, which is not consistent with officers uniformly inferring something about the driver from their race (Anwar and Fang, 2006). The fact that minority and female officers are less discriminatory on average suggests that the discrimination we observe is a function of preferences rather than statistical inference. We also provide evidence that officers are not statistically discriminating on the basis of whether drivers will contest their ticket, nor on whether they are deterred from future speeding by getting the full ticket. Therefore, for the remainder of the paper, we use *discrimination* and *bias* interchangeably.

This paper contributes to a growing literature on methods of testing for the presence of discrimination in criminal justice and beyond. Popular approaches include audit studies that vary individual race (Bertrand and Mullainathan, 2004; Edelman et al., 2017; Agan and Starr, 2016), studies that vary the observability of race (Goldin and Rouse, 2000; Grogger and Ridgeway, 2006; Donohue, 2014), and studies of settings with rich controls for underlying behavior and context (Fryer, 2016). A prevailing approach to testing for bias is the "hit rate test," pioneered by Becker (1957). This test posits that a treator who is maximizing some objective regardless of race will have equal marginal success regardless of the race of the treated. Knowles et al. (2001) apply this logic to police motor vehicle searches and argue that a race-blind police force should be equally successful in finding contraband in white and minority vehicles. The strength of such an approach is the ability to differentiate between bias and statistical discrimination. Relative to our approach, the hit rate test is less applicable to estimating individual-level discrimination, as the success rates of most policing activities are hard to measure well at the individual level.[5]

Another popular set of methods for detecting racial bias are *benchmarking* procedures, whereby the behavior of one agent is compared to a proposed control group.[6] Ridgeway and MacDonald (2009) compare the racial makeup of NYPD officers' stop and frisks to those of nearby officers and are able to identify a small number of officers with a disproportionately high share of minority stops. Price and Wolfers (2010) compare the

[5]For an example of the hit-rate test outside of criminal justice, see Chandra and Staiger (2010).
[6]See Ridgeway and MacDonald (2010) for a review of the benchmarking literature.

racial makeup of fouls given by different NBA referees and find that the share of black fouls given by white referees is higher than for black referees and thereby reject that all referees are monolothic and unbiased in their behavior. Abrams et al. (2012) compare the treatment of race across judges and find significant heterogeneity, suggesting that at least some judges are discriminatory. In a paper closely related to ours, Anbarci and Lee (2014) study the discounting behavior of officers and find that the racial makeup of discounted tickets is whiter for white officers than for minority officers, suggesting that at least one group is biased in favor of their own race. Our approach broadly falls into the benchmarking literature, as we use the set of non-lenient officers as a benchmark for the behavior of other officers.

This paper also falls into a broad category of recent research using "bunching" estimators to recover behavioral parameters (Kleven, 2016). Predominantly used in the literature on taxation, these studies traditionally attempt to estimate the hypothetical distribution of interest in the absence of bunching by looking at the distribution outside a region around the manipulated area and inferring out-of-sample how the distribution should look at the discontinuity (Chetty et al., 2011; Saez, 2010). They then estimate bunching to be the difference between the true and hypothetical distribution around the bunch point. In contrast, 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 specification checks, and Section 6 discusses applications of our officer-level measures of discrimination. In Section 7, we present and estimate a model of officer behavior and perform counterfactuals. Section 8 considers alternative interpretations of our results, and Section 9 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.[7] 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.

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.[8] 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

[7] The actual fine schedule depends on the county in Florida, though the jump point is the same across all counties and always includes at least a $50 jump in fine.

[8] 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.

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. This fact has been confirmed by officers we spoke with, 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 traffic citations issued by the Florida Highway Patrol (FHP) 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, model, and year of the stopped automobile is provided, but this information is recorded inconsistently. In the final sample of citations, 69% of tickets list the vehicle make and year. The citing officer is identified by name, rank, troop number, and badge number.[9] While we see the speed charged by the officer, we do not see the original speed recorded by the officer. We also do not see stops and interactions that do not result in a traffic citation.[10]

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. It further includes 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

[9]The full data from the FCC contain all traffic citations for 2005-2015, including tickets not given by the highway patrol. We use these tickets to measure an individual's previous driving record. We do not use non-FHP tickets in our measures of bias, because officers are much harder to identify in these data. Further, many of the personnel information we collected is unique to the FHP.

[10]The problem of only seeing interactions that lead to enforcement is general in the discrimination literature. For a recent paper that addresses this issue, see West (2015).

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.

While the citations record the driver race, there appear to be inconsistencies in the recording of Hispanic. For example, Miami-Dade County issues less 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.

We restrict the sample to citations in which the main offense is speeding; 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 1,142,210 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 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). We link the geocoded tickets to a Florida Department of Transportation roadmap shapefile using ArcGIS.[11] 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

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

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. Throughout the analysis, we report results for the restricted sample of tickets with GPS with corresponding fixed effects at the road-segment level.

2.3 Summary Statistics

Table 1 presents summary statistics for the sample, broken out by driver race. 70% of drivers are white, 20% are black, and about 10% 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 0.04 times in the past year, though minorities have slightly 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. Minorities are also less likely to be ticketed for 9 MPH over the limit, which is just below the first jump in the fine schedule. 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.

In Table 2, we compare the racial distribution of speeding tickets with the racial distribution of residents and drivers in Florida using the 2006-2010 American Community Survey (ACS) 1% samples.[12] These data demonstrate that whites account for about 62.5% of Florida's population, and 60% of its drivers (an ACS respondent is considered a driver if they indicate that they drive to work), and about 58% of tickets. Blacks represent around 14% of the population and driving population, but 18% of tickets. Similarly, Hispanics are 20% of the population, almost 22% of the driving population, and 24% of tickets. In columns 4 and 5, we present the racial distribution of black, white, and Hispanic drivers involved in crashes and crashes with injuries over the 2006-2010 period. These

[12]We obtained these data from Integrated Public Use Microdata Series (IPUMS). So that the samples are parallel, we use only citations 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.

shares are computed from records provided by the Florida Division of Motorist Services that contain information on all auto accidents known to police. Relative to the citations data, Hispanics are slightly underrepresented in crashes, which may suggest that Hispanic drivers are targeted for citations relatively more often. Blacks are slightly overrepresented in crashes relative to citations, while white shares in citations and serious crashes are nearly identical. However, overall we do not have the impression that minorities are severely overrepresented or underrepresented in the tickets data relative to the population.

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 observed speed x' is drawn from some discrete distribution $F_r(\cdot)$, which can be a function of the driver's race r. We assume a simple discontinuous fine structure, in which the fine for speeding depends on the charged speed x according to

$$\text{Fine}(x) = \begin{cases} \pi_L & \text{if } x \leq x_d \\ \pi_H & \text{if } x > x_d \end{cases}$$

with $\pi_H > \pi_L$. If the driver's speed is above x_d, the officer has the choice to reduce the charged speed to x_d to reduce the fine the driver will face; otherwise, the speed is set to x'. 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 $P_j(x, r(i))$ of discounting the individual, which may be a function of the driver's race r and the driver's speed x.

In this framework, it is natural to define discrimination in the following way: We say that officer j is *discriminatory* if $P_j(x, r(i) = w) > P_j(x, r(i) = m)$ for a given speed x. 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. We discuss statistical discrimination in Section 8 and why we believe the observed discrimination in discounting is taste-based.

The first empirical step we take is to model the likelihood of an individual's appearing at the discount point and above, given his observables. In our model, 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:

$$\Pr(X_i = x_d | i, j) = F_{r(i)}(x_d) + \sum_{k > x_d} F_{r(i)}(k) \cdot P_j(x, r(i)) \tag{1}$$

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

$$\Pr(X_i = x > x_d | i, j) = F_{r(i)}(x) \cdot \Big(1 - P_j(x, r(i))\Big) \tag{2}$$

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 difference in likelihood of appearing at the bunch point and above can arise from either differences in speeds $F_r(x)$ or differences in speed-specific discounting, $P_j(x, r(i))$. 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 one-third 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. About 16% of officers write no tickets for exactly 9 MPH over, while 30% write less than 1% of tickets for this bunching speed. Panel B plots the distribution of officer lenience after residualizing county and month-of-stop fixed effects and driver characteristics. The observed disparity suggests

that the heterogeneity across officers is not due to differences in location or characteristics of the stopped drivers.

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 second-most stops (y-axis) against his residualized lenience in his year with the most stops (x-axis). 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 lenience in the county where the officer has made the second most stops against lenience in the county where he has made the most stops, confirming that officer lenience is highly correlated over space.

We treat the 33% of officers with fewer than 2% of their tickets issued at 9 MPH over as non-lenient officers, and we use these officers for two purposes.[13] First, we suppose that these officers' ticketing distribution reflects the true distribution of speeds within their location and shift and use them to uncover the true racial difference in speeding. Secondly, we use these officers as a control group in a difference-in-difference 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 of whether a driver is stopped at a given speed S, where the race of the driver is interacted with the lenience of the officer:

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

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.

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

[13] An alternative approach is to explicitly test for the presence of bunching officer-by-officer. When we do so using the Frandsen (2017) test, the set of officers identified as non-lenient remain very similar and the regression results do not change.

9 MPH over, but we allow the coefficients for Lenient_j and $\text{White}_i \cdot \text{Lenient}_j$ to vary by individual officer:

$$\begin{aligned} S^9_{ij} &= \beta_0 + \beta_1 \cdot \text{White}_i + \beta^j_2 \cdot \text{Lenient}_j \\ &\quad + \beta^j_3 \cdot \text{White}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij} \end{aligned}$$

The coefficients of interest, β^j_3, 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^j_3 as officer j's degree of discrimination. For the purpose of reporting the distribution of discrimination across officers, we treat non-lenient officers as having $\beta^j_3 = 0$, since by definition they cannot be discriminatory.

The intuition for our differences-in-differences procedure is shown in Figure 7. 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 shift 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. 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.

While we cannot directly check the assumption of conditional random assignment of officer, we perform various specification checks to test for randomization. Figure 4 shows that officer lenience is not predictive of share minority, implying that officers do not vary systematically in their lenience toward minorities in terms of whether to give a warning. Figure 5 shows that lenient officers have no more or fewer daily tickets than non-lenient officers, corroborating the assumption that lenient officers in discounting are

not more lenient in the margin of whether to write a ticket. Figure 6 confirms that officers of different lenience do not vary in their share of tickets written for speeding or share of tickets written for seatbelt violations.

To test for selection on observables, Table 3 estimates how officer lenience varies with driver characteristics. The F-tests report a joint test of the hypothesis that all driver characteristics have zero correlation with officer lenience. The first column reports results with no controls for location or time. Here officer lenience varies significantly with driver characteristics. Hispanic drivers and in-state-license drivers are ticketed in areas where officers are less lenient to everyone. Columns 2 and 3 restrict attention to variation within location and location plus time, respectively. With these controls, officer type varies much less significantly with driver characteristics. A joint F-test fails to reject at 5% significance that all driver characteristics are equal to zero, though specification (3) rejects at the 10% level. Columns (4) and (5) report results for our GPS'ed sample. Both with and without fixed effects for the road-segment of the stop, we find that officer lenience is uncorrelated with driver characteristics.

5 Results

Our first use of non-lenient officers is to test whether minorities truly drive faster than white drivers. Figure 8 reports the distribution of ticketed speeds for non-lenient officers. Unconditional on any covariates, minorities drive 1.5 MPH faster than whites. However, when controlling for county and individual covariates, this disparity shrinks to .39 MPH, and the disparity is barely perceptible visually. The majority of the reduction comes from accounting for county fixed effects, since minorities tend to drive in counties in which all drivers are stopped at faster speeds. The fact that the county-specific disparity is so small suggests that the racial disparity in discounting cannot be explained by differences in driving speed. In Appendix Figure A-3, we show that this small gap is consistent across various specifications for time and location controls.

Figure 9 and Table 4 report the results of the difference-in-difference test of discrimination. The figure reports regression coefficients from both a specification with no controls and our preferred specification with individual covariates and fixed effects for county by year by month by shift by highway. As indicated by the interaction variable for white

drivers and lenient officers encountered at 9 MPH over, white drivers are significantly more likely to receive a discount than minority drivers. Off a mean probability of 45%, white drivers stopped by lenient officers are encountered at the bunch point 6-8.4pp more often than minorities, and this disparity persists regardless of the specification. In Columns 4 and 5 of Table 4, we perform the same regression for the restricted sample with GPS ticket location, and the results continue when allowing for stretch-of-road fixed effects.

The interpretation of these coefficients tell us 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%.

The interaction coefficients for speeds above 9 MPH shown in Figure 9 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 all speeds between 12 and 20 MPH, suggesting that at these speeds minorities are less likely to receive a break.

A natural question to ask is how this estimate aggregates to a total cost of discrimination. There are about 590,000 Florida tickets given to speeding drivers for 9 MPH or greater, 130,000 of which are given to black and Hispanic drivers. The jump from 9MPH over to 10MPH over leads to $75 fine increase. Using our estimate that minority drivers are 4 percentage points less likely to be discounted, we calculate the cost of discrimination toward minority drivers to be $389,000 per year. Scaled up to the entire US population, that figure increases to $6.1 million.[14]

Officer-level results are reported in Figure 10 and Table 5. 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 probability difference in percentage points. An officer whose discrimination against minorities is 0.1,

[14]Florida's 2016 population is 20.6 million, and the US population is 323.1 million, so we multiply our figure by 323.1/20.6

for example, is 10 percentage points more likely to offer a fine reduction to a white than a minority driver.

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 a base of 30%. 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 .014 – less than one-fourth the standard deviation of $\hat{\beta}_3^j$ across officers, .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 noise. [15]

5.1 Share of Officers Who Are Discriminatory

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 discrim-

[15] 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.

ination measure is an additive function of his true discrimination plus estimation error, $\hat{\theta}_j = \theta_j + \epsilon_j$, where ϵ_j is asymptotically normally distributed and σ_j^2 is known from the model estimation. We can assume an officer's discrimination can take on a finite set of values on a fine grid, $\theta_j \in \{\theta^k\}$,[16] 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. Using this approach, and calculating $1 - \hat{F}(0)$ as the share, we find that 41% (CI 38.5-43.7%) of officers are discriminatory.[17] In contrast, we find that only 7% (CI 5.6-8.7%) of officers have $\theta_j < 0$, i.e., practice reverse discrimination.[18]

5.2 Additional Specification Check

In Section 4, we reported various specification checks for the randomization of officer lenience. An additional test for the random assignment of officer to driver is that officer *discrimination* is not correlated with driver characteristics. We report such regressions in Table 6. As before, Column 1 reports the regression with no controls, and the F-test indicates that some driver characteristics are correlated with officer discrimination, statewide. All other regressions, which include controls for county, report no relationship between officer discrimination and driver demographics.

6 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. The first insight we gain from this distribution is that discrimination varies greatly from officer to officer. However,

[16] The grid is 99 points spanning the 1st to 99th percentiles of the empirical distribution of $\hat{\theta}_j$.

[17] Confidence intervals are calculated through bootstrapping by performing 100 draws of the set $\{\hat{\theta}_j\}$ and performing MLE on each draw.

[18] This approach is a discretized version of a deconvolution procedure (Delaigle et al., 2008). Doing the continuous deconvolution leads to an identical estimate for the share of officers who are discriminatory.

estimating the degree of discrimination of individual officers allows us to address various previously unanswerable questions. How does discrimination vary by officer demographics? Are early measures of discrimination predictive of long-term discrimination? And which personnel policies can mitigate the effect of discrimination? We answer the first two questions in this section and the third in Section 7.

6.1 Do Officer Characteristics Predict Discrimination?

Given an officer-level measure of racial discrimination, a natural question is how it correlates with other officer characteristics and behaviors. We can tackle this question using the personnel records collected from the FDLE and the FHP.

Figure 11 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 officer are still, on average, discriminatory against minority drivers. Among black officers, a very small percentage are discriminatory in favor of minorities.

Some of the disparity in discrimination across officer race is driven by minority officers' being less likely to be lenient overall. However, this fact is due mostly to minority officers' working in troops in which all officers are less lenient. In Figure 12, we show the distribution of discrimination only for lenient officers. White officers' distribution continues to shift farther to the right.

The ability to identify discrimination separately by officer race is another 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.

In Table 8, 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. 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 drivers of their own race, as we might expect. Female officers appear less biased against both black and Hispanic drivers. Older officers exhibit more bias against both groups, though the standard errors are large. Officers with any higher education are less likely to be biased toward blacks but not Hispanics. Neither complaints nor seeking promotion have consistent predictive power on bias.

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. Specifically, we ask whether our measures of lenience and discrimination can be used to predict an officer's propensity to receive a civilian complaint or use force on the job. To make the analysis at the officer-level – but still account for differences in years and locations worked – we run regressions of the following form:

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

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 9, 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 .19 fewer civilian complaints and a 5.5% decreased likelihood of receiving any complaints. Similarly, a 1 SD increase in lenience is associated with .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.[19]

[19]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. By doing so, the coefficients on discrimination increase overall in magnitude, though the standard errors remain too large to definitively say whether there is a true relationship.

