Would Eliminating Racial Disparities in Motor Vehicle Searches Have Efficiency Costs?*

Benjamin Feigenberg

Conrad Miller

May 2021

Abstract

During traffic stops, police search black and Hispanic motorists more than twice as often as white motorists, yet those searches are no more likely to yield contraband. We ask whether equalizing search rates by motorist race would reduce contraband yield. We use unique administrative data from Texas to isolate variation in search behavior across and within highway patrol troopers and find that search rates are unrelated to the proportion of searches that yield contraband. We find that troopers can equalize search rates across racial groups, maintain the status quo search rate, and increase contraband yield. Troopers appear to be limited in their ability to discern between motorists that are more or less likely to carry contraband. *JEL Codes*: J15, K42.

*We thank the editors, Larry Katz and Stefanie Stantcheva, five anonymous referees, Ian Ayres, Felipe Goncalves, Peter Hull, Patrick Kline, Jonathan Leonard, David Levine, Dan O'Flaherty, Steven Rivkin, Evan Rose, Yotam Shem-Tov, Chris Walters, and seminar participants at USC, University of Illinois, Chicago, the Online Economics of Crime seminar, the Online Economics of Racism seminar, NBER Summer Institute, Yale SOM, Harvard, Stanford, and MIT for helpful comments. We thank researchers at the Stanford Open Policing Project for providing data on Texas Highway Patrol stops. We thank the Fisher Center for Real Estate and Urban Economics at Haas for providing computing resources.

I. Introduction

During routine traffic stops, black and Hispanic motorists are more than twice as likely to be searched for contraband by police than white motorists (Pierson et al., 2020). These stark disparities invite allegations that police engage in *racial profiling*, using race as one factor when deciding whether to search someone. Many regard the practice as unjust: perceived profiling undermines trust in police (Epp et al., 2014) and racial disparities in search rates likely contribute to racial differences in arrests and exposure to police use of force (Fryer, 2019). However, equalizing search rates by race may reduce the effectiveness of policing if race is an informative predictor of criminal behavior (Persico, 2002).

In this paper we evaluate whether racial profiling in fact poses an equity-efficiency trade-off. The answer has important practical implications; recent legal scholarship argues that profiling is not legally permissible in the absence of "legitimate law-enforcement-related necessity" (Tiwara, 2019). Researchers have found that the percentage of searches that yield contraband—known as the *(average) hit rate*—among black and Hispanic motorists is typically equal to or lower than the hit rate for white motorists (e.g., Pierson et al., 2020). Some argue that this pattern indicates equal or lower rates of offending among black and Hispanic motorists (Harcourt, 2004). This reasoning suggests that equalizing search rates across motorist racial groups would not decrease overall contraband yield.

This argument implicitly assumes that equalizing search rates across motorist racial groups would not change group-specific hit rates. However, this assumption fails to hold if troopers face diminishing returns to search. If racial profiling is (accurate) statistical discrimination applied by police to maximize the proportion of their searches that yield contraband, then the hit rate for the marginal black or Hispanic motorist—the last black or Hispanic motorist deemed suspicious enough to be searched—will be equal to that for the marginal white motorist. This logic motivates the Becker (1957, 1993) *outcome test*: to evaluate whether police are on the efficient frontier, test whether *marginal* hit rates are equal across motorist racial groups. If troopers face diminishing returns to search, the hit rate for the average and marginal search may differ significantly, and comparisons of average hit rates between motorist racial groups may be uninformative—the well-known *inframarginality problem* (Ayres, 2002).

We test for an equity-efficiency trade-off using data on traffic stops for speeding violations conducted by Texas Highway Patrol troopers, where we define equity as the equalization of search rates across motorist racial groups. We assess whether search rate equalization would reduce contraband yield by exploiting variation in search behavior across troopers. In our setting, the identity of the trooper conducting a speeding stop is plausibly exogenous conditional on the location and time of the stop. We measure variation across troopers in the rate at which they search motorists—their *search rate*. Across troopers, we estimate the relationship between search rates and the percentage of stops that yield contraband (the *unconditional* hit rate), where we calculate these rates separately by motorist racial group. Strikingly, this relationship is approximately linear within each motorist group, implying approximately constant returns to search across troopers. In other words, troopers who search motorists twice as often find contraband twice as often. We show that, under conditions consistent with our setting, this result implies that there is no inframarginality problem because average and marginal hit rates are similar. Among motorists searched with positive probability, troopers appear unable to distinguish between those who are more or less likely to carry contraband. Our findings imply that it is feasible for troopers to (1) search all motorist racial groups at the same rate, (2) maintain the status quo overall search rate, and (3) increase contraband yield.

Our analysis proceeds in four steps. First, we summarize racial disparities in search rates and hit rates. A unique feature of the data is that they contain identifying information on stopped motorists, which allows us to (a) track motorists across multiple stops and to (b) merge in additional data on motorist characteristics, including criminal history and neighborhood income. Conditional on stop location and time, we find that black and Hispanic motorists are about 170% and 70% more likely to be searched than white motorists. However, searches of these motorists are about 15% and 30% less likely to yield contraband, a pattern consistent with prior work (Pierson et al., 2020). Controlling for stop history, criminal history and neighborhood income reduces the black-white and Hispanic-white disparities in search rates by about 50% and 35%. The residual black-white difference in search rates is similar in magnitude to the increase in search likelihood associated with multiple previous non-drug arrests and half of the increase associated with a prior drug arrest. Among stopped motorists with no arrest record at the end of our sample period, black and Hispanic motorists are about 200% and 80% more likely to be searched than white motorists. We investigate whether these stark racial disparities in search rates, even among motorists who do not engage in crime, are a necessary feature of contraband yield maximization.

Second, we introduce a simple model of trooper search behavior to frame our analysis. We build on Anwar and Fang (2006), where troopers decide whether to search a stopped motorist using a noisy signal for the motorist's guilt. Prior research assumes that troopers can strictly rank motorists by contraband risk and hence face strictly diminishing returns to search. This implies that troopers equalizing marginal hit rates across groups leads to an equity-efficiency trade-off. By contrast, we allow signals to be coarse in the sense that, among those motorists searched with positive probability, troopers are unable to distinguish between those who are more or less likely to carry contraband. In this case, the returns to search are constant, and equalizing hit rates at the margin no longer implies an equity-efficiency trade-off.

Third, we use between-trooper variation in search behavior to trace out the relationship between trooper search rates and unconditional hit rates—the between-trooper *search productivity curve (SPC)*—separately by motorist group. Our identifying assumption is that, conditional on the location and time of a stop, the identity of the trooper conducting the stop is as good as random. For a speeding violation at a given time of the week and on a given stretch of highway, the identity of the trooper making the stop will vary due to week-to-week variation in shift schedules and to within-shift variation in exact trooper location. We show that troopers vary in the rate at which they search motorists following speeding stops. While motorists are searched in about 1% of stops, search rates range across troopers from 0% at the 10th percentile to 3% at the 90th percentile. Conditional on the location and time of the stop, motorists with different characteristics are stopped by troopers with similar search rates. We find that the between-trooper SPC is approximately linear for each motorist group. The slopes of groups-specific SPCs—which give the between-trooper returns to search—are similar for black and white motorists, while the returns to search are lower for Hispanic motorists.

The key threat to our approach is that, conditional on our measures of stop location and time, troopers vary in the composition of motorists they stop. This variation may exist because our location and time measures are not sufficiently granular or because, in the same environment, troopers vary in the motorists they decide to stop. We address this concern using several approaches. We show that baseline SPC estimates do not change if we use more granular location and time measures and control directly for observable motorist characteristics. Baseline SPC estimates are invariant to removing troopers that stop motorists with unusual observable characteristics from the analysis. We also corroborate our baseline SPC estimates using two alternative research designs. First, we find similar patterns when we rely only on within-motorist variation in outcomes among motorists stopped multiple times. This approach nets out time-invariant unobservable motorist characteristics. Second, we compare stop outcomes on opposite sides of trooper patrol area borders in a spatial regression discontinuity (RD) design. Along the same highway route, the composition of troopers making traffic stops changes sharply across patrol area borders. We use this feature to validate trooper search rates as measures of the causal effect of trooper assignment on search likelihood and to confirm that average and marginal hit rates are approximately equal.

Fourth, we construct two types of policy counterfactuals that reduce racial disparities in search rates. In one type, we reallocate troopers that search at high rates to patrol areas where motorists are disproportionately white and troopers that search at low rates to patrol areas where motorists are disproportionately black and Hispanic. The key identification challenge is that, for the troopers we reallocate, we must infer what their search behavior would be in locations where we do not observe them. We make assumptions about counterfactual search behavior that are informed by the troopers that we do observe in multiple locations, whose search behavior is highly correlated across locations. We find that reallocating troopers can reduce the black-white disparity in search rates by more than 60% and eliminate the Hispanic-white disparity while increasing contraband yield.

In the second policy counterfactual we require troopers to equalize their search rates across motorist racial groups. Determining what would happen in this scenario requires knowledge of the *within-trooper* SPC: the within-trooper relationship between search rates and contraband yield. If the between-trooper and average within-trooper SPCs are similar, the between-trooper SPC estimates imply that troopers can equalize search rates while increasing contraband yield. However, the between-trooper SPC and average within-trooper SPC may differ if, for example, troopers that are better at screening also search motorists at higher rates. The two SPCs are identical if trooper screening ability and search propensity are independent.[1]

We document three pieces of evidence consistent with skill-propensity independence. First, we find that low and high search rate troopers search motorists with similar observable characteristics. This suggests that low and high search rate troopers are applying a similar, coarse screening of motorists. Second, we find that between-trooper SPCs do not materially vary with observable trooper characteristics, including experience, stop rate, and race. Third, we provide direct evidence that the between-trooper and average within-trooper SPC are similar. Estimating the within-trooper SPC is complicated by the fact that within-trooper variation in search rates in part derives from changes in trooper screening ability and motorist composition. We use variation in search rates across locations, where effective search costs vary due to factors like manager preferences, peer composition, and staffing levels. We instrument for search using location by year leave-out search rates in a model with both trooper and motorist fixed effects. The within-trooper SPC slope we estimate is statistically indistinguishable from the between-trooper slope.

We conclude that, in partial equilibrium, racial profiling does not present an equity-efficiency trade-off. Moreover, we present suggestive evidence that motorist racial group-specific *deterrence* effects are negligible at the margin, implying that predicted changes in search productivity are unlikely to be offset by changes in contraband carrying behavior. We exploit the large-scale reallocation of troopers to the border region in 2014 as part of "Operation Strong Safety." We show that this well-publicized influx of troopers resulted in a dramatic increase in the number of stops and searches conducted with little corresponding change in the hit rate.[2] An important caveat that underpins our policy conclusions is that we abstract away from policy-specific behavioral responses. As we discuss in Section V.C, it is possible that troopers would respond to systematic reallocations or to mandated changes in search rates in unanticipated ways (for example, driven by trooper resistance to associated policy aims).

Previous researchers have argued that disparate policing behavior is driven by racial bias (e.g., Pierson et al., 2020), and this channel offers one potential explanation for our findings. We examine how search patterns vary with several proxies for trooper preferences and beliefs: trooper race (Anwar and Fang, 2006; Close and Mason, 2007; Antonovics and Knight, 2009; West, 2018), local political preferences (Cohen and Yang, 2019), and citation behavior (Goncalves and Mello, forthcoming). We find that all trooper racial groups are more likely to search black and Hispanic motorists than white motorists, but the black-white disparity is smaller for black troopers. The black-white search disparity is also substantially larger in counties with higher Republican vote shares in the 2016 presidential election. While we find no clear link between black-white disparities in search and citation rates, in counties where troopers cite Hispanic motorists more often, troopers are also more likely to search Hispanic motorists.

[1] Throughout the paper, by *search propensity* we mean searches per stop. By *screening ability* we mean contraband yield per search, conditional on search propensity. We discuss these terms in more detail in the model presented in Section III.

[2] Our findings are consistent with MacDonald and Fagan (2019) who find that a New York Police Department policy that increased search and frisk rates during civilian stops in specific locations, particularly of black and Hispanic civilians, did not affect hit rates.

Our paper relates closely to a series of papers that apply the reasoning of the Becker (1957, 1993) outcome test to investigate racial bias. Two seminal papers develop tests of racial bias that attempt to circumvent the inframarginality problem. Both papers document large racial disparities in search rates and similar or lower hit rates for black and Hispanic motorists, but conclude there is no evidence of racial bias against black motorists. Knowles et al. (2001) develop a model in which all motorists must carry contraband with equal probability in equilibrium if troopers are not racially biased. While we find that marginal and average hit rates are similar empirically, the variation in screening ability and lack of deterrence effects that we document are inconsistent with the Knowles et al. (2001) framework. Anwar and Fang (2006) argue that if troopers are not racially biased, the ranking of search and hit rates by white and black troopers should be unaffected by the motorist's race. Applying this test to our data, we do not find evidence of relative bias among black, white, and Hispanic troopers.

More recent work has used quasi-experimental variation in decision-maker assignments to address inframarginality without fully specifying underlying decision-making models (Hull, 2021). Arnold et al. (2018) use the quasi-random assignment of defendants to bail judges and assume *strict monotonicity*—that bail judges share the same ranking of defendants by misconduct risk.[3] Under strict monotonicity—an assumption that does not hold in our setting—our results would imply that Texas state troopers exhibit racial bias against Hispanic, but not black, motorists, yet troopers search black motorists more often than white motorists with no associated efficiency gains. This highlights a limitation of the Becker (1957, 1993) outcome test: when the returns to search are constant, equalized marginal hit rates do not imply an equity-efficiency trade-off.[4]

Methodologically, we build on Arnold et al. (2020) and Chan et al. (2020), who do not assume monotonicity and identify variation in both preferences and screening ability across decision-makers facing similar cases.[5] Arnold et al. (2020) use the quasi-random assignment of defendants to bail judges to non-parametrically identify differential treatment by race conditional on pretrial misconduct risk and to estimate a structural model of judge behavior that recovers the distribution of bias and screening ability across judges. Chan et al. (2020) exploit the quasi-random assignment of patients to physicians to estimate a structural model that recovers the distribution of diagnostic skill and preferences across physicians. A common feature of these two papers and ours is that they use *between-agent* variation in behavior to make inferences about policy counterfactuals that require *within-agent* changes in behavior. A key distinguishing feature of our analysis is that we use quasi-experimental within-agent variation in behavior to support the validity of this extrapolation.

A second branch of the economics literature considers whether profiling is justified, either on

[3] Marx (forthcoming) applies a logic similar to that of Arnold et al. (2018) to the policing context to bound marginal hit rates for each motorist racial group.

[4] In the Hull (2021) framework, our findings reflect a broader form of discrimination, termed *unwarranted disparities*, whereby black and white motorists are subject to differing rates of search despite comparable "qualification" for search (i.e., likelihood of carrying contraband). Our findings are also consistent with recent work demonstrating that, due to poor targeting, racial disparities in criminal justice punishment in several contexts can be substantially reduced without increasing crime (Kleinberg et al., 2018; Rose, 2021; Arnold et al., 2020).

[5] Simoiu et al. (2017) also address the inframarginality problem by jointly estimating decision thresholds and risk distributions in a parametric model, but do not isolate variation in behavior across decision-makers.

efficiency or ethical grounds. Several papers argue that even profiling that relies on unbiased statistical discrimination to maximize contraband yield may be inefficient in the presence of deterrence effects if the actual social goal is to minimize crime (Persico, 2002; Dominitz and Knowles, 2006). We provide suggestive evidence that deterrence effects are limited in our setting, and so efficiency given the social goal of crime minimization can be assessed on the basis of contraband yield.[6]

II. Context and Data

In this section, we discuss the institutional setting and describe the legal framework pertaining to the use of race as a factor in trooper decisions to search a motorist or vehicle. We then describe the combination of three datasets that we use to characterize patterns in search rates and outcomes by motorist racial group: (1) administrative data on traffic stops conducted by the Texas Highway Patrol, (2) administrative data on individual criminal histories in Texas, and (3) commercial address history data. Lastly, we present descriptive statistics for stops and searches across all motorists and separately by motorist racial group.

II.A. Institutional Setting

We study the search behavior of highway patrol troopers. In Texas, the primary responsibility of highway patrol troopers is to enforce state traffic laws on highways and state roads, but they have authority to enforce state criminal law throughout the state. During a typical shift, troopers conduct an average of eight traffic stops. When conducting each traffic stop, a trooper will give either a warning or citation for the original traffic violation. Troopers may also decide to further investigate if they suspect that a motorist may be carrying contraband, such as illicit drugs or weapons. As part of their investigation, troopers may search the motorist or vehicle for contraband. Troopers typically work alone, but may wait for support when conducting searches.

In our setting, there are four types of searches: consent, probable cause, incident to arrest, and inventory. Inventory searches are searches that occur after a vehicle is ordered impounded. In these instances, police are free to search the inventoried vehicle subject to departmental search policy. Incident to arrest searches are searches that occur following an arrest. After an arrest, troopers can search the arrested individual for contraband and, under broad conditions, search the vehicle. Alternatively, troopers have the right to conduct a search if they have probable cause to believe a law has been broken. Last, in a consent search, a trooper conducts a search only after receiving permission from the motorist to do so. In our sample, 84% of searches are consent and probable

[6] It is also the case, however, that under realistic conditions racial profiling requires that 'innocent' black and Hispanic motorists—those without contraband—are searched more often than similarly 'innocent' white motorists (Durlauf, 2006). Indeed, as we note, black and Hispanic motorists in our setting with no arrest record by the end of the sample period are much more likely to be searched after a stop than white motorists. These disproportionate searches of 'innocent' black and Hispanic motorists may impose significant social costs. In assessing the trade-offs associated with racial profiling, an extensive literature at the intersection of economics and philosophy also emphasizes that ethical factors, such as fairness, merit consideration independent of their relationship to individual utility (Durlauf, 2006; Sen, 1979; Hahn, 1982).

cause searches. When contraband is discovered following a search, the motorist may be arrested on charges related to the contraband discovered.

