A Few Bad Apples? Racial Bias in Policing†

By Felipe Goncalves and Steven Mello*

We estimate the degree to which individual police officers practice racial discrimination. Using a bunching estimation design and data from the Florida Highway Patrol, we show that minorities are less likely to receive a discount on their speeding tickets than White drivers. Disaggregating this difference to the individual police officer, we estimate that 42 percent of officers practice discrimination. We then apply our officer-level discrimination measures to various policy-relevant questions in the literature. In particular, reassigning officers across locations based on their lenience can effectively reduce the aggregate disparity in treatment. (JEL H76, J15, K42)

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 2015), convicted of a crime (Anwar, Bayer, and Hjalmarsson 2012), denied bail (Arnold, Dobbie, and Yang 2018), and issued a lengthy prison sentence (Rehavi and Starr 2014) relative to observably similar Whites. In light of these disparities, a literature has developed to test whether these outcomes can be explained by discrimination on the part of police officers, judges, and other criminal justice agents (Knowles, Persico, and Todd 2001; Anwar and Fang 2006; Grogger and Ridgeway 2006; Antonovics and Knight 2009; Persico 2009; Abrams, Bertrand, and Mullainathan 2012; Horrace and Rohlin 2016; Fryer 2019; Arnold, Dobbie, and Yang 2018). The view that discrimination is responsible for these disparate outcomes has gained traction in recent years, particularly within minority communities, following several highly publicized police killings of minorities. A 2013 Gallup poll found that one-half of Black adults agreed that racial
differences in incarceration rates are “mostly due to discrimination,” while only 19 percent of White respondents agreed.1

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

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

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

As shown in Figure 1, the distribution of speeds ticketed by the Florida Highway Patrol between 2005 and 2015 shows substantial excess mass at speeds just below the

---

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.”(https://www.washingtonpost.com/graphics/2017/investigations/police-fired-rehired/).
3 This practice is similar to teachers’ bunching up of grades on high-stakes exams (Dee et al. 2019, Diamond and Persson 2016).
first fine increase, where speeds are reported relative to the speed limit. Meanwhile, a remarkably small portion of tickets are issued for speeds just above. We take this bunching as evidence that officers systematically manipulate the charged speed, commonly charging speeds just below fine increases after observing a higher speed, perhaps to avoid an onerous punishment for the driver. However, when disaggregated by driver race in Figure 2, we see that minorities are significantly less likely to be found at the bunch point.

The first task of this paper is to confirm that this disparity is evidence of officer 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. While our data record the speed that is charged on a ticket, we do not observe the true stopped speed of the drivers in our data. To deal with this challenge, we use the fact that one-third of officers practice no lenience. Namely, they exhibit no bunching in their distribution of ticketed speeds. For these officers, we argue that their distribution of ticketed speeds reflects the true distribution of driven speeds among stopped and ticketed drivers. We show that, conditional on location and time, driver characteristics are not predictive of whether the officer he encounters is lenient. Non-lenient officers do not write fewer tickets than lenient officers, and a similar share of their tickets are for speeding offenses. These facts suggest that lenient and non-lenient officers are pulling over similar types of drivers, and thus non-lenient officers can be used to identify the “true” distribution of speeds. The speed distributions for non-lenient officers indicate that, while minorities are driving slightly faster speeds

\[ \text{Notes: Connected line shows histogram of tickets. Dashed line plots fine schedule for Broward County. Felony speeding is 30 mph over and carries a fine to be determined following a court appearance.} \]
when stopped, the racial gap in speeding is minimal. We therefore argue that our estimates of differences in discounting can be interpreted as evidence of discrimination. Further, we provide a simple approach to adjusting our estimates to account for any racial differences in speeding.

Using a difference-in-differences framework, we then find that White drivers differentially benefit from being stopped by a lenient officer. White drivers stopped by lenient officers are six percentage points more likely to be discounted than minority drivers off a base of 42 percent. This gain stems from the fact that minorities are treated less leniently when stopped for speeds ranging from 10 to 22 mph over the limit.

The central contribution of our paper is to further provide an estimate of the discrimination of each individual officer. Specifically, we compute an officer’s lenience toward minorities relative to his own treatment of White drivers, differencing out the treatment of each race by non-lenient officers and adjusting for other features of the stop, and treat that difference as the officer’s discrimination. Disaggregating to the officer level reveals significant heterogeneity in the degree of discrimination. An officer at the ninetieth percentile of discrimination is more than 10 percentage points more likely to discount a White driver than a minority driver. The modal officer practices no discrimination, and 42 percent of officers explain the entirety of the aggregate disparity in treatment. Correlating officer-level discrimination to demographics, we find that minority and female officers tend to practice less discrimination than other officers.

The remainder of the paper exploits our officer-level measures of lenience and discrimination to understand the mechanisms that lead to the disparity in

![Figure 2. Charged Speed Distributions by Driver Race](image-url)

*Notes:* Connected line shows histogram of ticketed speeds, separately by driver race; 34.3 percent of tickets to white drivers are given at 9 mph over compared to 25.2 percent of tickets for minority drivers.
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 81 percent of the gap in discounting. Only 19 percent of the gap is due to minorities driving faster. Across locations, a large share of the disparity in treatment is due to the fact that minorities drive in areas where officers are less lenient to all motorists.

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

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

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

Our approach falls broadly into a literature using benchmarking procedures, where the behavior of one agent is compared to a proposed control group, to identify discrimination. In the paper most closely related to ours, Anbarci and Lee (2014) study

---

5 See Ridgeway and MacDonald (2010) for a review of the benchmarking literature.
the discounting behavior of traffic officers using a benchmarking design and find that at least one racial group of officers is biased in favor of their own race. To our knowledge, the only existing study aiming to identify discrimination by individual criminal justice agents is Ridgeway and MacDonald (2009), who compare the racial makeup of NYPD officers’ stops and frisks to those of nearby officers and are able to identify a set of officers with a disproportionately high share of minority stops.

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

Methodologically, our approach builds on recent research using “bunching” estimators to recover behavioral parameters (Kleven 2016). In contrast with most applications (Chetty et al. 2011, Saez 2010), which infer a hypothetical true distribution by examining areas away from the manipulation region, our approach is similar to Best et al. (2020) 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 I provides institutional background on the Florida Highway Patrol and describes the data. Section II presents a conceptual framework, and Section III describes our empirical strategy. Section IV presents the central findings, and Section V considers specification checks and alternative interpretations of our results. In Section VI, we present and estimate a model of officer behavior and perform counterfactuals, and Section VII concludes.

