NBER WORKING PAPER SERIES

# POLICE DISCRETION AND PUBLIC SAFETY

Felipe M. Gonçalves
Steven Mello

Working Paper 31678
http://www.nber.org/papers/w31678

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2023

We are grateful to Leila Agha, David Arnold, Eric Chyn, Will Dobbie, Ben Hansen, Peter Hull, Elisa Jacome, Louis Kaplow, Ilyana Kuziemko, Emily Leslie, Adriana Lleras-Muney, Erzo Luttmer, Alex Mas, Michael Mueller-Smith, Emily Nix, Sam Norris, Rodrigo Pinto, Bruce Sacerdote, Yotam Shem-Tov, Chris Snyder, Doug Staiger, Megan Stevenson, and Emily Weisburst, as well as seminar participants at Princeton, UCLA, Dartmouth, Case Western, UBC, Claremont McKenna, SDSU, Duke, Brown, Michigan, FRBNY, the SEA Annual Meeting, and NBER Summer Institute for helpful comments. We thank Jeffrey Bissainthe, Kiara Guzzo, Wilton Johnson, Timothy Kutta, Stacy Lehmann, Brenda Paige, and especially Beth Allman for assistance with acquiring the data from various Florida agencies. Jacklyn Pi provided excellent research assistance. The Princeton Industrial Relations Section provided generous financial support. Any errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Police Discretion and Public Safety
Felipe M. Gonçalves and Steven Mello
NBER Working Paper No. 31678
September 2023
JEL No. D73,J45,K42

**ABSTRACT**

We study the implications of police discretion for public safety. Highway patrol officers exercise discretion over fines by deviating from statutory fine rules. Relying on variation across officers in this discretionary behavior, we find that harsher sanctions reduce future traffic offending and crash involvement. We then show that officer discretion over sanctions decreases public safety by comparing observed reoffending rates with those in a counterfactual without discretion, estimated using an identification at infinity approach. About half the safety cost of discretion is due to officer decisions which result in harsh sanctions for motorists who are least deterred by them. We provide evidence that this officer behavior is attributable to a preference for allocating harsh fines to motorists with higher recidivism risk, who are also the least responsive to harsher sanctions.

Felipe M. Gonçalves
Department of Economics
University of California, Los Angeles
8283 Bunche Hall
Los Angeles, CA 90095
and NBER
fgoncalves@econ.ucla.edu

Steven Mello
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
steve.mello@dartmouth.edu

# 1 Introduction

Ensuring public safety is a central function of the state. To that end, policy regimes specify sanctions for socially undesirable behaviors (Becker, 1968). However, beyond the set of official policies, the state also relies on bureaucrats tasked with allocating sanctions in practice. As in much of the public sector, these actors typically wield significant discretion (Lipsky 1980; Wilson 1989). For example, patrolling police officers can overlook minor crimes, prosecutors can reduce criminal charges, and judges can issue a wide range of sentences. Permitting this type of discretion can have important benefits if agents internalize the goals of the state and can effectively discern the allocation of sanctions that maximizes public safety (e.g., Banfield 1975; Kang & Silveira 2021). On the other hand, these decision-makers may have poor information, attend to alternative objectives such as personal gain, or hold competing notions of the optimal allocation when making decisions (e.g., Prendergast 2007).

A critical question raised by the pervasive discretion practiced by criminal justice agents, then, is the extent to which that discretion improves public safety. We study this question in the context of traffic enforcement, focusing on highway patrol officers who exercise discretion over speeding sanctions by deviating from statutory fine rules. To start, we document that these discretionary choices have important implications for public safety: harsher fines reduce a motorist's future traffic offending and likelihood of future crash involvement. Next, we show that officer discretion reduces public safety by estimating reoffending rates in a counterfactual scenario without discretion. We find that officer choices of *whom* to sanction harshly play an important role in explaining the net safety costs of discretion. In particular, eliminating discretion results in harsh fines for motorists who typically see lenience when officers use their discretion, and deterrence effects are particularly large in the subsample. We conclude by exploring potential explanations for this officer behavior, including officer preferences and imperfect information.

Speeding enforcement is a high stakes setting in terms of public safety. In 2020, there were nearly twice as many traffic fatalities ($\sim$ 39,000) as homicides ($\sim$ 22,000) in the United States. Economic costs associated with motor vehicle accidents have been estimated at nearly $250 billion per year, higher than annual costs of crime victimization (Blincoe et al. 2015; Chalfin 2016). Standard estimates suggest that at least one third of fatal crashes are caused by speeding and existing studies have found strong associations between average driving speeds and traffic fatalities (NHTSA 2014; Ashenfelter & Greenstone 2004).

While speeding sanctions are statutorily based only on a driver's speed relative to the posted limit, officers can manipulate fines by writing down a slower speed than was observed on the actual citation, resulting in a discounted fine (Anbarci & Lee 2014; Goncalves & Mello 2021). In Florida, the setting of our study, nearly one third of all speeding citations are issued for exactly nine miles per hour (MPH) over the limit, just below a $75 increase in the statutory fine amount. Less than one percent are issued for either eight or ten MPH. Officers patrolling the same beat-shifts vary considerably in the degree of bunching in their charged speed distributions, highlighting that officer discretion, rather than driver behavior, explains the bunching in cited speeds (Goncalves & Mello, 2021).

Our empirical approach leverages this variation across officers in the propensity to bunch motorists below fine increases. We use an instrumental variables (IV) framework where the treatment and outcome of interest are whether a driver receives a harsh (versus discounted) fine and whether a cited driver commits a new traffic offense in the following year. Our instrument is the citing officer's propensity *not to bunch* other drivers, which we call stringency. Our design mirrors a growing literature leveraging randomly assigned judges for identification (e.g., Kling 2006; Maestas et al. 2013; Dahl et al. 2014; Dobbie & Song 2015, Mueller-Smith 2015), with the caveat that, in our setting, citing officers are not randomly assigned to drivers. An important concern for our approach, then, is whether an officer's bunching propensity is correlated with the characteristics of her sample of cited drivers. We show that, conditional on beat-shift fixed effects, stringency is uncorrelated with an officer's ticketing frequency, driver characteristics that predict reoffending, and past traffic offending.

First, we use this IV framework to study the deterrence effects of sanctions. We estimate the local average treatment effect (LATE) of harsher fines on the future driving behavior of cited drivers, instrumenting harsh fines with officer stringency. We find that a 125 dollar increase in fine amounts reduces the likelihood of any new traffic offense in the following year by 1.6 percentage points ($\epsilon = -0.07$). We document a stronger effect on speeding offenses ($\epsilon = -0.13$) and a statistically significant, but less precise, effect on the likelihood that a driver is involved in a traffic accident in the next year ($\epsilon = -0.04$).

In this setting, motorists receiving harsh and lenient fines face identical sanctions for future offenses. Hence, the ex-post response we document represents a specific deterrence effect, or the impact of the experience of punishment on offending (Nagin, 2013). Our fine elasticity estimates advance a literature focused on isolating specific deterrence effects (Gehrsitz 2017; Hansen 2015, Dusek & Traxler 2021, Finlay et al. 2022) and the broader literature on the deterrence effects of traffic enforcement efforts (e.g., Makowsky & Stratmann 2011; DeAngelo & Hansen 2014; Luca 2014; Traxler et al. 2018). Our ability to measure individual-level impacts on crash involvement, an outcome clearly associated with social costs and not subject to enforcement-based measurement, also adds to the literature on specific deterrence.

Next, we quantify the overall contribution of officer discretion to public safety by estimating the reoffending rate in a counterfactual scenario without officer discretion. Because the specific type of discretion we study takes the form of officers departing from fine rules by reducing charges, our aim is to estimate the reoffending rate if all cited drivers were instead issued harsh fines. The empirical challenge we face is that this counterfactual is not identified by the LATE when treatment effects are heterogeneous, especially if officers strategically use their discretion. For example, suppose that officers allocate harsh fines to maximize public safety. In this case, the motorists currently facing harsh fines should be the most responsive to those harsh fines in terms of driving behavior, and the safety gains associated with issuing harsh fines to those currently not receiving them could be much smaller than suggested by the LATE. Note that this is a pervasive issue in the program evaluation literature, where extrapolating beyond the LATE is an important step in assessing the effects of counterfactual policy changes (e.g., Heckman & Vytlacil 2007, Cornelissen et al. 2016, Brinch et al. 2017).

Hence, our desired counterfactual requires a dedicated estimate of the average treatment effect on the untreated (ATU). While this parameter can be recovered from estimates of marginal treatment effects (e.g., Bjorklund & Moffitt 1987; Heckman & Vytlacil 2007; Mogstad & Torgovitsky 2018), identification of the MTE requires a strict monotonicity assumption, which implies that all officers share a common ranking of motorists when allocating harsh sanctions. Instead, we estimate the ATU using a novel two-step procedure that does not require any monotonicity assumptions. We first extend our IV framework using an identification at infinity approach (e.g., Heckman 1990; Hull 2020; Arnold et al. 2022). Specifically, examining the reoffending rates of motorists cited by officers who always and never issue harsh fines identifies the average treated and untreated potential outcomes. We use parametric extrapolation techniques as in Arnold et al. (2022), but our setting is unique in that we actually observe officers in the tails of the stringency distribution. As such, our extrapolation-based estimates of the average treated and untreated potential outcomes are insensitive to functional form assumptions. We then combine these estimates of average potential outcomes with observed reoffending rates in the data in the data to estimate the average treatment effect (ATE), average treatment effect on the treated (ATT), and average treatment effect on the untreated (ATU).

Our approach indeed reveals important differences between the LATE and the ATU. We find that motorists who currently face lenient fines exhibit treatment effects which are 40 percent larger (more negative) than the estimated LATE. Our estimated ATU implies that the overall reoffending rate would decline by 2.2 percent in a counterfactual scenario without officer discretion. To put this number in perspective, we estimate that switching *all drivers* from lenient to harsh fines would reduce the reoffending rate by 3.4 percent. Eliminating discretion thus achieves over sixty percent of the feasible deterrence on this margin.

The safety gains from removing officer discretion can be decomposed into two channels. The first, which we call the average component, is the safety gains associated with increasing the likelihood of harsh fines for the *average motorist*. The second, which we call the sorting component, is the relative safety gain from issuing harsh fines to the specific *subgroup* of motorists who typically face lenience when officers exercise discretion. The second component captures differences in treatment effects for the average motorist (ATE) and the subset of motorists whose treatment status is changed without discretion (ATU). We find that each channel explains about half of the overall safety gains from eliminating officer discretion.

In other words, increasing the harshness rate for the average driver improves safety, and allocating harsh fines to motorists who typically receive lenience when officers use discretion further reduces the reoffending rate, because these motorists are especially responsive to harsh fines. Note that this is exactly the opposite of what would be expected under safety maximization by officers, which would predict the smallest treatment effects for the motorists typically issued lenient fines.

We find similar patterns when considering alternative measures of safety. Examining a motorist's number of new traffic offenses in the following two years, those issued lenient fines represent over 80 percent of the feasible deterrence. In terms of future crash involvement, untreated motorists comprise nearly all the feasible deterrence. For this group of drivers, we estimate that each additional dollar in fines issued from counterfactually imposing harsh sanctions yields a public safety gain of nearly two dollars in the form of lower social costs attributable to traffic accidents.

Finally, we consider potential explanations for the officer behavior we document. Along with the *reverse selection on gains* discussed above, we find striking *positive selection on levels*: if issued a lenient fine, the motorists who face harsh sanctions would reoffend about 15 percent more often than motorists who receive lenience. We find that both reverse selection on gains and selection on levels persist when examining patterns within motorist covariates or focusing only on first-time offenders and cannot be explained by officer decisions based on the underlying stopped speed. We also find similar selection patterns when examining experienced officers. Hence, the available evidence is consistent with the view that the current allocation of fines reflects the goals of officers. In particular, these patterns align well with an intuitive notion of fairness, with officers allocating harsh fines to the "worst" drivers in terms of recidivism risk. These motorists are the least deterred by harsh fines, and hence eliminating these "sorting" decisions by officers has important public safety benefits.

Our central contribution is to a large literature on the implications of bureaucratic discretion for state effectiveness (e.g., Prendergast 2007; Ash & MacLeod 2015; Best et al. 2017; Bandiera et al. 2021) and to a rapidly growing literature on the role of discretion in the criminal justice system (e.g., Weisburst 2017; Ba et al. 2021; Chalfin & Goncalves 2021; Goncalves & Mello 2021; Abrams et al. 2021; Feigenberg & Miller 2022; Norris 2022; Angelova et al. 2023). We quantify the net effect of officer discretion over sanctions on public safety and show that a misalignment between the deterrence objectives of the state and the discretionary behavior of police officers has important implications for road safety.

Our findings on this dimension also speak to a largely theoretical literature on fairness-efficiency tradeoffs in the design of legal institutions (e.g, Polinsky & Shavell 2000; Kaplow & Shavell 2006; O'Flaherty & Sethi 2019; Moore 2019) by documenting the empirical relevance of a tradeoff between efficiency and other potential law enforcement objectives, such as the targeting of sanctions to the "worst" offenders. Importantly, we do not take a firm stand on the overall welfare implications of this officer behavior, as this behavior could reflect a legitimate law enforcement goal or the desired allocation of the state.

We also contribute to a broad literature on the allocation choices of economic agents. A common practice in this literature is to examine selection patterns in settings where efficient allocations should exhibit selection on gains (e.g., Carneiro et al. 2011; Abaluck et al. 2016; Van Dijk 2019; Chandra & Staiger 2020). Several such studies have nonetheless found sorting based on levels, such as parents choosing school districts for their children (Abdulkadiroglu et al., 2020) and hospitals opting into a Medicare reform (Einav et al., 2022). We document similar behavior in a new setting, law enforcement, and we do so using a novel method that can characterize selection patterns under minimal assumptions.

Our paper proceeds as follows. Section 2 describes our data and setting. We lay out our empirical framework in section 3 and estimate the causal effect of sanctions in section 4. Section 5 quantifies and decomposes the net contribution of officer discretion to public safety and section 6 considers potential explanations for the officer behavior we find. Section 7 concludes.

# 2 Data and setting

## 2.1 Data sources

The Florida Clerks and Comptrollers provided administrative records of the universe of traffic citations issued in Florida for the years 2005–2018 from Florida's Uniform Traffic Citation (UTC) database. These records include the date and county of the citation as well as information on the cited violation. When the violation is speeding, this information includes the charged speed and posted speed limit (e.g., 74 MPH in a 65 MPH zone). The UTC data also include all information provided on a stopped motorist's driver license (DL): name, date of birth, address, race, gender, as well as DL state and number. Using the driver license number, we are able to link drivers across citations and construct our primary measures of past and future traffic offending.

We augment the motorist information in the UTC data with four auxiliary data sources. First, we match drivers on zip code of residence to estimated per-capita income at the zip code level from the IRS Statistics of Income files. Second, we construct estimated vehicle values based on a database of online vehicle resale values. Third, we recode a motorist's race as Hispanic if, based on census records, their surname is associated with Hispanic status for more than 80 percent of individuals.[1] Finally, we link drivers on full name and date of birth to prison spell records from the Florida Department of Corrections to construct a measure of prior incarceration.

In the citations data, the ticketing officer is identified by name. We construct a consistent officer identifier by linking the officer name with data on Florida Highway Patrol (FHP) employment spells provided by the Florida Department of Law Enforcement. We focus on tickets issued by the FHP both because we can more consistently the citing officer and because speeding enforcement is a central duty of the FHP. However, we measure past and future offending using all citations, not just the FHP-issued citations in our focal sample.

We also obtained administrative crash reports covering the universe of automobile accidents known to police over the period 2006–2018 from the Florida Department of Transportation (FDOT). These data are collected during a police response or investigation and include the date and county of the incident and include the DL numbers of involved drivers, which we use to link drivers with the citations data.

The Florida Clerks and Comptrollers provided records from the Traffic Citation Accounting Transition System (TCATS) database, which includes information on the traffic court disposition associated with about 80 percent of the citations in our sample. We use these records to construct a measure of whether a citation was contested in traffic court and, based on the disposition, to construct measures of accrued, rather than statutory, sanctions. For additional details, see appendix D.

## 2.2 Sample construction

To construct our sample of focal citations, we first restrict attention to tickets written by the Florida Highway Patrol over 2007–2016 where the citing officer is identified.[2] We further restrict the sample to include tickets where speeding is the only violation, no crash is indicated, and the charged speed is between nine and twenty-nine miles per hour over the posted speed limit. We choose twenty-nine as our upper limit because (i) the available evidence suggests that motorists are still bunched with positive probability when their true speed is as high

[1]As discussed in Goncalves & Mello (2021), there are clear inconsistencies in the recording of Hispanic status in the UTC data. Officers frequently write down race = $H$ (for Hispanic). But in Miami-Dade county, where the population is over 60 percent Hispanic, less than one percent of citations are coded as being issued to a Hispanic motorist.

[2]We focus on 2007–2016 so that we can measure other offending (including crash activity) for least one full year prior and one full year after the focal citation. Over this period, the ticketing officer is identifiable for 85 percent of FHP-issued speeding tickets.

as twenty-nine MPH over the limit (see figure 1) and (ii) thirty MPH over the limit is the threshold for criminal speeding.

We also restrict to drivers with a valid Florida driver license number, so that we can reliably measure past and future offending, and require that officers have at least fifty citations meeting the above criteria to compute our instrument. Ultimately, our focal sample is comprised of 1,693,457 speeding citations issued by 1,960 FHP officers. There are 1.4M unique drivers in the sample. Table 1 presents summary statistics.

Again, reoffending and past offending are measured using *all* citations issued in the state rather than just the citations that comprise our sample of focal FHP tickets. Worth noting here is the fact that our main outcome measure will capture whether a motorist is caught and ticketed for a new traffic offense, which itself could be subject to officer discretion. If anything, we expect that officer discretion at the recidivism stage will bias our specific deterrence effect estimates towards zero. We compare reoffending rates for individuals (randomly) receiving harsh and lenient fines and find that those receiving harsh fines differentially reduce their offending rates. If officers are more likely to let drivers with less severe offending histories off with formal or informal warnings, that would bias our estimates towards zero by deflating the reoffending rates of those who are issued lenient sanctions.

## 2.3 Florida highway patrol

State-level patrols are the primary enforcers of traffic laws on interstates and many highways, especially those in unincorporated areas. On patrol, officers are given an assigned zone over which they can combine roving patrol and parked observation patrol. Florida Highway Patrol (FHP) officers are divided into one of nine assigned troops, almost all of which patrol six to eight counties each. Officer assignments operate on eight-hour shifts and cover an assignment region that roughly corresponds to a county, though the size of a "beat" can vary based on an area's population density. In practice, we use counties to proxy for assignment regions.

The FHP is comprised of approximately 1,500 full-time officers. Speeding enforcement is a primary duty of FHP officers and the FHP collectively issues between 150,000 and 200,000 speeding citations each year. Other responsibilities include enforcing a wide array of other traffic laws, investigating crashes, and responding to and assisting with highway emergencies. The FHP officer handbook reads *"Members should take the enforcement action they deem necessary to ensure the safety of the motoring public, reduce the number and severity of traffic crashes, and reduce the number of criminal acts committed on highways of this state,"* highlighting that officers are explicitly given discretion over enforcement decisions.

In Florida, speeding sanctions are based on an offender's speed relative to the posted speed limit. Speeding 1-5 MPH over the limit carries a statutory warning but no sanctions, while speeding 30 or more MPH over the limit is a misdemeanor offense requiring the offender to appear in court. Between 6 and 29 MPH over the limit, the statutory fine is a step function, plotted as a red dotted line in figure 1.

Speeding offenses are also associated with "points" on an offender's driver license (DL). Point assessments are also based on speed; speeding 6-15 MPH over the limit is associated with 3 points while speeding 16+ MPH over the limit is associated with 4 points. Points are used by car insurers to adjust premiums and offenders that collect a sufficient number of points (12 points in 12 months; 18 points in 18 months; 24 points in 36 months) have their license suspended for 30 days (6 months; 1 year).

After a citation has been issued, a driver can either submit payment to the county clerk or request a court date to contest the ticket. If the ticket goes to court, a judge or hearing officer typically decides either to uphold the original charge, reduce the charge, or dismiss the citation. At the time of payment, a subset of drivers can elect to attend an optional traffic school, completion of which combined with on-time payment will remove the citation from a driver's record and prevent the accrual of the associated DL points.

### 2.4 Discretion over sanctions

Panel (a) of figure 1 shows the speeding fine schedule in Florida and a histogram of charged speeds on FHP-issued speeding citations. Over one third of all citations are issued for exactly 9 MPH over the posted limit, just below a $75 increase in the associated fine. Less than one percent of all citations are issued for eight or ten MPH over the limit. The dramatic bunching in the speed distribution suggests systematic manipulation by officers. Specifically, the distribution implies the practice of speed discounting, where officers observe drivers traveling at higher speeds but write down nine MPH on the citation as a form of lenience (Anbarci & Lee 2014; Goncalves & Mello 2021). An officer's decision whether to bunch a driver, resulting in either a discounted or full fine, is the focus of our study.

We rely on several pieces of evidence to demonstrate that bunching in the speed distribution is generated by the behavior of officers rather than drivers (e.g., Traxler et al. 2018). First, following Goncalves & Mello (2021), figure 1 shows that all bunching is attributable to a subset of *lenient* officers.[3] About 25 percent of officers, whom we term the non-lenient officers, almost never write tickets for nine MPH. In appendix A, we further document the variation across officers in bunching propensity, show that an officer's bunching propensity

[3]See appendix D-2 for details on the classification of officers as lenient versus non-lenient, which is based on the manipulation test from Frandsen (2017). To ensure that the pattern in figure 1 is not mechanical and to avoid the reflection problem in IV estimates, we randomly partition an officer's stops into two groups, classify each officer × partition as lenient versus not, and then use the officer's classification in the *other* partition.

is strongly correlated across space and time, and find that the identity of the stopping officer is significantly more predictive of a bunched ticket than locations or motorist characteristics.

A natural question here is why do officers bunch drivers? First, we note that fine revenue is routed to the county government where the citation was issued. Hence, neither the officers themselves, nor the FHP or state government, have any financial stake in fine amounts. Officers may 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 (Goncalves & Mello, 2021).

Based on the available evidence, our view is that distaste for traffic court best explains officer lenience in this context. After receiving a traffic ticket, the cited driver has the option to contest the citation in traffic court, with the citing officer expected to attend the associated court hearing. Using the same identification strategy that we exploit to assess the causal effect of sanctions on offending, we find that a 125 dollar increase in fine (causally) increases the likelihood that a driver contests a ticket in court by about 40 percent (see table 2). Hence, distaste for appearing in traffic court generates an incentive to bunch drivers and heterogeneity in distaste for traffic court could explain the observed variation in lenience across officers.

## 3 Empirical framework

Our empirical approach leverages the variation across officers in the propensity to bunch drivers within an instrumental variables framework. The outcome of interest, $Y_i$, is whether cited driver $i$ commits a new traffic offense in the following year. The treatment of interest is whether driver $i$ receives a harsh fine (as opposed to a lenient one), which we denote by $D_i = \mathbf{1}[\text{speed}_i \geq 10]$. The instrument, which we call officer stringency, is computed as:

$$Z_{ij} = 1 - \left( \frac{1}{N_j - 1} \sum_{k \neq i} \mathbf{1}[speed_{kj} = 9] \right) \equiv \text{stringency}$$

where $i$ indexes motorists and $j$ indexes officers. In words, $D_i$ is an indicator for whether a motorist is not bunched at 9 MPH over and $Z_{ij}$ is the fraction of officer $j$'s citations issued to all *other* drivers that are for speeds of 10 MPH or more over the limit; or in other words, the fraction of citations that are not bunched at 9 MPH.

To adjust for differential exposure of officers to groups of motorists based on patrol shift assignments, we condition on detailed beat-shift fixed effects, denoted by $\psi_s$, in all our analyses. These beat-shift effects are at the level of the county $\times$ $\mathbf{1}$[highway] $\times$ year

$\times$ month $\times$ $\mathbf{1}$[weekend] $\times$ shift. A county is approximately a patrol area for each officer. Officers rotates shift (day of week and time of day) monthly.