Within these constraints, troopers have broad discretion when deciding whether to pursue or conduct a search (e.g., Goldstein, 1963; Kelling, 1999; Mastrofski, 2004). This is reflected in the substantial variation in search behavior across troopers that we document below.

II.B. Legal Framework

Whether police officers can legally use race as a factor in deciding to engage in routine activities, such as vehicle and motorist searches, remains controversial. In an early review of the relevant case law, Knowles et al. (2001) concludes that the legality of racial profiling is complex and context-specific. Legal scholars have also noted that constitutional challenges to racial profiling have largely been unsuccessful, often requiring plaintiffs to show evidence of "discriminatory purpose" (i.e., racial animus) underpinning the profiling behavior being challenged. An alternative avenue for redress is offered by Legal Code 34 U.S.C. § 12601 (Section 12601), which authorizes the Department of Justice to pursue cases against police departments engaged in unconstitutional practices. Indeed, the Department of Justice has historically taken action against a number of police departments for racially-targeted stops of pedestrians and motorists on this basis (Anderson, 2020), although documented transgressions in these departments were particularly egregious, including false arrests, illegal searches, and excessive use of force. In any case, Section 12601 cannot be used by private individuals seeking legal remedy for mistreatment, and its use by the Department of Justice is discretionary; since 2017, no new Section 12601 investigations have been initiated.

Recent legal scholarship, including Tiwara (2019), has challenged the legality of racial profiling on the basis of disparate impact liability arising under the Omnibus Crime Control and Safe Streets Act of 1968. Under this framework, a practice that has a disparate impact on minorities "may be permissible only if the police can demonstrate that it has a legitimate law-enforcement-related necessity for the use of the practice at issue" (Tiwara, 2019). In our context, evidence that state troopers search black and Hispanic motorists more frequently than white motorists without any associated efficiency gain would likely constitute a discriminatory practice on this basis.

II.C. Administrative Traffic Stop Data

The primary data source we use is a comprehensive dataset of 16 million traffic stops of motor vehicles conducted by the Texas Highway Patrol between 2009 and 2015. For each stop, the data include the date, time, location (including GPS coordinates), motorist's race and ethnicity, motorist's gender, information on the motor vehicle (including make, model, and year), the associated violation(s), whether a search was conducted, the rationale for each search, whether contraband was found, and the ID number of the trooper who conducted the stop. The data include both stops that result in warnings and citations. The data are similar to those used in earlier studies of racial profiling (see Anwar and Fang, 2006). A unique feature of the Texas data is that they include the motorist's full name and address. This identifying information allows us to augment the data in three ways: (1) we match multiple traffic stops to the same motorist, (2) we merge in criminal histories for each motorist using data described below, and (3) we use each motorist's address to identify their neighborhood (Census block group) median income.

The data only cover motorists who are stopped and not all motorists who could potentially be stopped. This constraint will be particularly relevant when we study variation across troopers in their search behavior because troopers may also vary in whom they decide to stop. If the composition of stopped motorists varies across troopers in ways that we cannot observe in the data, this will complicate our interpretation of between-trooper differences in search behavior. Due to this concern, we focus our analysis on what Epp et al. (2014) and Baumgartner et al. (2018) classify as *safety stops*, which they distinguish from *investigatory stops*. The goal of their classification is to distinguish stops by the trooper's motivation for the stop. In safety stops the trooper's motivation for conducting the stop is the traffic violation itself and not the characteristics of the motorist or vehicle. By contrast, in investigatory stops, troopers use minor traffic offenses as a pretext for pulling motorists over and potentially searching them or their vehicles. Troopers use more discretion in deciding whether to conduct an investigatory stop, and hence there is more potential for motorist characteristics to vary across troopers for these stops.

Our data do not identify the trooper's reason for the stop directly, but we infer this from the traffic or criminal violation(s) associated with the stop. To proxy for safety stops, we follow prior work (Epp et al., 2014; Baumgartner et al., 2018) and limit the sample to stops that include a speeding violation. This includes 61% of all stops. Consistent with our interpretation of speeding stops as predominantly safety stops, we show in Online Appendix B that variation across troopers in cited speeds is limited. In addition, in Section IV we measure and account for variation in the composition of stopped motorists across troopers.

We also limit our analysis to stops where the motorist has a Texas address and where the motorist is black, Hispanic, or white.[7] A prior investigation found that Texas state troopers incorrectly recorded many Hispanic drivers as white (Collister [2015]; see also Luh [2020]). Following Pierson et al. (2020), we recategorize motorists as Hispanic if they have a surname such that at least 75% of people with that surname identify as Hispanic in the 2010 Census.[8]

Finally, to reduce variation in stop location across troopers, we limit our analysis to stops made on state and interstate highways. This restriction excludes stops made on farm-to-market roads, ranch-to-market roads, county roads, and city streets, which account for about 26% of stops but have far fewer stops per miles of road than state and interstate highways. Online Appendix Table A1 summarizes the number of observations we drop with each sample restriction. After applying these restrictions, our sample includes 4,931,332 stops.

We divide motorists into four categories based on their history of previous traffic stops.[9] We

[7] We exclude Asian and American Indian motorists from the analysis because they make up less than 2% of stops.

[8] For the subsample of motorists with arrest records, the correlation between this constructed measure of Hispanic ethnicity and the measure included in Texas administrative criminal history data is 0.74 (0.75 for males and 0.70 for females).

[9] Stop histories are constructed using all stops, not just those meeting our sample criteria.

assign all motorists who have not had a previous stop to the first category. Motorists with a prior stop but no prior search are assigned to the second category. Motorists with a prior search but no prior search that yielded contraband are assigned to the third category. Motorists with a prior search that yielded contraband are assigned to the fourth category.

II.D. Administrative Criminal History Data

We construct motorist criminal histories using data from the Texas Computerized Criminal History System. These data are maintained by the Texas Department of Public Safety. State troopers have access to these same data when conducting stops. The data track state felony and misdemeanor criminal charges from arrest through sentencing up to 2015. Agencies are required to report data for all offenses that are Class B misdemeanors or greater, including all offenses that would potentially lead to a confinement sentence. The data include information on each criminal charge, including the original arrest charge, date of arrest, final court charge, charge disposition, and, if the charge results in conviction, the final sentence. The data include arrest charges that are ultimately dropped. We use these data to create summary measures of each motorist's criminal history at the time they are stopped. The data also include an individual's full name, address, race and ethnicity, gender, and a unique individual ID.

We construct two criminal history indices, one based on all drug offense arrests and another for non-drug offense arrests. For the drug offense index, we divide motorists into three categories. The first category is motorists with no prior drug-related arrests. Among motorists with any prior drug-related arrest, the median number of prior drug-related arrests is one. We assign remaining motorists to the second and third criminal history categories if their number of prior drug-related arrests is one and greater than one, respectively. We construct an analogous index for non-drug offense arrests. Among motorists with any prior non-drug offense arrest, the median number of prior non-drug offense arrests is two.

II.E. Commercial Address History Data

One shortcoming of the traffic stop data is that it does not include a unique motorist ID. The problem this presents is that for two traffic stops with the same associated motorist name but different addresses, we do not know whether these stops correspond to the same person. The criminal history data includes an individual identifier and allows us to construct a partial address history for a given person. But the addresses we observe in those data only correspond to the points in time when that person is arrested, if they have any criminal history at all.

To facilitate matching traffic stops and criminal history to a given motorist, we use commercial data on address history from Infogroup. These data are similar to address history data used in prior research, including Diamond et al. (2019) and Phillips (2020). For each individual, the data include their full name and street addresses at which the individual lived with estimated dates of residence. The data extract we use includes the addresses histories for all individuals in the database with a Texas residence between 2005 and 2016.

We merge traffic stops and criminal history to individuals using full name and address, incorporating address history data to account for address changes. The data merge is described in more detail in Online Appendix A. Note that we do not require a match with the address history data to include a traffic stop in the analysis.

II.F. Descriptive Statistics

We present descriptive statistics for our merged dataset in Table I. We report descriptive statistics for all stops and subset the data by motorist race. We do the same for all stops that lead to searches. The motorist is female in 36% of stops, white in 55% of stops, Hispanic in 35.9% of stops, and black in 9.1% of stops.[10] For about 43% of stops, the motorist has been stopped previously. For about 1.2% of stops, the motorist has been stopped and searched previously, and for about 30% of those stops, the motorist has also been found with contraband in a previous search. For about 9% of stops, the motorist has a previous non-drug arrest, and in about 3% of stops the motorist has a previous drug arrest. Troopers search motorists in 1.06% of stops and find contraband in 0.34% of stops.

[Table 1 about here.]

Black motorists are slightly less likely than white motorists to have been stopped in the past, but they are more likely to have been searched in the past. They are also more likely to have an arrest history. Consistent with past research on racial profiling (see Pierson et al., 2020), black motorists are nearly three times more likely to be searched than white motorists (corresponding search rates are 2.202% and 0.755%). For Hispanic motorists, stop history, criminal history, and search rates generally fall between white and black motorists. Hispanic and black motorists reside in neighborhoods with similar median incomes, while median neighborhood incomes for white motorists are higher.

Compared to all stopped motorists, searched motorists are about 18 percentage points more likely to be male and come from neighborhoods with median incomes that are 13 log points lower. Searched motorists are more than seven times more likely to have been searched previously, three times more likely to have a previous arrest unrelated to drugs, and six times more likely to have a previous drug-related arrest.

In Online Appendix B we summarize the joint determinants of search in a series of logistic regressions. Conditional on stop location and time, black and Hispanic motorists are about 170% and 68% more likely to be searched than white motorists. Conditioning further on motorist neighborhood income, expected neighborhood income given vehicle type, criminal history, and stop history reduces black-white and Hispanic-white odds ratios to 1.86 and 1.44. This rich set of controls can statistically account for about half of the black-white and 35% of the Hispanic-white disparities in search rates we estimate by conditioning on only stop time and location. The residual black-white

[10]For comparison, in 2010 the age 15 and above Texas population was 51% female, 49% non-Hispanic white, 34% Hispanic, and 12% non-Hispanic black.

difference in search rates is similar in magnitude to the increase in search likelihood associated with multiple previous non-drug arrests and half of the increase associated with a prior drug arrest.

The percentage of searches that yield contraband (the hit rate) is 31.9%. The hit rate for white motorists (37.4%) exceeds the hit rate for black motorists (34.0%), which exceeds the hit rate for Hispanic motorists (25.9%). This ranking is consistent with past research on racial profiling (see Pierson et al., 2020). In Online Appendix B we summarize the joint determinants of contraband yield among searches. Conditional on stop location and time, searches of black and Hispanic motorists are about 15% and 30% less likely to yield contraband than searches of white motorists. Conditioning further on both motorist income proxies, criminal history, and stop history attenuates these differences to about 10% and 25%.

Online Appendix Table B1 describes the distributions of search types, contraband types, and arrest outcomes. Black motorists are more likely to be subject to probable cause searches, and less likely to be subject to consent and inventory searches. Drugs make up 51.8% of contraband found, weapons make up 3.8%, and currency makes up 0.6%. In the remaining 44% of cases, contraband is characterized as "Other", a category that includes drug paraphernalia and open containers of alcohol. Across motorist racial groups, the most salient difference is that black motorists are about four and two percentage points more likely to be found with drugs and weapons than the pooled average, and are less likely to be found with "Other" contraband. We find that only 24.5% of stops that yield contraband lead to an arrest, similar to the rate documented in North Carolina by Baumgartner et al. (2018). This percentage is similar across motorist racial groups. The severity of arrest charges, as proxied by the average incarceration sentence associated with conviction, is also similar across groups.

III. A Model of Trooper Search Behavior

In this section, we first present our benchmark model, which considers the search behavior of a single trooper. We then extend the model to consider heterogeneity in trooper preferences and screening ability and to allow for multiple motorist groups (e.g., black, Hispanic, and white motorists).

III.A. Benchmark Model

We model trooper search behavior using a modified version of the model developed in Anwar and Fang (2006). Troopers decide whether to search a stopped motorist using a noisy signal for the motorist's guilt. The modification we make is to allow this signal to be coarse over some range so that troopers are unable to distinguish between more or less suspicious motorists in this range.

We begin with a continuum of motorists and we first consider the behavior of a single trooper. Suppose fraction π of motorists carry contraband. For each stopped motorist i, the trooper observes a noisy signal for the motorist's guilt, $\theta_i \in [0, 1]$. If the motorist is carrying contraband, the index θ is randomly drawn from a distribution with continuous probability density function (PDF) $f_g(\cdot)$ and cumulative distribution function (CDF) $F_g(\cdot)$; if the motorist is not carrying contraband, θ is randomly drawn from a continuous PDF $f_n(\cdot)$ and CDF $F_n(\cdot)$. (The subscripts g and n stand for "guilty" and "not guilty," respectively.)

We assume that $f_g(\cdot)$ and $f_n(\cdot)$ satisfy a modified version of the standard monotone likelihood ratio property (MLRP): $f_g(\theta)/f_n(\theta)$ is strictly increasing in θ for $\theta < \bar{\theta}$ and is **constant** for $\theta \geq \bar{\theta}$. The MLRP assumption on the signal distributions provides that a higher index θ signals that a motorist is more likely to be guilty. However, in our formulation, for sufficiently "suspicious" signals, there is a point at which signals are no longer informative at the margin about a motorist's likelihood of carrying contraband. In other words, signals are coarse in the sense that troopers can identify the riskiest motorists but, within this group, are unable to distinguish between those who are more or less likely to carry contraband. We make this assumption because, as we show in Section IV, it is consistent with the data.[11]

Let G denote the event that a motorist is found with contraband if searched. When a trooper observes a motorist with signal θ, the posterior probability that the motorist is guilty of carrying contraband, $Pr(G|\theta)$, is given by Bayes's rule:

$$P(G|\theta) = \frac{\pi f_g(\theta)}{\pi f_g(\theta) + (1-\pi) f_n(\theta)}.$$

From the MLRP, we have that $P(G|\theta)$ is strictly increasing in θ for $\theta < \bar{\theta}$. For $\theta \geq \bar{\theta}$, this probability is constant and is given by

$$P(G|\theta \geq \bar{\theta}) = \frac{\pi f_g\left(\bar{\theta}\right)}{\pi f_g\left(\bar{\theta}\right) + (1-\pi) f_n\left(\bar{\theta}\right)}.$$

Following the literature, we assume that the trooper's objective is to maximize the rate that traffic stops yield contraband, net of search costs. We further assume that search costs are a convex function, $C(\cdot)$, of the overall proportion of stops that result in searches, σ.

Given this cost structure, troopers will choose some threshold θ^* where troopers will search any motorist with $\theta_i \geq \theta^*$. Given this search threshold, the overall search rate is

$$\sigma(\theta^*) = \pi(1 - F_g(\theta^*)) + (1-\pi)(1 - F_n(\theta^*)).$$

The trooper's problem is to choose θ^* that maximizes their objective function

$$\int_{\theta^*}^{1} P(G|\theta^*) f(\theta) d\theta - C(\sigma(\theta^*)),$$

where $f(\theta) = \pi f_g(\theta) + (1-\pi) f_n(\theta)$. Hence, the trooper will set a threshold θ^* to equalize the

[11] We allow for one region for the most suspicious motorists where $f_g(\theta)/f_n(\theta)$ is constant because this fits the pattern we observe in the data. The framework can be readily extended to allow for alternative locations of this "flat" region or multiple such regions.

marginal costs and benefits of search for the marginal searched motorist:

$$P(G|\theta^*) = C'(\sigma(\theta^*)).$$

Given search threshold θ^*, the trooper's unconditional hit rate is

$$\eta(\theta^*) = \pi(1 - F_g(\theta^*)).$$

We define the contraband *yield* rate (or *hit rate*) as $\frac{\eta(\theta^*)}{\sigma(\theta^*)}$, the share of searches that yield contraband.

We denote the relationship between $\eta(\theta^*)$ and $\sigma(\theta^*)$ as the trooper's SPC. Equivalently, we define SPC implicitly as $\tilde{\eta}(\sigma) = \eta(\sigma(\theta^*))$. We present a theoretical example of this SPC in Figure I. By the MLRP, this relationship is linear where $\theta^* \geq \bar{\theta}$, and hence $\sigma(\theta^*)$ is low. As θ^* declines below $\bar{\theta}$, the relationship becomes concave, as the marginal searched motorist is less likely to have contraband than inframarginal searched motorists.

[Figure 1 about here.]

III.B. Trooper Heterogeneity

In practice, we will estimate feasible combinations of search rates and unconditional hit rates using variation in outcomes across troopers. More formally, let p index troopers. We will identify the set of outcomes for all troopers, $\{(\sigma_p^*, \eta_p^*)\}_{p \in P}$. We use this set to calculate the *between-trooper* SPC, the conditional expectation function for η_p^* given σ_p^*, which can be expressed as

$$\tilde{\eta}^{\mathrm{BT}}(\sigma) \equiv E[\eta_p^* | p \text{ s.t. } \sigma_p^* = \sigma].$$

This may differ from the SPC that an individual trooper faces if trooper-specific SPCs—the set of feasible outcomes for a specific trooper—are heterogeneous.

In our setting, troopers may vary in their search rates and unconditional hit rates because they face different search costs, $C_p(\cdot)$, which would lead to varying search thresholds, θ_p^*. Troopers may also vary in their ability to infer the contraband risk for each motorist in the sense that the signal distributions $f_g(\cdot)$ and $f_n(\cdot)$ may vary across troopers. In this case, troopers may vary in the unconditional hit rates they can achieve for a given search rate, leading to variation in trooper-specific SPCs. A uniformly higher SPC—meaning a trooper can achieve a (weakly) higher hit rate for every given search rate—corresponds to greater screening ability.

If troopers vary only in search costs, SPCs will not vary across troopers, and the between-trooper SPC we identify will correspond to each trooper's own SPC. This condition follows from the strict monotonicity assumption of Arnold et al. (2018). But if troopers vary in screening ability, SPCs will vary across troopers, and the between-trooper SPC we identify may no longer correspond to any particular trooper's SPC.

If trooper-specific SPCs vary, we can still define the *average within-trooper* SPC. We define the average within-trooper SPC as the average of trooper-specific SPCs,

$$\tilde{\eta}^{\mathrm{WT}}(\sigma) \equiv E_p[\tilde{\eta}_p(\sigma)].$$

For any given search rate σ, the value for the average within-trooper SPC is the average unconditional hit rate troopers would achieve if all troopers search at that rate. The between-trooper SPC we identify will correspond to the average within-trooper SPC if variation in trooper screening ability is independent of trooper search rates, σ_p^*. More formally, suppose there exists a function that assigns a skill α_p to each trooper j such that $\tilde{\eta}_p(\sigma) = \tilde{\eta}_{p'}(\sigma)$ for all search rates σ where $\alpha_p = \alpha_{p'}$. Then the between-trooper SPC identifies the average within-trooper SPC if α_p is independent of σ_p^*. This condition corresponds to the skill-propensity independence condition in Chan et al. (2020), and is implied by the average monotonicity condition of Frandsen et al. (2020). The condition is weaker than the strict monotonicity assumption of Arnold et al. (2018), which would require that any motorist searched by a given trooper would have also been searched by any trooper with a higher search propensity, and any motorist not searched by a given trooper would not have been searched by any trooper with a lower search propensity.

III.C. Disparities between Motorist Groups

The focus of this paper is on racial disparities in search rates and whether equalizing search rates across motorist racial groups would reduce contraband yield. Accordingly, we extend the model to allow for multiple motorist groups (e.g., black, Hispanic, and white motorists). In particular, we index groups by $r \in \{b, h, w\}$ and allow for group-specific signal distributions ($f_g^r(\cdot)$ and $f_n^r(\cdot)$) and search thresholds (θ_r^*), which imply group-specific SPCs. We also allow for search costs to depend on the search rates for each group so that costs are defined as

$$C(\sigma^b(\theta_b^*), \sigma^h(\theta_h^*), \sigma^w(\theta_w^*)).$$

By characterizing group-specific SPCs and identifying where troopers locate along those group-specific SPCs, we can determine whether troopers face an equity-efficiency trade-off.

There are two scenarios where no trade-off exists. In the first scenario, search productivity at the margin is unequal across motorist groups, and marginal productivity is lower for the group with the higher search rate. This corresponds to Panel B of Figure II. In this case, the Becker (1957, 1993) outcome test would identify troopers as biased.

In the second scenario, search productivity at the margin is equalized across groups, but $\theta_r^* \geq \bar{\theta}_r$ for $r \in \{A, B\}$, and search rates are unequal across groups. This scenario is depicted in Panel C of Figure II. Note that, in this scenario, troopers are unbiased in the sense of Becker (1957, 1993).[12] For comparison, Panel A of Figure II depicts a scenario in which an equity-efficiency trade-off is

[12] An alternative notion of bias is based on whether search rates are equal **among motorists with** $\theta_r^* \geq \bar{\theta}_r$. We are unable to test for this form of bias, however, because we cannot measure the total number of motorists meeting this condition.

present because equalizing marginal hit rates requires unequal search rates.

[Figure 2 about here.]

IV. Estimating the between-Trooper Search Productivity Curve

We have shown that black and Hispanic motorists are more likely to be searched than white motorists, while searches of black and Hispanic troopers are equally or less likely to yield contraband. The central question of this paper is whether equalizing search rates across motorist racial groups would decrease contraband yield. To answer this question, we first estimate the relationship between trooper search rates and unconditional hit rates. We present evidence that different troopers search equivalent groups of motorists at varying rates and examine how troopers' search rates relate to their search productivity.

An essential requirement of our approach is that we measure how outcomes vary across troopers for equivalent stops. There is no random assignment of troopers to stops in our context. Instead, we will make a *selection on observables* argument—conditional on the time and location of the stop, the identity of the trooper who conducts the stop is unrelated to other determinants of search and search outcomes. In practice, cross-trooper variation arises from week-to-week variation in trooper shift schedules and within-shift variation in trooper locations. Our primary analysis relies on between-trooper variation within assigned patrol areas ("sergeant areas") to isolate variation in search rates conditional on location. We will bolster the argument that we are identifying how different troopers treat equivalent stops by showing that our SPC estimates are robust to varying the set of included controls and troopers and are corroborated by two alternative research designs that rely on different identifying assumptions.

IV.A. Conceptualizing the between-Trooper Search Productivity Curve

For each stop, let i denote the motorist and t denote the specific time. The functions $\ell(i,t)$ and $\tau(t)$ map each stop to its associated location and time category. Let P_ℓ denote the set of troopers working in location ℓ. We limit the analysis to trooper-by-location combinations where the trooper has conducted stops in that location for each time category. For each stop, the associated trooper must decide whether to conduct a search. Let $\text{SEARCH}_{itp} \in \{0,1\}$ denote the potential (search) outcome of the stop, which indicates whether trooper $p \in P_{\ell(i,t)}$ would conduct a search if they were conducting stop (i,t). Let $G_{it} \in \{0,1\}$ indicate whether the motorist is carrying contraband at the time of the stop. Hence, trooper p would find contraband in stop (i,t) if $G_{it} \times \text{SEARCH}_{itp} \equiv \text{CONTRABAND}_{itp} = 1$.[13] Finally, the function $p(i,t)$ maps a stop (i,t) to the trooper who conducts the stop in practice.

[13]Note that we assume no variation across troopers in their ability to identify contraband during a search.

For every trooper $p \in P_\ell$, we can define what their search rate and unconditional hit rate would be if they conducted all of the searches conducted in location ℓ:

$$\sigma_{p\ell} \equiv E[\text{SEARCH}_{itp}|\ell(i,t) = \ell], \tag{1}$$

$$\eta_{p\ell} \equiv E[G_{it}\text{SEARCH}_{itp}|\ell(i,t) = \ell]. \tag{2}$$

We refer to these objects as search *propensities* and unconditional hit *propensities*.

We define our between-trooper SPC as

$$\tilde{\eta}^{\text{BT}}(\sigma) \equiv E_\ell[E_p[\eta_{p\ell}|p \text{ s.t. } \sigma_{p\ell} = \sigma]]. \tag{3}$$

In words, the between-trooper SPC is the relationship between trooper search propensities and unconditional hit propensities across troopers within a location, averaging across locations.

We are also interested in the between-trooper SPC for specific racial groups of motorists (black, Hispanic, and white). Let $r(i)$ indicate the race of motorist i, where $r \in \{b, h, w\}$. We define $\sigma^r_{p\ell}$ and $\eta^r_{p\ell}$ analogously as a trooper's search propensity and unconditional hit propensity for motorists from group r and $\tilde{\eta}^{\text{BT}}_r$ as the between-trooper SPC for motorists from group r.

In practice, we do not observe search and unconditional hit propensities. Instead, for trooper p, we only observe the stop outcomes for stops conducted by trooper p in practice. To recover propensities, we rely on the following conditional independence assumption:

Conditional Independence (CI) Assumption. *Conditional on location $\ell(i,t)$ and time category $\tau(t)$, the race $r(i)$, guilt G_{it}, and potential search decisions $\{SEARCH_{itp}\}_{p \in P_{\ell(i,t)}}$ are independent of the trooper associated with the stop $p(i,t)$.*

We assess the plausibility of this assumption in Section IV.C. Let $S^r_{p\ell\tau}$ denote the set of stops conducted by trooper p in location ℓ at time category t of motorists from group r. Under this assumption, we can construct estimates for $\sigma^r_{p\ell}$ and $\eta^r_{p\ell}$ using the following weighted averages of observed trooper search rates and unconditional hit rates, $s^r_{p\ell}$ and $h^r_{p\ell}$:

$$s^r_{p\ell} \equiv \sum_\tau \underbrace{\left(\frac{1}{|S^r_{p\ell\tau}|} \sum_{(i,t) \in S^r_{p\ell\tau}} \text{SEARCH}_{it} \right)}_{p-\ell-\tau-r \text{ search rate}} \times \underbrace{\left(\frac{|\{(i,t)|\ell(i,t) = \ell; \tau(t) = \tau; r(i) = r\}|}{|\{(i,t)|\ell(i,t) = \ell; r(i) = r\}|} \right)}_{\tau \text{ share for } \ell - r}, \tag{4}$$

$$h^r_{p\ell} \equiv \sum_\tau \underbrace{\left(\frac{1}{|S^r_{p\ell\tau}|} \sum_{(i,t) \in S^r_{p\ell\tau}} \text{CONTRABAND}_{it} \right)}_{p-\ell-\tau-r \text{ unconditional hit rate}} \times \underbrace{\left(\frac{|\{(i,t)|\ell(i,t) = \ell; \tau(t) = \tau; r(i) = r\}|}{|\{(i,t)|\ell(i,t) = \ell; r(i) = r\}|} \right)}_{\tau \text{ share for } \ell - r}. \tag{5}$$

IV.B. Measuring between-Trooper Variation in Search Propensities

We begin by documenting substantial between-trooper variation in search rates among stops with similar locations and times. The notion of location we use at baseline is the sergeant area. The Texas Highway Patrol Division defines six primary regions, which encompass a total of 21 districts and 157 sergeant areas. Outside of the state's most populous areas, sergeant areas typically cover one to two counties in their entirety. In contrast, there are multiple sergeant areas associated with each of the state's most populous counties. Our data identify the exact location of the stop, and we geocode the corresponding sergeant area using boundary shape files received in response to a Texas Public Information Act request.

We infer the sergeant area to which each trooper is assigned based on the trooper-specific distribution of stop locations. In a given calendar year (month), troopers conduct 85% (90%) of all stops in the same sergeant area, on average. Assigning troopers to the modal sergeant area in which they conduct stops in each year, we observe that roughly two-thirds of troopers are assigned to the same sergeant area during the entirety of the sample period. Within a given year, however, temporary reassignments are quite common. When sergeant area assignments are determined on a monthly basis, we calculate that over 70% of troopers experience a change in assignment at least once during the sample period. Roughly half of these reassignments are within the trooper's home region, with the remaining reassignments disproportionately concentrated in the set of districts adjacent to the US-Mexico border (the redeployment of troopers to the border region is discussed in more detail in Section V.D).

The notion of time we use at baseline is the combination of quarter of day and whether the stop was conducted on a weekday or the weekend.

We apply additional sample restrictions that limit the analysis to troopers who have made a sufficient number of stops in a given location. For our pooled analysis, which pools motorists from all racial groups, we limit the analysis to trooper-by-location and time cells with at least 5 stops. We further limit to trooper and location cells with at least 100 stops. Panel A of Online Appendix Figure B1 depicts the number troopers meeting these criteria in each sergeant area. Finally, we limit to locations with at least 10 troopers meeting these criteria, leaving us with 1,951 troopers in 133 locations with 2,657 combinations of trooper and location accounting for 71% of stops. There are an average of 1,235 stops per trooper and location combination.

Online Appendix Figure B2 plots the distribution of $s_{p\ell}$ where each trooper-by-location combination is weighted equally. While the median trooper-by-location search rate is only 0.7%, there is a long right tail, indicating that a small number of troopers search at particularly high rates. The 90th percentile search rate is 3.4%.

We next look at search rates separately by motorist race. We denote the race-specific search rates by s^r_{pl} where $r \in \{w, b, h\}$. When examining race-specific search behavior, we apply different sample criteria to ensure that the troopers we include have made a sufficient number of stops for a specific motorist group. We limit the analysis to trooper-by-location-by-motorist-race cells with least 100 stops. We then limit to locations where, for each motorist racial group, there are at least

5 troopers meeting the sample criteria, leaving us with 1,084 troopers in 77 locations accounting for 53% of black motorist stops, 29% of Hispanic motorist stops, 45% of white motorist stops, and 40% of stops overall. The sample includes 736 troopers for black motorists, 991 for Hispanic motorists, and 1,080 for white motorists. For the sergeant areas we include in the race-specific analysis, Panel B of Online Appendix Figure B1 depicts the number of troopers who satisfy the sample criteria, averaging across motorist racial groups. There are an average of 350, 544, and 1,163 stops per trooper for black, Hispanic, and white motorists.[14]

Mirroring racial differences in overall search rates, between-trooper variation in search rates is larger for non-white motorists. For white, black, and Hispanic motorists, the difference in search rates between the 10th and 90th percentiles of the trooper distribution is 2.8, 6.2, and 4.2 percentage points.

Our goal is to identify variation in how different troopers treat equivalent stops. However, for a fixed sergeant area and time category, the composition of stopped motorists may still vary across troopers. Sergeant area and time category may not fully capture variation in stop context. Even in the same environment, troopers may vary in the composition of motorists they decide to stop. For example, troopers may vary in whether they racially profile when deciding whom to stop for speeding (Horrace and Rohlin, 2016). We examine how search rates change when we condition on a larger set of stop and motorist characteristics. The objective is to learn whether a significant portion of the variation in search rates is due to differences in motorist composition and to isolate variation due to differences in trooper search behavior holding motorist composition fixed.[15]

To calculate search rates that adjust for differences in stop and motorist characteristics, we estimate the following linear probability model, separately by location:

$$\text{SEARCH}_{it} = \phi_{p(i,t)\ell(i,t)\tau(t)} + X_{it}\gamma + \delta_{m(t)} + \rho_{r(i,t)} + \epsilon_{it}, \tag{6}$$

where $\phi_{p(i,t)\ell(i,t)\tau(t)}$ are fixed effects for trooper-by-location and time combinations, $\delta_{m(t)}$ are fixed effects for the month of the stop, and $\rho_{r(i,t)}$ are fixed effects for the (highway) road of the stop. X_{it} is a vector of motorist characteristics, including race, gender, log of neighborhood median income, vehicle-based expected log neighborhood income, stop history, non-drug arrest history, and drug arrest history. We use this model to calculate search rates for each trooper-by-location-by-time combination, adjusting for motorist characteristics and stop month. We use these predicted search rates to construct an overall search rate for a trooper-by-location combination using the same weights as above. We denote this adjusted trooper search rate as $\tilde{s}_{p\ell}$.[16]

[14] Descriptive statistics for the stops included in the pooled and race-specific analyses are presented in Online Appendix Tables B3 and B4, respectively.

[15] Note that we take potential bias in whom troopers decide to stop in the first place (see Knox et al., 2020) as given. For example, troopers may use different criteria for different motorist racial groups when deciding which motorists to search. We consider variation in trooper search behavior, holding the composition of stopped motorists fixed.

[16] Formally, $\tilde{s}_{p\ell}$ is given by

$$\tilde{s}_{p\ell} = \sum_{\tau} \left(\hat{\phi}_{p\ell\tau} + E[X_{it}\hat{\gamma} + \hat{\delta}_{m(t)} + \hat{\rho}_{r(i,t)} | \ell(i,t) = \ell; \tau(t) = \tau] \right) P(\tau(t) = \tau | \ell(i,t) = \ell).$$

Online Appendix Figure B3 compares adjusted trooper search rates ($\tilde{s}_{p\ell}$) to unadjusted trooper search rates ($s_{p\ell}$) across trooper-by-location combinations after partialling out location fixed effects. The slope of the fitted line is 0.97, and the correlation is 0.99. Observable motorist characteristics explain virtually none of the variation in search rates across troopers. Instead, the variation is attributable to differences in trooper search behavior for observably similar stops. This finding provides support for our interpretation of trooper search rates as characterizing causal trooper search propensities.

IV.C. Trooper Search Rates and Motorist Characteristics

To further probe the CI assumption, we investigate the degree to which troopers with high and low search rates stop different types of motorists. Specifically, we examine how motorist characteristics predict the search rate of the trooper conducting the stop. As a benchmark, we estimate the analogous relationship between the same motorist characteristics and whether the stop leads to a search. In particular, we estimate linear regression models of the form

$$Y_{it} = \lambda_{\ell(i,t)} + \omega_{\tau(t)} + \delta_{m(t)} + X_{it}\gamma + \epsilon_{it}, \tag{7}$$

where Y_{it} is either SEARCH_{it} or leave-out trooper-by-location search rates, unadjusted ($s_{p\ell}^{-it}$) or adjusted ($\tilde{s}_{p\ell}^{-it}$). $\lambda_{\ell(i,t)}$ and $\omega_{\tau(t)}$ are location and time category fixed effects.

Table II shows that motorist characteristics predict trooper-by-location search rates, but the magnitude of the relationship is small.[17] Column (1) uses a linear probability model to examine how motorist characteristics predict whether a stop leads to search. Columns (2) and (3) use identical specifications to assess the extent to which these same motorist characteristics predict the unadjusted ($s_{p\ell}^{-it}$) and adjusted leave-out search rate ($\tilde{s}_{p\ell}^{-it}$) of the trooper conducting the stop. Where the outcome is the trooper search rate, the coefficients on all motorist characteristics are one to two orders of magnitude smaller. The relative magnitudes are particularly small for stop and arrest history, which would be difficult for troopers to observe before conducting the stop. Given the large size of our sample, many of these coefficients are statistically significant. In Online Appendix B we show the relationship between motorist characteristics and trooper search rates is similar for stops made at night, when observing the motorist prior to the stop is particularly difficult (Grogger and Ridgeway, 2006). We also construct a trooper stop rate measure that is defined as the average time between stops (including stops for speeding and all other violations) across all sequential pairs of within-shift stops. We show that trooper cited speeds and this stop rate measure are similarly unrelated to motorist characteristics. In Section IV.D we show that our SPC estimates are quantitatively similar whether or not we control for motorist characteristics directly. We also corroborate our baseline SPC estimates using two alternative research designs that rely on different identifying assumptions.

[17] Online Appendix Table B5 presents analogous estimates for trooper-by-location unconditional hit rates. The findings are similar.

[Table 2 about here.]

As a robustness check, we repeat our main analyses after excluding troopers with the most selected set of stopped motorists. For varying κ, we remove the κ% of troopers with compositions of stopped motorists who deviate most from their expected composition given the time and location of their stops. We discuss how we identify these troopers in more detail in Online Appendix B. Columns (4)–(6) replicate (1)–(3) after removing stops conducted by the 20% of troopers with the most selected set of stopped motorists. In columns (5) and (6), which relate trooper search rates to stopped motorist characteristics, removing these troopers from the analysis further reduces the magnitude of coefficients on motorist characteristics. All results presented in the remainder of the paper are insensitive to excluding troopers with the most selected set of stopped motorists.

IV.D. Baseline Search Productivity Curve Estimates

For each location, we observe multiple troopers with varying search rates. We next calculate and compare unconditional hit rates across troopers within a location. We calculate trooper-by-location-specific unconditional hit rates analogous to the search rates constructed above, replacing the outcome with CONTRABAND_{it}. We denote the unadjusted and adjusted trooper-by-location-specific unconditional hit rates as $h_{p\ell}$ and $\tilde{h}_{p\ell}$. Each trooper demonstrates a feasible combination of search rate and unconditional hit rate, which we use to construct a between-trooper SPC. We then pool these location-specific SPCs across locations to construct an aggregate SPC.

Note that trooper-by-location-specific search rates and unconditional hit rates are, in principle, only directly comparable across troopers within a location. Hence, we aggregate location-specific SPCs without relying on between-location comparisons (the 'quantile approach'). Within locations we divide troopers into quantiles by search rate, group quantiles across locations, and then plot the relationship between search rates and unconditional hit rates across quantiles. The interpretation of the slope of this relationship is the change in the aggregate unconditional hit rate associated with a change in the search rate quantile for all locations. We use deciles for the pooled analysis because each location has at least 10 troopers by construction. Similarly, we use quintiles for the race-specific analysis.[18]

Figure III summarizes the relationship between adjusted trooper search rates ($\tilde{s}_{p\ell}$) and unconditional hit rates ($\tilde{h}_{p\ell}$) using two approaches: a simple binscatter (Cattaneo et al., 2019) and the quantile approach described above. In both approaches, we weight trooper-by-location combinations by number of stops. From the binscatter approach, this figure includes 95% confidence bands for the local linear relationship between adjusted trooper search rates and unconditional hit rates, the best linear fit in the depicted range, and the slope of the best linear fit, labeled as β. From the quantile approach, the figure includes the mean values for each decile, the best linear fit, and the

[18] As an alternative approach to aggregation, we plot the relationship between search rates and unconditional hit rates while adjusting for location fixed effects using the method of Cattaneo et al. (2019). This approach generates similar results, which can be found in Online Appendix B.

slope of the best linear fit, labeled as β^{Q}. Bootstrap standard errors for the estimated slopes are provided in parentheses.

[Figure 3 about here.]

There are two main findings to note. First, for both approaches the SPC is approximately linear. For the binscatter approach, we conduct a formal test for linearity following Cattaneo et al. (2019) and fail to reject the null hypothesis that the relationship is linear. Second, the two approaches provide similar SPC estimates: β is 0.343 and β^{Q} is 0.342. These slopes indicate that a 1 percentage point increase in search rate is associated with about a 0.34 percentage point increase in the unconditional hit rate. Note that the ratio of the unconditional hit rate to the search rate gives us the percentage of searches that yield contraband, the hit rate. The fact that the relationship between the unconditional hit rate and the search rate is linear implies that hit rates are approximately **constant** across quantiles of trooper search rates and hence are unrelated to trooper search rates.

In Figure IV we repeat the analysis separately by motorist racial group. Estimates in Panel A are derived using only white motorists, Panel B using only black motorists, and Panel C using only Hispanic motorists. There are two main findings to note. First, as above, each SPC is approximately linear. Second, the slopes for white and black motorists are comparable, while the slope for Hispanic motorists is smallest in magnitude. Using either approach, we cannot reject the null hypothesis of equal slopes for white and black motorists, while the slope for Hispanic motorists is significantly smaller than the slope for white motorists at the 1% level. Using the binscatter approach, the estimated SPC slopes for white, black, and Hispanic motorists are 0.396, 0.352, and 0.298. These slope estimates are comparable to overall group-specific hit rates described in Online Appendix Table B4. Using the quantile approach, the estimated SPC slopes for white, black, and Hispanic motorists are 0.385, 0.389, and 0.305.

[Figure 4 about here.]

IV.E. Search Productivity Curve Estimates: Robustness Checks

The CI assumption underlying our approach is violated if, conditional on the sergeant area and time category of a stop, troopers vary in the composition of motorists they stop. This variation may exist because our location and time measures are not sufficiently granular or because, in the same environment, troopers vary in the motorists they decide to stop. In this section we present several robustness checks for the baseline results presented in Section IV.D.

In Section IV.E.1 we reestimate pooled and race-specific SPCs using only within-motorist variation in stop outcomes among motorists involved in multiple speeding stops. In Section IV.E.2 we compare stop outcomes on opposite sides of trooper patrol area borders in a spatial RD design to validate trooper search rates as estimates of causal trooper search propensities. In both sections we conduct additional tests for whether between-trooper SPCs are linear.

In Online Appendix Figure B7 we show that the slope of the pooled between-trooper SPC is stable if we exclude a varying proportion of troopers with compositions of stopped motorists who deviate most from their expected composition given the time and location of their stops. In Online Appendix Figure B8 we conduct a similar exercise for race-specific SPC slopes and find that slope estimates and their ordering across groups are stable when we vary the set of included troopers. In Online Appendix Figure B10 we show that SPC estimates are similar if we restrict the sample to stops conducted at night (Grogger and Ridgeway, 2006).

Another concern with our approach is that $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$, as estimates of their population analogs, $\sigma_{p\ell}$ and $\eta_{p\ell}$, are subject to correlated sampling error. This sampling error may bias our estimate of β. We account for sampling error in two ways in Online Appendix B. First, we apply empirical Bayes adjustments to $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ (Morris, 1983). Second, we take a split-sample IV approach to estimation. We randomly split stops into two samples and estimate $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ separately in each sample. In each sample, we regress $\tilde{h}_{p\ell}$ on $\tilde{s}_{p\ell}$ and location fixed effects, instrumenting for $\tilde{s}_{p\ell}$ using its pair estimate from the other sample. Reassuringly, both approaches yield β estimates that are statistically indistinguishable from the OLS estimates.

Finally, we verify that the key features of the pooled and race-specific SPCs are not sensitive to the particular hit rate definition employed. In our main specifications, we measure hit rates using an indicator for whether the trooper finds any contraband as recorded in the traffic stop data. This measure may mask heterogeneity in the significance of the contraband discovered across stops. In Online Appendix Figures B11 and B12 we replicate Figure III and Figure IV using two alternative outcomes: (1) an indicator for whether the contraband found leads to an arrest and (2) the average incarceration sentence associated with conviction for those arrest charges, where the latter outcome is set to zero for stops that do not lead to an arrest. We interpret the latter outcome as a measure of the severity of any associated arrest charges. The patterns are qualitatively similar. The SPCs are approximately linear, the slopes for white and black motorists are comparable, and the slope for Hispanic motorists is smaller in magnitude.

IV.E.1. Search Productivity Curve Estimates: Exploiting within-Motorist Variation

High and low search rate troopers may stop motorists who differ on unobservables correlated with contraband risk. To address this concern, we take advantage of the fact that we can match multiple stops to the same motorist. We look at sequential pairs of stops for the same motorist and measure the relationship between differences in stop outcomes and differences in the search behavior of the troopers conducting those stops. By looking at differences in stop outcomes for the same motorist, we net out time-invariant motorist-level determinants of search and contraband risk.

Consider a group of motorists stopped by two sets of troopers, one with high search costs and the other with low search costs. Suppose that the distribution of screening skill across troopers and the probability that a given motorist is carrying contraband are equal across sets. With diminishing returns to search, we expect the hit rate to be lower for the low search cost (and hence high search rate) set. Moreover, we expect this difference in hit rates to be increasing in the difference in search rates between the trooper sets. By contrast, if troopers are searching on the linear portion of the SPC, we expect constant hit rates.

To implement this idea, we first group sequential pairs of stops of the same motorist into deciles by their difference in trooper-by-location search rates, $\Delta_{it}\tilde{s}_{p\ell} = \tilde{s}_{p(i,t)\ell(i,t)} - \tilde{s}_{p(i,t')\ell(i,t')}$, where $t > t'$ and stops at t' and t are sequential for motorist i. Descriptive statistics for the sequential pairs of stops are presented in Online Appendix Table B7 and Online Appendix Table B8.

Panel A of Figure V summarizes the characteristics of motorists involved in sequential pairs of stops, grouped into deciles on the horizontal axis based on their value of $\Delta_{it}\tilde{s}_{p\ell}$. The vertical axis depicts the average value of $P(SEARCH|X_{it})$, the predicted search rate for a motorist given their characteristics at the time of the initial stop.[19] The figure includes the best linear fit and a bin scatter. The measured relationship is flat. Motorists stopped by different troopers, as characterized by their search rates, do not markedly vary in their characteristics.

[Figure 5 about here.]

In Panel B of Figure V we plot the relationship between $\Delta_{it}\tilde{s}_{p\ell}$ and Δ_{it}SEARCH. The relationship is linear with a slope of 0.733. If we suppose that Δ_{it}SEARCH is an unbiased measure of differences in trooper search propensities for motorist i, then the fact that this slope is less than one indicates that $\Delta_{it}\tilde{s}_{p\ell}$ exhibits forecast bias. This could be explained by one or a combination of two factors: (1) $\tilde{s}_{p\ell}$ is a biased estimate of $\sigma_{p\ell}$ or (2) trooper search propensities vary with motorist characteristics, and motorists stopped multiple times differ from typical motorists stopped in either location.[20] The overall search rate in the pooled SPC sample is 1.11% while the search rate for the subset of stops analyzed here is 0.93%.

Panel C of Figure V plots the relationship between $\Delta_{it}\tilde{s}_{p\ell}$ and Δ_{it}CONTRABAND by decile. Again, the relationship is strikingly linear. If marginal motorists—motorists more likely to be searched by high search rate troopers—are less likely to carry contraband, we would expect the slope to be declining in $\Delta_{it}\tilde{s}_{p\ell}$. Instead, linearity is consistent with marginal motorists who are as likely to carry contraband as inframarginal motorists.

The patterns in Panels B and C of Figure V imply a particular relationship between a given increase in search rates and an increase in the unconditional hit rate if we frame $\Delta_{it}\tilde{s}_{p\ell}$ as an instrument for Δ_{it}SEARCH. The slope we estimate is essentially the slope of an SPC in first differences. More concretely, we estimate the following model via just-identified two-stage least

[19]We construct $P(SEARCH|X_{it})$ using the logistic regression model

$$P(\text{SEARCH}_{it} = 1|X_{it}) = \frac{e^{(X_{it}\beta)}}{1 + e^{(X_{it}\beta)}}$$

where X_{it} is a vector of motorist characteristics including motorist race, gender, log of neighborhood income, expected log income given vehicle, stop history, non-drug arrest history, and drug arrest history.

[20]A third possibility is that $\tilde{s}_{p\ell}$ differs from $\sigma_{p\ell}$ due to sampling error, leading to attenuation bias. However, the fact that empirical Bayes and split sample adjustments of $\tilde{s}_{p\ell}$ explored in Online Appendix B do not make a material difference indicates that sample sizes are sufficiently large for sampling error to not be a substantive issue.

squares (2SLS), separately by motorist race:

$$\Delta_{it}\text{CONTRABAND} = \beta\Delta_{it}\text{SEARCH} + \epsilon_{it}, \tag{8}$$

where the first stage is

$$\Delta_{it}\text{SEARCH} = \pi\Delta_{it}\tilde{s}_{p\ell} + \zeta_{it}. \tag{9}$$

We repeat this exercise pooling all motorists and separately by motorist racial group. When we estimate the model for a specific motorist racial group, we replace $\Delta_{it}\tilde{s}_{p\ell}$ with its race-specific analog, $\Delta_{it}\tilde{s}^{r}_{p\ell}$.[21]

We report β estimates in Table III. For all motorists pooled and each motorist group, the slopes are somewhat smaller than, but comparable to, the baseline SPC slopes. The SPC slope for white motorists exceeds the slopes for black and Hispanic motorists in magnitude. Time-invariant motorist unobservable characteristics cannot explain the pattern of SPC slopes we identify in Section IV.D.

[Table 3 about here.]

IV.E.2. Search Productivity Curve Estimates: Exploiting Sergeant Area Borders

We derive an alternative estimate for the marginal returns to search by comparing outcomes of stops conducted on opposite sides of sergeant area borders. Along the same highway route, the composition of troopers making traffic stops changes sharply across sergeant area borders, which designate the areas that troopers are assigned to patrol. If troopers assigned to one sergeant area search motorists at higher rates than troopers in a neighboring sergeant area, then motorists crossing from one sergeant area to the other will face sharp changes in their chances of being searched. This spatial feature of search rates suggests a natural RD research design. By comparing search and unconditional hit rates for speeding stops on either side of sergeant area borders, we measure the causal effect of changing from one set of troopers to another with higher search rates. This comparison provides another test for whether trooper search rates characterize the causal effect of trooper assignment on search likelihood and another measure of the between-trooper SPC slope.

Our identifying assumption is that the composition of motorists evolves continuously through sergeant area borders. This assumption is reasonable because sergeant area borders are defined only for administrative purposes; there is little reason to think the composition of motorists traveling on a given stretch of highway would change discontinuously at these boundaries. One possible exception is that some motorists are aware that their chances of being subject to a search change at sergeant area borders and adjust their travel or contraband carrying behavior accordingly. In practice, we find little evidence of a deterrence effect at sergeant area borders, a point we discuss in more detail in Section V.D.

[21] Note that the set of stops included in the race-specific analysis is a subset of the stops included in the pooled analysis because $\tilde{s}^{r}_{p\ell}$ is measured for a smaller set of trooper-by-location combinations, as discussed in Section IV.B.

For this exercise, we apply sample restrictions that differ from the sample restrictions described in Section IV.B.[22] There are 202 state and interstate highways crossing 319 distinct sergeant area borders. Each highway by border pair is a potential RD. We limit the RD analysis to highway and border intersections with at least 100 stops made in each corresponding sergeant area between 2 and 7 miles from the intersection. We are left with 424 intersections. These intersections as well as the corresponding highways and sergeant areas are shown in Online Appendix Figure B13.

For each intersection, we limit the analysis to speeding stops made in each corresponding sergeant area within seven miles from the intersection. We use the distance between the location of a stop and the intersection as the running variable. For each intersection, we set the distance as negative for the sergeant areas with the lower average trooper search rate.[23]

While sergeant area borders generate a discrete change in trooper patrol areas, in practice these borders do not provide a discontinuity in where troopers search. Troopers conduct some stops outside of their patrol area, and they are particularly likely to do so just outside sergeant area borders. Figure VI pools all intersections and plots the share of stops conducted by troopers assigned to each adjacent sergeant area by distance from the intersection. More than two miles from the border, the share of stops conducted by troopers assigned to that corresponding sergeant area generally exceeds 70%. Approaching the border, this share falls to about 40%.

[Figure 6 about here.]

To add substantial statistical power to our test, we take a "donut" approach and exclude stops that occur within a two-mile window around the intersection, denoted by the dashed vertical lines in Figure VI (Barreca et al., 2016). The trade-off is that, by excluding stops in this range, we can no longer take a non-parametric approach to identification. Instead we assume that, in the absence of contamination near the intersection, potential search outcomes would continue to evolve as they do outside of this range.[24] We use a bandwidth of seven miles, leaving us with 1,480,372 stops conducted between two and seven miles from the border.

Figure VII plots (leave-out) trooper search rates in Panel A and motorist characteristics in Panel B as a function of distance from the intersection. For each RD plot, within each set of stops corresponding to an intersection, we demean the outcome and then stack observations across intersections. Panel A shows that the search rates for troopers conducting stops are approximately constant across distances within each sergeant area. Across the boundary, extrapolated trooper search rates jump by 0.231 percentage points using a constant extrapolation and 0.242 percentage points using a linear extrapolation.

[22] Descriptive statistics for the stops included in this analysis are presented in Online Appendix Table B9.

[23] We measure trooper search rates using all speeding stops conducted by a trooper, not just those made in the RD window or in a specific location.

[24] In contrast with typical applications of the donut RD design, we are not concerned about manipulation or error in the running variable (Barreca et al., 2016). Instead, we apply this approach because the change in treatment—in our case, the search rates of the troopers conducting the stops—at the border is muted by the fact that troopers make some stops just outside their assigned patrol area.

By contrast, as indicated in Panel B, the characteristics of stopped motorists vary only slightly across the threshold. Note that the statistical significance of this discontinuity does not necessarily indicate a discontinuity in the composition of motorists on either side of the border. Instead, it may simply reflect minor differences in trooper stop behavior.

[Figure 7 about here.]

Figure VII plots search rates (Panel C) and unconditional hit rates (Panel D) by stop location. In Panel C the pattern is noisier, but there is again a clear jump in extrapolated search rates at the boundary. The magnitude of the jump is 0.229 percentage points using a constant extrapolation and 0.424 percentage points using a linear extrapolation. In the constant case, the magnitude of the jump is comparable to the magnitude of the corresponding change in trooper search rates. This result can be interpreted as a validation of trooper search rates as measures of causal trooper search propensities. In the linear case, the jump in search rates is larger, though the estimate is relatively imprecise.

The ratio of the increase in the unconditional hit rate to the increase in the search rate is 0.352 (with a standard error of 0.035) in the constant case and 0.334 (0.068) in the linear case.[25] This ratio provides an alternative estimate of the marginal slope of the between-trooper SPC. Both estimates are consistent with our SPC slope estimates of 0.343 and 0.342 in Section IV.D. They are also statistically indistinguishable from the overall hit rate, consistent with a linear between-trooper SPC.

V. Policy Counterfactuals

In this section we consider policy counterfactuals that reduce racial disparities in search rates. We consider two types of counterfactuals. First, we reallocate troopers across sergeant areas. We present these counterfactuals in Section V.A.

Second, we force troopers to equalize their search rates across motorist racial groups. In Section V.B we describe the distinction between the between-trooper SPC and the *within-trooper* SPC and why the latter is central to constructing these counterfactuals. We then provide support for skill-propensity independence and direct evidence that the between-trooper and average within-trooper SPCs are similar. Section V.C presents the counterfactuals.

In Section V.D, we present evidence that motorist group-specific deterrence effects are negligible at the margin, implying that such deterrence effects are unlikely to substantively influence our conclusions about counterfactuals.

V.A. Counterfactuals: Reallocating Troopers Across Locations

The first counterfactual we consider is one where troopers are reallocated across sergeant areas. We ask how much racial disparities in search rates would be reduced if troopers that search at high

[25] We estimate the ratios and associated standard errors by instrumenting for SEARCH_{it} using threshold crossing in a linear regression model for CONTRABAND_{it}.

rates were reallocated to sergeant areas where white motorists make up a large share of stopped motorists and troopers that search at low rates were reallocated to sergeant areas with many black and Hispanic motorists. The key challenge is that we must specify what trooper outcomes would be if troopers were reallocated to different sergeant areas.

We consider two specifications of counterfactual trooper outcomes. First, we assume that troopers have the same race-specific search rates if they are reallocated. Second, we assume that troopers' counterfactual race-specific search rates are a weighted average of their observed search rates and the observed search rates of the sergeant area to which they are reallocated. To choose this weighting, we take the top two locations for each trooper and regress trooper-location search rates for one location on the other while including race fixed effects. This yields a coefficient on the other-location search rate of 0.75 (see Online Appendix Figure B14), the weight on the trooper's observed search rate that we use. We assume that group-specific hit rates are unchanged.

We reallocate troopers as follows. First, we apply the same sample restrictions used in the race-specific analysis described in Section IV.B and further restrict to trooper-location combinations with at least 100 stops for each motorist racial group.[26] We then order trooper-location combinations by their search rate for white motorists. We use the top and bottom X% of trooper-location combinations for reallocation. For each combination, we calculate the black share of stopped motorists (or Hispanic share, or the nonwhite share). We then reallocate the trooper-location combination with the highest search rate to the set of stops corresponding to the trooper-location combination with the lowest black share of stops, the combination with the second highest search rate to the set of stops corresponding to the combination with the second lowest black share of stops, and so on. We repeat this exercise where we reallocate based on the Hispanic share of stops or the nonwhite share of stops.

The results are reported in Table IV. The top panel reports results where we assume observed and counterfactual search rates are the same. In the bottom panel we adjust counterfactual search rates to depend partially on the allocated sergeant area, as described above. Column (1) reports observed search rates for the status quo trooper-location combinations. In columns (2)-(4) we reallocate troopers to minimize black search rates. If we reassign 20% of troopers, we can reduce black search rates from 2.16% to 1.85% using observed search rates and to 1.89% using reweighted search rates. If we reassign all troopers, we can reduce black search rates further to 1.37% and 1.61%. The Hispanic search rate is roughly unchanged with these reallocations.

[Table 4 about here.]

In columns (5)-(7) we reallocate to minimize Hispanic search rates. If we reassign 20% of troopers, we can reduce Hispanic search rates from 1.37% to 1.22% using either unadjusted or adjusted search rates. If we reassign all troopers, we can reduce Hispanic unadjusted and adjusted search rates further to 0.98% and 1.09%. In columns (8)-(10) we reallocate to minimize nonwhite

[26] Descriptive statistics for the stops included in this reallocation exercise are presented in Online Appendix Table B10.

search rates combined. This leads to smaller reductions in black search rates, but similar reductions in Hispanic search rates. At baseline, the search rate for black motorists is 167% higher than the search rate for white motorists. Using unadjusted search rates, reallocating all troopers to minimize the nonwhite search rate reduces this proportional gap to 60%, a decline of 64%. The Hispanic-white gap in search rates is essentially eliminated. Moreover, reallocations that trade searches of Hispanic motorists for searches of black and white motorists increase contraband yield.

V.B. Skill-Propensity Independence

We have documented the between-trooper relationship between search rates and unconditional hit rates. We find that this relationship is linear with racial group-specific slopes which suggest it is feasible for troopers to (1) search all motorist racial groups at the same rate, (2) maintain the status quo overall search rate, and (3) increase overall contraband yield. This counterfactual requires that individual troopers change their search behavior, so determining what would happen in this scenario requires knowing the SPCs faced by individual troopers. Yet, the *between-trooper* SPC need not be the same as *within-trooper* SPCs. In particular, a linear between-trooper SPC may still be consistent with troopers facing diminishing returns to search if troopers with more screening skill search at higher rates.[27]

Two recent papers closely related to ours, Arnold et al. (2020) and Chan et al. (2020), face a similar issue. Those papers identify cross-sectional variation in behavior across bail judges and doctors, respectively, and then conduct counterfactual exercises where they consider what would happen if decision-makers were made to change their behavior in some way. To infer how judge and doctor outcomes would change under counterfactuals, Arnold et al. (2020) and Chan et al. (2020) make parametric assumptions about the form of heterogeneity across agents and then use features of the cross-sectional distribution of agent behavior to identify parameters that characterize this heterogeneity.

By contrast, we argue that the between-trooper and average within-trooper SPCs are similar, consistent with skill-propensity independence. We document three pieces of supporting evidence. First, we find that low and high search rate troopers search motorists with similar observable characteristics; high search rate troopers just search more often. This suggests that low and high search rate troopers are applying a similar, coarse screening of motorists. Second, we find that observable **trooper** characteristics—including trooper experience, stop rates, and race—are unrelated to trooper hit rates. Moreover, between-trooper SPCs do not materially vary with observable trooper characteristics. Hence, skill-propensity independence holds approximately between observ-

[27]As we discuss in Section III.B, the distinction between the between-trooper SPC and within-trooper SPCs is related to the monotonicity conditions of Arnold et al. (2018), Chan et al. (2020), and Frandsen et al. (2020). Under the strict monotonicity condition of Arnold et al. (2018), each trooper faces the same SPC, which also corresponds to the between-trooper SPC. We document in Appendix B that strict monotonicity does not hold in our setting; in particular, we find systematic variation in hit rates between troopers. However, the between-trooper SPC and average within-trooper SPC are equal under a substantively weaker condition: independence of trooper screening ability and search propensity. This condition corresponds to the skill-propensity independence condition in Chan et al. (2020) and is implied by the average monotonicity condition of Frandsen et al. (2020).

able trooper groups. Third, we estimate the slope of the within-trooper SPC directly, using the fact that search rates vary systematically across sergeant areas, even for the same trooper and motorist. We find that the estimated slopes of the between-trooper and within-trooper SPCs are similar.

We begin by comparing the determinants of search for low and high search rate troopers. A linear SPC indicates that troopers are coarsely screening motorists, identifying a segment that are "at risk" of search, and searching at risk motorists with some probability that varies across troopers. This suggests that low and high search rate troopers are searching similar motorists. However, the linear between-trooper SPC could be masking trooper heterogeneity, in which case there is no reason to expect low and high search rate troopers to be searching observably similar motorists. Indeed, if troopers faced declining returns to search at the margin, low and high search rate troopers would seem likely to search observably different motorists.[28]

To test whether low and high search rate troopers search observably similar motorists, we estimate the following logistic regression models, separately by trooper search rate quartile:

$$P(\text{SEARCH}_{it} = 1 | \ell(i,t), \tau(t), X_{it}) = \frac{e^{(\lambda_\ell + \omega_\tau + X_{it}\gamma)}}{1 + e^{(\lambda_\ell + \omega_\tau + X_{it}\gamma)}}, \tag{10}$$

where λ_ℓ and ω_τ are fixed effects for stop location and time category; and X_{it} is a vector of motorist characteristics, including race, gender, log of neighborhood median income, expected log neighborhood median income given vehicle type, stop history, non-drug arrest history, and drug arrest history. The results are presented in Online Appendix Table B11. Across quartiles, odds ratios are nearly identical. To summarize the similarity across models, we calculate the predicted probability of search for each stop and for each model, and correlate predicted values across models. The correlations range from 0.95 to 0.99. High search rate troopers appear to search the same types of motorists as low search rate troopers, they just search more often.

Next, we test whether observable trooper characteristics correlate with both trooper search rates and hit rates. Skill-propensity independence implies that, if trooper skill varies with observable trooper characteristics, that variation is unrelated to search rates. We examine three trooper characteristics: trooper experience measured in years; trooper average time between within-shift stops; and trooper race. We identify trooper experience and race using 2015 personnel records for 2,469 troopers accounting for 84% of stops.[29] We regress adjusted trooper search rates, $\tilde{s}_{p\ell}$, on trooper characteristics and adjusted unconditional hit rates, $\tilde{h}_{p\ell}$, on both trooper characteristics and search rates. The results are shown in Online Appendix Table B13. We find that trooper search rates are increasing in time between stops and lower for black and Hispanic troopers. Yet, conditional on search rate, unconditional hit rates are unrelated to observable trooper characteristics. In Online Appendix Figure B17 we construct and compare between-trooper SPCs estimated using trooper subgroups. Between-trooper SPCs are similar across trooper experience quartiles,

[28]For example, as we show in Appendix Section B.7, the search behavior of troopers in the top quartile of search rates indicates that those troopers could achieve a 23% higher hit rate at the average search rate simply by forgoing searches of observable motorist groups with the least productive searches.

[29]We limit the analysis to black, Hispanic, and white troopers, which account for 98% of matched troopers.

stop rate quartiles, and racial groups.

Third, we estimate the slope of the average within-trooper SPC directly. To estimate the average within-trooper SPC, we require an instrument for search that shifts a troopers' effective search costs, yet is orthogonal to motorist composition and trooper screening ability. We use the sergeant area for the stop as an instrument for search in a model with both trooper **and** motorist fixed effects. Location may influence the effective costs of search through the manager (sergeant) or peers associated with a location, or through staffing levels, which may alter the opportunity cost of time-intensive searches. We hold the composition of motorists fixed by including motorist fixed effects. The key threats to the exclusion restriction are that: (1) for the same motorist, contraband carrying behavior varies across locations; (2) trooper screening ability varies across locations.

Let $s^{-(i,t)}_{\ell(i,t)y(t)}$ denote the search rate for all stops in location ℓ in the year corresponding to t, excluding stop (i,t). We estimate a 2SLS system with first stage

$$\text{SEARCH}_{it} = \pi s^{-(i,t)}_{\ell(i,t)y(t)} + \mu_i + \phi_{p(i,t)} + \delta_{y(t)} + \epsilon_{it}, \tag{11}$$

where μ_i are motorist fixed effects, $\phi_{p(i,t)}$ are trooper fixed effects, and $\delta_{y(t)}$ are year fixed effects. The second stage is

$$\text{CONTRABAND}_{it} = \beta\text{SEARCH}_{it} + \mu_i + \phi_{p(i,t)} + \delta_{y(t)} + \zeta_{it}. \tag{12}$$

Table V correlates the instrument with observable motorist characteristics and presents first stage and 2SLS estimates. In column (1) we estimate (11) but with $P(\text{SEARCH}_{it}|X_{it})$ as the outcome. The estimated coefficient on $s_{\ell(i,t)y(t)}$ is 0.032, indicating that a 1 percentage point increase in leave-out search rates is associated with a 0.03 percentage point increase in $P(\text{SEARCH}_{it}|X_{it})$.[30] This coefficient estimate is statistically significant but is small in magnitude. Columns (2) and (3) present first stage estimates with and without controls for observable motorist characteristics. The coefficient on $s_{\ell(i,t)y(t)}$ is 0.422 in column (2), indicating that a 1 percentage point increase in coworker search rates is associated with a 0.42 percentage increase in a trooper's own search rate. Consistent with the limited relationship between coworker search rates and motorist characteristics shown in column (1), adding controls for motorist characteristics in column (3) only slightly attenuates this coefficient to 0.394. Column (4) replaces controls for observable motorist characteristics with motorist fixed effects. Including motorist fixed effects reduces the coefficient on $s_{\ell(i,t)y(t)}$ to 0.213, but the first stage remains highly statistically significant, with a t-statistic of nearly 7.

Columns (5) through (7) present IV estimates for β. Without controlling directly for motorist characteristics, the 2SLS estimate for β of 0.374 with a standard error of 0.033, in line with our between-trooper estimates. Adding motorist controls X_{it} has no material effect on this estimate. Adding motorist fixed effects (column (7)) increases the coefficient to 0.391 and standard error to 0.058. In Online Appendix Figure B15 we show that both the first stage and reduced form relation-

[30] For reference, the standard deviation of $s_{\ell(i,t)y(t)}$ after partialling out trooper and year fixed effects is 0.004, or 0.4 percentage points. The standard deviation of $s_{\ell(i,t)y(t)}$ after partialling out trooper, year, and motorist fixed effects is 0.003.

ships are approximately linear. We obtain similar, though less precise estimates if we instrument for search using $s^{-p(i,t)}_{\ell(i,t)y(t)}$, the search rate for all stops in location ℓ in the year corresponding to t, excluding **trooper** $p(i,t)$ (see Online Appendix Table B14).

[Table 5 about here.]

Overall, these results suggest that the marginal slope of the within-trooper SPC, averaged across troopers, is similar to the between-trooper SPC slope.[31]

Moreover, as we show in Online Appendix C, if the marginal hit rate averaged across troopers—which our within-trooper design is intended to identify—is equal to the (average) hit rate averaged across troopers, then trooper-specific SPCs are linear at least between a search rate of 0 and the trooper's observed search rate. In other words, for all troopers with search rates equal to or above some value σ, each trooper's SPC is linear between 0 and σ. We cannot reject the null hypothesis that this condition is satisfied in our context. Moreover, given that the between-trooper SPC is linear, it follows that for troopers with search rates equal to or above some value σ, the average within-trooper SPC is equal to the between-trooper SPC below σ.

This is a useful result for our counterfactuals. In particular, equalizing search rates requires a significant reduction in search rates for black motorists, a moderate reduction for Hispanic motorists, and a moderate increase in search rates for white motorists. We can rely on the above linearity argument to credibly construct counterfactual hit rates for black and Hispanic motorists. For constructing counterfactual hit rates for white motorists, we can rely on the fact that the prescribed increase in search rates for white motorists is a reasonable extrapolation from the IV-induced variation in search rates.

V.C. Counterfactuals: Equalizing Trooper Search Rates

Based on the group-specific SPCs we estimate, it is straightforward to assess whether search rates can be equalized across motorist racial groups while maintaining search efficiency at the status quo overall search rate. This corresponds to moving troopers along curves we estimate to fit the relationship between search rates and unconditional hit rates depicted in Figure IV to predict their counterfactual, motorist race-specific hit rates. We construct counterfactual hit rates in this way using both between-trooper SPCs that pool all troopers together, and between-trooper SPCs where we divide troopers into subgroups based on an observable characteristic, either experience, race, or stop rate. For comparability, we limit the analysis to stops conducted by troopers we are able to match to personnel records. While, as we have argued, between-trooper SPCs are well-approximated by a linear function, for this exercise we fit a quadratic curve via constrained OLS using all troopers or subsets of troopers. We set the intercept to zero to match the fact that

[31] A potential concern is that the IV weights are not constant across troopers and the average of trooper hit rates, appropriately weighted, differs substantially from the between-trooper SPC slope. In Online Appendix Table B15 we divide troopers in half by their hit rate on a leave out sample and show that the magnitude of the first stage is similar for troopers with below and above average hit rates. This suggests the IV weights are not a first order concern.

unconditional hit rates are mechanically zero when search rates are zero and constrain the curve to be weakly concave.

Counterfactual estimates are provided in Table VI. To provide a benchmark, column (1) of Table VI summarizes observed search and hit rates, both pooling all motorists and by motorist race. For each SPC specification, pooling all troopers or by some subgroup, we provide two hit rates. The first is the predicted hit rate if each trooper searches motorists by racial group at their status quo rate. The second is the predicted hit rate if each trooper searches each motorist racial group at the pooled search rate in this sample, 1.12%. Bootstrap standard errors are provided in parentheses.

The estimates indicate that equalizing search rates would modestly **increase** search productivity. We reach the same conclusion if we use arrests or charge severity as alternative outcomes (see Online Appendix Tables B16 and B17). Under the assumption that marginal changes in search rates do not influence contraband carrying behavior, this summary finding implies that there is no equity-efficiency trade-off present. In other words, it is feasible for troopers to search all motorist racial groups at the same rate, maintain the status quo overall search rate, and increase overall contraband yield. Identifying how to best incentivize troopers to adjust search rates represents a worthwhile research question in its own right, but one that is outside the scope of this paper.

Relatedly, an important caveat that underpins both sets of policy counterfactuals is that we abstract away from policy-driven behavioral responses. Specifically, in the counterfactual relying on trooper reallocations, we allow for counterfactual search rates to adjust as they do for those troopers that we observe in multiple locations during the sample period. In our second policy counterfactual that requires troopers to adjust their own search rates, we assume that troopers would respond to mandated changes as predicted by our SPC estimates. Given the lack of prior reforms aimed at reallocating troopers or adjusting trooper search rates to improve equity, we cannot assess how policy-induced behavioral responses (for example, driven by trooper resistance to search-related mandates) would be expected to influence our conclusions in practice.

[Table 6 about here.]

V.D. Deterrence Effects

To assess how contraband yield would change under counterfactual search rates, it is also important to gauge the responsiveness of contraband carrying behavior to motorist racial group-specific search intensity. While our analysis is not structured to explicitly characterize motorist racial group-specific deterrence effects, a number of factors suggest that changes in contraband carrying rates are likely to be negligible. First, if drivers respond to the overall search rate rather than to racial group-specific search rates (e.g., if drivers are uninformed regarding racial group-specific changes in search), then equalizing search rates across racial groups while keeping the overall search rate constant will not influence contraband carrying rates. To the extent that drivers do respond to racial group-specific search rate changes, then the relative (racial group-specific) elasticities of contraband carrying with respect to the search rate will determine the net impact of search rate equalization (see Bjerk [2007] for a related discussion). In practice, changes in aggregate contraband carrying rates will be quite limited unless the difference in racial group-specific elasticities is large. Low observed rates of search and the fact that searches can only occur if a motorist is first stopped suggest that these elasticities are likely to be low in the neighborhood of status quo search rates.

To provide support for the assertion that deterrence effects are likely to be negligible on the margin in our setting, we undertake three complementary exercises. The logic behind these exercises is that motorists cannot influence the trooper who conducts a stop within a sergeant area, but they can choose to not drive through a particular sergeant area while carrying contraband. Hence, if higher search rates deter motorists from carrying contraband, we expect motorist behavior to be responsive to variation in search rates between sergeant areas or within sergeant areas over time, but not to within-sergeant area, between-trooper variation in search rates.

In the first exercise, we compare SPC slopes estimated using within-sergeant area variation (as in Section IV.D) to SPC slopes estimated using cross-sergeant area boundary variation (as in Section IV.E.2). With significant deterrence effects, we would expect the SPC slope to be smaller when derived from cross-sergeant area boundary variation because the estimate would reflect the fact that motorists are less likely to carry contraband in high search rate sergeant areas. Instead, we find that the two slope estimates are statistically indistinguishable, suggesting that motorist behavior is not responding to sergeant area search rates.

In the second exercise, we compare the SPC slopes estimated using within-motorist variation (as in Section IV.E.1) among two sets of sequential pairs of stops. In one set, both stops are in the same sergeant area. In the other set, the two stops are in different sergeant areas. With significant deterrence effects, we would expect SPC slope derived using the second set of stops to be smaller. Instead, we find the two slopes to be statistically indistinguishable.[32] In other words, we find no evidence of deterrence.

In the third exercise, we examine deterrence effects in the context of a large-scale reallocation of troopers to the border region in 2014. In June of that year, state authorities in Texas moved troopers to the Rio Grande Valley as part of Operation Strong Safety, which was intended to "address the significant influx of undocumented aliens including undocumented minors coming across the border" (TDPS, 2014). Although there was no associated change in the criteria for conducting stops or searches, this influx of troopers resulted in a dramatic increase in the number of stops and searches conducted in the three sergeant areas located in the Rio Grande Valley (Benning and Chavez, 2016). Despite the increased risk of search faced by motorists, there was little change in the contraband finding rate in response to the influx of troopers (as shown in Online Appendix Figure B18). This lack of motorist responsiveness to variation in search risk provides further evidence that deterrence effects on the margin are likely negligible in our setting.

Importantly, the lack of measured deterrence effects is applicable only for the range of sergeant-

[32]The within-sergeant area estimate is 0.294 (with standard error 0.020). The between-sergeant area estimate is 0.292 (0.012).

area-level search rates we observe. Deterrence effects may be non-linear so that motorists are substantially more responsive to larger changes in search risk. For example, while we find that searches of white motorists are weakly more productive than searches of black or Hispanic motorists at the margin, suggesting that troopers could increase contraband yield by only searching white motorists, deterrence effects may be more relevant in that counterfactual. In addition, it is possible that motorists are insensitive to temporal or between-sergeant area variation in search rates because they are not aware of that variation. In that case, deterrence effects may be more relevant in counterfactuals where changes in motorist racial group-specific search rates are more salient. Nonetheless, the lack of a change in motorist behavior in response to the well-publicized trooper surge associated with Operation Strong Safety suggests that even salient changes in search rates may have limited deterrence effects.

VI. Are Search Disparities Driven by Racial Bias?

Black and Hispanic motorists are more likely to be searched, but racial disparities in search rates are not justified on efficiency grounds. As such, an outstanding question is why troopers elect to search black and Hispanic motorists more frequently. Two candidate explanations are that (1) troopers hold inaccurate beliefs regarding differences in contraband carrying behavior by race or (2) troopers search black and Hispanic motorists more frequently due to racial bias. Understanding the precise reason why black and Hispanic motorists are searched more frequently in the absence of efficiency gains is not central to our analysis, since our counterfactual conclusions hold regardless of the source of search rate disparities. Moreover, the legality of search rate disparities in the absence of efficiency gains does not likely depend on the specific mechanism that explains these disparities. Nonetheless, we briefly investigate the potential role of racial bias by assessing whether trooper-level racial disparities in search rates are associated with three factors: trooper race, local political preferences, and local disparities in citation rates.

A common test for racial bias in the policing literature is to compare the behavior of officers from different racial groups (Anwar and Fang, 2006; Close and Mason, 2007; Antonovics and Knight, 2009; West, 2018; Goncalves and Mello, forthcoming). The typical approach is to test whether black-white search rate disparities are smaller or reversed for black troopers, or whether Hispanic-white search disparities are smaller or reversed for Hispanic troopers. The premise is that if search disparities are driven by racial bias, we should expect biased troopers to favor motorists from the same racial group. Online Appendix Table B18 documents search rates and hit rates by both motorist and trooper race. Online Appendix Table B19, discussed in more detail in Online Appendix B, summarizes differences in black-white and Hispanic-white search odds ratios by trooper race that account for other stop and motorist characteristics. We find that all trooper racial groups are more likely to search black and Hispanic motorists than white motorists, but the black-white disparity is smaller for black troopers. The Hispanic-white disparity is similar for white and Hispanic troopers and smaller for black troopers.

We next examine whether, at the sergeant area level, racial group search rate disparities are associated with two proxies for trooper preferences and beliefs: racial disparities in citation rates (Goncalves and Mello, forthcoming) and the local Republican vote share in the 2016 presidential election (Cohen and Yang, 2019).[33] For the 79 sergeant areas that we include in our estimation of race-specific SPCs, we calculate black-white and Hispanic-white search odds ratios using the following logistic regression model:

$$P(\text{SEARCH}_{it} = 1 | X_{it}) = \frac{e^{\left(X_{it}\beta + \omega_{\tau(t)} + \delta_{m(t)}\right)}}{1 + e^{\left(X_{it}\beta + \omega_{\tau(t)} + \delta_{m(t)}\right)}} \tag{13}$$

where X_{it} includes indicators for whether the motorist is female, black, and Hispanic.

To measure racial disparities in citation rates, for each sergeant area we estimate logistic regression models analogous to equation (13) where the outcome is replaced with an indicator for whether the stop led to a speeding citation. Overall, white motorists are cited in 28.2% of stops, while black and Hispanic motorists are cited in 34.9% and 37.5% of stops. We next calculate the Republican vote share in each sergeant area in the 2016 presidential election. For sergeant areas that cover multiple counties, we take a weighted average of the county-level Republican vote shares where weights reflect the share of sergeant area stops conducted in each county.

Online Appendix Table B20 summarizes the joint, sergeant area-level relationship between search disparities, citation disparities, and Republican vote share. The relationship between black-white search disparities and vote share is both economically and statistically significant. A 10 percentage point increase in the Republican vote share is associated with an increase in the black-white odds ratio of 0.32. By contrast, there is no detectable relationship between black-white search disparities and citation disparities. We find a marginal positive relationship between Hispanic-white search and citation odds ratios, but no detectable relationship between the Hispanic-white search disparity and Republican vote share.

VII. Conclusion

In this paper we use unique administrative data on traffic stops conducted by the Texas Highway Patrol to evaluate whether racial profiling poses an equity-efficiency trade-off. As in previous analyses, we find that troopers are more likely to search black and Hispanic motorists than white motorists following stops for speeding, yet these searches are equally or less likely to yield contraband. In general, this finding does not imply that troopers can equalize search rates without sacrificing contraband yield due to the inframarginality problem: average search productivity may differ from search productivity at the margin. However, we show that the inframarginality problem is not empirically relevant in our setting. We exploit variation across and within troopers in search behavior and find that the relationship between trooper search rates and the proportion of stops that yield contraband is approximately linear. This finding suggests that, among motorists

[33] Cohen and Yang (2019) find that Republican-appointed district court judges exhibit larger black-white disparities in prison sentences than Democrat-appointed district court judges.

searched with positive probability, troopers are unable to distinguish between those who are more or less likely to carry contraband.

We consider two types of policy counterfactuals that reduce racial disparities in search rates. In one we reallocate troopers across patrol areas; in another, we require troopers to equalize their search rates across motorist racial groups. We find that it is feasible for troopers to (1) search all motorist racial groups at the same rate, (2) maintain the status quo overall search rate, and (3) increase overall contraband yield.

Our findings highlight a limitation of the Becker (1957, 1993) outcome test: when the returns to search are constant, as may be the case when troopers are limited in their ability to discern between motorists that are more or less likely to carry contraband, equalized marginal hit rates do not imply an equity-efficiency trade-off. More generally, our findings demonstrate that even if racial disparities in treatment cannot be definitively attributed to racial bias, such disparities may not be a necessary consequence of efficient decision-making.

Ultimately, understanding precisely why trooper search rates differ so dramatically by motorist racial group is beyond the scope of our paper, and remains an important area for future research. The answer will inform how policymakers can most effectively induce troopers to eliminate these disparities.

University of Illinois at Chicago
Haas School of Business, University of California, Berkeley, and NBER

References

Anderson, April J., "Racial Profiling: Constitutional and Statutory Considerations for Congress," July 2020. Congressional Research Service.

Antonovics, Kate and Brian G. Knight, "A New Look at Racial Profiling: Evidence fom the Boston Police Department," *Review of Economics and Statistics*, February 2009, *91* (1), 163–177.

Anwar, Shamena and Hanming Fang, "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Review*, March 2006, *96* (1), 127–151.

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

_ , _ , and Peter Hull, "Measuring Racial Discrimination in Bail Decisions," October 2020. Unpublished manuscript.

Ayres, Ian, "Outcome Tests of Racial Disparities in Police Practices," *Justice Research and Policy*, 2002, *4*, 131–142.

Barreca, Alan, Jason Lindo, and Glen Waddell, "Heaping-Induced Bias in Regression Discontinuity Designs," *Economic Inquiry*, 2016, *54* (1), 268–293.

Baumgartner, Frank R., Derek A. Epp, and Kelsey Shoub, *Suspect Citizens: What 20 Million Traffic Stops Tell Us About Policing and Race*, Cambridge, United Kingdom: Cambridge University Press, 2018.

Becker, Gary S., *The Economics of Discrimination*, Chicago: University of Chicago Press, 1957.

_ , "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, June 1993, *101* (3), 385–409.

Benning, Tom and Andrew Chavez, "After 2 years and $800M, Texas' Border Boost Has one Solid Outcome: More Traffic Tickets," The Dallas Morning News, November 22, 2016. `https://www.dallasnews.com/news/politics/2016/11/22/after-2-years-and-800m-texas-border-boost-has-one-solid-outcome-more-traffic-tickets/` [Accessed: 2021-04-06].

Bjerk, David, "Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate," *Journal of Public Economic Theory*, 2007, *9* (3), 521–546.

Cattaneo, Matias, Richard Crump, Max Farrell, and Yingjie Feng, "On Binscatter," 2019. Unpublished manuscript.

Chan, David C., Matthew Gentzkow, and Chuan Yu, "Selection with Variation in Diagnostic Skill: Evidence from Radiologists," April 2020. Unpublished manuscript.

Close, Billy R. and Patrick Leon Mason, "Searching for Efficienct Enforcement: Officer

Characteristics and Racially Biased Policing," *Review of Law and Economics*, 2007, *3* (2), 263–321.

Cohen, Alma and Crystal S. Yang, "Judicial Politics and Sentencing Decisions," *American Economic Journal: Economic Policy*, February 2019, *11* (1), 160–191.

Collister, Brian, "Texas Troopers Ticketing Hispanic Drivers as White," Kxan Investigates, November 6, 2015. `http://www.kxan.com/news/investigations/texas-troopers-ticketing-hispanics-motorists-as-white/1156475533` [Accessed: 2021-04-06].

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco," *American Economic Review*, September 2019, *109* (9), 3365–3394.

Dominitz, Jeff and John Knowles, "Crime Minimisation and Racial Bias: What Can We Learn From Police Search Data?," *Economic Journal*, November 2006, *116* (515), F368–F384.

Durlauf, Steven N., "Assessing Racial Profiling," *Economic Journal*, November 2006, *116* (515), F402–F426.

Epp, Charles R., Steven Maynard-Moody, and Donald P. Haider-Markel, *Pulled Over: How Police Stops Define Race and Citizenship*, Chicago: University of Chicago Press, 2014.

Frandsen, Brigham R., Lars J. Lefgren, and Emily C. Leslie, "Judging Judge Fixed Effects," November 2020. Unpublished manuscript.

Fryer, Roland, "An Empirical Analysis of Racial Differences in Police Use of Force," *Journal of Political Economy*, June 2019, *127* (3), 1210–1261.

Goldstein, Herman, "Police Discretion: The Ideal versus the Real," *Public Administration Review*, September 1963, *23* (3), 140–148.

Goncalves, Felipe and Steve Mello, "A Few Bad Apples? Racial Bias in Policing," *American Economic Review*, forthcoming.

Grogger, Jeffrey and Greg Ridgeway, "Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness," *Journal of the American Statistical Association*, 2006, *101* (475), 878–887.

Hahn, Frank, "On Some Difficulties of the Utilitarian Economist," in Amartya Sen and Bernard Williams, eds., *Utilitarianism and Beyond*, Cambridge University Press, 1982, chapter 9, pp. 187–198.

Harcourt, Bernard E., "Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature, and of Criminal Profiling More Generally," *University of Chicago Law Review*, 2004, *71* (4), 1275–1381.

Horrace, William C. and Shawn M. Rohlin, "How Dark Is Dark? Bright Lights, Big City, Racial Profiling," *Review of Economics and Statistics*, May 2016, *98* (2), 226–232.

Hull, Peter, "What Marginal Outcome Tests Can Tell Us About Racially Biased Decision-Making," February 2021. National Bureau of Economic Research Working Paper 28503.

Kelling, George L., "'Broken Windows' and Police Discretion," 1999. US Department of Justice, Office of Justice Programs, National Institute of Justice.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainanthan, "Human Decisions and Machine Predictions," *Quarterly Journal of Economics*, February 2018, *133* (1), 237–293.

Knowles, John, Nicola Persico, and Petra Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, February 2001, *109* (1), 203–229.

Knox, Dean, Will Lowe, and Jonathan Mummolo, "Administrative Records Mask Racially Biased Policing," *American Political Science Review*, 2020, *114* (3), 619–637.

Luh, Elizabeth, "Not So Black and White: Uncovering Racial Bias from Systematically Misreported Trooper Reports," 2020. Unpublished manuscript.

MacDonald, John M. and Jeffrey Fagan, "Using Shifts in Deployment and Operations to Test for Racial Bias in Police Stops," *AEA Papers and Proceedings*, May 2019, *109*, 148–51.

Marx, Philip, "An Absolute Test of Racial Prejudice," *Journal of Law, Economics, and Organization*, forthcoming.

Mastrofski, Stephen D., "Controlling Street-Level Police Discretion," *The Annals of the American Academy of Political and Social Science*, 2004, *593* (1), 100–118.

Morris, Carl N., "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of the American Statistical Association*, 1983, *78* (381), 47–55.

Persico, Nicola, "Racial Profiling, Fairness, and Effectiveness of Policing," *American Economic Review*, December 2002, *92* (5), 1472–1497.

Phillips, David, "Measuring Housing Stability with Consumer Reference Data," *Demography*, 2020, *57* (4), 1323–1344.

Pierson, Emma, Camelia Simoiu, Jan Overgoor, Sam Corbett-Davies, Daniel Jenson, Amy Shoemaker, Vignesh Ramachandran, Phoebe Barghouty, Cheryl Phillips, Ravi Shroff, and Sharad Goel, "A Large-scale Analysis of Racial Disparities in Police Stops Across the United States," *Nature Human Behaviour*, 2020, *4*, 736–745.

Rose, Evan K., "Who Gets a Second Chance? Effectiveness and Equity in Supervision of Criminal Offenders," *Quarterly Journal of Economics*, May 2021, *136* (2), 1199–1253.

Sen, Amartya, "Utilitarianism and Welfarism," *Journal of Philosophy*, 1979, *76* (9), 463–489.

Simoiu, Camelia, Sam Corbett-Davies, and Sharad Goel, "The Problem of Infra-Marginality in Outcome Tests for Discrimination," *The Annals of Applied Statistics*, 2017, *11* (3), 1193–1216.

Texas Department of Public Safety (TDPS), "Public Safety Commission Meeting Minutes," October 2014. https://www.dps.texas.gov/public_safety_commission/documents/2014/1016/minutes20141016.pdf [Accessed: 2021-04-06].

Tiwara, Alisa, "Disparate-Impact Liability for Policing," *The Yale Law Journal*, 2019, *129* (1), 252–306.

West, Jeremy, "Racial Bias in Police Investigations," 2018. Unpublished manuscript.

TABLE I

TRAFFIC STOP DESCRIPTIVE STATISTICS

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	9.09	100	0	0	18.91
% Hispanic	0	100	0	35.88	0	100	0	41.82
% White	0	0	100	55.03	0	0	100	39.26
% Female	38.95	32.59	37.67	35.96	16.32	15.58	20.80	17.77
Log Median Income	10.70	10.71	10.95	10.84	10.59	10.60	10.87	10.71
	(0.500)	(0.495)	(0.466)	(0.495)	(0.492)	(0.469)	(0.471)	(0.493)
Expected Log Income Given	-0.161	-0.083	0.081	0.000	-0.524	-0.456	-0.371	-0.436
Vehicle (Standardized)	(1.013)	(0.966)	(1.012)	(1.000)	(0.889)	(0.841)	(0.893)	(0.873)
Stop History (%):								
No Prior Stops	60.32	57.31	56.53	57.15	58.12	56.72	53.60	55.76
Prior Stop, No Search	37.92	41.13	42.61	41.66	32.01	35.12	37.50	35.47
Prior Search, No Contraband	1.130	1.205	0.534	0.829	4.669	4.429	3.796	4.226
Prior Search, Contraband	0.626	0.356	0.322	0.362	5.195	3.728	5.098	4.544
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	87.02	89.41	93.48	91.43	64.46	72.42	71.80	70.67
1-2 Prior Non-Drug Arrests	7.158	6.830	4.296	5.465	14.88	14.46	14.24	14.45
3+ Prior Non-Drug Arrests	5.824	3.760	2.223	3.102	20.66	13.12	13.96	14.88
Drug Arrest History (%):								
No Prior Drug Arrests	93.92	96.29	97.55	96.77	73.48	82.39	79.42	79.54
1 Prior Drug Arrest	2.814	2.161	1.363	1.781	10.13	8.588	9.426	9.208
2+ Prior Drug Arrests	3.263	1.547	1.088	1.451	16.40	9.019	11.16	11.25
Search Rate (%)	2.202	1.234	0.755	1.059	100	100	100	100
Unconditional Hit Rate (%)	0.757	0.324	0.285	0.342	34.00	25.90	37.35	31.93
Observations	448,337	1,769,369	2,713,626	4,931,332	9,874	21,832	20,497	52,203

Sample restrictions are described in Section II. All outcome values, excluding 'Log Median Income' and 'Expected Log Income Given Vehicle', are expressed as percentage points. 'Log Median Income' refers to the median household income for the Census block group of the motorist's residential address as measured in the 2009-2013 5-year American Community Survey. 'Expected Log Income Given Vehicle' is the average Log Median Income associated with a vehicle, where vehicles are classified as a combination of make, type (passenger car, pick-up truck, SUV), and age (above and below median given make and type), generating 204 total vehicle categories. We standardize Expected Log Income Given Vehicle to have mean zero and standard deviation one in our sample of stops.

Table II

Motorist Selection into Stops by Trooper Search Rate

				Excluding Most Selective Troopers		
	100 × SEARCH_{it} (1)	100 × $s^{-it}_{p\ell}$ (2)	100 × $\tilde{s}^{-it}_{p\ell}$ (3)	100 × SEARCH_{it} (4)	100 × $s^{-it}_{p\ell}$ (5)	100 × $\tilde{s}^{-it}_{p\ell}$ (6)
Black	0.888	0.106	0.066	0.774	0.050	0.026
	(0.058)	(0.020)	(0.020)	(0.054)	(0.013)	(0.012)
Hispanic	0.299	0.028	0.009	0.274	0.007	-0.005
	(0.022)	(0.008)	(0.008)	(0.022)	(0.006)	(0.006)
Female	-0.513	-0.023	-0.018	-0.474	-0.010	-0.008
	(0.021)	(0.004)	(0.004)	(0.021)	(0.003)	(0.003)
Log Median Income	-0.312	-0.007	-0.006	-0.281	0.003	-0.001
	(0.019)	(0.005)	(0.005)	(0.019)	(0.004)	(0.004)
Expected Log Income Given	-0.295	-0.027	-0.025	-0.258	-0.011	-0.011
Vehicle (Standardized)	(0.012)	(0.003)	(0.003)	(0.011)	(0.002)	(0.002)
1-2 Prior Non-Drug Arrests	0.805	0.011	0.008	0.744	-0.003	-0.004
	(0.053)	(0.005)	(0.005)	(0.055)	(0.004)	(0.004)
3+ Prior Non-Drug Arrests	1.700	0.022	0.019	1.531	0.008	0.008
	(0.099)	(0.008)	(0.008)	(0.101)	(0.006)	(0.006)
1 Prior Drug Arrest	3.304	0.022	0.017	3.106	0.002	0.001
	(0.145)	(0.007)	(0.007)	(0.146)	(0.006)	(0.006)
2+ Prior Drug Arrests	5.324	0.029	0.020	4.823	-0.000	-0.004
	(0.213)	(0.010)	(0.010)	(0.218)	(0.008)	(0.008)
Prior Stop, No Search	-0.221	-0.015	-0.008	-0.192	-0.010	-0.006
	(0.017)	(0.004)	(0.004)	(0.017)	(0.004)	(0.004)
Prior Search, No Contraband	2.943	0.072	0.065	2.806	0.033	0.031
	(0.201)	(0.017)	(0.017)	(0.201)	(0.012)	(0.012)
Prior Search, Contraband	10.025	0.158	0.154	8.722	0.100	0.105
	(0.642)	(0.025)	(0.024)	(0.653)	(0.023)	(0.022)
Location by Time FEs	✓	✓	✓	✓	✓	✓
Joint F-Statistic	84.88	9.31	8.65	74.03	6.30	5.84
Observations	3,280,250	3,280,250	3,280,171	2,739,955	2,739,955	2,739,899

This table presents coefficients from estimates of equation (7), where in columns (2) and (3) we replace the outcome SEARCH_{it} with $s^{-it}_{p(i,t)\ell(i,t)}$ and $\tilde{s}^{-it}_{p(i,t)\ell(i,t)}$, leave-out trooper search rates corresponding to the trooper who conducted the stop. SEARCH_{it} is defined as an indicator variable and $s^{-it}_{p(i,t)\ell(i,t)}$ takes on values between zero and one. $\tilde{s}^{-it}_{p(i,t)\ell(i,t)}$ takes on values between zero and one before it is residualized. Columns (4)–(6) exclude stops conducted by the 20% of troopers with the most selected set of stopped motorists. Standard errors are clustered at the trooper level. 'Joint F-Statistic' refers to an F-test for the joint significance of all motorist characteristics.

Table III
Search Productivity Curve Estimates, within-Motorist Design

Outcome:	Δ_{it}SEARCH	Δ_{it}CONTRABAND	
	First Stage	Reduced Form	2SLS
	(1)	(2)	(3)
Pooled			
$\Delta_{it}\tilde{s}_{p\ell}$	0.733	0.214	
	(0.016)	(0.009)	
Δ_{it}SEARCH			0.292
			(0.011)
Observations	694,246	694,246	694,246
White Motorists			
$\Delta_{it}\tilde{s}^r_{p\ell}$	0.737	0.263	
	(0.032)	(0.020)	
Δ_{it}SEARCH			0.357
			(0.022)
Observations	246,070	246,070	246,070
Black Motorists			
$\Delta_{it}\tilde{s}^r_{p\ell}$	0.759	0.242	
	(0.041)	(0.025)	
Δ_{it}SEARCH			0.319
			(0.028)
Observations	45,489	45,489	45,489
Hispanic Motorists			
$\Delta_{it}\tilde{s}^r_{p\ell}$	0.694	0.221	
	(0.040)	(0.020)	
Δ_{it}SEARCH			0.318
			(0.028)
Observations	76,047	76,047	76,047

This table presents coefficients for the two-stage least squares (2SLS) system described by equations (8) and (9). Each observation is a pair of sequential stops for a given motorist. SEARCH$_{it}$ and CONTRABAND$_{it}$ are defined as indicator variables and $\tilde{s}_{p(i,t)\ell(i,t)}$ and $\tilde{s}^r_{p(i,t)\ell(i,t)}$ take on values between zero and one (before it is residualized). Standard errors are clustered at the motorist level.

Table IV
Counterfactual Search Rates with Reallocated Troopers

	Status Quo	By Black Share			By Hispanic Share			By Nonwhite Share		
% Reassigned:	0%	20%	50%	100%	20%	50%	100%	20%	50%	100%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Observed Search Rates (%)										
All Motorists	1.16	1.14	1.13	1.12	1.16	1.15	1.14	1.14	1.13	1.13
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black Motorists	2.16	1.85	1.62	1.37	2.11	2.02	1.91	1.90	1.76	1.63
		(0.03)	(0.02)	(0.03)	(0.05)	(0.05)	(0.03)	(0.03)	(0.02)	(0.03)
Hispanic Motorists	1.37	1.32	1.34	1.35	1.22	1.10	0.98	1.25	1.14	1.03
		(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
White Motorists	0.81	0.87	0.90	0.94	0.88	0.93	0.99	0.89	0.95	1.02
		(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reweighted Search Rates (%)										
All Motorists	1.16	1.12	1.12	1.16	1.13	1.13	1.18	1.12	1.12	1.16
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Black Motorists	2.16	1.89	1.71	1.61	2.08	2.02	2.02	1.92	1.83	1.79
		(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)
Hispanic Motorists	1.37	1.30	1.31	1.37	1.22	1.13	1.09	1.25	1.16	1.13
		(0.03)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
White Motorists	0.81	0.83	0.86	0.93	0.84	0.89	0.97	0.85	0.90	0.99
		(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

This table presents observed and simulated search rates (in percentage points) for counterfactuals where troopers are reallocated across sergeant areas. This exercise is described in more detail in Section V.A. We restrict to trooper-location combinations with at least 50 stops for each motorist racial group. We then order trooper-location combinations by their search rate for white motorists. We use the top and bottom X% of trooper-location combinations for reallocation. For each combination we calculate the black share of stopped motorists (or Hispanic share, or the nonwhite share). We then reallocate the trooper-location combination with the highest search rate to the stops for the trooper-location combination with the lowest black share of stops, the combination with the second highest search rate to the stops for the combination with the second lowest black share of stops, and so on. We then repeat this exercise where we reallocate based on the Hispanic share of stops (columns (5)-(7)) or the nonwhite share of stops (columns (8)-(10)). We construct counterfactual search rates in two ways. In the top panel, we assume that troopers would have the same race-specific search rates if they are reallocated. In the bottom panel, we assume that troopers' counterfactual race-specific search rates are a weighted average of their observed search rates and the observed search rates of the sergeant area to which they are allocated.

Table V

Within-Trooper Returns to Search

Outcome:	$P(\text{SEARCH}_{it} \mid X_{it})$	SEARCH_{it} First Stage			CONTRABAND_{it} 2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SEARCH_{it}					0.374	0.373	0.391
					(0.033)	(0.034)	(0.058)
$s_{\ell(i,t)y(t)}^{-(i,t)}$	0.032	0.422	0.394	0.213			
	(0.004)	(0.040)	(0.039)	(0.031)			
Trooper FEs	✓	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓	✓
Motorist Controls			✓			✓	
Motorist FEs				✓			✓
Kleibergen-Paap F-Stat		110.78	103.83	47.69			
Observations	4,351,217			2,059,851	4,351,217		2,059,851

This table presents estimates of equations (11) and (12). $s_{\ell(i,t)y(t)}^{-(i,t)}$ denotes the search rate for all stops in location ℓ in the year corresponding to t, excluding stop (i, t). SEARCH_{it} and CONTRABAND_{it} are defined as indicator variables and $s_{\ell(i,t)y(t)}^{-(i,t)}$ takes on values between zero and one. Motorist characteristics include race, gender, log of neighborhood median income, vehicle-based expected log income, stop history, non-drug arrest history, and drug arrest history.

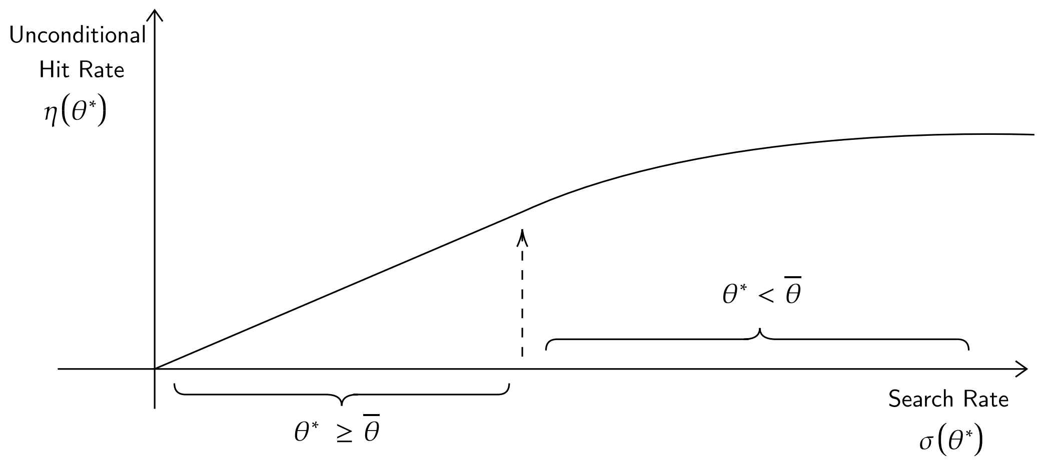
TABLE VI

COUNTERFACTUALS WHERE TROOPERS EQUALIZE THEIR RACE-SPECIFIC SEARCH RATES

	Observed	Estimated SPC-Based							
		Pooled		By Experience		By Trooper Race		By Trooper Stop Rate	
	Status Quo (1)	Status Quo (2)	Equalized (3)	Status Quo (4)	Equalized (5)	Status Quo (6)	Equalized (7)	Status Quo (8)	Equalized (9)
All Motorists									
Search Rate (%)	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12
Hit Rate (%)	34.70	35.59	36.41	35.13	36.66	34.93	35.91	35.09	36.60
		(0.52)	(0.58)	(0.56)	(0.58)	(0.52)	(0.61)	(0.54)	(0.64)
$H_0 : SQ = E$		0.036		0.000		0.023		0.000	
White Motorists									
Search Rate (%)	0.80	0.80	1.12	0.80	1.12	0.80	1.12	0.80	1.12
Hit Rate (%)	37.90	37.93	37.93	38.12	38.13	37.94	37.38	38.12	38.01
		(0.76)	(0.74)	(0.80)	(0.75)	(0.73)	(0.75)	(0.77)	(0.82)
Black Motorists									
Search Rate (%)	2.22	2.22	1.12	2.22	1.12	2.22	1.12	2.22	1.12
Hit Rate (%)	34.23	34.25	34.25	34.39	34.85	34.18	33.98	34.54	34.58
		(1.24)	(1.29)	(1.23)	(1.41)	(1.23)	(1.38)	(1.27)	(1.53)
Hispanic Motorists									
Search Rate (%)	1.35	1.35	1.12	1.35	1.12	1.35	1.12	1.35	1.12
Hit Rate (%)	30.64	33.40	33.88	31.56	34.09	31.34	33.40	31.32	34.28
		(0.85)	(1.45)	(1.17)	(1.23)	(0.89)	(1.53)	(0.97)	(1.50)

This table presents observed and simulated counterfactual search rates and hit rates (in percentage points) by motorist racial group. To calculate observed search rates and hit rates we take the sample of stops used to estimate race-specific search productivity curves (SPCs) (sample criteria described in Section IV.B) and further limit the analysis to stops conducted by troopers we are able to match to personnel records. We estimate racial group SPCs and simulate counterfactual search rates and hit rates as described in Section V.C. In columns (2) and (3) we estimate race-specific SPCs pooling all troopers. In columns (4) and (5) we allow race-specific SPCs to vary within trooper experience quartile. In columns (6) and (7) we allow race-specific SPCs to vary with trooper race. In columns (8) and (9) we allow race-specific SPCs to vary with trooper stop rate. Bootstrap standard errors are provided in parentheses.

Figure I
A Theoretical Search Productivity Curve

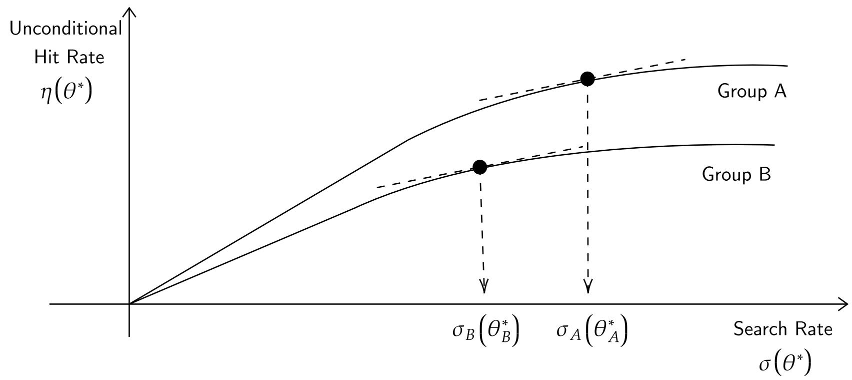


Note: This figure depicts a theoretical example of a trooper's search productivity curve (SPC), the relationship between the trooper's search rate, $\sigma(\theta^*)$, and unconditional hit rate, $\eta(\theta^*)$. For signal thresholds $\theta^* \geq \bar{\theta}$, the SPC is linear. For $\theta^* < \bar{\theta}$, the relationship is concave, as the marginal searched motorist is less likely to have contraband than inframarginal searched motorists.

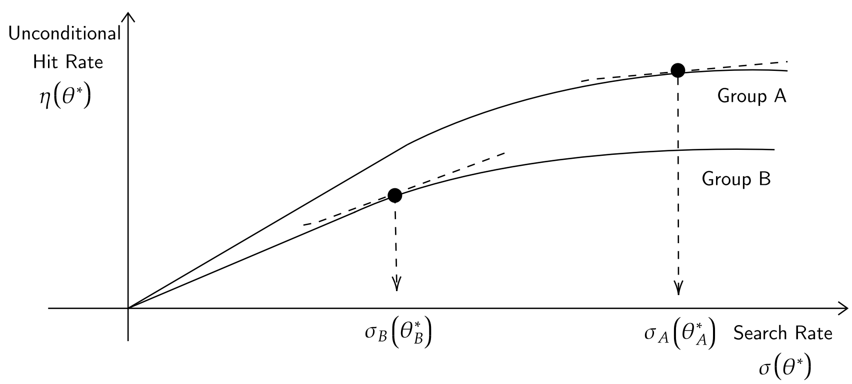
Figure II

Theoretical Cases with and without an Equity-Efficiency Trade-Off

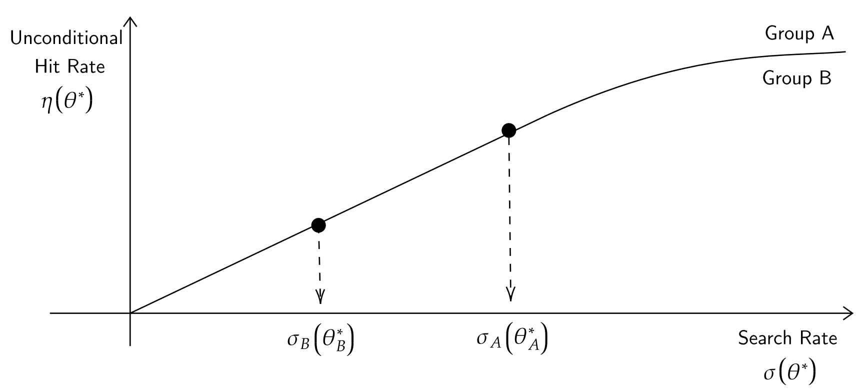
(a) Trade-Off



(b) No Trade-Off and $\theta_r^* < \bar{\theta}_r$

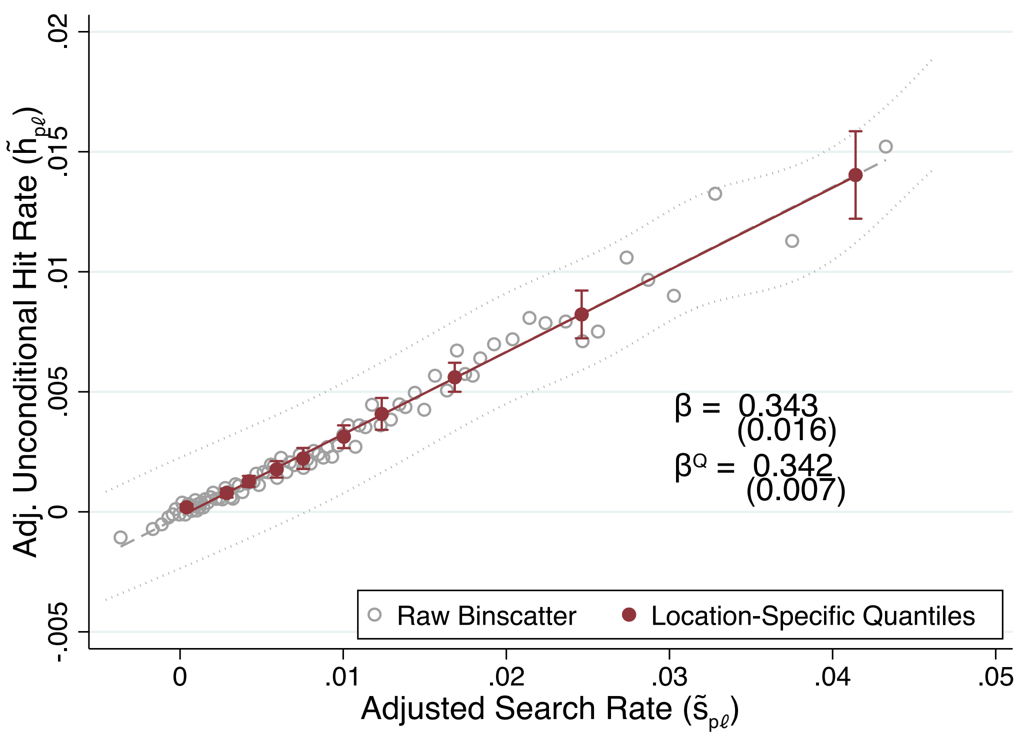


(c) No Trade-Off and $\theta_r^* \geq \bar{\theta}_r$



Note: These figures depict three scenarios for search productivity curves (SPCs) for two groups of motorists, Group A and Group B. In all scenarios, the search rate for Group A exceeds the search rate for Group B. In Panel A, the trooper faces diminishing returns to search within each group and equalizes marginal hit rates between groups. Equalizing search rates between groups while maintaining the overall search rate would decrease the hit rate. In Panel B, the trooper faces diminishing returns to search within each group but does not equalize marginal hit rates. In Panel C, the trooper equalizes marginal hit rates between groups but faces constant returns to search. In Panels B and C, the trooper can equalize search rates without reducing the overall hit rate.

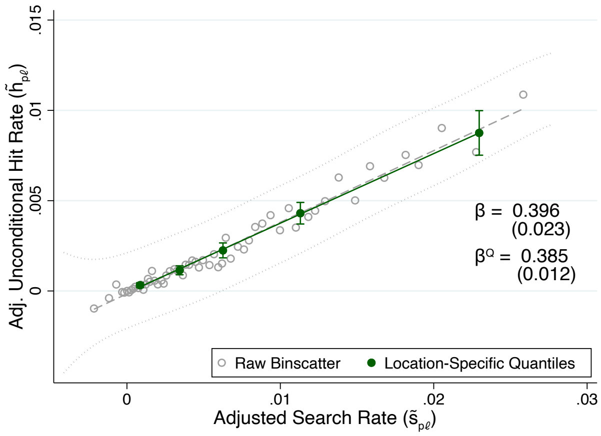
Figure III
Between-Trooper Search Productivity Curve



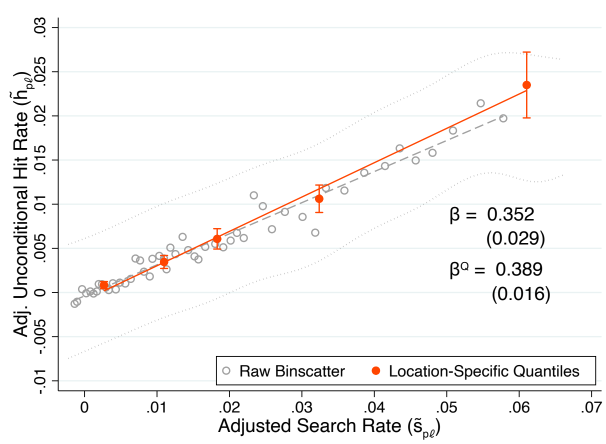
Note: In this figure we plot adjusted trooper unconditional hit rates ($\tilde{h}_{p\ell}$) against trooper search rates ($\tilde{s}_{p\ell}$) using two approaches described in Section IV.D, where $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ take on values between zero and one (before each is residualized). The first approach is a simple binscatter, where we choose the integrated mean square error-optimal number of bins as in Cattaneo et al. (2019) (using the Stata package `binsreg`). The figure includes 95% confidence bands for the local linear relationship between adjusted trooper search rates and unconditional hit rates and the best linear fit and its slope. The local linear fit is derived using a Gaussian kernel with a rule-of-thumb bandwidth. A bootstrap standard error for the estimated slope, where we stratify resampling by trooper and location, is provided in parentheses. In the second approach we divide troopers into quantiles by search rate within locations, group quantiles across locations, and then plot the relationship between search rates and unconditional hit rates across quantiles. From this approach, the figure includes the mean values for each decile and the best linear fit and its slope. A bootstrap standard error for the estimated slope is provided in parentheses.

Figure IV
Between-Trooper Search Productivity Curve, by Motorist Race

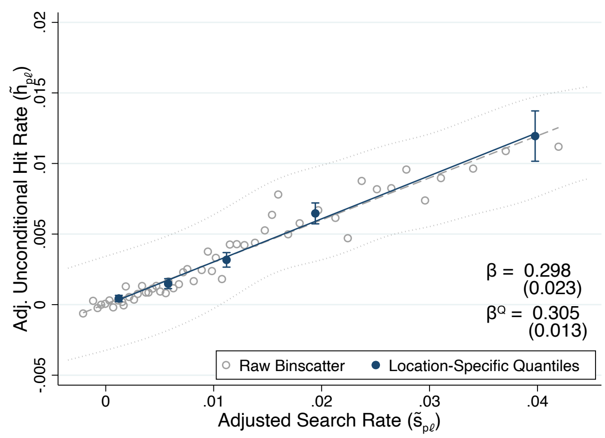
(a) White Motorists



(b) Black Motorists



(c) Hispanic Motorists

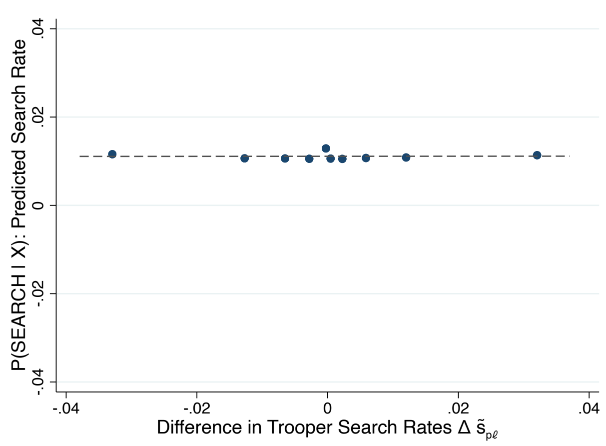


Note: In this figure we plot adjusted trooper unconditional hit rates ($\tilde{h}_{p\ell}$) against trooper search rates ($\tilde{s}_{p\ell}$) using two approaches described in Section IV.D, where $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ take on values between zero and one (before each is residualized). The first approach is a simple binscatter, where we choose the integrated mean square error-optimal number of bins as in Cattaneo et al. (2019) (using the Stata package `binsreg`). The figure includes 95% confidence bands for the local linear relationship between adjusted trooper search rates and unconditional hit rates and the best linear fit and its slope. The local linear fit is derived using a Gaussian kernel with a rule-of-thumb bandwidth. Bootstrap standard errors for the estimated slopes, where we stratify resampling by trooper and location, are provided in parentheses. In the second approach we divide troopers into quantiles by search rate within locations, group quantiles across locations, and then plot the relationship between search rates and unconditional hit rates across quantiles. From this approach, the figure includes the mean values for each quintile and the best linear fit and its slope. Bootstrap standard errors for the estimated slopes are provided in parentheses. Panel A, Panel B, and Panel C plot the search productivity curve (SPC) for white motorists, black motorists, and Hispanic motorists, respectively.

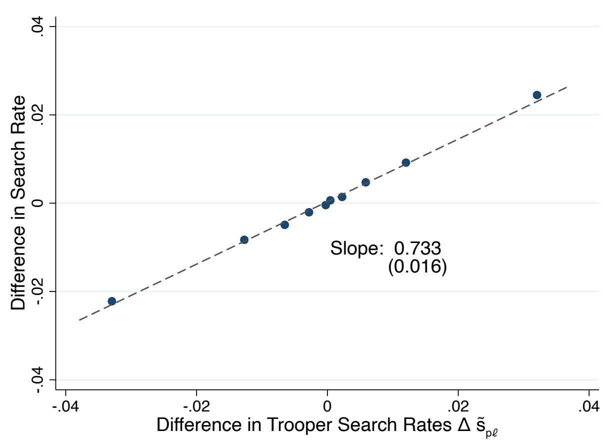
Figure V

Selection and within-Motorist Differences in Trooper Search Rates

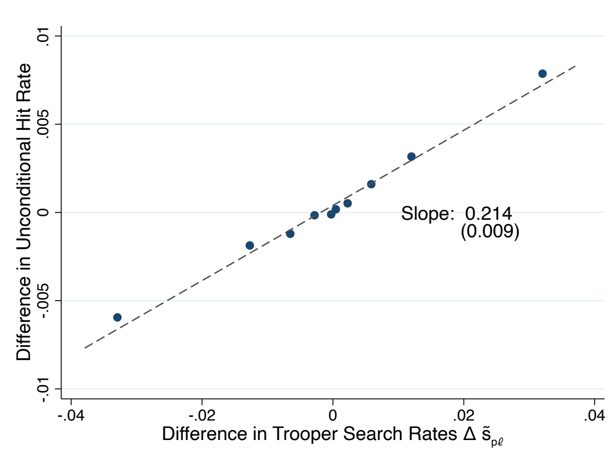
(a) Motorist Characteristics



(b) ΔSearch Rate

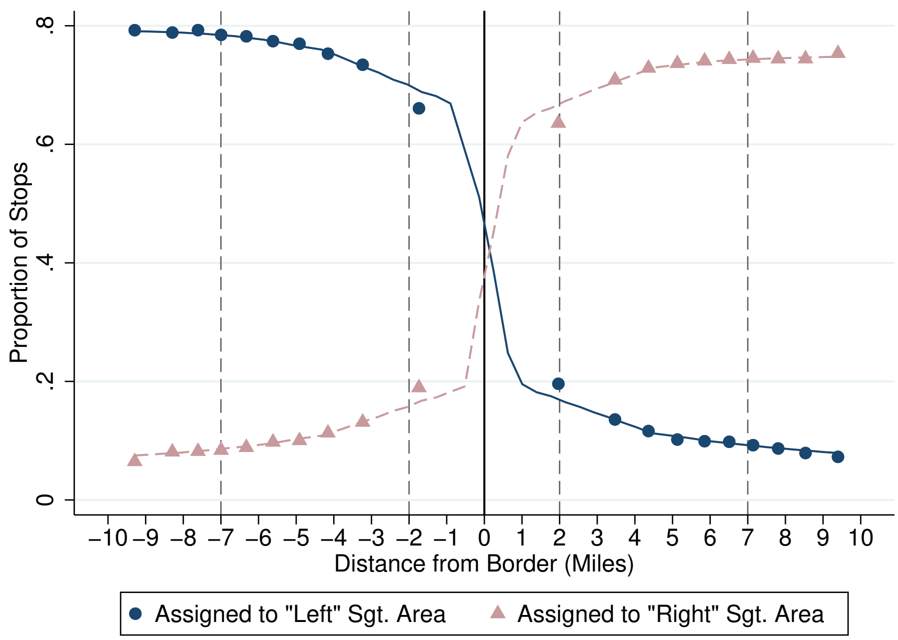


(c) ΔUnconditional Hit Rate



Note: This figure plots the relationship between the difference in trooper search rates associated with sequential pairs of stops for the same motorist, $\Delta_{it}\tilde{s}_{p\ell}$, and three variables: (1) motorist characteristics, (2) the difference in search rates between stops, and (3) the difference in unconditional hit rates between stops. Motorist characteristics are summarized by the probability of search given motorist characteristics at the time of the initial stop, $P(\text{SEARCH}_{it}|X_{it})$. Sequential pairs of stops are grouped by their decile value of $\Delta_{it}\tilde{s}_{p\ell}$. $P(\text{SEARCH}_{it}|X_{it})$, the Search Rate, and the Unconditional Hit Rate all take on values between zero and one. $\tilde{s}_{p\ell}$ takes on values between zero and one before it is residualized.

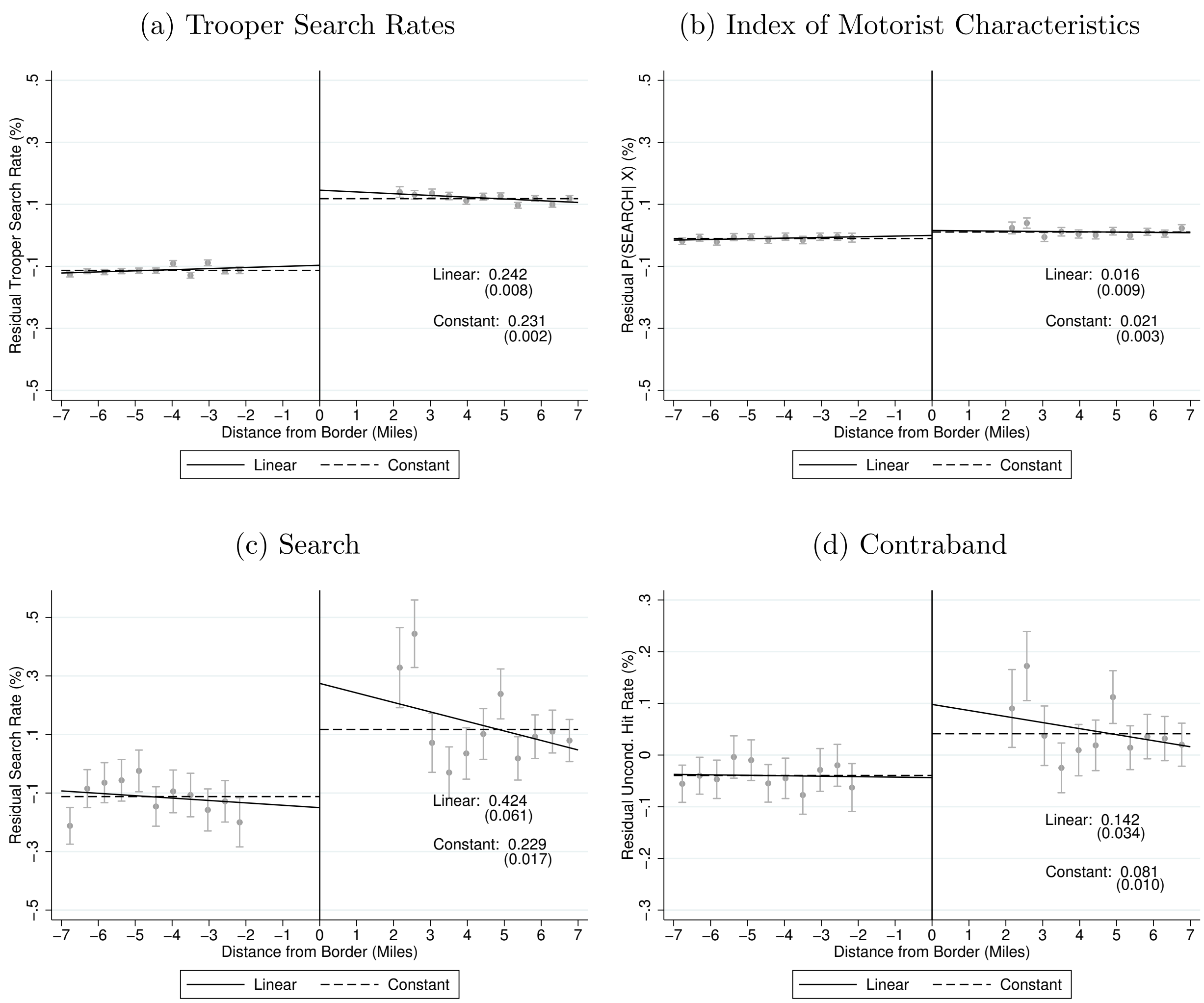
Figure VI
Trooper Assignments by Stop Location



Note: This figure plots the share of stops conducted by troopers assigned to each adjacent sergeant area by travel distance from highway and sergeant area border intersections as described in Section IV.E.2. The data are limited to stops within 10 miles of the intersection. The figure includes a bin scatter, where stops are grouped by side of the border and into deciles by distance from the intersection. Stops conducted between 2 and 7 miles of the intersections are included in the regression discontinuity (RD) analysis.

Figure VII

Stop Characteristics and Outcomes by Stop Distance from Border



Note: These figures are stacked regression discontinuity plots for four outcomes: leave-out trooper search rates (Panel A); $P(\text{SEARCH}_{it}|X)$ (Panel B), an index of motorist characteristics; search (SEARCH_{it}) (Panel C); and contraband (CONTRABAND_{it}) (Panel D). All outcome variables take on values between zero and one (before each is residualized). The running variable is the travel distance from a stop to its corresponding highway by sergeant area border intersection. Sample selection criteria are described in Section IV.E.2. The figures include discontinuity estimates using linear and constant extrapolation.