I. Institutional Background and Data

A. Institutions of the Florida Highway Patrol

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

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

Patrol officers in the FHP are divided into one of nine 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

6 The actual fine schedule depends on the county in Florida, though 10 mph over always includes at least a $50 increase in the fine. The point schedule is identical statewide and does not contain a jump at 10 mph over.
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. Officers do, however, potentially have a promotion incentive to write a certain number of tickets, as the number of tickets they write appears on their performance evaluations. We believe these set of institutional factors contribute to an environment in which officers are encouraged to write tickets but also have the freedom to write reduced charges, which is ideal for our research design.

While all speeding beyond 5 mph over the limit commands a statutory fine, the evidence suggests that drivers are not regularly pulled over for less than 10 mph over, and the data show very few tickets for 8 mph over and 10 mph over. As we will reiterate in Section III, 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.

B. Data

From the Florida Court Clerks and Comptrollers, we obtained data on all traffic citations issued in Florida for the years 2005–2015 (FCCC 2016b). These data include all information provided on the stopped motorist’s driver’s license—name, address, race, gender, height, and date of birth, as well as driver’s license state and number. The make and year of the stopped automobile is provided in 79 percent of tickets, and we link this information to an online database of vehicle price estimates (TrueCar.com 2017). While we see the speed charged on the citation, we do not see the original speed of the stop. We also do not see stops and interactions that do not result in a traffic citation.

The citing officer is identified by name, rank, troop number, and badge number. To supplement the citations data, we obtained officer demographic information from the Florida Department of Law Enforcement (FDLE 2016). These data include officer race, sex, age, education level, and the Florida law enforcement employment history of all law enforcement officers employed in the State of Florida. We focus on stops by the Florida Highway Patrol (FHP), both because officers are more consistently identified in the data in FHP stops and because traffic enforcement is the primary responsibility of FHP.

While the citations record the driver race, there appear to be inconsistencies in the recording of Hispanic. For example, Miami-Dade County issues fewer than 1 percent 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 (Census

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

8 The problem of only seeing interactions that lead to enforcement is common in the discrimination literature. We discuss this point further in Section VA and online Appendix Section F. For a recent paper that addresses this issue directly, see West (2018).
If an individual in our data has a name that is more than 80 percent Hispanic, we record them as such.

Using the driver’s license number, we are able to link individuals across tickets. Doing so, we construct for each ticket a measure of the number of tickets received in the previous three years, including from non-FHP stops. Using the driver’s date of birth and full name, we also link each individual to prison spell records from the Florida Department of Corrections and construct an indicator for any past incarceration (FDOC 2019). At the point of a stop, the officer is able to see a driver’s full criminal history, including arrest history that does not lead to prison, so we consider tickets and incarceration as the best available approximation to an individual’s criminal history at the point of the stop.

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

The final sample includes 2,122,555 citations issued by 1,851 officers, from an initial sample of 3,864,034 speeding citations. The two most binding restrictions are requiring that race be specified (80 percent of tickets) and requiring that the officer be linkable to the FDLE (73 percent). In online Appendix Section A we include a table that documents the sample reduction from each restriction we make. In all of our analyses, we consider speed relative to the speed limit (or posted speed) rather than absolute speed. We often refer to this quantity as mph over or simply as “the speed.”

Beginning in 2013, about 40 percent of tickets are geocoded with the latitude and longitude of a stop (271,164 observations). Using ArcGIS, we link the geocoded tickets to road “segments,” which are on average 6.7 miles long and roughly correspond to entire streets within cities and uninterrupted stretches of road on interstates and highways. Throughout the analysis, we also provide results for the restricted sample of geocoded tickets with corresponding fixed-effects at the road-segment level. The road-segment analysis allows us to consider a more granular comparison of drivers.

C. Summary Statistics

Table I presents summary statistics for the sample, broken out by driver race. 59 percent of drivers are White, 18 percent are Black, and about 23 percent are Hispanic. Drivers are 37 percent female and about 36 years old on average, with Hispanics less likely to be female and minority drivers typically younger. In-state

Data provided by FDOT (2020). See online Appendix A for further details.
drivers account for 84 percent of tickets. The average driver has been cited about 0.53 times in the past year, though Black and Hispanic drivers have 0.19 more prior tickets than White drivers. On average, minority drivers are charged with higher speeds than Whites: just over 1 mph higher for Blacks and almost 3 mph higher for Hispanics. Consistent with Figures 1 and 2, drivers of all races have a high probability of being ticketed at 9 mph over the limit, which is just below the first jump in the fine schedule. However, minority drivers are also less likely to be charged this

Table 1—Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>White (1)</th>
<th>Black (2)</th>
<th>Hispanic (3)</th>
<th>All (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Tickets</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.37</td>
<td>0.42</td>
<td>0.32</td>
<td>0.37</td>
</tr>
<tr>
<td>(0.48)</td>
<td>(0.49)</td>
<td>(0.47)</td>
<td></td>
<td>(0.48)</td>
</tr>
<tr>
<td>Age</td>
<td>37.59</td>
<td>34.65</td>
<td>34.59</td>
<td>36.38</td>
</tr>
<tr>
<td>(14.96)</td>
<td>(12.37)</td>
<td>(12.12)</td>
<td></td>
<td>(13.98)</td>
</tr>
<tr>
<td>FL license</td>
<td>0.82</td>
<td>0.84</td>
<td>0.89</td>
<td>0.84</td>
</tr>
<tr>
<td>(0.39)</td>
<td>(0.37)</td>
<td>(0.32)</td>
<td></td>
<td>(0.37)</td>
</tr>
<tr>
<td>Zip code income</td>
<td>64.95</td>
<td>45.71</td>
<td>53.18</td>
<td>58.81</td>
</tr>
<tr>
<td>(49.97)</td>
<td>(26.68)</td>
<td>(38.92)</td>
<td></td>
<td>(44.86)</td>
</tr>
<tr>
<td>Vehicle price/1,000</td>
<td>18.82</td>
<td>17.94</td>
<td>19.01</td>
<td>18.70</td>
</tr>
<tr>
<td>(9.73)</td>
<td>(8.44)</td>
<td>(9.71)</td>
<td></td>
<td>(9.51)</td>
</tr>
<tr>
<td>Citations in past year</td>
<td>0.45</td>
<td>0.64</td>
<td>0.64</td>
<td>0.53</td>
</tr>
<tr>
<td>(0.95)</td>
<td>(1.17)</td>
<td>(1.15)</td>
<td></td>
<td>(1.05)</td>
</tr>
<tr>
<td>mph over</td>
<td>15.40</td>
<td>16.43</td>
<td>18.11</td>
<td>16.21</td>
</tr>
<tr>
<td>(6.45)</td>
<td>(6.96)</td>
<td>(6.94)</td>
<td></td>
<td>(6.75)</td>
</tr>
<tr>
<td>Discount</td>
<td>0.35</td>
<td>0.32</td>
<td>0.21</td>
<td>0.31</td>
</tr>
<tr>
<td>(0.48)</td>
<td>(0.47)</td>
<td>(0.41)</td>
<td></td>
<td>(0.46)</td>
</tr>
<tr>
<td>Imputed fine</td>
<td>179.10</td>
<td>186.50</td>
<td>197.87</td>
<td>184.67</td>
</tr>
<tr>
<td>(75.76)</td>
<td>(78.74)</td>
<td>(77.64)</td>
<td></td>
<td>(77.10)</td>
</tr>
<tr>
<td>Share</td>
<td>0.59</td>
<td>0.18</td>
<td>0.23</td>
<td>1</td>
</tr>
<tr>
<td>Observations</td>
<td>1,253,149</td>
<td>379,268</td>
<td>490,138</td>
<td>2,122,555</td>
</tr>
</tbody>
</table>