Our empirical framework requires that the stringency instrument satisfy the local average treatment effect (LATE) assumptions of Imbens & Angrist (1994):

1. *Relevance.* $D(Z)$ is a nontrivial function of $Z$.
2. *Exogeneity.* $\{Y_{i1}, Y_{i0}, D_i(Z)\} \perp Z \mid \psi$
3. *Exclusion.* $Y_i(D, Z) = Y_i(D)$
4. *Monotonicity.* $\forall w, j \in J$, either $D_i(w) \geq D_i(h)\ \forall i$ or $D_i(w) \leq D_i(h)\ \forall i$

where $J$ denotes the set of officers and $\{Y_{i1}, Y_{i0}\}$ are the potential outcomes of driver $i$ when sanctioned harshly ($D = 1$) and leniently ($D = 0$).

The relevance assumption requires a statistical relationship between stringency and harsh fines, which is empirically testable. Panel (c) of figure 2 plots the probability of harsh fines against officer stringency, conditional on beat-shift fixed effects, laid over a histogram of stringency, net of beat-shift effects. The figure documents a linear and statistically precise relationship, with an estimated first stage coefficient of $\hat{\beta} = 0.944$ ($se = 0.006$) and associated $F \approx 22,000$. In figure A-3, we show the first stage estimates for other sanction measures. In terms of fine amounts, shown in panel (a), the estimated first stage is $\hat{\beta} = \$122$. We further discuss the exogeneity, exclusion, and monotonicity assumptions in turn below.

***Exogeneity***. Existing studies using examiner designs (e.g., Kling 2006, Dobbie & Song 2015, Maestas et al. 2013, Bhuller et al. 2020) have appealed to the institutional quasi-random assignment of examiners (e.g., bail judges) to satisfy the exogeneity assumption. Citing officers in our setting are, of course, not randomly assigned to drivers. Instead, officers can select their own samples by choosing (i) whom to pull over versus whom to let pass and (ii) whom to cite versus whom to issue a formal or informal warning. We cannot observe formal or informal warnings in our data and cannot observe the full population of drivers passing by an officer during a given shift.

An especially salient threat to our empirical design would be a correlation between stringency on the citing margin (whom to cite versus not) and the charging margin (whom to bunch versus not). To help illustrate this point, suppose there were two officers, $j \in \{a, b\}$, with $a$ an officer who bunches most drivers and $b$ an officer who bunches fewer drivers. Suppose that $a$ is also very lenient on the citing margin; that is, she lets most motorists pass with no citation, while $b$ is very stringent on the ticketing margin, citing most drivers. If $a$ restricts her sample by only citing "worse" drivers, then $E(Y_{i0} \mid j = a) > E(Y_{i0} \mid j = b)$, violating exogeneity.

There are two testable implications of the hypothesis that lenience on the intensive (bunch versus not) and extensive (ticket versus not) margins are correlated. First, our instrument $Z$ should be correlated with an officer's citation frequency. Holding constant the supply of offenders, officers with higher ticketing thresholds should have "missing" tickets relative to officers with lower ticketing thresholds. Second, $Z$ should be correlated with driver characteristics that predict reoffending. We test both these predictions in figure 2. Panel (a) plots the relationship between officer stringency and an officer's average monthly citations, both adjusted for beat-shift fixed effects. For both all citations and speeding citations, regression coefficients are quantitatively small and statistically indistinguishable from zero. Panel (b) illustrates that there is no relationship between stringency and either past offending or predicted reoffending based on driver covariates.[4]

Hence, the evidence suggests that exogeneity violations generated by sample selection are unlikely. Nonetheless, we take sample selection concerns seriously and subject our treatment effect estimates to a series of associated robustness checks, described further in section 4.

***Exclusion***. The exclusion restriction requires that officer stringency affects future offending only through sanctions. Note that our strategy allows other (non-sanction) officer behaviors to affect drivers as long as those behaviors are uncorrelated with our stringency measure (Frandsen et al., 2019). On the other hand, features of the officer-driver interaction other than the sanction that cause a driver to change behavior would violate exclusion if those features are correlated with stringency.

Another plausible source of exclusion violations is downstream involvement in the traffic court system. As previously mentioned, stringency increases the likelihood that a driver contests a ticket in court and might influence traffic school elections. If anything about the court experience changes driver behavior, that could be considered an exclusion violation. However, whether traffic court involvement constitutes a violation of exclusion or simply a mechanism for the fine's impact is subject to interpretation. When viewed from the officer's perspective, downstream events that are (i) caused by harsher sanctions and (ii) reduce reoffending still could be interpreted as a causal effect of sanctions themselves.[5]

Finally, the choice to bunch a driver indirectly affects the statutory "points" a driver

[4]Appendix table A-2 shows the relationship between all driver characteristics and recidivism, charged fines, and officer stringency. Driver covariates have substantial joint predictive power over reoffending ($F = 1734$) and reduced charges ($F = 29$), but considerably less power to predict officer stringency ($F = 2.7$). While our test rejects the null hypothesis of no statistical relationship between observables and officer stringency, a joint $F = 2.7$ is remarkable small in a setting without institutional random assignment and $N \approx 1.7$M.

[5]Moreover, the evidence is largely inconsistent with the court system playing an important role in generating the treatment effects. As shown in figure B-4, treatment effects are very similar for local and non-local drivers. Because drivers need to travel to the citation county to attend court, local drivers are more likely to contest citations.

receives on their license, as discussed in section 2.3. Appendix figure A-3 illustrates that officer stringency slightly affects statutory points ($\beta_{FS} = 0.7$). However, drivers can mitigate their point exposure through the court system, and we find that, taking into account those downstream behaviors, there is almost no relationship between stringency and points, again shown in figure A-3. Hence, the burden of accrued license points cannot explain the effects we observe.

***Monotonicity***. Monotonicity violations are a natural concern in our setting given evidence of racial bias in officer leniency decisions (e.g., Goncalves & Mello 2021). Importantly, Frandsen et al. (2019) show that 2SLS estimates in examiner designs recover the appropriate local average treatment effect under a weaker average monotonicity condition, which requires only that counterfactual reassignment to a more stringent officer increases the probability of harsh sanctions in expectation. In table A-3, we provide evidence for average monotonicity by showing that the first stage is statistically strong and similar in magnitude across motorist subsamples.[6]

To assess the overall impact of discretion, we estimate a set of treatment effect parameters beyond the LATE. Our method for recovering these parameters, described in section 5, does not rely on a monotonicity assumption.

# 4 Deterrence effects

Given our interest in the safety implications of discretion, characterizing the average causal effect of sanctions on motorist behavior is an important first step in our analysis. Our approach follows directly from the empirical framework discussed above. We estimate the regression $Y_{ijs} = \beta D_{ij} + \psi_s + u_{ijs}$ by two-stage least squares, instrumenting for $D_i$ with officer stringency $Z_{ij}$. Given the assumptions discussed in section 3, our deterrence IV estimate will recover a local average treatment effect (LATE), which is a positive weighted average of treatment effects for individuals whose treatment status is shifted by the instrument (Imbens & Angrist, 1994).

Note that our stringency instrument solves an important identification challenge arising from the nonrandom assignment of punishments. Not only do statutory sanctions increase with offense severity, as shown in figure 1, but officers further manipulate fines, as discussed in section 2.4. Naive OLS estimates, presented in table B-1, illustrate both dimensions of the identification challenge well. A regression of one-year reoffending on the charged fine (in \$100's) and beat-shift fixed effects gives $\hat{\beta} = 0.043$ ($se = 0.002$), suggesting that harsher fines increase reoffending. Adding officer fixed effects increases the estimate to $\hat{\beta} = 0.055$ ($se = 0.002$), highlighting the nonrandom sorting of motorists into sanctions by officers.

[6] See appendix B-1 for additional discussion of monotonicity concerns.

## 4.1 Results

In figure 3, we show the dynamic relationship between officer stringency and traffic offending. Specifically, we plot estimated coefficients (and confidence bands) from regressions of the form:

$$Y_{ijs\tau} = \beta_\tau Z_{ij} + \psi_s + u_{ijs}$$

where $Y_{ijs\tau}$ is an indicator for whether driver $i$ receives a traffic citation in quarter $\tau$, which are quarters *relative* to the focal FHP citation. In the figure, $\tau = 0$ corresponds to the exact date of the focal FHP citation and $\tau = k$ corresponds to $k$ quarters before or after the focal citation. The figure illustrates that the stringency of the citing officer at $\tau = 0$ has no ability to predict offending over the previous eight quarters but predicts a stark decline offending immediately after the focal citation. Impacts persist over the first four quarters and fade out thereafter. Over the year following the focal citation, the reduced form estimate is $\hat{\beta} = -0.017$ ($se = 0.005$).

Table 2 presents 2SLS estimates for the full set of one-year offending outcomes.[7] Column 1 reports the lenient officer mean of the outcome, while columns 2 and 3 report IV estimates excluding and including controls for driver characteristics. To help interpret magnitudes, column 4 reports the implied fine elasticity, which is computed by regressing the outcome on the (continuous) fine amount, driver controls, and beat-shift fixed effects, instrumenting the fine amount with stringency, and then scaling the IV estimate by the ratio of the average fine and average reoffending rates.

We find that harsh fines reduce the likelihood of a new traffic offense in the following year by about 1.6 percentage points ($\epsilon = -0.07$). The majority of this effect is attributable to reductions in speeding offenses; a harsh fine reduces the likelihood of a new speeding offense in the next year by about 1.4 percentage points. The 2SLS estimate is precisely estimated, with a 95 confidence interval of $(-0.017, -0.012)$. Our point estimate for speeding offenses represents an 8.5 percent decline relative to the lenient officer mean and implies a fine elasticity of $-0.13$. In other words, our estimate implies that a doubling of the fine amount would reduce the likelihood of speeding recidivism by 13 percent.

Estimated impacts of harsh fines on non-speeding offenses are also statistically significant but less pronounced ($\epsilon = -0.06$). The finding that speeding sanctions reduce other traffic offenses is consistent with Gehrsitz (2017), who finds specific deterrence effects of short-term license suspensions imposed on speeders in Germany on all forms of traffic offending.

Consistent with reductions in traffic offending implying a true behavioral response on

[7] In the appendix, we present graphical versions of the reduced form estimates (figure B-1), dynamic versions of the reduced form for other outcomes (figure B-2), and the full set of first stage and reduced form estimates with and without controls (table B-2).

the part of drivers, we also find that harsh fines reduces the likelihood of crash involvement over the following year by between 0.2 and 0.3 percentage points ($\epsilon = -0.04$). While less precisely estimated than the effects on traffic offenses, the IV estimates for crash involvement are statistically significant at the 10 percent level.

Finally, following our discussion in section 2.4, the last row of table 2 reports IV estimates of the impact of harsh fines on the likelihood that a driver contests a ticket in court. Relative to a lenient officer mean of 0.26, we find that a harsh fine increases the likelihood of a contested citation by about 11 percentage points, or about 42 percent, consistent with our hypothesis that court aversion motivates officer lenience.

## 4.2 Robustness

In the appendix, we present results from a battery of robustness checks. Table B-3 shows that estimated deterrence impacts are robust to alternative methods for computing the stringency instrument. Figure B-3 illustrates that our findings cannot be explained by differential selection of motorists across officers. We also show that results are similar when further interacting our beat-shift effects with stretch-of-road fixed effects, constructed by mapping the subset of geocoded tickets ($N = 244,858$) to Florida roads. In section B-1, we discuss additional robustness tests to address concerns about monotonicity.

## 4.3 Interpretation

As highlighted in section 2, motorists issued harsh and lenient fines face the same sanctions for future offenses. Hence, our estimates capture a pure *specific deterrence* effect (e.g., Nagin 2013), or a behavioral responses to the experience of punishment, rather than the effects of statutorily higher sanctions for future offenses.

In figure B-4, we show that treatment effects are similar for local residents and out-of-county drivers, suggesting a minimal role for the traffic court system in explaining treatment effects, since drivers need to travel to the citation county to attend traffic court. Figure B-4 also illustrates that offending responses are nearly identical for motorists with higher and lower incomes, proxied by the zip code of residence, which runs counter to a financial distress mechanism based on Mello (2021).

A mechanism that seems particularly consistent with the dynamic patterns in figure 3 is drivers updating their beliefs about sanctions (Dusek & Traxler, 2021). In figure B-5, we show that offending responses are stronger in the county of the focal ticket and larger for motorists that have been exposed to stringent officers in the past, both of which are consistent with a learning hypothesis.

# 5 The impact of discretion

Having shown that officer's discretionary choices of sanctions matter for public safety by estimating the local average treatment effect of harsher fines on reoffending, we now examine the overall implications of officer discretion over sanctions for public safety. Our approach is to compare the observed reoffending rates with reoffending rates in a counterfactual scenario without officer discretion. In other words, our aim is to identify reoffending rates were all motorists issued harsh fines. This counterfactual reoffending rate is identified by an estimate of the average treatment effect for motorists currently issued lenient fines.

Our strategy for estimating average treatment effects for treated and untreated motorists again exploits across-officer variation in stringency and proceeds in two steps. First, we rely on an identification at infinity approach to identify the average treated and untreated potential outcomes for the entire sample of cited motorists. Second, we combine these estimates with the observed reoffending outcomes for treated and untreated motorists to identify average treatment effects on the treated (ATT) and untreated (ATU).

## 5.1 Estimating treated and untreated potential outcomes

To build intuition for our first step, consider a supremely stringent officer $j$ who always issues harsh fines, $Z_{ij} = 1$. Motorists cited by such an officer always receive the harsh fine, $D_{ij} = 1$. Hence, the observed reoffending rates for motorists stopped by officer $j$ represent the average treated potential outcomes for that subset of motorists, $E(Y_i|J = j) = E(Y_{i1}|J = j)$. Assuming that officers are as-good-as-randomly assigned, the average treated potential outcomes for this group of motorists are the same as those for the entire sample of motorists, $E(Y_{i1}|J = j) = E(Y_{i1})$. By the same logic, the outcomes of motorists stopped by a supremely lenient officer with $Z_{ij} = 0$ can identify $E(Y_{i0})$. This reasoning follows Hull (2020) and Arnold et al. (2022) and is in the spirit of identification at infinity in selection models (e.g., Chamberlain 1986; Heckman 1990; Andrews & Schafgans 1998).

In terms of implementing this idea in practice, we face two relevant issues. First, although our data indeed include supremely stringent and lenient officers who always and never issue harsh fines (as shown in panel a of figure A-1), our sample size is greatly reduced when focusing only on these extreme officers. In particular, only 1,245 citations, or 0.07 percent of our sample, are issued by supremely lenient officers. Hence, as our baseline approach, we follow Arnold et al. (2022) and rely on extrapolation from the data away from the tails, described further below.

The second practical issue is that as-good-as-random assignment of officers holds only within beat-shifts, $\psi_s$, and hence, only within-$\psi$ variation should be used for extrapolation. Our baseline approach is to simply include these fixed effects in our extrapolation regressions.

In other words, we estimate regressions of the form:

$$E(Y_i|Z_{ij},\psi_s) = \alpha\psi_s + f(Z_{ij}) + u_{ijs} \tag{1}$$

and then compute the conditional expectation at the values $Z_{ij} = 0$ and $Z_{ij} = 1$ at the average value of the beat-shift effects, $E(Y_{i0}) = \hat{\alpha}\bar{\psi}_i + f(0)$ and $E(Y_{i1}) = \hat{\alpha}\bar{\psi}_i + f(1)$. Below, we consider various functional forms for $f(\cdot)$, including polynomials, splines, and a fully non-parametric specification that relies only on the ($\psi$-adjusted) average outcomes for officers with $Z_{ij} < 0.01$ and $Z > 0.99$ to estimate $E(Y_{i0})$ and $E(Y_{i1})$.

Including the beat-shift fixed effects in this way requires an auxiliary linearity assumption, discussed in detail in Arnold et al. (2022). As robustness, we consider alternative approaches to adjusting our extrapolation estimates for differences across beat-shifts. First, we estimate the above regression separately by troop, and then aggregate up the troop-level estimates. We also estimate a version of the extrapolation based on the within-locations approach of Feigenberg & Miller (2022), described further in section E-1. Both approaches directly address the concern that our estimates of $E(Y_{i0})$ and $E(Y_{i1})$ are identified off officers patrolling different beats by relying only on comparisons between officers patrolling the same locations to estimate both.

## 5.2 Recovering treatment effect parameters

After identifying average potential outcomes, our second step notes that the potential outcomes of treated and untreated groups must average to the sample-wide average potential outcomes. For the average treated outcome,

$$\begin{aligned} E(Y_{i1}) &= pE(Y_{i1}|D_i = 1) + (1-p)E(Y_{i1}|D_i = 0) \\ &= pE(Y_i|D_i = 1) + (1-p)E(Y_{i1}|D_i = 0) \end{aligned}$$

We estimate the left-hand side via extrapolation in the first step. The first term on the right-hand side can be estimated directly from the data. Rearranging gives an expression for the final term, the treated outcome for untreated motorists, $E(Y_{i1}|D_i = 0) = \frac{1}{1-p}E(Y_{i1}) - \frac{p}{1-p}E(Y_i|D_i = 1)$. Combined with the observed outcomes for untreated motorists, we can identify the average treatment effect on the untreated:

$$\begin{aligned} ATU =& E(Y_{i1}|D_i = 0) - E(Y_{i0}|D_i = 0) \\ =& \frac{1}{1-p}E(Y_{i1}) - \frac{p}{1-p}E(Y_i|D_i = 1) - E(Y_i|D_i = 0) \\ =& \frac{1}{1-p}[E(Y_{i1}) - E(Y_i)] \end{aligned} \tag{2}$$

where the last equation uses the identity $pE(Y_i|D_i = 1) + (1 - p)E(Y_i|D_i = 0) = E(Y_i)$.

The final equation provides useful intuition for how we identify the ATU. Once we have estimated the average treated potential outcome, the difference between that value and the observed average reoffending outcome $E(Y_i)$ reflects the difference in outcomes due to individuals who have not been treated, and the division by $(1 - p)$ scales this difference for the share of individuals who are untreated.

By the same logic, we can identify the untreated potential outcome for treated drivers from the relationship $E(Y_{i0}|D_i = 1) = \frac{1}{p}E(Y_{i0}) - \frac{1-p}{p}E(Y_i|D_i = 0)$, which gives us the average treatment effect on the treated,

$$\begin{aligned} ATT =& E(Y_{i1}|D_i = 1) - E(Y_{i0}|D_i = 1) \\ =& E(Y_i|D_i = 1) - \frac{1}{p}E(Y_{i0}) + \frac{1-p}{p}E(Y_i|D_i = 0) \\ =& \frac{1}{p}[E(Y_i) - E(Y_{i0})] \end{aligned} \tag{3}$$

Again, the final equation makes clear how the ATT is identified. The difference between the average reoffending rate and the untreated potential outcome reflects the difference in outcomes due to individuals who have been treated, and the division by $p$ scales this difference for the share of individuals who are treated.

To compute standard errors, we rely on a Bayesian bootstrap (Rubin, 1981), clustering at the officer-level and bootstrapping the entire procedure (both steps). The Bayesian bootstrap is a special case of the standard bootstrap procedure, where instead of resampling with replacement, random weights are drawn in each iteration. This procedure has the advantage of preserving support of the stringency instrument and the beat-shift fixed effects in each iteration.

Notice that, while our procedure for recovering treatment effect parameters of interest continues to rely on the assumptions of officer instrument relevance, exogeneity, and exclusion discussed in section 3, our approach does not require making a monotonicity assumption about the treatment behavior of officers. In that sense, our estimation of the ATE, ATT, and ATU requires fewer assumptions than were needed to recover the LATE above. Implementation of the extrapolation procedure for estimating the sample average potential outcomes requires specifying a functional form for the conditional expectation function, $E(Y|Z, \psi)$, but our setting permits nonparametric specifications, and none of the conclusions we discuss below are sensitive to functional form assumptions.

## 5.3 Extrapolation results

Panel (a) of figure 4 illustrates our extrapolation-based estimation of the sample average potential outcomes. The figure plots the probability of reoffending on the vertical axis against officer stringency on the horizontal axis. We show both a nonparametric binscatter, adjusted for beat-shift effects using the method of Cattaneo et al. (2021), as well as the fitted line from an estimate of equation 1. We begin with a simple quadratic specification, which fits the data quite well, especially in the tails of the stringency distribution, and discuss sensitivity to functional form assumptions below. The value of the quadratic fit at $Z = 0$ and $Z = 1$ provide estimates of the average potential outcomes, $E(Y_{i0}) = 0.355$ and $E(Y_{i1}) = 0.343$, as well as an estimate of the average treatment effect, $-0.012$.

In panel (b) of figure 4, we present our estimates of the ATT and ATU, which are obtained by combining our extrapolation estimates of $E(Y_{i0})$ and $E(Y_{i1})$ with treatment-specific observed outcomes, as described above. The vertical axis plots, for each group of motorists, the untreated potential outcome, $E(Y_{i0})$, which we refer to as the "reoffending rate," as well each group's treatment effect, $E(Y_{i1} - Y_{i0})$. Although treated motorists reoffend at the highest rates, $E(Y_{i0}|D_i = 1) = 0.371$, they exhibit small and statistically insignificant treatment effects, $E(Y_{i1} - Y_{i0}|D_i = 1) = -0.006$ ($se = 0.007$). On the other hand, untreated motorists reoffend significantly less often, $E(Y_{i0}|D_i = 0) = 0.325$ and exhibit sizable and statistically significant treatment effects, $E(Y_{i1} - Y_{i0}|D_i = 0) = -0.023$ ($se = 0.009$). To summarize, we find positive selection on levels and reverse selection on gains: relative to untreated motorists, treated individuals reoffend more often but are less responsive to harsh fines.

Notice that the treatment effect for currently untreated motorists is about twice as large (more negative) as the ATE and forty percent larger than the LATE estimate from section 4, highlighting the importance of our approach for assessing the overall impacts of discretion. Relative to the untreated (potential) reoffending levels, untreated motorists reduce their reoffending by 7.1 percent when issued harsh fines, whereas the comparable figures for the average motorist and treated motorists are 3.4 percent and 1.6 percent, respectively.

Given the specific feature of our setting, where officers deviate from fine rules by giving breaks to a subset of motorists, the ATU identifies the impact of removing officer discretion over sanctions. Currently, the probability of a harsh fine is $p = E(D_i) = 0.657$ and the reoffending rate is $E(Y_i) = 0.351$. Without discretion, the probability of a harsh fine would become one and the reoffending rate would decline by $(1-p) \cdot \text{ATU} = 0.78$ percentage points, or about 2.2 percent. In a typical year in our data ($N = 172,000$ citations), this corresponds to about 4,000 additional traffic offenses attributable to officer discretion. Note that these are additional offenses due to just one margin of discretion in one police department. We discuss some alternative interpretations of this magnitude below.

### 5.4 Decomposing safety gains

Conceptually, we can think of safety gains from eliminating discretion as working through two distinct channels. The removal of discretion increases the share of harsh fines from $p = 0.657$ to $p = 1$, which we would expect to reduce reoffending based on our analyses in section 4, but also changes which drivers face harsh fines. A simple decomposition based on estimated treatment effect parameters can shed light on the quantitative importance of these two channels.

Specifically, the net safety gain from eliminating discretion is given by:

$$\begin{aligned} \Delta Y &= (1-p) \cdot \text{ATU} \\ &= \underbrace{(1-p) \cdot \text{ATE}}_{\text{average}} + \underbrace{(1-p) \cdot (\text{ATU} - \text{ATE})}_{\text{sorting}} \end{aligned}$$

The first term captures the safety gains from increasing the likelihood that the average motorist faces harsh fines, which we call the *average* component. The second term captures additional safety gains (or losses) attributable to the differences in treatment effects for the average motorist and currently untreated motorists, which we call the *sorting* component.

The sorting component equals zero if the ATT and ATU are equal. This would be the case with homogeneous treatment effects or if, for example, officers randomized motorists into sanctions. With selection on gains, the average and sorting components are opposite-signed. The realized safety gain from eliminating discretion is less than would be predicted by the ATE only, because the newly punished motorists are less-responsive than the average driver. With reverse selection on gains, the two components work in the same direction. There is a safety benefit from increasing harshness for the average driver, and then an *additional* safety benefit from reallocating fines to a particularly responsive group of motorists.