6.2 Are Our Estimates Usable by a Police Department?

We argued above 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 complicated measure we have constructed for each officer is actually usable by a department to identify discriminatory officers. Specifically, we ask whether an individual's discrimination, as calculated from his first 100 tickets, is close to his measure from the full sample. The median officer takes 400 calendar days to write that many tickets.

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 interaction, including 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.

$$D_j^{\text{early}} = \overline{\epsilon_{ij}^{\text{white}}} - \overline{\epsilon_{ij}^{\text{min}}}$$

We report in Table 10 the relationship between this early measure and our full-sample estimates of discrimination. We find that D_j^{early} has significant value for policy. Its correlation with our full measure β_3^j is .45. The top panel reports how the percentiles of the two distributions correspond. Among the 2% of officers with the most discrimination in our early measure, the median percentile in the full sample is 3.2. The 5% and 10% most discriminatory also mostly consist of officers who are discriminatory in the full sample. However, the 95th percentile of "early discrimination" officers are quite nondiscriminatory when calculated in the full sample. This fact implies that some officers who are discriminatory in their early patrolling grow out of this practice in later years.

This "mistake" in the early measure is confirmed in the bottom panel of Figure 10, which reports Type I and Type II error. Among the 398 (25%) officers whose early measures indicate discrimination with 95% confidence, 32.2% are found not to be discriminatory at 5% significance in the full sample. Restricting attention to officers whose z-statistic in the early measure exceeds 3 (99.8% confidence) barely reduces Type I error, to 28%. The stubborness of this error suggests that the early measures are somewhat incorrect – not because of imprecision, but because officers change in their ticketing practice past their

first year in policing. The Type II error column indicates the share of officers who, in the full sample, are found to be discriminatory at the 5% level but were not detected in the early measure. This number is greater than 50% in all columns, suggesting that early detection can catch no more than half of discriminatory officers.

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 they often identify officers who are non-discriminatory in the totality of their careers. An early warning system is also not a panacea, as it fails to identify more than 50% of officers who will practice discrimination in their later careers.

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. 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. By doing so, we can also perform counterfactuals that change the distribution of discrimination across officers. The model setup is as follows.

Officer j encounters individual i driving at a speed drawn from a Poisson distribution $x \sim P_{\lambda_i}(x)$, where the Poisson parameter depends on the driver's county and race, $\lambda_i = \lambda_{rc}$. The officer faces the choice to either charge the driver his measured speed x or, if the speed is above the jump in fine, discount the speed to x_d. He makes this decision by weighing a cost to discounting, which we impose to have the form $c(x) = b \cdot x$, against the value of discounting, $t_{ij} = t_{rj} + \epsilon_{ij}$. So the driver has her speed reduced to x_d if

$$t_{rj} + \epsilon_{ij} > a + bx_i$$

The officer's preference is allowed to vary by race r. We pool black and Hispanic drivers into a single nonwhite, or minority, group when estimating model parameters. The noise term ϵ_{ij} is assumed to be a standard normal variable. Thus, for an individual driving at

speed x, her probability of discount is

$$Pr(\text{Discount}|x) = \Phi(t_{rj} - b \cdot x_i)$$

Conditional on officer, county, and driver race, the likelihood for each speed, after the manipulation by the officer, is the following:

$$Pr(X = x) = \begin{cases} P_{\lambda_{rc}}(x) & \text{if } x < x_d \\ P_{\lambda_{rc}}(x_d) + \sum_{k=x_d+1}^{X} P_{\lambda_{rc}}(k) \cdot \Phi(t_{rj} - b \cdot k) & \text{if } x = x_d \\ P_{\lambda_{rc}}(x) \cdot \Phi(t_{rj} - b \cdot x) & \text{if } x > x_d \end{cases}$$

7.1 Identification

In principle, our model can be identified using only aggregate information, as if all data came from one officer. Intuitively, the tickets provide 40 moments (for each 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 in similar to our differences-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 estimation of speeds is primarily estimated using officers who exhibit no lenience, from which we get an estimate of the true distribution of speeds.

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}; 1592×2 officer average racial preferences, t_{jr}; and the slope of the cost function b, totaling 2,789 parameters. Details of how the estimation is carried out in practice are provided in the Appendix.

7.2 Model Estimates

Table 11 presents estimates of the model parameters. Columns 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 .023. 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 0.1 higher than for minorities. While the values of t are by themselves hard to interpret, dividing them by the slope parameter b gives the interpretation of the valuation in terms of MPH driven. The difference of 0.1 in valuation between whites and minorities, scaled by .0228, tells us that the average officer treats a minority driver like a white driver stopped for driving 4 MPH faster.

These differences in treatment are more easily understood in terms of the probability of discount (i.e., fine reduction). *P(Discount)* represents the likelihood of receiving a reduced ticket if the driver is at the speed right above the bunching speed. 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 13 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.

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 .5 to .7 MPH. Figure 15 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.

7.2.1 Decomposing the Gaps in Speeding and 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 bias. Tables 12 and 13 seek to answer this question by decomposing the measured racial speed and discounting disparities, respectively, into discrimination by officers, sorting of officers across counties, and differential speeding by racial groups.

To do so, we simulate 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 x from their race-and-county specific distribution $P_\lambda(x)$. 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 $t + \epsilon > b \cdot x$.

The "Baseline" row of Table 12 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 speeding by 15%. 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 speeding 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 46%. The remaining gap is due exclusively to the fact that minorities are driving faster speeds. In the second panel of Table 12 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.

In Table 13, we report our decomposition of the gap in bunching. Consistent with the previous discussion, removing both the discrimination and sorting of officers leads to reductions in the aggregate disparity. As reported in the "Neither" rows, only 9-13% of the gaps in bunching are due to differences in speeding by white and minority drivers.

7.3 Counterfactual Analysis

The model above estimates officer discrimination and driver speeds using a complementary approach to our differences-in-differences regressions, but corroborates the same story. While minority drivers on average drive slightly faster, officer are less likely to discount minorities at any given speed, and this disparity varies significantly across officers.

We now use the estimates to conduct a series of policy counterfactuals to explore how best to curb discrimination in speeding tickets. Because we provide a nonparametric estimate of the distribution of behavior across officers and speeds across race and locations, we can explore a rich set of counterfactuals whose outcomes depend on the full distribution of lenience and discrimination. 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.

It is important to note here that we steer clear of making normative statements in these counterfactuals; the only outcome we consider is the statewide disparity in discounting. 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, 2015).

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 .4% of officers).

The statewide disparity in treatment barely changes in response to removing these officers, falling by less than 2%. 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 in nearly equal measure for both white and black 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. Results are reported in Table 14, in which we display the probabilities of discount for white and minority drivers unweighted, to represent the statewide disparity. 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 reduction in the discount gap from .094 to .082, a 14% reduction. An increase in minority officers from the empirical share of 35% to 45% reduces the gap from .094 to .088, an 8% reduction.

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 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 14 present the results of such a policy. The first column is the baseline simulation of the model to match the true data. The second column sorts officers within a troop such that the least biased officers are in counties with the most minorities. The third column 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

deleterious effect on the treatment gap. The least biased officers are also the least lenient on average, leading minorities to be treated poorly relative to whites, now exposed to lenient and biased officers. The gap in probability of discount increases from .094 to .097. Much more effective in reducing the gap in treatment is assigning the most *lenient* officers to minority counties. This policy reduces the treatment gap from .094 to .01.

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; 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.

We have also assumed that officers are uniform in their treatment of drivers across counties, which we impose for modeling simplicity. If this uniformity does not hold, it may be the case that reassigning lenient officers to minority communities will reduce their lenience, weakening the efficacy of our proposed reassignment policy.

8 Interpretation

We have shown that officers discriminate when setting punishments for speeding, that this discrimination is concentrated among 40% of officers, and that our officer-level measures of discrimination can inform the best policy to mitigate the disparity in ticketing. In this section we consider potential threats to our methodology and alternative interpretations of our results.

8.1 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.

We do not believe that this issue is a serious concern in our setting. In Section 4 we show that officer lenience does not correlate with the frequency of tickets written and, in Section 5, that discrimination does not correlate with the share of tickets written for minorities. We therefore believe that lenience and discrimination on the stopping margin are not correlated with behavior on the discounting margin.

We further believe that any discrimination on the stopping margin would 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, and they are disproportionately at the discount point. Thus the disparity in discounting without the disparity in ticketing would be even greater than what we observe.[20] We formalize this argument in the Appendix and present a method to estimate the potential effect of selection on our results.

8.2 Racial Difference in Requesting a Break

One key insight of our analysis is that while whites and minorities do not seem to be differentially exposed to police through traffic enforcement, the quality of the interaction can vary significantly. This insight has also been confirmed by research that 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

[20] 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. We therefore use a trimming procedure in the Appendix that provides bounds on our discrimination estimates under any relationship between ticketing and discounting.

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. As in most discrimination studies, we do not have direct information on the quality and content of the interaction between officer and driver, so we cannot directly test for whether drivers differ in their propensity to request a break.

We do not believe, however, that differences in requests for a break can explain the disparity in discounting we observe. For a given level of lenience toward whites, we still see differences in discrimination across officers. If officers are simply receiving solicitations for a discount from the drivers (and whites ask more often), we should expect that for a given level of lenience toward whites, lenience toward minorities is a fixed fraction of that lenience. This pattern is not borne out in the data. We find that 50% of the variance in discrimination remains after conditioning for lenience against whites.

Additional evidence that requests for a break are not driving our results are presented in Table 15. Here we estimate linear probability models of receiving a discount with driver fixed effects and officer fixed effects separately, and then we estimate with both sets jointly. We restrict attention to tickets where the driver appears in our data at least twice. As evidenced by the adjusted R-squared measures, both sets of fixed effects have explanatory power for whether a driver is discounted. We should expect this finding, since we have shown that officer lenience and driver characteristics matter for discounting. Surprisingly, however, driver fixed effects have very little explanatory power after accounting for officer fixed effects. If it is the case that drivers vary systematically in their propensity to request a ticket, we should expect that the identity of the driver matters more than the identity of the officer, but that is not what we find.

Understanding differences in likelihood to request a break remains a challenge for all discrimination papers. The hit-rate test and approaches that benchmark the behavior of white officers against black officers are as susceptible to labeling as discrimination a propensity of white drivers to differentially request a break from white officers. So, while we do not have a perfect solution, this issue is unresolved in the literature.

8.3 Statistical Discrimination v. Taste-Based Discrimination

Throughout the paper, we have defined discrimination as the differential treatment of drivers by race who are stopped for the same speed. This definition is not innocuous, as there may be some reasons for differential treatment unrelated to observed driving speed that, while contentious in their use, are not specifically racial animus. For example, officers may be choosing treatment on the basis of how an individual's future driving responds to punishment (Gehrsitz, 2015; Hansen, 2015) or the likelihood of paying a ticket (Rowe, 2010; Makowsky and Stratmann, 2009). If individuals systematically differ by race in these characteristics, the racial disparities we observe may reflect the fact that officers are statistically discriminating by using race as a proxy for deterrability.

In the Appendix, we present a simple test for whether officers are attempting to minimize court contesting or maximize deterrence. We find no evidence that they are doing either, which suggest that they are not statistically discriminating to optimize these objectives.

Further, as noted in Anwar and Fang (2006), statistical discrimination can only explain behaviors that are uniform across officers, as they are due to the relationship between race and unobserved heterogeneity rather than anything specific to the officer. The median amount of bias in our setting is small, but we find a significant and skewed right tail. Such a distribution of disparate treatment is hard to explain with statistical discrimination.

We do, however, believe that our results are consistent with some officers incorrectly discriminating in the belief that it is statistical. As shown in Table A-2, drivers with fewer previous tickets are treated more leniently. And in Figure 17, we show that the white-minority gap in treatment is smaller among drivers with more previous tickets.

These facts suggest that officers decide who to discount on the basis of whether they are a responsible driver, and they use race as a signal for responsibility when there is no previous driving record. We argue that this statistical discrimination is incorrect because whites and minorities in truth have quite similar driving records, as reported in our summary statistics, Table 1. This finding is consistent with the result of Arnold et al. (2017), who find that bail judges discriminate against black offenders, but do so out of an inaccurate belief that black defendants are differentially more likely to reoffend.

9 Conclusion

The large racial disparities in the criminal justice system have led many to claim discrimination as the root cause. We argue in this paper that identifying discrimination at the level of the individual criminal justice agent is crucial for understanding the best policy for mitigating the disparities in outcomes. We study speeding tickets and the choice of officers to discount drivers to a speed just below an onerous punishment.

By using a bunching estimator approach that allows for officer-by-race measures of lenience in tickets, we can explore the entire distribution of both lenience and discrimination on the part of officers. We find that 90% of the gap in discounting can be attributed to discrimination. The rest of the gap is due to underlying differences in driving speeds across races. Officers are very heterogeneous in their degree of discrimination, with 40% of members explaining the entirety of the aggregate discrimination.

We explore whether discrimination is predictable by regressing individual officers' bias on demographic and personnel characteristics. We find that officers tend to favor their own race, older officers are more racially biased, and women and college-educated officers are less biased on average. Personnel information, such as failing an entry exam, receiving civilian complaints, and seeking a promotion, are not strongly informative about bias.

Using a model of driver speeding and officer decision-making, we confirm that while minorities drive faster on average, our officer-level estimates of bias are not confounded by differences in speeding across groups. We find that setting discrimination to zero across officers fails to remove the majority of the treatment gap, due to the fact that minorities tend to live in regions where officers are less lenient toward all drivers. Because of this fact, policies directed at reducing discrimination directly have a significant but modest effect on the treatment gap. Policies that instead target officers' lenience, by reassigning lenient officers to minority neighborhoods, are much more effective at reducing the aggregate treatment disparity.

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.

Anwar, S. 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. (2017). Racial bias in bail decisions. Technical report, National Bureau of Economic Research.

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.

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.

Chandra, A. and Staiger, D. O. (2010). Identifying provider prejudice in healthcare. Technical report, National Bureau of Economic Research.

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*, pages 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*, pages 1–67.

Delaigle, A., Hall, P., and Meister, A. (2008). On deconvolution with repeated measurements. *The Annals of Statistics*, pages 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*, pages 281–315. Emerald Publishing Limited.

Fryer, R. G. (2016). An Empirical Analysis of Racial Differences in Police Use of Force. *NBER Working Paper*, 22399:1–63.

Gehrsitz, M. (2015). Speeding, Punishment, and Recidivism Evidence from a Regression Discontinuity Design. *Unpublished Manuscript*, pages 1–45.

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*, pages 1–48.

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.

Hansen, B. (2015). Punishment and Deterrence: Evidence from Drunk Driving †. *American Economic Review*, 105(4):1581–1617.

Heckman, J. J., 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.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Makowsky, M. D. and Stratmann, T. (2009). Political Economy at Any Speed: What Determines Traffic Citations? *American Economic Review*, 99(1):509–527.

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.

Najdowski, C. J., 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 Rice, S. and White, M., editors, *Race, Ethnicity, and Policing New and Essential Readings*, pages 180–204.

Ridgeway, G. 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.

Rowe, B. (2010). Discretion and Ulterior Motives in Traffic Stops: The Detection of Other Crimes and the Revenue from Tickets. *Unpublished Manuscript*, pages 1–27.

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*, pages 1–407.

Trinkner, R. and Goff, P. A. (2016). The color of safety: The psychology of race & policing. *The SAGE Handbook of Global Policing*, pages 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*, page 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. (2015). Racial Bias in Police Investigations. *Working Paper*, pages 1–29.