| **Panel B. Officers** |       |           |              |         |
| Female               | 0.08   | 0.14      | 0.09         | 0.09    |
| (0.27)               | (0.35) | (0.28)    |              | (0.29)  |
| Bachelor’s degree    | 0.52   | 0.60      | 0.36         | 0.51    |
| (0.50)               | (0.49) | (0.48)    |              | (0.50)  |
| Total tickets        | 1,145.50 | 1,332.72 | 954.66       | 1,146.71 |
| (1,282.75)           | (1,606.03)| (1,653.01)|              | (1,473.05)|
| Yearly tickets       | 164.63 | 189.21    | 163.50       | 169.62  |
| (158.68)             | (189.01) | (188.02) |              | (174.99)|
| Tenure               | 11.99  | 12.19     | 7.30         | 11.00   |
| (10.13)              | (9.82) | (7.21)    |              | (9.72)  |
| Share                | 0.63   | 0.16      | 0.20         | 1       |
| Observations         | 1,168  | 292       | 364          | 1,851   |

Notes: Standard deviations in parentheses. Zip code income is missing for 3.0 percent of White stops, 2.7 percent of Black stops, and 4.1 percent of Hispanic stops. Vehicle information is missing for 26.1 percent of White stops, 26.4 percent of Black stops, and 34.2 percent 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. In all tables, Discount is an indicator for a charged speed of 9 mph over the posted limit.
speed. As we show in online Appendix Tables A.1 and A.2, these disparities in speed and ticketing below the jump persist after controlling for all stop characteristics and time and location fixed effects.

A notable feature of the distribution of tickets is the heaping of charged speeds at multiples of five above the bunch point. This heaping occurs because, in many instances, officers do not use a radar gun, and their recording of the speed may be approximate. In cases where a “method of arrest” is recorded, 74 percent of tickets indicate the officer observed the speed through “visual inspection,” and the remaining 26 percent report the use of a radar or laser gun. We report in online Appendix Figure A.1 the distribution of ticketed speeds for this latter subsample, where we find no heaping at multiples of five.\(^{10}\)

The bottom panel of Table 1 presents summary statistics for the set of officers in our sample. We note here the large number of tickets written by the average officer (1,147) and the fact that the population of officers is slightly more White (63 percent) than our sample of drivers (59 percent).

II. 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 III and our modeling exercise in Section VI.

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

In this framework, it is natural to define discrimination in the following way. We say that officer \(j\) is discriminatory at speed \(s^*\) if \(P_{jw}(s^*) > P_{jm}(s^*)\). While we describe the officers’ preferences as potentially reflecting bias, we are not yet taking a stand on whether any disparity in treatment is taste-based versus statistical. For example, it is possible that some officers prefer Whites because they believe the likelihood of having to go to court later is lower. In Section V, we discuss how to differentiate between taste-based and statistical discrimination. Note further that we define discrimination specifically through action. If an officer exhibits

\(^{10}\)In Table 4, we also show that our main result is not changed when restricting attention to this subsample.
\( P_{jr}(s^*) = 0 \) for all \( s^* \) and \( r \), we treat him as not discriminatory by definition, despite the potential for internal animus. We discuss this issue further in Section III.

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

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

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

\[
Pr(S_i = s > s_d | i, j) = f_r(s^*) \cdot \left(1 - P_{jr}(s^*)\right),
\]

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

### III. Empirical Strategy

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

In Figure 3, we motivate this approach by documenting the significant heterogeneity in discounting across officers. Panel A plots the officer-level distribution of lenience, defined as the share of tickets written for 9 mph or above that are for exactly 9 mph. A large share of officers appear to exhibit very little lenience, with 30 percent writing fewer than 1 percent of tickets for this bunching speed. Further, this heterogeneity across officers cannot be fully explained by the locations and times when they are patrolling. Panel B plots the share of tickets in each county and shift that is written by officers for whom fewer than 2 percent of their tickets are for 9 mph over. Of the 247 county-shifts in our data, 226 have a share strictly between 0 and 1, indicating significant overlap in the patrolling of lenient and non-lenient officers.\(^{11}\)

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

\(^{11}\) Another way to see the importance of officer behavior in generating the observed bunching is to consider the explanatory power of various predictors of a discounted charge. In a regression of a discount indicator on officer fixed effects and location-time fixed effects (described below), the \( R^2 \) attributable to the officer effects is 0.31, while that attributable to the location-time effects is 0.26. This finding provides further evidence that officers exhibit significant variability in discount behavior after accounting for location and time.
lenience in his year with the most stops (x-axis), where we residualize county and month-of-stop fixed effects and driver characteristics. A strong correlation is evident: an officer who charges 9 mph relatively more often in one year also does so in other years. In panel D, we plot residualized lenience in the county where the officer has made the second most stops against residualized lenience in the county where he has made the most stops, confirming that officer lenience is highly correlated over space.

