




In the format provided by the authors and unedited.

A large-scale analysis of racial disparities in police stops across the United States

Emma Pierson¹, Camelia Simoiu², Jan Overgoor², Sam Corbett-Davies¹, Daniel Jenson², Amy Shoemaker ², Vignesh Ramachandran², Phoebe Barghouty², Cheryl Phillips³, Ravi Shroff⁴ and Sharad Goel ² 

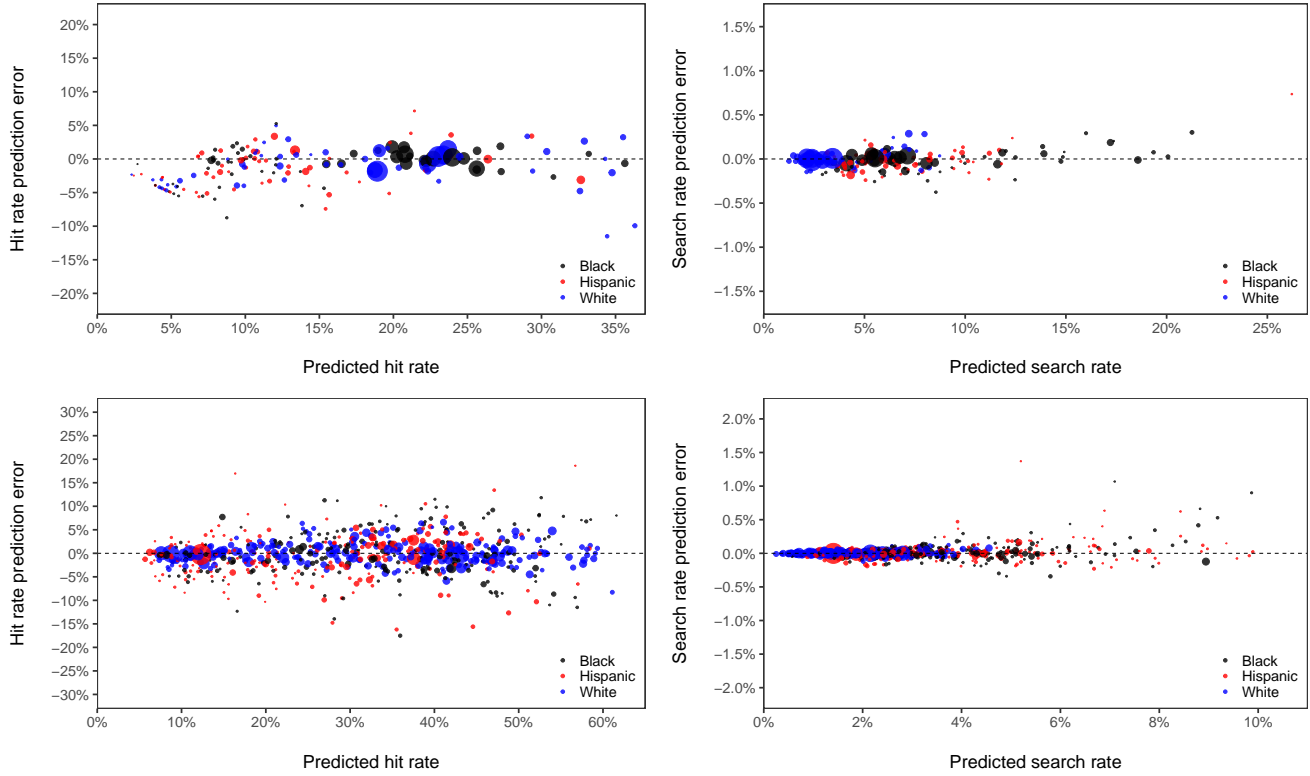
¹Computer Science, Stanford University, Stanford, CA, USA. ²Management Science & Engineering, Stanford University, Stanford, CA, USA.

³Communication, Stanford University, Stanford, CA, USA. ⁴Applied Statistics, Social Science, and Humanities, New York University, New York, NY, USA.

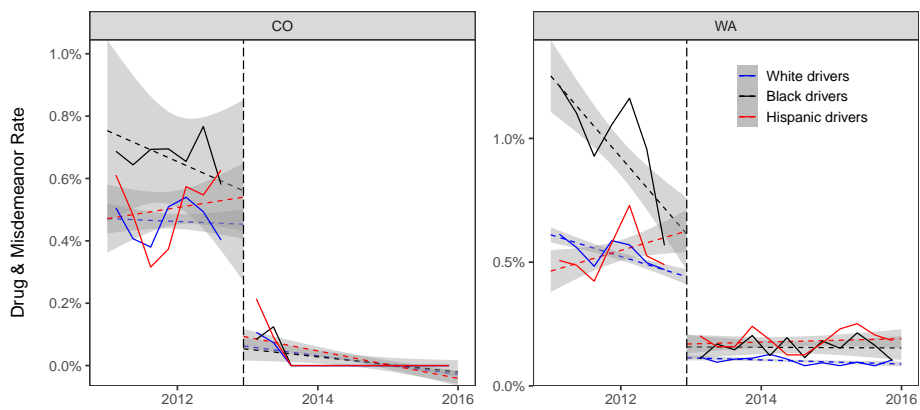
e-mail: scgoel@stanford.edu

Supplementary Information

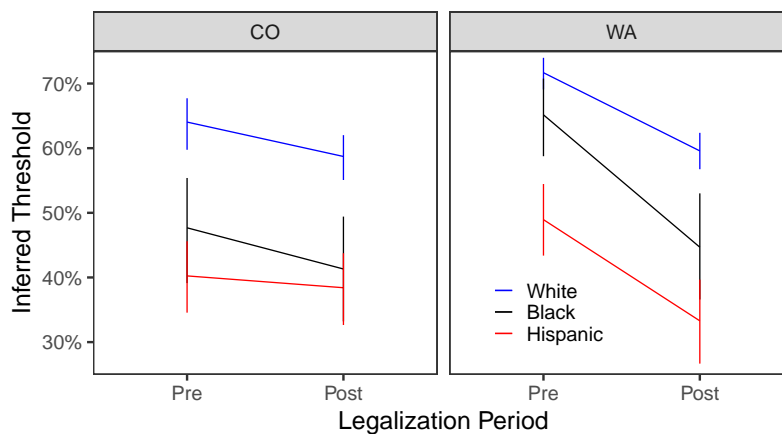
Supplementary Figures



Supplementary Figure 1. Posterior predictive checks for threshold test models for municipal police department (top) and state patrol (bottom) data. Prediction errors for both hit rates (left) and search rates (right) are minimal and consistent across race groups, validating the models' fit. In the left column, point size corresponds to number of searches per location (per county, above, and per district, below); in the right column, point size corresponds to the number of stops per location (per county, above, and per district, below).



Supplementary Figure 2. The proportion of stops that result in a drug-related infraction or misdemeanor before and after recreational marijuana was legalized in Colorado and Washington at the end of 2012 (indicated by the vertical lines). Subsequent to legalization, there is a substantial drop in offense rates. Data from the fourth quarter of 2012 are excluded, since that period includes stops both before and after legalization. The rates in CO (left panel) are based on 1,534,893 stops; the rates in WA (right panel) are based on 3,985,677 stops. In Colorado, we consider only offenses for marijuana possession; in Washington, we include all drug-related misdemeanors, as more detailed information is not available, so there are still some recorded drug violations post-legalization.



Supplementary Figure 3. Inferred thresholds faced by white (blue lines), black (black lines), and Hispanic (red lines) drivers before and after marijuana legalization. Error bars show the 95% credible intervals. The threshold values in CO (left panel) are inferred from 1,674,619 stops; the threshold values in WA (right panel) are inferred from 3,985,677 stops. In all cases black and Hispanic drivers face a lower threshold than white drivers.

Supplementary Tables

Supplementary Table 1. Coefficient estimates for different veil-of-darkness model specifications, varying the inclusion of separate state/city intercepts, the inclusion of an interaction term between time and location, and the degrees of freedom in the natural spline over time. All models are based on stops conducted during the two 60-day windows centered on the beginning and end of daylight savings time. The “coef.” column estimates differences in the racial composition of stopped drivers before sunset and after dark, after adjusting for time and location. The results presented in the main text are in bold, and correspond to a model which includes state and city intercepts, has no interaction term between time and location, and has six degrees of freedom in the natural spline over time. All coefficient estimates are negative and statistically significant (all p-values are two-tailed), suggesting that evidence of racial bias against black drivers is robust to different model specifications.

	spline d.f.		coef.	s.e.	95% CI	p value
	1		-0.040	0.002	(-0.045, -0.035)	<0.001
Models with	2		-0.035	0.003	(-0.040, -0.030)	<0.001
no state/city intercept	3		-0.034	0.003	(-0.039, -0.029)	<0.001
and no interaction term	4		-0.036	0.003	(-0.041, -0.031)	<0.001
between time and location	5		-0.036	0.003	(-0.041, -0.031)	<0.001
	6		-0.036	0.003	(-0.041, -0.031)	<0.001
	1		-0.032	0.003	(-0.037, -0.028)	<0.001
Models with	2		-0.039	0.003	(-0.044, -0.034)	<0.001
no state/city intercept	3		-0.041	0.003	(-0.046, -0.035)	<0.001
and with an interaction term	4		-0.041	0.003	(-0.046, -0.036)	<0.001
between time and location	5		-0.041	0.003	(-0.046, -0.036)	<0.001
	6		-0.041	0.003	(-0.046, -0.035)	<0.001
	1	city	-0.044	0.003	(-0.049, -0.038)	<0.001
	1	state	-0.036	0.003	(-0.041, -0.030)	<0.001
	2	city	-0.038	0.003	(-0.044, -0.032)	<0.001
	2	state	-0.032	0.003	(-0.038, -0.026)	<0.001
Models with	3	city	-0.037	0.003	(-0.043, -0.031)	<0.001
state/city intercept and	3	state	-0.032	0.003	(-0.038, -0.027)	<0.001
no interaction term	4	city	-0.039	0.003	(-0.045, -0.033)	<0.001
between time and location	4	state	-0.033	0.003	(-0.039, -0.027)	<0.001
	5	city	-0.039	0.003	(-0.045, -0.033)	<0.001
	5	state	-0.033	0.003	(-0.039, -0.027)	<0.001
	6	city	-0.039	0.003	(-0.045, -0.033)	<0.001
	6	state	-0.033	0.003	(-0.039, -0.027)	<0.001
	1	city	-0.021	0.003	(-0.028, -0.015)	<0.001
	1	state	-0.043	0.003	(-0.050, -0.037)	<0.001
	2	city	-0.030	0.004	(-0.037, -0.023)	<0.001
	2	state	-0.049	0.004	(-0.056, -0.042)	<0.001
Models with	3	city	-0.033	0.004	(-0.040, -0.026)	<0.001
state/city intercept and	3	state	-0.049	0.004	(-0.057, -0.042)	<0.001
with an interaction term	4	city	-0.033	0.004	(-0.040, -0.026)	<0.001
between time and location	4	state	-0.050	0.004	(-0.057, -0.042)	<0.001
	5	city	-0.033	0.004	(-0.040, -0.026)	<0.001
	5	state	-0.049	0.004	(-0.057, -0.042)	<0.001
	6	city	-0.033	0.004	(-0.040, -0.026)	<0.001
	6	state	-0.049	0.004	(-0.057, -0.042)	<0.001

Supplementary Table 2. Summary of the data for the 35 municipal police departments (top) and 21 state patrol agencies (bottom) used in our analyses. A solid circle ● signifies that data are available for at least 50% of stops. An × signifies that while data are available, they appeared untrustworthy and thus were not used. Geographic subdivision typically means county (for state patrol agencies) or beat/precinct (for municipal police departments). A star * in the geographic subdivision column indicates that a subdivision other than county was available (department id or zone), but county information was not present.

	State	City	Stops	Date Range	Date	Time	Geographic Subdivision	Subject Race	Subject Age	Subject Gender	Search Conducted	Contraband Found
1	AR	Little Rock	13,072	2017-2017	●	●	●	●	●	●		
2	CA	Bakersfield	140,544	2011-2018	●	●	●	●	●	●		
3	CA	Los Angeles	3,328,632	2011-2018	●	●	●	●	●	●		
4	CA	Oakland	95,441	2013-2017	●	●	●	●	●	●	●	●
5	CA	San Diego	312,175	2014-2017	●	●	●	●	●	●	●	●
6	CA	San Francisco	326,717	2011-2016	●	●	●	●	●	●	●	●
7	CA	San Jose	75,577	2013-2018	●	●	●	●	●	●	●	●
8	CO	Aurora	160,080	2012-2016	●	●	●	●	●	●		
9	CT	Hartford	18,219	2013-2016	●	●	●	●	●	●	●	●
10	KS	Wichita	352,287	2011-2016	●	●	●	●	●	●		
11	KY	Louisville	105,055	2015-2018	●	●	●	●	●	●	●	●
12	LA	New Orleans	259,353	2011-2018	●	●	●	●	●	●	●	●
13	MN	Saint Paul	127,214	2011-2016	●	●	●	●	●	●	●	●
14	NC	Charlotte	604,544	2011-2015	●	●	●	●	●	●	●	●
15	NC	Durham	133,242	2011-2015	●	●	●	●	●	●	●	●
16	NC	Fayetteville	215,552	2011-2015	●	●	●	●	●	●	●	●
17	NC	Greensboro	206,492	2011-2015	●	●	●	●	●	●	●	●
18	NC	Raleigh	326,840	2011-2015	●	●	●	●	●	●	●	●
19	NC	Winston-Salem	174,775	2011-2015	●	●	●	●	●	●	●	●
20	NJ	Camden	138,142	2013-2018	●	●	●	●	●	●		
21	NV	Henderson	101,962	2011-2018	●	●	●	●	●	●		
22	NY	Albany	21,805	2011-2015	●	●	●	●	●	●		
23	OH	Cincinnati	200,742	2011-2018	●	●	●	●	●	●		
24	OH	Columbus	121,043	2012-2016	●	●	●	●	●	●	●	●
25	OK	Oklahoma City	531,313	2011-2017	●	●	●	●	●	●		
26	OK	Tulsa	198,576	2011-2016	●	●	●	●	●	●		
27	PA	Philadelphia	1,114,213	2014-2018	●	●	●	●	●	●	●	●
28	PA	Pittsburgh	169,964	2011-2018	●	●	●	●	●	●	●	●
29	TN	Nashville	2,362,727	2011-2018	●	●	●	●	●	●	●	●
30	TX	Arlington	102,901	2016-2016	●	●	●	●	●	●	●	●
31	TX	Garland	148,009	2012-2018	●	●	●	●	●	●		
32	TX	Plano	101,542	2013-2015	●	●	●	●	●	●	●	●
33	TX	San Antonio	845,733	2012-2018	●	●	●	●	●	●	●	●
34	VT	Burlington	28,718	2012-2017	●	●	●	●	●	●	●	●
35	WI	Madison	195,595	2011-2017	●	●	●	●	●	●		
<hr/>												
1	AZ	–	3,202,567	2011-2017	●	●	●	●	●	●	●	×
2	CA	–	20,642,287	2011-2016	●	●	●	●	●	●	●	●
3	CO	–	2,232,609	2011-2017	●	●	●	●	●	●	●	●
4	CT	–	432,122	2013-2015	●	●	●	●	●	●	●	●
5	FL	–	6,348,386	2011-2018	●	●	●	●	●	●	●	●
6	GA	–	927,071	2012-2016	●	●	●	●	●	●		
7	IL	–	1,813,162	2012-2017	●	●	*	●	●	●	●	●
8	MA	–	1,773,546	2011-2015	●	●	●	●	●	●	●	×
9	MT	–	642,200	2011-2016	●	●	●	●	●	●	●	●
10	NC	–	3,500,180	2011-2015	●	●	●	●	●	●	●	●
11	ND	–	250,525	2011-2015	●	●	●	●	●	●		
12	NH	–	160,794	2014-2015	●	●	●	●	●	●		
13	NY	–	6,318,577	2011-2017	●	●	●	●	●	●		
14	OH	–	5,630,180	2011-2017	●	●	●	●	●	●	●	●
15	RI	–	229,691	2011-2015	●	●	*	●	●	●	●	●
16	SC	–	4,288,166	2011-2016	●	●	●	●	●	●	●	●
17	TN	–	1,953,299	2011-2016	●	●	●	●	●	●		
18	TX	–	14,210,473	2011-2017	●	●	●	●	●	●	●	●
19	VT	–	250,949	2011-2015	●	●	*	●	●	●	●	●
20	WA	–	5,858,474	2011-2018	●	●	●	●	●	●	●	●
21	WI	–	754,451	2011-2016	●	●	●	●	●	●	●	●
Total			94,778,505									