10 Tables and Figures

Table 1: Summary Statistics

	White	Black	Hispanic	Total
Female	0.362	0.397	0.301	0.354
	(0.481)	(0.489)	(0.459)	(0.478)
Age	37.39	34.15	34.20	36.03
	(14.89)	(12.10)	(11.93)	(13.83)
Florida License	0.818	0.853	0.893	0.842
	(0.386)	(0.355)	(0.309)	(0.365)
Zip Code Income	56.96	39.70	46.49	51.13
	(53.20)	(31.18)	(42.81)	(47.79)
Citations in Past Year	0.0391	0.0482	0.0459	0.0424
	(0.209)	(0.231)	(0.232)	(0.219)
MPH Over	15.49	16.67	18.32	16.38
	(6.518)	(7.046)	(6.972)	(6.829)
Tickets at 9 MPH Over	0.352	0.316	0.206	0.311
	(0.478)	(0.465)	(0.405)	(0.463)
Fine Amount	182.3	190.2	200.7	188.0
	(76.40)	(80.50)	(79.33)	(78.23)
Share	.584	.184	.231	1
N	667,447	210,368	264,374	1,142,189

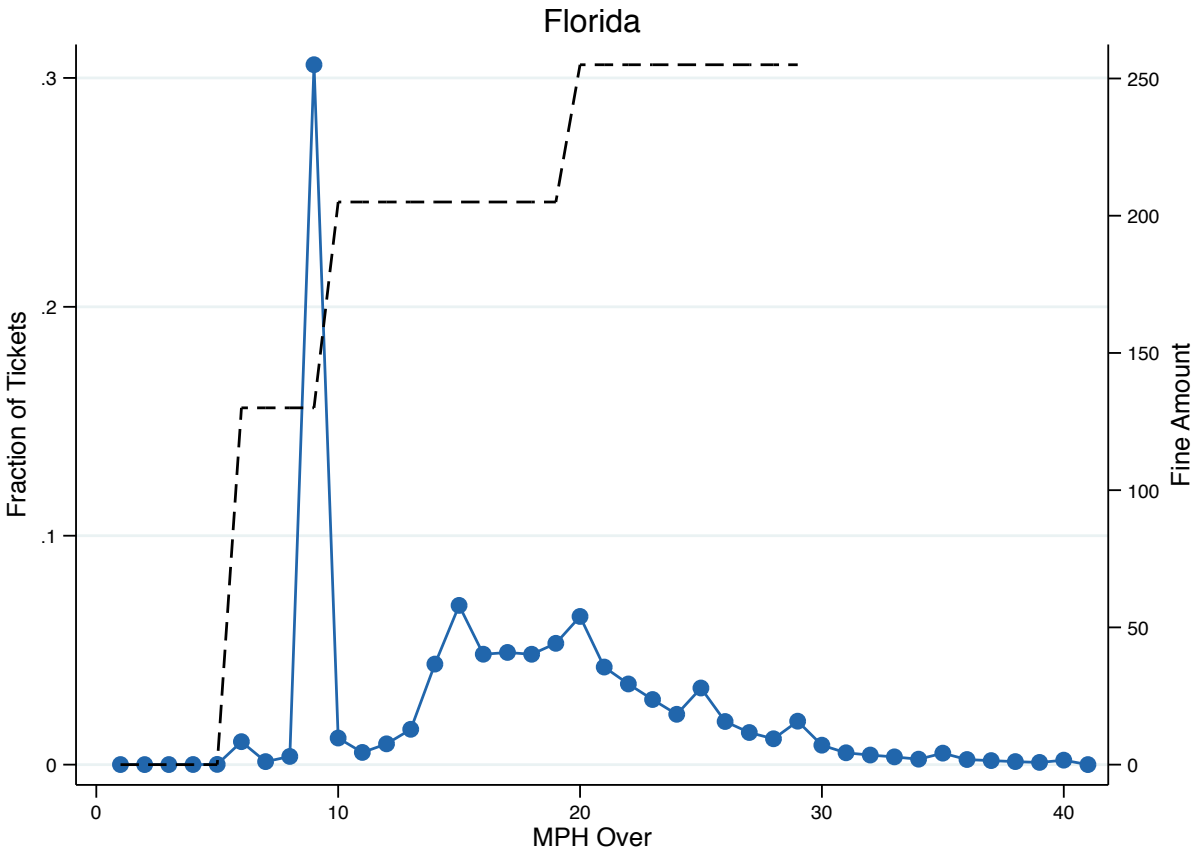
Notes: Number of observations is 571,751 (White); 184,567 (Black); 234,323 (Hispanic); 990,641 (Total). Standard deviations in parentheses. Zip code income is missing for 42% of White stops, 40% of Black stops, 37% of Hispanic stops. To account for the fact that a large share of fine amounts are missing or zero in our data, we impute the fine amount with the modal non-zero fine for each county × speed over the limit cell.

Table 2: Characteristics of Cited Drivers Relative to Other Data Sources

	Citations	ACS - Any	ACS - Drivers	Crash - Any	Crash - Injury
Female	0.356	0.515	0.474	0.424	0.441
	(0.479)	(0.500)	(0.499)	(0.494)	(0.497)
Age	34.90	47.46	41.70	39.65	39.77
	(13.45)	(19.39)	(13.72)	(16.78)	(17.11)
White	0.578	0.625	0.606	0.556	0.576
	(0.494)	(0.484)	(0.489)	(0.497)	(0.494)
Black	0.181	0.138	0.136	0.189	0.193
	(0.385)	(0.345)	(0.343)	(0.391)	(0.394)
Hispanic	0.241	0.200	0.217	0.233	0.211
	(0.428)	(0.400)	(0.412)	(0.423)	(0.408)

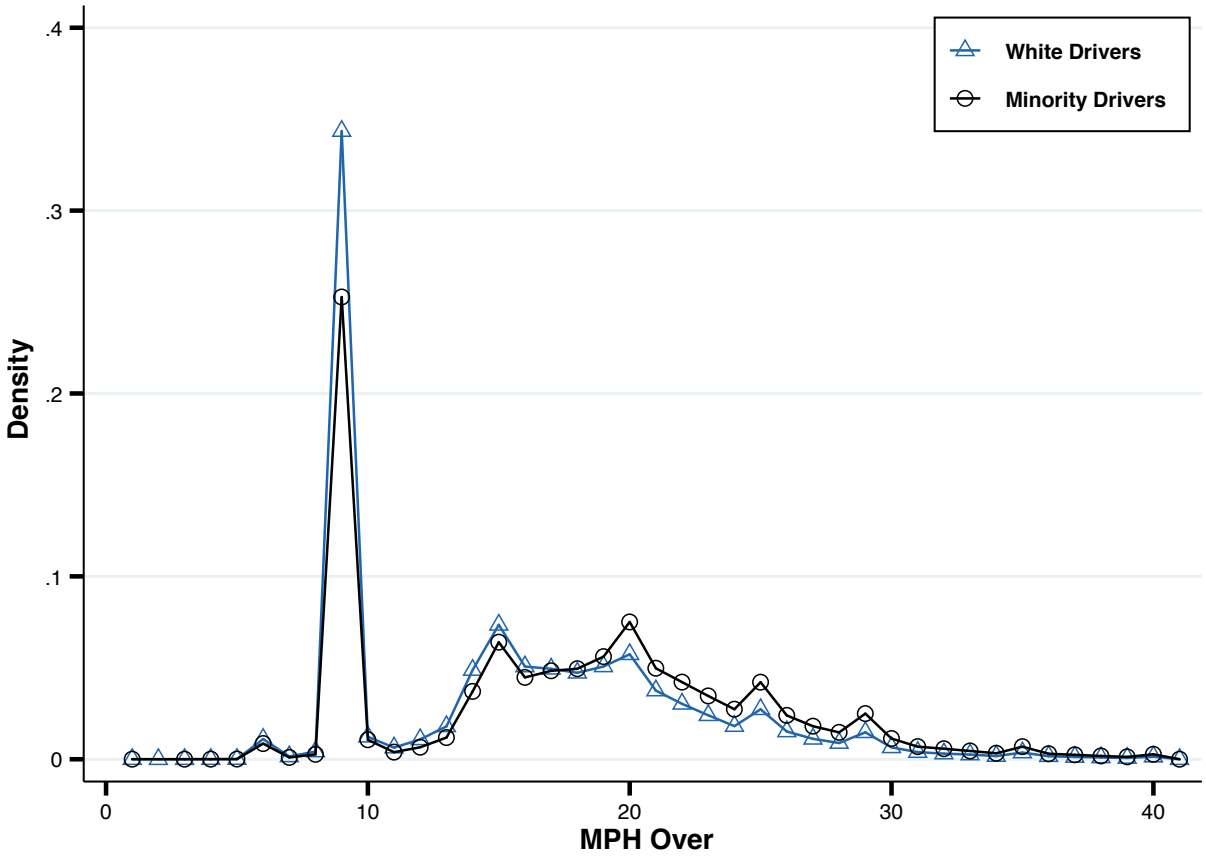
Notes: Standard deviations in parentheses. ACS data include individuals aged 16 or older and sampling weights are used.

Figure 1: Distribution of Charged Speeds and Fine Schedule



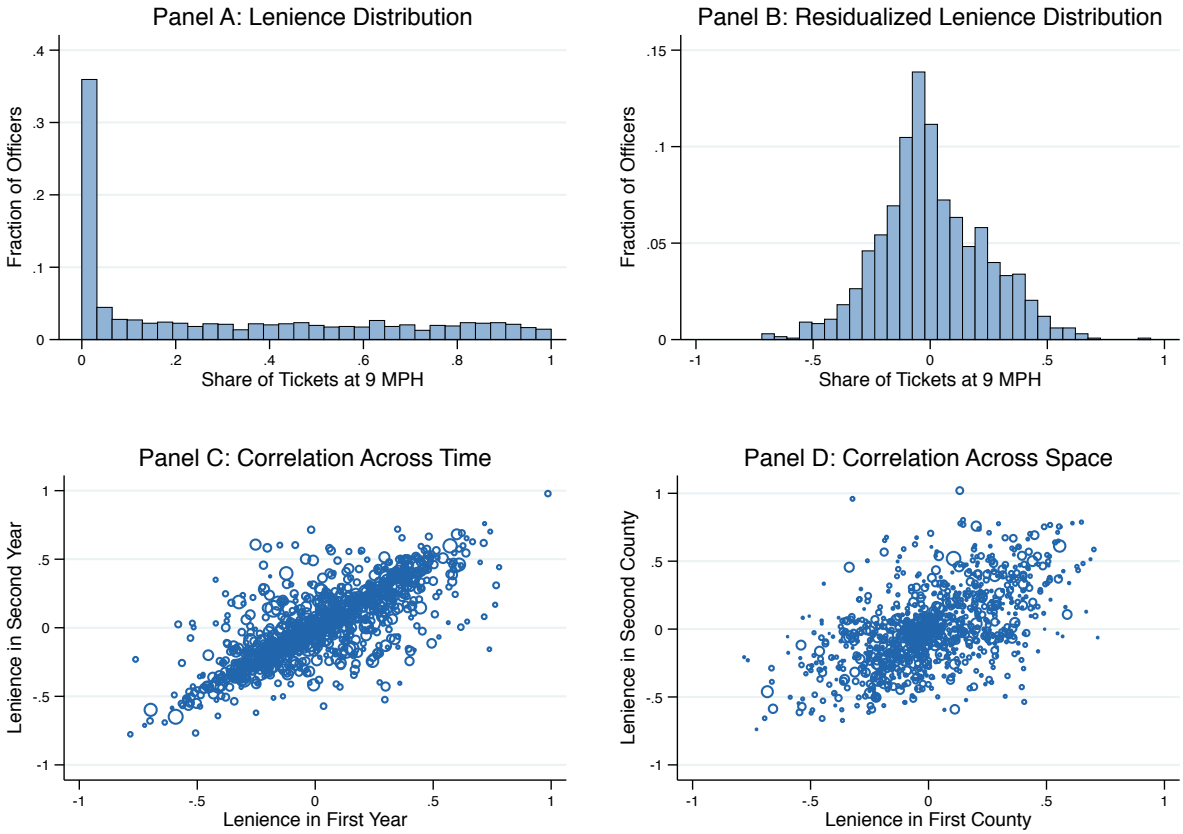
Notes: Connected line shows histogram of tickets. Dashed line plots fine schedule for Broward County. 30 MPH over is felony speeding and carries a fine to be determined following a court appearance.

Figure 2: Charged Speed Distributions by Driver Race



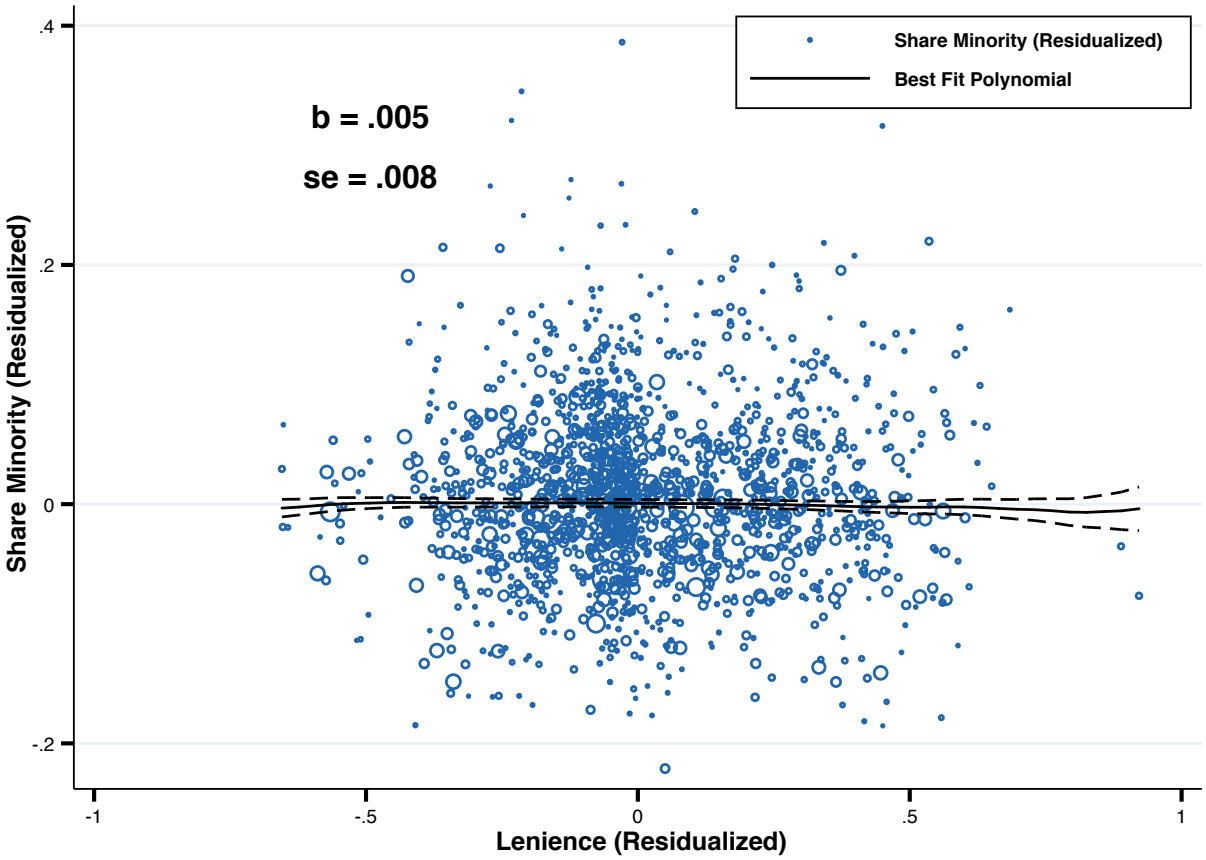
Notes: Connected line shows histogram of ticketed speeds, separately by driver race. 34.3% of tickets to white drivers are given at 9 MPH over compared to 25.2% of tickets for minority drivers.

Figure 3: Evidence of Officer Lenience



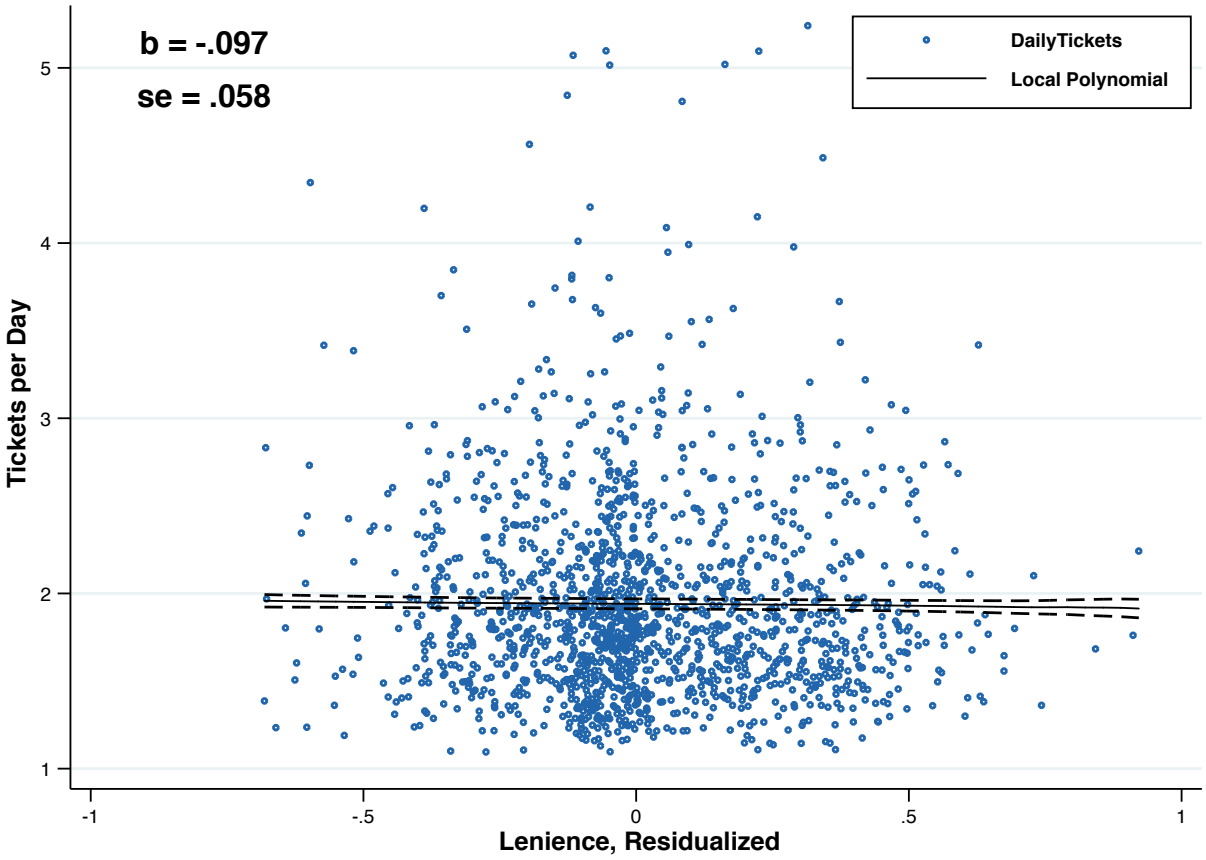
Notes: Panel A plots the across-officer distribution of lenience, calculated as the share of tickets given for 9 MPH over the limit. Panel B plots the across-officer distribution of residualized lenience. Panel C plots officers' residualized lenience in the years with the most and second most citations. 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. See text for additional details.