To identify the officers who practice no lenience, we use the Frandsen (2017) test for manipulation in bunching. In our setting, this test implies that, under the null hypothesis of no manipulation, the conditional probability of being found at the bunch point is in a range around one-third, \( \Pr(X = 9|X \in [8,10]) \in \left[(1 - k)/(3 - k), (1 + k)/(3 + k)\right] \), where \( k \) is a restriction on the second finite difference, \( \Delta^{(2)} \Pr(S = 9) \equiv \Pr(S = 8) - 2\Pr(S = 9) + \Pr(S = 10) \), such that \( |\Delta^{(2)} \Pr(S = 9)| \leq k(\Pr(S = 8) - \Pr(S = 10)) \). Intuitively, if the distribution of ticketed speeds is unmanipulated, the share of tickets at 9 mph over among those charged between 8 and 10 mph over should be approximately one-third, where the deviation \( k \) is due to curvature in the distribution of speeds. We calculate \( k \) by assuming the distribution \( \Pr(S) \) is Poisson and estimate the mean parameter \( \lambda \) using the
empirical mean of ticketed speeds. We say that an officer is non-lenient if we fail to reject that Pr(S = 9 | S ∈ [8, 10]) ≤ (1 + k)/(3 + k) at the 99 percent confidence level. Out of 1,851 officers, we identify 562 as non-lenient.

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

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

\[(3) \quad S^k_{ij} = \beta_0 + \beta_1 \cdot \text{White}_i + \beta_2 \cdot \text{Lenient}_j + \beta_3 \cdot \text{White}_i \cdot \text{Lenient}_j + X_i \gamma + X_i \cdot \text{Lenient}_j \alpha + \epsilon_{ij}.\]

For all regressions, the primary coefficient of interest is \(\beta_3\), the interaction between White driver and lenient officer. For the bunch point of 9 mph over the limit, \(\beta_3\) 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. The term \(X_i\) contains the set of all observable characteristics of the drivers, including gender, age, age squared, number of previous tickets, any prior incarceration, whether the driver is in-state, the log average income of the driver’s home zip code, vehicle age and age squared, estimated vehicle price, and indicators for vehicle make (where missing is its own category). We also include interactions between the indicator for officer lenient and the above driver covariates (excluding vehicle make fixed effects) to isolate lenience based on race from lenience based on other characteristics that are potentially correlated with race.

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

To calculate each officer’s individual discrimination coefficient, we take a similar approach and use non-lenient officers as a control for the baseline frequency of

---

12 An alternative approach is to follow up with a second step where \(\lambda\) is calculated using the non-lenient officers, and then rerun the test of manipulation for each officer. In practice, this approach leads to no difference in which officers are labeled as lenient.

13 All of our results are qualitatively similar if we instead identify lenient officers as those with 2 percent or more of tickets at 9 mph over.

14 Acquired from IRS (2016).
tickets at 9 mph over, but we allow the coefficients for $\text{Lenient}_j$ and $\text{White}_i \cdot \text{Lenient}_j$ to vary by individual officer:

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

The coefficients of interest, $\beta_3^j$, are identified from each officer’s difference in discounting between Whites and minorities, differencing out the disparity in ticketing for non-lenient officers. We denote $\beta_3^j$ as officer $j$’s degree of discrimination.

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

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

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

In Figure A.2, we evaluate how the characteristics of an officer’s stops vary with whether the officer is lenient or not, where both variables have been residualized with location-time fixed effects. The first three panels show that an officer’s lenience is uncorrelated with the share of tickets written to minorities, uncorrelated with the share of tickets where race is missing, and only minimally correlated with the share of tickets that are for speeding. The final panel shows that an officer’s lenience is not predictive of the number of tickets written per day. For this figure we calculate both measures at the annual level, during which officers write most of their tickets in one county, allowing us to control for county-by-year fixed effects.
To further test for selection on observables, Table 2 estimates how officer lenience varies with driver characteristics. Column 1 reports a regression of whether a driver is ticketed on location-by-time fixed effects and driver characteristics. The joint $F$-test of the null that all driver coefficients are zero has an $F$-value of 34.5 and a $p$-value 0.000. When the outcome is changed to the officer’s lenience indicator in column 2, the $F$-value declines to 0.798, and the $p$-value increases to 0.673, consistent with the view that officer-driver matches are conditionally random. Columns 3 and 4 replicate column 2 for our GPS sample. The $F$-value for both columns is similarly small, though the inclusion of fixed effects for the road-segment of the stop leads to a marginally significant $p$-value of 0.021.\footnote{In online Appendix Table A.7, we present a similar check for randomization, but using the officer’s estimated discrimination coefficient as the outcome variable. Although we reject the null hypothesis of no randomization for the full sample, the $F$-value (2.57) is small. In the GPS sample, we have a similarly small $F$-value (1.86) but marginally significant $F$-test using our standard location-time fixed effects and an insignificant $F$-test ($F$-value 1.14) after the inclusion of road-segment fixed effects.}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.pdf}
\caption{Difference-in-Differences Raw Data Plot}
\end{figure}

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

A. Interpreting the Difference-in-Differences Coefficient

The difference-in-differences regression coefficient $\beta_3$ in equation (3) for 9 mph over reflects the difference in probabilities of discounting between White and minority drivers among lenient officers relative to non-lenient officers. The use of non-lenient officers as a control group is designed to account for the possibility of different driving speeds across races. In practice, however, this approach only corrects for differences in the likelihood of driving exactly 9 mph over the limit (as we discuss in detail below). Drivers may differ in their speed distributions above 9 mph over, possibly leading to a bias in our difference-in-differences measure of discrimination. We document here how our estimates may be biased by racial differences in speeding and how to correct for such a bias.
Using the simple model from Section II, we can decompose this coefficient into a weighted average of discrimination at different speeds and a term that reflects racial differences in speeding:

\[
\beta_3 = \left[ \Pr(S_i = s_d | W_i, \text{Lenient}_j) - \Pr(S_i = s_d | W_i, \text{Non-Lenient}_j) \right] \\
- \left[ \Pr(S_i = s_d | M_i, \text{Lenient}_j) - \Pr(S_i = s_d | M_i, \text{Non-Lenient}_j) \right] \\
= \sum_{s > 9} [f_w(s^*) \times \left[ \Pr_w(s^*) - \Pr_m(s^*) \right]] + \sum_{s > 9} \left[ (f_w(s^*) - f_m(s^*)) \times \Pr_m(s^*) \right].
\]

The first term in the last equation is a weighted average of discrimination among lenient officers at each speed, where the weights are the speed distribution for
White drivers, and is our object of interest. The second term represents a bias that is non-zero when White and minority drivers have different speed distributions. In other words, our estimate of the difference-in-differences coefficient will deviate from the true discrimination coefficient we care to identify when drivers differ in underlying speed in such a way that the second summation above differs from zero.

