Online Appendix: Would Eliminating Racial Disparities in Motor Vehicle Searches Have Efficiency Costs?

Benjamin Feigenberg

Conrad Miller

June 2021

Part I

Table of Contents

A. Appendix: Data Construction

We merge traffic stop data to commercial address history data from Infogroup using full name and address. We first use an address standardization algorithm, the Stata function `stnd_address`, to ensure that addresses are structured analogously across the two data sets, with separate fields for street address, unit number, etc. We also extract the address number. In addition, we manually standardize Texas city and town names in the traffic stop data. We standardize full names and extract suffixes. We then use the Stata command `reclink2` to perform a probabilistic linkage across the two data sources. We fuzzy match using the following fields: last name, first name, middle name, suffix, address number, street name, city, and zip code. We require that observations match exactly on the first letter of the first name and the first letter of the last name. For zip code, we define agreement discretely based on whether the fields match exactly. For all other fields, we utilize the bigram string comparator to assess the degree of agreement. The address history data includes an identifier that matches the same individual to multiple addresses. We use this identifier to match multiple stops to the same person. We are able to match 75% of stops to the address history data. For stops that we are unable to match, we create identifiers based on full name, street address, and zip code.

We then match the criminal history data to traffic stops using the full set of addresses associated with each person. We apply the same address and name standardization to the criminal history data, and apply the same fuzzy match.

Though Diamond et al. (2019) and Phillips (2020) find that similar address history data from Infutor are of high quality, we are unable to match every stop to the address history data and these data may be incomplete. Hence, we may not correctly associate all stops and criminal history with the corresponding motorist.

To match geocoded stops to sergeant patrol areas, we use the sergeant area boundaries shapefile received in response to a Texas Public Information Act request. This shapefile includes two sergeant area identifiers: *sgt_area* and *sgt_area_n*. In practice, the *sgt_area* identifier includes a significant number of unique values corresponding to identical geographies and the same value of *sgt_area_n*. For example, the boundaries for *sgt_area* 1B03 and *sgt_area* 1B05 are identical; both objects are assigned to the same value of *sgt_area_n* ($1B03\_1B05$). As such, we rely on the *sgt_area_n* identifier to map stops to sergeant areas, and we reassign stops associated with the small number of remaining *sgt_area_n* values that are themselves unique but correspond to identical geographies. There are also instances in which distinct *sgt_area_n* objects are partially overlapping. In cases in which a stop is associated with multiple distinct but partially overlapping *sgt_area_n* values, we include one observation for each unique *sgt_area_n* value associated with the stop. The sergeant area(s) associated with each geocoded stop were identified using the `Spatial Join` analysis tool in ArcGIS.

ArcGIS was also used to construct the set of highway and sergeant area border intersections included in the RD analysis (Section IV.E.2) and to construct the associated driving distances between

stops and intersections. Highway and border intersections were identified using the `Intersect` analysis tool in conjunction with the sergeant area boundaries and Texas roadways shapefiles. To construct driving distances, we first used the `Origin-Destination Cost Matrix` analysis tool from the `Network Analyst` extension in ArcGIS to identify the driving distance to each highway and border intersection point within 30km from each geocoded stop. We then identified the set of stops ultimately included in the RD analysis (based on distance and sample size restrictions) in Stata.

A.1. Appendix A Tables

TABLE A1
SAMPLE SELECTION

	Observations	
Sample step	Dropped	Remaining
1. All stops conducted by Texas Highway Patrol between 2009 and 2015		15,956,460
2. Drop duplicate rows for the same stop	195,161	15,761,299
3. Retains stops made on state and interstate highways	4,059,903	11,701,396
4. Drop stops with missing location information	262,098	11,439,298
5. Drop stops in the state capitol region	51	11,439,247
6. Retain stops of motorists with Texas addresses	1,659,604	9,779,643
7. Retain stops of passenger cars, pick-up trucks, and SUVs	785,942	8,993,701
8. Drop stops with missing motorist information	599,832	8,393,869
9. Retain stops of motorists that are white, black, or Hispanic	172,977	8,220,892
10. Retain stops with at least one associated speeding violation	3,208,619	5,012,273
11. Drop stops with missing trooper ID or stop outcomes	724	5,011,549

B. Appendix: Additional Analyses

B.1. Variation in Cited Speeds Across Troopers Is Limited

We limit our analysis to traffic stops associated with speeding violations because, as previous researchers have argued (see, for example, Baumgartner et al., 2018), we believe these stops are more likely to be motivated by the traffic violation itself, rather than some investigatory motive. For stops that are motivated by the traffic violation itself, we expect the composition of stopped motorists to be more similar across troopers, conditional on the location and time of the stop.

In this section we document the extent that cited speeds vary across troopers. Each traffic stop is associated with a speeding warning or a speeding citation. There is a citation in 34% of stops. Actual speeds are observed for citations but not warnings.

We rescale cited speeds by taking the difference between the log cited speed and log posted speed limit. We refer to this rescaled speed as the *log speed above limit*.

The average log speed above limit is 0.208, meaning the average cited speed is about 21% over the posted speed limit. The standard deviation of log speed above limit is 0.076. 99% of cited speeds are at least 10% above the speed limit. Cited speeds are similar for white and Hispanic motorists (about 20.7% above the posted speed limit), while cited speeds are slightly higher for black motorists (22.4% above the posted speed limit).

For each trooper-by-location combination, we calculate the citation rate and average log speed above limit. Within locations, the standard deviation of average log speed above limit across troopers is 0.040. The difference in average log speed above limit between the 10th and 90th percentile of troopers is only about 10% of the speed limit, which for the average speed limit is about 6 miles per hour.

Troopers that cite more often have lower average cited speeds, but the differences are minor. For every 10 percentage point increase in the citation rate, cited speeds decrease by 0.4%.

B.2. Detailed Search Outcomes by Motorist Race

Detailed outcomes of searches are summarized in Table B1.

B.3. Descriptive Analysis of What Predicts Search and Contraband Yield

We use the uniquely rich merged data set to answer two descriptive questions: (1) what motorist characteristics predict trooper search? And (2) among those searched, what motorist characteristics predict whether a search yields contraband? The answers to these questions clarify the degree to which race-based differences in search and hit rates can be explained by factors correlated with race but that have been unobservable to previous researchers.

B.3.1. Racial Disparities in Search Rates

One advantage of our setting relative to prior analyses is that we have a much richer set of motorist covariates. It is potentially the case that racial differences in search rates documented previously–and interpreted as evidence of racial profiling–could be explained, at least in a statistical sense, by other motorist characteristics that are observed by troopers but typically not observed by researchers. We investigate this possibility by examining whether conditioning on criminal history, stop history, and income affects measured race-based differences in search rates.

For each stop, let i denote the motorist and t denote the specific time. The functions $\ell(i,t)$, $\tau(t)$, and $m(t)$ map each stop to its associated location, time category, and month (e.g., June 2013), respectively. We categorize time by the combination of quarter of day and whether the stop was conducted on a weekday or weekend. We estimate logistic regressions of the form

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

where SEARCH_{it} is an indicator whether the stop of motorist i at time t led to a search; λ_ℓ, ω_τ, and δ_m are fixed effects for stop location, time category, and month of the stop; and X_{it} is a vector of motorist characteristics, including race, gender, log of neighborhood median income, stop history, non-drug arrest history, and drug arrest history. We also construct a second proxy for motorist income based on the vehicle involved in the stop. We classify vehicles by make, type (passenger car, pick-up truck, SUV), and age (above and below median given make and type), generating 204 total vehicle categories. We then calculate the mean of log of median neighborhood income among stopped motorists for each vehicle category. To our knowledge, this is the first paper to examine the relationship between trooper search behavior and motorist's criminal history, stop history, and neighborhood income.

Odds ratios for estimates of equation (B.1) are presented in Table B2. Across specifications, we vary the set of covariates included in the model, moving from more parsimonious specifications to more saturated models. In column (1) we include only a subset of motorist characteristics (X_{it}): motorist race and gender. The baseline search rate for white motorists is 0.76 percent. The coefficient for *black* of 2.99 indicates that, controlling only for gender, odds of search are 3 times higher for black motorists. Given the low probabilities in this context, odds and probabilities are similar, meaning search rates are also approximately 3 times higher for black motorists. For Hispanic motorists, search rates are about 57% higher. Conditional on motorist race, women are about 62% less likely to be searched. In column (2) we add separate fixed effects for stop location, time category, and month. Doing this reduces the *black* odds ratio slightly to 2.64, while the *Hispanic* odds ratio increases to 1.68. The coefficient for *female* is unaffected. In column (3) we add our income proxies as covariates. The coefficient for *log median income* is 0.68, indicating that a one standard deviation increase in neighborhood income (about 35 log points) is associated with about a 11% decrease in search rates. The association with *vehicle-based expected log income* is similar. Including the income proxies as controls reduces the *black* odds ratio to 2.16 and the

Hispanic odds ratio to 1.45.

Column (4) adds motorist arrest history indices as explanatory variables. Previous arrests also predict searches, particularly drug arrests. The increase in search likelihood associated with black motorists relative to white motorists 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. Column (5) adds motorist stop history. Conditional on motorist demographics and arrest history, motorists who have been stopped previously but not searched previously are 30% less likely to be searched than motorists who have not been stopped previously, the omitted category. Motorists who have been searched previously but not found with contraband are about 160% more likely to be searched, while motorists who have been previously found with contraband are about 140% more likely to be searched.

Controlling for criminal and stop history reduces the black and Hispanic odds ratios to 1.81 and 1.42. Comparing columns (2) and (5), motorist income and criminal/stop history can statistically explain about 50% and 35% of the black-white and Hispanic-white disparities in search rates, respectively. Note that racial differences in stop history and likely criminal history already incorporate racial differences in police treatment. Hence, we think of these percentages as upper bounds on the share of black-white and Hispanic-white disparities that can be explained by these factors.

B.3.2. Racial Disparities in Hit Rates

Next, we estimate logistic models identical to (B.1) except that we replace the outcome with CONTRABAND_{it}, an indicator for whether a search yields contraband.[1] We limit estimation to stops that led to a search (i.e., where $\text{SEARCH}_{it} = 1$).

The results are presented in columns (6) through (10) of Table B2. The model specifications are analogous to those in columns (1) through (5).

There are four main findings. First, controlling for only motorist race and gender, searches of black and Hispanic motorists are about 15% and 40% less likely to yield contraband.

Second, hit rates are increasing in motorist income, and the magnitude of the relationship is economically significant. In columns (8) through (10) the coefficient for log neighborhood median income is 1.29, indicating that a one standard deviation increase in neighborhood income is associated with about a 10% increase in the hit rate. Interestingly, hit rates are unrelated to our vehicle-based proxy for motorist income.

Third, while previous drug arrests predict about a 45% increase in the hit rate, hit rates are weakly *lower* for motorists with previous non-drug arrests. For those with one or two previous non-drug arrests, the hit rate is the same as for those without any non-drug arrests; for those with more than two previous non-drug arrests, the hit rate is about 13% lower. This finding is particularly

[1] We show in Table B1 that the percent of stops yielding contraband that lead to an arrest and the severity of arrest charges, as proxied by the average incarceration sentence associated with conviction, are similar across motorist racial groups.

interesting given that previous non-drug arrests significantly *increase* a motorist's likelihood of being searched in the first place.

Fourth, for motorists who have been previously searched, the outcomes of those previous searches are highly predictive of contemporaneous outcomes. Relative to motorists with no stop history, searches of motorists who have been previously searched but not found with contraband are about 33% less likely to yield contraband. Searches of motorists who have been previously found with contraband are about 30% more likely to yield contraband.

B.4. Trooper Stop Rates, Cited Speeds, and Motorist Characteristics

Figure B6 plots motorists characteristics, as summarized by $P(\text{SEARCH}|X_{it})$, the search probability for each stop based on observable motorist characteristics, as a function of four trooper characteristics: search rate (Panel A), search rate at night (Panel B), time between stops (Panel C), and average cited speeds (Panel D). Time between stops (i.e., trooper stop rate) is measured as the within-shift number of minutes between sequential recorded stops, averaged over all sequential pairs of stops for a given trooper, location (sergeant area) and time bin (quarter of day by weekday/weekend).[2] Trooper stop rates and cited speeds are standardized to have mean zero and standard deviation one. We partial out location by time fixed effects for both $P(\text{SEARCH}|X_{it})$ and each trooper. Panel A is the figure analog to Table II.

For all trooper characteristics, the relationship is essentially flat. Troopers with varying search rates, stop rates, and cited speeds are stopping motorists with similar observable characteristics.

B.5. Aggregating Across Locations Using Location Fixed Effects

In the main text we aggregate location-specific SPCs using what we call 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. Here we use an alternative approach that we refer to as the *fixed effects* approach. In particular, 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). A shortcoming of this approach is that because the distribution of trooper search rates varies across locations, different portions of the SPC are estimated using varying sets of locations.

Pooled and race-specific SPCs using the fixed effects approach are provided in Figure B9.

B.6. Sampling Error in Estimating Between-Trooper Search Productivity Curves

We estimate search productivity curve slopes using various specifications in Table B6.

One 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 sampling error, and those errors are correlated. This correlated sampling error may bias our estimate of β.

[2] In the absence of shift schedule data, we define shifts to include sequences of stops for which the time between stops never exceeds seven hours.

As one approach to accounting for this measurement error, we adjust estimates of trooper-location search propensities using an Empirical Bayes (EB) approach (Morris, 1983, Aaronson et al., 2007). We observe trooper-location search rates, which are estimates of search propensities. Some trooper-location estimates are derived from more observations and are thus more precise than others. The EB estimate for trooper-location $p\ell$ is a weighted average of the trooper-location search rate and overall search rate of the location, where the weight is a function of the reliability of the trooper-location $p\ell$ estimate. We follow the approach of Chandra et al. (2016) and use their Stata code to construct EB estimates for trooper-location search rates, $s_{p\ell}^{EB}$. We construct an analogous EB estimate of conditional search propensities, $h_{p\ell}^{EB}$, using the same weighting.

In Table B6 we show alternative estimates for the SPC slope from regressing $h_{p\ell}^{EB}$ on $s_{p\ell}^{EB}$ with location fixed effects. The slope we estimate is indistinguishable from the slope we get using unadjusted search and unconditional hit rates.[3]

To account for sampling error, we also 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, as shown in Table B6, this procedure yields β estimates that are statistically indistinguishable from the OLS estimates.

B.6.1. Excluding Selective Troopers

A key concern with our research design is that stopped motorists are not randomly assigned to troopers. We take a 'conditional on observables' approach and argue that, conditional on stop time and location, the identity of the trooper conducting the stop is as good as randomly assigned. However, even conditional on these stop contextual characteristics, we see motorist characteristics that predict search (e.g., race, income, criminal history) also predict the search propensity of the troopers that stop them. This relationship is quite weak (as discussed in Section IV.B), and controlling directly for observable motorist characteristics does not affect any of our conclusions. Nonetheless, this selection may introduce bias.

Here we take a complementary approach to assess whether our results are sensitive to this form of selection. We exclude troopers for whom we find the most evidence of motorist selection, and then repeat our analysis using this selected sample of troopers.

We first describe how we identify the troopers to exclude. The goal is to identify troopers that have a composition of stopped motorists that deviates most from what one would expect based on the time and location of their stops alone. We estimate the following logistic regression model:

$$P(\text{SEARCH}_{it} = 1|X_{it}) = \frac{e^{(X_{it}\beta)}}{1 + e^{(X_{it}\beta)}} \tag{B.2}$$

[3]Note that the SPC slopes presented here differ somewhat from the slopes presented in Figure **??** and Figure **??**. The slopes in the main text are fit to local linear estimates for the relationship between search rates and unconditional hit rates over a more limited range of search rates. The slopes in Table B6 are derived from a linear regression of $h_{p\ell}^{EB}$ on $s_{p\ell}^{EB}$ with location fixed effects using all trooper-location combinations.

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. From this we calculate the search probability for each stop based on observable motorist characteristics, P(SEARCH | X_{it}). Figure B4 depicts a histogram of P(SEARCH | X_{it}) across stops.

We then characterize troopers by their mean value of P(SEARCH | X_{it}) after conditioning on stop time and location. We estimate the following Poisson pseudo-likelihood regression model:

$$\log(E(P(\text{SEARCH}|X_{it})|\psi_{p(i,t)\ell(i,t)}, \ell(i,t), \tau(t), m(t)) = \psi_{p(i,t)\ell(i,t)} + \lambda_{\ell(i,t)} + \omega_{\tau(t)} + \delta_{m(t)} + \epsilon_{it} \tag{B.3}$$

where $\psi_{p(i,t)\ell(i,t)}$ are trooper by location fixed effects. We estimate the model using the pseudo-maximum likelihood estimator of Correia et al. (2019). Figure B5 depicts a histogram of $\psi_{p\ell}$ across troopers. If the assignment of motorists to troopers conditional on stop time and location were indeed as good as random, $\psi_{p\ell}$ would only vary across troopers due to chance. Troopers with large and positive (negative) values of $\psi_{p\ell}$ are stopping motorists with characteristics that predict high (low) search rates (e.g. non-white, low-income men with criminal histories) relative to other troopers making stops at the same times and in the same locations. We rank trooper by location combinations by $|\psi_{p\ell}|$, where combinations with the largest absolute values are 'most selective'.

In 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 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.

B.6.2. Troopers Vary in Screening Ability

Our finding that average and marginal hit rates are similar is consistent with Knowles et al. (2001), who develop an equilibrium model where troopers decide whether or not to search motorists and motorists decide whether or not to carry contraband. They show that if troopers are not racially biased, all motorists must, in equilibrium, carry contraband with equal probability. In this model there is no inframarginality problem because there is no difference between hit rates for the marginal and average searched motorists.

However, there are at least two features of our setting that are inconsistent with the Knowles et al. (2001) framework. First, as we document in Section V.D, we find little evidence that motorists respond to variation in search risk by adjusting contraband carrying rates, at least in the range of search rates we observe. Second, as we show in this section, troopers vary systematically in their hit rates, implying variation in screening ability. This is inconsistent with Knowles et al. (2001), which assumes there is no screening.

Figure B16 documents that troopers vary in screening ability in two ways. Panel A plots adjusted search rates $\tilde{s}_{p\ell}$ against adjusted unconditional hit rates $\tilde{h}_{p\ell}$ for each trooper by location combination. Conditional on search rate, there is significant variation in unconditional hit rates. This variation is not due to statistical noise alone. Panel B plots trooper by location hit rates in one randomly selected half of stops against the same trooper by location hit rate in the remaining half of stops. The estimated slope is 0.376, indicating that while some variation in hit rates is due to chance, there is systematic variation in hit rates across trooper by location combinations.

B.7. Constructing a Feasible Search Productivity Curve

In this section we show that the search behavior of troopers with the highest search rates implies that troopers could, in principle, achieve higher hit rates and lower search rates. The basic argument is as follows: consider two groups of motorists, group A and group B, that are both searched with some probability. Suppose the observed hit rate is higher for group A than group B. Then troopers can increase their hit rate by searching group A as before, but not searching group B.

We take the search behavior of troopers in the top quartile by search rate. We use equation (10) to construct the probability of search for each motorist as a function of that motorist's observable characteristics, $P(SEARCH_{it}|X_{it})$. We construct a predicted hit rate for each motorist analogously. We then construct a hypothetical SPC as follows. We first rank motorists by their predicted hit rate. We then allocate searches by this rank, so that at the lowest trooper search rates, troopers search only motorists with the highest predicted hit rate. Troopers search the motorists with the highest predicted hit rate up to probability $P(SEARCH_{it}|X_{it})$, and then move on to motorists with the next highest predicted hit rate.

The SPC we construct through this procedure is show in Figure B19. At the average search rate, 1.1%, the hit rate is about 20% higher than the observed hit rate (39% versus 32%).

B.7.1. Racial Search Disparities by Trooper Race

In this section we examine differences in search behavior by trooper race. We identify trooper race using 2015 personnel records for 2,469 troopers accounting for 84% of stops. Table B18 documents search rates and hit rates by both motorist and trooper race.

We next measure differences in black-white and Hispanic-white search odds ratios by trooper race, accounting for stop and other motorist characteristics. We estimate logistic regression models analogous to equation (B.1) that include fixed effects for trooper race and interactions between motorist and trooper race. We limit the analysis to stops conducted by troopers that we identify as black, Hispanic, or white.

Table B19 presents coefficient estimates, where columns (1) through (5) are analogous to the same columns in Table B2. The black-white search disparity for black troopers is about 20% smaller than the same disparity for white troopers, and about 35% smaller than the same disparity for Hispanic troopers. The Hispanic-white disparity is similar for white and Hispanic troopers and smaller for black troopers.

B.8. Appendix B Tables

TABLE B1
DETAILED SEARCH OUTCOMES BY MOTORIST RACE

	Black	Hispanic	White	All
Consent	45.61	58.08	49.01	52.16
Incident to Arrest	4.216	5.171	5.062	4.948
Inventory	8.262	11.80	11.97	11.20
Probable Cause	41.91	24.95	33.96	31.69
Conditional on Contraband:				
Currency	0.511	1.148	0.170	0.568
Drugs	56.09	48.54	50.89	51.14
Weapon	5.020	2.743	3.606	3.599
Other	38.38	47.57	45.34	44.69
Arrest	23.20	23.11	24.54	23.79
Felony Arrest	8.837	7.549	6.820	7.471
Charge Severity (Days)				
Mean	91.37	83.49	83.59	85.12
90th Percentile	192.31	188.04	224.82	212.50

This table summarizes detailed search outcomes by motorist race. 'Charge Severity' refers to the average incarceration sentence associated with conviction for that arrest charge. All outcome values, excluding 'Charge Severity', are expressed as percentage points. Charge Severity is set to zero for searches that do not lead to an arrest.

Table B2
What Predicts Odds of Searches and Contraband Yield?

Outcome:	Motorist/Vehicle Searched					Contraband Found | Search				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black	2.99	2.64	2.16	1.84	1.81	0.86	0.84	0.91	0.90	0.89
	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Hispanic	1.57	1.68	1.45	1.44	1.42	0.58	0.68	0.73	0.73	0.73
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Female	0.38	0.39	0.42	0.51	0.50	0.99	0.99	0.99	1.02	1.01
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Log median income			0.68	0.74	0.74			1.29	1.29	1.29
			(0.01)	(0.01)	(0.01)			(0.03)	(0.03)	(0.03)
Expected log income given			0.65	0.68	0.69			1.01	1.01	1.01
vehicle (standardized)			(0.00)	(0.00)	(0.00)			(0.01)	(0.01)	(0.01)
1–2 prior non-drug arrests				1.82	1.85				1.01	1.02
				(0.03)	(0.03)				(0.03)	(0.03)
3+ prior non-drug arrests				2.12	2.13				0.86	0.87
				(0.04)	(0.04)				(0.03)	(0.03)
1 prior drug arrest				3.00	2.84				1.47	1.47
				(0.05)	(0.05)				(0.05)	(0.05)
2+ prior drug arrest				3.96	3.56				1.43	1.42
				(0.07)	(0.07)				(0.05)	(0.05)
Prior stop, no search					0.70					0.90
					(0.01)					(0.02)
Prior search, no contraband					1.62					0.67
					(0.04)					(0.04)
Prior search, contraband					2.39					1.32
					(0.07)					(0.08)
Time FEs		✓	✓	✓	✓		✓	✓	✓	✓
Location FEs		✓	✓	✓	✓		✓	✓	✓	✓
Month FEs		✓	✓	✓	✓		✓	✓	✓	✓
White Mean	0.755	0.755	0.755	0.755	0.755	36.51	36.51	36.51	36.51	36.51
Observations	5,011,549	5,011,549	5,011,549	5,011,549	5,011,549	52,932	52,932	52,932	52,932	52,932

This table presents odds ratio estimates for the logistic regression equation (B.1). In columns (1) through (5) the outcome is SEARCH_{it}, an indicator of whether the stop of motorist i at time t led to a search. In columns (6) through (10) the outcome is CONTRABAND_{it}, an indicator for whether a search yields contraband. For these specifications, the sample is limited to stops that result in a search (i.e., where $\text{SEARCH}_{it} = 1$). Standard errors are clustered at the motorist level.

Table B3
Traffic Stop Descriptive Statistics, Pooled SPC Sample

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	10.11	100	0	0	20.43
% Hispanic	0	100	0	35.29	0	100	0	39.93
% White	0	0	100	54.60	0	0	100	39.64
% Female	39.15	32.65	37.97	36.21	17.07	15.73	21.11	18.14
Log Median Income	10.71	10.73	10.96	10.85	10.59	10.62	10.89	10.72
	(0.501)	(0.496)	(0.466)	(0.494)	(0.489)	(0.470)	(0.471)	(0.493)
Expected Log Income Given	-0.150	-0.068	0.082	0.005	-0.521	-0.458	-0.368	-0.435
Vehicle (Standardized)	(1.022)	(0.979)	(1.019)	(1.009)	(0.887)	(0.850)	(0.895)	(0.878)
Stop History (%):								
No Prior Stops	61.88	59.43	57.33	58.53	65.21	63.07	59.22	61.98
Prior Stop, No Search	36.48	39.19	41.82	40.36	27.61	31.37	34.54	31.86
Prior Search, No Contraband	1.079	1.065	0.535	0.777	3.989	3.497	3.224	3.489
Prior Search, Contraband	0.558	0.311	0.310	0.335	3.196	2.064	3.008	2.670
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	87.22	89.45	93.41	91.39	65.24	72.12	71.75	70.57
1-2 Prior Non-Drug Arrests	7.162	6.906	4.369	5.547	14.82	14.79	14.52	14.69
3+ Prior Non-Drug Arrests	5.620	3.640	2.217	3.063	19.93	13.09	13.73	14.74
Drug Arrest History (%):								
No Prior Drug Arrests	94.03	96.33	97.48	96.73	73.55	82.23	79.42	79.34
1 Prior Drug Arrest	2.780	2.149	1.408	1.808	10.28	8.837	9.515	9.401
2+ Prior Drug Arrests	3.188	1.521	1.109	1.464	16.17	8.930	11.06	11.26
Search Rate (%)	2.245	1.257	0.806	1.111	100	100	100	100
Unconditional Hit Rate (%)	0.751	0.335	0.303	0.360	33.22	26.41	37.28	32.11
Observations	392,980	1,371,919	2,112,950	3,887,849	8,824	17,245	17,120	43,189

Sample restrictions are described in Section II and Section IV.B. 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 B4
TRAFFIC STOP DESCRIPTIVE STATISTICS, RACE-SPECIFIC SPC SAMPLE

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	12.00	100	0	0	23.76
% Hispanic	0	100	0	26.50	0	100	0	31.71
% White	0	0	100	61.50	0	0	100	44.53
% Female	40.54	33.21	39.11	37.72	17.04	15.45	21.26	18.41
Log Median Income	10.70	10.81	10.96	10.89	10.59	10.68	10.89	10.75
	(0.500)	(0.495)	(0.469)	(0.489)	(0.487)	(0.474)	(0.473)	(0.494)
Expected Log Income Given	-0.163	-0.018	0.077	0.023	-0.530	-0.430	-0.370	-0.427
Vehicle (Standardized)	(1.023)	(1.019)	(1.026)	(1.027)	(0.881)	(0.875)	(0.902)	(0.891)
Stop History (%):								
No Prior Stops	61.32	64.21	57.75	59.89	64.27	67.54	59.15	63.03
Prior Stop, No Search	37.02	34.80	41.40	39.13	28.31	28.02	34.34	30.90
Prior Search, No Contraband	1.093	0.707	0.537	0.649	4.095	2.531	3.429	3.303
Prior Search, Contraband	0.568	0.291	0.316	0.340	3.333	1.904	3.078	2.767
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	87.19	90.78	93.45	91.99	65.52	73.79	71.86	70.96
1-2 Prior Non-Drug Arrests	7.229	6.236	4.354	5.198	15.44	14.38	14.66	14.76
3+ Prior Non-Drug Arrests	5.584	2.981	2.196	2.811	19.04	11.83	13.49	14.28
Drug Arrest History (%):								
No Prior Drug Arrests	94.08	96.89	97.47	96.91	73.76	83.00	79.16	79.09
1 Prior Drug Arrest	2.780	1.810	1.414	1.683	10.51	8.434	9.658	9.472
2+ Prior Drug Arrests	3.139	1.296	1.116	1.407	15.74	8.569	11.18	11.44
Search Rate (%)	2.199	1.330	0.805	1.111	100	100	100	100
Unconditional Hit Rate (%)	0.754	0.408	0.303	0.385	34.05	30.47	37.46	34.43
Observations	304,209	671,471	1,558,483	2,534,163	6,691	8,928	12,539	28,158

Sample restrictions are described in Section II and Section IV.B. 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 B5
Motorist Selection into Stops by Trooper Unconditional Hit Rate

				Excluding Most Selective Troopers		
	$100 \times \text{CONTRABAND}_{it}$	$100 \times h_{p\ell}^{-it}$	$100 \times \tilde{h}_{p\ell}^{-it}$	$100 \times \text{CONTRABAND}_{it}$	$100 \times h_{p\ell}^{-it}$	$100 \times \tilde{h}_{p\ell}^{-it}$
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.263	0.043	0.031	0.223	0.019	0.012
	(0.024)	(0.009)	(0.009)	(0.023)	(0.005)	(0.005)
Hispanic	0.015	0.011	0.007	0.015	0.001	-0.001
	(0.008)	(0.003)	(0.003)	(0.008)	(0.003)	(0.002)
Female	-0.169	-0.008	-0.006	-0.153	-0.004	-0.004
	(0.009)	(0.002)	(0.002)	(0.009)	(0.001)	(0.001)
Log Median Income	-0.040	-0.002	-0.002	-0.032	0.001	-0.000
	(0.008)	(0.002)	(0.002)	(0.008)	(0.002)	(0.002)
Expected Log Income Given	-0.096	-0.011	-0.010	-0.079	-0.005	-0.005
Vehicle (Standardized)	(0.005)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)
1-2 Prior Non-Drug Arrests	0.300	0.006	0.005	0.261	0.000	-0.000
	(0.026)	(0.002)	(0.002)	(0.027)	(0.001)	(0.001)
3+ Prior Non-Drug Arrests	0.434	0.010	0.009	0.377	0.003	0.003
	(0.047)	(0.003)	(0.003)	(0.048)	(0.002)	(0.002)
1 Prior Drug Arrest	1.535	0.014	0.012	1.376	0.003	0.002
	(0.083)	(0.003)	(0.003)	(0.080)	(0.002)	(0.002)
2+ Prior Drug Arrests	2.328	0.016	0.014	1.981	0.005	0.004
	(0.119)	(0.004)	(0.004)	(0.111)	(0.003)	(0.003)
Prior Stop, No Search	-0.156	-0.007	-0.004	-0.132	-0.004	-0.002
	(0.009)	(0.002)	(0.001)	(0.009)	(0.001)	(0.001)
Prior Search, No Contraband	0.247	0.012	0.010	0.310	0.002	0.002
	(0.066)	(0.004)	(0.004)	(0.071)	(0.003)	(0.003)
Prior Search, Contraband	2.457	0.028	0.028	2.164	0.013	0.014
	(0.194)	(0.008)	(0.008)	(0.198)	(0.007)	(0.007)
Location by Time FEs	✓	✓	✓	✓	✓	✓
Month FEs	✓	✓	✓	✓	✓	✓
Highway by Location FEs	✓	✓	✓	✓	✓	✓
Joint F-Statistic	54.75	9.07	8.04	49.45	6.07	5.41
Observations	3,887,758	3,887,758	3,887,758	3,266,320	3,266,320	3,266,320

This table presents coefficients from estimates of equation (7), where we replace the outcome SEARCH_{it} with CONTRABAND_{it} in column (1). In columns (2) and (3), we replace the outcome CONTRABAND_{it} with $h_{p(i,t)\ell(i,t)}^{-it}$ and $\tilde{h}_{p(i,t)\ell(i,t)}^{-it}$, leave-out trooper unconditional hit rates corresponding to the trooper who conducted the stop. SEARCH_{it} and CONTRABAND_{it} are defined as indicator variables and $h_{p(i,t)\ell(i,t)}^{-it}$ takes on values between zero and one. $\tilde{h}_{p(i,t)\ell(i,t)}^{-it}$ 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 motorist level. 'Joint F-Statistic' refers to an F-test for the joint significance of all motorist characteristics.

Table B6
Robustness of Search Productivity Curve Slope Estimates

	Pooled	White Motorists	Black Motorists	Hispanic Motorists
Unadjusted Rates	0.330	0.378	0.383	0.317
	(0.015)	(0.027)	(0.024)	(0.020)
Covariate-Adjusted Rates	0.328	0.378	0.381	0.316
	(0.015)	(0.026)	(0.025)	(0.019)
EB-Adjusted Rates	0.331	0.377	0.385	0.314
	(0.015)	(0.029)	(0.024)	(0.020)
Split Sample (2SLS), First	0.332	0.388	0.371	0.310
	(0.017)	(0.034)	(0.027)	(0.025)
Split Sample (2SLS), Second	0.324	0.365	0.401	0.332
	(0.017)	(0.031)	(0.034)	(0.025)

This table presents the slope of the relationship between trooper search rates and unconditional hit rates conditional on location fixed effects for several specifications and for varying samples of motorists. Trooper-by-location combinations are weighted by number of stops. For the split sample models, 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 $\tilde{s}_{p\ell}$, instrumenting for $\tilde{s}_{p\ell}$ using its pair estimate from the other sample.

Table B7
Traffic Stop Descriptive Statistics, Pooled Within-Motorist Sample

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	9.238	100	0	0	18.91
% Hispanic	0	100	0	32.93	0	100	0	37.78
% White	0	0	100	57.83	0	0	100	43.32
% Female	34.71	27.35	33.58	31.63	14.48	12.84	17.85	15.32
Log Median Income	10.68	10.70	10.92	10.83	10.59	10.61	10.88	10.72
	(0.492)	(0.482)	(0.446)	(0.475)	(0.481)	(0.474)	(0.445)	(0.482)
Expected Log Income Given	-0.190	-0.085	0.052	-0.016	-0.529	-0.425	-0.343	-0.409
Vehicle (Standardized)	(0.996)	(0.936)	(0.961)	(0.960)	(0.909)	(0.844)	(0.868)	(0.869)
Stop History (%):								
No Prior Stops	49.39	46.29	46.60	46.75	49.86	45.60	48.52	47.67
Prior Stop, No Search	48.45	51.76	52.41	51.83	40.41	45.88	43.77	43.93
Prior Search, No Contraband	1.526	1.570	0.649	1.033	6	5.799	4.124	5.111
Prior Search, Contraband	0.632	0.384	0.344	0.384	3.724	2.727	3.582	3.286
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	84.94	86.36	91.88	89.42	58.14	66.62	67.52	65.41
1-2 Prior Non-Drug Arrests	8.666	8.953	5.465	6.909	18.55	18.26	18.06	18.23
3+ Prior Non-Drug Arrests	6.393	4.683	2.656	3.669	23.31	15.12	14.42	16.36
Drug Arrest History (%):								
No Prior Drug Arrests	93.45	95.48	97.13	96.25	68.62	77.29	76.82	75.45
1 Prior Drug Arrest	3.148	2.712	1.622	2.122	11.52	10.80	10.20	10.68
2+ Prior Drug Arrests	3.401	1.811	1.244	1.630	19.86	11.91	12.97	13.87
Search Rate (%)	1.709	0.958	0.625	0.835	100	100	100	100
Unconditional Hit Rate (%)	0.480	0.214	0.213	0.238	27.86	22.23	33.77	28.30
Observations	84,856	302,529	531,214	918,599	1,450	2,897	3,322	7,669

Sample restrictions are described in Section II, Section IV.B, and Section IV.E.1. These statistics refer to the first stop for each sequential pair of stops included in the analysis. 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 B8
Traffic Stop Descriptive Statistics, Race-Specific Within-Motorist Sample

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	11.70	100	0	0	23.60
% Hispanic	0	100	0	20.49	0	100	0	23.19
% White	0	0	100	67.81	0	0	100	53.21
% Female	37.29	30.23	35.81	34.84	14.40	14.76	18.30	16.56
Log Median Income	10.65	10.81	10.89	10.85	10.57	10.76	10.88	10.78
	(0.486)	(0.475)	(0.443)	(0.462)	(0.478)	(0.460)	(0.446)	(0.473)
Expected Log Income Given	-0.223	-0.042	0.033	-0.013	-0.557	-0.364	-0.354	-0.404
Vehicle (Standardized)	(0.990)	(0.983)	(0.964)	(0.975)	(0.887)	(0.902)	(0.870)	(0.885)
Stop History (%):								
No Prior Stops	48.15	50.90	46.64	47.69	48.77	51.10	48.43	49.13
Prior Stop, No Search	49.64	47.85	52.34	51.10	41.05	43.04	43.22	42.67
Prior Search, No Contraband	1.578	0.879	0.662	0.814	6.070	2.723	4.655	4.541
Prior Search, Contraband	0.630	0.374	0.359	0.393	4.115	3.141	3.697	3.667
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	84.88	88.67	91.97	90.46	57.20	70.05	67.82	65.83
1-2 Prior Non-Drug Arrests	8.738	7.749	5.384	6.261	20.68	16.75	17.66	18.16
3+ Prior Non-Drug Arrests	6.384	3.579	2.649	3.277	22.12	13.19	14.51	16.00
Drug Arrest History (%):								
No Prior Drug Arrests	93.57	96.50	97.12	96.58	68.83	76.86	76.27	74.65
1 Prior Drug Arrest	3.169	2.017	1.621	1.883	12.76	10.37	10.63	11.07
2+ Prior Drug Arrests	3.258	1.481	1.261	1.539	18.42	12.77	13.10	14.28
Search Rate (%)	1.632	0.916	0.635	0.809	100	100	100	100
Unconditional Hit Rate (%)	0.465	0.267	0.215	0.255	28.40	29.11	33.68	31.37
Observations	59,554	104,304	345,237	509,095	972	955	2,191	4,118

Sample restrictions are described in Section II, Section IV.B, and Section IV.E.1. These statistics refer to the first stop for each sequential pair of stops included in the analysis. 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 B9
TRAFFIC STOP DESCRIPTIVE STATISTICS, SPATIAL RD SAMPLE

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	11.17	100	0	0	23.06
% Hispanic	0	100	0	33.22	0	100	0	38.65
% White	0	0	100	55.61	0	0	100	38.29
% Female	39.62	33.80	39.00	37.35	15.84	14.77	21.17	17.47
Log Median Income	10.73	10.76	10.98	10.88	10.60	10.64	10.92	10.74
	(0.503)	(0.509)	(0.472)	(0.502)	(0.488)	(0.478)	(0.468)	(0.498)
Expected Log Income Given	-0.128	-0.028	0.115	0.040	-0.519	-0.437	-0.353	-0.424
Vehicle (Standardized)	(1.036)	(1.017)	(1.048)	(1.040)	(0.876)	(0.870)	(0.925)	(0.895)
Stop History (%):								
No Prior Stops	62.68	61.68	58.70	60.14	66.07	64.64	59.68	63.07
Prior Stop, No Search	35.70	37.03	40.46	38.79	26.82	30.69	33.73	30.96
Prior Search, No Contraband	1.084	0.996	0.531	0.747	3.954	2.964	3.281	3.314
Prior Search, Contraband	0.542	0.290	0.304	0.326	3.153	1.705	3.313	2.655
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	87.66	90.34	93.52	91.81	65.54	74.90	72.99	72.01
1-2 Prior Non-Drug Arrests	6.911	6.392	4.330	5.303	13.92	13.48	13.46	13.58
3+ Prior Non-Drug Arrests	5.433	3.270	2.149	2.888	20.55	11.62	13.54	14.41
Drug Arrest History (%):								
No Prior Drug Arrests	94.38	96.66	97.50	96.87	75.63	83.35	78.46	79.70
1 Prior Drug Arrest	2.662	1.982	1.412	1.741	9.057	8.367	9.844	9.091
2+ Prior Drug Arrests	2.954	1.363	1.091	1.389	15.31	8.287	11.69	11.21
Search Rate (%)	2.249	1.268	0.751	1.090	100	100	100	100
Unconditional Hit Rate (%)	0.742	0.302	0.280	0.339	32.62	23.67	36.95	30.82
Observations	166,435	494,850	828,268	1,489,553	3,743	6,275	6,217	16,235

Sample restrictions are described in Section II and Section IV.E.2. 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 B10
Traffic Stop Descriptive Statistics, Trooper Reallocation Sample

	All Stops				All Searches			
	Black	Hispanic	White	All	Black	Hispanic	White	All
% Black	100	0	0	14.29	100	0	0	27.06
% Hispanic	0	100	0	26.29	0	100	0	31.12
% White	0	0	100	59.41	0	0	100	41.82
% Female	40.52	33.06	39.28	37.82	16.88	15.38	21.69	18.42
Log Median Income	10.70	10.81	10.95	10.88	10.59	10.68	10.89	10.74
	(0.500)	(0.496)	(0.471)	(0.491)	(0.488)	(0.474)	(0.472)	(0.495)
Expected Log Income Given	-0.161	-0.017	0.080	0.020	-0.529	-0.429	-0.364	-0.429
Vehicle (Standardized)	(1.025)	(1.023)	(1.029)	(1.030)	(0.884)	(0.877)	(0.906)	(0.894)
Stop History (%):								
No Prior Stops	61.45	64.26	57.69	59.95	64.04	67.86	59.31	63.25
Prior Stop, No Search	36.89	34.74	41.45	39.04	28.43	27.70	34.22	30.62
Prior Search, No Contraband	1.085	0.705	0.542	0.662	4.102	2.515	3.428	3.327
Prior Search, Contraband	0.573	0.289	0.319	0.348	3.427	1.928	3.041	2.799
Non-Drug Arrest History (%):								
No Prior Non-Drug Arrests	87.19	90.82	93.45	91.87	65.39	74.11	72.26	70.98
1-2 Prior Non-Drug Arrests	7.219	6.251	4.358	5.265	15.29	14.38	14.53	14.69
3+ Prior Non-Drug Arrests	5.588	2.929	2.187	2.868	19.32	11.51	13.21	14.34
Drug Arrest History (%):								
No Prior Drug Arrests	94.08	96.93	97.47	96.84	73.53	83.05	79.08	78.81
1 Prior Drug Arrest	2.780	1.792	1.416	1.710	10.58	8.380	9.836	9.584
2+ Prior Drug Arrests	3.140	1.279	1.116	1.448	15.89	8.571	11.09	11.60
Search Rate (%)	2.151	1.344	0.800	1.136	100	100	100	100
Unconditional Hit Rate (%)	0.738	0.419	0.306	0.397	34.08	30.92	38.00	34.74
Observations	295,816	544,101	1,229,572	2,069,489	6,362	7,315	9,831	23,508

Sample restrictions are described in Section II and Section V.A. 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 B11
HIGH AND LOW SEARCH RATE TROOPERS SEARCH OBSERVABLY SIMILAR MOTORISTS

Outcome: $SEARCH_{it}$	Trooper Search Rate Quartile			
	Q1	Q2	Q3	Q4
Female	0.43	0.50	0.48	0.53
	(0.04)	(0.02)	(0.01)	(0.01)
Black	1.99	1.65	1.76	1.69
	(0.22)	(0.08)	(0.05)	(0.03)
Hispanic	1.63	1.28	1.45	1.34
	(0.15)	(0.05)	(0.04)	(0.02)
Log Median Income	0.83	0.69	0.73	0.76
	(0.07)	(0.02)	(0.02)	(0.01)
Expected Log Income Given	0.68	0.69	0.70	0.70
Vehicle (Standardized)	(0.03)	(0.01)	(0.01)	(0.01)
Prior Stop, No Search	0.65	0.63	0.66	0.72
	(0.05)	(0.02)	(0.02)	(0.01)
Prior Search, No Contraband	1.30	1.50	1.80	1.64
	(0.28)	(0.14)	(0.10)	(0.06)
Prior Search, Contraband	2.23	2.26	2.38	2.50
	(0.52)	(0.24)	(0.17)	(0.12)
1-2 Prior Non-Drug Arrests	1.76	2.13	1.74	1.89
	(0.21)	(0.11)	(0.06)	(0.04)
3+ Prior Non-Drug Arrests	2.19	2.42	2.01	2.17
	(0.29)	(0.14)	(0.08)	(0.05)
1 Prior Drug Arrest	2.79	2.71	3.37	2.88
	(0.41)	(0.18)	(0.14)	(0.07)
2+ Prior Drug Arrests	3.76	3.35	4.02	3.61
	(0.55)	(0.22)	(0.17)	(0.10)
Time FEs	✓	✓	✓	✓
Location FEs	✓	✓	✓	✓
Observations	941,667	972,244	972,770	968,943

This table presents odds ratio estimates for the logistic regression model (10), separately by trooper search rate quartile. The outcome is $SEARCH_{it}$, an indicator of whether the stop of motorist i at time t led to a search. Standard errors are clustered at the motorist level.

Table B12
Trooper Descriptive Statistics

	All Matched	Black/Hispanic/White Troopers
% Black	9.71	9.90
% Hispanic	28.92	29.48
% White	59.48	60.62
% Male	95.83	95.84
Age	38.3	38.3
	(8.6)	(8.6)
Experience	10.3	10.3
	(5.8)	(5.9)
Stops per Hour	0.52	0.52
	(0.13)	(0.13)
Observations	2,327	2,283

This table presents descriptive statistics for the troopers we match to personnel records from 2015. Column (2) restricts to black, Hispanic, and white troopers. 'Age' is trooper age in 2015. 'Experience' is the difference in years between 2015 and the year the trooper began their position.

Table B13
Trooper Search and Unconditional Hit Rates by Observable Characteristics

Outcome:	Search Rate ($\tilde{s}_{p\ell}$) (%)	Unconditional Hit Rate ($\tilde{h}_{p\ell}$) (%)
	(1)	(2)
Experience: Q2	0.214	0.001
	(0.094)	(0.019)
Experience: Q3	0.067	-0.016
	(0.087)	(0.020)
Experience: Q4	0.009	0.010
	(0.090)	(0.021)
Stop Rate: Q2	0.140	-0.025
	(0.090)	(0.019)
Stop Rate: Q3	0.129	-0.026
	(0.082)	(0.021)
Stop Rate: Q4	0.386	-0.032
	(0.091)	(0.020)
Black	-0.390	-0.042
	(0.144)	(0.033)
Hispanic	-0.207	0.012
	(0.106)	(0.019)
Adjusted Search Rate		33.515
		(1.609)
Location FEs	✓	✓
Joint F-Test (P-Value)	0.000	0.449
DV Mean	1.280	0.396
Within R^2	0.023	0.656
Observations	2,283	2,283

This table presents coefficient estimates from simple regressions of adjusted trooper search rates ($\tilde{s}_{p\ell}$) on location fixed effects and trooper characteristics (column 1) and of adjusted trooper unconditional hit rates ($\tilde{h}_{p\ell}$) on location fixed effects, trooper characteristics, and adjusted trooper search rates (column 2). Standard errors are clustered at the trooper level. $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ take on values between zero and one before each is residualized. 'Joint Test' refers to a test for whether the coefficients on trooper characteristics, excluding adjusted search rate, are jointly significant.

Table B15
First Stage Heterogeneity in Within-Trooper Design

	SEARCH_{it}					
	Below Median Hit Rate			Above Median Hit Rate		
	(1)	(2)	(3)	(4)	(5)	(6)
$s^{-(i,t)}_{\ell(i,t)y(t)}$	0.307	0.297	0.219	0.300	0.265	0.116
	(0.057)	(0.056)	(0.059)	(0.040)	(0.039)	(0.050)
Trooper FEs	✓	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Motorist Controls		✓			✓	
Motorist FEs			✓			✓
Kleibergen-Paap F-Stat	29.19	28.25	13.75	56.50	47.16	5.33
Observations	1,331,935		510,335	1,269,522		462,341

This table presents estimates of equation (11) for troopers with below and above average hit rates on a leave out sample. We randomly split motorists into two samples (sample A and sample B). We measure trooper hit rates on sample A and divide troopers in half based on this hit rate. We randomly allocate troopers with no searches in sample A. We then estimate equation (11) using sample B, separately for each half of troopers. Trooper hit rates on sample B are 20% and 40%. $s^{-(i,t)}_{\ell(i,t)y(t)}$ denotes the search rate for all stops in location ℓ in the year corresponding to t, excluding stop (i,t). 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. $s^{-(i,t)}_{\ell(i,t)y(t)}$ takes on values between zero and one.

TABLE B16
COUNTERFACTUAL ARREST 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 (%)	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10
Arrest Rate (%)	7.71	8.08	8.31	8.12	8.32	8.13	8.52	8.26	8.58
		(0.27)	(0.31)	(0.30)	(0.34)	(0.27)	(0.32)	(0.26)	(0.31)
$H_0 : SQ = E$		0.103		0.104		0.017		0.054	
White Motorists									
Search Rate (%)	0.79	0.79	1.10	0.79	1.10	0.79	1.10	0.79	1.10
Arrest Rate (%)	8.57	9.03	9.03	9.02	8.89	9.05	9.23	9.33	9.36
		(0.45)	(0.45)	(0.45)	(0.45)	(0.43)	(0.45)	(0.43)	(0.44)
Black Motorists									
Search Rate (%)	2.22	2.22	1.10	2.22	1.10	2.22	1.10	2.22	1.10
Arrest Rate (%)	8.07	8.04	8.04	8.12	8.39	8.08	8.16	8.19	8.42
		(0.53)	(0.67)	(0.58)	(0.70)	(0.54)	(0.73)	(0.54)	(0.80)
Hispanic Motorists									
Search Rate (%)	1.31	1.31	1.10	1.31	1.10	1.31	1.10	1.31	1.10
Arrest Rate (%)	6.16	6.74	6.74	6.83	6.95	6.85	7.01	6.78	6.81
		(0.45)	(0.44)	(0.52)	(0.57)	(0.45)	(0.46)	(0.44)	(0.46)

This table presents observed and simulated counterfactual search rates and arrest rates by motorist racial group (expressed as percentage points). Observed search rates and arrest rates are calculated using the sample of stops used to estimate race-specific search productivity curves (SPCs). Sample criteria are described in Section IV.B. Simulated counterfactual search rates and arrest rates are calculated using racial group-specific SPC estimates from Section IV.D.

TABLE B17
COUNTERFACTUAL CHARGE SEVERITY

	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 (%)	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.10
Charge Severity	28.66	29.22	29.23	29.16	31.05	28.64	29.87	29.40	30.87
		(1.58)	(1.75)	(1.98)	(1.92)	(1.61)	(1.84)	(1.67)	(1.93)
$H_0 : SQ = E$		0.982		0.007		0.180		0.079	
White Motorists									
Search Rate (%)	0.80	0.80	1.10	0.80	1.10	0.80	1.10	0.80	1.10
Charge Severity	30.43	29.19	29.17	30.27	30.52	29.50	29.81	30.86	30.74
		(2.19)	(2.18)	(2.51)	(2.32)	(2.22)	(2.34)	(2.33)	(2.34)
Black Motorists									
Search Rate (%)	2.22	2.22	1.10	2.22	1.10	2.22	1.10	2.22	1.10
Charge Severity	31.92	32.91	35.82	29.65	39.58	29.74	37.88	29.58	37.44
		(3.27)	(6.88)	(5.02)	(7.03)	(3.39)	(6.83)	(3.32)	(6.48)
Hispanic Motorists									
Search Rate (%)	1.31	1.31	1.10	1.31	1.10	1.31	1.10	1.31	1.10
Charge Severity	23.56	26.34	26.34	27.16	28.34	26.52	26.31	27.17	28.15
		(2.60)	(2.72)	(2.92)	(3.05)	(2.63)	(2.78)	(2.68)	(2.93)

This table presents observed and simulated counterfactual search rates and charge severity by motorist racial group. Observed search rates and charge severity are calculated using the sample of stops used to estimate race-specific search productivity curves (SPCs). Sample criteria are described in Section IV.B. Simulated counterfactual search rates and charge severity are calculated using racial group-specific SPC estimates from Section IV.D. Search rates are expressed as percentage points.

Table B18
Search and Hit Rates by Motorist and Trooper Race

	White Troopers	Black Troopers	Hispanic Troopers
All Motorists			
Search Rate (%)	1.14	0.84	0.89
Hit Rate (%)	33.4	27.4	25.9
White Motorists			
Search Rate (%)	0.80	0.62	0.62
Hit Rate (%)	37.3	29.9	36.5
Black Motorists			
Search Rate (%)	2.36	1.52	2.20
Hit Rate (%)	34.1	28.2	31.0
Hispanic Motorists			
Search Rate (%)	1.40	0.86	0.96
Hit Rate (%)	28.4	23.8	19.4
Number of Troopers	1,465	216	745

This table presents search and hit rates by motorist and trooper race. We identify trooper race from 2015 personnel records.

Table B19
Racial Disparities in Search Rates by Trooper Race

Outcome:	Motorist/Vehicle Searched				
	(1)	(2)	(3)	(4)	(5)
Black	3.02	2.65	2.18	1.87	1.83
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Hispanic	1.68	1.68	1.46	1.47	1.44
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Black Trooper	0.77	0.75	0.76	0.75	0.75
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Hispanic Trooper	0.76	0.82	0.83	0.84	0.84
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Black × Black Trooper	0.84	0.85	0.84	0.84	0.84
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Black × Hispanic Trooper	1.20	1.17	1.16	1.15	1.15
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Hispanic × Black Trooper	0.80	0.75	0.75	0.76	0.75
	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Hispanic × Hispanic Trooper	0.89	1.00	0.97	0.94	0.94
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Female	0.39	0.40	0.42	0.51	0.50
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Time FEs		✓	✓	✓	✓
Location FEs		✓	✓	✓	✓
Month FEs		✓	✓	✓	✓
Income			✓	✓	✓
Arrest History				✓	✓
Stop History					✓
Observations	4,279,349	4,279,349	4,279,349	4,279,349	4,279,349

This table presents odds ratio estimates for the logistic regression model (B.1) augmented with fixed effects for trooper race and interactions between motorist and trooper race. We identify trooper race from 2015 personnel records. We limit to stops conducted by black, Hispanic, and white troopers. Standard errors are clustered at the motorist level.

TABLE B20
SEARCH DISPARITIES, CITATION DISPARITIES, AND LOCAL POLITICAL PREFERENCES

Black-White Gap

	Black-White Search Odds Ratio		
	(1)	(2)	(3)
Citation Odds Ratio	1.126		0.370
	(0.445)		(0.556)
Republican Vote Share		3.576	3.210
		(0.583)	(0.939)
Observations	83	83	83
Adjusted R^2	0.100	0.240	0.248

Hispanic-White Gap

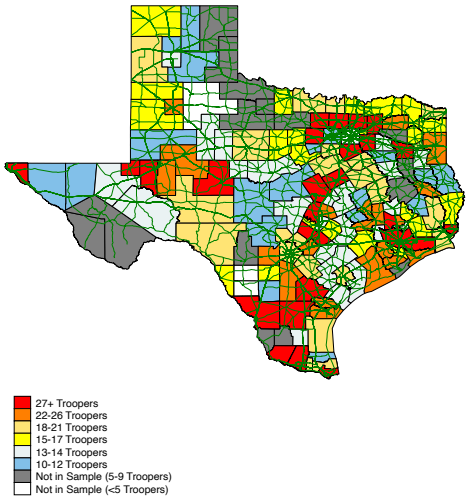
	Hispanic-White Search Odds Ratio		
	(4)	(5)	(6)
Citation Odds Ratio	0.540		0.330
	(0.213)		(0.236)
Republican Vote Share		1.077	0.860
		(0.316)	(0.348)
Observations	83	83	83
Adjusted R^2	0.062	0.092	0.112

This table presents estimates of linear regression models where the outcome is the sergeant area-specific black-white (Panel A) or Hispanic-white search odds ratio (Panel B) derived from equation (13). 'Citation Odds Ratio' refers to sergeant-area specific black-white (Panel A) or Hispanic-white citation odds ratio (Panel B) derived from equation (13) where the outcome is replaced with an indicator for whether the stop results in a citation rather than a warning. 'Republican Vote Share' refers to the Republican vote share 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. Robust standard errors are reported in parentheses.

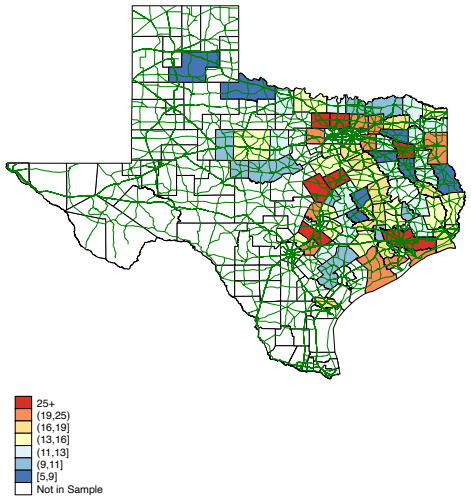
B.9. Appendix B Figures

FIGURE B1
NUMBER OF TROOPERS SATISFYING SAMPLE CRITERIA BY SERGEANT AREA

(a) Pooled SPC Criteria

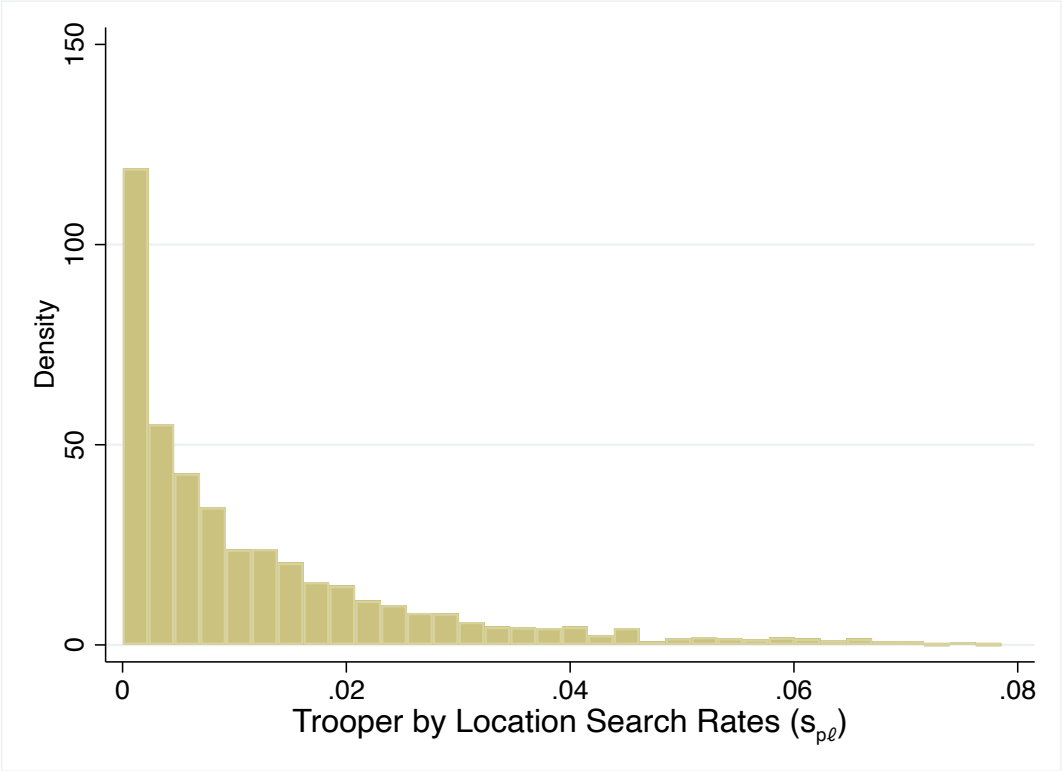


(b) Race-Specific SPC Criteria



Note: These maps depict the number of troopers in each sergeant area that satisfy the sample criteria described in Section IV.B. State and interstate highways are superimposed in green. Panel A depicts the number of troopers who satisfy sample criteria for estimating the pooled search productivity curve (SPC). We include sergeant areas in the analysis if they have at least ten troopers meeting the sample criteria. For sergeant areas included in the estimation of race-specific SPCs, Panel B depicts the number of troopers who satisfy the sample criteria, averaging across motorist racial groups (white, black, Hispanic). Sergeant areas included in the estimation of race-specific SPCs have at least five troopers meeting the sample criteria for each motorist racial group.

FIGURE B2
DISTRIBUTION OF SEARCH RATES ACROSS TROOPERS AND LOCATIONS

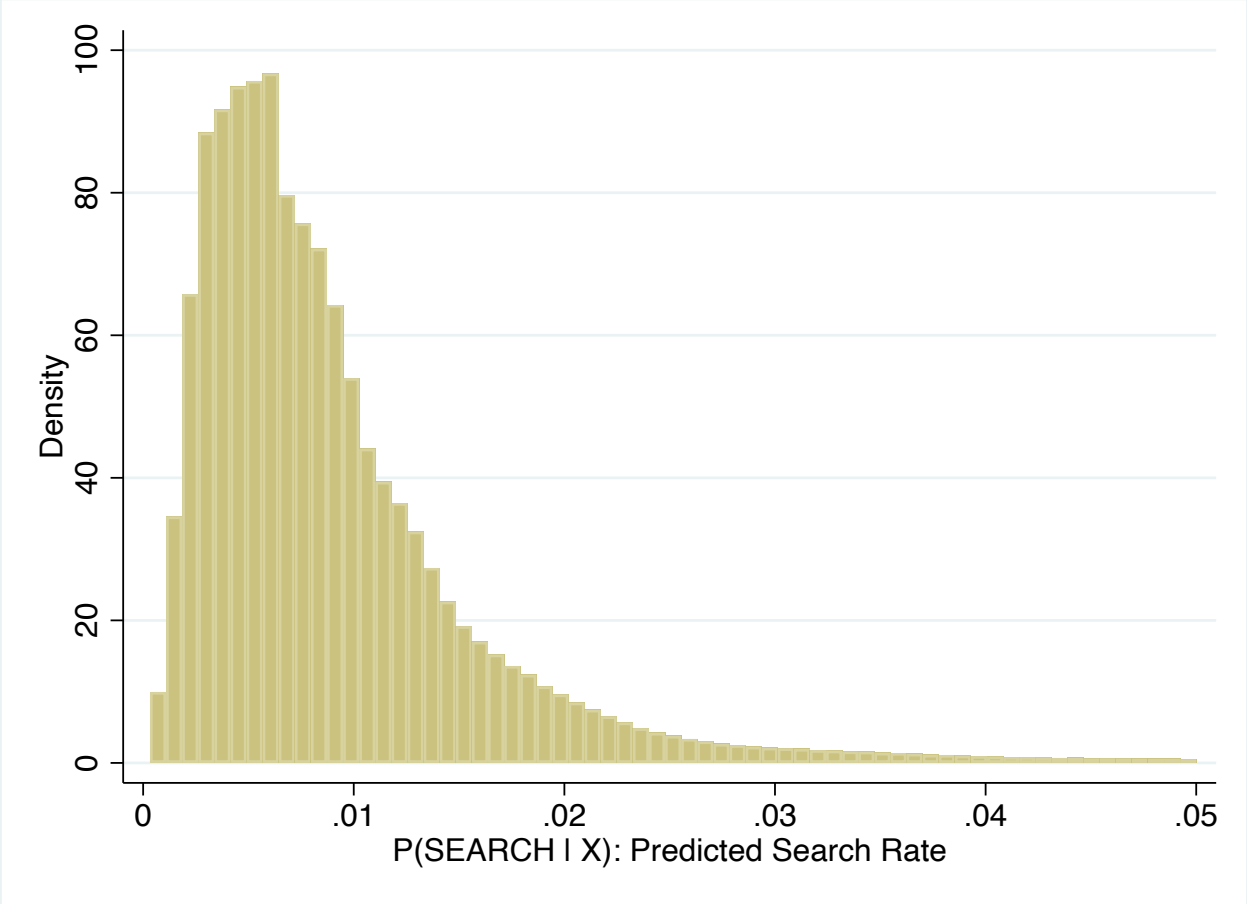


Note: This figure plots the distribution of search rates across trooper-by-location combinations ($s_{p\ell}$). $s_{p\ell}$ takes on values between zero and one. Sample restrictions are described in Section IV.B.

FIGURE B3
COMPARING ESTIMATES OF TROOPER SEARCH RATES

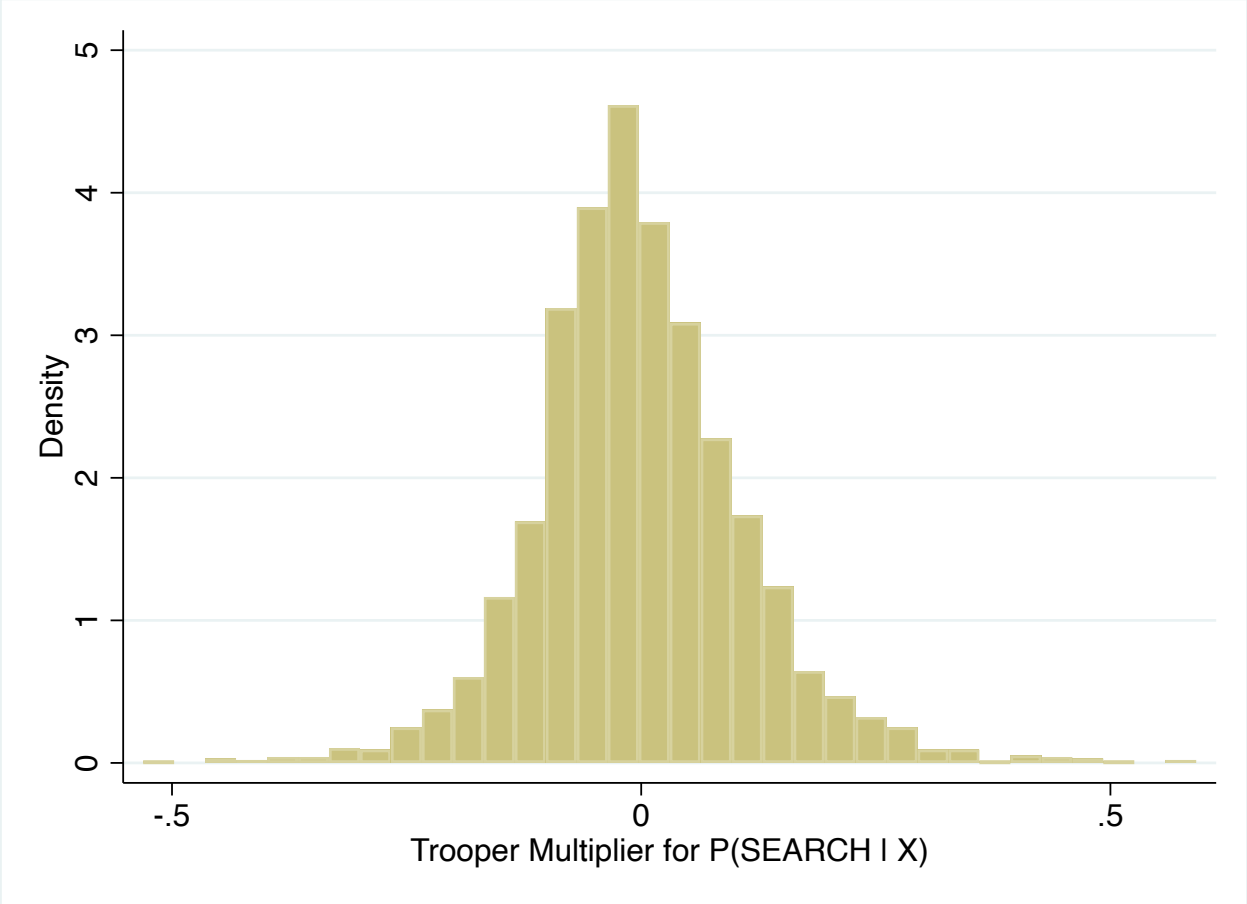
Note: This figure compares unadjusted trooper search rates ($s_{p\ell}$) to estimates of trooper search rates that adjust for additional stop and motorist characteristics ($\tilde{s}_{p\ell}$). $s_{p\ell}$ and $\tilde{s}_{p\ell}$ take on values between zero and one (before the latter is residualized). Stops characteristics include the month and specific highway of the stop. 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. The construction of trooper search rates is described in Section IV.B. The red dashed line is a 45° line. The slope of the best fit line is 0.99.

FIGURE B4
DISTRIBUTION OF $P(\text{SEARCH}|X_{it})$ ACROSS STOPS



Note: This figure plots the distribution of P(SEARCH | X_{it}), the search probability for each stop based on observable motorist characteristics. P(SEARCH | X_{it}) is estimated from equation (B.2) described in Section B.6.1 and takes on values between zero and one.

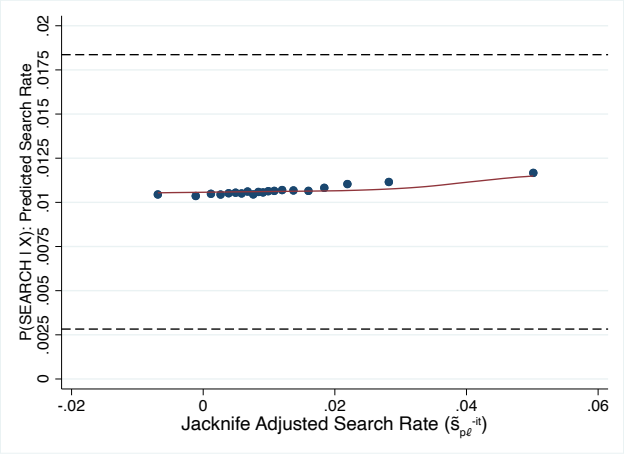
FIGURE B5
DISTRIBUTION OF MOTORIST SELECTION ACROSS TROOPERS



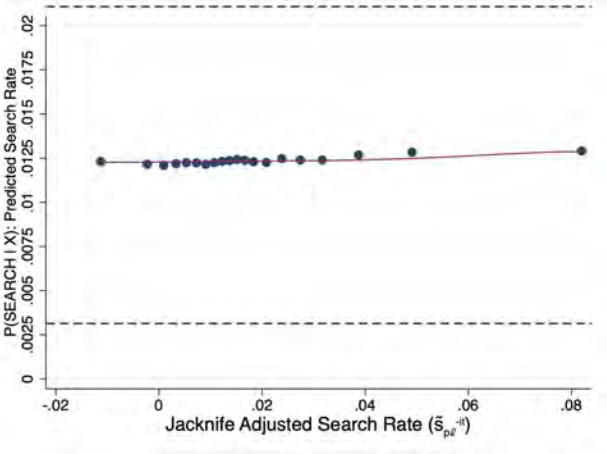
Note: This figure plots the distribution of $\psi_{p\ell}$ estimates derived from equation (B.3), described in more detail in Section B.6.1. The $\psi_{p\ell}$ terms are trooper by location fixed effects from a Poisson regression model for $P(\text{SEARCH}|X_{it})$, the search probability for each stop based on observable motorist characteristics. They summarize the degree to which motorist characteristics for those stopped by a given trooper in a given location deviate from what one would expect based on the time and location of their stops alone.

FIGURE B6
TROOPER CHARACTERISTICS AND MOTORIST COMPOSITION

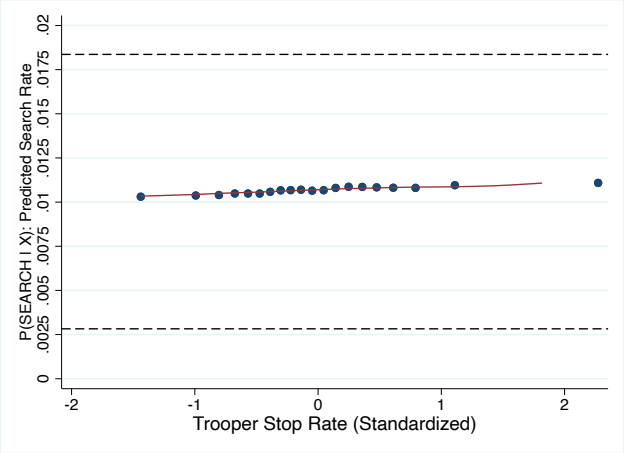
(a) By Search Rate



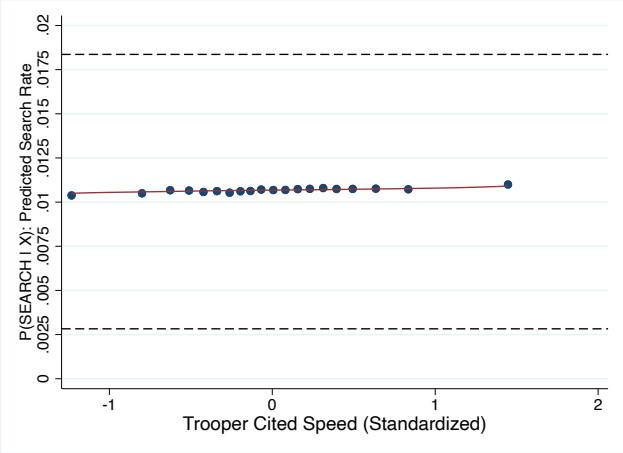
(b) By Search Rate at Night



(c) By Time Between Stops

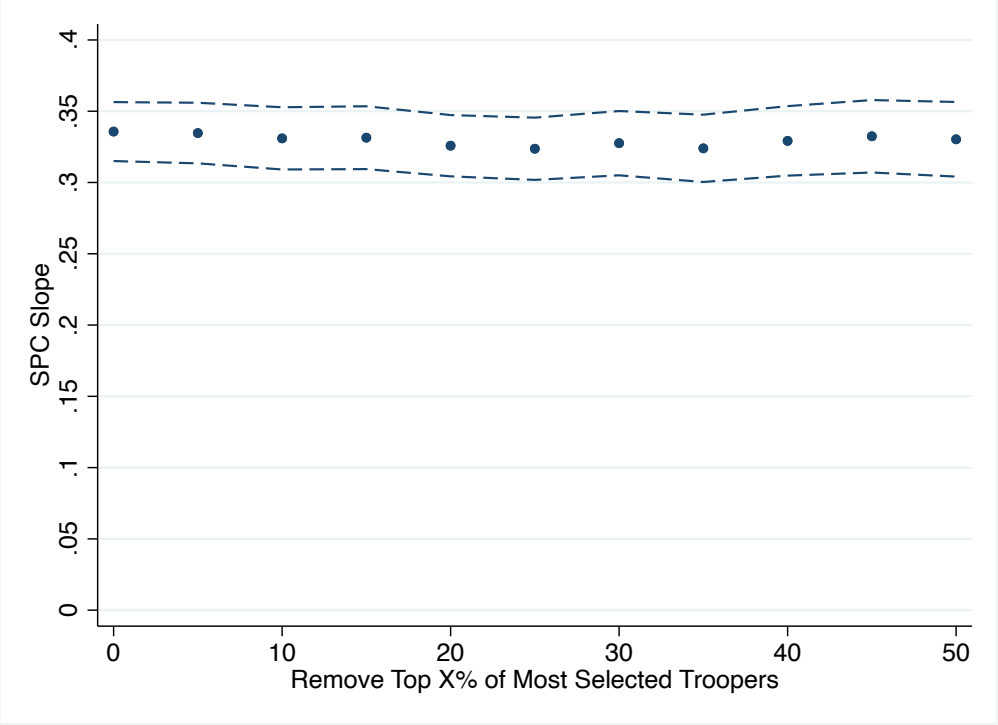


(d) By Cited Speed



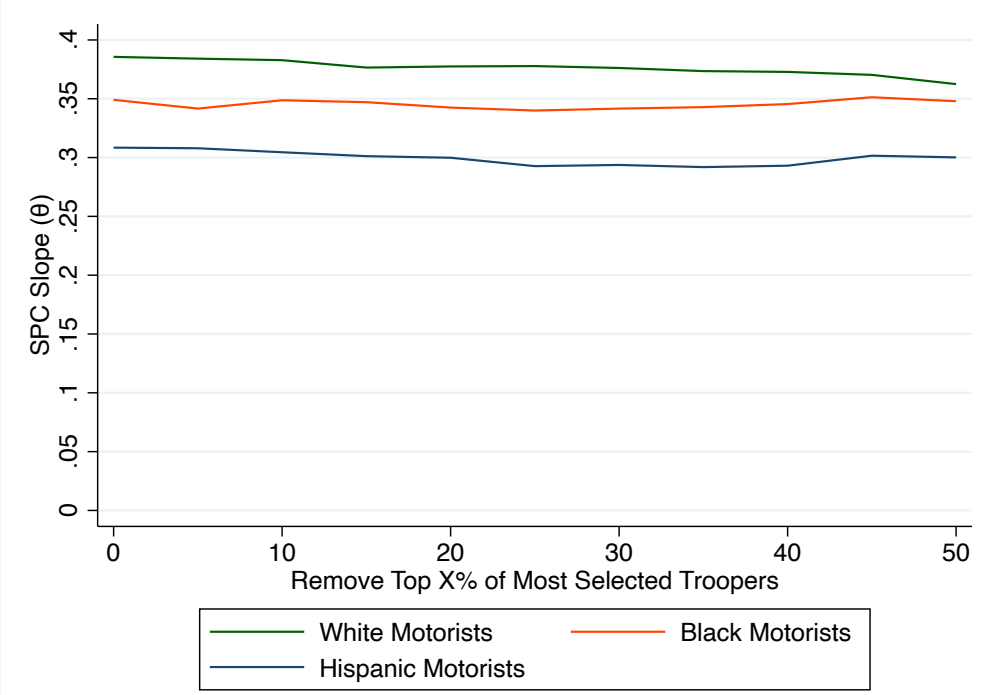
Note: These figures plot $P(\text{SEARCH}|X_{it})$, the search probability for each stop based on observable motorist characteristics, as a function of four trooper characteristics: search rate (Panel A), search rate at night (Panel B), time between stops (Panel C), and average cited speeds (Panel D). $P(\text{SEARCH}|X_{it})$ takes on values between zero and one. We partial out location by time fixed effects for both $P(\text{SEARCH}|X_{it})$ and trooper characteristics. In Panel B we restrict to stops made between 9pm and 5am.

Figure B7
Stability of Pooled Search Productivity Curve Slope For Varying Trooper Exclusions



Note: This figure plots the slope of the relationship between trooper search rates ($\tilde{s}^r_{p\ell}$) and unconditional hit rates ($\tilde{h}^r_{p\ell}$) for varying samples of troopers. For varying X, we remove the X% of troopers with compositions of stopped motorists that 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 Section B.6.1.

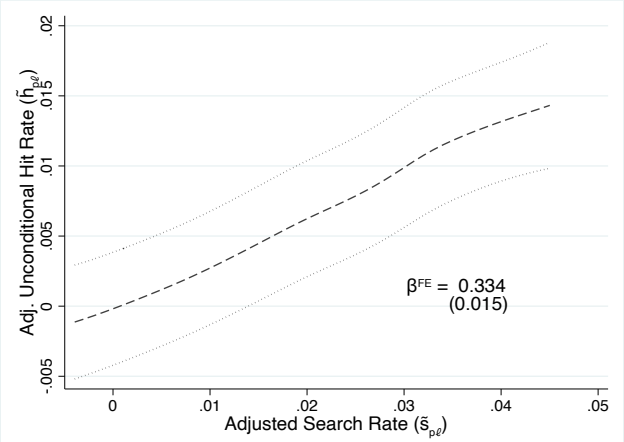
FIGURE B8
STABILITY OF RACE-SPECIFIC SEARCH PRODUCTIVITY CURVE SLOPES FOR VARYING TROOPER EXCLUSIONS



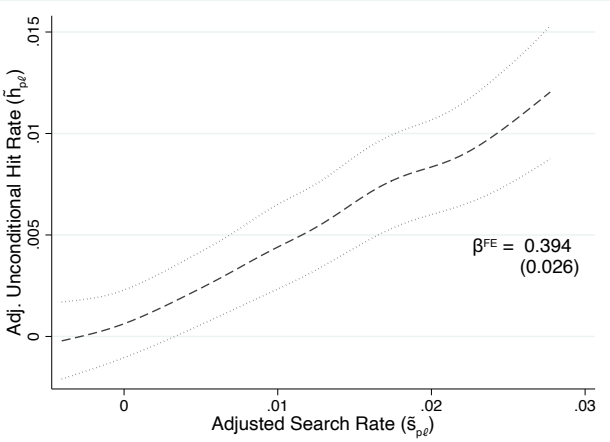
Note: This figure plots the slope of the relationship between trooper search rates ($\tilde{s}^r_{p\ell}$) and unconditional hit rates ($\tilde{h}^r_{p\ell}$) by motorist race and for varying samples of troopers. For varying X, we remove the $X\%$ of troopers with compositions of stopped motorists that 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 Appendix B.6.1.

FIGURE B9
BETWEEN-TROOPER SEARCH PRODUCTIVITY CURVES, LOCATION FIXED EFFECTS APPROACH

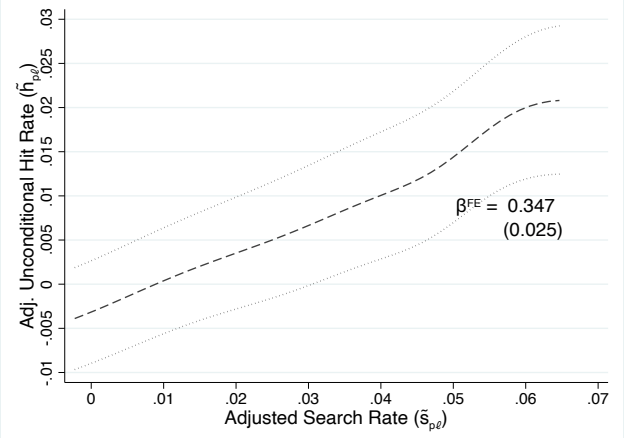
(a) Pooled



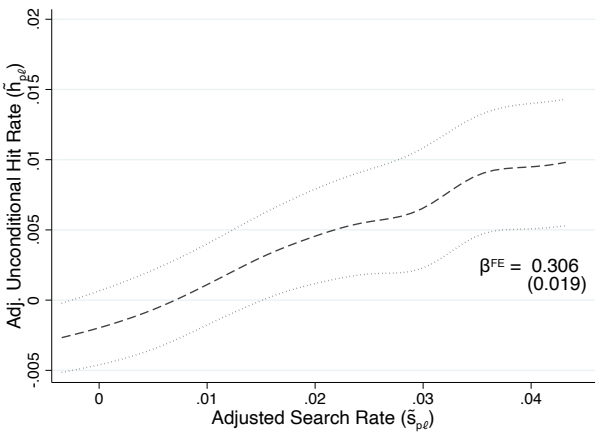
(b) White Motorists



(c) Black Motorists

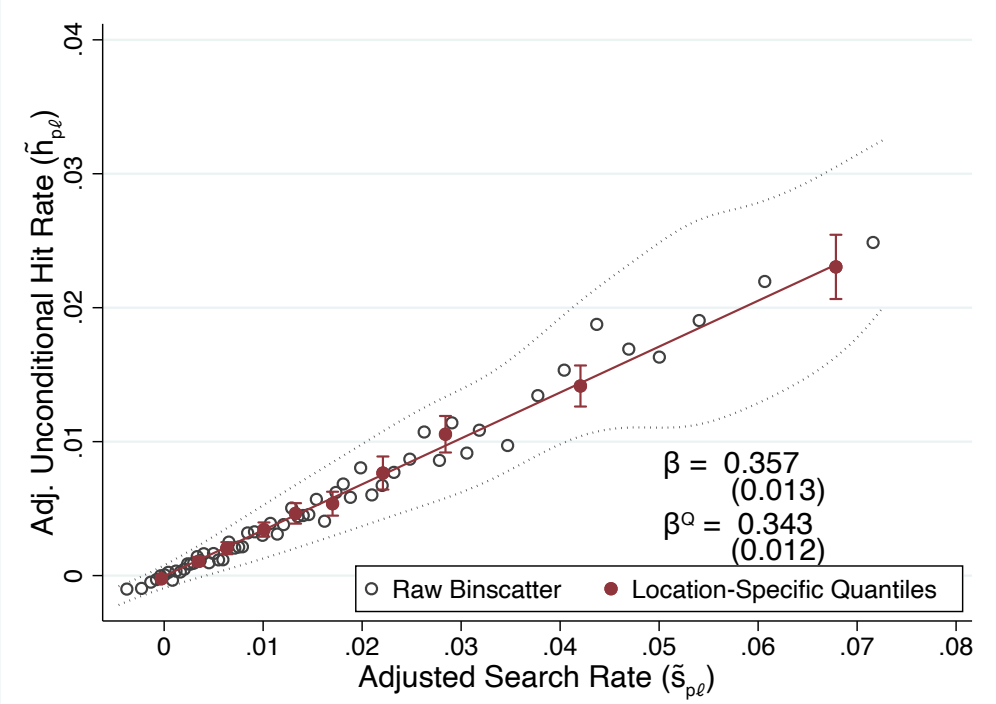


(d) Hispanic Motorists



Note: In this figure we plot adjusted trooper unconditional hit rates ($\tilde{h}_{p\ell}$) against trooper search rates ($\tilde{s}_{p\ell}$) adjusting for location fixed effects as in Cattaneo et al. (2019). $\tilde{h}_{p\ell}$ and $\tilde{s}_{p\ell}$ take on values between zero and one (before each is residualized). 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 are provided in parentheses. Panel A, Panel B, Panel C, and Panel D plot the search productivity curve (SPC) for all motorists, white motorists, black motorists, and Hispanic motorists, respectively.

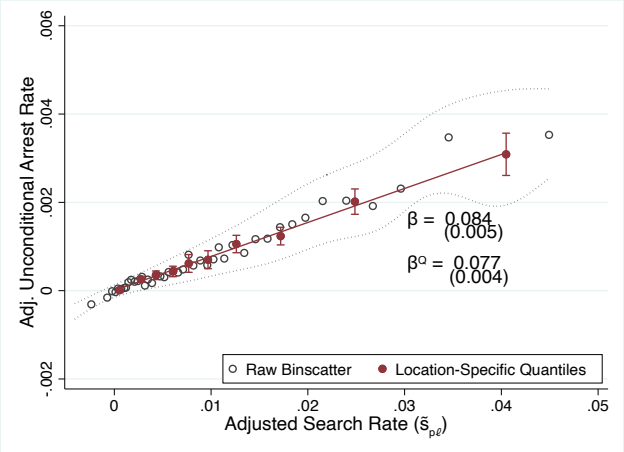
FIGURE B10
BETWEEN-TROOPER SEARCH PRODUCTIVITY CURVE, NIGHT STOPS



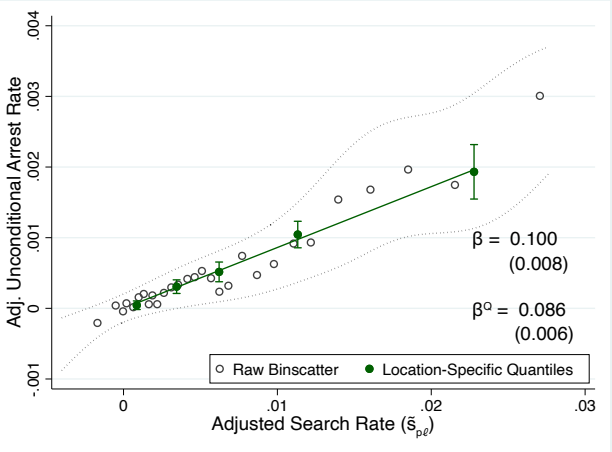
Note: This figure plots adjusted trooper unconditional hit rates ($\tilde{h}_{p\ell}$) against trooper search rates ($\tilde{s}_{p\ell}$), restricting to stops made between 9pm and 5am. $\tilde{h}_{p\ell}$ and $\tilde{s}_{p\ell}$ take on values between zero and one (before each is residualized). We use two approaches described in Section IV.D. 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 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 B11
BETWEEN-TROOPER SEARCH PRODUCTIVITY CURVE, ARRESTS

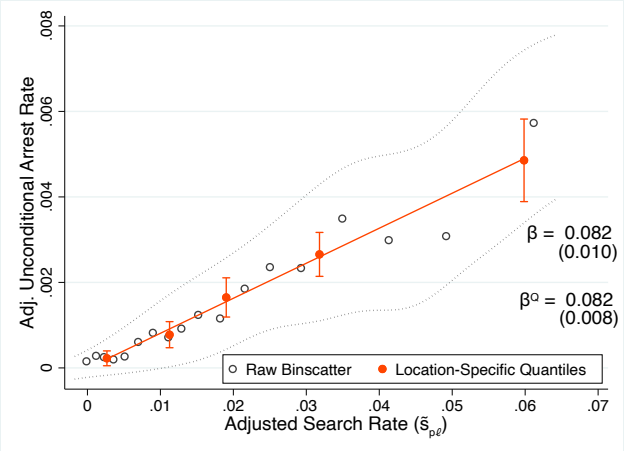
(a) Pooled



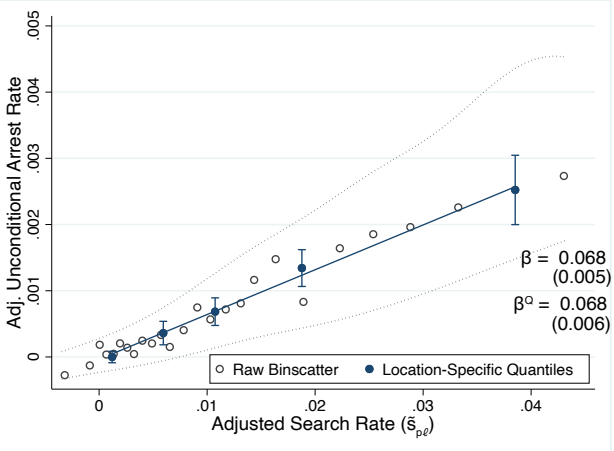
(b) White Motorists



(c) Black Motorists



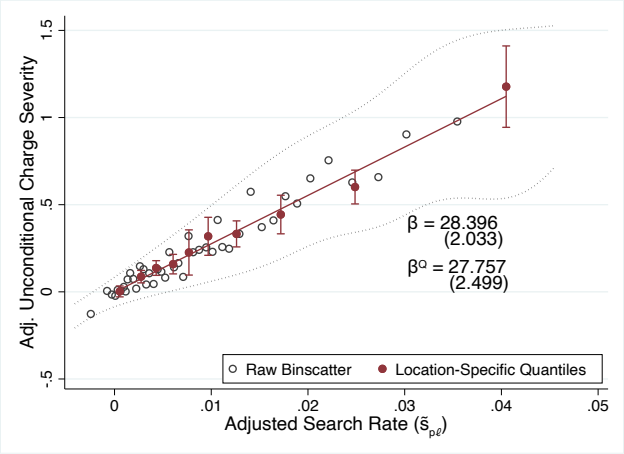
(d) Hispanic Motorists



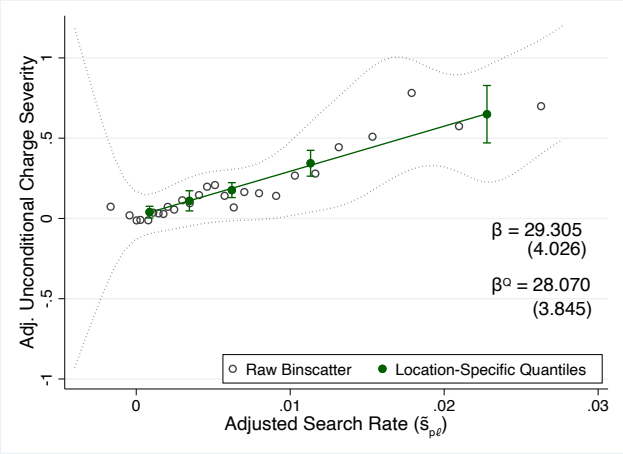
Note: In this figure we plot adjusted trooper unconditional arrest rates against trooper search rates ($\tilde{s}_{p\ell}$) using two approaches described in Section IV.D. The first approach is a simple binscatter, where we choose the integrated mean square error-optimal number of bins as in Cattaneo 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 arrest 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 slope, 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 arrest rates across quantiles. From this approach, the figure includes the mean values for each decile and the best linear fit and its slope. Bootstrap standard errors for the estimated slope are provided in parentheses. Panel A, Panel B, Panel C, and Panel D plot the search productivity curve (SPC) for all motorists, white motorists, black motorists, and Hispanic motorists, respectively. Trooper unconditional arrest rates and trooper search rates take on values between zero and one (before each is residualized).

FIGURE B12
BETWEEN-TROOPER SEARCH PRODUCTIVITY CURVE, CHARGE SEVERITY

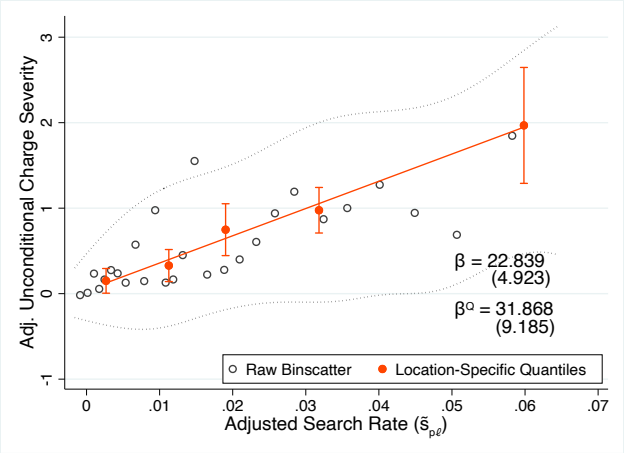
(a) Pooled



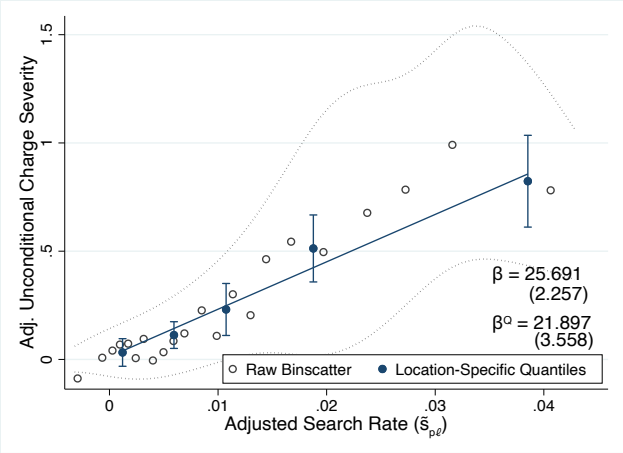
(b) White Motorists



(c) Black Motorists

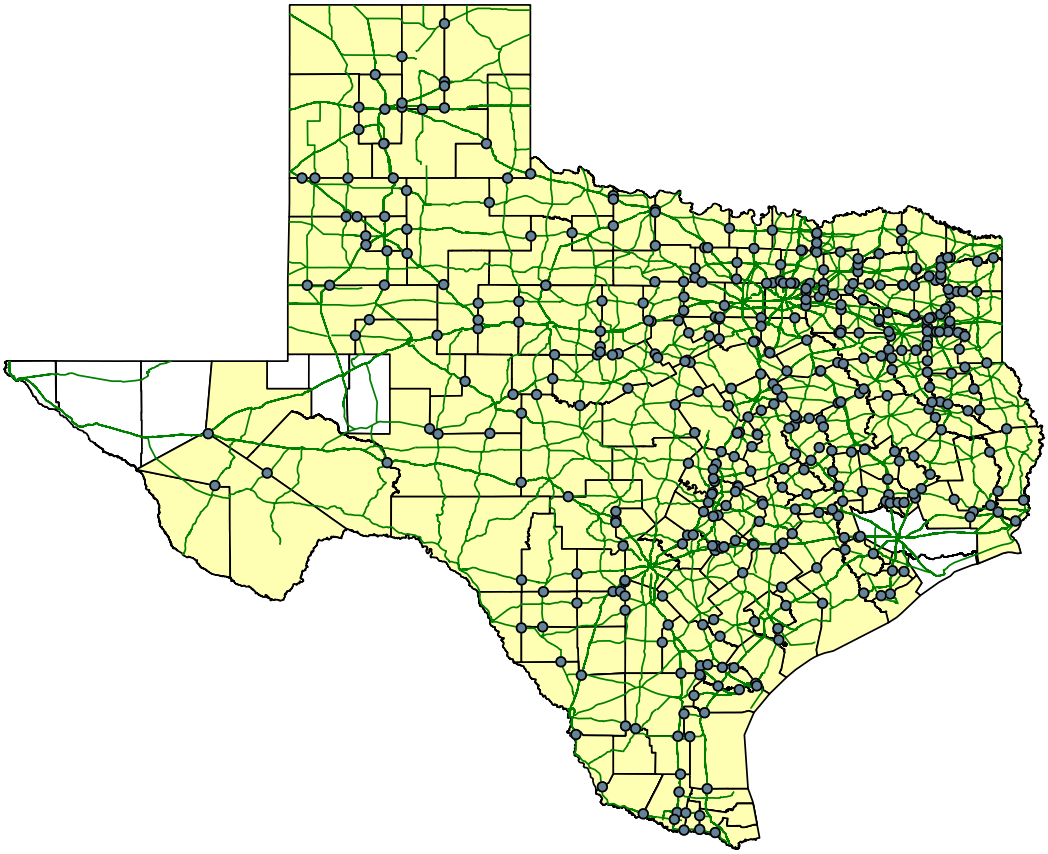


(d) Hispanic Motorists



Note: In this figure we plot adjusted trooper unconditional charge severity, measured as the average incarceration sentence associated with conviction for any arrest charges and otherwise set to zero, against trooper search rates ($\tilde{s}_{p\ell}$) using two approaches described in Section IV.D. 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 charge severity 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 slope, 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 charge severity across quantiles. From this approach, the figure includes the mean values for each decile and the best linear fit and its slope. Bootstrap standard errors for the estimated slope are provided in parentheses. Panel A, Panel B, Panel C, and Panel D plot the search productivity curve (SPC) for all motorists, white motorists, black motorists, and Hispanic motorists, respectively.

Figure B13
Map of Highway-Border Intersections Included in RD Analysis



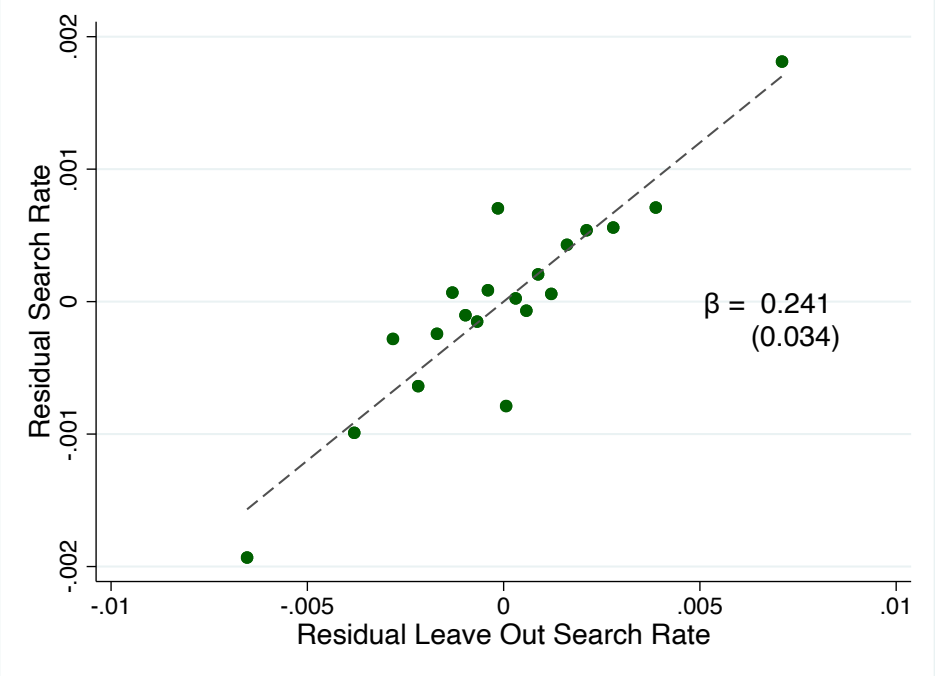
Note: This map depicts in blue the 424 intersections between state/interstate highways and sergeant area borders that define the spatial RD sample. The set of state and interstate highways associated with these intersections is superimposed in green. Sergeant areas included in the RD sample are shaded yellow.

FIGURE B14
COMPARING TROOPER SEARCH RATES BETWEEN LOCATIONS

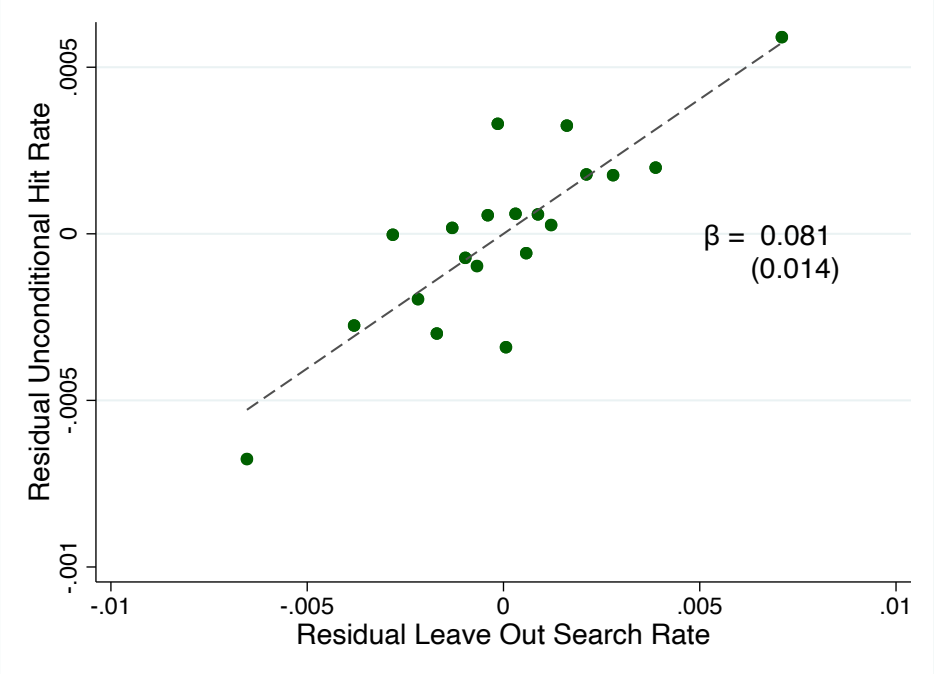
Note: This figure plots a binscatter for the relationship between a trooper's race-specific residual search rates in the two locations where they have conducted the most stops. Residual search rates are constructed after partialing out motorist race fixed effects. We limit to trooper by motorist race combinations where the trooper has conducted at least 50 stops for that racial group in both locations. Trooper search rates take on values between zero and one (before they are residualized).

FIGURE B15
WITHIN-TROOPER VARIATION IN SEARCH RATES

(a) First Stage



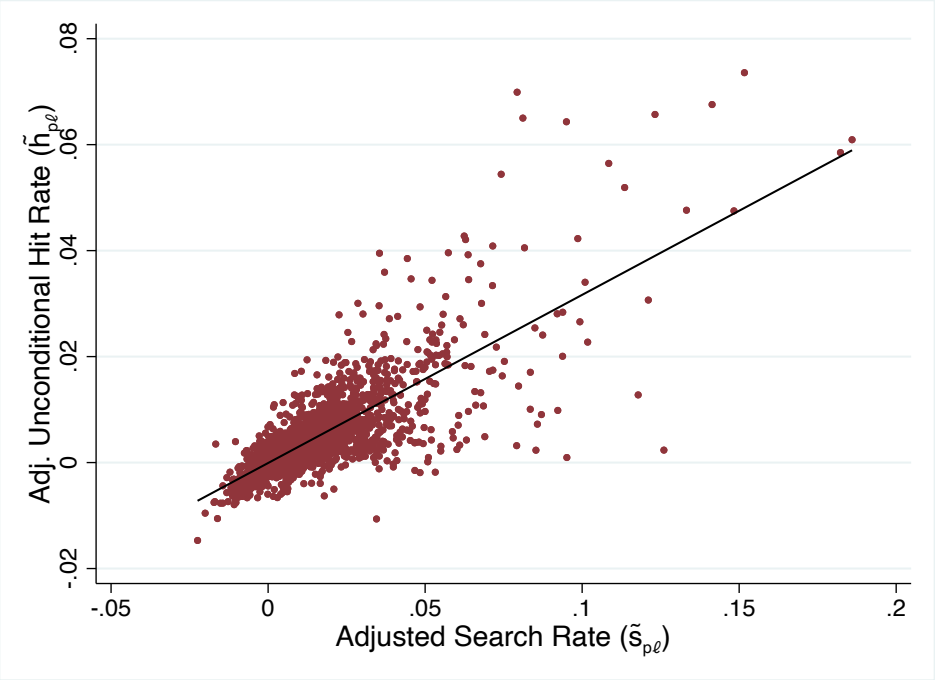
(b) Reduced Form



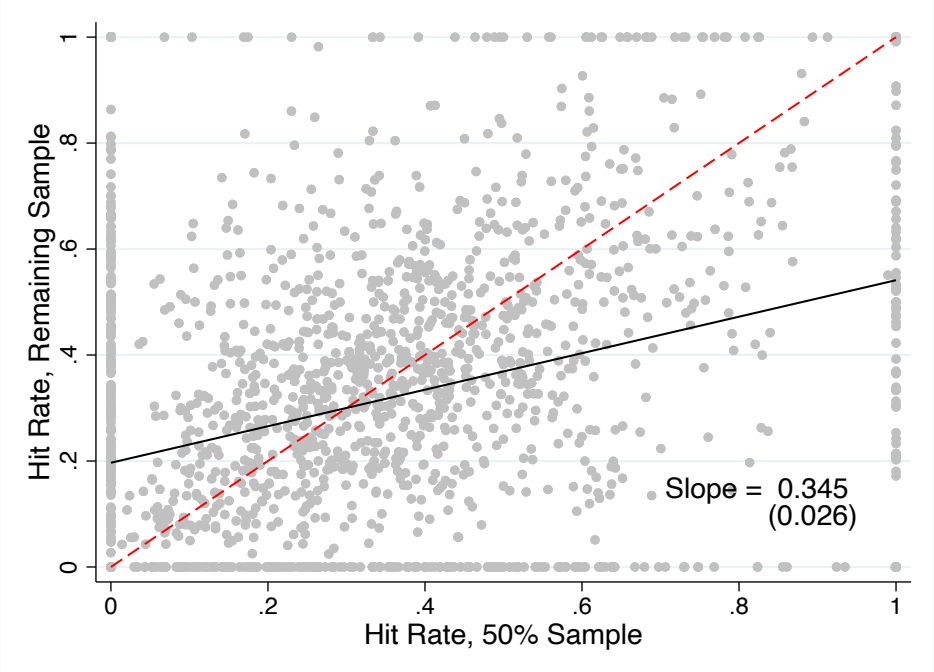
Note: These figures summarize the relationship between residual leave out search rates and residual search rates (Panel A) and residual unconditional hit rates (Panel B). Observations are at the stop level. Leave out search rates, search rates, and unconditional hit rates take on values between zero and one (before each is residualized). Both plots include bin scatters where observations at grouped into ventiles based on the deviation in coworker search rates. The construction of leave out search rates is described in Section **??**. Residuals partial out trooper, motorist, and year fixed effects.

FIGURE B16
TROOPERS VARY IN SCREENING ABILITY

(a)



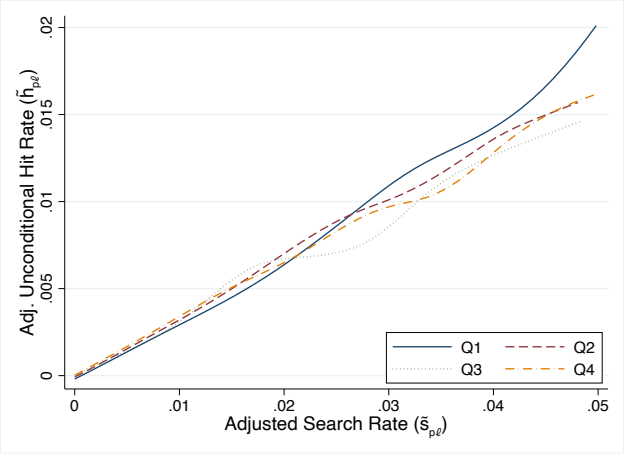
(b) Split Sample Hit Rates



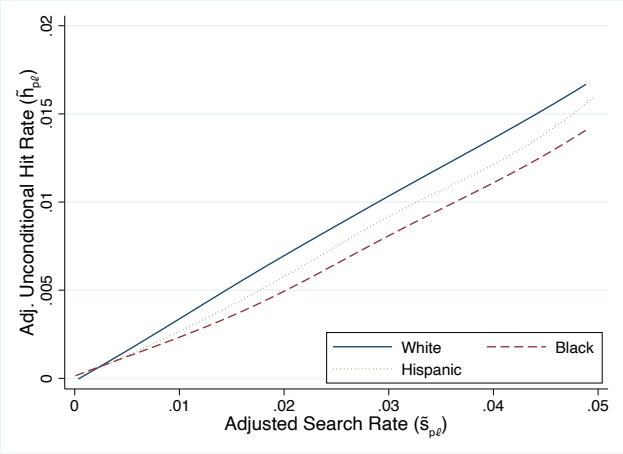
Note: These figures establish that hit rates vary systematically across troopers. Panel A plots adjusted search rates ($\tilde{s}_{p\ell}$) against adjusted unconditional hit rates ($\tilde{h}_{p\ell}$), where each marker represents a trooper by location pair. Panel B plots trooper by location hit rates in one randomly selected half of stops against the same trooper by location hit rates in the remaining half of stops. $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ take on values between zero and one (before each is residualized).

FIGURE B17
BETWEEN-TROOPER SPC BY TROOPER SUBGROUP

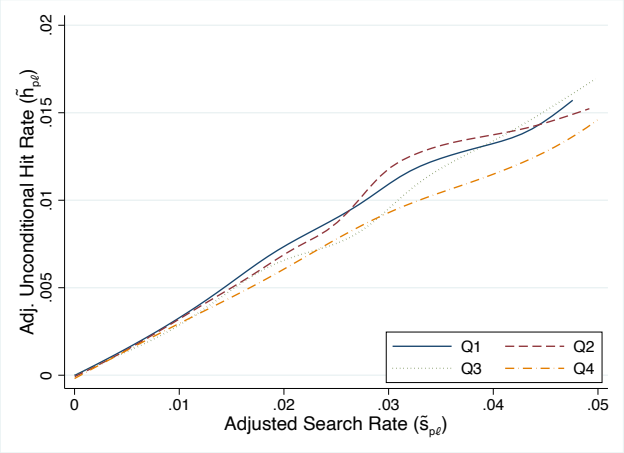
(a) By Experience



(b) By Race



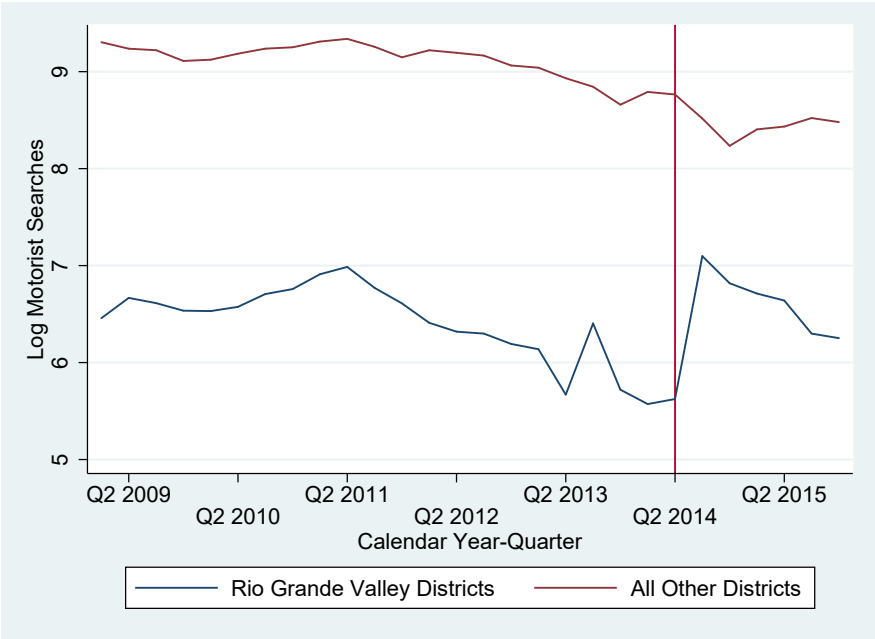
(c) By Stop Rate



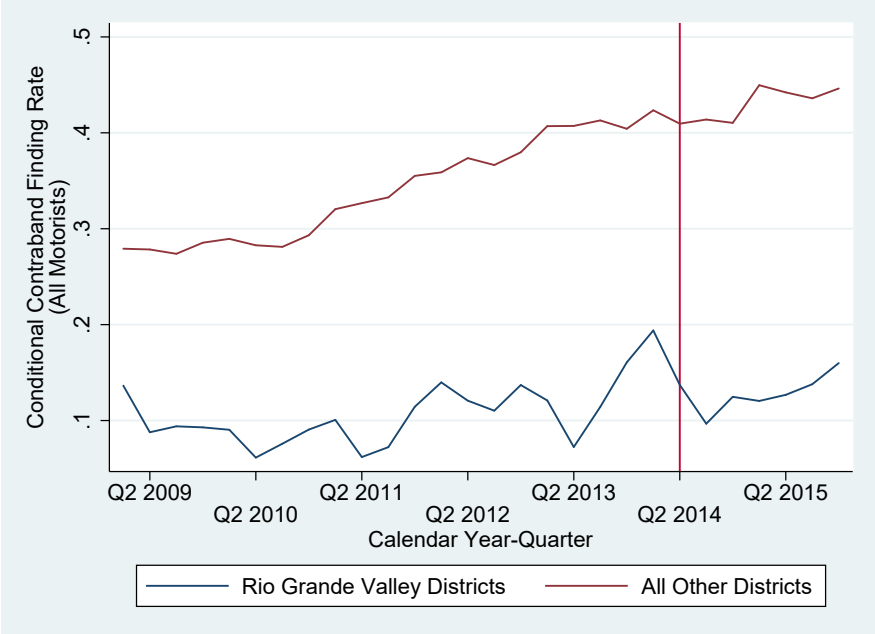
Note: In this figure we plot adjusted trooper unconditional hit rates ($\tilde{h}_{p\ell}$) against trooper search rates ($\tilde{s}_{p\ell}$) for subgroups of troopers. Panel A splits troopers into quartiles by experience. Panel B splits trooper by race. Panel C splits troopers into quartiles by stop rate. $\tilde{s}_{p\ell}$ and $\tilde{h}_{p\ell}$ take on values between zero and one (before each is residualized).

FIGURE B18
DETERRENCE EFFECT ESTIMATES: OPERATION STRONG SAFETY

(a) Log Searches Conducted by Region, Quarterly

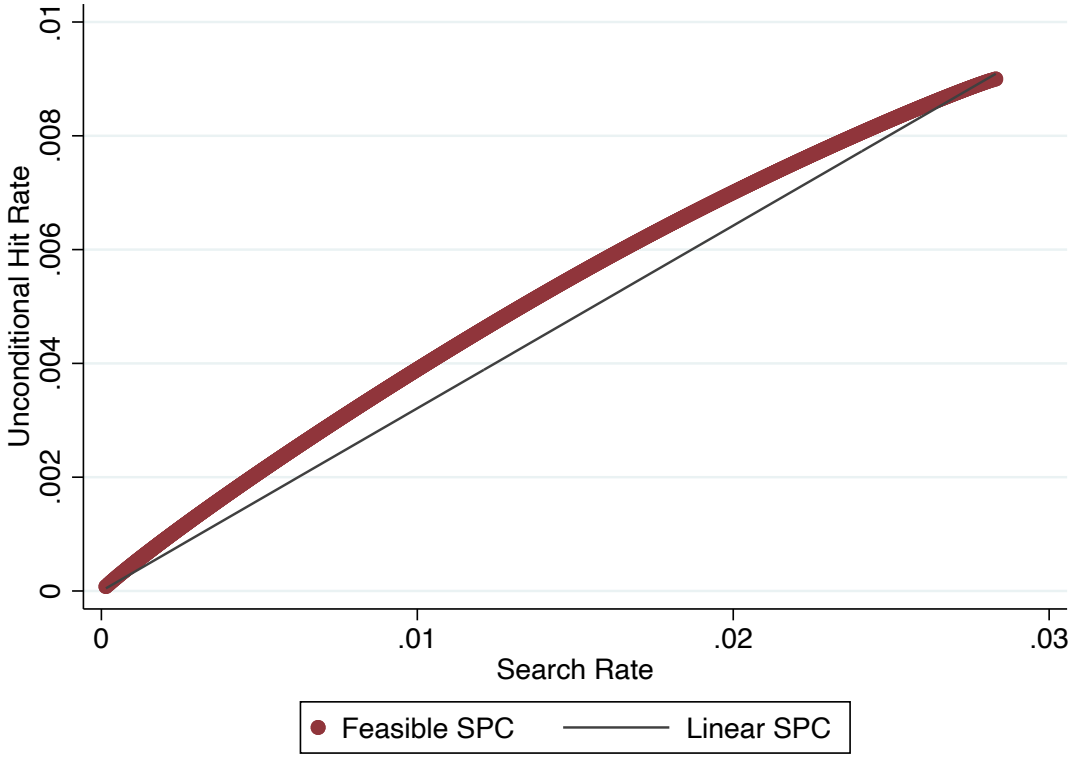


(b) Hit Rate by Region, Quarterly



Note: These figures compare the evolution of quarterly logged counts of motorist searches and of the quarterly conditional contraband finding rate by region (Rio Grande Valley districts as compared to all other districts). The red vertical line denotes the start of Operation Strong Safety in June 2014. The conditional contraband finding rate takes on values between zero and one.

Figure B19
A Feasible Constructed Search Productivity Curve



Note: This figure plots the feasible search productivity curve (SPC) we construct using observed search and hit rates for troopers in the top quartile by search rate. The construction of this SPC is described in detail in Section B.7. The search rates and hit rates take on values between zero and one.

C. Appendix: Additional Proofs

As in Section III, let $\tilde{\eta}_i(\sigma)$ denote the SPC for trooper i. The theory described in Section III implies that this function is weakly concave, so $\tilde{\eta}_i'(\sigma) \leq \tilde{\eta}_i'(\sigma')$ for $\sigma > \sigma'$. We do not observe the function $\tilde{\eta}_i(\sigma)$ but we observe it for at least one point, trooper i's observed search rate, σ_i^0. We also know that, by construction, $\tilde{\eta}_i(0) = 0$. The average hit rate for trooper i is therefore

$$\frac{\tilde{\eta}_i(\sigma_i^0)}{\sigma_i^0}$$

and marginal hit rate is $\tilde{\eta}_i'(\sigma_i^0)$.

Suppose we know that, for trooper i, average and marginal hit rates are equal, or

$$\frac{\tilde{\eta}_i(\sigma_i^0)}{\sigma_i^0} = \tilde{\eta}_i'(\sigma_i^0).$$

Claim: this implies that $\tilde{\eta}_i(\sigma_i^0)$ is linear for $\sigma \in [0, \sigma_i^0]$.

$$\begin{aligned}
\frac{\tilde{\eta}_i(\sigma_i^0)}{\sigma_i^0} &= \frac{\int_0^{\sigma_i^0} \tilde{\eta}_i'(\sigma) d\sigma}{\sigma_i^0} \\
&\geq \frac{\int_0^{\sigma_i^0} \tilde{\eta}_i'(\sigma_i^0) d\sigma}{\sigma_i^0} \\
&= \tilde{\eta}_i'(\sigma_i^0)
\end{aligned}$$

where the equality only holds if $\tilde{\eta}_i'(\sigma)$ is constant for $s \in [0, \sigma_i^0]$.

Now take an average of SPCs across troopers, $\sum \pi_i \tilde{\eta}_i(\sigma)$. Now the average hit rate is

$$\sum_i \pi_i \frac{\tilde{\eta}_i(\sigma_i^0)}{\sigma_i^0}$$

and the average marginal hit rate is

$$\sum_i \pi_i \tilde{\eta}_i'(\sigma_i^0).$$

We have by the same argument that

$$\sum_i \pi_i \frac{\tilde{\eta}_i(\sigma_i^0)}{\sigma_i^0} = \sum_i \pi_i \tilde{\eta}_i'(\sigma_i^0)$$

if and only if $\tilde{\eta}_i'(\sigma)$ is constant for $s \in [0, \sigma_i^0]$ for every i.

References

Aaronson, Daniel, Lisa Barrow, and William Sander, "Teacher and Student Achievement in the Chicago Public Schools," *Journal of Labor Economics*, 2007, *25* (1), 95–135.

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

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

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson, "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector," *American Economic Review*, August 2016, *106* (8), 2110–2144.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, "Verifying the existence of maximum likelihood estimates for generalized linear models," August 2019. Unpublished manuscript.

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.

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.

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

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