Figure 4: Officer Lenience and Share Drivers Minority



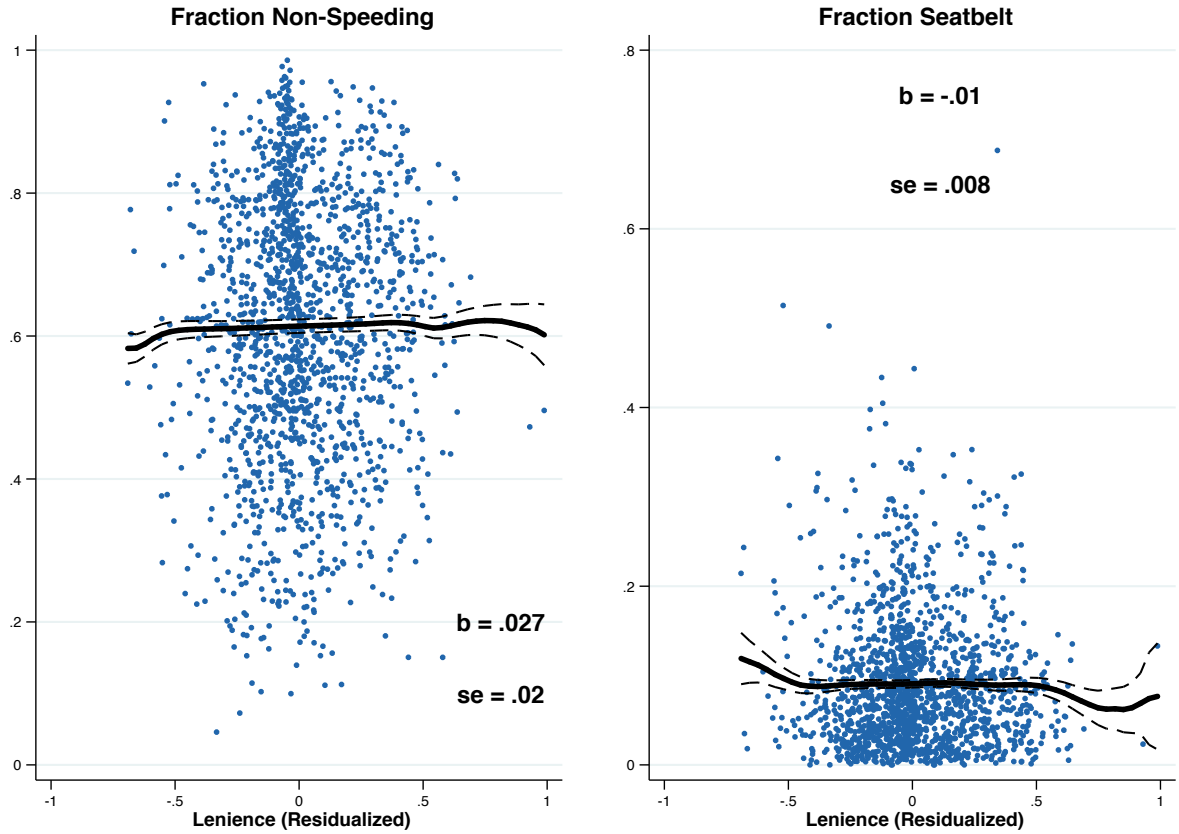
Notes: Figure plots each officer's lenience against the share of officer's drivers who are minority, residualized. The y-axis variable is constructed by regressing whether an individual is a minority on county and year fixed effects and other demographics and averaging the residual within an officer.

Figure 5: Officer Lenience and Average Daily Tickets



Notes: Figure plots each officer's lenience against the average number of tickets per day given by the officer.

Figure 6: Officer Lenience and Other Tickets



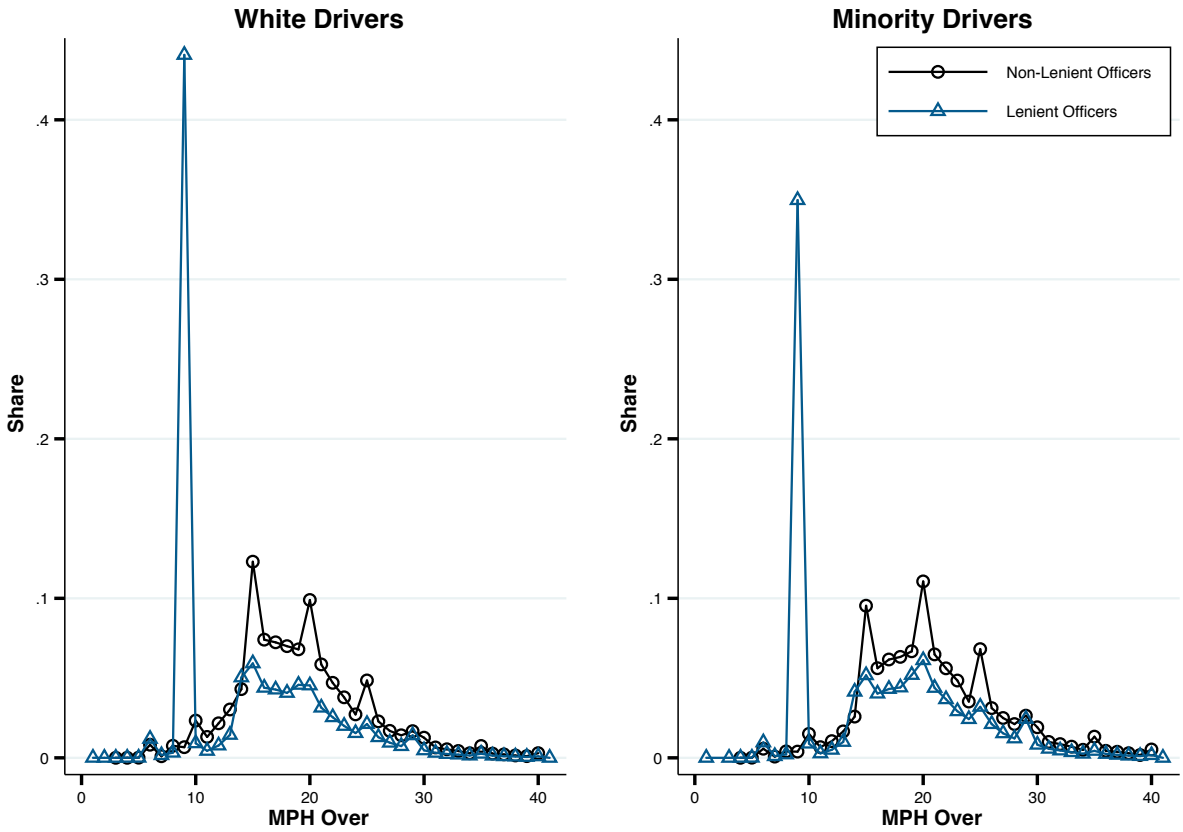
Notes: Figure plots each officer's lenience against the share of tickets written for non-speeding infractions and share of tickets written for seatbelt violations.

Table 3: Officer Lenience Randomization Check

	Full Sample			GPS Sample	
	(1)	(2)	(3)	(4)	(5)
	Lenience	Lenience	Lenience	Lenience	Lenience
Driver Black	0.00228	0.00268	-0.000275	-0.00688	0.000165
	(0.0166)	(0.00305)	(0.00323)	(0.00485)	(0.00518)
Driver Hispanic	-0.0900***	-0.00557	-0.00690	-0.0220*	-0.00230
	(0.0285)	(0.00476)	(0.00452)	(0.0123)	(0.00335)
Driver Female	0.0194***	0.00490**	0.00272	0.00127	0.00128
	(0.00434)	(0.00240)	(0.00182)	(0.00128)	(0.00198)
Florida License	-0.0775***	0.000557	0.000560	0.00695	-0.00226
	(0.0249)	(0.00334)	(0.00312)	(0.00749)	(0.00372)
Driver Age	-0.167	0.243*	0.0881	0.156	0.0352
	(0.161)	(0.142)	(0.115)	(0.106)	(0.0751)
1 Prior Ticket	-0.0104*	-0.0000658	-0.000576	0.00136	-0.000708
	(0.00620)	(0.000811)	(0.000875)	(0.00127)	(0.00241)
2+ Prior Tickets	-0.0223*	-0.000594	-0.000175	0.00427**	0.00138
	(0.0122)	(0.00110)	(0.00129)	(0.00208)	(0.00377)
Log Zip Code Income	-0.00178	0.00445	-0.00402**	-0.000417	-0.0000911
	(0.00826)	(0.00288)	(0.00193)	(0.00294)	(0.00223)
F-test	0	.401	.085	.311	.957
Mean	.305	.305	.305	.326	.326
Location FE		X			
Location + Time FE			X	X	
GPS FE					X
Observations	1142189	1142189	1142189	135565	135565

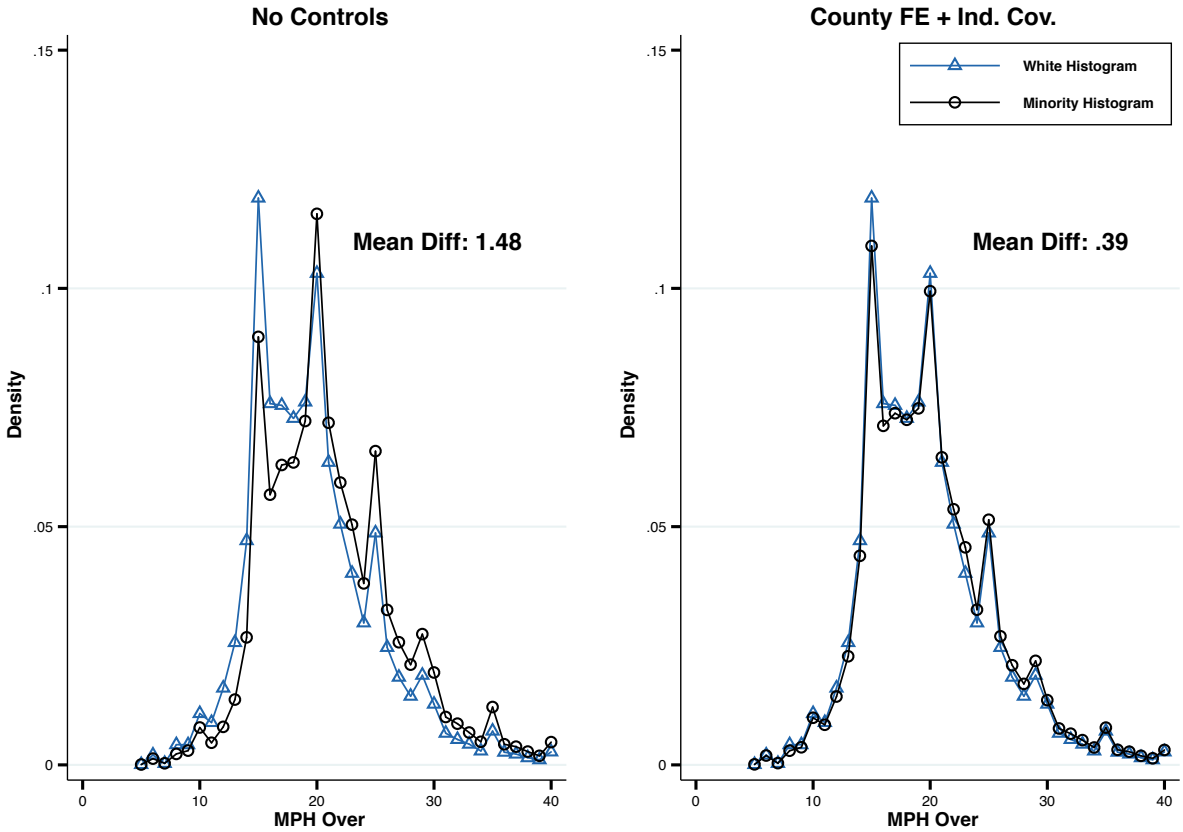
Notes: All regressions includes vehicle type fixed effects and county fixed effects. The F-test reports the joint hypothesis test that variables Driver Black through Log Zip Code Income are zero. Standard errors are clustered at the county level. "Location FE" includes county by highway fixed effects. "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.

Figure 7: Difference-in-Difference Raw Data Plot



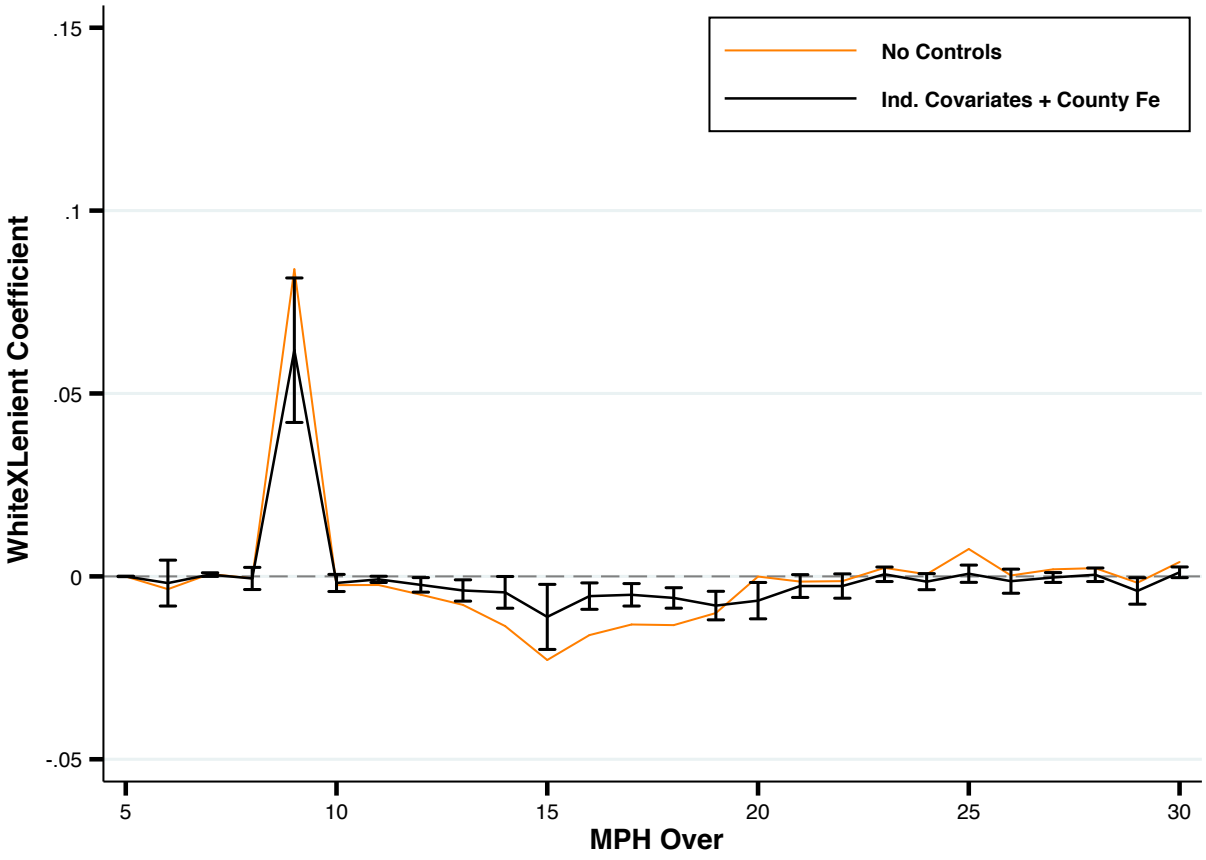
Notes: Figure plots the difference-in-difference regression results for each speed. The y-axis plots the interaction between driver being white and the officer being lenient. Standard errors are at the 5% level.