Using the drivers stopped by non-lenient officers, we can directly measure each race’s speed distribution, which we plot in Figure 5. The left panel controls for the time and location of stops, and the right panel additionally controls for all other demographic characteristics. The speed distributions are remarkably similar across races, with average speed gaps of 0.33 and 0.50 mph over with and without demographic controls, respectively. In online Appendix Table A.3, we show that this small gap is consistent across various specifications for time and location controls.

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

To estimate \( f_w(s^*) - f_m(s^*) \) for each speed \( s^* \), we run a regression among the non-lenient officer sample where the outcome is that the driver is ticketed speed \( s^* \) and includes location and time fixed effects, driver covariates, and an indicator for driver race being White. We take the coefficient on White driver to be our estimate for

\[ f_w(s^*) - f_m(s^*) \]
the difference in density. To estimate $P_{rm}(s^*)$, we use that the probability a minority driver is ticketed at $k$ is $f_m(s^*)$ for non-lenient officers and $(1 - P_{rm}(s^*)) \cdot f_m(s^*)$ for lenient officers. We restrict attention to minority drivers and regress whether the driver is ticketed speed $s^*$ on whether the officer is lenient, in addition to shift-time fixed effects and driver covariates. Our estimate for $P_{rm}(s^*)$ is then derived by dividing the negative of the coefficient on lenient officer by the statewide share of minority drivers at $s^*$ for non-lenient officers (our estimate of $f_m(s^*)$).

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

### IV. Results

Table 3 reports the results of our difference-in-differences test of discrimination where the outcome of interest is whether the driver is ticketed at the discount speed. Columns 1 and 2 show estimates from a specification with location-time fixed effects only and our preferred specification that includes location-time fixed effects, covariates, and covariates interacted with lenience. The coefficient on the interaction between White drivers and lenient officers indicates that White drivers are 5.8 to 6.5 percentage points more likely to receive a discount than minority drivers, relative to a mean probability of 41.6 percent for lenient officers. Columns 3 and 4 show that our discrimination estimates persist, but shrink slightly to 4.6 percentage points, when adjusting for stretch-of-road fixed effects using our GPS sample. These coefficients indicate how much more likely a lenient officer is to discount a White driver. To calculate a differential probability of discount by an average officer, we use the fact that three-fourths of tickets are written by lenient officers and scale accordingly, finding that an average encounter leads to a 3.4 to 4.9 percentage point higher discount probability for White drivers, off a base of 31 percent.
Worth noting is the statistical similarity between the discrimination coefficients in columns 1 and 2 of Table 3. This finding allows us to rule out a form of “discrimination by proxy.” In other words, the racial disparity in treatment we observe is not driven by lenience towards other characteristics that are correlated with race.

Implementing the correction procedure from Section IIIA to these estimates, we find the degree of bias in our difference-in-differences coefficient induced by racial differences in speeding to be quantitatively small. Our point estimate for the bias expression in equation (5) is 0.0065 (standard error 0.000617), which is 11.1 percent of our baseline coefficient. This measure of bias indicates to us that our difference-in-differences estimator is not meaningfully different from the true behavioral parameter we would like to identify, and the small gap is due to the fact that White and minority drivers have very similar speeds in our data. However, the above correction allows for our approach to be used in other settings where the distribution of driving speed (or other measure of criminality) varies substantially across races.

The graphical analog to Table 3 is presented in Figure 6. Here we report the coefficients on interactions between White driver and lenient officer from regressions where an indicator for being charged speed $s$ (x-axis) is the outcome of interest. The two specifications correspond to columns 1 and 2 of Table 3. These coefficients indicate where minority drivers are disproportionately being ticketed, and thus the speeds at which White drivers are being differentially discounted. The interaction coefficient is negative and significant for almost all speeds between 10 and 22 mph, suggesting that at these speeds minorities are less likely to receive a break. Note from equation (5) that, under equal speed distributions across races, our diff-diff coefficient for 9 mph over is equal to a weighted sum of coefficients for higher speeds. Because of the insignificant estimate we find for the bias component from speed differences, we think of the coefficient for 9 mph over approximately reflecting this weighted sum of higher coefficients and conclude that the majority of the disparity in discounting is driven by discrimination at 10 to 22 mph over.

A. Officer-Level Heterogeneity

Officer-level results from estimating equation (4) are reported in Figure 7. The figure displays the across-officer distribution of the interaction coefficient $\hat{\beta}_3$, where non-lenient officers are assigned $\hat{\beta}_3 = 0$. The line represents a kernel density plot of our measure of discrimination against minority drivers, so that the farther right an officer is in the distribution of discrimination, the greater his level of discrimination. The unit of our measure is difference in percentage points probability of discounting. An officer whose discrimination against minorities is 0.1, for example, is 10 percentage points more likely to offer a fine reduction to a White than a minority driver. The percentiles of officer discrimination are also reported in online Appendix Table A.5.

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 tenth and ninetieth percentiles

---

17 We calculate bootstrapped standard errors, where we draw random sets of observations from our data with replacement, calculate the statistic, and iterate 100 times.
of discrimination have a 11.0 percentage point difference in their racial disparity. Relative to the overall rate of discounting, officers at the ninetieth percentile are over 35 percent less likely to discount minority drivers.

The second notable fact is that the average level of discrimination is quite small, three percentage points off of a base of 31 percent of tickets at 9 mph over. While this disparity is comparable to the Black-White wage gap (Neal and Johnson 1996), it is possible that the officer in question is not aware of such a disparity. A large literature has explored the role of implicit bias as a source of discrimination (Greenwald and Krieger 2006; Banks, Eberhardt, and Ross 2006), and in many cases the individual in question is not aware of his bias. The level of discrimination of an average officer is 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.

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 $\hat{\beta}_3^{ij}$ is 0.008—less than one-seventh the standard deviation of $\hat{\beta}_3^{ij}$ across officers, 0.061. In the scenario in which true discrimination is uniform, these numbers would be similar in magnitude. We thus conclude that the majority of the variation is due to true officer differences in discrimination rather than estimation error.

$$18$$

One way to calculate officer heterogeneity’s accounting for noise is to do a Bayes shrinkage procedure. When we replicate the approach of Aaronson, Barrow, and Sander (2007), our distribution of discrimination looks nearly identical to the unshrunk version.