As mentioned above, we find reverse selection on gains, with treatment effects larger among the currently untreated motorists. Our estimates indicate that eliminating discretion would reduce the recidivism rate by 0.78 percentage points. The decline explained by increasing fine harshness for the average driver is given by $(1-p)\cdot\text{ATE} = 0.41$ percentage points (53 percent of the total), while the decline explained by the reallocation of fines is given by $(1-p)\cdot(\text{ATU}-\text{ATE}) = 0.37$ percent points (47 percent of the total). Hence, we find that each channel is equally important in explaining reduced reoffending rates without discretion.

This is an especially important point that warrants further discussion. Our decomposition reveals that officer discretion is actually more harmful to public safety than it might otherwise be, specifically because of which motorists are currently sorted into harsh sanctions. A useful way to see this point is to note that the total feasible deterrence in this setting is given by the ATE $= -0.012$, which is the per-person decline in reoffending induced by harsh fines if every driver were issued a harsh fine. Currently, officers achieve just $p \cdot \text{ATT} = -0.004$ in deterrence, or only about one third of the total feasible deterrence.

We can also note that, based on the ATT and ATU, safety improvements are feasible even holding the share of harsh fines constant at the current rate. Specifically, consider a counterfactual reallocation of harsh fines which holds constant the share of harsh fines at $p = 0.657$ but reallocates harsh fines to all currently untreated motorists. Under this reallocation, currently untreated motorists would comprise 52 percent of the treated sample, with the remainder a random sample of the currently treated motorists. The achieved deterrence in this reallocation is $p \cdot (0.52 \cdot \text{ATU} + 0.48 \cdot \text{ATT}) = -0.0098$, or 83 percent of the feasible deterrence. If officers had instead sorted motorists in this way, the safety gain from eliminating discretion would amount to just a 0.22 percentage point decline in the reoffending rate.

## 5.5 Robustness

Table 3 explores the robustness of our extrapolation-based estimates. In panel (a), we present results based on polynomial versions of equation 1, with the first row corresponding to our baseline quadratic specification. In panel (b), we show estimates based on non-parametric specifications, which take the $\psi$-adjusted average reoffending rate for officers in the left and right tails of the stringency distribution as our estimates of $E(Y_{i0})$ and $E(Y_{i1})$, using different bandwidths to define these "extreme" officers. In panel (c), we replicate our baseline quadratic specification, replacing our measure of officer stringency with a propensity score, estimated by regressing harsh fines on stringency and beat-shift effects and constructing predicted values.[8]

The key parameter for computing our no-discretion counterfactual, the average treatment effect on the untreated (ATU), is shown in column 7. Our baseline estimate is $-0.023$ ($se = 0.009$) and the estimates from the various specifications range from $-0.015$ ($se = 0.008$) to $-0.28$ ($se = 0.009$). Importantly, the estimate from our most flexible specification, which

[8]See appendix figure C-2 for a graphical depiction of extrapolation-based estimates of $E(Y_{i0})$ and $E(Y_{i1})$ under various functional form assumptions. An important takeaway from figure C-2 is that our estimate of $E(Y_{i1})$ is completely insensitive functional form assumptions because we observe many officers who treat all motorists harshly. Many fewer officers treat all motorists leniently and, accordingly, our estimates of $E(Y_{i0})$ vary more across specifications (although nearly all estimates are within the 95 percent confidence interval of our baseline quadratic estimate). This fact, combined with equations 2 and 3, explains why our estimates of the ATU are much more consistent across specifications than our estimates of the ATT.

uses the $\psi$-adjusted average outcomes for officers with $Z < 0.01$ and $Z > 0.99$ and is shown in the third row of panel (c), is nearly identical to our baseline estimate (ATU $= -0.024$; $se = 0.009$).

Table 3 also illustrates the robustness of the broader selection patterns discussed above. In particular, column (8) summarizes the degree of reverse selection on gains by reporting the ATT$-$ATU difference, our estimates of which range from 0.009 to 0.029 and are statistically significant in some, but not all, specifications. In our most flexible specification, the associated $p$-value is 0.106, just outside of conventional significance levels. Note that in section 5.7 below, we briefly discuss estimates of this difference based on an alternative marginal treatment effects approach, which requires additional assumptions but delivers meaningful precision gains.

Column 4 summarizes the degree of positive selection on levels across specifications by reporting the difference in untreated reoffending rates for treated and untreated motorists These estimates are quite consistent across approaches, ranging from 0.032 to 0.046, and are statistically significant at conventional levels in all specifications.

In panel (d) of table 3, we report estimates from variations of our baseline quadratic specification which use alternative approaches for adjusting for beat-shift effects. In the first (second) row, we estimate our baseline quadratic specification, including beat-shift effects, separately for troops (counties), and then aggregate up the estimates, weighting by sample shares. In the third row, we replicate the within-locations approach of Feigenberg & Miller (2022), described in detail in appendix E-1. In all three specifications, we obtain similar estimates for the ATU and observe comparable degrees of selection on levels and reverse selection on gains.

## 5.6 Other offending outcomes

In table 4, we present extrapolation-based estimates from our baseline quadratic specification for other safety outcomes of interest. Focusing first on panel (a), the second row shows that patterns are similar when we focus only on reoffenses for speeding, a subset of our baseline outcome. While reoffending rates for speeding alone are slightly lower, the treatment effect on speeding recidivism for untreated motorists is actually even larger than the ATU in our baseline estimate when rescaled by the untreated reoffending rate. In the third row, we again find similar patterns for crash involvement. When issued lenient fines, motorists treated harshly are involved in a crash with probability 0.081, whereas motorists receiving lenient fines are involved in a crash with probability 0.77 (difference $= 0.004$; $se = 0.002$). The treatment effect on treated drivers is a small and statistically insignificant 0.001 ($se = 0.002$). For untreated drivers, the effect is $-0.005$ ($se = 0.003$), which is marginally significant but represents a sizable decline, about 6.5 percent, relative to their overall rate of crash involvement.[9]

Crash involvement has two key strengths as an outcome measure. First, as highlighted in section 4, this measure does not depend on police enforcement activity, so there is no concern that the outcome is sensitive to reporting. And second, car crashes carries direct social costs. A large literature has sought to identify the monetary cost from various forms of injury, using data from labor markets and other settings (see Viscusi & Aldy 2003 for an extensive review). We use the value of a statistical injury (VSI) to get a sense of the monetary benefit of these abated car accidents. We borrow a recent VSI estimate from Guardado & Ziebarth (2019) of $45,000, and multiplying this estimate with our ATU gives $225. To put this dollar value in perspective, giving a harsh ticket to the average untreated driver would increase their fine by approximately $125. So each additional dollar in fines issued to this group yields over one dollar and eighty cents in social value from reduced car accidents.

In panel (b) of table 4, we replace these binary outcome measures with counts of each type of incident for each driver over the two years following the focal FHP citation. In all cases, selection patterns are similar when using counts instead of binary outcomes. Focusing on the first row, columns 2–4 illustrate that selection on levels is comparable when using this alternative notion of reoffending: when issued lenient fines, treated motorists commit 0.24 ($se = 0.038$) more traffic offenses over the following two years than untreated motorists. As a percentage of the recidivism rate for untreated motorists (22 percent), this difference is slightly larger than the difference in our baseline specification (14 percent).

Columns 6–8 illustrate that reverse selection on gains is significantly more pronounced when examining the number of future traffic offenses. Here, the ATT is nearly identical to that when using the binary measure (first row), but the ATU is *four times larger*. Repeating our baseline counterfactual calculation, we find that eliminating discretion would result in $(1 - p) \cdot \text{ATU} = 0.095$ fewer offenses per cited motorist in the following two years. Based on the average number of traffic citations annually in our sample, this amounts to about 16,000 fewer expected traffic offenses over the next two years when removing discretion for all tickets in a single year.

## 5.7 Comparison with marginal treatment effects

An alternative approach to identifying treatment effect parameters beyond the LATE is to utilize the marginal treatment effects framework (Bjorklund & Moffitt, 1987; Heckman &

[9] In the appendix, figure C-5 graphically depicts the extrapolations underlying these estimates and and table C-1 reports the corresponding estimates when using local means to estimate $E(Y_{i0})$ and $E(Y_{i0})$ instead of the quadratic specification.

Vytlacil, 2007). In our setting, this approach supposes that selection into treatment can be modeled as a threshold crossing rule, $D = \mathbf{1}[\mu_D(Z) > U_D]$, which depends on characteristics $Z$, including both $X$ and our stringency instrument, and unobservable $U_D$. Each individual has a fixed value of $U_D$, their *resistance to treatment*. The higher one's value of $U_D$, the greater their realization of $\mu(Z)$ must be for that individual to take up treatment. Without loss of generality, we can assume $U_D$ has a uniform distribution, so that $\mu_D(Z)$ reflects the probability of treatment at $Z$ (i.e., the propensity score).

The marginal treatment effect is defined as $MTE(u) = E(Y_1 - Y_0|U_D = u)$, the treatment effect for individuals who are induced into treatment at propensity score $u$. In this framework, the ATU is defined as a weighted average of MTE's, ATU $= \int_0^1 MTE(u)h_{UT}(u)du$, where the weights, $h_{UT}(u) = \frac{Pr(\mu_D(Z) \leq u_D)}{E(1-\mu_D(Z))}$, reflect the probability that an individual at resistance to treatment $u$ has a sufficiently low propensity score to avoid treatment. The ATT is defined analogously, with weight $h_{TT} = \frac{Pr(\mu_D(Z) > u_D)}{E(\mu_D(Z))}$, and the ATE is defined with a uniform weight across MTE's.

The key assumption of the marginal treatment effects approach is that treatment is characterized by the threshold crossing rule: individuals have a fixed resistance to treatment $U_D$, and their propensity score $\mu_D(Z)$ depends only on observables and the instrument but is otherwise invariant across individuals. This rule implies the strict monotonicity of Imbens & Angrist (1994), since all individuals who take up treatment at a given value of $\mu_D(Z)$ would also take up treatment at greater values (Vytlacil, 2002). However, a strength of the marginal treatment effects approach is that, with assumptions on the functional form for $MTE(u)$, various treatment effects can be identified beyond LATE without necessarily having full support of the instrument (Mogstad et al., 2018; Mogstad & Torgovitsky, 2018).

In appendix figure C-9, we present estimated MTE curves for the effect of fines on reoffending. We show both polynomial parametric specifications for the marginal treatment effect, $MTE = \sum_{k=0}^{K} \theta_k v^k$, as well as a non-parametric specification.[10] In all cases, we find a pattern of reverse selection on gains, consistent with our extrapolation estimates: drivers with higher levels of resistance to treatment (i.e., a lower likelihood of facing harsh fines) exhibit more negative treatment effects.

In appendix table C-4, we use these MTE curves to construct estimates of the ATE, ATT, and ATU, based on the formulas above. The estimates are remarkably similar to the extrapolation-based estimates. In particular, the ATU estimates range from $-0.023$ to $-0.027$, aligning with our extrapolation estimate of $-0.023$. While the MTE approach

[10]We estimate the MTE as follows: we first estimate a first-stage regression of $D$ on $X$, $Z$, and beat-shift fixed effects $W$, and then construct an estimated propensity score $\hat{p}$ for each driver. We then estimate a conditional expectation of $Y$ on $X$, $W$, and a polynomial in $\hat{p}$. The derivative of the polynomial in $\hat{p}$ is the marginal treatment effect (Carneiro et al., 2011). We calculate standard errors using the officer-level bootstrap procedure described in Section 5.2.

requires a potentially strong assumption of strict monotonicity, our results offer reassurance that this approach provides reasonable estimates for a wide range of important treatment effects.

Moreover, table C-4 highlights a consequential benefit of imposing the additional assumptions required for the MTE approach: statistical precision. While ATU estimates based on extrapolation and MTE approaches are remarkably similar, standard errors are about half as large for the MTE-based estimates. These precision benefits are even more pronounced when examining our summary measure of reverse selection on gains (the ATT−ATU difference). While the $p$-values associated with this parameter hover between 0.1 and 0.2 in some extrapolation specifications (see table 3), this difference is statistically significant at the one percent level in all MTE specifications.

## 5.8 General deterrence effects

Our analysis of counterfactual reoffending rates without officer discretion focuses only on the *specific deterrence* effects of changing sanctions for cited motorists, holding that sample of cited motorists constant. An important caveat associated with our calculations, then, is that eliminating discretion may also have effects on offending through a *general deterrence* channel. Specifically, eliminating officer lenience would increase the expected fine associated with speeding, which may deter potential speeders.

General deterrence effects associated with eliminating officer lenience should, if anything, increase public safety, further increasing the net social benefit of eliminating discretion that we discuss above. Unfortunately, our data and setting do not permit credible identification of general deterrence effects associated with changing fine amounts. Moreover, as discussed in section 5.4, our findings imply that safety gains are feasible even holding general deterrence effects constant.

Importantly, the literature on criminal deterrence suggests that general deterrence effects associated with modest increases in fine amounts may be negligible. While a large literature has documented that potential offenders are deterred by increases in the probability of apprehension (e.g., Makowsky & Stratmann 2011, Mello 2019), evidence on the deterrence effects of changes in sanctions is considerably more mixed (Chalfin & McCrary, 2017). Even dramatic and highly salient changes in sanctions, such as the increased probability of carceral punishment at the age of majority or changes in capital punishment regimes have been shown to have minimal general deterrence effects (e.g., Lee & McCrary 2017; Chalfin & McCrary 2017).

# 6 Understanding officer behavior

Our analyses above yield the clear takeaway that current officer behavior generates reverse selection on gains: motorists who currently face harsh fines are less responsive to those harsh sanctions than those who currently receive lenient fines. We also find stark selection on levels: treated motorists reoffend at much higher rates than untreated motorists when issued lenient fines. In this section, we consider potential explanations for this officer behavior. Importantly, our central conclusion that officer discretion over sanctions harms public safety does not depend on the underlying reasons for this officer behavior. However, the potential for alternative policy tools, other than the altogether removal of trooper discretion, to improve safety could depend on the rationale underlying officer decisions.

## 6.1 Driver characteristics

One possibility is that the selection patterns we observe are driven by officers making sanction decisions based on salient driver characteristics, such as offending history or race, that are incidentally positively correlated with reoffending levels and negatively correlated with deterrability. Evidence that officers consider driver characteristics when making discretionary sanction choices can be seen easily in column 2 of table A-2, which illustrates that our full set of driver covariates are jointly predictive of a harsh charge ($F = 28.8$) after conditioning on beat-shift effects. Hence, a natural question is to what extent decisions based on driver observables can explain the patterns of selection on levels and reverse selection on gains that we document.

Figure 5 explores this question. We split motorists into 32 cells based on gender $\times$ race $\times$ $\mathbf{1}[\text{age} \geq 35]$ $\times$ $\mathbf{1}[\text{any citation in the past year}]$ and, within each cell, compute the average reoffending level $E(Y_{i0})$ and average treatment effect $E(Y_{i1} - Y_{i0})$ using a quadratic extrapolation. We then explore the relationship between these estimated group-specific parameters and the group-specific harshness rate (the rate at which motorists with those characteristics face harsh fines). As shown in panels (a) and (b), the types of motorists who are punished harshly more often also exhibit significantly higher reoffending rates and slightly less negative treatment effects. These patterns are consistent with our main findings and suggest that easily observable motorist characteristics can explain at least some share of the positive selection on levels and reverse selection on gains that we observe.

However, panel (c) of figure 5 also presents extrapolation-based estimates which are computed *within* these 32 covariate cells, with our baseline estimates also shown for comparison. The figure clearly illustrates that selection on levels and reverse selection on gains persist in these within-covariate specifications, although both are attenuated slightly, consistent with the discussion above. In our baseline specification, the difference in $E(Y_{i0})$ for treated and untreated motorists is 0.046, while the corresponding within-covariate difference is 0.033. Moving from our baseline specification to a within-covariate specification reduces the ATT−ATU difference from 0.016 to 0.1. Hence, officer sanction decisions based on driver characteristics can only explain about 30-35 percent of the overall selection patterns.[11]

Considering that offending history is likely an especially salient characteristic for officers when making their sanction decisions, figure 6 further explores heterogeneity based on driving history. Specifically, we compute extrapolation-based estimates of average potential outcomes and treatment effects for motorists with and without at least one traffic citation in the past two years (because we can observe at least two years of driving history for every motorist in our sample). Panel (a) plots the group-specific quadratic extrapolations and panel (b) plots the associated group-specific untreated offending levels and treatment effects by treatment status.

Note that, in both panels, we adjust the level of the axes, but keep the scale the same, because those with prior offenses reoffend at significantly higher rates. Despite these differences in the level of offending, however, we find strikingly similar selection patterns across groups.[12] Interestingly, this figure suggests that even among motorists without prior offenses, officer treatment decisions select motorists with a higher likelihood of reoffending but who are less deterrable.

## 6.2 The importance of stopped speed

Another possible explanation for the observed selection patterns is that officers may punish drivers based primarily on offense severity, or, in this case, the true rather than manipulated speed.. If driving speed were positively correlated with the likelihood of reoffending but negatively correlated with treatment effects, punishments based on speed could explain the selection patterns that we document.

While we can observe the true stopped speed of drivers who are given the harsh fine, this information is not directly observed for drivers given the lenient fine. However, we can adapt our extrapolation approach to estimate the average stopped speed of these drivers (which we explain in more detail in appendix E-2). While the average driver issued a harsh fine is stopped and ticketed for a speed of 19.1 MPH over the limit, we estimate that drivers given

[11] Appendix figures C-7 and C-8 probe the robustness of panels (a) and (b) of 5. See appendix figure C-3 for graphical depictions of the within-cell extrapolations underlying the treatment effect estimates in panel (c) of figure 5.

[12] In section 4.3, we discuss suggestive evidence that driver learning can partly explain the deterrence effects of harsh fines. However, figure 6 rules out that the observed selection patterns can be explained solely by a learning story, whereby untreated motorists are more responsive because of less information from prior citations. Reverse selection on gains persists even among motorists with the same citation history and thus similar access to information about speeding fines.

a lenient fine were stopped on average at 18.2 MPH over. We interpret this difference as evidence that, at least to some extent, officers exhibit retributivist preferences (e.g., Kaplow & Shavell 2006). In other words, officers prefer to issue harsher sanctions to more serious offenders.

How important is this preference for punishment based on "guilt" to the overall selection patterns we document? To answer this question, we first simulate a simple counterfactual under the assumption that, within each county, officers consider *only* stopped speed when issuing harsh sanctions, holding constant average harshness. Under this counterfactual, the average stopped speed for harshly and leniently treated drivers would be 21.5 and 14.8 MPH over, respectively, a much more pronounced difference than we empirically observe. Hence, speed alone can only partially explain how officers set punishment levels.

Next, we ask whether decisions based on offense severity can explain the fact that harshly treated drivers reoffend at higher rates. To do so, we use the supremely stringent officers to identify the reoffending rates of drivers at different speeds. We then estimate the distribution of stopped speeds for leniently treated drivers, using a variation of our baseline extrapolation approach. We find that while harshly treated drivers have a predicted reoffending rate of 0.362 based on their stopped speeds, this figure is 0.355 for leniently treated drivers. This gap is only about 15 percent of the estimated true difference in reoffending, shown in table 3. Overall, these estimates suggest that, while stopped speed is an important factor in officer's punishment decisions, it can only explain a small portion of the overall selection patterns.

## 6.3 Information

Another potential explanation for the patterns we document is that officers attempt to allocate harsh sanctions to more deterrable motorists but have poor information about how to achieve that goal. Under this hypothesis, we might expect selection patterns to change as officers acquire more information, and we present a simple test of this hypothesis in panels (a) and (b) of appendix figure C-6. We split officers based on experience as of January 2007, the start of our sample (officers who begin after 2007 are included in the low experience group), and estimate the ATE, ATT, and ATU for more and less experienced officers using our extrapolation-based approach. Although the figure suggests slightly weaker selection on levels among more experienced officers, the extent of reverse selection on gains is, if anything, more extreme among officers with at least five years of experience. We should note, however, that officers are not explicitly provided information on the reoffending outcomes associated with their citations, so this test yields only suggestive evidence at best.

On the other hand, officers receive explicit feedback on whether their citations are contested in court and, as discussed in section 2.4, we posit a distaste for traffic court as a central motivation for issuing lenient citations. Hence, a natural question is whether officers seem to allocate sanctions strategically to minimize their time in traffic court. Panels (c) and (d) of figure C-6 replicates the analysis in panel (a) and (b), replacing the outcome with whether the citation was contested in court. Note that harsh fines *increase* the likelihood that a citation is contested, hence the treatment effects go in the opposite direction of those in panel (b). This figure shows minimal evidence of selection on gains in either direction and no evidence of differential selection on gains by officer experience.

## 6.4 Discussion

Above, we show that that selection patterns cannot be explained away by officers making sanction decisions based on salient observables or guilt. We also find suggestive evidence against the hypothesis that officers allocate sanctions to maximize other objectives or that officers strive for safety maximization but act on poor information. This leaves two potential explanations for the allocation of hash fines that we observe.

First, officers may allocate harsh fines based on unobservable characteristics which are (incidentally) positively correlated with the likelihood of recidivism and negatively correlated with deterrability. For example, troopers likely issue harsh sanctions to motorists who behave confrontationally during the traffic stop. If these motorists tend to reoffend at higher rates and are undeterred by harsh fines, this would explain the selection patterns we find. Or alternatively, officers may explicitly prioritize issuing harsh sanctions to the offenders they perceive as having the highest risk of recidivism. These two explanations are observationally equivalent from the perspective of the econometrician and thus cannot be distinguished. But in either case, the current allocation of sanctions, which results in harsh fines for high-risk but undeterred motorists, likely reflects the desired allocation of officers.

On one hand, we might conclude that a mismatch between the deterrence objectives of the state and the goals of the police is an important reason why officer discretion harms public safety. On the other hand, we should acknowledge that the goal of allocating harsh sanctions to the highest-risk offenders may be a legitimate law enforcement goal or may, in fact, reflect the preferences of the state. In particular, this objective accords well with an intuitive notion of fairness: the "worst" drivers "deserve" the most severe punishments.

To the extent that the state values this notion of fairness, a key takeaway from our empirical analyses is the existence of a tradeoff between fairness and efficiency. Because the subgroup of motorists with the highest recidivism risk are the least responsive to harsh fines, officers face a tradeoff between achieving fairness and maximizing safety when sorting some, but not all, motorists into harsh sanctions. While the literature has discussed the potential importance of this tradeoff for the design of legal and criminal justice institutions from a theoretical perspective (e.g., Kaplow & Shavell 2006), we are the first to document its empirical relevance.

## 7 Conclusion

In this paper, we study the public safety implications of police discretion over sanctions for speeding offenses. First, relying on variation across officers in the propensity to issue harsh fines, we show that sanctions decisions have important deterrence effects. Comparing motorists cited in the same beat-shifts by officers of varying stringency, we find that higher fines reduce the likelihood of a new traffic offense, a new speeding offense, and crash involvement in the following year.

We then assess the overall contribution of discretion to public safety by comparing observed reoffending rates to those in a counterfactual scenario without officer discretion. This counterfactual is identified by the average treatment effect for motorists currently issued lenient fines (i.e., the average treatment effect on the untreated or ATU), which may differ from the local average treatment effect recovered by our 2SLS estimates. We rely on a novel, two-step approach to estimate the ATU which leverages identification at infinity (e.g., Hull 2020) and, importantly, does not require an instrument monotonicity assumption.

Based on our estimated ATU, eliminating officer discretion would reduce the reoffending rate by about two percent. We show that about half of this safety improvement can be attributed to increasing fine harshness for the average driver, while the other half can be attributed to changing which motorists receive harsh fines. The importance of the second channel stems from the fact that current officer decisions generate *reverse selection on gains*: the average treatment effect for motorists that currently receive lenient fines are nearly four times larger than those for motorists currently issued harsh fines. A back-of-the-envelope calculation suggests that every $125 in fines issued to the population of currently untreated motorists results in $225 in social welfare gains from reduced auto accidents.

We conclude by discussing some potential explanations for the officer behavior we document. To accompany the reverse selection on gains we find, we also show stark, positive selection on levels: motorists facing harsh fines reoffend at significantly higher rates than those facing lenient fines. These selection patterns (both levels and gains) persist within driver covariates, cannot be explained by motorist "guilt," and are present among experienced officers. Though not definitive, the preponderance of evidence aligns well with the view that the current allocation reflects the goals of officers. Specifically, officers target harsh sanctions to motorists with high recidivism risk, which accords with an intuitive notion of fairness. These motorists are also the least deterrable, and hence the "undoing" of this sorting when removing discretion gives rise to meaningful safety gains.