Figure 8: Histogram for Non-Lenient Officers



Notes: Figure plots the difference-in-difference regression results for each speed. The y-axis plots the interaction between driver being white and the officer being lenient. Standard errors are at the 5% level.

Figure 9: Difference-in-Difference Results



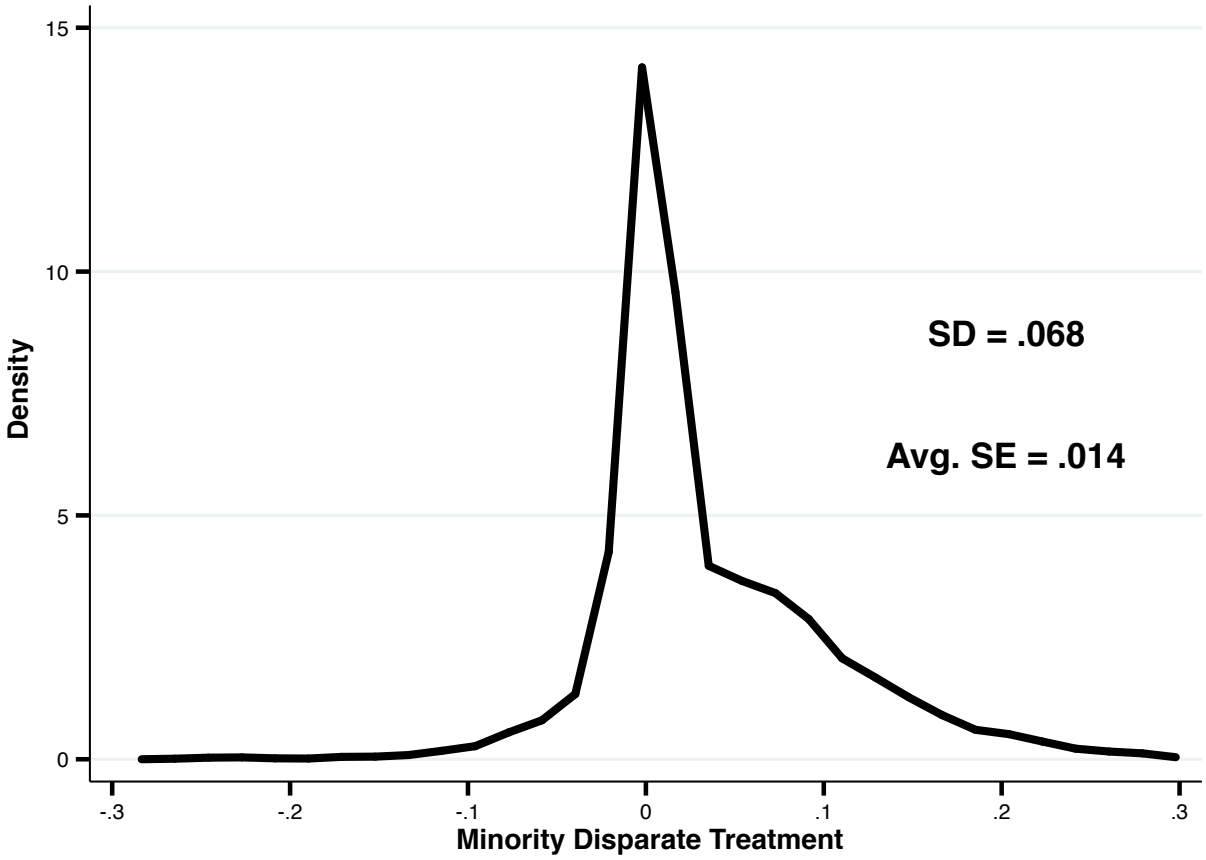
Notes: Figure plots the difference-in-difference regression results for each speed. The y-axis plots the interaction between driver being white and the officer being lenient. Standard errors are at the 5% level.

Table 4: Difference-in-Difference Results

	Full Sample			GPS Sample	
	(1)	(2)	(3)	(4)	(5)
	Discount	Discount	Discount	Discount	Discount
Driver White	0.00126***	-0.0210***	-0.0127**	-0.00827*	-0.00715*
	(0.000326)	(0.00617)	(0.00586)	(0.00464)	(0.00395)
Officer Lenient	0.396***	0.304***	0.294***	0.243***	0.199***
	(0.0355)	(0.0377)	(0.0440)	(0.0194)	(0.0322)
Driver White	0.0840***	0.0673***	0.0620***	0.0689***	0.0561***
× Officer Lenient	(0.0167)	(0.0111)	(0.0101)	(0.00800)	(0.00602)
Mean	.305	.305	.305	.322	.322
Covariates		X	X	X	X
Location FE		X			
Location + Time FE			X	X	
GPS FE					X
Observations	1142189	1142189	1142189	135565	135565

Notes: Table reports linear probability estimates where the outcome variable is whether an individual is ticketed for 9 MPH over the limit. County-level clustered standard errors in parentheses. Dependent variable is an indicator for charged speed equals 9 MPH above the limit. GPS indicates fixed effects at the level of road segment by county.

Figure 10: Diff-in-Diff Officer-Level Results



Notes: Figure plots each officer's β_3^j from the regression
$S_{ij}^9 = \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2^j \cdot \text{Lenient}_j + \beta_3^j \cdot \text{White}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$.
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.

Table 5: Diff-in-Diff Officer-Level Results

	Discrimination Percentile					
	10 %	25%	50%	75%	90%	N
All Officers	-.0113	0	.0053	.0681	.1275	1591
White Officers	-.0076	0	.0231	.0835	.1386	998
Black Officers	-.0339	0	0	.0228	.0637	250
Hispanic Officers	-.0112	0	0	.0403	.1199	318

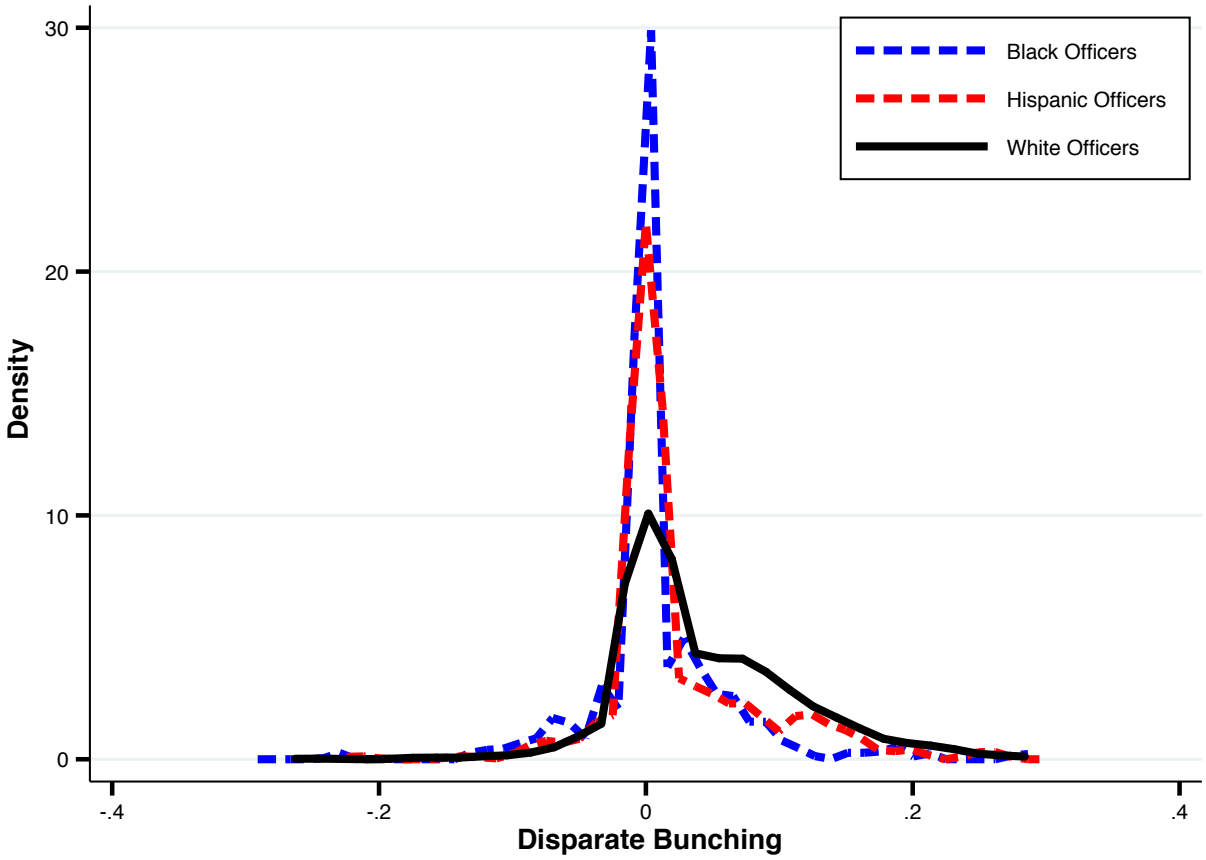
Notes: Each cell reports the average probability of being observed at the bunch point, broken down by officer-driver race, where the probability is derived from the officer lenience and discrimination estimates from the regression
$S^9_{ij} = \beta_0 + \beta_1 \cdot \text{White}_i + \beta^j_2 \cdot \text{Lenient}_j + \beta^j_3 \cdot \text{White}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$.
Observations are weighted so that officers have equal share across troops.

Table 6: Officer Discrimination Randomization Check

	Full Sample			GPS Sample	
	(1) Discrimination	(2) Disc	(3) Disc	(4) Disc	(5) Disc
Driver Black	0.000891	0.00152**	0.000708	0.00227	-0.0000482
	(0.00190)	(0.000604)	(0.000575)	(0.00190)	(0.000884)
Driver Hispanic	-0.00616*	0.000174	-0.000665	-0.00101	-0.000385
	(0.00319)	(0.000849)	(0.000677)	(0.00266)	(0.000635)
Driver Female	0.00125**	-0.000174	-0.000157	0.000534	0.000172
	(0.000614)	(0.000181)	(0.000210)	(0.000546)	(0.000440)
Florida License	-0.0111***	-0.000233	-0.000532	0.000171	0.000211
	(0.00304)	(0.000715)	(0.000681)	(0.00223)	(0.000487)
Driver Age	-0.0140	-0.00259	-0.000642	-0.0104	-0.00229
	(0.0274)	(0.0168)	(0.0144)	(0.0276)	(0.0139)
1 Prior Ticket	-0.000770	0.0000159	0.0000403	0.000661	0.000373
	(0.000694)	(0.000174)	(0.000170)	(0.000663)	(0.000452)
2+ Prior Tickets	-0.00198	0.0000713	0.000125	0.00111	0.000302
	(0.00144)	(0.000224)	(0.000210)	(0.000754)	(0.000522)
Log Zip Code Income	0.00115	0.00150**	0.0000822	0.000799	0.000106
	(0.00128)	(0.000720)	(0.000365)	(0.00108)	(0.000610)
F-test	0	.177	.19	.274	.785
Mean	.305	.305	.305	.326	.326
Location FE		X			
Location + Time FE			X	X	
GPS FE					X
Observations	1142189	1142189	1142189	135565	135565

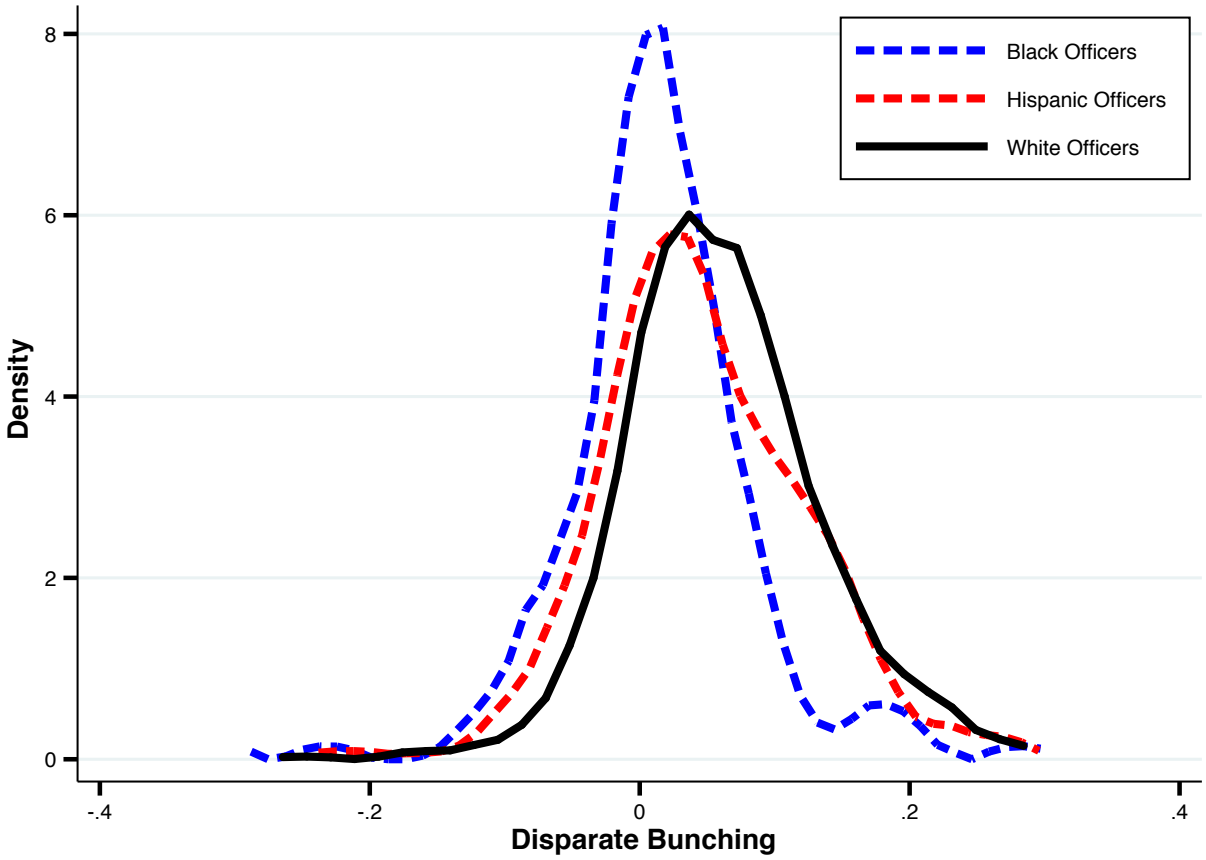
Notes: All regressions includes vehicle type fixed effects and county fixed effects. The F-test reports the joint hypothesis test that variables Driver Black through Log Zip Code Income are zero. Standard errors are clustered at the county level. "Location FE" includes county by highway fixed effects. "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.

Figure 11: Diff-in-Diff Officer-Level Results



Notes: Figure plots each officer's β_3^j from the regression
$S_{ij}^9 = \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2^j \cdot \text{Lenient}_j + \beta_3^j \cdot \text{White}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$.
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 12: Officer-Level Results, Lenient Officers



Notes: Figure plots the discrimination coefficient β_3^j for all officers who are lenient.

Table 7: Lenience by Officer-Driver Race

	White Officer	Black Officer	Hispanic Officer	Total
White Driver	0.321	0.291	0.319	0.310
	(0.303)	(0.295)	(0.298)	(0.299)
Black Driver	0.282	0.289	0.278	0.283
	(0.286)	(0.292)	(0.279)	(0.285)
Hispanic Driver	0.266	0.278	0.279	0.274
	(0.279)	(0.294)	(0.285)	(0.286)

Notes: Each cell reports the average probability of being observed at the bunch point, broken down by officer-driver race, where the probability is derived from the officer lenience and discrimination estimates from the regression $S^9_{ij} = \beta_0 + \beta_1 \cdot \text{White}_i + \beta^j_2 \cdot \text{Lenient}_j + \beta^j_3 \cdot \text{White}_i \cdot \text{Lenient}_j + X_{ij}\gamma + \epsilon_{ij}$. Observations are weighted so that officers have equal share across troops.

Table 8: Predicting Officer Bias

	(1) Black Bias	(2) Hispanic Bias
Black Officer	-0.0232***	-0.0203***
	(0.00328)	(0.00460)
Hispanic Officer	-0.00731**	-0.0148***
	(0.00326)	(0.00457)
Other Race	0.0103	-0.00375
	(0.0108)	(0.0131)
Female Officer	-0.0124***	-0.0112**
	(0.00366)	(0.00476)
Age	0.0000700	0.000496
	(0.00166)	(0.00277)
Experience	-0.00223	0.00242
	(0.00212)	(0.00345)
Failed Entrance Exam	-0.00647	-0.00484
	(0.00431)	(0.00550)
Any College	-0.00273	0.00227
	(0.00262)	(0.00383)
Any Complaints	-0.00101	0.00115
	(0.00367)	(0.00460)
Sought Promotion	-0.00123	-0.00294
	(0.00262)	(0.00390)
Dep Var	Bias	Bias
Mean	.02	.028
Observations	1587	1587

Notes: Robust standard errors in parentheses. In Columns (1) and (4), dependent variable is our measure of the officer's bias. In Columns (2) and (5), it is an indicator for whether our bias estimate is positive. In Columns (3) and (6), it is an indicator for whether the estimate is positive and statistically significant. Columns (1)-(3) refer to bias against blacks while Columns (4)-(6) refer to bias against Hispanics.

Table 9: Predicting Officer Complaints/Force

	(1) # Complaints	(2) Any Complaints	(3) # Use of Force	(4) Any Use of Force
Lenience	-0.721***	-0.206***	-0.246	-0.113**
	(0.223)	(0.0588)	(0.150)	(0.0530)
Minority Bias	0.355	0.321	0.444	0.0793
	(0.642)	(0.216)	(0.462)	(0.183)
Black	0.112	0.00156	-0.190**	-0.0847**
	(0.176)	(0.0400)	(0.0905)	(0.0336)
Hispanic	-0.00683	0.0170	0.0340	0.00730
	(0.144)	(0.0372)	(0.0993)	(0.0373)
Other	0.184	0.0261	-0.234	-0.0720
	(0.380)	(0.0998)	(0.181)	(0.0939)
Female	-0.296*	-0.110**	-0.0145	0.0135
	(0.158)	(0.0496)	(0.104)	(0.0442)
Age	-0.121	0.161*	-0.735***	-0.193**
	(0.332)	(0.0898)	(0.213)	(0.0794)
Age Squared	0.0166	-0.0247*	0.0608**	0.0136
	(0.0478)	(0.0131)	(0.0266)	(0.0109)
Experience	-0.0636	-0.0821	-0.559*	-0.0108
	(0.413)	(0.130)	(0.331)	(0.117)
Exp Squared	-0.0265	0.0332	-0.00401	-0.00118
	(0.0767)	(0.0249)	(0.0460)	(0.0196)
Failed Entrance Exam	0.256	0.0429	-0.104	-0.00344
	(0.205)	(0.0483)	(0.109)	(0.0458)
Any College	-0.186*	-0.0263	0.102	0.0135
	(0.104)	(0.0294)	(0.0946)	(0.0264)
Sought Promotion	-0.199*	-0.0680**	-0.0407	0.0243
	(0.113)	(0.0294)	(0.0887)	(0.0277)
Mean	1.259	.551	.560	.294
Observations	1401	1401	1401	1401
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 fixed effects for years when and districts where the officer worked.

Table 10: Early Discrimination

Early Measure Cutoff	Full Sample Percentiles		
	N	Median	95th percentile
2% most discriminatory	28	3.2	23.6
5% most discriminatory	76	6.5	60.2
10% most discriminatory	153	9.2	82.6

Early Sample	Full Sample		
	N	Type I Error	Type II Error
$\theta_j > 1.96 \cdot \text{SE}(\hat{\theta}_j)$	398	32.2%	54.6%
$\theta_j > 2.33 \cdot \text{SE}(\hat{\theta}_j)$	329	31.0%	61.8%
$\theta_j > 3 \cdot \text{SE}(\hat{\theta}_j)$	236	28.0%	71.4%

Notes: This table presents the relationship between early measures of discrimination (using first 100 tickets) and discrimination using all an officer's data. The first panel reports how different cutoffs in the percentile of early discrimination translate to percentiles in the full sample. For example, the median percentile of full-sample discrimination for an officer who is in the top 2% of early discrimination is 3.2. The 95th percentile among those from the early 2% cutoff is 23.6. The bottom panel reports how often the early measures mislabels an officer as discriminatory and how often it misses a discriminatory officer. Type I Error reports the percentage of officers identified as discriminatory in the early sample who are *not* discriminatory at the 5% level in the full sample. Type II Error reports the percentage of officers who are discriminatory in the full sample at the 5% level who are not identified as discriminatory in the early sample.

Table 11: Model Parameter Estimates

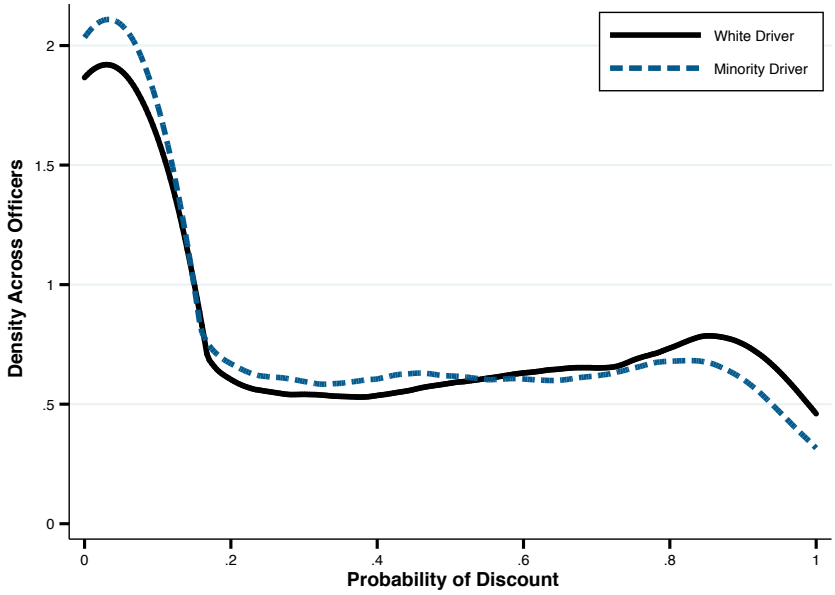
	White			Minority			
	μ	σ^2	# Param	μ	σ^2	# Param	Mean Diff
b, slope	.0228***	—	1	—	—	—	—
	(9.38 X 10^{-5})						
t, officer valuations	-2.41***	20.12***	1591	-2.52***	19.50***	1591	0.10***
	(.046)	(0.781)		(.057)	(0.757)		(.013)
λ, *speeds*	18.387***	2.905***	67	19.107***	2.186***	67	-0.720***
	(0.002)	(0.506)		(0.003)	(0.381)		(0.004)
Pr(Discount)	0.357***	0.125***	1591	0.325***	0.112***	1591	0.033***
	(6.90 X 10^{-4})	(0.005)		(7.98 X 10^{-4})	(0.004)		(0.001)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

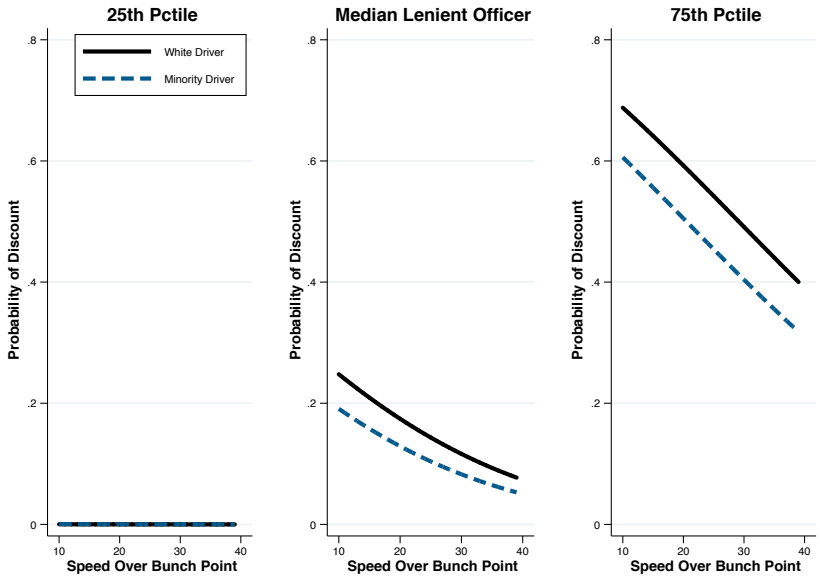
Notes: This table presents estimates of the model introduced in section 7. b is the slope parameter for how officers weight the speed of drivers in choosing to discount, t is each officer's mean valuation of a racial group in choosing to discount, and λ is the poisson speed parameter for each race by county. Pr(Discount) = $\Phi(t - 10b)$, i.e. the probability of being discounted when driving right above the bunch point. Note that the discount probabilities are not technically parameters but rather are calculated using the estimated t's and b. The variances are empirical variances of the estimates, not adjusted for sampling error.

Figure 13: Model Estimates: Officer Lenience by Race



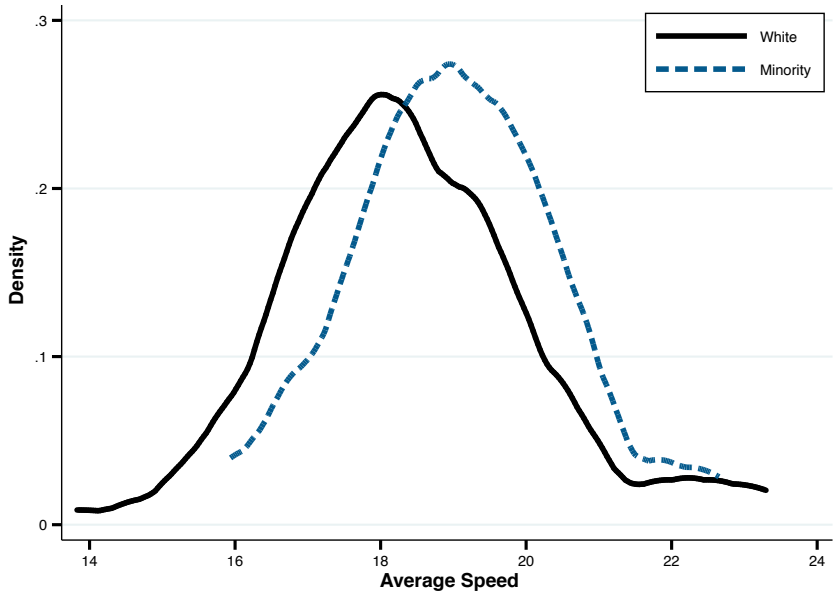
Notes: $P_{rj} \equiv P_j(\text{Discount}|X = 10, \text{ Driver Race} = r)$

Figure 14: Model Estimates: Percentiles of Officer Lenience



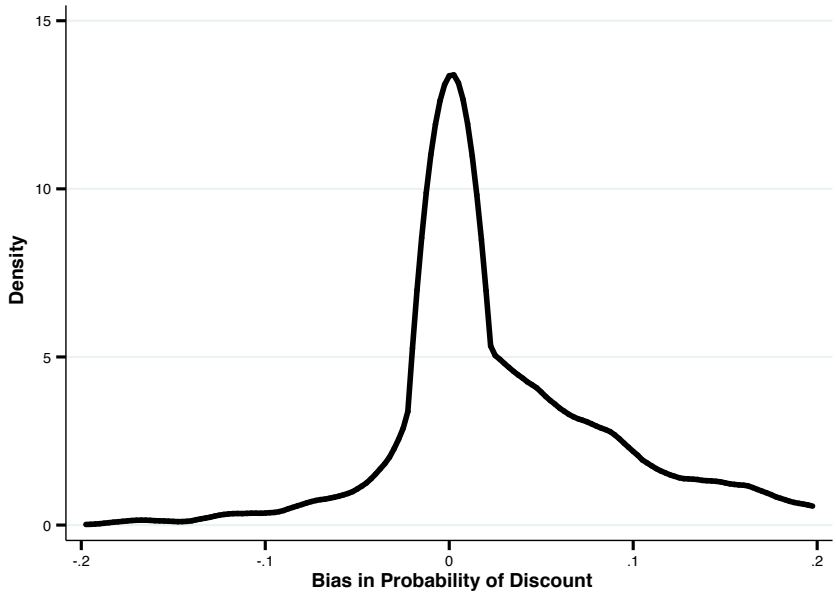
Notes: $P_{rj} \equiv P_j(\text{Discount}|X = 10, \text{ Driver Race} = r)$

Figure 15: Model Estimates: Speed Distribution



Notes: Figure plots the distribution of speed parameters λ across counties, separately by race of the driver.

Figure 16: Model Estimates: Racial Discrimination by Officer



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

Table 12: Speed Gap Decomposition

	State-Wide Disparity			
	White Mean (MPH)	Minority Mean	Difference	Percent
Baseline	15.533	17.545	2.011	100
	(.539)	(0.643)	(0.351)	
No Discrimination	15.511	17.218	1.707	84.9
	(.563)	(0.670)	(0.363)	
No Sorting	15.633	17.086	1.453	72.3
	(.501)	(.572)	(.276)	
Neither	15.520	17.063	1.543	54.6
	(.505)	(.577)	(.279)	

	County-Level Disparity			
	White Mean (MPH)	Minority Mean	Difference	Percent
Baseline	15.533	16.43	.897	100
	(.576)	(.586)	(.103)	
No Discrimination	15.511	16.091	.58	64.7
	(.563)	(.568)	(.076)	
No Sorting	15.633	16.55	.917	102.3
	(.501)	(.504)	(.054)	
Neither	15.622	16.199	.577	64.3
	(.505)	(.508)	(.0550)	

Notes: Table presents how the racial gap in discounting and speeds change 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 white drivers' outcome minus minority 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 13: Discounting Gap Decomposition

	State-Wide Disparity			
	White Mean	Minority Mean	Difference	Percent
Baseline	.346	.257	.090	100
	(.043)	(.048)	(.021)	
No Discrimination	.348	.287	.061	67.8
	(.041)	(0.047)	(.022)	
No Sorting	.325	.283	.042	46.1
	(.037)	(.041)	(.017)	
Neither	.325	.317	.009	9.2
	(.037)	(.041)	(.017)	

	County-Level Disparity			
	White Mean	Minority Mean	Difference	Percent
Baseline	.346	.3097	.037	100
	(.043)	(.043)	(.006)	
No Discrimination	.354	.341	.007	17.2
	(.041)	(0.041)	(.004)	
No Sorting	.325	.287	.038	101.7
	(.037)	(.037)	(.002)	
Neither	.325	.32	.0049	13.1
	(.037)	(.037)	(.0009)	

Notes: Table presents how the racial gap in discounting and speeds change 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 white drivers' outcome minus minority 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.

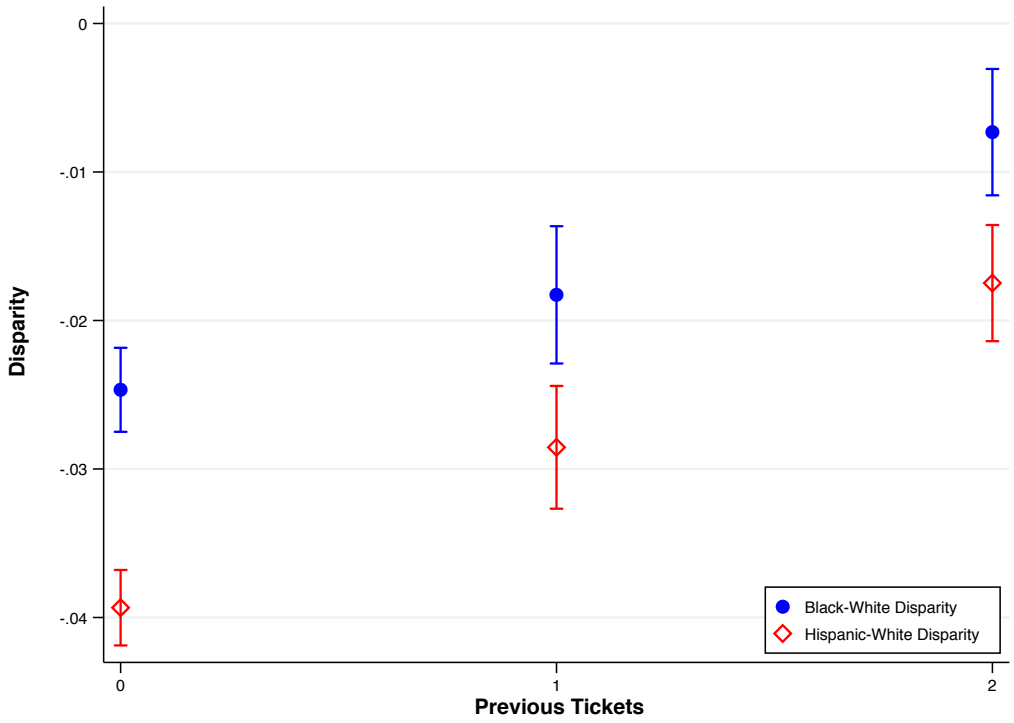
Table 14: Model Counterfactuals

	Hiring & Firing			
	White Mean	Minority Mean	Difference	Percent
Baseline	.353	.259	.094	100
	(.041)	(.046)	(.02)	
Fire 5% Most Biased Officers	.347	.254	.094	98.7
	(.043)	(.04)	(.02)	
Increase Female Share 10pp	.333	.251	.082	86.2
(Base of 8%)	(.038)	(.043)	(.02)	
Increase Minority Share 10pp	.353	.266	.088	92.3
(Base of 35%)	(.039)	(.044)	(.021)	

	Resorting Officers			
	White Mean	Minority Mean	Difference	Percent
Exposing Minorities	.350	.253	.097	102.6
To *Least Biased*	(.045)	(.05)	(.022)	
Exposing Minorities	.308	.299	.01	10.1
To *Most Lenient*	(.057)	(.065)	(.03)	

Notes: Results are reporting the probability of being ticketed 9MPH over. Observations are not weighted, so the disparities reflect statewide racial differences in lenience.

Figure 17: Racial Disparity by Previous Tickets



Notes: Table reports the white-black and white-hispanic disparity in discounting, where the race of the driver is interacted with the number of previous tickets a driver has at the point of the stop.

Table 15: Officer and Driver Fixed Effects

	(1) Discount	(2) Discount	(3) Discount
R-Squared	.558	.471	.728
Adj. R-Squared	.178	.466	.486
Driver FE	X		X
Officer FE		X	X
N	192380	192384	192380

Notes: Data is at the citation level and the outcome is whether a driver is found at the discount point. Sample is restricted to drivers with two or more tickets in the citations data. Columns vary by level of fixed effects.

Data Appendix

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 (1,677,177 observations, 64% of previous sample)
2. no crash is involved (1,676,141, 99.9%)
3. speed is between 0 and 40 over the limit (1,665,699, 99.4%)
4. posted speed limit is between 25MPH and 75MPH (1,664,570, 99.9%)
5. citations not from an airplane (1,660,355, 99.7%)
6. race/ethnicity is not missing (1,408,355, 84.8%)
7. race/ethnicity is white, black or Hispanic (1,335,887, 94.8%)
8. not missing driver's license state, gender, or age (1,333,720, 99.8%)
9. officer is identifiable (1,024,631, 76.8%)
10. officer has at least 100 tickets, and at least 20 for minorities and 20 for whites (988,096, 96.5%)

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 1,677,177 speeding tickets in our data, 1,651,933 have at least an officer last name and first initial listed. Of these, 1,260,900 match successfully to the FDLE data.

Hours and Shifts of Tickets

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

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

Estimation of Model

While the setup of the model is simple, non-parametrically estimating the distribution of bias is computationally complex. The model parameters to be identified are the 67×2 county-race speeds λ_{rc}, 1327×2 officer average racial preferences, t_{jr}, and the slope of the cost function b, totaling 2,789 parameters. This complexity makes estimation directly through maximum likelihood challenging.

We estimate the model by iterating across the groups of parameters until the solution converges. We solve first for a set of initial guesses by assuming that speeds and officer preferences are uniform and performing maximum likelihood on b, λ, t. We set $\hat{b}$ to be from the aggregate MLE. We then calculate the county-specific speeds and officer-specific preferences in the following way:

1. Set initial values $\lambda_{rc}^{(0)} = \hat{\lambda}$ from the aggregate MLE estimation.
2. Solve for officer preferences $t_{ij}^{(0)}$ by maximizing likelihood $\mathcal{L}\left(t_{jr}|x_{jr}, \{\lambda_{rc}^{(0)}\}, \hat{b}\right)$ conditional on speed parameters and slope estimate.
3. Solve for speed parameters $\lambda_{rc}^{(1)}$ by maximizing likelihood $\mathcal{L}\left(\lambda_{rc}|x_{jr}, \{t_{jr}^{(0)}\}, \hat{b}\right)$ conditional on officer preferences and slope estimate.
4. return to step 2 and repeat until parameter guesses converge.

Because in every iteration the total likelihood increases, the process will converge at least to a local optimum. We check different starting values to confirm that our results achieve a global maximum.

Conditional on speed and slope, the officer parameters are separable and thus can be easily solved, and similarly for the speed parameters when conditioning on officer parameters. Further, the conditional likelihood functions are unimodal in the parameters; this means that the score functions only cross zero once, simplifying the search for an optimum. Standard errors are calculated by estimating the information matrix via the variance of the parameters' score functions.

Accounting for Stopping Margin Selection

As discussed in Section 8, one concern we face is that we do not observe the 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 discount discrimination. Here we write down a simple selection model to think about potential bias and present a trimming procedure for bounding the effect of officer differences in ticketing on our estimates of discrimination.

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

$$\begin{aligned} Y^*_{ij} &= \theta^W_j + \theta^B_j \cdot B_i + \epsilon_{ij} \\ Z_{ij} &= \alpha^W_j + \alpha^B_j \cdot B_i + \eta_{ij} \end{aligned}$$

Y^*_{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. We observe Y_{ij} if Z_{ij} crosses zero and the officer chooses to ticket the driver.

$$Y_{ij} = \begin{cases} \mathbf{1}(Y^*_{ij} \geq 0) & \text{if } Z_{ij} \geq 0 \\ \text{missing} & \text{otherwise} \end{cases}$$

When we estimate our probit/LPM, the object of interest is θ^B_j, each officer's degree of discrimination in discounting. Our estimation of θ^B_j is based on comparing treatment of white and non-white drivers *conditional on appearing in the data*:

$$\begin{aligned} & E[Y^*_{ij}|B_i = 1, Z_{ij} > 0] - E[Y^*_{ij}|B_i = 0, Z_{ij} > 0] \\ = \; & \theta^B_j + E[\epsilon_{ij}|\eta_{ij} > -\alpha^W_j - \alpha^B_j] - E[\epsilon_{ij}|\eta_{ij} > -\alpha^W_j] \end{aligned}$$

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

Ideally we could compare $E[Y^*_{ij}|B_i = 1, \eta_{ij} > -\alpha^W_j] - E[Y^*_{ij}|B_i = 0, \eta_{ij} > -\alpha^W_j]$, i.e. drivers

of both races who would be ticketed regardless of race. We present here a trimming procedure based on Lee (2009) that provides bounds on this difference.

We know that

$$\begin{aligned} & E[Y^*_{ij}|B_i = 1, \eta_{ij} > -\alpha^W_j - \alpha^B_j] \\ =p_j & \cdot E[Y^*_{ij}|B_i = 1, \eta_{ij} > -\alpha^W_j] \\ +(1-p_j) & \cdot E[Y^*_{ij}|B_i = 1, -\alpha^W_j > \eta_{ij} > -\alpha^W_j - \alpha^B_j] \end{aligned}$$

where

$$p_j = \frac{P(\eta_{ij} > -\alpha^W_j)}{P(\eta_{ij} > -\alpha^W_j - \alpha^B_j)}$$

In words, the observed treatment is a linear combination of treatment of marginal and inframarginal black drivers, where $1 - p_j$ is the share of marginal drivers.

The Lee (2009) approach is the following: we do not know $E[Y^*_{ij}|B_i = 1, \eta_{ij} > -\alpha^W_j]$, but we can bound it by assuming the marginal drivers are either at the bottom or top of the distribution $f(Y^*_{ij}|B_i = 1, \eta_{ij} > -\alpha^W_j - \alpha^B_j)$ and "trimming" them from the sample, where the amount that needs to be trimmed is $1 - p_j$.

In a traditional case of censored data, the researcher directly observes p_j, the differential degree of censoring across groups. In our setting, we do not observe p_j, because we do not see who does not appear in the data. To remedy this issue, we do the following: we take each officer's share of tickets issued to black (Hispanic) drivers and regress it on his discrimination against blacks (Hispanics), residualizing both for all driver characteristics and location and time of stop:

$$\text{Black}_i = b_0 + b_1 \cdot \hat{\theta}^B_j + X_{ij} \cdot \beta + \epsilon_{ij}$$

If we find that $b_1 > 0$ ($b_1 < 0$), we trim the pool of black drivers from officers with high (low) θ^B_j so that the relationship in the trimmed sample is zero. We trim observations from both the top and the bottom of the distribution of $f(Y^*_{ij}|B_i = 1, Z_{ij} = 1)$ to get our bounds. This procedure is meant to remove drivers that seem to appear in the data because the officer is discriminatory who would otherwise not appear in the data.

As shown in Section 5, we find no statistically significant relationship when this regression is aggregated to share of minority drivers and discrimination against all minorities.

Figure A-1 shows the results of the regression above, where the left panel is for black drivers and the right panel is for Hispanic drivers. We graph a nonparametric relationship between the variables but report the coefficient from a linear regression. Consistent with our previous results, there is almost no relationship, and the coefficient for black drivers is not significant. There is a small but statistically significant relationship between share hispanic drivers and Hispanic discrimination.

Figure A-2 shows the results of the trimming procedure that sets the above relationships to zero. We then rerun the difference-in-difference regressions from before for both the trimmed-above and trimmed-below samples, and the bounds are shown all possible values between these two estimates. For both black and Hispanic discrimination, the disparity in discounting persists after accounting for differential selection into the data.

Testing for Statistical Discrimination

We argue that the discrimination we find in traffic discounting is taste-based rather than statistical. In other words, there is no policing objective that is correlated with race that is driving the disparity in treatment. In this section, we present a method for testing whether officers are maximizing a certain objective, and the two outcomes we consider are future recidivism and court contestations. The approach borrows from the selection-on-gains test from Heckman et al. (2010), among others. The notation flips depending on the approach, so we will start with court contesting.

Officer j stops driver i for speeding and sets punishment $\pi_{ij} \in \{0, 1\}$, where 1 indicates the harsher ticket. Each individual is identified by his response to both ticket types, $\{Y_0(i), Y_1(i)\}$, which indicate the probability they will contest their ticket. An individual's probability of this action can be written as

$$Y(i) = Y_0(i) + \pi \cdot (Y_1(i) - Y_0(i)) \equiv Y_0(i) + \pi \cdot \Delta Y(i)$$

For simplicity, we assume drivers are drawn from a unit mass, where $i \in [0, 1]$, indicates where in the distribution a driver is drawn. Drivers are ordered by their responsiveness to receiving a ticket:

$$i < i' \Rightarrow \Delta Y(i) \geq \Delta Y(i')$$

i.e. $\Delta Y(i)$ is monotonically decreasing in the ordering of drivers.

The officer's objective is to maximize the total number of drivers who receive the harsh ticket, subject to a constraint on the amount of time he can be in court C. In practice, because of the ordering of drivers, he will choose some cutoff θ below which drivers are treated leniently and above which drivers are treated harshly. Because the objective is increasing in number of tickets, as is the constraint, the constraint will be binding and we can solve the problem by solving

$$\int_0^1 Y_0(v)dv + \int_{\theta^*}^1 \Delta Y(v)dv = C \tag{3}$$

In our data, we observe θ^* and C for each officer. The Marginal Treatment Effect (MTE) of lenience on court time at a given lenience can be calculated using the Implicit Function Theorem and Leibniz's rule:

$$MTE(\theta^*) \equiv \frac{\partial C}{\partial \theta^*} = -\Delta Y(\theta^*) \tag{4}$$

This fact tells us that the MTE across officers should be strictly declining in magnitude by officer lenience. In other words, the officer with the least lenience should have the most negative MTE, and the officer with the most lenience should have the least negative MTE.

The intuition for the result is the following: imagine an officer who is very lenient toward his drivers. If he is going to be harsh to one driver, it's going to be because they're not very responsive to a harsh ticket and will not contest. We will thus see that officer have a small MTE. Imagine an officer who is harsh toward nearly all drivers. If he is going to give a break to someone, it is because they would have a large response to it, and giving him a break would give him a large return in reduced court time. We should thus expect a large MTE for him.

The formulation for recidivism is similar, except that harsher tickets lead to less recidivism, so $\Delta Y(i) < 0$. In that case, the officer who is most lenient will have the most positive MTEs, and the officer who is least lenient will have the least positive MTEs.

To take this test to the data, we first check whether ticket harshness affects court contestation. Of course, each driver's ticket harshness is potentially endogenous, so we use the stopping officer's average lenience as an instrument for the driver's ticket. 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 consider whether a driver who encounters a harsh officer is more likely to contest her ticket. Figure A-4 reports the reduced form relationship between residualized officer lenience and contest rate, and we find a strong relationship between officer lenience and court contest rate.

More importantly for our analysis, we next consider whether the effect of ticketing harshness on driver contest rate varies by the level of officer lenience. To do so, we take the approach of Doyle (2007) and calculate each driver's probability of getting a break by regressing whether he receives a harsh ticket solely on the officer's residualized lenience, generating a driver-level propensity score. We then regress court contest rate on driver propensity score, separately by bins of propensity level. The result is reported in Figure

A-5.

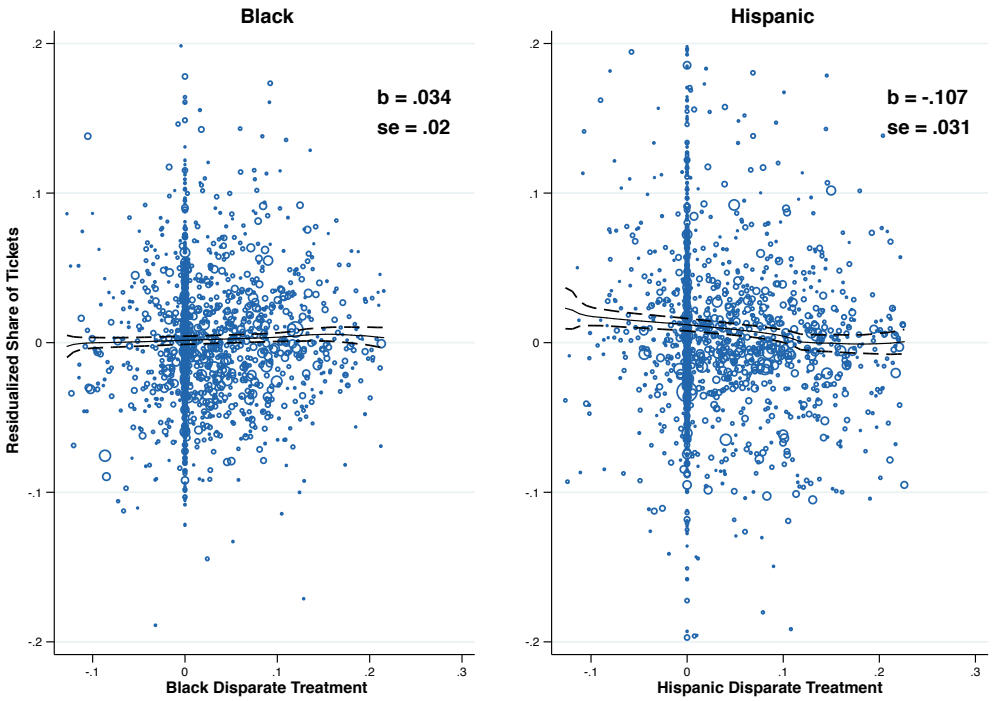
The prediction of our theory is that drivers encountering officers with low lenience will have the most negative treatment effect, and the drivers encountering the officers with high lenience will have the least negative treatment effect. We do not see this borne out in our analysis. While we reject that the treatment effects across bins are all the same, the treatment effects do not have a strict monotonic direction, nor do they appear to be generally increasing with officer lenience. This fact suggests that officers are not choosing ticket harshness so as to minimize court time.

We perform the same analysis for traffic deterrence, reported in Figures A-6 and A-7. As with contest rates, we find that recidivism responds to ticketing severity. A driver who encounters a lenient officer is more likely to receive a ticket in the following year relative to one encountering a harsh officer. The prediction of our framework is that, if officers are maximizing deterrence, the marginal treatment effect for lenient officers is least positive, and the marginal treatment effect for harsh officers is most positive. Again, we find that this prediction is not consistent with the data. If anything, it seems that officers who are harshest have the largest treatment effects.

We therefore conclude that officers are not ticketing optimally to minimize court time or maximize deterrence, ruling out that they are discriminating on the basis of these outcomes.

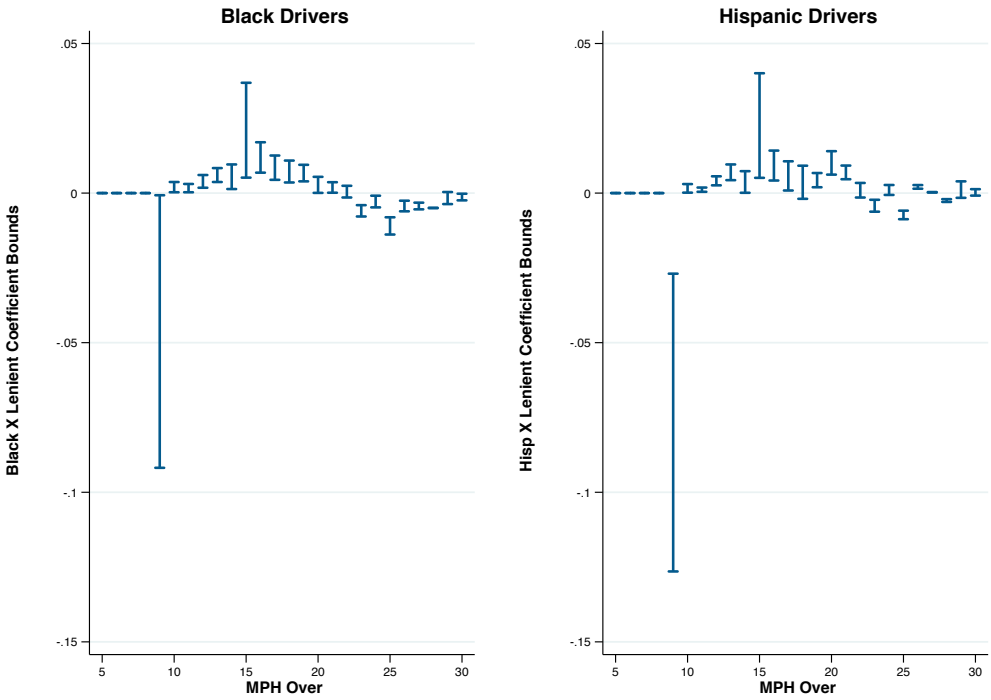
Appendix Figures

Figure A-1: Share Minority By Officer Discrimination, by Race



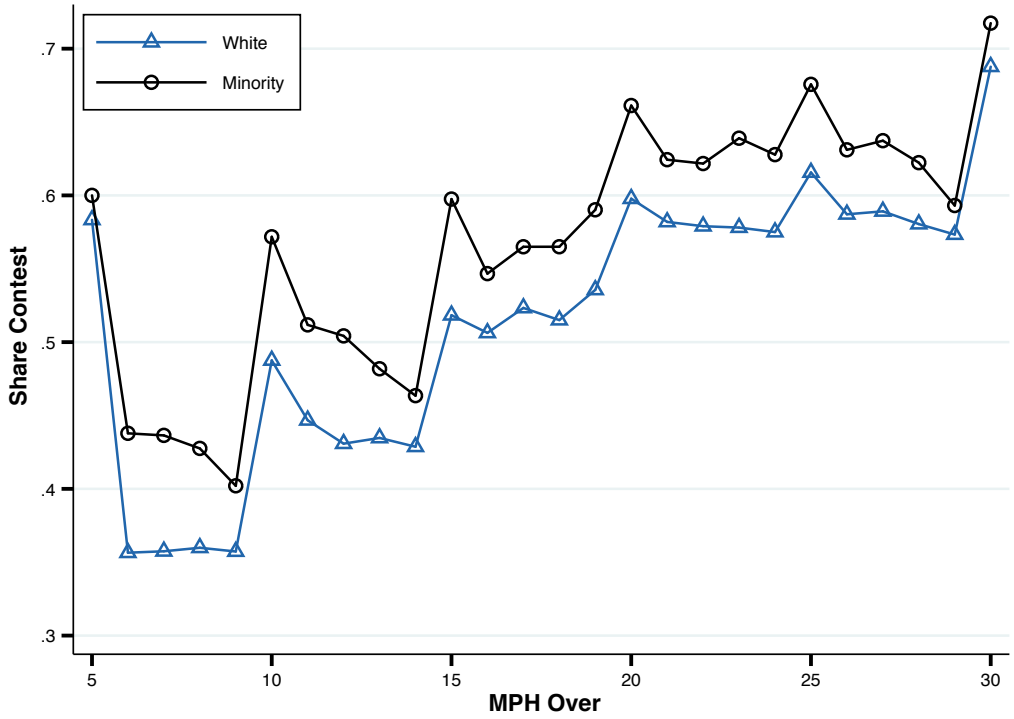
Notes: Left panel reports an officer's residualized share of tickets that are for black drivers against that officer's measure of discrimination. The same is reported on the right hand side, but for Hispanic drivers and discrimination against Hispanics.

Figure A-2: Trimmed Bounds on Difference-in-Difference Regressions



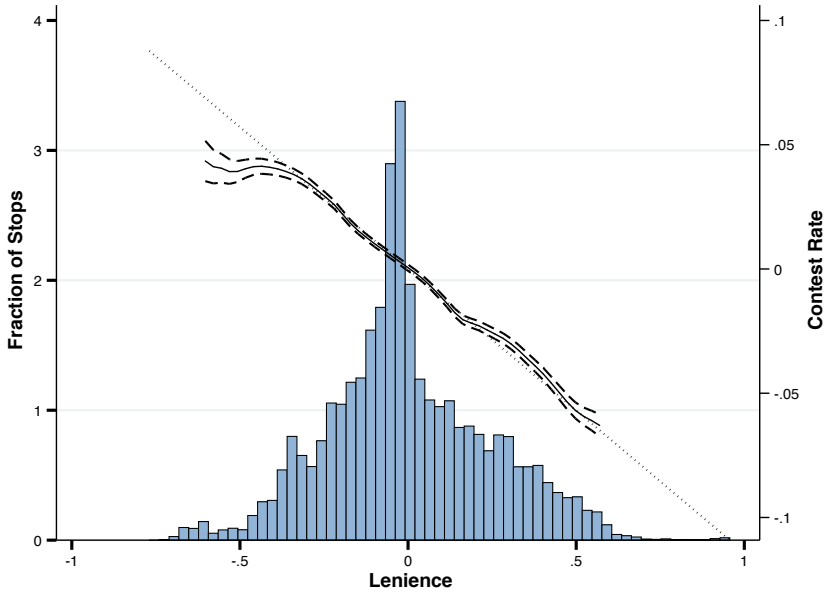
Notes: Figures report regression coefficients from regression where outcome is whether a driver is encountered at speed s and driver race Black_i and Hispanic_i are interacted with whether the officer is lenient. The left panel reports the interaction term for black drivers for each speed, and the right panel reports the interaction for Hispanic drivers.

Figure A-3: Court Contest Rate



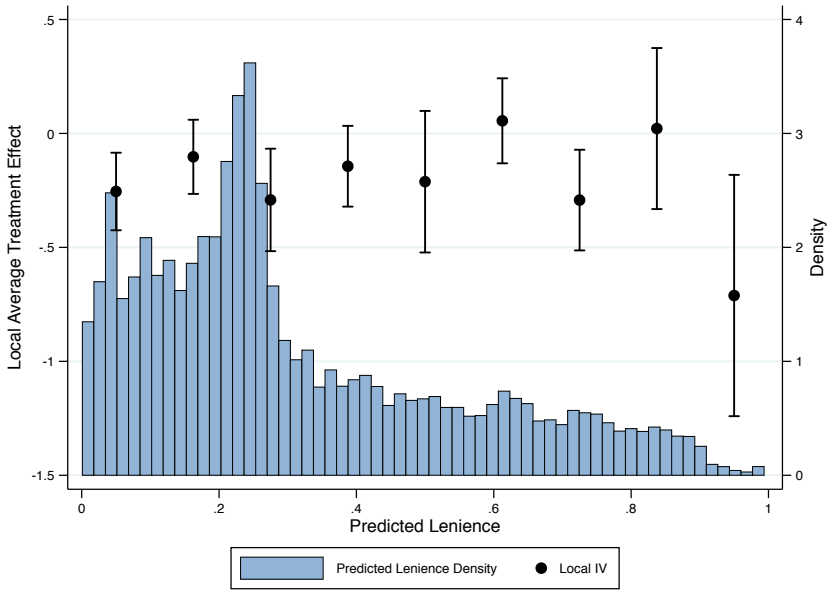
Notes: Figure reports the rate at which individuals contest their ticket, for each speed and race of driver.

Figure A-4: Officer Lenience and Court Contest Rate



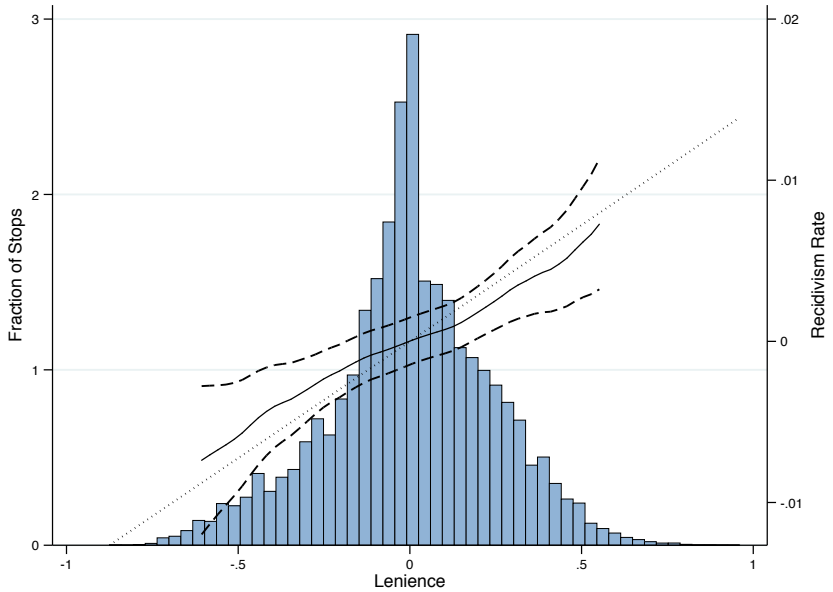
Notes: Figure reports the rate at which individuals contest their ticket, for each speed and race of driver.

Figure A-5: Marginal Treatment Effect of Officer Lenience on Court Contest Rate



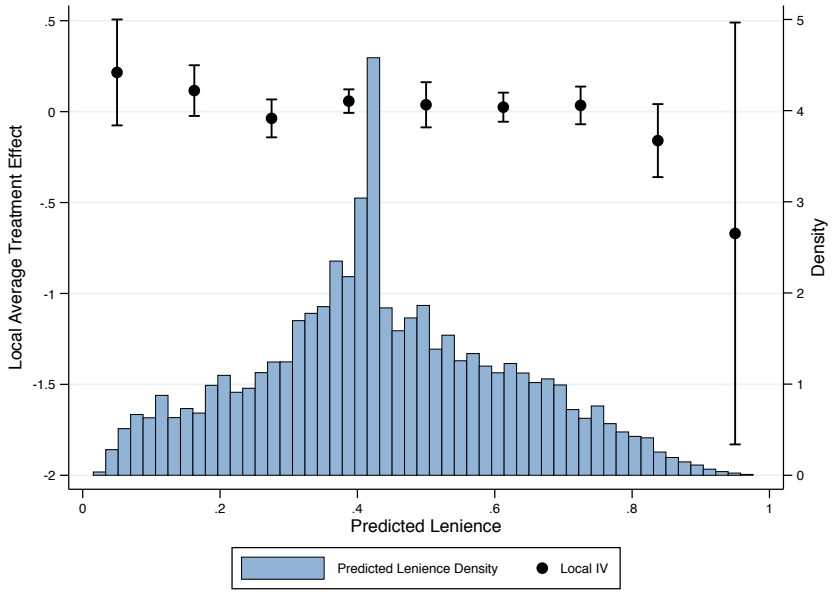
Notes: Figure reports the Appendix analysis of how the effect of officer lenience on contest rate varies by officer level of lenience.

Figure A-6: Officer Lenience and Recidivism Rate



Notes: Figure reports the rate at which individuals contest their ticket, for each speed and race of driver.

Figure A-7: Officer Lenience and Recidivism Rate



Notes: Figure reports the Appendix analysis of how the effect of officer lenience on recidivism rate varies by officer level of lenience.

Table A-1: Racial Disparity in Speeding

	Full Sample					GPS Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MPH Over	MPH Over	MPH Over	MPH Over	MPH Over	MPH Over	MPH Over
Driver Black	1.072***	0.809***	0.727***	0.636***	0.620***	0.897***	0.785***
	(0.267)	(0.0876)	(0.0845)	(0.0831)	(0.0762)	(0.0666)	(0.0750)
Driver Hispanic	2.765***	0.875***	0.791***	0.646***	0.650***	1.044***	0.785***
	(0.526)	(0.128)	(0.133)	(0.136)	(0.134)	(0.213)	(0.125)
Driver Female				-0.610***	-0.554***	-0.393***	-0.348***
				(0.0462)	(0.0408)	(0.0536)	(0.0540)
FL License				-0.179**	-0.350***	-0.683***	-0.521***
				(0.0804)	(0.0799)	(0.151)	(0.124)
Driver Age				-0.0439***	-0.0416***	-0.0379***	-0.0337***
				(0.00131)	(0.00125)	(0.00166)	(0.00208)
1 Prior Ticket					0.283***	0.287***	0.296***
					(0.0244)	(0.0457)	(0.0685)
2+ Prior Tickets					0.792***	0.686***	0.737***
					(0.0366)	(0.0715)	(0.0618)
Log Zip Code Income					0.122**	0.0879*	0.0402
					(0.0481)	(0.0466)	(0.0501)
Mean	16.554	16.554	16.554	16.554	16.554	15.979	15.979
Vehicle FE					X	X	X
Location FE		X				X	
Location + Time FE			X	X	X	X	
GPS FE							X
Observations	1125068	1125068	1125068	1125068	1125068	133927	133927

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)	(2)	(3)	(4)	(5)	(6)	(7)
	Discount	Discount	Discount	Discount	Discount	Discount	Discount
Driver Black	-0.0315*	-0.0270***	-0.0244***	-0.0220***	-0.0231***	-0.0386***	-0.0313***
	(0.0179)	(0.00452)	(0.00477)	(0.00482)	(0.00450)	(0.00679)	(0.00642)
Driver Hispanic	-0.143***	-0.0401***	-0.0390***	-0.0344***	-0.0355***	-0.0571***	-0.0395***
	(0.0331)	(0.00883)	(0.00937)	(0.00890)	(0.00880)	(0.0124)	(0.00844)
Driver Female				0.0282***	0.0262***	0.0188***	0.0178***
				(0.00416)	(0.00389)	(0.00374)	(0.00390)
FL License				0.00796*	0.0142***	0.0326***	0.0211**
				(0.00410)	(0.00433)	(0.00774)	(0.00864)
Driver Age				0.00134***	0.00127***	0.00124***	0.00103***
				(0.000232)	(0.000221)	(0.000199)	(0.000208)
1 Prior Ticket					-0.0122***	-0.0101***	-0.0125***
					(0.00260)	(0.00367)	(0.00456)
2+ Prior Tickets					-0.0289***	-0.0215***	-0.0269***
					(0.00577)	(0.00611)	(0.00667)
Log Zip Code Income					-0.00938***	-0.00506	-0.00203
					(0.00208)	(0.00373)	(0.00472)
Mean	.31	.31	.31	.31	.31	.326	.326
Vehicle FE					X	X	X
Location FE		X				X	
Location + Time FE			X	X	X	X	
GPS FE							X
Observations	1125068	1125068	1125068	1125068	1125068	133927	133927

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 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-3: Racial Disparity in Speeding, Non-lenient Officers

	Full Sample					GPS Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MPH Over	MPH Over	MPH Over	MPH Over	MPH Over	MPH Over	MPH Over
Driver Black	1.230***	0.701***	0.622***	0.503***	0.474***	0.751***	0.599***
	(0.184)	(0.160)	(0.158)	(0.144)	(0.138)	(0.154)	(0.204)
Driver Hispanic	1.578***	0.486***	0.416***	0.262***	0.258***	0.445***	0.401**
	(0.277)	(0.0676)	(0.0602)	(0.0655)	(0.0719)	(0.152)	(0.163)
Driver Female				-0.525***	-0.468***	-0.397***	-0.308***
				(0.0760)	(0.0694)	(0.102)	(0.0896)
FL License				-0.199**	-0.364***	-0.550***	-0.490**
				(0.0807)	(0.0833)	(0.160)	(0.197)
Driver Age				-0.0449***	-0.0424***	-0.0324***	-0.0319***
				(0.00253)	(0.00233)	(0.00259)	(0.00256)
1 Prior Ticket					0.265***	0.261***	0.180*
					(0.0321)	(0.0667)	(0.0935)
2+ Prior Tickets					0.708***	0.667***	0.617***
					(0.0431)	(0.113)	(0.141)
Log Zip Code Income					0.0248	0.0704	0.0576
					(0.0483)	(0.0740)	(0.0716)
Mean	20.377	20.377	20.377	20.377	20.377	19.99	19.99
Vehicle FE					X	X	X
Location FE		X				X	
Location + Time FE			X	X	X	X	
GPS FE							X
Observations	366277	366277	366277	366277	366277	32131	32131

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 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.