---

Figure 6. Difference-in-Differences Results

Notes: Figure plots the difference-in-difference regression results for each speed from equation (3). The y-axis plots the interaction between driver being white and the officer being lenient. Standard errors are at the 5 percent level. The legend indicates the controls in each set of regressions.
Another approach to understanding the variance in discrimination across officers is to estimate what share of officers are discriminatory. We know that each officer’s discrimination measure is an additive function of his true discrimination plus estimation error, $\theta_j^\hat{} = \theta_j + \epsilon_j$. We can assume an officer’s discrimination can take on a finite set of values on a fine grid, $\theta_j \in \{\theta^k\}$, and calculate the likelihood of observing each officer’s discrimination measure $\theta_j^\hat{}$, given the noise in the measure and the true distribution $f(\theta^k)$:

$$
Pr(\theta_j^\hat{}) = \sum_{\theta^k} f(\theta^k) \cdot Pr(\epsilon_j = \hat{\theta}_j - \theta^k).
$$

We then estimate $\{f(\theta^k)\}$ by maximum likelihood, where we impose that $\epsilon_j$ is normally distributed with $\sigma_j^2$ taken from the officer-level regression. Using this approach, and calculating $1 - \sum_{\theta^k<0} f(\theta^k)$ as the share, we find that 42.0 percent (standard error 1.2) of officers are discriminatory. In contrast, we find that only 8.4 percent (0.8) of officers have $\theta_j < 0$, i.e., practice reverse discrimination. When removing officers who are non-lenient and thus cannot practice discrimination, our estimate of share discriminatory increases to 60.3 percent (1.7). We can also apply our correction from Section IIIA here and subtract from each officer’s discrimination coefficient our

![Figure 7. Difference-in-Differences Officer-Level Results](image)

Notes: Figure plots each officer’s $\beta_j^3$ from equation (3). Officers who are non-lenient are assigned $\beta_j^3 = 0$. SD reports the standard deviation across $\beta_j$, and average SE reports the average standard error for each individual $\beta_j$.

---

19 This approach is a discretized version of a deconvolution procedure (Delaigle, Hall, and Meister 2008). Doing the continuous deconvolution leads to an identical estimate for the share of officers who are discriminatory. The grid is 99 points spanning the first to ninety-ninth percentiles of the empirical distribution of $\theta_j$. Confidence intervals are calculated through bootstrapping by performing 100 draws of the set $\{\hat{\theta}_j\}$ and performing MLE on each draw.

20 Table A.6 presents the share discriminatory and reverse discriminatory separately by officer race.
estimate of statistical bias induced by racial differences in speeding. Doing so slightly decreases our baseline estimate from 42.0 percent to 38.5 percent (1.1).

The left panel of Figure 8 shows how our measure of discrimination varies by officer race. Perhaps consistent with intuition, White officers are much more likely to be discriminatory against minority drivers, with a greater rightward skewness in their distribution. However, the average Hispanic officer is also discriminatory against minority drivers. On the other hand, Black officers exhibit no discrimination on average, with a roughly even share of officers with positive and negative discrimination coefficients. Some of the disparity in discrimination across officer race is driven by minority officers being less likely to be lenient overall, in part due to minority officers working in troops in which all officers are less lenient. In the right panel of Figure 8, we show the distribution of discrimination only for lenient officers, and the White officers’ distribution continues to be shifted farther to the right.

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

Note that the “hit-rate” approach to testing for discrimination, most recently used in Marx (2018) and Arnold, Dobbie, and Yang (2018), also provides an “absolute” measure of discrimination that does not use a comparison group and could in principle be used to estimate discrimination separately by race of the criminal justice agent.
In online Appendix Section D.1, we conduct a further analysis of how our full set of officer demographics predict discrimination and lenience. While we find that officer experience and education are unrelated to discrimination, female officers are on average less discriminatory. Online Appendix Section D.2 also probes the stability of our officer-level estimates across an officer’s career and documents a significant positive relationship between estimates drawn from an officer’s early and late samples.

While our main analysis focuses on the degree of discrimination against Black and Hispanic drivers as a composite, an important follow-up question is how our measures of discrimination change when estimated separately for Black and Hispanic drivers. We conduct these analyses in online Appendix Section E. We estimate a significant degree of discrimination against both minority groups, though our estimates are somewhat larger for Hispanic drivers.

V. Robustness Checks and Alternative Explanations

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

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

One concern with our baseline estimates is the sometimes-significant coefficient on White driver. In principle, this coefficient characterizes speed differences between White and non-White drivers among non-lenient officers and should be small based on our analysis in Figure 5. When we re-weight the data in column 4 so that the share minority is the same in each county, the coefficient on White driver diminishes substantially and is no longer significant, suggesting that model misspecification could play a role in the occasional significance of the White driver coefficient. An important candidate type of misspecification is the attribution of heterogeneity in the lenience coefficients to the location-time fixed effects. We correct for this type of specification error in column 5 by overweighting the non-lenient officers and find a greatly reduced and insignificant coefficient on White driver relative to our baseline results.22

22 This misattribution issue arises in difference-in-differences models where parts of the treatment group (here, the lenient officers) operate as a control group for other treated observations by helping to estimate the fixed effects (Borusyak and Jaravel 2017, Goodman-Bacon 2018). In column 5, we re-run the baseline regression where the non-lenient officer observations are given weight of 1,000 and the lenient officer observations are given weight of 1. By doing so, only the non-lenient officers are used to identify the fixed effects and only the lenient officers are used to identify the lenient officer variable coefficients.
One feature of the data discussed earlier is that the histogram of ticketed speeds exhibits jumps at multiples of five, and we find that this heaping only occurs among “visual” stops. In column 6 of Table 4, we find that our baseline regression is essentially unchanged when restricting to the sample of tickets from a radar or laser gun. In all these specifications, the interaction coefficient between officer lenient and driver race is significant and quantitatively similar to the baseline specification. The largest disparity is evident in the re-weighted specification, where the coefficient reduces from 5.8 to 3.8 percentage points. This difference suggests that some of the gap in treatment between Whites and minorities is due to minorities disproportionately driving in counties where officers are less lenient overall. These differences across counties could be due to differences in how much drivers exceed the speed limit. In our model in Section VI, we explicitly account for the possibility that counties and races differ in speeds and continue to find a disparity in discounting between races.