## References


Abaluck, J., Agha, L., Kabrhel, C., Raja, A., & Venkatesh, A. (2016). The determinants of productivity in medial testing: Intensity and allocation of care. *American Economic Review*, *106*(2), 3730–64.

Abdulkadiroglu, A., Pathak, P., & Schellenberg, J. (2020). Do parents value school effectiveness? *American Economic Review*, *110*(5), 1502–1539.

Abrams, D., Goonetilleke, P., & Fang, H. (2021). Do cops know who to stop? assessing optimizing models of police behavior wih a natural experiment. *Unpublished manuscript*.

Anbarci, N. & Lee, J. (2014). Detecting racial bias in speed discounting: Evidence from speeding tickets in Boston. *International Review of Law and Economics*, *38*, 11–24.

Andrews, D. W. & Schafgans, M. M. (1998). Semiparametric estimation of the intercept of a sample selection model. *The Review of Economic Studies*, *65*(3), 497–517.

Angelova, V., Dobbie, W., & Yang, C. (2023). Algorithmic recommendations and human discretion. *Unpublished manuscript*.

Arnold, D., Dobbie, W., & Hull, P. (2022). Measuring racial discrimination in bail decisions. *American Economic Review*, *112*(9), 2992–3038.

Ash, E. & MacLeod, W. B. (2015). Intrinsic motivation in public service: Theory and evidence from state supreme courts. *The Journal of Law and Economics*, *58*(4), 863–913.

Ashenfelter, O. & Greenstone, M. (2004). Using mandated speed limits to measure the value of a statistical life. *Journal of Political Economy*, *112*(S1), S226–S267.

Ba, B., Knox, D., Mummolo, J., & Rivera, R. (2021). Diversity in policing: The role of officer race and gender in police-civilian interactions in Chicago. *Science*, *371*(6530), 696–702.

Bandiera, O., Best, M. C., Khan, A. Q., & Prat, A. (2021). The allocation of authority in organizations: A field experiment with bureaucrats. *The Quarterly Journal of Economics*, *136*(4), 2195–2242.

Banfield, E. (1975). Corruption as a feature of governmental organization. *Journal of Law and Economics*, *18*(3), 587–605.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, *76*(2), 169–217.

Best, M. C., Hjort, J., & Szakonyi, D. (2017). Individuals and organizations as sources of state effectiveness. Technical report, National Bureau of Economic Research.

Bhuller, M., Dahl, G., Loken, K., & Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, *128*(4), 1269–1324.

Bjorklund, A. & Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *Review of Economics and Statistics*, *69*(1), 42–49.

Blincoe, L., Miller, T., Zaloshnja, E., & Lawrence, B. (2015). The economic and societal impact of motor vehicle crashes, 2010. *National Highway Safety Administration Technical Report*.

Brinch, C. N., Mogstad, M., & Wiswall, M. (2017). Beyond late with a discrete instrument. *Journal of Political Economy*, *125*(4), 985–1039.

Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, *101*(6), 2754–81.

Cattaneo, M., Crump, R., Farrell, M., & Feng, Y. (2021). On Binscatter. *Unpublished Manuscript*.

Chalfin, A. (2016). The economic cost of crime. *The Encyclopedia of Crime and Punishment*.

Chalfin, A. & Goncalves, F. (2021). The professional motivations of police officers. *Unpublished manuscript*.

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

Chamberlain, G. (1986). Asymptotic efficiency in semi-parametric models with censoring. *journal of Econometrics*, *32*(2), 189–218.

Chandra, A. & Staiger, D. (2020). Identifying sources of inefficiency in healthcare. *Quarterly Journal of Economics*, *135*(2), 785–843.

Cornelissen, T., Dustmann, C., Raute, A., & Schönberg, U. (2016). From late to mte: Alternative methods for the evaluation of policy interventions. *Labour Economics*, *41*, 47–60.

Dahl, G., Kostol, A., & Mogstad, M. (2014). Family welfare cultures. *Quarterly Journal of Economics*, *129*(4), 1711–1752.

DeAngelo, G. & Hansen, B. (2014). Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy*, *6*(2), 231–257.

Dobbie, W., Goldin, J., & Yang, C. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, *108*(2), 201–40.

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

Dusek, L. & Traxler, C. (2021). Learning from law enforcement. *Journal of the European Economic Association, forthcoming.*

Einav, L., Finkelstein, A., Ji, Y., & Mahoney, N. (2022). Voluntary regulation: Evidence from medicare payment reform. *Quarterly Journal of Economics*, *137*(1), 565–618.

Feigenberg, B. & Miller, C. (2022). Would eliminating racial disparities in motor vehicle searches have efficiency costs? *Quarterly Journal of Economics*, *137*(1), 49–113.

Finlay, K., Gross, M., Luh, E., & Mueller-Smith, M. (2022). The impact of financial sanctions in the us justice system: Regression discontinuity evidence from michigans driver responsibility program. *Unpublished manuscript*.

Frandsen, B., Lefgren, L., & Leslie, E. (2019). Judging judge fixed effects. *NBER Working Paper 25528*.

Frandsen, B. R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. In *Regression Discontinuity Designs: Theory and Applications* (pp. 281–315). Emerald Publishing Limited.

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

Goncalves, F. & Mello, S. (2021). A few bad apples? racial bias in policing. *American Economic Review*, *111*(5), 1406–1441.

Guardado, J. R. & Ziebarth, N. R. (2019). Worker investments in safety, workplace accidents, and compensating wage differentials. *International Economic Review*, *60*(1), 133–155.

Hansen, B. (2015). Punishment and deterrence: Evidence from drunk driving. *American Economic Review*, *105*(4), 1581–1617.

Heckman, J. (1979). Sample selection bias as specification error. *Econometrica*, *71*(1), 53–161.

Heckman, J. (1990). Varieties of selection bias. *The American Economic Review*, *80*(2), 313–318.

Heckman, J. J. & Vytlacil, E. J. (2007). Chapter 71 econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. volume 6 of *Handbook of Econometrics* (pp. 4875–5143). Elsevier.

Hull, P. (2020). Estimating hospital quality with quasi-experimental data.

Imbens, G. W. & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Kang, K. & Silveira, B. S. (2021). Understanding disparities in punishment: Regulator preferences and expertise. *Journal of Political Economy*, *129*(10), 2947–2992.

Kaplow, L. & Shavell, S. (2006). *Fairness versus welfare*. Harvard University Press.

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

Lee, D. & McCrary, J. (2017). The deterrence effec of prison: Dynamic theory and evidence. *Advances in Econometrics*, *38*(3).

Lipsky, M. (1980). Street-level bureaucracy: dilemmas of the individual in public services.

Luca, D. (2014). Do traffic tickets reduce motor vehicle accidents? Evidence from a natural experiment. *Journal of Policy Analysis and Management*, *34*(1), 85–106.

Maestas, N., Mullen, K., & Strand, A. (2013). Does disability insurance disourage work? using examiner assignment to estimate causal effects of ssdi receipt. *The American Economic Review*, *103*(5), 1797–1829.

Makowsky, M. & Stratmann, T. (2011). More tickets, fewer accidents: How cash-strapped towns make for safer roads. *The Journal of Law and Economics*, *54*(4), 863–888.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics*, *172*, 174–200.

Mello, S. (2021). Fines and financial wellbeing. *Working paper*.

Mogstad, M., Santos, A., & Torgovitsky, A. (2018). Using instrumental variables for inference about policy relevant treatment parameters. *Econometrica*, *86*(5), 1589–1619.

Mogstad, M. & Torgovitsky, A. (2018). Identification and extrapolation of causal effects with instrumental variables. *Annual Review of Economics*, *10*, 577–613.

Moore, M. S. (2019). The moral worth of retribution. In *Retribution* (pp. 337–388). Routledge.

Morris, C. (1983). Parametric empirical bayes inference: Theory and applications. *Journal of the American Statistical Association*, *78*(381), 47–55.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. *Unpublished working paper*, *18*.

Nagin, D. (2013). Deterrence: A review of the evidence by a criminologist for economists. *Annual Review of Economics*, *5*, 83–105.

NHTSA (2014). Traffic safety facts: 2012 data. *National Highway Safety Administration Technical Report*.

Norris, S. (2022). Examiner inconsistency: Evidence from refugee decisions. *Unpublished manuscript*.

O'Flaherty, B. & Sethi, R. (2019). *Shadows of doubt: Stereotypes, crime, and the pursuit of justice*. Harvard University Press.

Polinsky, A. M. & Shavell, S. (2000). The fairness of sanctions: Some implications for optimal enforcement policy. *American Law and Economics Review*, *2*(2), 223–237.

Prendergast, C. (2007). The motivation and bias of bureaucrats. *American Economic Review*, *97*(1), 180–196.

Rubin, D. (1981). The Bayesian bootstrap. *The Annals of Statistics*, *9*(1), 130–134.

Traxler, C., Westermeier, F., & Wohlschlegel, A. (2018). Bunching on the Autobahn? Speeding responses to a notched penalty scheme. *Journal of Public Economics*, *157*, 78–94.

Van Dijk, W. (2019). The socioeconomic consequences of housing assistance. *Unpublished manuscript*.

Viscusi, W. K. & Aldy, J. E. (2003). The value of a statistical life: a critical review of market estimates throughout the world. *Journal of risk and uncertainty*, *27*, 5–76.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, *70*(1), 331–341.

Weisburst, E. (2017). Whose help is on the way? the importance of individual police officers in law enforcement outcomes. *Unpublished manuscript*.

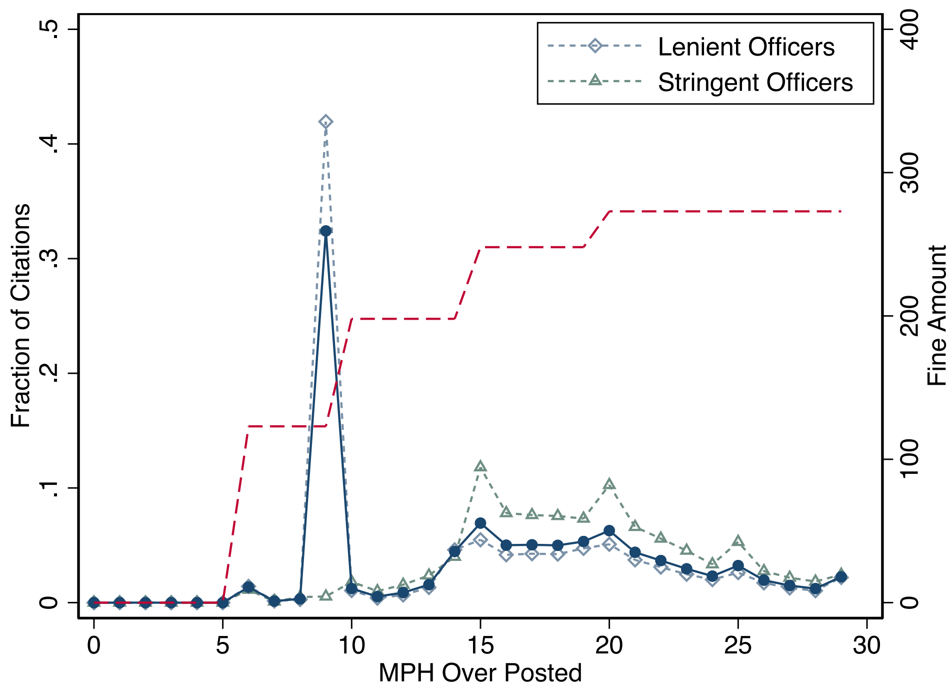
Wilson, J. (1989). Bureaucracy: What government agencies do and why they do it.

Table 1: Summary Statistics

| | (1) All | By Fines (2) Discounted | (3) Harsh |
|---|---|---|---|
| *Panel A: Demographics* | | | |
| Female | 0.384 | 0.415 | 0.368 |
| Age | 36.47 | 36.98 | 36.20 |
| Age Missing | 0.0002 | 0.0002 | 0.0002 |
| Race = White | 0.474 | 0.525 | 0.448 |
| Race = Black | 0.154 | 0.157 | 0.152 |
| Race = Hispanic | 0.187 | 0.144 | 0.209 |
| Race = Other | 0.041 | 0.034 | 0.044 |
| Race = Unknown | 0.144 | 0.140 | 0.147 |
| *Panel B: Socioeconomic Status* | | | |
| Zip Income | 57962 | 56459 | 58745 |
| Zip Income Missing | 0.101 | 0.107 | 0.097 |
| Vehicle Value | 17807 | 17297 | 18073 |
| Vehicle Info Missing | 0.143 | 0.139 | 0.145 |
| *Panel C: Offending History* | | | |
| Prior Prison Spell | 0.009 | 0.009 | 0.009 |
| Citation Past Year | 0.350 | 0.317 | 0.368 |
| Speeding Past Year | 0.179 | 0.158 | 0.189 |
| Other Past Year | 0.253 | 0.226 | 0.268 |
| Crash Past Year | 0.071 | 0.067 | 0.074 |
| *Panel D: Offense Characteristics* | | | |
| MPH Over Posted | 15.62 | 9.00 | 19.06 |
| Fine Amount | 207.70 | 123.00 | 251.79 |
| Contest in Court | 0.289 | 0.222 | 0.323 |
| Observations | 1,693,457 | 579,760 | 1,113,697 |

*Notes:* This table reports means for the analysis sample. See table A-1 for officer characteristics.

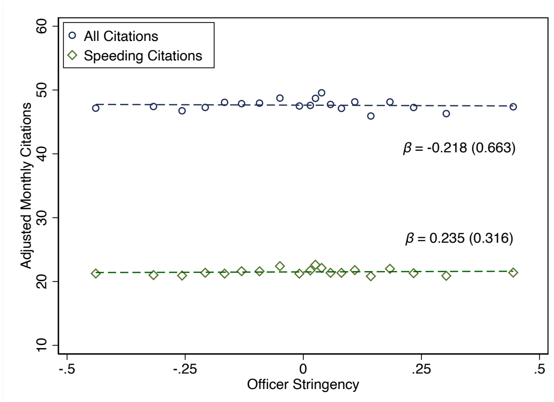
Figure 1: Fine Schedule and Charged Speed Distribution



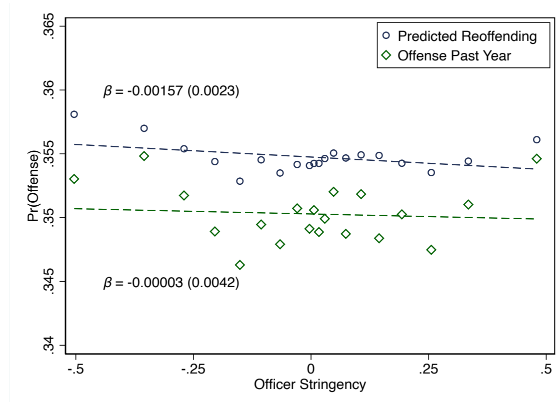

*Notes:* This figure plots the distribution of charged speeds on FHP-issued speeding citations in Florida. Dashed red line shows the fine schedule (right axis). Solid line and circles shows the aggregate distribution. Dashed lines with hollow diamonds and triangles plot the distribution for lenient and stringent officers, respectively, using the method described in section D.

Figure 2: Instrument Validity

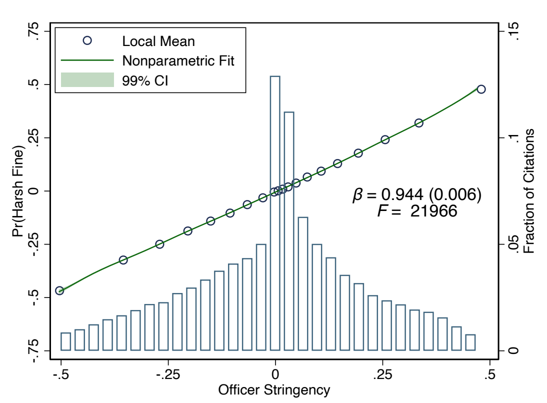
(a) Ticketing Frequency




(b) Predicted and Past Offending

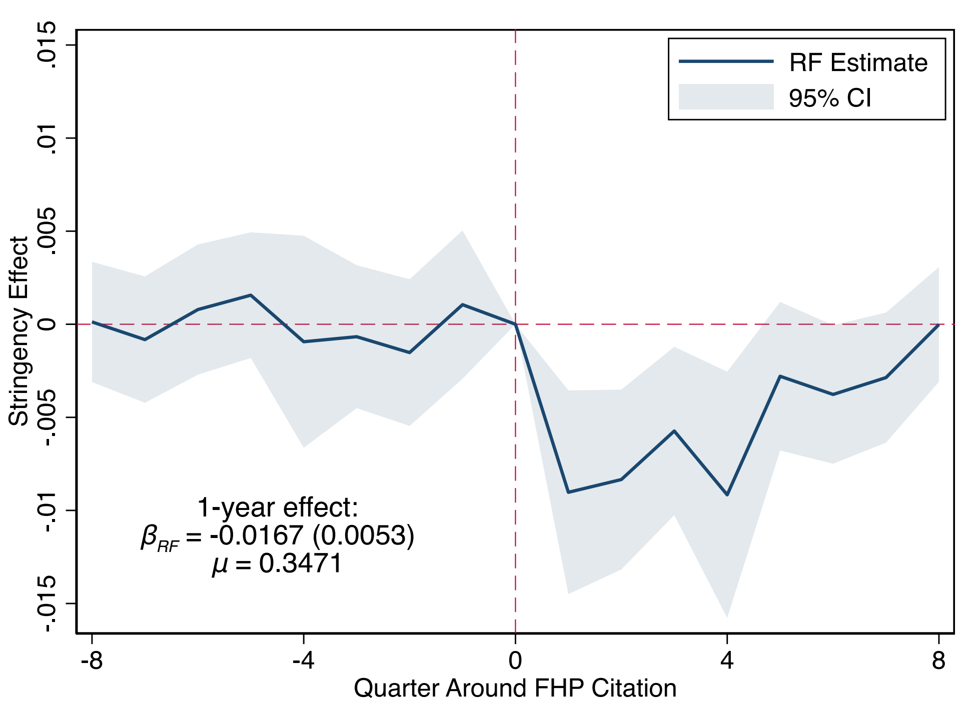



(c) First stage estimate




*Notes:* Panel (a) reports the relationship between officer stringency, residualized of beat-shift effects, and an officer's average monthly number of citations, adjusted for beat-shift effects. Blue circles report the relationship for all citations and green diamonds report the relationship for only speeding citations. Panel (b) reports the relationship between officer stringency, residualized of beat-shift effects, and predicted reoffending based on covariates and past offending, both residualized of beat-shift effects. Predicted reoffending is computed in opposite sample partitions using only tickets issued by non-bunching officers (see section D-3 for further details). Panel (c) shows the first stage relationship between stringency and the probability of a harsh fine, both residualized of beat-shit fixed effects (left axis). Local binscatter means are denoted by blue circles and the green line shows a non-parametric fit, with a 99 percent confidence interval indicated by the shaded region. Figure also illustrates a histogram of the officer stringency instrument, residualized of beat-shit fixed effects (right axis). Figure reports the linear first stage estimate and associated $F$-statistic.

Figure 3: Reduced Form Over Time




*Notes:* This figure reports coefficients on officer stringency from regressions where the outcome of interest is an indicator for whether the driver received a traffic citation in each quarter relative to the date of their focal FHP citation. $\tau = 0$ denotes the exact date of the focal FHP citation (one day only, where all motorists receive a citation so the effect of stringency is zero by construction). Regressions also include beat-shift fixed effects. Shaded region denotes 95 percent confidence intervals, constructed from standard errors clustered at the officer-level. Identical figures for other outcomes are shown in figure B-2. Figure reports the reduced form coefficient for one-year reoffending as well as the mean one-year reoffending rate for lenient officers.

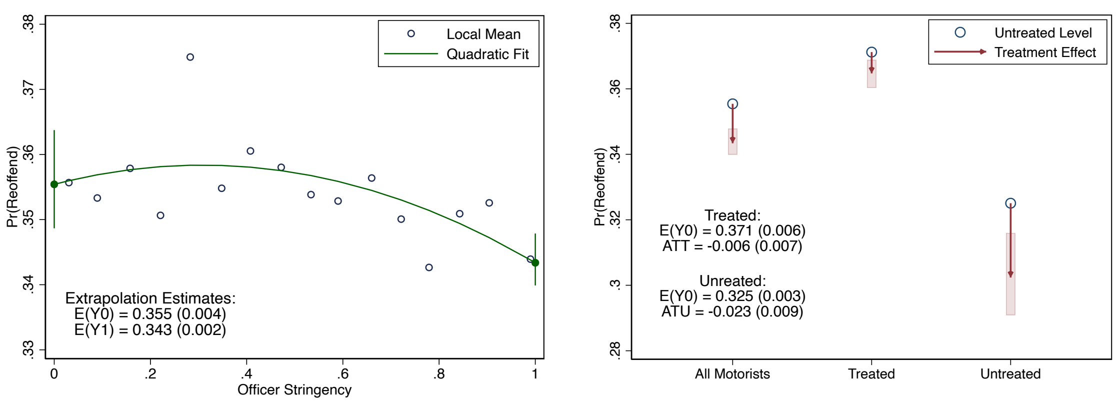
Table 2: Effect of Harsh Fines, LATE Estimates

| | (1) | IV Estimates (2) | (3) | (4) |
|---|---|---|---|---|
| | Lenient Mean | $\beta_{IV}$ | $\beta_{IV}$ | $\epsilon$ |
| Any Violation | 0.347 | -0.0177 | -0.0159 | -0.069 |
| | | (0.0017) | (0.0016) | (0.007) |
| Speeding Violation | 0.170 | -0.0146 | -0.0144 | -0.127 |
| | | (0.0013) | (0.0013) | (0.012) |
| Other Violation | 0.256 | -0.0119 | -0.0097 | -0.057 |
| | | (0.0016) | (0.0015) | (0.009) |
| Crash Involvement | 0.080 | -0.0029 | -0.0021 | -0.040 |
| | | (0.0010) | (0.0010) | (0.018) |
| Contest in Court | 0.262 | 0.1125 | 0.1093 | 0.626 |
| | | (0.0014) | (0.0014) | (0.008) |
| Controls | | No | Yes | Yes |
| Beat-Shift FE | | Yes | Yes | Yes |
| Observations | | 1693457 | 1693457 | 1693457 |

*Notes:* This table reports 2SLS estimates of the impact of receiving a harsh fine on one-year reoffending. Standard errors clustered at the officer -level are in parentheses. First stage estimates are $\beta = 0.944$ (0.006) without controls and $\beta = 0.943$ (0.006) with controls. See table B-2 for the full set of first stage and reduced form estimates with and without controls. Implied elasticities are computed as $\hat{\beta}_{IV} \times \bar{fine}/\bar{y}$, where $\hat{\beta}_{IV}$ is estimated using the statutory fine as the treatment variable and the means are the lenient officer means.

Figure 4: Extrapolation and treatment effect estimates

(a) Extrapolation Estimates (b) Treatment Effects




*Notes*: Panel (a) plots the relationship between officer stringency and driver reoffending, adjusted for beat-shift effects. Circles illustrate a binscatter, adjusted for beat-shift effects using the method of Cattaneo et al. (2021) and solid line denotes the quadratic fit. Solid circles indicate the implied estimates of the sample average potential outcomes, $E(Y_{i0})$ and $E(Y_{i1})$, with associated 95 confidence intervals. Panel (b) illustrates estimated of the untreated reoffending levels, $E(Y_{i0})$, and treatment effects, $E(Y_{i1} - Y_{i0})$, for all motorists, harshly treaed motorists and untreated (leniently treated) motorists, constructed from the extrapolation estimates in panel (a).

Table 3: Robustness of extrapolation estimates

| | $Y_0$ | | | | $Y_1 - Y_0$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) All | (2) $D=1$ | (3) $D=0$ | (4) Diff | (5) All (ATE) | (6) $D=1$ (ATT) | (7) $D=0$ (ATU) | (8) Diff |
| *Panel A: Polynomials* | | | | | | | | |
| $q=2$ | 0.355 | 0.371 | 0.325 | 0.046 | -0.012 | -0.006 | -0.023 | 0.016 |
| | (0.004) | (0.006) | (0.003) | (0.007) | (0.005) | (0.007) | (0.009) | (0.012) |
| $q=3$ | 0.350 | 0.363 | 0.325 | 0.038 | -0.006 | 0.001 | -0.019 | 0.021 |
| | (0.006) | (0.009) | (0.003) | (0.009) | (0.008) | (0.009) | (0.010) | (0.012) |
| $q=4$ | 0.346 | 0.357 | 0.325 | 0.032 | -0.002 | 0.008 | -0.021 | 0.029 |
| | (0.007) | (0.011) | (0.003) | (0.011) | (0.008) | (0.012) | (0.010) | (0.016) |
| $q=8$ | 0.351 | 0.364 | 0.325 | 0.039 | -0.008 | 0.001 | -0.024 | 0.025 |
| | (0.013) | (0.019) | (0.003) | (0.019) | (0.014) | (0.019) | (0.012) | (0.023) |
| *Panel B: Local means* | | | | | | | | |
| $bw=0.1$ | 0.355 | 0.370 | 0.325 | 0.045 | -0.009 | -0.006 | -0.015 | 0.009 |
| | (0.003) | (0.005) | (0.003) | (0.005) | (0.004) | (0.006) | (0.008) | (0.010) |
| $bw=0.05$ | 0.355 | 0.371 | 0.325 | 0.046 | -0.011 | -0.006 | -0.019 | 0.013 |
| | (0.005) | (0.007) | (0.003) | (0.007) | (0.006) | (0.008) | (0.009) | (0.012) |
| $bw=0.01$ | 0.353 | 0.367 | 0.325 | 0.042 | -0.010 | -0.003 | -0.024 | 0.021 |
| | (0.006) | (0.009) | (0.003) | (0.009) | (0.007) | (0.010) | (0.009) | (0.013) |
| *Panel C: Polyomials in propensity score* | | | | | | | | |
| $q=2$ | 0.355 | 0.370 | 0.325 | 0.045 | -0.013 | -0.005 | -0.028 | 0.022 |
| | (0.004) | (0.006) | (0.003) | (0.006) | (0.006) | (0.007) | (0.009) | (0.011) |
| *Panel D: Within-locations* | | | | | | | | |
| Troops | 0.358 | 0.375 | 0.330 | 0.045 | -0.011 | -0.010 | -0.022 | 0.012 |
| | (0.006) | (0.008) | (0.003) | (0.009) | (0.006) | (0.009) | (0.010) | (0.014) |
| Counties | 0.366 | 0.382 | 0.333 | 0.049 | -0.019 | -0.016 | -0.028 | 0.012 |
| | (0.009) | (0.010) | (0.003) | (0.010) | (0.009) | (0.011) | (0.024) | (0.028) |
| FM (2022) | 0.355 | 0.371 | 0.326 | 0.045 | -0.012 | -0.004 | -0.026 | 0.021 |
| | (0.006) | (0.007) | (0.003) | (0.007) | (0.006) | (0.007) | (0.008) | (0.009) |

*Notes:* This table reports estimated untreated reoffending levels $E(Y_{i0})$ and treatment effects $E(Y_{i1} - Y_{i0})$ for all motorists, treated motorists, and untreated motorists, using different specifications for obtaining extrapolated estimates of $E(Y_{i0})$ and $E(Y_{i1})$. In panel (a), we show estimates from estimating equation 1 with different polynomials in $Z$. In panel (b), we estimate equation $Z$ using a piecewise linear specification with rule-of-thumb bandwidth = 0.2. In panel (c), we show estimates that use the beat-shift adjusted average $Y$ for officers in the tails of the stringency distribution, with $bw$ denoting the bandwidth used to define the tails. In panel (d), we present estimates from polynomial specifications of 1 that replace $Z$ with the estimated propensity score. Graphical analogues for the underlying extrapolation estimates are shown in figure C-2.

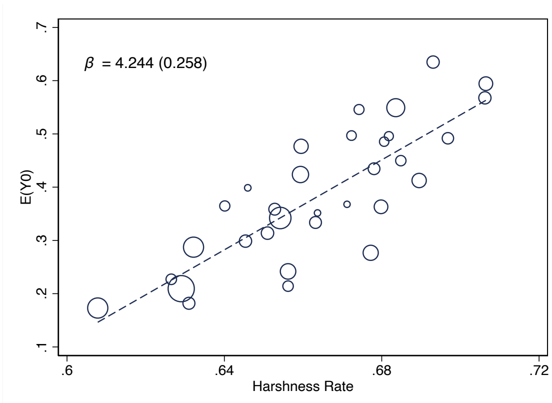
Table 4: Extrapolation estimates for other outcomes

| | $Y_0$ | | | | $Y_1 - Y_0$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) All | (2) $D=1$ | (3) $D=0$ | (4) Diff | (5) All (ATE) | (6) $D=1$ (ATT) | (7) $D=0$ (ATU) | (8) Diff |
| *Panel A: Any offense in following year* | | | | | | | | |
| Any Offense | 0.355 | 0.371 | 0.325 | 0.046 | -0.012 | -0.006 | -0.023 | 0.016 |
| | (0.004) | (0.006) | (0.003) | (0.007) | (0.005) | (0.007) | (0.009) | (0.012) |
| Speeding | 0.179 | 0.188 | 0.161 | 0.028 | -0.013 | -0.011 | -0.015 | 0.004 |
| | (0.003) | (0.004) | (0.002) | (0.004) | (0.004) | (0.005) | (0.005) | (0.007) |
| Crash | 0.080 | 0.081 | 0.077 | 0.004 | -0.001 | 0.001 | -0.005 | 0.006 |
| | (0.001) | (0.002) | (0.001) | (0.002) | (0.002) | (0.002) | (0.003) | (0.004) |
| *Panel B: Number of offenses in following two years* | | | | | | | | |
| Any Offense | 1.278 | 1.361 | 1.118 | 0.243 | -0.035 | -0.005 | -0.095 | 0.090 |
| | (0.025) | (0.036) | (0.016) | (0.038) | (0.027) | (0.039) | (0.046) | (0.066) |
| Speeding | 0.399 | 0.421 | 0.355 | 0.067 | -0.025 | -0.020 | -0.035 | 0.015 |
| | (0.007) | (0.011) | (0.004) | (0.011) | (0.008) | (0.011) | (0.013) | (0.018) |
| Crash | 0.157 | 0.160 | 0.151 | 0.009 | 0.001 | 0.004 | -0.005 | 0.009 |
| | (0.002) | (0.004) | (0.002) | (0.004) | (0.003) | (0.004) | (0.005) | (0.007) |

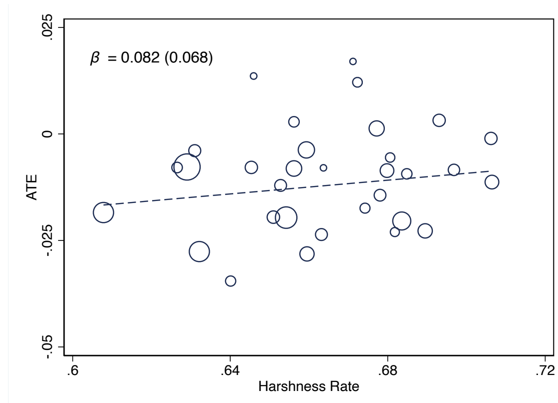
*Notes:* This table reports estimated untreated reoffending levels $E(Y_{i0})$ and treatment effects $E(Y_{i1} - Y_{i0})$ for all motorists, treated motorists, and untreated motorists, based on quadratic extrapolations using equation 1. In panel (a), we show estimates for binary outcomes indicating whether a motorist commits any new traffic offense, a new speeding offense, and is involved in an auto accident in the following year (the first row of panel (a) is identical to the first row of table 3). In panel (b), we replace the outcomes with counts of the number of offenses over the following two years. See appendix figure C-5 for graphical depictions of the underlying extrapolations and appendix table C-1 for estimates which alternatively use a local mean estimator for $E(Y_{i0})$ and $E(Y_{i1})$.

Figure 5: Reoffending levels and treatment effects by motorist characteristics

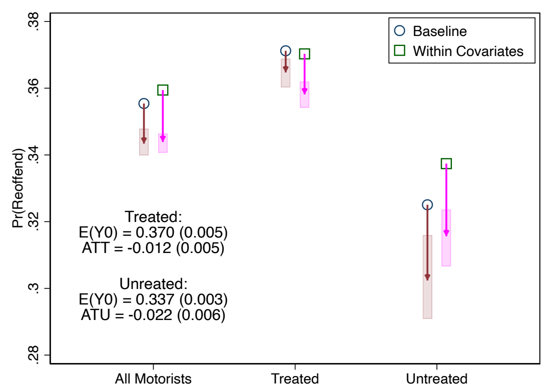
(a) $E(Y_{i0}|X)$




(b) $E(Y_{i1} - Y_{i0}|X)$



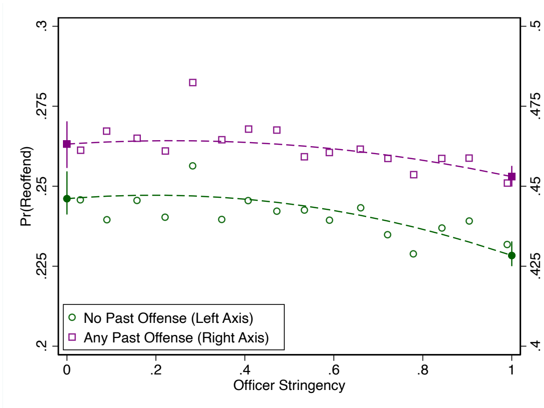

(c) Levels and gains within covariates



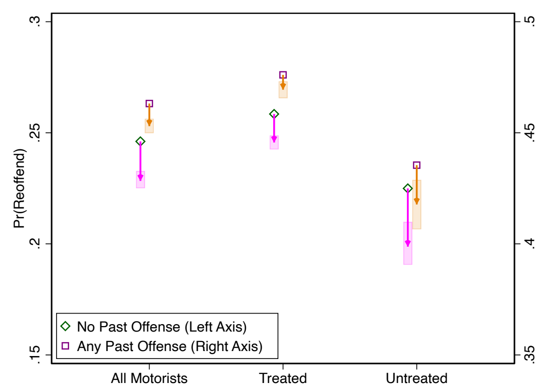

*Notes:* This figure explores selection patterns between and within motorist covariates. We divide motorists into 32 covariate cells at the level of gender × race × **1**[age≥ 35] × **1**[past offense] and estimate a quadratic extrapolation within each cell. Panel (a) plots the estimated $E(Y_{i0})$ in each cell against the cell-level harshness rate (the share of motorists in that cell that face harsh fines, adjusted for beat-shift fixed effects). Panel (b) plots the estimated cell-level treatment effect, $E(Y_{i1} - Y_{i0})$ against the same cell-level harshness rate. Both panels report an estimated linear slope and associated bootstrapped standard error. Panel (c) replicates panel (b) of figure 4 and then additionally shows estimated reoffending levels and treatment effects which are estimated *within* covariate cells, obtained by estimating parameters via extrapolation for each group and then aggregating up, weighting each group by its sample share. See appendix figure C-3 for a graphical depiction of the within-group extrapolations underlying the treatment effect estimates in panel (c).

Figure 6: Extrapolation estimates by offending history

(a) Extrapolation estimates




(b) Treatment effect estimates




*Notes:* Same as figure 4 except estimates are shown separately for motorists with (51.3 percent) and without (48.7 percent) at least one traffic citation in the past two years. As described in the text, we use a two-year lookback period because two years of pre-stop data are available for all citations in our main sample. Note that the estimates for motorists with and without past offenses are shown using different axes for ease of exposition (because reoffending rates are much higher for those with past offenses). The estimates illustrated in panel (b) are also reported in appendix table C-2, and alternative versions of these estimates using a local mean estimator for $E(Y_{i1})$ and $E(Y_{i0})$ are presented in appendix table C-3.

# FOR ONLINE PUBLICATION: APPENDICES

## A Stringency Instrument

### A-1 Evidence that bunching represents officer behavior

Here we discuss additional evidence that the bunching in the distribution of charge speeds presented in figure 1 is the result of discretionary behavior by officers. Figure 1 shows that about 25 percent of officers almost never write tickets for exactly nine MPH over the limit. Moving beyond a binary split of officers, figure A-1 illustrates significant variation across officers in the propensity to bunch drivers. Panel (a) demonstrates full support across officers in bunching propensity, while panel (b) shows that this variation persists after netting out location and time fixed effects. Such variation is inconsistent with bunching due to driver behavior; if drivers systematically bunch below fine increases, then officers patrolling the same beat-shift should exhibit a similar degree of bunching.

However, this across-officer variation could alternatively be due to noise or estimation error. To confirm that the across-officer variation in bunching propensity is "true" variation (in a statistical sense), we estimate the following regression:

$$\mathbf{1}[bunch_{ijs}] = \gamma X_i + \psi_s + \alpha_j + u_{ijs}$$

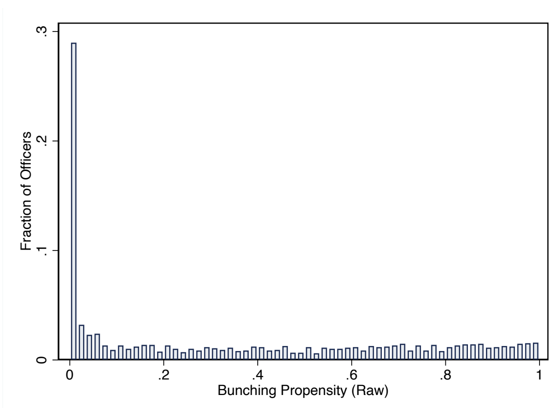
where $i$ indexes citations, $j$ indexes officers, and $s$ indexes beat-shifts; $X_i$ is a vector of driver covariates, $\psi_s$ is a beat-shift fixed effect, and $\alpha_j$ is an officer fixed effect.[13] This regression has an $R^2 = 0.55$, with 0.32 (58 percent) attributable to the officer effects, 0.22 (41 percent) attributable to the beat-shift effects, and less than one percent attributable to the driver $X$'s. In other words, the identity of the citing officer is significantly more predictive of a bunched citation than the beat-shift of the stop or the full set of driver characteristics. Moreover, there is significant variation in the estimated $\hat{\alpha}_j$'s ($\sigma^2 = 0.076$). Applying Empirical Bayes shrinkage (Morris, 1983) to adjust for estimation error has minimal impact on the dispersion of the estimated officer effects ($\sigma^2 = 0.71$). See panel (c) of figure A-1 for further details.

Finally, we show in figure A-2 that an officer's bunching propensity is highly correlated across space and time. First, we randomly partition an officer's citations into two location (county) groups and regress an officer's bunching propensity, adjusted for beat-shift fixed effects, in one set of locations on the same officer's adjusted bunching propensity in the other set of locations. This regression yields $\hat{\beta} = 0.68$ ($se = 0.02$). Next, we split an officer's citations in half temporally and perform the same exercise, which gives $\hat{\beta} = 0.85$ ($se = 0.01$).

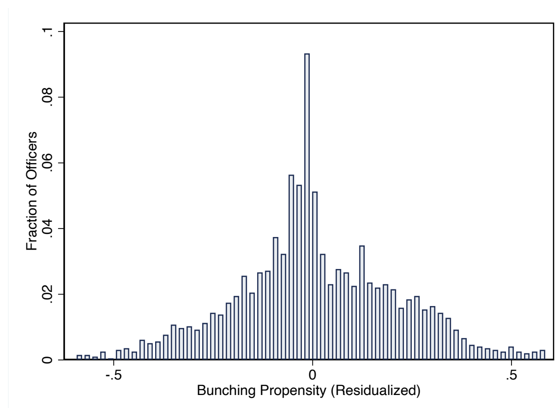
[13] The $\psi$'s are the same fixed effects we use in our main analysis, described in section 3. They are at the level of county × $\mathbf{1}$[highway] × year × month × $\mathbf{1}$[weekend] × shift.

Figure A-1: Across-Officer Distribution of Bunching Propensity

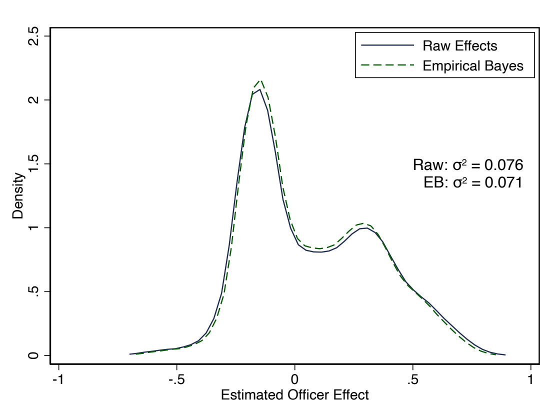
(a) Raw Bunching Propensity




(b) Adjusted Bunching Propensity



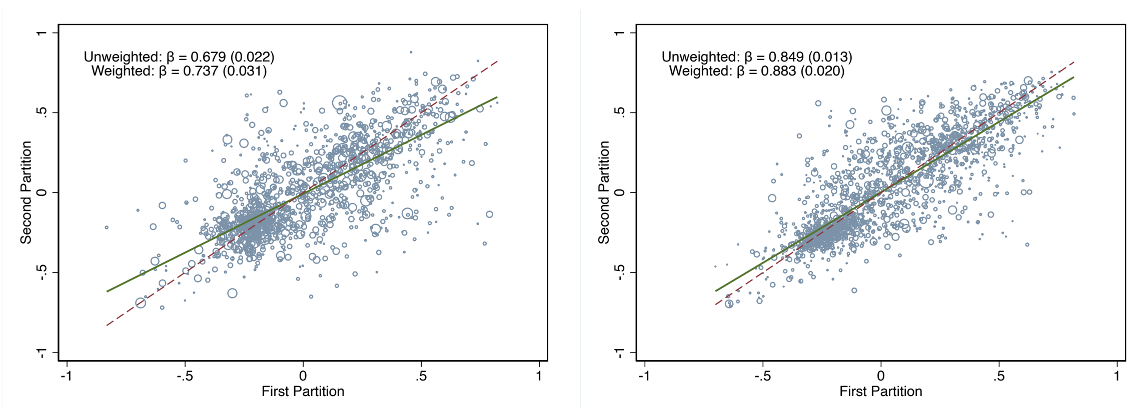

(c) Estimated Officer Effects




*Notes:* Panel (a) plots the officer-level distribution of the share of tickets bunched. Panel (b) plots the officer-level distribution of bunching propensity, residualized of beat-shift fixed effects. Panel (c) reports estimated officer effects from a regression of $\mathbf{1}[bunch_{ijs}]$ on officer fixed effects, beat-shift fixed effects, and the full set of driver covariates, as described in section 2.4. The solid blue illustrates the distribution of raw officer effects and the dashed green line illustrates the distribution of effects after applying Empirical Bayes shrinkage (Morris, 1983).

Figure A-2: Within-Officer Correlation in Bunching Propensity

(a) Locations (b) Time Periods




*Notes:* Each circle corresponds to an officer. Dashed red line is the 45-degree line. This figure splits each officer's sample of citations into two groups and illustrates the correlation in (residualized) bunching propensity across groups. In panel (a), the groups are constructed as location partitions, with each partition comprised of half of an officer's patrol locations. In panel (b), the groups are constructed as time partitions, with the $x$ and $y$-axes corresponding to the officer's first and second half of tickets over time, respectively. Each figure reports the raw linear regression coefficient as well as the linear regression coefficient when weighting by the total number of citations. Another way to note the stability over time in an officer's bunching propensity is to regress $\mathbf{1}[bunch_{ijs}]$ on beat-shift fixed effects, officer fixed effects, and a quadratic in officer experience (in months). The $p$-value on each experience term is $> 0.45$ and the joint test $p$-value $= 0.7855$. In other words, after conditioning on officer identity, there is no experience profile in the likelihood of a bunched ticket.

Table A-1: Relationship between Lenience and Officer Characteristics

| | | Binary | Continuous | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | Mean | Lenient | Raw | Adjusted | Weighted |
| Female | 0.0893 | -0.0704 | -0.0508 | -0.0356 | -0.00340 |
| | | (0.0366) | (0.0266) | (0.0161) | (0.00388) |
| Race = Black | 0.143 | -0.0916 | -0.0289 | -0.00799 | 0.00550 |
| | | (0.0297) | (0.0231) | (0.0142) | (0.00291) |
| Race = Hispanic | 0.169 | -0.0933 | -0.0771 | 0.0117 | 0.00921 |
| | | (0.0693) | (0.0564) | (0.0359) | (0.00707) |
| Race = Other | 0.191 | -0.0120 | -0.0193 | -0.0186 | -0.00135 |
| | | (0.0655) | (0.0544) | (0.0350) | (0.00724) |
| Age | 34.06 | -0.0203 | -0.0236 | 0.00982 | 0.00741 |
| | | (0.0561) | (0.0478) | (0.0288) | (0.00631) |
| Experience | 7.09 | -0.117 | -0.0778 | -0.0345 | -0.000936 |
| | | (0.0388) | (0.0324) | (0.0214) | (0.00616) |
| Any College | 0.319 | -0.00798 | -0.0114 | -0.00229 | 0.00490 |
| | | (0.0213) | (0.0171) | (0.0108) | (0.00279) |
| Mean | — | 0.753 | 0.353 | 0.005 | 0.006 |
| Officers | 1960 | 1960 | 1960 | 1960 | 1958 |

*Notes:* Robust standard errors in parentheses. Age and experience are in years/10 and are computed as of January 2007. Raw lenience is the fraction of an officer's tickets that are bunched and adjusted lenience is the fraction of an officer's tickets that are bunched, residualized of location-time fixed effects. In column 4, the regression is weighted by one over the variance of adjusted lenience. Regressions also included quadratic terms in age and experience, which are are statistically insignificant in all cases.

Table A-2: Randomization Test

| | (1) Reoffend | (2) Harsh Fine | (3) Stringency | (4) **1**[Stringent] |
|---|---|---|---|---|
| Female | -0.0645 | -0.0233 | -0.00119 | -0.00161 |
| | (0.000840) | (0.00173) | (0.000928) | (0.00130) |
| Age | -0.00631 | -0.00229 | 0.000280 | 0.000540 |
| | (0.000151) | (0.000273) | (0.000172) | (0.000257) |
| Age Squared | 0.0000190 | 0.0000151 | -0.00000303 | -0.00000586 |
| | (0.00000168) | (0.00000268) | (0.00000166) | (0.00000259) |
| Race = Black | 0.0710 | 0.0202 | -0.000326 | -0.00407 |
| | (0.00127) | (0.00270) | (0.00176) | (0.00242) |
| Race = Hispanic | 0.0324 | 0.0345 | 0.00633 | -0.000643 |
| | (0.00117) | (0.00290) | (0.00212) | (0.00295) |
| Race = Other | 0.000888 | 0.0348 | 0.00581 | -0.00101 |
| | (0.00202) | (0.00274) | (0.00176) | (0.00250) |
| Race = Unknown | 0.0120 | 0.00505 | 0.00377 | -0.00131 |
| | (0.00299) | (0.00550) | (0.00266) | (0.00374) |
| Prior Prison Spell | 0.139 | 0.0107 | -0.00315 | -0.000259 |
| | (0.00393) | (0.00359) | (0.00207) | (0.00275) |
| County Resident | 0.00583 | -0.0186 | -0.00586 | 0.00310 |
| | (0.00127) | (0.00309) | (0.00265) | (0.00441) |
| Log Zip Income | -0.0161 | 0.0103 | 0.00486 | 0.00180 |
| | (0.000992) | (0.00211) | (0.00164) | (0.00194) |
| Log Vehicle Price | -0.00846 | 0.0199 | 0.00487 | 0.00159 |
| | (0.000992) | (0.00160) | (0.00122) | (0.00191) |
| Speeding Past Year | 0.127 | 0.0247 | 0.000948 | -0.000595 |
| | (0.00103) | (0.00157) | (0.000653) | (0.00110) |
| Other Past Year | 0.154 | 0.0149 | -0.000758 | -0.00201 |
| | (0.00103) | (0.00115) | (0.000727) | (0.00118) |
| Crash Past Year | 0.0393 | 0.00736 | 0.000990 | -0.000721 |
| | (0.00148) | (0.00132) | (0.000769) | (0.00113) |
| Mean | .351 | .658 | .658 | .763 |
| F-Stat | 5883.84 | 28.8 | 2.7 | .92 |
| F-test | <.0001 | <.0001 | .0002 | .5484 |
| Beat-Shift FE | Yes | Yes | Yes | Yes |
| Officers | 1960 | 1960 | 1960 | 1960 |
| Observations | 1693457 | 1693457 | 1693457 | 1693457 |

*Notes:* Standard errors clustered at the officer-level in parentheses. Regressions also include indicators for missing age (<1%), missing zip code income ($\approx$ 10%), and missing vehicle information ($\approx$ 14%); joint significance tests include these variables.

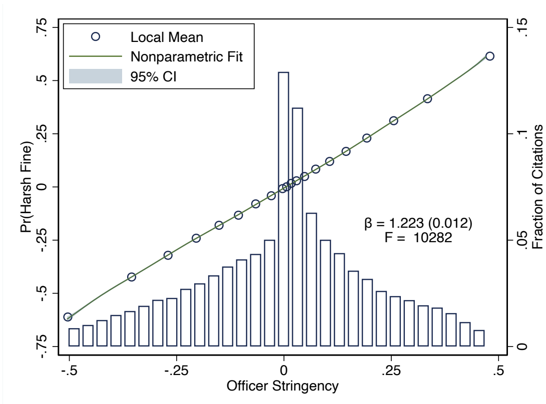
Table A-3: First Stage Estimates Across Subsamples

| | Subgroup | |
|---|---|---|
| | (1) | (2) |
| | = 1 | = 0 |
| Female | 0.970 | 0.928 |
| | (0.007) | (0.007) |
| Age > 30 | 0.957 | 0.927 |
| | (0.007) | (0.007) |
| Race = White | 0.954 | 0.934 |
| | (0.008) | (0.007) |
| Race = Black | 0.922 | 0.948 |
| | (0.010) | (0.006) |
| Race = Hispanic | 0.923 | 0.948 |
| | (0.008) | (0.007) |
| Race = Other | 0.916 | 0.945 |
| | (0.015) | (0.006) |
| Race = Unknown | 0.964 | 0.941 |
| | (0.016) | (0.007) |
| County Resident | 0.972 | 0.925 |
| | (0.007) | (0.007) |
| Zip Income > $50,000 | 0.946 | 0.942 |
| | (0.007) | (0.007) |
| Vehicle > $20,000 | 0.916 | 0.953 |
| | (0.009) | (0.006) |
| Citation Past Year | 0.913 | 0.961 |
| | (0.007) | (0.007) |

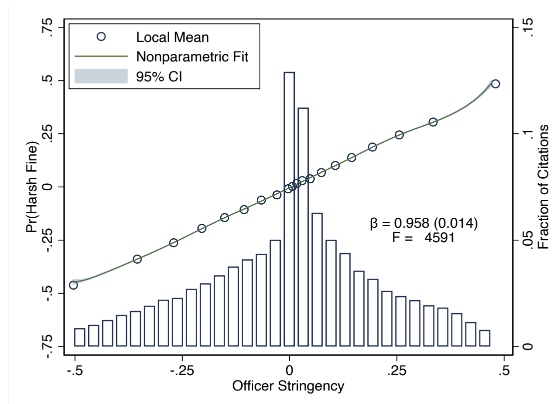
*Notes:* This table reports first stage estimates for subsamples. Each coefficient is from a separate regression of $\mathbf{1}[harsh]$ on the stringency instrument and beat-shift fixed effects using only the denoted subgroup of drivers, where the subgroups are the groups for which the denoted indicator variable = 1 (column 1) and = 0 (column 2). Standard errors clustered at the officer-level in parentheses. For reference, the first stage estimate in the full sample is $\beta = 0.944$ (0.006).

Figure A-3: First Stage Estimates, Sanction Measures

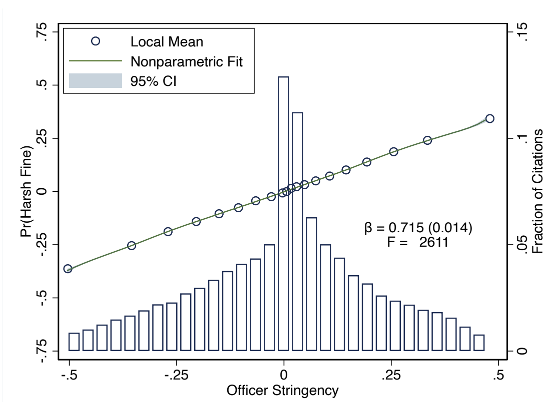
(a) Statutory Fine Amount



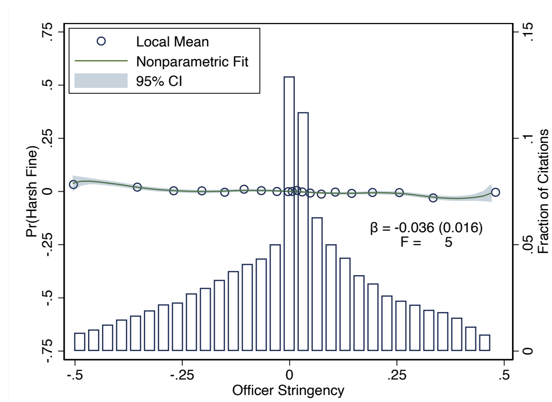

(b) Paid Fine Amount




(c) Statutory DL Points




(d) Accrued DL Points




*Notes:* Each panel shows an identical to panel (c) of figure 2 but replaces the outcome variable with a different sanctions measure. In panel (a), the outcome is the statutory fine based on the charged speed. In panel (b), the outcome is the effective fine amount, taking into account the *ex-post* court outcomes of offenders. In panel (c), the outcome is statutory driver license points based on the points schedule. In panel (d), the outcome is accrued DL points, taking into account the *ex-post* court outcomes of offenders. See appendix section D-1 for details on the computation of effective sanction measures (paid fines and accrued points).

# B Deterrence effects

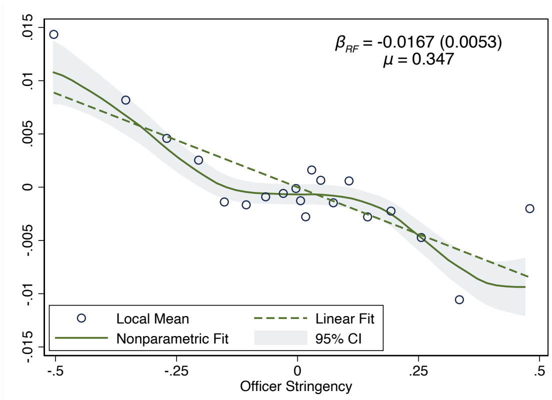
Table B-1: Naive OLS Estimates

| | (1) Reoffend | (2) Reoffend | (3) Reoffend | (4) Reoffend |
|---|---|---|---|---|
| Fine ($100s) | 0.0426 | 0.0548 | 0.0228 | 0.0286 |
| | (0.00241) | (0.00178) | (0.00160) | (0.00127) |
| Mean | 0.325 | 0.325 | 0.325 | 0.325 |
| Controls | No | No | Yes | Yes |
| Officer FE | No | Yes | No | Yes |
| Beat-Shift FE | Yes | Yes | Yes | Yes |
| Observations | 1693457 | 1693457 | 1693457 | 1693457 |

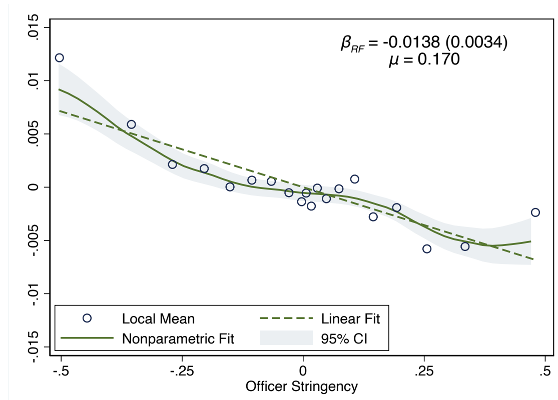
*Notes:* Standard errors clustered at the officer-level in parentheses. Dependent variable is an indicator for a new traffic offense in the next year. The reported mean is the mean for drivers cited at 9 MPH over the limit.

Figure B-1: Reduced Form Estimates

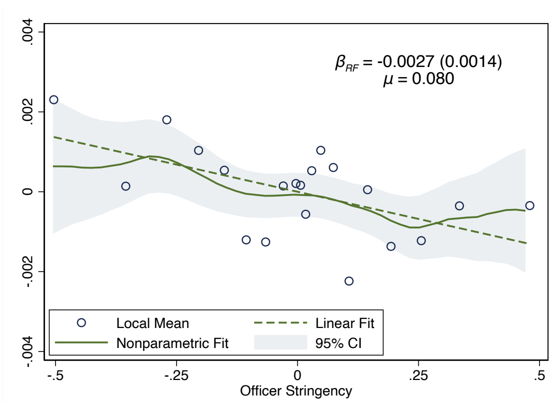
(a) Any Violation



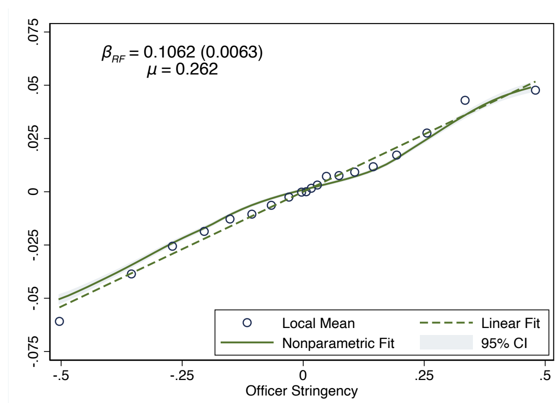

(b) Speeding Violation




(c) Crash Involvement

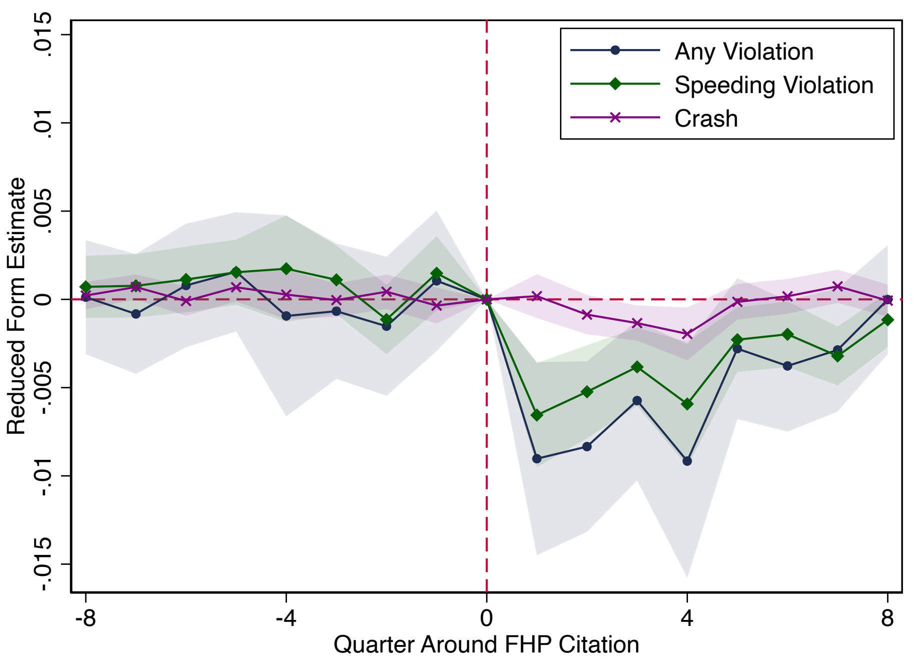



(d) Contested Citation




*Notes:* Same as figure A-3 except for reduced form outcomes.

Figure B-2: Dynamic Reduced Form Estimates




*Notes:* Same as figure 3 using any moving violation in a given quarter (blue circles), any speeding violation in a given quarter (green diamonds), and any crash involvement in a given quarter (purple $x$'s) as the outcome variable. Shaded regions denote 95 percent confidence intervals obtained from standard errors clustered at the officer-level.

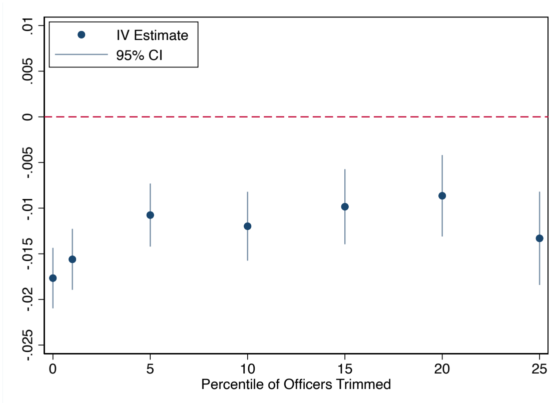
Table B-2: First Stage and Reduced Form Estimates

| | (1)<br>Lenient Mean | (2)<br>$\beta$ | (3)<br>$\beta$ |
|---|---|---|---|
| *Panel A: First Stage* | | | |
| Harsh Fine | 0.5573 | 0.9441 | 0.9432 |
| | | (0.0064) | (0.0064) |
| Fine Amount | 194.308 | 122.340 | 122.210 |
| | | (1.206) | (1.190) |
| Fine Amount (Paid) | 167.187 | 95.819 | 96.066 |
| | | (1.414) | (1.406) |
| DL Points | 3.416 | 0.7152 | 0.7142 |
| | | (0.0140) | (0.0138) |
| DL Points (Accrued) | 1.684 | -0.0362 | -0.0273 |
| | | (0.0164) | (0.0154) |
| *Panel B: Reduced Form* | | | |
| Any Violation | 0.3471 | -0.0167 | -0.0150 |
| | | (0.0053) | (0.0034) |
| Speeding Violation | 0.1702 | -0.0138 | -0.0136 |
| | | (0.0034) | (0.0025) |
| Other Violation | 0.2563 | -0.0112 | -0.0092 |
| | | (0.0045) | (0.0028) |
| Moving Violation | 0.2801 | -0.0135 | -0.0128 |
| | | (0.0047) | (0.0031) |
| Non-Moving Violation | 0.1602 | -0.0117 | -0.0097 |
| | | (0.0035) | (0.0022) |
| Crash Involvement | 0.0799 | -0.0027 | -0.0020 |
| | | (0.0014) | (0.0011) |
| Contest in Court | 0.2620 | 0.1062 | 0.1031 |
| | | (0.0063) | (0.0060) |
| Controls | | No | Yes |
| Beat-Shift FE | | Yes | Yes |
| Officers | | 1960 | 1960 |
| Observations | | 1693457 | 1693457 |

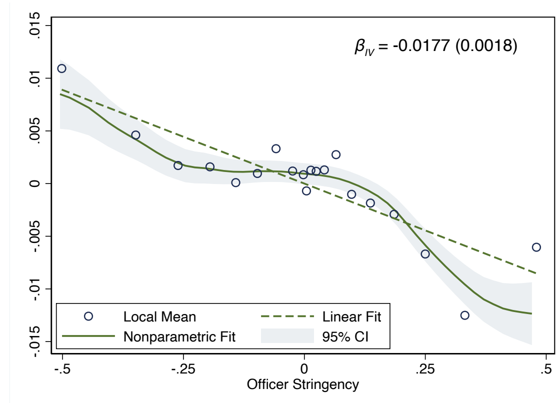
*Notes:* Standard errors clustered at the officer-level in parentheses. This table reports first stage and reduced form regression estimates with and without covariates. Each coefficient is from a separate regression of the denoted outcome on the stringency instrument and beat-shift effects, with (column 2) and without (column 3) controls

Figure B-3: Robustness, Sample Selection

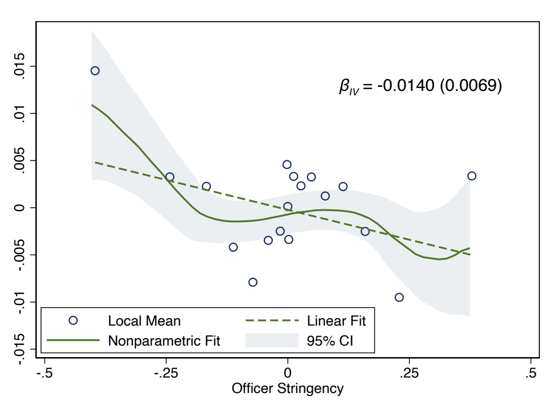
(a) Trimming Officers




(b) Selection Correction




(c) GPS FE



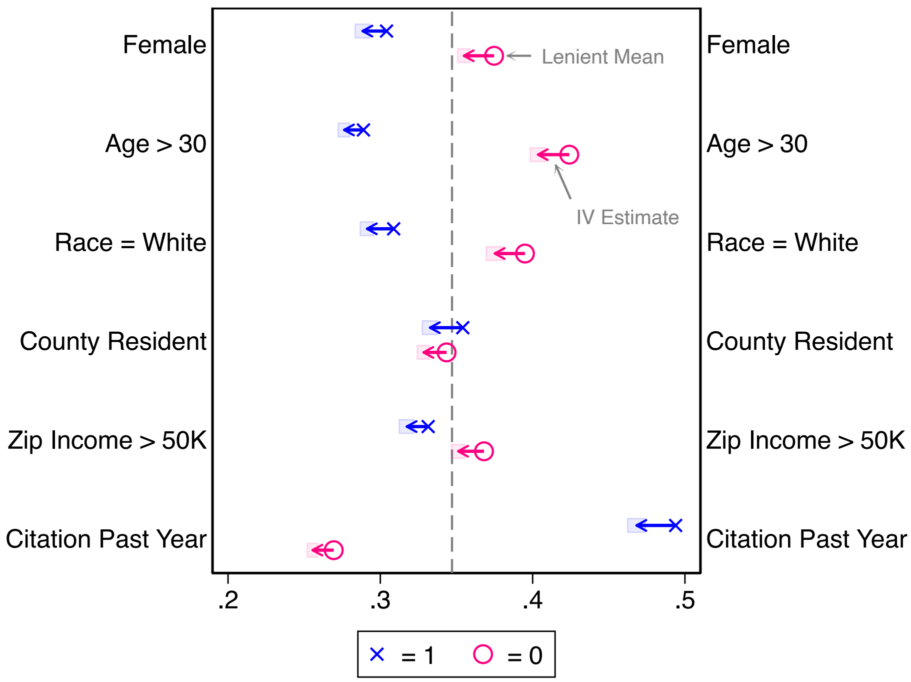

*Notes:* For comparison, our main IV estimate is $\beta_{IV} = -0.0177$ (0.0017). Panel (a) shows the sensitivity of our IV estimate to trimming officers with the most selected samples. Panel (b) plots the reduced form and reports the IV estimate using a Heckman (1979) selection correction based on officer ticketing frequency. Panel (c) plots the reduced form and reports the IV estimate using GPS road segment fixed effects for the subset of citations including GPS coordinates ($N = 244,858$).

Table B-3: Robustness, Alternative Instruments

| Instrument | (1) N | $F$-Stat (2) Balance | $F$-Stat (3) FS | (4) $\beta_{IV}$ |
|---|---|---|---|---|
| Leave-out | 1693457 | 2.702 | 21966 | -0.0159 |
| | | | | 0.0016 |
| Leave-Out (Residualized) | 1693457 | 2.691 | 25278 | -0.0191 |
| | | | | 0.0016 |
| Leave-County-Out | 1500479 | 1.475 | 626 | -0.0176 |
| | | | | 0.0024 |
| Binary | 1693457 | 0.921 | 289 | -0.0268 |
| | | | | 0.0039 |
| Officer Dummies | 1693404 | 2.778 | 323 | -0.0183 |
| | | | | 0.0016 |
| *Within Demographics* | | | | |
| Race | 1689414 | 5.111 | 24453 | -0.0168 |
| | | | | 0.0016 |
| Race × Gender × Age | 1651212 | 4.579 | 21505 | -0.0176 |
| | | | | 0.0016 |
| Race × Gender × Age × History | 1587971 | 5.175 | 18469 | -0.0182 |
| | | | | 0.0017 |
| Race × Gender × Age × History × Income | 1471826 | 3.884 | 14912 | -0.0189 |
| | | | | 0.0017 |

*Notes:* This table shows how results vary under different computations of the stringency instrument. Columns 2 and 3 report F-statistics associated with a joint balance test and the first stage; column 4 reports the IV estimate for one-year speeding recidivism. Row 1 reports results corresponding to the main instrument. In row 2, the instrument is the leave-out-mean after residualizing out beat-shift effects (e.g., Dobbie et al. 2018). Row 3 computes the instrument as the leave-county-out mean. Row 4 uses a binary instrument and row 5 uses the full set of officer dummies as instruments. Rows 6-9 show results when the instrument is computed as the leave-out mean within demographic cells, defined according to four race groups (white, Black, Hispanic, other/unknown), gender, $\mathbf{1}[\text{age} \geq 35]$, $\mathbf{1}[\text{any citation in past year}]$ and $\mathbf{1}[\text{zip code income} \geq \$50{,}000]$. Regressions using the by-group instrument also include fixed effects at the relevant demographic cell-level.

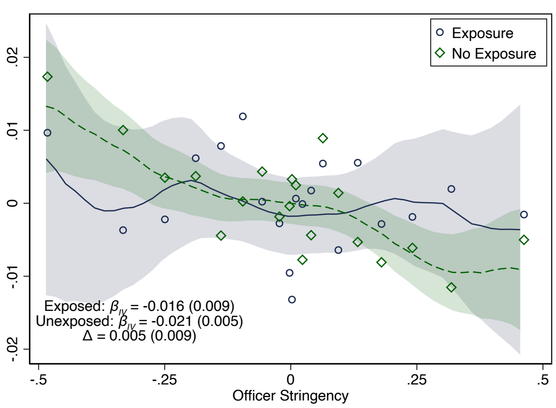
Figure B-4: IV Estimate Heterogeneity by Driver Characteristics



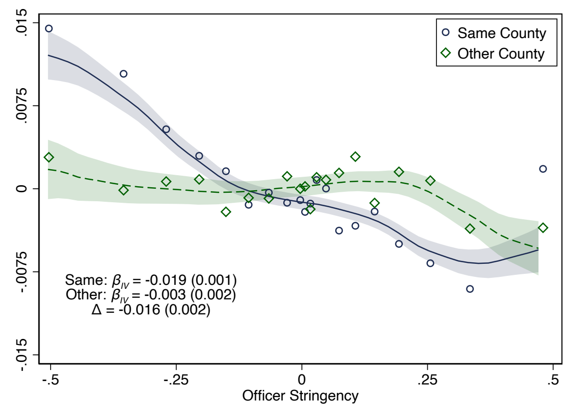

*Notes:* This figures shows heterogeneity in IV estimates for one-year recidivism by driver characteristics. Each characteristic is denoted as a binary category; the $x$'s plot lenient means for the category = 1 subgroup and the $o$'s plot lenient means for the category = 0 subgroup. Arrows pointing away from the means indicate the IV estimate, and shaded region around the arrow denotes the 95 percent confidence interval. Vertical dashed line denotes the lenient officer mean recidivism rate for the full sample.

Figure B-5: Evidence of Driver Learning

(a) Exposure to Stringent Officers




(b) Reoffense Locations




*Notes:* Panel (a) illustrates reduced form relationships and reports IV estimates for motorists with and without past exposure to stringent (non-bunching) officers. To mitigate selection issues resulting from the fact that past exposure to stringency reduces the likelihood a driver reappears in the data, we focus on exposure at least one year in the past because treatment effects fade out after one year (see figure 3) . Specifically, we take the subset of drivers with an FHP-issued citation at least one year prior and compare treatment effects for those with and without past tickets issued by stringent officers. In panel (b), we report treatment effects of harsh fines on the likelihood that drivers reoffend in the same county they were ticketed and in different counties from the one they were ticketed in. Both panels report the estimated difference (and associated standard error) in treatment effects.

## B-1 Additional robustness: monotonicity

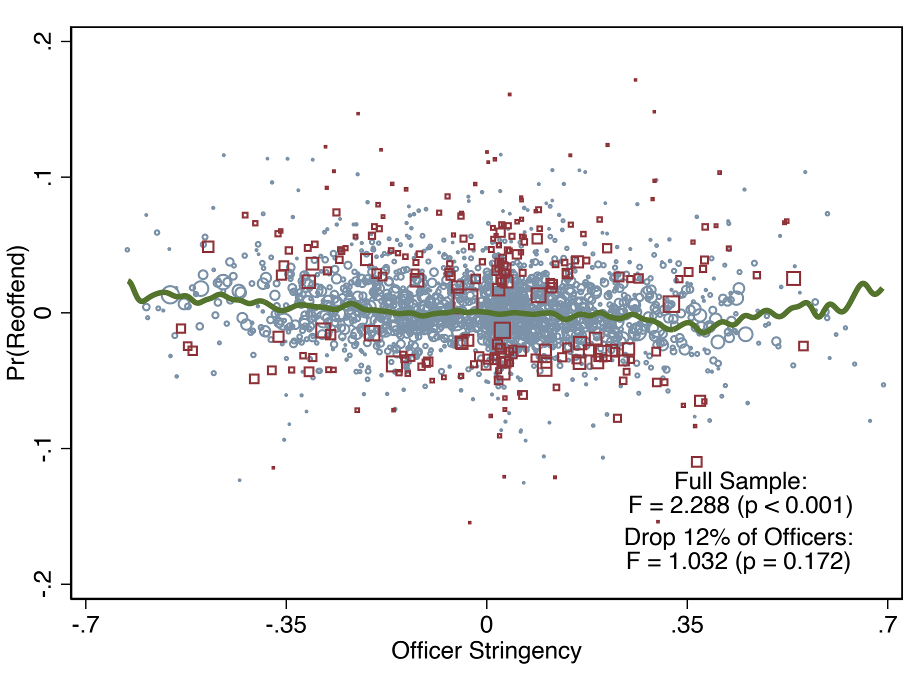
In addition to the result that average monotonicity is sufficient for 2SLS estimates in examiner designs, Frandsen et al. (2019) also provide a test for the joint assumptions of *strict* monotonicity and exclusion. We were unable to use their code due to computing constraints, but here we provide an ad-hoc version of their test. Specifically, our version replicates the "fit" component of their test by flexibly fitting reoffending rates to the stringency instrument (conditional on beat-shift effects), computing residuals, and then testing the ability of officer effects to explain the residuals. We focus only on the "fit" component of the test at the suggestion of the authors, who note that the "slope" component is unpowered with many judges (as there are in our setting, $N = 1,960$).

Because we do not account for estimation error in the construction of the residuals, our ad-hoc version of the test yields a $p$-value which is *biased towards zero.* In other words, we will over-reject the joint null hypothesis of monotonicity and exclusion. Hence, while qualitatively useful, our version of the test is somewhat challenging to interpret. We depict the results of this Frandsen et al. (2019) test graphically below. The joint test statistic, which summarizes the ability of officer fixed effects to explain the residuals, is $F = 2.3$. On the one hand, this test statistic is reassuringly small; on the other hand, the (biased) $p$-value suggests that deviations from exclusion and strict monotonicity could be a salient concern for our empirical results.

The above figure also notes that the outcome of this test hinges on a relatively small share of officers. Specifically, if we drop the top 12 percent of officers in terms of their associated $t$-statistics in the second stage regression of the residuals on the officer fixed effects, the joint $F$-statistic falls to $F = 1.03$, with a biased-downward $p = 0.172$. To that end, in table B-4 below, we replicate our main specific deterrence estimate using only the subsample of officers passing the Frandsen et al. (2019) test, finding similar 2SLS estimates.

In table B-4, we also show 2SLS estimates which recompute officer stringency within 32 motorist covariate cells in column 2 (same as row 8 of table B-3), which estimate effects separately for 16 officer covariate cells (gender $\times$ $\mathbf{1}$[white] $\times$ $\mathbf{1}$[any college] $\times$ $\mathbf{1}$[experienced]) and then aggregates up, weighting by sample shares in column 3, and a combined version of columns 2 and 3 in column 4. The specification in column 4 only requires monotonicity to hold within each combination of motorist and officer characteristics.

Figure B-6: Fit Test from Frandsen et al. (2019)




*Notes:* This figure illustrates results of the joint test of monotonicity and exclusion from Frandsen et al. (2019). Each circle (or square) represents an officer ($N = 1,960$) and plots the officer's average stringency and reoffending rates, residualized of covariates and beat-shift fixed effects. Green line denotes a non-parametric fit and the figure reports the results of computing residuals from the non-parametric fit, regressing the residuals on officer dummies, and performing a joint significance test.

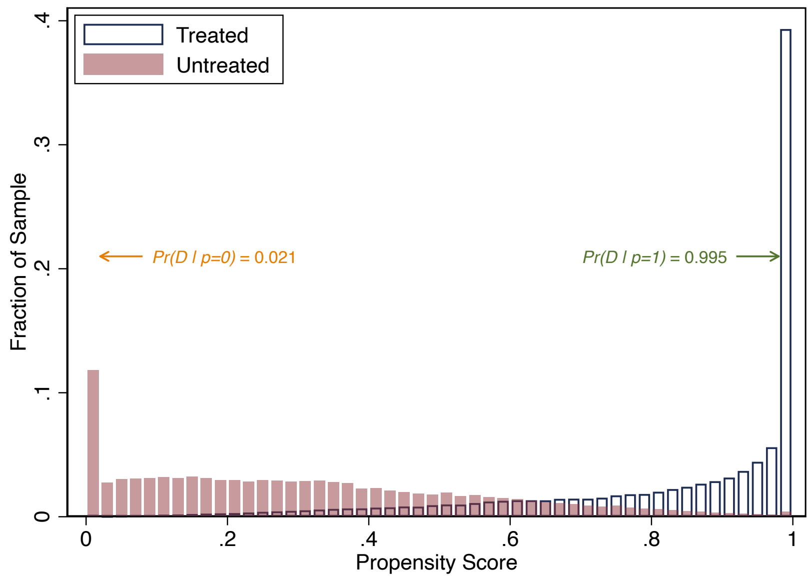
Table B-4: Monotonicity robustness for IV estimates

| | (1) Baseline | (2) Driver Cells | (3) Officer Cells | (4) Both |
|---|---|---|---|---|
| Full Sample | -0.0177 | -0.0183 | -0.0164 | -0.0179 |
| | (0.0056) | (0.0038) | (0.0038) | (0.0037) |
| FLL Officers | -0.0152 | -0.0164 | -0.0152 | -0.0166 |
| | (0.0037) | (0.0026) | (0.0025) | (0.0025) |

*Notes:* This table reports 2SLS estimates. The outcome variable is any new traffic offense in the following year. In column 1, we use our baseline leave-out stringency measure. In column 2, we use a stringency measure which is computed separately for each of 32 motorist covariate cells. In column 3, we estimate a 2SLS coefficient separately for each of 16 officer cells and then compute the weighted average of the coefficients, weighting by sample shares. In column 4, we repeat the exercise in column 3 using the cell-specific stringency instrument from column 2. In the second row, we drop the 12 percent of officers indicated in table B-6 ($N$ officers $= 1,725$, $N = 1,314,635$). In columns 3 and 4, we compute standard errors using a Bayesian bootstrap, clustered at the officer-level.

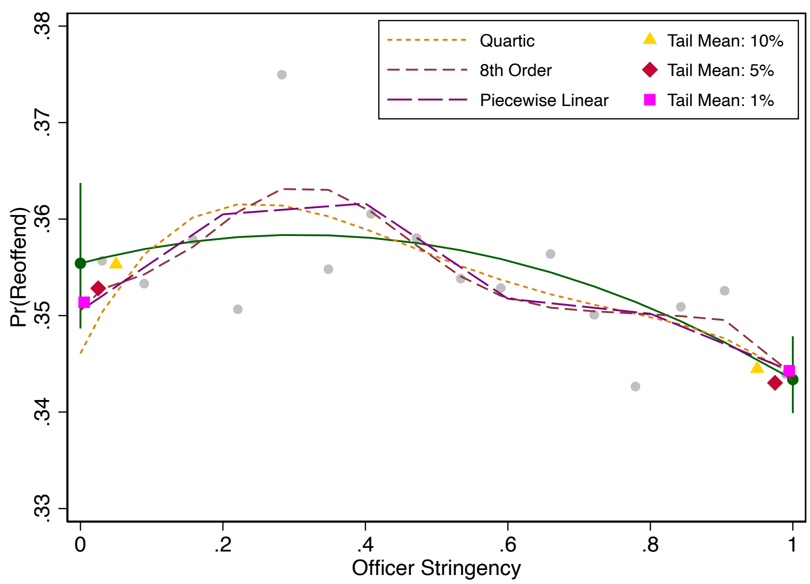
# C Extrapolation estimates

Figure C-1: Common Support of Officer Stringency Instrument




*Notes:* Figure plots the distribution of propensity scores for the treated (66%) and untreated (34%) subsets of sample, where treatment is defined as $\mathbf{1}[harsh]$. Following the text, the propensity score is estimated from a linear regression of $\mathbf{1}[harsh]$ on officer stringency and beat-shift fixed effects.

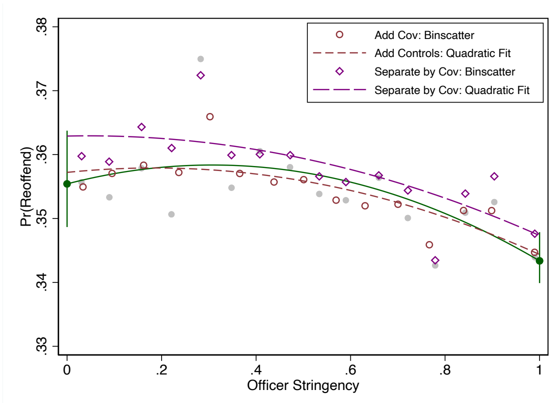
Figure C-2: Extrapolation under different functional forms



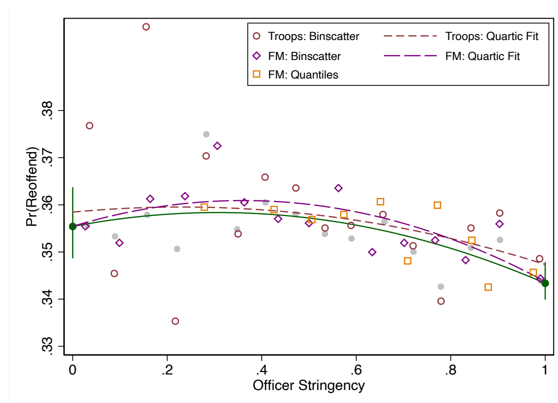

*Notes:* This figure illustrates extrapolation estimates of the sample $E(Y_{i0})$ and $E(Y_{i1})$ under different functional forms. Gray circles illustrates a binscatter and green solid line shows the quadratic fit (both are the same as figure 4). Orange dotted line presents a quartic fit, maroon dashed line shows an 8th order polynomial, and purple long-dashed line shows a piecewise linear fit based on a bin width of 0.2. Gold triangles, maroon diamons, and pink squares show the beat-shift adjusted averages for officers with $Z < 0.1$ and $Z > 0.9$, $Z < 0.05$ and $Z > 0.95$, and $Z < 0.01$ and $Z > 0.99$, respectively.

Figure C-3: Within-group extrapolation estimates

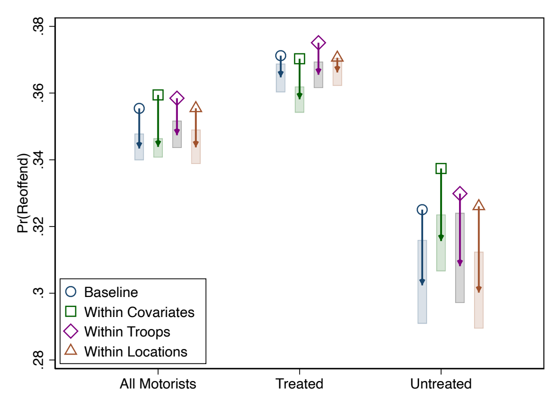
(a) By driver covariates




(b) By stop locations

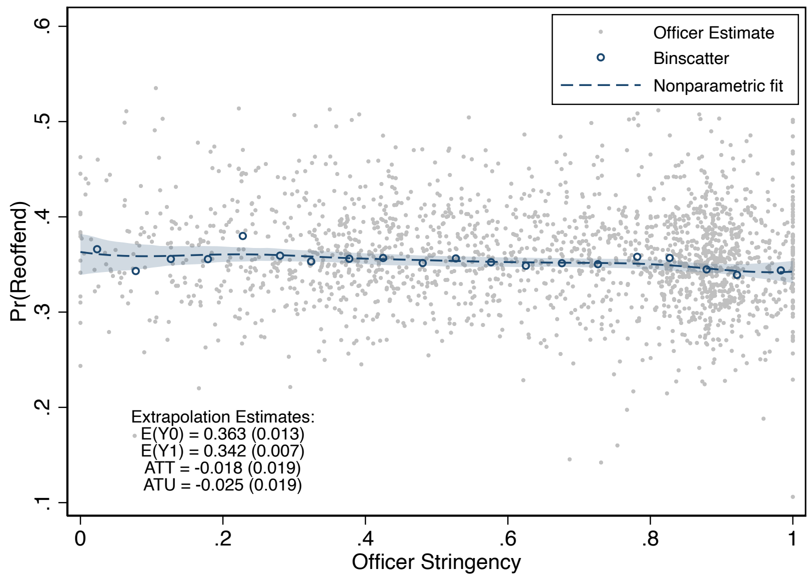



(c) Treatment effects



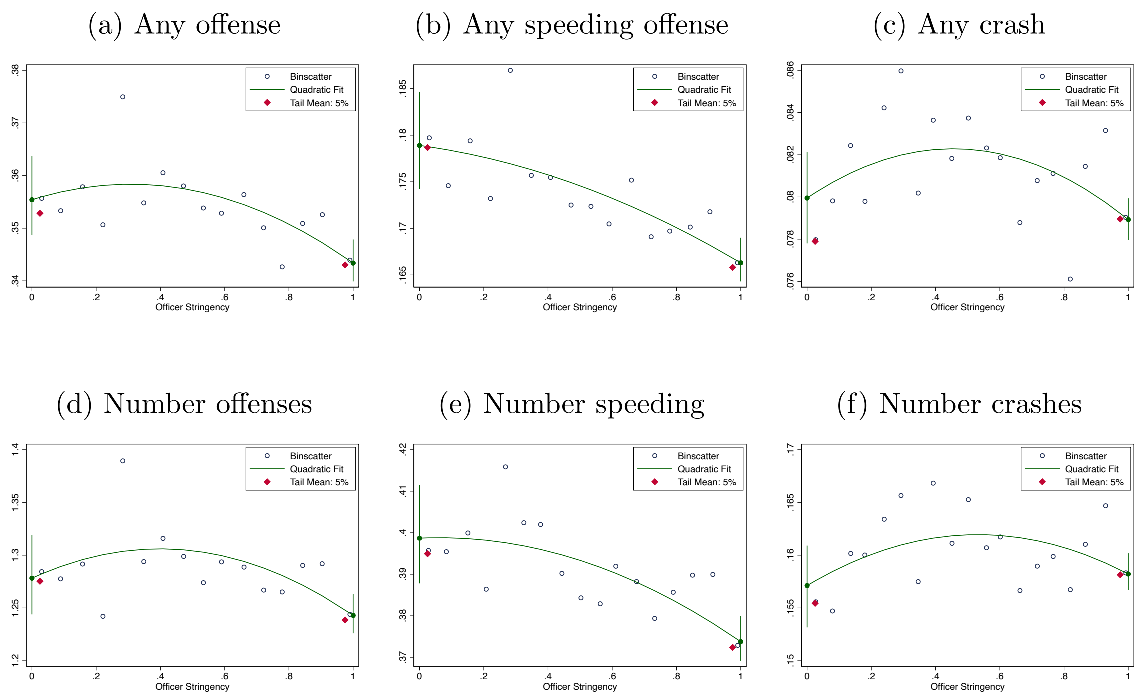

*Notes:* In both panels, gray dots and green solid line represent our baseline binscatter and baseline quadratic fit, as in figure 4. Panel (a) illustrates extrapolation estimates adjusted for driver covariates, parameterized as 32 categories (gender $\times$ race $\times$ $\mathbf{1}[\text{age} \geq 35]$ $\times$ $\mathbf{1}[\text{any citation in the past year}]$). Red circles and dashed line show the binscatter and quadratic fit when also conditioning on these cell fixed effects, while the purple diamonds and dashed line show the binscatter and quadratic fit when estimating separately for each driver group and then aggregating up, weighting by sample shares. Panel (b) illustrates extrapolation estimates constructed within-locations. Red circle and dashed line show the binscatter and quadratic fit when estimating separately by troop and then aggregating up, weighting by sample shares. Purple dots and dashed line show the binscatter and quadratic fit when using the within-locations approach of Feigenberg & Miller (2022), described in appendix E-1. Orange squares illustrate the binscatter corresponding to the within-location quantiles approach of Feigenberg & Miller (2022). Panel (c) replicates panel (b) of figure 4, illustrating treatment effects based on the within-group extrapolations depicted in panels (a) and (b).

Figure C-4: Extrapolation via local linear regression, Arnold et al. (2022) approach




*Notes:* This figure illustrates extrapolated estimates of $E(Y_{i0})$ and $E(Y_{i1})$, as well as reports the associated ATT and ATU estimates, when using the local linear regression approach to extrapolation from Arnold et al. (2022). Specifically, for outcome $Y=\mathbf{1}$[one year reoffending], we estimate the regression $Y_{ijs} = \alpha_j + \psi_s + u_{ijs}$ and construct officer-level estimates $\theta_j^Y = \hat{\alpha}_j + E(\hat{\psi}_s)$. We then do the same procedure, replacing the outcome with $D = \mathbf{1}$[harsh fine] to get officer level estimates $\theta_j^D$. The figure plots the estimated $\theta_j^Y$'s (vertical axis) against the estimated $\theta_j^D$'s (horizontal axis) in the small gray dots. We then fit the $\theta_j^Y$'s to the $\theta_j^D$'s using a local linear regression with a guassian kernel and rule-of-thumb bandwidth = 0.1, weighting by each officer's number of stops, and use the fitted values at $\theta_j^D = 0$ and $\theta_j^D = 1$ as our estimates of $E(Y_{i0})$ and $E(Y_{i1})$. For inference, we use a bootstrap clustered at the officer-level.

Figure C-5: Extrapolation estimates for other outcomes




*Notes:* This figure graphically depicts the extrapolation estimates underlying tables 4 and C-1. Each panel is the same as panel (a) of figure 4 for a different outcome. In panels (a)-c(c), the outcome is whether a motorist reoffends in the following year. In panels (d)-(f), the outcome is the number of offenses in the following two years. Red diamonds depict the 5 percent tail mean estimates of $E(Y_{i0})$ and $E(Y_{i1})$ used to construct the estimates reported in table C-1.

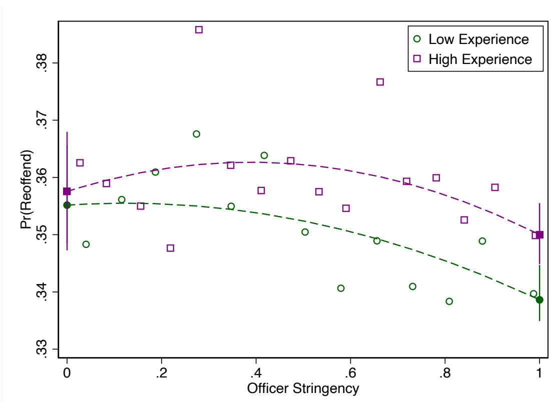
Table C-1: Extrapolation estimates for other outcomes
(using 5 percent tail mean estimates)

| | $Y_0$ | | | | $Y_1 - Y_0$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) All | (2) $D = 1$ | (3) $D = 0$ | (4) Diff | (5) All (ATE) | (6) $D = 1$ (ATT) | (7) $D = 0$ (ATU) | (8) Diff |
| *Panel A: Any offense in following year* | | | | | | | | |
| Any Offense | 0.353 | 0.367 | 0.325 | 0.042 | -0.010 | -0.003 | -0.024 | 0.021 |
| | (0.006) | (0.009) | (0.003) | (0.009) | (0.007) | (0.010) | (0.009) | (0.013) |
| Speeding | 0.179 | 0.188 | 0.161 | 0.027 | -0.013 | -0.011 | -0.017 | 0.006 |
| | (0.004) | (0.006) | (0.002) | (0.006) | (0.005) | (0.007) | (0.005) | (0.008) |
| Crash | 0.078 | 0.078 | 0.077 | 0.001 | 0.001 | 0.004 | -0.005 | 0.009 |
| | (0.002) | (0.003) | (0.001) | (0.003) | (0.002) | (0.003) | (0.003) | (0.004) |
| *Panel B: Number of offenses in following two years* | | | | | | | | |
| Any Offense | 1.275 | 1.357 | 1.118 | 0.239 | -0.037 | -0.000 | -0.107 | 0.107 |
| | (0.027) | (0.040) | (0.016) | (0.044) | (0.031) | (0.043) | (0.051) | (0.072) |
| Speeding | 0.395 | 0.416 | 0.355 | 0.061 | -0.023 | -0.014 | -0.039 | 0.024 |
| | (0.008) | (0.013) | (0.004) | (0.014) | (0.010) | (0.014) | (0.014) | (0.021) |
| Crash | 0.155 | 0.157 | 0.151 | 0.006 | 0.003 | 0.007 | -0.005 | 0.012 |
| | (0.003) | (0.005) | (0.002) | (0.005) | (0.004) | (0.005) | (0.005) | (0.008) |

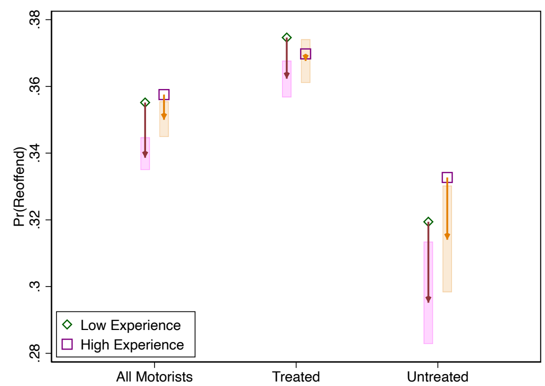
*Notes:* Same as table 4 except that we use the beat-adjusted average outcomes for officers with $Z < 0.05$ and $Z > 0.95$ as our estimates of $E(Y_{i0})$ and $E(Y_{i1})$ instead of our baseline quadratic extrapolation specification of equation 1.

Figure C-6: Extrapolation estimates by officer experience

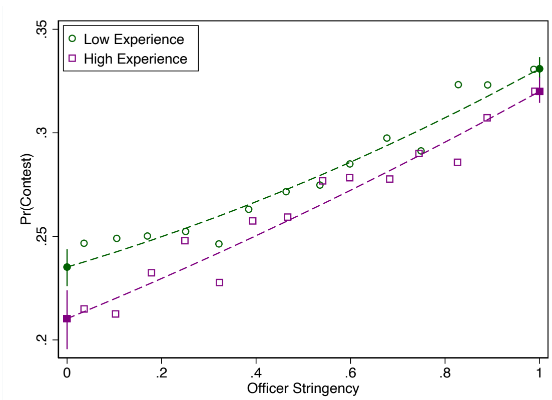
(a) Extrapolation: reoffending



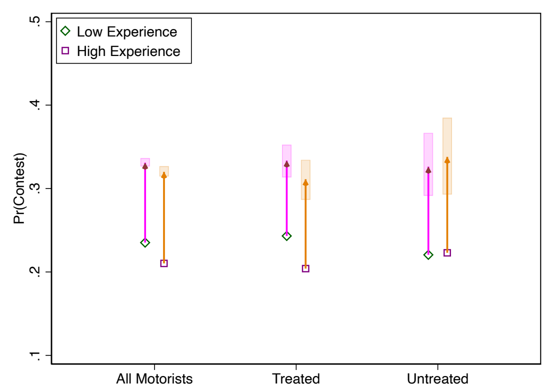

(b) Treatment effects: reoffending




(c) Extrapolation: contest in court




(d) Treatment effects: contest in court




*Notes:* Same as figure 4 except estimates are shown separately for officers with at least five years experience as FHP patrol officers as of January 1, 2007 (the start of our sample). In panels (a) and (b), the outcome of interest is whether a motorist commits a new traffic offense in the following year and in panels (c) and (d), the outcome of interest is whether a motorist contests a citation in traffic court. Treatment effect estimates shown in panels (b) and (d) are constructed from the quadratic extrapolations of $E(Y_{i0})$ and $E(Y_{i1})$ shown in panels (a) and (c). The estimates illustrated in panels (b) and (c) are also reported in appendix table C-2, and alternative versions of these estimates using a local mean estimator for $E(Y_{i1})$ and $E(Y_{i0})$ are presented in appendix table C-3.

Table C-2: Extrapolation estimates by subgroup

| | $Y_0$ | | | | $Y_1 - Y_0$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) All | (2) $D = 1$ | (3) $D = 0$ | (4) Diff | (5) All (ATE) | (6) $D = 1$ (ATT) | (7) $D = 0$ (ATU) | (8) Diff |
| *Panel A: By motorist offense history (prior two years)* | | | | | | | | |
| No offense | 0.246 | 0.258 | 0.225 | 0.033 | -0.018 | -0.013 | -0.026 | 0.013 |
| | (0.004) | (0.006) | (0.002) | (0.006) | (0.005) | (0.006) | (0.007) | (0.010) |
| Any offense | 0.463 | 0.476 | 0.435 | 0.041 | -0.010 | -0.007 | -0.018 | 0.011 |
| | (0.004) | (0.006) | (0.003) | (0.006) | (0.005) | (0.007) | (0.008) | (0.011) |
| *Panel B: By officer experience* | | | | | | | | |
| Low | 0.355 | 0.375 | 0.319 | 0.055 | -0.017 | -0.012 | -0.024 | 0.012 |
| | (0.005) | (0.008) | (0.004) | (0.008) | (0.006) | (0.009) | (0.011) | (0.016) |
| High | 0.358 | 0.370 | 0.333 | 0.037 | -0.008 | -0.002 | -0.019 | 0.017 |
| | (0.006) | (0.009) | (0.005) | (0.010) | (0.007) | (0.010) | (0.012) | (0.017) |
| *Panel B: By officer experience ($Y$ = Pr(contest))* | | | | | | | | |
| Low | 0.235 | 0.243 | 0.221 | 0.023 | 0.096 | 0.090 | 0.106 | -0.015 |
| | (0.006) | (0.011) | (0.013) | (0.022) | (0.006) | (0.018) | (0.031) | (0.047) |
| High | 0.210 | 0.204 | 0.223 | -0.019 | 0.110 | 0.107 | 0.115 | -0.008 |
| | (0.008) | (0.015) | (0.014) | (0.025) | (0.009) | (0.021) | (0.033) | (0.051) |

*Notes:* This table reports estimated untreated reoffending levels $E(Y_{i0})$ and treatment effects $E(Y_{i1} - Y_{i0})$ for all motorists, treated motorists, and untreated motorists, using different specifications for obtaining extrapolated estimates of $E(Y_{i0})$ and $E(Y_{i1})$. In panel (a), we show estimates from estimating equation 1 with different polynomials in $Z$. In panel (b), we estimate equation $Z$ using a piecewise linear specification with rule-of-thumb bandwidth = 0.2. In panel (c), we show estimates that use the beat-shift adjusted average $Y$ for officers in the tails of the stringency distribution, with $bw$ denoting the bandwidth used to define the tails. In panel (d), we present estimates from polynomial specifications of 1 that replace $Z$ with the estimated propensity score. Graphical analogues for the underlying extrapolation estimates are shown in figure C-2.

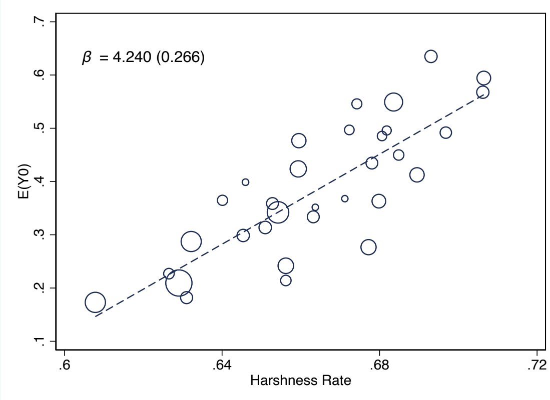
Table C-3: Extrapolation estimates by subgroup
(using 5 percent tail mean estimates)

| | $Y_0$ | | | | $Y_1 - Y_0$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | All | $D = 1$ | $D = 0$ | Diff | All (ATE) | $D = 1$ (ATT) | $D = 0$ (ATU) | Diff |
| *Panel A: By motorist offense history (prior two years)* | | | | | | | | |
| No offense | 0.247 | 0.260 | 0.225 | 0.035 | -0.020 | -0.015 | -0.029 | 0.014 |
| | (0.005) | (0.008) | (0.002) | (0.008) | (0.006) | (0.008) | (0.007) | (0.010) |
| Any offense | 0.458 | 0.468 | 0.435 | 0.032 | -0.004 | 0.002 | -0.017 | 0.019 |
| | (0.007) | (0.009) | (0.003) | (0.009) | (0.007) | (0.010) | (0.008) | (0.013) |
| *Panel B: By officer experience* | | | | | | | | |
| Low | 0.347 | 0.363 | 0.319 | 0.043 | -0.009 | -0.000 | -0.024 | 0.023 |
| | (0.006) | (0.010) | (0.004) | (0.010) | (0.008) | (0.010) | (0.013) | (0.016) |
| High | 0.361 | 0.374 | 0.333 | 0.042 | -0.011 | -0.007 | -0.020 | 0.013 |
| | (0.009) | (0.013) | (0.005) | (0.014) | (0.010) | (0.014) | (0.013) | (0.020) |
| *Panel B: By officer experience ($Y = Pr(contest)$)* | | | | | | | | |
| Low | 0.242 | 0.254 | 0.221 | 0.033 | 0.087 | 0.080 | 0.101 | -0.021 |
| | (0.007) | (0.013) | (0.013) | (0.023) | (0.008) | (0.019) | (0.031) | (0.047) |
| High | 0.218 | 0.215 | 0.223 | -0.008 | 0.102 | 0.096 | 0.115 | -0.019 |
| | (0.008) | (0.015) | (0.014) | (0.026) | (0.010) | (0.022) | (0.033) | (0.052) |

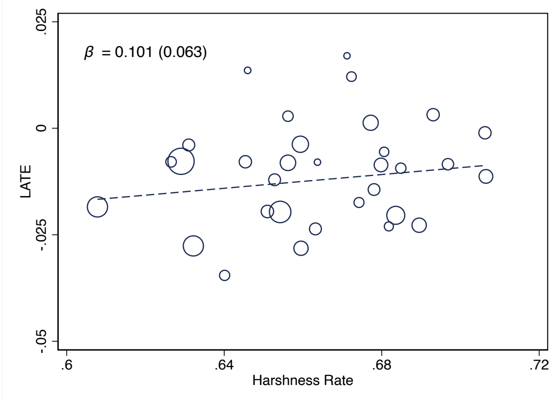
*Notes:* Same as table C-2 except that we use the beat-adjusted average outcomes for officers with $Z < 0.05$ and $Z > 0.95$ as our estimates of $E(Y_{i0})$ and $E(Y_{i1})$ instead of our baseline quadratic extrapolation specification of equation 1.

Figure C-7: Reoffending levels and treatment effects by motorist characteristics (Alternative estimates)

(a) Lenient officer mean



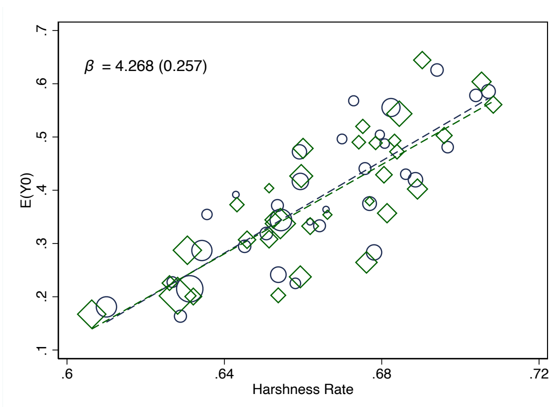

(b) LATE



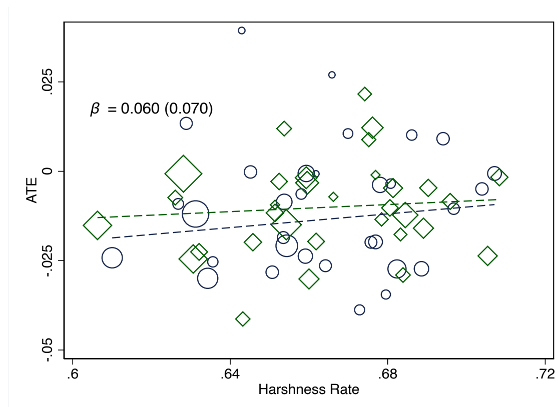

*Notes:* Same as figure 5 except that the untreated reoffending rate, reported in panel (a) is estimated by taking the average reoffending rate for lenient (bunching) officers; in panel (b), we show the group-specific LATE, estimated by 2SLS, rather than the extrapolation-based ATE estimate.

Figure C-8: Reoffending levels and treatment effects by motorist characteristics (Partition-based estimates)

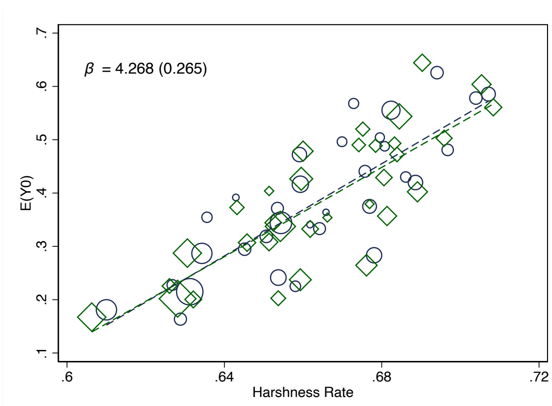
(a) Within officers: $E(Y_0)$



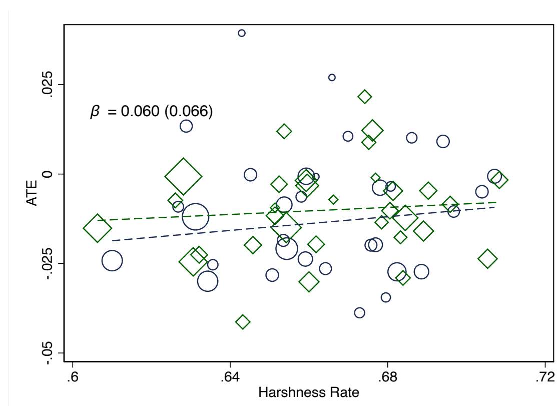

(b) Within officers: ATE




(c) Across officers: $E(Y0)$

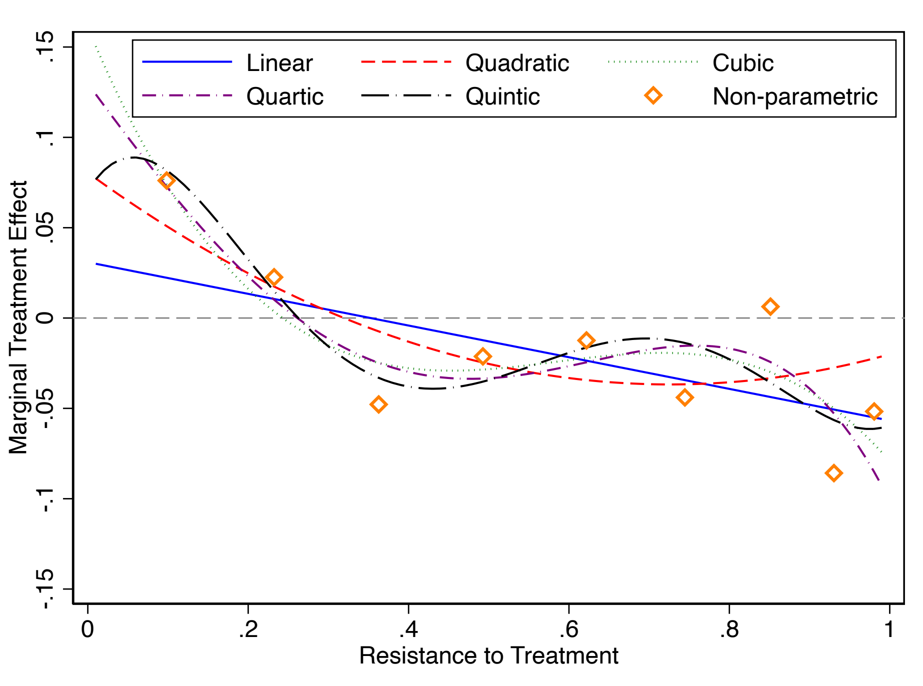



(d) Across officers: ATE




*Notes:* Same as figure 5 except that the harshness rate (horizontal axis) and the reoffending rate or treatment effect are estimated in separate partitions of the data. In panels (a) and (b), we construct partitions within officer and estimate the harshness rate, as well as $E(Y_{i0})$ and $E(Y_{i1})$ in each partition, and show correlations across partitions. In panels (c) and (d), we instead construct partitions across officers. Each figure shows the cross-partition correlations for one partition with blue circles and the other with green diamonds. Each figure reports the average of the linear slope in each partition and the boostrapped standard error.

Figure C-9: Marginal treatment effect estimates




*Notes:* Marginal treatment effect estimates of impact of harsh fine on any reoffending in following year, as described in Section 5.7. Legend reports the order of polynomial used for the MTE specification. The non-parametric specification estimates the MTE separately across bins of the propensity score via the binscatter approach.

Table C-4: Treatment effect estimates based on marginal treatment effect (MTE) curves

| *MTE Specification* | Average (ATE) | Treated (ATT) | Untreated (ATU) | Diff |
|---|---|---|---|---|
| Linear | -0.013 | -0.005 | -0.027 | 0.021 |
| | ( 0.004) | ( 0.004) | ( 0.004) | ( 0.004) |
| Quadratic | -0.007 | 0.001 | -0.023 | 0.024 |
| | ( 0.004) | ( 0.005) | ( 0.004) | ( 0.004) |
| Cubic | -0.006 | 0.005 | -0.026 | 0.031 |
| | ( 0.004) | ( 0.005) | ( 0.004) | ( 0.005) |
| Quartic | -0.007 | 0.004 | -0.026 | 0.030 |
| | ( 0.004) | ( 0.005) | ( 0.004) | ( 0.005) |
| Quintic | -0.007 | 0.003 | -0.026 | 0.029 |
| | ( 0.004) | ( 0.005) | ( 0.004) | ( 0.005) |
| Non-Parametric | -0.008 | 0.002 | -0.027 | 0.029 |
| | ( 0.004) | ( 0.005) | ( 0.005) | ( 0.005) |

*Notes:* Estimates of the Average Treatment Effect (ATE), Treatment Effect on the Treated (ATT), and Treatment Effect on the Untreated (ATU), based on the MTE curves presented in Figure C-9. ATE is defined as $E(Y_1 - Y_0)$ and is estimated by summing equally across a uniform grid of MTE estimates from resistance to treatment points $u = 0.01, 0.02, ..., 0.99$. ATT is defined as $E(Y_1 - Y_0|U_D < \mu_D(Z))$ and is estimated by summing across the same grid of points and weighting by $Pr(u < \hat{p})$, the probability of having a propensity score greater than that resistance to treatment. ATU is defined as $E(Y_1 - Y_0|U_D \geq \mu_D(Z))$ and is estimated by the same summation but with weights $Pr(u \geq \hat{p})$.

# D Data Appendix

## D-1 Traffic courts data

Traffic court dispositions associated with the citations from the *TCATS* database were also shared by the Florida Clerk of Courts. Citations were matched to disposition information using county codes and alphanumeric citation identifiers (which are unique within counties). Some citations have no associated disposition in the *TCATS* database, while others have multiple associated entries. Disposition verdicts can take on the following values:

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

In practice, the verdicts 1, 3, 4, A, and C account for the vast majority of citations. Moreover, as confirmed in a phone conversation with Beth Allman at the Florida Clerk of Courts on July 24, 2018, several of the violation codes are difficult to interpret. In particular, it is very difficult in practice to infer the precise outcome of tickets with disposition codes *1*, *3*, *A*, or those with multiple dispositions in the TCATS database.

To construct an approximate measure of court contesting, we use any disposition not equal to 4 or $C$, which both imply that the individual paid their fine without contest, as an indicator that the driver contested a citation. To construct measures of *effective* sanctions, termed *paid* fines and *accrued* points in figure A-3, we adjust the statutory sanctions as follows:

- Replace fine = fine/2 if verdict = A
- Replace fine = 0 if verdict = 3
- Replace points = 0 if verdict $\in$ {3,A,C}

Note that our measure of *paid* fines is likely conservative as it ignores court fees. Drivers contesting their tickets in court face a $75 court fee in addition to their fine (the court fee can also be waived during the court process). See Goncalves & Mello (2021) and Mello (2021) for further discussion of the issues associated with working with the TCATS data.

## D-2 Binary stringency measure

To identify officers who do not bunch, we use the Frandsen (2017) test for manipulation. In our setting, this test implies that, under the null hypothesis of no manipulation, the conditional probability of being found at the bunching speed is in a range around one third, $Pr(X = 9|x \in [8, 10]) \in [(1 - k)/(3 - k), (1 + k)/(3 + k)]$ where $k$ is a restriction on the second finite difference, $\Delta^{(2)}Pr(S = 9) \equiv Pr(S = 8) - 2Pr(S = 9) + Pr(S = 10)$, such that $|\Delta^{(2)}Pr(S = 9)| \leq k(Pr(S = 9) - Pr(S = 10))$. Intuitively, if the distribution of ticketed speeds is unmanipulated, the share of tickets at 9 MPH among those between 8 and 10 MPH 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 estimating the mean parameter $\lambda$ using the empirical mean of ticketed speeds. We say that an officer is stringent (non-bunching) if we fail to reject that $Pr(S = 9|S \in [8, 10]) \leq (1 + k)/(3 + k)$ at the 99 percent confidence level.

To avoid the reflection problem, we randomly partition an officer's stops into two halves and compute the binary measure separately for each half of the sample. We then use the officer's binary measure in the *other half* as our binary stringency measure.

## D-3 Predicted recidivism

At a few points in our analysis, we rely on a predicted reoffending measure based on covariates (e.g., figure 2). An important concern in constructing this measure is the possibility of contamination from treatment effects (e.g., if certain covariates are highly correlated with receiving harsh fines, the predicted recidivism for these covariate groups will be too low because of the treatment effect of harsh fines on future offending). To circumvent this concern, we construct our predicted recidivism index as follows. First, we randomly partition each officer's stop into halves. Each officer $\times$ partition is coded as lenient or nonlenient using the Frandsen (2017) described above. In each partition, we regress $\mathbf{1}$[any new traffic offense] on driver covariates using only stops made by non-lenient (non-bunching) officers. We then use the coefficients from this regression to construct our $\hat{Y}$ measure in the other partition.

# E Technical Appendix

## E-1 Within-locations approach from Feigenberg & Miller (2022)

Let $j$ index officers, $\ell$ index counties, and $\tau$ index time categories, defined as shift $\times$ $\mathbf{1}$[weekend]. Let $t$ index time, defined as calendar year $\times$ month. We construct an adjusted stringency measure for each officer $\times$ location by estimating the regression:

$$D_{ij\ell\tau} = \phi_{j\ell\tau} + \gamma X_{ij\ell\tau} + \delta_\tau + u_{ij\ell\tau}$$

where the $\phi$'s are fixed effects for each officer $\times$ location $\times$ shift category and $D = \mathbf{1}$[harsh fine]. We estimate this regression separately for each location. We then aggregate to the officer $\times$ location level as follows:

$$\tilde{D}_{j\ell} = \sum_{\tau} \left( \hat{\phi}_{j\ell\tau} + E[\hat{\gamma} X_{ij\ell\tau} + \hat{\delta}|\ell, \tau] \right) P(\tau|\ell)$$

We then repeat the exact same procedure, replacing $D$ with $Y = \mathbf{1}$[reoffend] to obtain an adjusted probability of reoffending for each officer $\times$ location, $\tilde{Y}_{j\ell}$.

Following Feigenberg & Miller (2022), we document the relationship between $\tilde{Y}_{j\ell}$ and $\tilde{D}_{j\ell}$ in two ways. First, we show a simple binscatter, conditional on location fixed effects and weighting each officer $\times$ location by number of stops via the `binsreg` command. We also estimate a linear slope corresponding to this approach by regressing $\tilde{Y}_{j\ell}$ on $\tilde{D}_{j\ell}$ and location fixed effects, weighting by the number of stops. Second, we construct location-specific quantiles of each and plot the relationship between quantiles.

In figure C-3, we show a version of our extrapolation based on this Feigenberg & Miller (2022) approach. Specifically, we fit the relationship between $\tilde{Y}_{j\ell}$ and $\tilde{D}_{j\ell}$ using a quadratic, conditioning on location fixed effects, and use the fitted values at $\tilde{D} = 0$ and $\tilde{D} = 1$ as our estimates of $E(Y_{i0})$ and $E(Y_{i1})$. As shown in figure C-3, this quadratic approximates both the binscatter and the quantiles estimates quite well.

## E-2 Calculating stopped speed for lenient tickets

We will use a version of the extrapolation approach to identify the stopped speed of drivers who are given a lenient ticket ($D = 0$). Denoting stopped speed by $S_i^*$ and ticketed speed by $S_i$, we know by design that $S_i(D = 1) = S_i^*$ and $S_i(D = 0) = 9$.

We first estimate a version of Equation 1, where stopped speed is on the left-hand side: $E(S_i|W_i, Z_{ij}) = \alpha W_i + f(Z_{ij})$. The extrapolation for the treated outcome gives us the average stopped speed, $\hat{\alpha}\bar{W}_i + f(1) = E(S_{i1}) = E(S_i^*)$. Next, we know that the stopped speeds of harshly and leniently ticketed drivers have to weight to the average, $E(S_i^*) = pE(S_i^*|D = 1) + (1 - p)E(S_i^*|D = 0) = pE(S_i|D = 1) + (1 - p)E(S_i^*|D = 0)$. Rearranging gives us the stopped speed of leniently ticketed drivers, $E(S_i^*|D = 0) = \frac{1}{1-p}E(S_i^*) - \frac{p}{1-p}E(S_i^*|D = 1)$.

Next, we want to understand whether the observed positive selection on levels can be explained by officers' preference for punishing drivers stopped at faster speeds. In other words, if we let $E(Y_1|S_i^*)$ be the expected offending of a driver (when treated harshly) stopped at speed $S^*$, what are $E(E(Y_1|S_i^*)|D_i)$ for $D_i = 0$ and 1 and how does the difference compare with the true difference $E(Y_{1i}|D_i)$ for $D_i = 0$ and 1?

There are two steps to this estimation. First, we need to estimate $E(Y_{1i}|S_i^*)$, which is the predicted reoffending rate for drivers stopped at a given speed. We can do this by examining supremely stringent officers, so that $S = S^*$ for practically all drivers. So we estimate $E(Y_{1i}|S_i^*) = E(Y_i|Z_i \approx 1)$.

The second step is to identify the distribution of stopped speeds separately by treatment status, $Pr(S^* = s|D_i)$. This distribution is directly observed for drivers treated harshly, and we can directly calculate $E(E(Y_{1i}|S_i^*)|D_i = 1)$ by taking the average of $E(Y_{1i}|S_i^*)$ over those drivers.

For drivers with $D_i = 0$, we need to estimate the distribution of stopped speeds. For each speed $s$, we run a version of our extrapolation where equation 1 is estimated with the left-hand side replaced with an indicator for whether the driver is ticketed at $S_i = s$. The extrapolation from the most stringent officers gives us an estimate of $Pr(S^* = s)$, the overall probability a driver is stopped at this speed. We can then infer the frequency of this speed for the leniently treated drivers from $Pr(S^* = s|D_i = 0) = \frac{1}{1-p}Pr(S^* = s) + \frac{p}{1-p}Pr(S^* = s|D_i = 1)$.

Our estimate for the predicted reoffending rate for leniently treated drivers is then $E(E(Y_{1i}|S_i^*)|D_i = 0) = \sum_s Pr(S^* = s|D_i = 0)E(Y_{1i}|S_i^*)$.