23 In online Appendix Table A.4, we show that our discrimination estimates are also unchanged when restricting to subsamples where speeding is either the only offense or one of at least two offenses, alleviating concerns that minorities are treated more harshly because they are also facing charges for other offenses.
A. Selection into the Data

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

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

Furthermore, any discrimination on the stopping margin would likely bias our results toward finding less discrimination in discounting. To see this argument, imagine a minority driver who is on the margin of being ticketed, such that if he were White he would have been let off with a warning. This driver appears in our data only because he is a minority. Because he is at this margin, it is very likely the officer will give him a discount. Therefore, discrimination on the ticketing margin places too many minority drivers in our sample who are disproportionately at the discount point. Thus, the disparity in discounting would be even greater without a hypothetical disparity in ticketing.24

In online Appendix Section F, we formalize this logic with a simple selection model that allows for officer differences in propensity to let drivers off with a warning. Using this model, we implement a sample selection correction, as in Heckman (1979), that accounts for officer-by-race differences in propensity to appear in the data. We report the results of this regression in column 7 of Table 4, and all of our primary coefficients look identical to our baseline specification.25

B. Racial Difference in Requesting a Break

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

---

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

25 In online Appendix Table A.11, we document how our results vary by daytime versus nighttime, when officers are less able to discern race (Grogger and Ridgeway 2006). In the potentially less selected sample of nighttime stops, we find a slightly larger coefficient, consistent with the argument that discrimination in the ticketing margin biases our estimates downwards.
We argue that racial differences in propensity to request a break cannot explain our findings. Relative to existing studies in the discrimination literature, one strength of our data is that individuals can be linked across tickets, allowing us to evaluate whether there are individual-level differences in propensity to receive a discount. To probe this question, we restrict attention to individuals with at least two tickets, comprising 172,810 observations. Running a regression of discounting on individual characteristics has an $R^2$ of 0.31, and the addition of officer fixed effects leads to an increase to 0.54. In contrast, the further addition of individual-fixed effects only increases the $R^2$ to 0.55. This small increase shows that, beyond individual covariates and the stopping officer, the specific individual has little explanatory power for whether a discount is given, indicating that individual differences in propensity to request a break is likely not a substantial factor in the disparity in discounting.

C. Statistical Discrimination versus Taste-Based Discrimination

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

In online Appendix Section G, we present a simple test for whether officers are attempting to minimize court contesting or maximize deterrence, which we report in Table 5.26 To evaluate the impact of a discounted ticket, we instrument for receiving a discount using the stopping officer’s persistent (leave-out) level of leniency. Our test then follows the logic of Heckman, Schmierer, and Urzua (2010) and claims that nonlinearities in the relationship between the outcome and the propensity score reflect sorting of individuals on the basis of their responsiveness.

We find no evidence that officers choose who to discount on the basis of deterrability: columns 1 and 2 show that the impact of a discount on future speeding is positive and that the effect is largest for low values of the propensity score. This nonlinearity indicates that drivers who are at the margin of being discounted by officers who are not very lenient (i.e., the “first” drivers to be discounted) have the largest increase in recidivism from being discounted, suggesting that officers

---

26 See online Appendix Section A.4 for more information on our measure of court contesting, also acquired from the Florida Clerk of Courts (FCCC 2016a) and online Appendix Section A.5 for additional background on the institutional details regarding fine repayment.

27 This procedure is very commonly used in the criminal justice literature when judges differ in their punitiveness (Kling 2006, Dobbie and Song 2015) We use this approach to evaluate how individuals respond to their ticket in a follow-up paper, Goncalves and Mello (2017).
are actually sorting individuals in the opposite manner from what would maximize deterrence.

However, we do find that officers choose who to discount based on whether they will contest their ticket in court: among officers who are not very lenient, the marginal impact of giving a driver a discount is a large reduction in likelihood of contesting the ticket. In contrast, more lenient officers have a marginal impact of a discount on court contestation that is significantly smaller, suggesting that more responsive drivers are discounted first. We then perform in column 5 a hit-rate test similar to Arnold, Dobbie, and Yang (2018) and find that officers’ statistical discrimination on court contestation cannot explain the racial disparity in discounting.

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

VI. Model and Counterfactuals

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

Individual $i$ drives at a speed $s^*$ that is drawn from a Poisson distribution $P_{\lambda_i}(s^*)$, where

$$\lambda_i = \lambda_r + \gamma Z^{(1)}$$

is a function of the county-by-race of the driver and other demographics $Z^{(1)}$. We include in $Z^{(1)}$ the driver’s gender, age, and number of tickets in the previous three years.

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

$$t_{ij} + \alpha Z^{(2)}_i + \epsilon_{ij} > a + b \cdot s^*_i.$$

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

Two simplifications of the model should be discussed here. First, we do not allow the driver’s distribution of speeds to respond to the lenience of the officers in their county. We are comfortable in making this restriction because we find that there is no cross-sectional relationship between the county lenience rate and the speeds charged.28

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

### A. Identification

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

28 In Goncalves and Mello (2017), we find that drivers do respond ex post to receiving a harsh ticket by speeding less. This should lead to a steady-state relationship between lenience and the frequency of traffic tickets. However, the magnitude of the deterrence effect is small enough that the racial gaps in the counterfactuals would not be meaningfully impacted. For example, in the 11 years of our sample, if all minority drivers were treated as White drivers, there would only be about 70 more car accidents and less than one more death in expectation.
In practice, our estimation is similar to our difference-in-differences regressions, in that it relies on the heterogeneity across officers in discount lenience. While all officers’ data enter the maximum likelihood equations, the speed parameters are primarily identified using officers who exhibit no lenience, from which we get an estimate of the true distribution of speeds. We do so based on the assumption that officers and drivers are randomly sorting within a county, allowing us to suppose that the underlying distribution of speeds are the same for non-lenient and lenient officers.

Our estimation also depends 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 $67 \times 2$ county-race speeds $\lambda_{rc}$; 3 demographic speed parameters $\gamma$; $1,851 \times 2$ officer average racial preferences, $t_{rj}$; 4 demographic preference parameters $\alpha$; and the slope of the cost function $b$, totaling 3,843 parameters. Details of how the estimation is carried out in practice are provided in online Appendix Section H.

### B. Model Estimates

The results of the model estimation are reported in online Appendix Table A.14. Because the estimates are closely aligned to the findings from our difference-in-differences approach, we leave our full discussion of these estimates to online Appendix Section H.1. In short, we find that the average officer practices substantial lenience, with a significant variance across officers. Off a baseline of 34 percent likelihood of discounting a driver from 15 mph to 9 mph, the average officer is 2 percentage points less likely to discount minority drivers. We find that minorities drive significantly faster than White drivers, as do males, younger drivers, and drivers with previous tickets. The average officer is also more generous to female drivers, older drivers, and drivers with fewer previous tickets. They are also less lenient to all drivers when ticketing in a county with more minorities.

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

The “Baseline” row of Table 6 shows how the charged speeds of drivers appear in a simulation of the model that does not change any of the parameters of the model. All of the decompositions are benchmarked to this baseline. In the “No discrimination” row, we remove discrimination by making each officer treat minority
drivers like they treat White drivers. This restriction reduces the gap in discounting by 24 percent. 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 percent of the gap in discounting is removed, consistent with the earlier finding that officers tend to be more lenient overall in neighborhoods with fewer minorities. Removing both sorting and discrimination, the gap in speeding is reduced by 54 percent. The remaining gap is due exclusively to the fact that minorities are driving faster speeds. In the second panel of Table 6 we report the same decompositions, where the gap is conditional on the county of the stop. Removing the sorting of officers no longer has any effect, since that only leads to differences across counties. Further, notice that over 80 percent of the within-county disparity can be explained by discrimination, leaving only 19 percent of the disparity to be explained by differences in speeding across races. In online Appendix Table A.15, we perform these same calculations where the outcome of interest is the average speed rather than share discounted and find similar results.

C. Policy Counterfactuals

We now use the model estimates to conduct a series of policy counterfactuals to explore how best to curb discrimination in speeding tickets, which we report in Table 7. As a benchmark, the first two rows are reproduced from the top panel of
Table 6, showing a baseline simulation that keeps the empirical pool of officers and their locations and a simulation where all officers treat minority drivers like White drivers and are randomly assigned to drivers throughout the state. As with Table 6, the calculation of standard errors is discussed in online Appendix Section H.

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 ninety-fifth percentile and above of discrimination and remove them from the pool of officers. This cutoff removes officers with a difference in discounting at 10 mph over of 14.5 percentage points or greater between Whites and minorities. For symmetry, we also remove officers who reverse discriminate by that amount (comprising only 1.4 percent of officers).

The statewide disparity in treatment barely changes in response to removing these officers, falling by less than 3 percent. The lack of effectiveness from this policy partly stems from the fact that discriminatory officers are on average very lenient. When they are removed, drivers are left to be stopped by officers who, while less discriminatory, are also less lenient overall. This fact can be seen by noting that the average discount rate goes down for White drivers.

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

Table 7—Model Counterfactuals

<table>
<thead>
<tr>
<th>Hiring and firing</th>
<th>White mean</th>
<th>Minority mean</th>
<th>Difference</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.355</td>
<td>0.274</td>
<td>−0.081</td>
<td>100</td>
</tr>
<tr>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No discrimination or sorting</td>
<td>0.335</td>
<td>0.299</td>
<td>−0.037</td>
<td>45.5</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Fire 5 percent most biased officers</td>
<td>0.353</td>
<td>0.274</td>
<td>−0.079</td>
<td>97.8</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>Increase female share 10 pp (base of 8 percent)</td>
<td>0.352</td>
<td>0.274</td>
<td>−0.078</td>
<td>95.9</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>Increase minority share 10 pp (base of 35 percent)</td>
<td>0.305</td>
<td>0.242</td>
<td>−0.063</td>
<td>77.8</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.026)</td>
<td></td>
</tr>
</tbody>
</table>

Resorting officers

| Exposing minorities to least biased | 0.339 | 0.267 | −0.072 | 89.2 |
| (0.002) | (0.002) | (0.002) | (0.020) | |
| Exposing minorities to most lenient | 0.307 | 0.300 | −0.007 | 8.6 |
| (0.001) | (0.001) | (0.001) | (0.010) | |

Notes: Results are reporting the probability of being ticketed 9 mph over, where the averages are statewide. In the bottom panel of counterfactuals, officers are resorted within troops.
Increasing the share of female officers from 8 percent to 18 percent of the force leads to a 4.1 percent reduction in the discount gap. An increase in minority officers from the empirical share of 35 percent to 45 percent reduces the gap by 22.2 percent.

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

**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 on average consist of seven counties. Within the troops, officers regularly vary in which locations they patrol. It may be potentially feasible for a senior officer to, for example, change the assignment of officers, such that minorities face less biased officers. The bottom two rows of Table 7 present the results of such a policy. Column 1 sorts officers within a troop, such that the least biased officers are in counties with the most minorities. Column 2 sorts officers within a troop, such that the most lenient officers are in counties with the most minorities.

Surprisingly, sorting officers to expose minorities to the least discriminatory has a very small effect on the treatment gap. The least biased officers are also not very lenient on average, dampening the impact of their equal discounting across races and reducing the gap in discounting by only 11 percent. Much more effective in reducing the gap in treatment is assigning the most lenient officers to minority counties, regardless of their discrimination in discounting. This policy reduces the treatment gap by 91 percent.

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.

**D. Caveats**

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

For example, increasing lenience uniformly may lead to increased speeding, which we show to be the case in a separate study (Goncalves and Mello 2017). Changing leniency standards may also lead officers to give drivers verbal warnings rather than a reduced charge. A full consideration of the welfare impact of the ensuing policies would likely consider additional outcomes, such as the speeding response to changes in enforcement (Gehrsitz 2017, Goncalves and Mello 2017, Chalfin and McCrary 2017) and the tradeoff between the level and inequality in lenience.
One additional concern is that officers will change their lenience behavior in response to being reassigned counties. We address this concern in part by allowing officer behavior to vary by the share of drivers who are minorities, though it is important to note that officers may respond in other ways.

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

VII. Conclusion

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

We study speeding tickets and the choice of officers to discount drivers to a speed just below an onerous punishment. Using a bunching estimator in a difference-in-differences framework, we document that minority drivers are significantly less likely to be given a discounted speed on their ticket. A key advantage of our approach is the ability to explore the entire distribution of both lenience and discrimination on the part of officers. Our estimates reveal significant heterogeneity in behavior across officers, with 42 percent of the force explaining the entire aggregate disparity. Estimates from our parametric model of driver speeding and officer decision-making confirm that, while minorities drive slightly faster on average, our officer-level estimates of discrimination, which leverage non-lenient officers as a control group, are not confounded by differences in speeding across racial groups. On net, we attribute 82 percent of the racial gap in discounting to discrimination by patrol officers.

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

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


This article has been cited by: