Statistical Transparency of Policing (S.T.O.P.)

December 1, 2025

Oregon Criminal Justice Commission

Ryan Keck
Interim Executive Director

The mission of the Oregon Criminal Justice Commission is to improve the legitimacy, efficiency, and effectiveness of state and local criminal justice systems.

Acknowledgements

The Oregon Criminal Justice Commission (CJC) would like to acknowledge and thank Oregon's law enforcement agencies, whose continued commitment is crucial to the success of the S.T.O.P. program, and the Oregon Department of Public Safety Standards and Training (DPSST) for providing support to agencies and to the CJC to continuously improve the program.

Authors:
Kelly Officer, Research Director
Emma Dunn, Research Analyst
Paige Frye-Holcomb, PhD, Senior Research Analyst
Carrie Glaser, Senior Research Analyst
Angel Pairan, STOP Program Liaison
Katherine Tallan, Senior Research Analyst
Michael Weinerman, PhD, Economist

Table of Contents

Data Tables

Figures

Appendix Tables & Figures

Executive Summary

House Bill 2355 (2017) mandates all Oregon law enforcement agencies to submit officer-initiated traffic and pedestrian stop data to the Oregon Criminal Justice Commission (CJC). The resulting Oregon Statistical Transparency of Policing (STOP) Program, housed at the CJC, was created with assistance from the Oregon State Police (OSP) and the Oregon Department of Public Safety Standards and Training (DPSST). This is the seventh annual report to the Oregon Legislature by the STOP Program examining data submitted by law enforcement agencies.

Table 0.1 reports descriptive statistics for Tier 1 (100+ officers), Tier 2 (25 – 99 officers), and Tier 3 (<25 officers) agency stops. Most drivers stopped were white and male. All minority races combined accounted for almost 30% of stops by larger (tier 1) agencies and a smaller portion, just over 20%, for smaller agencies, reflecting differences in urban and rural driving populations.

Table 0.1 Aggregate Year 7 Stop Data

Variable	Tier 1	Tier 2	Tier 3
Traffic Stop	98.4%	98.6%	98.9%
Race/Ethnicity			
Asian or PI	3.6%	2.9%	2.5%
Black	5.3%	3.2%	2.2%
Hispanic	17.8%	17.2%	15.2%
Middle Eastern	1.8%	1.2%	0.9%
Native	0.6%	0.3%	0.3%
White	70.4%	75.2%	78.3%
Gender			
Female	32.3%	35.6%	34.8%
Male	67.2%	64.2%	64.4%
Nonbinary	0.4%	0.2%	0.7%
Age			
Under 21	10.4%	12.3%	12.3%
21-29	21.6%	21.1%	18.9%
30-39	24.4%	23.3%	22.3%
40-49	18.6%	18.9%	18.4%
50+	24.3%	24.4%	28.0%
Disposition			
None	1.8%	3.7%	3.2%
Warning	60.8%	64.8%	71.1%
Citation	35.7%	28.5%	23.6%
Juv Summons	0.0%	0.0%	0.0%
Arrest	1.8%	1.5%	1.2%
Search Conducted	1.2%	0.9%	0.6%

Agencies with predominantly unreported values are excluded from the summaries shown here.

A majority of stops (61% - 71%) resulted in warnings with no further law enforcement action, while 24% - 36% resulted in a citation. Only a small fraction resulted in more serious law enforcement action including arrest, in part because the data does not include calls for service such as 911 calls.

STOP Program researchers use three analytical methods to examine traffic and pedestrian stop data for evidence of racial/ethnic disparities. The first method, the 'Decision to Stop' analysis, takes advantage of natural variations in daylight and darkness throughout the year to determine if minority individuals are more likely to be stopped when race/ethnicity is easier to detect because it is light out.

The second is the 'Stop Outcomes' analysis, which examines whether, after matching on all available stop characteristics (e.g., time of day and day of the week the stop was made, reason for the stop, gender, age), minority individuals are cited, searched, or arrested more often than similarly situated white individuals.

Finally, the 'Search Findings' analysis compares relative rates of successful searches (i.e., those resulting in the discovery of contraband) across racial/ethnic groups. This analysis works off the assumption that if search decisions by officers are made based on race/ethnicity neutral criteria, then success rates should be similar across different racial/ethnic categories.

For each of these analyses, any different outcomes by racial/ethnic group must meet the 95 percent confidence level for it to be statistically significant. Further, following best practices, for a law enforcement agency to be identified as one requiring further analysis as well as DPSST technical assistance, it must be identified as having a statistically significant difference in at least two of the three analytical tests performed on the STOP data. However, DPSST has and will continue to provide technical assistance to any agency upon request, regardless of the number of analyses that are statistically significant.

In this reporting year, Oregon State Police showed statistically significant differences in two analytical tests. Regardless of whether an agency is officially referred to DPSST, the CJC urges each agency to scrutinize their full set of results[1] and engage with DPSST on any results that show a statistically significant difference.

[1] The dashboard can be found here: https://public.tableau.com/app/profile/cjcdashboards/viz/S_T_O_P_StatisticalTransparencyofPolicing/Introduction

1. Background

In 2017, the Oregon Legislature mandated that by July 2020 all Oregon law enforcement agencies would collect data during officer-initiated traffic and pedestrian stops. The law required the Oregon Criminal Justice Commission (CJC) to analyze the data to determine whether evidence of racial disparities exists in officer-initiated stops or stop outcomes.

CJC, in collaboration with Oregon State Police (OSP), DPSST, and the STOP Steering Committee, developed a data collection strategy and free software for all state law enforcement agencies. The resulting data is one of the most robust traffic stop data sets in the United States; however, there are key limitations to note. First, the analyses presented here can only identify differences in police/citizen interactions during discretionary stops. Second, the data and analyses are only reported at the agency level. HB 2355, which created the STOP program, expressly forbids the collection of data that identifies either stopped individuals or officers. Taken together, this means that the analyses contained in this report cannot and do not address the personal experience of any individual who believes they may have been subjected to biased treatment, nor attribute a motivation of bias to any particular stop, officer, or agency.

STOP program researchers have selected highly respected, thoroughly vetted, and peer-reviewed methods for analyzing these stop data. Due to the rigor of these analyses, a statistically significant difference in stops made by a particular agency in at least two of the three tests indicates a need for that agency to engage with CJC and DPSST to identify the cause(s) of the statistically significant difference and seek to address them through technical assistance. The STOP program believes that the results presented herein can contribute to open and transparent dialogue between Oregonians inside and outside of law enforcement.

2. Characteristics of Year 7 Stop data

The data and analyses presented in this report have been compiled by CJC researchers from stop reports submitted by individual agencies. The data may not match other figures or reports of police activity due to different definitions, queries, time periods presented, and/or changes in availability and accuracy of records. For example, STOP program data generally excludes stops or activities made in response to a call for service such as a 911 call or an accident report.

For the current reporting period, Multnomah County Sheriff's Office and Springfield Police Department reports have been omitted from all tables, figures, and analyses due to data repository errors. An updated report will be provided in early 2026 with amended data and analyses for these agencies.

2.1 General Characteristics

While the analyses contained in sections 3, 4, and 5 of this report utilize two years of submitted data, the general characteristics described in this section are only for the most recent fiscal year, which includes stops made between July 1, 2024, and June 30, 2025. During that period, a total of 624,255 stops were made by 139 agencies. Over a third of all stops (219,443) were made by Oregon State Police.

Table 2.1.1, Table 2.1.2, and Appendix Table C.2 present the number of traffic and pedestrian stops for each tier 1, tier 2, and tier 3 agency, respectively. The 12 agencies in tier 1, the largest agencies in the state, accounted for 58% of all stops; the 39 tier 2 agencies accounted for 25% of all stops while the 90 tier 3 agencies made the remaining 17% of stops.

Table 2.1.1. Tier 1 Agency Stops by Stop Type

	Traffic		Pedestrian		
Agency	**Count**	**Pct**	**Count**	**Pct**	**Total**
Beaverton PD	12,500	93.6%	855	6.4%	13,355
Clackamas CO SO	27,601	95.6%	1,258	4.4%	28,859
Eugene PD	11,701	90.0%	1,307	10.0%	13,008
Gresham PD	4,464	99.3%	32	0.7%	4,496
Hillsboro PD	12,043	98.9%	131	1.1%	12,174
Marion CO SO	15,261	99.2%	116	0.8%	15,377
Medford PD	3,530	86.1%	570	13.9%	4,100
Oregon State Police	218,702	99.7%	741	0.3%	219,443
Portland PB	25,918	99.5%	138	0.5%	26,056
Salem PD	6,183	94.7%	345	5.3%	6,528
Washington CO SO	25,025	99.2%	205	0.8%	25,230
Tier 1 Total	**362,928**	**98.5%**	**5,698**	**1.5%**	**368,626**

Table 2.1.2. Tier 2 Agency Stops by Stop Type

	Traffic		Pedestrian		
Agency	**Count**	**Pct**	**Count**	**Pct**	**Total**
Albany PD	7,067	99.1%	62	0.9%	7,129
Ashland PD	2,347	93.0%	178	7.0%	2,525
Bend PD	3,539	99.7%	10	0.3%	3,549
Benton CO SO	8,622	99.8%	17	0.2%	8,639
Canby PD	3,920	97.9%	84	2.1%	4,004
Central Point PD	2,129	99.3%	16	0.7%	2,145
Corvallis PD	6,272	99.2%	50	0.8%	6,322
Deschutes CO SO	5,190	99.0%	50	1.0%	5,240
Douglas CO SO	1,215	99.9%	1	0.1%	1,216
Forest Grove PD	3,923	99.7%	11	0.3%	3,934
Grants Pass PD	1,880	95.0%	98	5.0%	1,978
Hermiston PD	9,033	98.8%	111	1.2%	9,144
Hood River CO SO	1,853	99.8%	4	0.2%	1,857
Jackson CO SO	11,261	98.7%	151	1.3%	11,412
Keizer PD	1,056	100.0%	0	0.0%	1,056
Klamath CO SO	84	100.0%	0	0.0%	84
Klamath Falls PD	1,138	99.9%	1	0.1%	1,139
Lake Oswego PD	6,575	99.0%	67	1.0%	6,642
Lane CO SO	5,533	98.7%	71	1.3%	5,604
Lebanon PD	1,687	100.0%	0	0.0%	1,687
Lincoln CO SO	1,840	99.9%	2	0.1%	1,842
Lincoln City PD	2,568	99.7%	7	0.3%	2,575
Linn CO SO	5,226	98.7%	68	1.3%	5,294
McMinnville PD	2,100	99.4%	13	0.6%	2,113
Milwaukie PD	5,576	97.9%	120	2.1%	5,696
Newberg-Dundee PD	5,161	99.9%	7	0.1%	5,168
OHSU PD	86	97.7%	2	2.3%	88
Oregon City PD	5,739	91.8%	510	8.2%	6,249
Polk CO SO	3,133	99.8%	7	0.2%	3,140
Port of Portland PD	803	94.7%	45	5.3%	848
Redmond PD	6,259	99.7%	18	0.3%	6,277
Roseburg PD	1,025	95.6%	47	4.4%	1,072
Tigard PD	7,917	98.4%	132	1.6%	8,049
Tualatin PD	5,124	99.8%	11	0.2%	5,135
UO PD	314	90.2%	34	9.8%	348
West Linn PD	3,646	99.6%	16	0.4%	3,662
Woodburn PD	1,481	99.8%	3	0.2%	1,484
Yamhill CO SO	5,859	99.8%	11	0.2%	5,870
Tier 2 Total	**148,181**	**98.6%**	**2,035**	**1.4%**	**150,216**

The demographic profile (race/ethnicity, gender, and age) of drivers and pedestrians stopped by agencies in each tier are shown in Table 2.1.3. Across all agency and stop types, the majority of those stopped were perceived as white. Males accounted for between 67.0 – 81.2% of stops depending on the group. The rates of stops by age group varied by category, with some agency groups skewing younger and others older.

Table 2.1.3. Demographics of Stopped Drivers & Pedestrians

	Tier 1		Tier 2		Tier 3	
	Traffic	Ped	Traffic	Ped	Traffic	Ped
Race/Ethnicity						
Asian or PI	3.7%	1.5%	2.9%	1.8%	2.5%	2.2%
Black	5.3%	7.8%	3.2%	4.7%	2.2%	4.4%
Hispanic	18.0%	11.2%	17.3%	10.7%	15.2%	10.8%
Middle Eastern	1.8%	0.5%	1.2%	0.6%	0.9%	0.2%
Native	0.6%	0.4%	0.3%	0.4%	0.3%	3.0%
White	70.3%	78.6%	75.1%	81.8%	78.3%	79.1%
Gender						
Female	32.5%	18.5%	35.8%	22.8%	34.3%	27.7%
Male	67.0%	81.2%	64.1%	76.7%	63.0%	72.0%
Nonbinary	0.4%	0.3%	0.2%	0.5%	2.6%	0.3%
Age						
Under 21	10.5%	5.5%	12.3%	8.0%	12.3%	17.7%
21-29	21.7%	13.4%	21.1%	15.2%	18.9%	15.5%
30-39	24.3%	33.9%	23.2%	29.2%	22.3%	24.8%
40-49	18.5%	26.1%	18.9%	24.0%	18.4%	20.1%
50+	24.3%	20.3%	24.4%	23.6%	28.1%	21.8%

Agencies with predominantly unreported values are excluded from the summaries shown here.

Table 2.1.4, Table 2.1.5, and Appendix Table C.1 present stops by race/ethnicity for each tier 1, tier 2, and tier 3 agency, respectively.

Table 2.1.4. Tier 1 Race/Ethnicity by Agency

Agency	Asian or PI	Black	Hispanic	Middle Eastern	Native	White
Beaverton PD	699	1,265	3,375	431	104	7,481
Clackamas CO SO	1,381	1,761	4,656	446	145	20,470
Eugene PD	321	836	1,228	0	0	10,535
Gresham PD	208	671	1,256	73	27	2,261
Hillsboro PD	824	759	3,798	398	55	6,340
Marion CO SO	330	456	4,114	156	17	10,304
Medford PD	83	179	836	30	4	2,968
Oregon State Police	5,949	6,926	33,282	3,308	1,445	167,010
Portland PB	1,514	5,039	4,540	552	129	14,282
Salem PD	202	309	1,914	63	29	4,151
Washington CO SO	1,825	1,425	6,813	1,064	107	13,996
Total Tier 1	**13,336**	**19,626**	**65,812**	**6,521**	**2,062**	**259,798**

Table 2.1.5. Tier 2 Race/Ethnicity by Agency

Agency	Asian or PI	Black	Hispanic	Middle Eastern	Native	White
Albany PD	126	176	978	24	16	5,809
Ashland PD	89	86	205	29	1	2,115
Bend PD	21	20	128	9	0	3,371
Benton CO SO	294	238	972	147	17	6,971
Canby PD	90	102	956	30	2	2,824
Central Point PD	55	44	336	9	0	1,701
Corvallis PD	359	243	640	141	45	4,894
Deschutes CO SO	97	83	655	30	5	4,370
Douglas CO SO	15	23	67	6	0	1,105
Forest Grove PD	108	132	1,231	41	12	2,410
Grants Pass PD	27	32	143	7	1	1,768
Hermiston PD	83	194	4,410	8	65	4,384
Hood River CO SO	71	25	432	32	0	1,297
Jackson CO SO	192	271	1,878	79	3	8,989
Keizer PD	34	51	352	13	0	606
Klamath CO SO	6	3	10	0	0	65
Klamath Falls PD	44	45	181	9	6	854
Lake Oswego PD	311	350	798	169	43	4,971
Lane CO SO	73	192	413	28	7	4,891
Lebanon PD	23	27	85	4	0	1,548
Lincoln CO SO	86	24	182	22	15	1,513
Lincoln City PD	110	59	391	23	0	1,992
Linn CO SO	69	84	446	31	27	4,637
McMinnville PD	44	39	460	10	1	1,559
Milwaukie PD	219	411	788	130	26	4,122
Newberg-Dundee PD	148	149	981	38	0	3,864
OHSU PD	9	12	11	2	0	54
Oregon City PD	148	278	676	57	24	5,066
Polk CO SO	104	82	637	34	17	2,266
Port of Portland PD	62	113	145	32	1	495
Redmond PD	139	79	948	28	1	5,108
Roseburg PD	10	20	58	5	0	979
Tigard PD	541	604	1,650	344	37	4,873
Tualatin PD	227	208	920	127	5	3,648
UO PD	22	22	30	2	0	272
West Linn PD	178	157	416	85	19	2,807
Woodburn PD	27	25	909	4	1	536
Yamhill CO SO	130	134	1,314	47	20	4,225
Total Tier 2	**4,391**	**4,837**	**25,832**	**1,836**	**417**	**112,959**

Table 2.1.6 displays the most serious disposition (i.e., outcome) of stops reported by law enforcement. Most stops result in a warning, with no further law enforcement action against the stopped individual. Arrests are more common outcomes for pedestrian stops than for traffic stops.

Table 2.1.6. Disposition by Stop Type and Agency Size (Tier)

	Tier 1		Tier 2		Tier 3	
	Traffic	**Ped**	**Traffic**	**Ped**	**Traffic**	**Ped**
None	1.6%	11.6%	3.5%	21.1%	3.2%	11.7%
Warning	60.8%	56.9%	65.9%	60.2%	71.8%	60.5%
Citation	36.0%	12.5%	29.2%	8.7%	23.9%	6.2%
Juv Summons	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%
Arrest	1.5%	19.0%	1.4%	10.0%	1.1%	21.5%

Table 2.1.7 provides information about searches resulting from traffic and pedestrian stops. Pedestrians are more likely to be searched than drivers, although they are still only searched 13% of the time. Overall, just over half of all pedestrian searches and about 42% of searches at traffic stops resulted in discovery of contraband, which most often consisted of illegal drugs.

Table 2.1.7. Search Results by Stop Type & Agency Size (Tier)

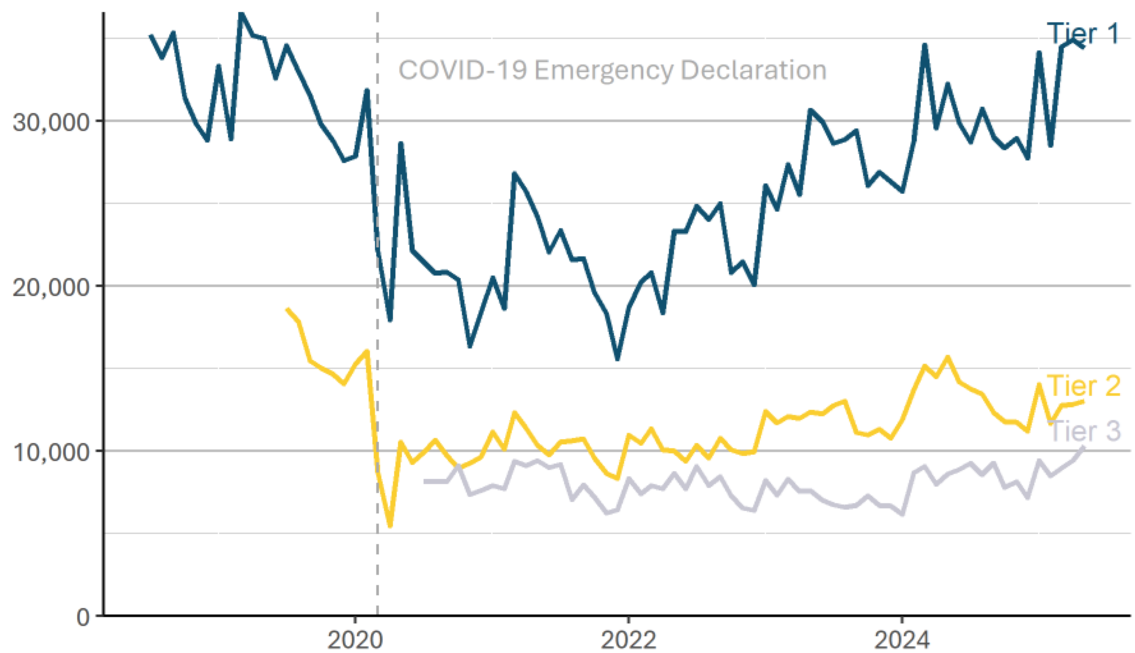
	Traffic	**Ped**
Percent of Stops with Search Conducted	0.9%	12.4%
Type of Search*		
Consent Search	33.6%	29.8%
Consent Search Denied	0.4%	0.2%
Other	66.0%	70.0%
Percent of Searches with Contraband	40.8%	52.2%
Percent of Searches with Item Seized**		
Alcohol	10.9%	1.9%
Drugs	19.3%	38.4%
Weapons	7.0%	6.6%
Stolen Property	1.2%	3.0%
Other Evidence	7.5%	6.1%
Other Non-Evidence	3.9%	5.9%
Nothing	59.2%	47.8%

*Officers may indicate multiple types; percentages may add to more than 100%.
**Multiple items may be seized; percentages may add to more than 100%.

2.2 Trends & Significant Events

Figure 2.2.1 displays all stops made by law enforcement agencies by tier since the beginning of data collection in 2018. Tier 1 agencies, the largest 12 agencies in the state, began reporting in July 2018, while medium-sized agencies (tier 2) began in 2019, and the smallest (tier 3) agencies began reporting in 2020. COVID-19 caused a significant drop in stops in 2020; from that time until the current reporting year, stops have steadily increased.

Figure 2.2.1. Monthly Stops by Tier



In March 2022, the Oregon legislature passed SB 1510[2], making several changes to public safety law. Sections 1 and 2 of the bill require officers to inform a person when they have the right to refuse a search request. Table 2.2.1 shows the impact of this on search rates.

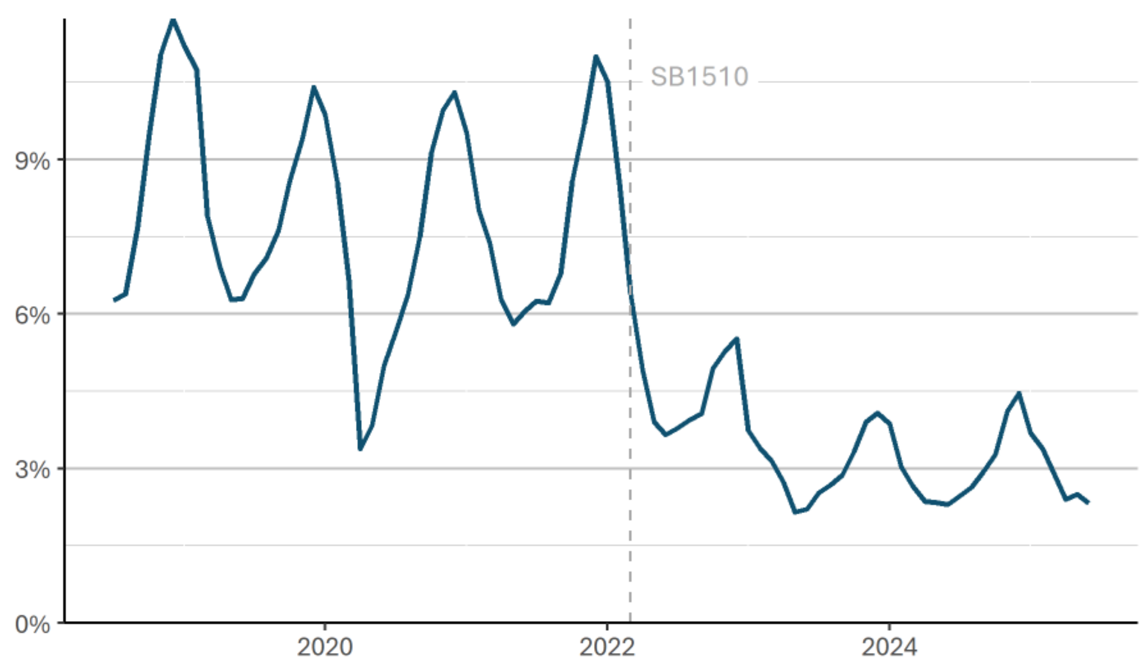
[2] The full text of SB 1510 can be found here: https://olis.oregonlegislature.gov/liz/2022R1/Downloads/MeasureDocument/SB1510/Enrolled

Table 2.2.1. Search Rates Before & After 1510

	Tier 1	Tier 2	Tier 3
Prior to SB1510			
Year 1 (18-19)	2.9%	-	-
Year 2 (19-20)	2.6%	2.8%	-
Year 3 (20-21)	2.5%	1.9%	1.4%
Year 4 (21-22)	2.2%	1.6%	0.9%
After SB1510			
Year 5 (22-23)	1.5%	1.3%	0.7%
Year 6 (23-24)	1.5%	1.2%	0.5%
Year 7 (24-25)	1.2%	0.9%	0.6%

Section 6 of SB 1510 prevents law enforcement officers from initiating a traffic stop solely because of lighting violations unless certain additional criteria are met. Figure 2.2.2 shows the percentage of traffic stops made for lighting violations since the beginning of the STOP program through the current year. Seasonal peaks during the winter and troughs during the summer are apparent for all years; since the passage of SB 1510, the overall percent of stops for lighting violations has declined from about 7% to about 3%, and seasonal peaks have become less pronounced.

Figure 2.2.2. Percentage of Stops by Month for Lighting Violations



3. Decision to Stop Analysis

Often referred to as the “gold standard” of statistical analyses examining the initial law enforcement decision to stop an individual[3], the Decision to Stop (DTS) analysis compares stops made by law enforcement officers during the day when it is light to those made at night when it is dark. The DTS analysis is built on the assumption that officers can better detect the race/ethnicity of an individual in daylight as compared to darkness and takes advantage of natural variations in daylight over the course of the year. A key advantage of this approach is that it does not require an estimated composition of the driving or residential population; however, it does assume that populations, driving behaviors, and commuting patterns are relatively consistent over the course of a year. To address potential weaknesses in these assumptions, the analysis incorporates additional control variables, such as age, gender, reason for stop, day of week, time of day, quarter or season, county stop volume, and agency stop volume.

3.1 Description of Decision to Stop Analysis

The DTS analysis compares the racial composition of people stopped during a combined inter-twilight window, which occurs during morning and evening commute times. The morning twilight window is defined as the earliest start of civil twilight to the latest sunrise, while the evening twilight window is defined as the earliest sunset to the latest end of civil twilight. Visibility during this time will vary throughout the course of the year, which makes it possible to compare stop decisions at the same time of day but in different lighting conditions. For example, the DTS analysis can compare stops made on January 10 when it was dark at 5:00pm to stops made two months later at the same time on March 10, when it was still light outside. Assuming that these two points in time should capture substantially similar driving populations, comparisons made between the race/ethnicity of stopped drivers in the light and darkness will detect whether stops are being made in a disparate fashion when race/ethnicity is visible.

The results of this analysis are presented as a ratio of the odds of a stop for a non-white driver during daylight to the odds for a similar stop in darkness compared to white drivers. If the odds ratio is not statistically different from 1.0, then the test finds no difference in stops made during daylight and darkness. If the odds ratio is greater than 1.0 and is statistically significant after accounting for additional control variables, the test concludes the odds of non-white drivers being stopped in daylight is higher than in darkness, which is taken as evidence of a racial disparity. Following best practices, the STOP Program identifies all agencies with odds ratios above 1.0 that are statistically significant at the 95 percent confidence level in any minority group at the agency level.

[3] *See* Barone et al. (2018) under Veil of Darkness analysis.

3.2 Decision to Stop Results

As shown in Table 3.2.1, three agencies (Philomath PD, The Dalles PD, and Brookings PD) showed statistically significant differences in the odds for stops of Hispanic drivers in daylight compared to darkness[4].

Table 3.2.1. Decision to Stop Results

Agency	Race/ Ethnicity	Odds Ratio
Brookings PD	Hispanic	2.29
Philomath PD	Hispanic	2.22
The Dalles PD	Hispanic	2.14

Only statistically significant results are shown.
For full results, please visit CJC's STOP dashboard.

[4] The odds ratio for Oregon State Police for Native American drivers (1.31) shows a p-value of 0.017. The odds ratio for Hispanic drivers (1.06) shows a p-value of 0.014. With the Bonferroni adjustment for five tests, these do not show statically significant differences. However, without the adjustment, the p-values are below the 0.05 threshold. For more information on the Bonferroni adjustment *see* Appendix B.

4. Stop Outcomes Analysis

The Stop Outcomes Analysis (SOA) determines what demographic and stop factors were statistically associated with citations, searches, and/or arrests resulting from a traffic or pedestrian stop. The analysis estimates whether each race/ethnic group is more likely than the white group to have a stop end in each type of outcome when controlling for all other measurable stop and demographic factors.

HB 2355 requires all law enforcement agencies to collect data regarding the disposition of stops. Because stops can have multiple dispositions (i.e., an individual could be both cited for a traffic violation and arrested for a crime) the STOP Program collects data on the most serious disposition that occurred within a single stop[5]. This means that if an individual was stopped for speeding, received a citation, and was subsequently arrested on a preexisting warrant, this individual would be recorded in the stop data as only having been arrested. This most serious disposition is the outcome used in the SOA for each stop.

4.1 Description of Stop Outcomes Analysis

Variation in enforcement outcomes could be due to time of day, day of the week, the conduct that led to the stop, or many other factors. During rush hour on a weekday, for instance, heavy traffic flows may prevent drivers from exceeding the speed limit which would reduce the likelihood of receiving a citation for speeding relative to other infractions. Variation could also be attributed to demographic factors including age or gender of the driver. The SOA uses propensity score and regression analysis to account for as many of these differences as possible and to isolate the effect, if any, of race on the disposition of the stop.

The SOA uses propensity score methods to “balance” the data before conducting statistical analyses. Propensity score methods use the estimated tendency to be included in the group of interest, or propensity score, to make that group and the comparison group look as similar as possible except for the characteristic in question, i.e., “balanced.” This approach enables STOP Program researchers to make the white comparison group statistically identical across all measured factors compared to the non-white group of interest. If all other measured variables (i.e., time of day, day of the week, gender, age, stop reason, disposition reason, and stop volume) are identical or “balanced” across the two groups then the remaining difference in outcomes is evidence of a disparity due to racial/ethnic differences (Ridgeway, 2006).

The SOA includes up to twenty sub-analyses for each agency: each possible outcome (citation, search, arrest, or any non-warning disposition) across each racial/ethnic group

[5] *See* Appendix E for more details on how the STOP Program research team determines the most serious disposition and the appropriate comparison outcomes for each type of disposition.

(Asian/PI, Black, Hispanic, Middle Eastern, and Native American). The comparison group is drawn from the group of white stops for the agency in question, balanced by propensity score as described above. Bonferroni adjustments are applied at the agency level based on the number of analyses completed for that agency.[6] For the prior 2024 report, two sets of analyses were conducted, one that accounted for the reason for disposition and one that did not, depending on the quality of data submitted by the agency. Starting this year, one set of analyses is conducted that includes the reason for the disposition if the agency submitted the additional data.

4.2 Stop Outcomes Results

As with the Decision to Stop analysis in the previous section, the SOA conducted here includes two years of data for all agencies. Table 4.2.1 reports agency-level results for agencies where a statistically significant difference is found for either a search or arrest outcome. These agencies may also have statistically significant differences for the citation and any non-warning outcomes tests, the results of which are presented in Table 4.2.2.

Three agencies (Gilliam CO SO, Morrow CO SO, and Portland PB) had statistically significant differences for searches of Hispanic drivers. Portland PB also showed a statistically significant difference for searches and arrests for Black drivers; Oregon State Police had a statistically significant difference for searches of Black drivers.

Table 4.2.1. Stop Outcome Analyses - Search & Arrest

		Search		Arrest	
Agency	Race/ Ethnicity	Actual	Pred	Actual	Pred
Gilliam CO SO	Hispanic	2.3%	0.5%	-	-
Morrow CO SO	Hispanic	1.1%	0.2%	-	-
Oregon State Police	Black	1.5%	1.1%	-	-
Portland PB	Black	6.8%	4.1%	5.2%	3.7%
Portland PB	Hispanic	5.7%	4.1%	-	-

Only statistically significant results are shown.
For full results, please visit CJC's STOP dashboard.

Table 4.2.2 reports agency-level results for agencies where a statistically significant difference is found for a citation or any non-warning disposition. Many agencies had statistically significant differences for Hispanic drivers, reflecting, in part, the fact that a larger number of Hispanic people in the driving population allows more agencies to meet

[6] Low sample sizes for certain groups or a lack of comparability between groups for a given agency could prevent some of these sub-analyses from being completed. In these cases, the Bonferroni adjustment is changed accordingly. For more details on the Bonferroni adjustment *see* Appendix B.

the sample-size threshold for the analysis. Clackamas CO SO, Oregon State Police, and Tigard PD also had statistically significant differences for Asian or Pacific Islander drivers. Oregon State Police also had a statistically significant difference for Middle Eastern drivers.

Table 4.2.2. Stop Outcome Analyses – Citation & Any Non-Warning Outcome

		Citation		Any Outcome	
Agency	**Race/ Ethnicity**	**Actual**	**Pred**	**Actual**	**Pred**
Astoria PD	Hispanic	33.6%	21.8%	34.1%	22.3%
Beaverton PD	Hispanic	28.8%	26.7%	-	-
Canby PD	Hispanic	36.8%	32.8%	38.7%	34.8%
Cannon Beach PD	Hispanic	25.4%	13.5%	25.4%	13.7%
Clackamas CO SO	Asian or PI	36.7%	31.6%	38.3%	34.0%
Deschutes CO SO	Hispanic	15.6%	12.1%	17.7%	14.3%
Eugene PD	Hispanic	44.2%	39.7%	46.4%	42.7%
Gilliam CO SO	Hispanic	69.6%	60.6%	70.3%	61.2%
Hermiston PD	Hispanic	28.3%	23.5%	29.8%	25.1%
Linn CO SO	Hispanic	45.6%	39.1%	43.5%	37.9%
Medford PD	Hispanic	22.8%	18.8%	25.0%	21.7%
Ontario PD	Hispanic	52.0%	36.4%	52.0%	36.5%
Oregon City PD	Hispanic	39.1%	34.8%	41.3%	36.6%
Oregon State Police	Asian or PI	38.0%	31.4%	38.4%	31.8%
Oregon State Police	Black	41.7%	38.8%	42.4%	39.4%
Oregon State Police	Hispanic	42.8%	39.2%	43.7%	39.8%
Oregon State Police	Middle Eastern	41.4%	31.5%	41.6%	31.8%
Phoenix PD	Hispanic	45.7%	38.3%	46.4%	39.2%
Seaside PD	Hispanic	15.5%	11.3%	15.8%	11.5%
Sutherlin PD	Hispanic	60.9%	49.0%	61.1%	49.3%
Talent PD	Hispanic	43.0%	34.9%	44.0%	36.2%
Tigard PD	Asian or PI	27.0%	21.5%	28.5%	23.6%
Tigard PD	Hispanic	33.8%	29.3%	35.9%	31.9%
Tualatin PD	Hispanic	50.2%	41.2%	51.1%	42.1%
Umatilla CO SO	Hispanic	25.0%	17.9%	27.2%	20.4%
Washington CO SO	Hispanic	22.9%	21.4%	25.6%	24.1%

Only statistically significant results are shown.
For full results, please visit CJC's STOP dashboard.

5. Search Findings Analysis

The Search Findings analysis (SFA) is based on a hit-rate model originally developed in the field of economics to study discrimination in mortgage loan decision making[7]. These models have been adapted to analyses of law enforcement for decades; the most common adaptation, and the one used in this report, is the KPT Hit-Rate model[8].

In this section, we refer to a search conducted by a law enforcement officer during a traffic or pedestrian stop that results in discovery of contraband as a "successful" search. The SFA examines whether the likelihood of a successful search differs across racial/ethnic groups. The model assumes that officers decide to conduct a search based on visual and other contextual clues (e.g. location, furtive movements, odors, etc.). The model further assumes that drivers will tend to decide whether or not carrying contraband is worth the risk based on their perceived likelihood of being searched. The model does not assume that all groups are equally likely to carry contraband; rather, if one group is particularly likely to carry contraband, then officers are more likely to search them. Collectively, members of that group will perceive a higher risk and will respond by reducing the rate at which they carry contraband until an equilibrium is reached. However, if officers are basing their decision to search on race/ethnicity and not consistent with the actual rate at which the group carries contraband, then their rate of successful searches will decrease; this is interpreted as evidence of a disparity due to racial/ethnic differences.

5.1 Agency-Level Search Findings Results

As in the previous two sections, analyses in this section utilized two years of data for all agencies. The Search Findings analyses were performed for each agency for up to five minority racial/ethnic groups (Black, Hispanic, Asian or PI, Middle Eastern, and Native), comparing the successful search rate of each to that of the white majority group. Significant results for these analyses are presented in Table 5.1.1.

Table 5.1.1. Search Findings Analysis

Agency	White	Asian or PI	Black	Hispanic	Middle Eastern	Native
Marion CO SO	13.6%	-	-	6.2%	-	-
Oregon State Police	61.3%	-	-	50.5%	-	-

Only statistically significant results are shown.
For full results, please visit CJC's STOP dashboard.

[7] *See* Becker (1957), Becker (1993)
[8] *See* Knowles, Persico, and Todd (2001)

While all agencies have differences in search success rates across racial/ethnic groups, most of these small differences can be attributed to random chance rather than policies or practices leading to disparate treatment of different groups. The Search Findings analyses for Marion CO SO and Oregon State Police found statistically significant differences in the success rate of searches conducted with Hispanic drivers compared to white drivers.

6. Conclusions

The data contained in this report are intended to be used as a tool for law enforcement, community members, researchers, policy makers, and others to focus training and provide technical assistance to agencies found to have evidence of disparities in stops for minority groups. As described previously, STOP program researchers utilized three rigorous statistical analyses, consistent with best practices, to identify differences in the treatment of drivers during police stops in Oregon. The use of these three tests allows the STOP program researchers to evaluate numerous decision points before and during a stop.

To determine if statistically significant results require further analysis and support from the STOP Program and its partners at the Department of Public Safety Standards and Training (DPSST), the following criteria must be met: (1) A difference in an individual analysis must have met the 95 percent confidence level for it to be statistically significant. This means STOP program researchers must be at least 95 percent confident that differences identified by the analyses were not due to random chance. And (2) following best practices, for a law enforcement agency to be identified as one requiring further analysis as well as DPSST technical assistance, it must be identified as having a statistically significant difference in two of the three analytical tests performed on the STOP data. However, DPSST has and will continue to provide technical assistance to any agency, regardless of the number of analyses that are statistically significant.

Using the above-mentioned analyses and thresholds, Oregon State Police had statistically significant results in two tests. This agency, as well as several other agencies with a statistically significant result in one test of this report, have initiated additional analysis of the STOP data. Regardless of whether an agency is officially referred to DPSST, the CJC urges each agency to scrutinize their full set of results[9] and engage with DPSST on any results that show a statistically significant difference.

[9] Full results may be found on the CJC's dashboards here: https://public.tableau.com/app/profile/cjcdashboards/viz/S_T_O_P_StatisticalTransparencyofPolicing/Introduction

7. Oregon Law Enforcement Contacts and Data Review Committee Report

7.1 LECC Background

The Oregon Law Enforcement Contacts and Data Review Committee (LECC) is a statewide committee tasked with assisting Oregon law enforcement agencies in creating equitable outcomes for Oregonians. The LECC was initially created in 2001 with the passage of SB 415. In 2015, HB 2002 created a standard definition of profiling[10], required agencies to adopt procedures for submitting copies of racial profiling complaints to the LECC, and tasked the LECC with establishing policies for receiving and forwarding profiling complaints to the general public (see ORS 131.915, ORS 131.920, and ORS 131.925). The administration of the LECC was transferred to Portland State University in 2007, where it remained until 2019 when it was transferred to the CJC by order of HB 5050, Section 13.

This report summarizes the information found in the profiling complaints the LECC received from Oregon law enforcement agencies in calendar years 2023 and 2024. This information is provided to meet the reporting requirements described above and is not used to refer an agency to DPSST for technical assistance.

7.2 Summary of 2023 & 2024 LECC Reports

Table 7.2.1 summarizes law enforcement agency reporting for 2023 and 2024. In 2023, 116 of 154 (76.0%) law enforcement agencies reported the number of profiling complaints they received and in 2024, 122 of 154 (79.2%) law enforcement agencies reported the number of profiling complaints they received for each respective calendar year. Of those agencies that reported in 2024, 29 (18.8%) reported at least one complaint, and across those 29 agencies there were a total of 84 complaints, compared to 75 complaints reported across 22 agencies in 2023.

Table 7.2.1. Law Enforcement Reporting Compliance

	2023	2024
Agencies Reporting	116	122
Total Reported Complaints	75	84
Agencies Reporting No Complaints	94	93
Agencies Reporting 1+ Complaints	22	29

[10] The law defines profiling as when “a law enforcement agency or a law enforcement officer targets an individual for suspicion of violating a provision of law based solely on the real or perceived factor of the individual’s age, race, ethnicity, color, national origin, language, gender, gender identity, sexual orientation, political affiliation, religion, homelessness or disability, unless the agency or officer is acting on a suspect description or information related to an identified or suspected violation of a provision of law.”

Table 7.2.2 shows the dispositions of complaints that were reported in 2023 and 2024. The most common disposition in both years was "unfounded", followed by "not sustained" in 2024 and "no basis for further investigation" in 2023. The disposition of one complaint in 2023 was unknown/not provided and is therefore excluded from the following table. Total dispositions per year may not be equal to total number of complaints per year, as agencies have the option to select more than one disposition per complaint.

Table 7.2.2. Reported Complaints by Disposition

Disposition	2023	2024
Exonerated	8	11
Not Sustained	4	6
Unfounded	43	50
Administrative Closure	7	7
No Basis for Further Investigation	11	6
Other	4	7

Table 7.2.3 shows the number of complaints reported by agency in 2023 and 2024. Across those two years, Oregon State Police had the highest complaint volume with 22 complaints, which is consistent with their position as the largest law enforcement agency by employed officers in the state. The agencies with the next highest report volume over that period were Portland PB with 19 reported complaints and Bend PD with 17 reported complaints.

Table 7.2.3. Complaints by Agency

Agency	2023	2024
Albany PD	1	1
Ashland PD	0	3
Beaverton PD	5	2
Bend PD	7	10
Canby PD	0	1
Clackamas CO SO	6	3
Corvallis PD	2	2
Deschutes CO SO	1	0
Eagle Point PD	1	0
Eugene PD	3	6
Hillsboro PD	4	7
Jackson CO SO	2	1
Keizer PD	1	2
Klamath CO SO	0	2
Lake Oswego PD	2	2
Lane CO SO	6	0
Marion CO SO	1	1
Medford PD	2	2
Milwaukie PD	0	1
Multnomah CO SO	2	1
Newberg-Dundee PD	0	2
Ontario PD	0	1
Oregon City PD	0	2
Oregon State Police	11	11
Portland PB	10	9
Rainier PD	0	1
Redmond PD	1	0
Salem PD	0	1
Springfield PD	4	4
The Dalles PD	0	3
Tigard PD	1	1
Washington CO SO	2	0
Woodburn PD	0	1
Yamhill CO SO	0	1
Total	**75**	**84**

Works Cited

Austin, Peter C. “Balance Diagnostics for Comparing the Distribution of Baseline Covariates between Treatment Groups in Propensity-Score Matched Samples.” *Statistics in Medicine* 28, no. 25 (2009): 3083–3107. https://doi.org/10.1002/sim.3697.

Austin, Peter C. “Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research.” *Communications in Statistics - Simulation and Computation* 38, no. 6 (September 2009): 1228–34. https://doi.org/10.1080/03610910902859574.

Austin, Peter C., and Elizabeth A. Stuart. “Moving towards Best Practice When Using Inverse Probability of Treatment Weighting (IPTW) Using the Propensity Score to Estimate Causal Treatment Effects in Observational Studies.” *Statistics in Medicine* 34, no. 28 (March 2015): 3661–79. https://doi.org/10.1002/sim.6607.

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

Gary S. Becker, “The Evidence against Banks Doesn’t Prove Bias.” Business Week, April 19, 1993.

Beninati, Nancy A., Tiffany Jantz, et al. “Racial and Identity Profiling Advisory Board Report.”, 2022. https://oag.ca.gov/system/files/media/ripa-board-report-2022.pdf

Engel, Robin S., Rob Tillyer, Jennifer Calnon Cherkauskas, and James Frank. 2007. Traffic Stop Data Analysis Study: Year 1 Final Report. November 1, 2007.

Fazzalaro, J., Ken Barone, et al. “Connecticut Racial Profiling Prohibition Project: State of Connecticut Traffic Stop Data Report October 1, 2013 – May 31, 2014.” Institute for Municipal and Regional Policy Central Connecticut State University (2014). https://www.ctrp3.org/analysis-reports/reports.

Grogger, Jeffrey, and Greg Ridgeway. “Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness.” *Journal of the American Statistical Association* 101, no. 475 (2006): 878–87. https://www.rand.org/pubs/reprints/RP1253.html.

“Racial Profiling in Decisions to Search: a Preliminary Analysis Using Propensity-Score Matching.” *Policing: An International Journal of Police Strategies & Management* 36, no. 4 (April 2013). https://doi.org/10.1108/pijpsm-07-2013-0074.

Higgins, George E., Wesley G. Jennings, et al. 2011. “Racial Profiling in Decisions to Search: A Preliminary Analysis Using Propensity-Score Matching.” *International*

Journal of Police Science & Management 13 (4): 336–47. https://doi.org/10.1350/ijps.2011.13.4.232.

Higgins, George E., Melissa L. Ricketts, et al. "Race and Juvenile Incarceration: A Propensity Score Matching Examination." *American Journal of Criminal Justice* 38, no. 1 (2013): 1–12. https://doi.org/10.1007/s12103-012-9162-6.

Niel, Roland, and Christopher Winship. "Methodological Challenges and Opportunities in Testing for Racial Discrimination in Policing." *Annual Reviews of Criminology* 12 (October 2018). https://doi.org/10.1146/annurev-criminol-011518-024731.

Knowles, John, Nicola Persico, and Petra Todd. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109, no. 1 (2001): 203–29. https://doi.org/10.1086/318603.

Persico, Nicola. "Racial Profiling? Detecting Bias Using Statistical Evidence." *Annual Review of Economics* 1, no. 1 (February 2009): 229–54. https://doi.org/10.1146/annurev.economics.050708.143307.

Ridgeway, Greg. "Assessing the Effect of Race Bias in Post-Traffic Stop Outcomes Using Propensity Scores." *Journal of Quantitative Criminology* 22, no. 1 (March 2006): 1–29. https://doi.org/10.1007/s10940-005-9000-9.

Simoiu, Camelia, Sam Corbett-Davies, and Sharad Goel. "The Problem of Infra-Marginality in Outcome Tests for Discrimination." *The Annals of Applied Statistics* 11, no. 3 (2017): 1193–1216. https://doi.org/10.1214/17-aoas1058.

"Stata Treatment-Effects Reference Manual: Potential Outcomes/Counterfactual Outcomes." 2019. StataCorp. https://www.stata.com/manuals16/te.pdf.

Stringer, Richard J., and Melanie M. Holland. "It's Not All Black and White: A Propensity Score Matched, Multilevel Examination of Racial Drug Sentencing Disparities." *Journal of Ethnicity in Criminal Justice* 14, no. 4 (2016): 327–47. https://doi.org/10.1080/15377938.2016.1187239.

"The Stanford Open Policing Project." openpolicing.stanford.edu. https://openpolicing.stanford.edu/.

Vito, Anthony G., Elizabeth L. Grossi, and George E. Higgins. "The Issue of Racial Profiling in Traffic Stop Citations." *Journal of Contemporary Criminal Justice* 33, no. 4 (2017): 431–50. https://doi.org/10.1177/1043986217724537.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second. MIT Press.

Appendix A List of Law Enforcement Agencies by Tier

Table A.1. Tier 1 Agencies

Beaverton PD	Hillsboro PD	Oregon State Police
Clackamas CO SO	Marion CO SO	Portland PB
Eugene PD	Medford PD	Salem PD
Gresham PD	Multnomah CO SO***	Washington CO SO

*No stops submitted in current report period.
**Agency no longer exists or staffed.
***Final data for current year is unavailable.

Table A.2. Tier 2 Agencies

Albany PD	Jackson CO SO	OHSU PD
Ashland PD	Keizer PD	Oregon City PD
Bend PD	Klamath CO SO	Polk CO SO
Benton CO SO	Klamath Falls PD	Port of Portland PD
Canby PD	Lake Oswego PD	Redmond PD
Central Point PD	Lane CO SO	Roseburg PD
Corvallis PD	Lebanon PD	Springfield PD***
Deschutes CO SO	Lincoln CO SO	Tigard PD
Douglas CO SO	Lincoln City PD	Tualatin PD
Forest Grove PD	Linn CO SO	UO PD
Grants Pass PD	McMinnville PD	West Linn PD
Hermiston PD	Milwaukie PD	Woodburn PD
Hood River CO SO	Newberg-Dundee PD	Yamhill CO SO

*No stops submitted in current report period.
**Agency no longer exists or staffed.
***Final data for current year is unavailable.

Table A.3. Tier 3 Agencies

Astoria PD	Hubbard PD	Powers PD*
Aumsville PD	Independence PD	Prineville PD
Baker CO SO	Jacksonville PD	Rainier PD
Baker City PD	Jefferson CO SO	Reedsport PD
Bandon PD	John Day PD**	Rockaway Beach PD**
Black Butte Ranch PD	Josephine CO SO	Rogue River PD
Boardman PD	Junction City PD	Sandy PD
Brookings PD	King City PD	Scappoose PD*
Burns PD	La Grande PD	Seaside PD
Butte Falls PD*	Lake CO SO*	Sherman CO SO
Cannon Beach PD	Madras PD	Sherwood PD
Carlton PD	Malheur CO SO	Silverton PD
Clatsop CO SO	Malin PD	St. Helens PD*
Coburg PD	Manzanita PD	Stanfield PD
Columbia CO SO	Merrill PD*	Stayton PD
Columbia City PD	Milton-Freewater PD	Sunriver PD
Coos Bay PD	Molalla PD	Sutherlin PD
Coos CO SO	Monmouth PD	Sweet Home PD
Coquille PD	Morrow CO SO	Talent PD
Cottage Grove PD	Mt. Angel PD	The Dalles PD
Crook CO SO	Myrtle Creek PD	Tillamook CO SO
Curry CO SO	Myrtle Point PD	Tillamook PD
Dallas PD	Newport PD	Toledo PD
Eagle Point PD	North Bend PD	Turner PD
Enterprise PD	Nyssa PD*	Umatilla CO SO
Florence PD	OSU PD	Umatilla PD
Gearhart PD	Oakridge PD	Union CO SO
Gervais PD*	Ontario PD*	Union Pacific Railroad PD
Gilliam CO SO	PSU CPS*	Vernonia PD
Gladstone PD	Pendleton PD	Wallowa CO SO
Gold Beach PD	Philomath PD	Warrenton PD
Grant CO SO*	Phoenix PD	Wasco CO SO
Harney CO SO	Pilot Rock PD	Wheeler CO SO
Hines PD*	Port Orford PD	Winston PD
Hood River PD	Portland FB*	Yamhill PD

*No stops submitted in current report period.
**Agency no longer exists or staffed.
***Final data for current year is unavailable.

Appendix B Background

B.1. House Bill (HB) 2355 (2017)

Efforts by the State of Oregon to collect data regarding stops of individuals made by law enforcement began with the passage of HB 2433 in 1997, which mandated that law enforcement agencies develop written policies related to traffic stop data collection. Following the passage of HB 2433, the Governor's Public Safety Policy and Planning Council recommended that a full statewide data collection effort be initiated legislatively. It was not until 2001, however, that the Legislature again considered the collection of police stop data. In Senate Bill (SB) 415 (2001), the Legislature created the Law Enforcement Contacts Policy & Data Review Committee (LECC), which provided for the voluntary collection of stop data by law enforcement agencies, and for analysis of collected data by the LECC.

Apart from a brief hiatus from 2003 to 2005, the LECC engaged with law enforcement agencies throughout the 2000s and 2010s to examine stop data. During this period, however, challenges were encountered related to the creation of a comprehensive database of stops, given that few agencies in Oregon collected stop data and/or elected to partner with the LECC for data analysis. As a remedy, the Legislature passed HB 2355 in 2017, which led to the creation of the STOP Program. The STOP Program represents the culmination of the process started in 1997 and is the first statewide data collection and analysis program focused on traffic and pedestrian stops in Oregon.

HB 2355, which is codified in ORS 131.930 et seq., created a statewide data collection effort for all officer-initiated traffic[11] and pedestrian[12] stops that are not associated with calls for service. The aim of HB 2355 was to collect data regarding discretionary stops, as opposed to stops where discretion was absent. The CJC, in partnership with the Oregon State Police and the Department of Justice, worked to develop a standardized method for collecting the data elements required by statute, which include data regarding both the stop itself as well as demographic characteristics of the stopped individual (for a description of the STOP Program data elements utilized in this report, see Section 2.3.1.).

[11] Officer initiated traffic stops are defined as any "detention of a driver of a motor vehicle by a law enforcement officer, not associated with a call for service, for the purpose of investigating a suspected violation of the Oregon Vehicle Code" (ORS 131.930 § 4). Included with traffic stops are stops made of individuals operating bicycles. Stops involving operators of watercraft, however, are not included in the stop database, as watercraft violations fall outside the Oregon Vehicle Code (see ORS Chapter 830).
[12] Officer initiated pedestrian stops are defined as "a detention of a pedestrian by a law enforcement officer that is not associated with a call for service. The term does not apply to detentions for routine searches performed at the point of entry to or exit from a controlled area" (ORS 131.930 § 3).

To implement the STOP Program, HB 2355 established a three-tiered approach, whereby the largest law enforcement agencies in the state would begin to collect data and report in the first year, followed by medium and small agencies in the next two years, respectively. Table B.1 reports the inclusion criteria for each tier as well as the data collection and reporting dates. A full list of agencies broken down by tier can be found in Appendix A.

Table B.1. Three-Tier Reporting Approach in HB 2355 (2017)

Tier	Officers per Agency	Data Collection Began	Reporting Began
Tier 1	100+	July 1, 2018	July 1, 2019
Tier 2	25 - 99	July 1, 2019	July 1, 2020
Tier 3	1 - 24	July 1, 2020	July 1, 2021

In the development of the standardized data collection method, the primary goals of the STOP Program were to ensure that (1) all data collected are as accurate and complete as possible, (2) data collection methods are minimally impactful to each agency's workload and free or affordable for each agency, and (3) data collection methods are minimally impactful on law enforcement personnel to ensure that officer safety is not negatively impacted during the data collection process. As such, the STOP Program contracted with a technology vendor to develop software that could both collect and receive stop data via multiple submission methods.

The STOP Program software solution includes three methods of data collection/input. First, the software can receive data from local agencies' records management systems. Under this approach, an agency with the ability to collect stop data through its own preexisting systems can integrate stop data collection requirements into their in-car or e-ticketing system, recording the data internally before submitting the required data fields to the STOP Program in electronic format via a secure data connection. Second, for agencies that either cannot or choose not to integrate the required stop data fields into their preexisting systems, the STOP Program provides a free web application that can be loaded on officers' in-car computers (or other similar devices, like iPads) and used when a stop is made that requires data collection under the requirements of HB 2355. Third, the STOP Program provides mobile applications free of charge for both iPhones and Android phones through which officers can submit stop data for qualifying police-citizen interactions under HB 2355.

B.2. Methodological Approach

Background

The formal examination of police traffic and pedestrian stop data began in the U.S. in the mid-1990s. Advocacy groups have long cited anecdotal evidence supporting the notion that law enforcement applies different standards to minority drivers and pedestrians. Specific and systematic measurement of police practices during citizen stops, however, did not occur until court cases alleging racial bias in policing were filed (see Wilkins v. Maryland State Police (1995) and State of New Jersey v. Soto et al. (1996)). Building on this foundation, the US Department of Justice and several other organizations began hosting conferences related to the improvement of police-community relationships with a specific focus on the collection, analysis, and public reporting of traffic and pedestrian stop data. In response, many states mandated the collection of traffic stop data. In states that had yet to require data collection, many local jurisdictions and departments started collecting and analyzing stop data on their own.

During the approximately three decades that stop data have been studied, the majority of analyses have relied on population-based benchmarks. This approach compares the demographic breakdown of stopped individuals to residential census data. Benchmarks are both intuitive and relatively simple to calculate, but the comparisons that result are overly simplistic and often biased or invalid (see Neil and Winship 2018). The concerns regarding population-based benchmarks are many and discussed at length in academic research as well as in a companion research brief released by the STOP Program in 2018[13]. The central thrust of these critiques is that the driving population in a given area (which forms the pool of individuals at risk for being stopped) is often unrelated to the residential population of that area. There are myriad reasons for this (e.g., commuting patterns and tourism), all of which lead to a disjuncture between residential demographics and driving population demographics in a given area.

Oregon STOP Program Analyses

To address the shortcomings of population-based benchmark analyses, researchers and statisticians have developed several statistical approaches that allow for more precise and less biased estimates of differential outcomes in stop data. The STOP Program relies on three of these analyses. The decision to utilize multiple tests was based on two factors.

First, there are multiple opportunities within a police-community member interaction where disparate treatment may be present. Initially, it is tempting to view a stop as a single instance of law enforcement-citizen contact that can be assessed for the presence or

---

[13] *See* STOP Program Research Brief: Analytical Approaches to Studying Stops Data (October 2018), which can be found at https://www.oregon.gov/cjc/stop/Documents/Traffic_Stop_Research_Memo_Final_Draft-10-16-18.pdf

absence of discriminatory behavior by a law enforcement agent. Race/ethnicity could be a factor in each decision to stop, search, cite, and/or arrest an individual. This distinction is critical, because both the data and analytical techniques required to analyze the various decision points found in a single stop differ. STOP Program researchers address each of these decision points separately.

Second, while the statistical tests utilized by the STOP Program represent the gold standard[14] in law enforcement stop data analyses, the application of multiple tests is also necessary to address the possibility that any single analysis could produce false positives or false negatives. Statistics are estimates and some degree of error could influence results, whether stemming from data collection practices, errors in reporting, or the like. The three analyses utilized by the STOP Program are[15]:

Decision to Stop Analysis. The Decision to Stop analysis takes advantage of natural variations in daylight and darkness throughout the year to examine the initial decision to stop an individual. Based on the assumption that it is easier for an officer to discern race/ethnicity during the day when it is light than during the night when it is dark, this analysis compares stop rates for minority individuals to those for white individuals during the time windows surrounding sunrise and sunset. If, as demonstrated by the statistics that result from the Decision to Stop analysis, minority individuals are more likely to be stopped in the daylight when race/ethnicity is easier to detect, then there is evidence of a disparity.

Stop Outcomes Analysis. The Stop Outcomes analysis examines matched groups using a statistical technique called propensity score analysis to explore whether disparities exist in stop outcomes (i.e., citations, searches, or arrests). This test matches stop data between two groups based on all available characteristics, only allowing race/ethnicity to vary between the two groups being compared. This means that the analysis compares white and Black groups, for example, who have identical proportions of gender, age, stop time of the day, stop day of the week, reason for the stop, season of the year, whether the stop was made in the daylight, and agency and county stop volumes. The test determines whether one group is cited more often, searched more often, or arrested more often. If, after matching on all the factors listed above and further controlling for these factors with regression analysis, minority individuals are either cited, searched, or arrested more often than similarly situated white individuals, then there is evidence of a disparity.

Search Findings Analysis. The Search Findings analysis compares relative rates of successful searches (i.e., those resulting in the seizure of contraband) across racial/ethnic

[14] The analytical approach utilized by the STOP Program is based on the work conducted by the Connecticut Racial Profiling Prohibition Project, which employs research and analytical techniques that have been peer reviewed by academics who specialize in the study of racial/ethnic disparities in law enforcement contacts.
[15] More detailed and technical descriptions of these analyses can be found in Appendices E, F, and G.

groups. It is based on the assumption that if search decisions by officers are based on race/ethnicity neutral criteria, then search success rates should be similar, if not identical, across different racial/ethnic categories. If, however, search success rates differ and the search success rates for minority individuals are significantly lower than those reported for white individuals, then there is evidence of a disparity.

Analytical Sample

A total of 624,255 records were submitted by 139 tier 1, tier 2, and tier 3 agencies during the seventh year of data collection. As required by HB 2355 (2017), agencies submit numerous data points, including information regarding the stop itself as well as information regarding the stopped individual. While HB 2355 is clear regarding the data elements the STOP Program is required to collect, it did not define these elements. To fill this gap, the Oregon State Police assembled a group of stakeholders, which included representatives from law enforcement, community groups, state agencies, and the Oregon Legislature, to formally define the following data elements required for submission by the statute:

Date and Time the Stop Occurred. Law enforcement personnel are required to record the date (month/day/year) and time that the stop occurred. The data is further categorized into day of the week and season. Stop times are recorded on a 24-hour clock ("military time") and converted to 12-hour clock time for this report.

Type of Stop. As required by HB 2355, both traffic and pedestrian stops are reported by law enforcement. Included in the database is a binary variable denoting whether the record is for a traffic or pedestrian stop. During the analysis of this data element, it was discovered that in a number of cases, stops were coded as "pedestrian" that were clearly for moving or other traffic violations. Similarly, some stops were coded as "traffic" that were clearly violations by pedestrians. These stops were recoded by STOP Program researchers to the appropriate categories[16].

Perceived Race/Ethnicity of Subject. Law enforcement officers are required by HB 2355 to record their perception of a subject's race/ethnicity (only the perceived race/ethnicity of the driver, not the passenger(s), is reported for traffic stops). The categories included in the data collection are: white, Black, Hispanic, Asian or Pacific Islander (hereinafter, Asian/PI), Native American, and Middle Eastern. The STOP data solution combines race and ethnicity into a single variable, and allows for one option to be selected. This differs from defined Census categories[17], and doesn't account for the additional nuance of multiple races and

---

[16] For instance, 212 Year 6 stops were labeled as traffic stops, but the citation code was ORS 814.070, which refers to a pedestrian improperly proceeding along a highway. These stops were reclassified by CJC researchers as pedestrian stops.

[17] *See* U.S. Census Bureau at https://www.census.gov/topics/population/race/about.html and https://www.census.gov/topics/population/hispanic-origin/about.html

individuals who are not white and Hispanic. However, to simplify the data collection process and in recognition of the challenges for law enforcement officers to record perceived race/ethnicity, a single combined variable is available.

Perceived Gender of Subject. Law enforcement officers are required by HB 2355 to record their perception of a subject's gender (for traffic stops, only the perceived gender of the driver, not the passenger(s) is reported). The categories included in the data collection are male, female, and nonbinary.

Perceived Age of Subject. Law enforcement officers are required by HB 2355 to record their perception of a subject's age, which is entered as a whole number (for traffic stops, only the perceived age of the driver, not the passenger(s) is reported).

Legal Basis for the Stop. The legal basis for each stop is reported to the STOP Program. This includes violations of an Oregon statute, a municipal traffic code, a municipal criminal code, a county code, TriMet rules/regulations, or a federal statute.

Oregon Statutory Violations Detail. For violations of an Oregon statute, which represent over 90 percent of all stops, law enforcement provides the specific ORS code corresponding to the violation. In this data element, over 700 different ORS codes were reported during the first year of data collection. To simplify the use of this information in the models conducted in the remainder of this report, the STOP Program research team aggregated these violations into the following categories: serious moving violations; minor moving violations; equipment, cell phone, and seat belt violations; registration and license violations; and "other" violations (e.g., criminal offenses, camping violations)[18].

Disposition of the Stop. The final disposition for each stop is reported by law enforcement officers. The categories included in the data collection are: nothing; warning; citation; juvenile summons; and arrest. It is important to note that stops can have multiple dispositions (e.g., an individual could be both cited for a traffic violation and arrested for a crime), however, only the final, or most serious, disposition is reported into the STOP Program database. This means that the categories for warnings, citations, and juvenile summons could be undercounted. For the analyses examining stop disposition in this report, the juvenile summons category was removed from the data set because the Year 6 data included only 125 juvenile summons (0.02 percent of all dispositions).

Whether a Search was Conducted. Law enforcement officers report whether or not a search was conducted, which is recorded as a binary in the STOP Program database. Searches incident to arrest and other non-discretionary searches are not recorded.

Justification for the Search. Law enforcement officers can provide several bases for a search using the following categories: consent search; consent search denied; or "other"

[18] Details on the offenses falling into each category are available upon request.

search. The “other” search category includes frisks, probable cause searches, and other administrative searches. Multiple data points are allowed so that the data can include several search justifications. For example, if an officer initially requests to search an individual but consent is not given, an officer may then perform a search based on probable cause. In this example, the officer could record both “consent search denied” as well as “other search” into the database.

Search Findings. Seven categories were predefined by the STOP Program stakeholder engagement group with regard to search findings. These categories are: nothing; alcohol; drugs; stolen property; weapon(s); other evidence; and other non-evidence. Officers are permitted to report up to six search findings to the STOP database so that searches resulting in the seizure of multiple types of contraband are properly documented.

Stop Location. Law enforcement officers are required by HB 2355 to record the location of the stop. The form in which these data are submitted varies by agency. Some agencies report latitude and longitude X,Y coordinates, while others submit textual descriptions of the location (e.g., 123 Main Street, intersection of Main and Maple Streets).

The STOP Program created four of its own variables for use in its analyses. Following best practices, variables representing both the daily agency stop volume and daily county stop volume were created. For agency stop volume, the aggregate number of stops for a single date are divided by the maximum number of daily stops for the agency unit in question. Thus, if an agency stopped 1,000 drivers on its busiest day, this would be the denominator against which all other days would be compared. A measure of the county stop volume would be calculated the same way, although all stops made by agencies within a single county would be included together. Additionally, variables representing sunrise time and sunset time were made for use in the Decision to Stop and Stop Outcomes analyses[19]. Every traffic stop is defined to have occurred in daylight or darkness based on the date, time, and location of the stop. Astronomical data from the United States Naval Observatory is used to determine the sunrise, sunset, and start and end of civil twilight.

In 2019 and 2021, the STOP program added two additional optional data categories. First, in July 2019, the STOP Program began collecting data on whether the stopped individual was perceived prior to the police stop. This data point is particularly valuable in the Decision to Stop analysis which relies on the assumption that the driver’s race will be harder for the officer to perceive in darkness. Data on whether the subject, and their race, was perceived prior to the stop enables analysts to test the Decision to Stop assumption. Second, beginning in February 2021, law enforcement agencies were able to start submitting additional data to the STOP Program on the reason for the most serious stop

[19] Sunrise time and sunset time were also used for analysis conducted for the 2019, 2020, and 2021 STOP reports. They were not explicitly listed in this section previously; however their construction is the same as in the past.

disposition. Previously, for example, if an officer stopped someone for a moving violation but the stop ended in arrest because of an outstanding warrant, analysts would only be able to see a moving violation ending in arrest. This additional data point allows the STOP program analysts to more accurately account for the reason for the stop disposition. These additional data points are submitted voluntarily by STOP agencies.

While the overall number of records was substantial, the STOP Program team faced challenges regarding sample size when the data were broken down into subsamples based on race/ethnicity and agency. In cases where the sample size is too small, statistical analyses cannot be conducted. Table B.2 lists the sample size thresholds for each test to be conducted; samples sizes are based on a minimum number of observations (stops by race) for the Decision to Stop[20], and Search Findings analyses (100 and 30 observations, respectively), and on model convergence[21] for the Stop Outcome analysis.

Table B.2. Sample Size Thresholds for Conducting Statistical Analyses

Statistical Test	Sample Size Threshold
Decision to Stop	Minimum of 100 observations for an individual racial/ethnic group
Stop Outcomes	Model convergence
Search Findings	Minimum 30 observations per racial/ethnic group analyzed; no cell with less than 5 observations

The sample size issue identified above had a significant impact on the STOP Program research team's ability to conduct analyses on each of the racial/ethnic groups found in the stop database. In several cases, even with two years of data, the total number of stopped individuals for certain racial/ethnic groups falls under the thresholds defined Table B.2. Further, once the STOP Program research team began to analyze subsets of the data (e.g., only those individuals who were searched, or arrested; those observations that

[20] Wilson, Voorhis, and Morgan (2007: 48) recommend that for regression equations where six or more variables are included in the model, "an absolute minimum of 10 participants per predictor variable is appropriate." While this is the minimum, if possible, they recommend 30 participants per predictor. Further, in instances where the outcome variable is skewed due to the small sizes of minority groups relative to the white group, larger sample sizes are needed. In this report, the STOP research team elected to use the 10-participant minimum, which when multiplied by 10 predictor variables sets the minimum number of observations for an individual racial/ethnic group at 100.
[21] All possible racial group and stop outcome models are estimated in Stata (a statistical software for data analysis). Models that did not converge are not included in the results.

met the standards to be included in the Decision to Stop), many of these counts fell under the requisite thresholds. To combat sample size issues, this report includes two years of data in all analyses.

A final concern is the prevalence of missing data. Resource limitations at some law enforcement agencies with a small number of staff are a challenge for STOP data submission and increases the potential for missing data. These resource and staffing limitations are likely exacerbated by the impacts of the COVID-19 pandemic, with tier 3 agencies beginning data collection in July 2020 shortly after the pandemic started. Missing data in the context of the STOP Program could come from two sources. First, a data point could be missing because it was never entered. Second, a data point could be submitted in an invalid format which lacks the information necessary to determine where it fits into the STOP Program data schema. Missing data attributable to both sources were found.

Threshold for Statistical Significance

To determine if statistically significant differences identified in this report warrant additional in-depth analysis and/or technical assistance from the DPSST, STOP Program researchers reviewed the results of each of the three analyses conducted on the STOP Program data. For each individual analysis, the estimated difference is considered statistically significant if it meets the 95% confidence level, meaning there is less than a 5% chance differences identified by the analyses were not due to random variation.

Table B.3. Bonferroni Adjustment by Analysis

Analysis	Number of Comparisons per Agency
Decision to Stop	Up to 5 comparisons
Stop Outcomes	Up to 20 comparisons
Search Findings	Up to 5 comparisons

When possible, multiple comparisons were made for each agency test. In situations where multiple tests are employed, all of which may indicate statistical significance, best practices require Bonferroni adjustments[22] to adjust for the likelihood of a given test yielding a false positive result. The Bonferroni adjustment differed for each agency test, contingent on the number of comparisons made. The number of comparisons is detailed in

[22] The Bonferroni Adjustment is a widely used statistical method that protects against the multiple comparison problem. For statistical tests that make multiple comparisons (for example, a single agency is tested for multiple race groups), the likelihood of finding a statistically significant result is higher. The Bonferroni Adjustment controls for that higher likelihood by raising the threshold for statistical significance for any one of the multiple comparisons, dependent upon the actual number of comparisons. See an example of how the Adjustment is used for the Search Findings Analysis in Appendix F.

Table B.3. Some agencies had too few stops of Asian/PI, Black, Hispanic, Middle Eastern, or Native American individuals to run tests for each group. Therefore, the magnitude of the Bonferroni adjustment may differ by agency, based on the number of tests run for that agency.

Beyond the 95% confidence threshold for each individual analysis, STOP Program researchers also established a threshold at which identified differences warrant further investigation and technical assistance from DPSST at the project level. Following best practices and the "gold standard" analyses conducted by the State of Connecticut[23], for a law enforcement agency to be identified as one requiring further analysis as well as DPSST technical assistance, it must be identified as having a statistically significant difference in at least two of the three analytical tests performed on the STOP data[24]. The justification for this approach mirrors the reasoning behind the utilization of multiple tests to examine the data acquired for this project. As discussed previously, given that the statistical output provided in this report in many instances are estimates which could lead to false positives or false negatives in any single analysis, best practices suggest that caution should be taken when examining and interpreting results from the statistical tests we performed.

[23] The Connecticut Racial Profiling Prohibition Project is located at http://www.ctrp3.org/.

[24] The State of Connecticut applies a sliding scale in its analyses, whereby a statistically significant difference identified via the Veil of Darkness analysis alone results in an agency being identified for further analysis. For its other analyses, two or more statistically significant differences result in further analysis. Unlike Connecticut, the Oregon STOP Program treats all three of its analyses as coequal while retaining the two-or-more-out-of-three threshold.

Appendix C Stop Characteristics for Tier 3 Agencies

Table C.1. Race/Ethnicity Reporting for Tier 3 Agencies

Agency	Asian or PI	Black	Hispanic	Middle Eastern	Native	White
Astoria PD	45	42	244	2	0	2,230
Aumsville PD	15	21	138	4	0	640
Baker CO SO	11	21	55	3	0	954
Baker City PD	6	8	22	0	1	412
Bandon PD	16	8	21	2	1	396
Black Butte Ranch PD	11	3	29	3	0	282
Boardman PD	3	14	820	4	13	560
Brookings PD	57	37	214	30	10	1,709
Burns PD	11	8	29	6	0	289
Cannon Beach PD	37	19	137	28	1	971
Carlton PD	6	5	49	6	0	337
Clatsop CO SO	70	61	286	15	0	2,411
Coburg PD	31	38	136	16	0	941
Columbia CO SO	42	76	175	30	4	2,764
Columbia City PD	6	3	21	3	0	183
Coos Bay PD	44	34	197	10	10	2,745
Coos CO SO	10	14	72	9	3	1,635
Coquille PD	2	1	9	0	0	311
Cottage Grove PD	6	7	74	6	0	682
Crook CO SO	24	17	222	5	0	1,967
Curry CO SO	1	1	0	0	0	48
Dallas PD	47	45	291	8	0	1,806
Eagle Point PD	26	25	172	5	0	1,284
Enterprise PD	0	0	1	0	0	6
Florence PD	12	8	31	2	0	680
Gearhart PD	16	9	52	9	0	300
Gilliam CO SO	22	35	157	6	1	1,011
Gladstone PD	118	197	502	80	11	2,915
Gold Beach PD	17	2	8	7	0	207
Harney CO SO	1	7	9	0	2	98
Hood River PD	45	35	300	5	3	688
Hubbard PD	24	28	473	5	0	543
Independence PD	36	66	514	7	2	1,105
Jacksonville PD	5	5	41	1	0	242
Jefferson CO SO	54	7	154	6	1	942

Continue on next page...

...Continued from previous page

Agency	Asian or PI	Black	Hispanic	Middle Eastern	Native	White
Josephine CO SO	13	15	97	4	1	970
Junction City PD	4	8	26	2	0	334
King City PD	86	108	395	78	2	925
La Grande PD	40	20	44	3	0	814
Madras PD	67	12	183	15	0	384
Malheur CO SO	14	21	121	23	0	445
Malin PD	2	3	28	2	0	78
Manzanita PD	13	7	23	3	0	238
Milton-Freewater PD	20	12	381	4	4	653
Molalla PD	28	31	356	9	4	1,542
Monmouth PD	41	40	245	13	0	784
Morrow CO SO	18	24	656	4	8	1,220
Mt. Angel PD	14	13	124	4	0	187
Myrtle Creek PD	16	12	49	6	0	1,452
Myrtle Point PD	2	0	6	1	1	153
Newport PD	15	10	104	5	0	544
North Bend PD	106	59	313	32	4	3,932
OSU PD	118	46	97	41	10	785
Oakridge PD	26	9	22	16	0	188
Pendleton PD	28	29	126	3	77	1,159
Philomath PD	86	60	151	21	7	1,289
Phoenix PD	31	41	214	6	0	984
Pilot Rock PD	6	4	19	3	0	293
Port Orford PD	14	4	9	8	0	105
Prineville PD	6	8	56	4	1	573
Rainier PD	45	26	130	6	1	1,126
Reedsport PD	0	0	4	0	0	13
Rogue River PD	6	3	48	1	0	290
Sandy PD	82	98	358	23	31	2,649
Seaside PD	76	60	351	19	3	1,978
Sherman CO SO	41	25	166	40	0	522
Sherwood PD	220	191	859	100	18	4,208
Silverton PD	38	51	496	4	3	2,218
Stanfield PD	13	40	430	10	26	872
Stayton PD	13	12	135	1	0	741

Continue on next page...

...Continued from previous page

Agency	Asian or PI	Black	Hispanic	Middle Eastern	Native	White
Sunriver PD	32	16	136	15	0	1,382
Sutherlin PD	28	24	107	42	0	908
Sweet Home PD	4	5	18	2	1	349
Talent PD	42	62	231	10	0	1,292
The Dalles PD	31	25	273	2	9	876
Tillamook CO SO	37	19	163	14	1	1,127
Tillamook PD	35	15	167	10	5	822
Toledo PD	4	5	37	2	8	304
Turner PD	1	3	41	2	0	184
Umatilla CO SO	6	27	452	9	20	828
Umatilla PD	21	47	1,215	4	30	1,221
Union CO SO	15	14	46	14	1	282
Union Pacific Railroad PD	17	34	76	1	4	293
Vernonia PD	6	1	13	3	0	185
Wallowa CO SO	2	0	11	3	0	144
Warrenton PD	28	23	159	0	2	1,314
Wasco CO SO	10	8	76	1	18	432
Wheeler CO SO	2	0	11	5	0	121
Winston PD	13	10	28	1	1	555
Yamhill PD	36	20	206	10	0	792
Total Tier 3	**2,595**	**2,367**	**15,943**	**987**	**364**	**82,353**

Table C.2. Tier 3 Agency Stops by Stop Type

	Traffic		Pedestrian		
Agency	Count	Pct	Count	Pct	Total
Astoria PD	2,562	100.0%	1	0.0%	2,563
Aumsville PD	818	100.0%	0	0.0%	818
Baker CO SO	1,043	99.9%	1	0.1%	1,044
Baker City PD	449	100.0%	0	0.0%	449
Bandon PD	444	100.0%	0	0.0%	444
Black Butte Ranch PD	328	100.0%	0	0.0%	328
Boardman PD	1,403	99.2%	11	0.8%	1,414
Brookings PD	2,057	100.0%	0	0.0%	2,057
Burns PD	343	100.0%	0	0.0%	343
Cannon Beach PD	1,190	99.7%	3	0.3%	1,193
Carlton PD	401	99.5%	2	0.5%	403
Clatsop CO SO	2,842	100.0%	1	0.0%	2,843
Coburg PD	1,161	99.9%	1	0.1%	1,162
Columbia CO SO	3,079	99.6%	11	0.4%	3,090
Columbia City PD	216	100.0%	0	0.0%	216
Coos Bay PD	3,040	100.0%	0	0.0%	3,040
Coos CO SO	1,743	100.0%	0	0.0%	1,743
Coquille PD	323	100.0%	0	0.0%	323
Cottage Grove PD	774	99.9%	1	0.1%	775
Crook CO SO	2,230	99.8%	5	0.2%	2,235
Curry CO SO	49	98.0%	1	2.0%	50
Dallas PD	2,193	99.8%	4	0.2%	2,197
Eagle Point PD	1,496	98.9%	16	1.1%	1,512
Enterprise PD	7	100.0%	0	0.0%	7
Florence PD	733	100.0%	0	0.0%	733
Gearhart PD	386	100.0%	0	0.0%	386
Gilliam CO SO	1,517	99.7%	4	0.3%	1,521
Gladstone PD	3,757	98.3%	66	1.7%	3,823
Gold Beach PD	241	100.0%	0	0.0%	241
Harney CO SO	117	100.0%	0	0.0%	117
Hood River PD	1,070	99.4%	6	0.6%	1,076
Hubbard PD	1,069	99.6%	4	0.4%	1,073
Independence PD	1,713	99.0%	17	1.0%	1,730
Jacksonville PD	294	100.0%	0	0.0%	294
Jefferson CO SO	1,162	99.8%	2	0.2%	1,164

Continue on next page...

...Continued from previous page

	Traffic		Pedestrian		
Agency	**Count**	**Pct**	**Count**	**Pct**	**Total**
Josephine CO SO	1,099	99.9%	1	0.1%	1,100
Junction City PD	374	100.0%	0	0.0%	374
King City PD	1,542	96.7%	52	3.3%	1,594
La Grande PD	921	100.0%	0	0.0%	921
Madras PD	659	99.7%	2	0.3%	661
Malheur CO SO	626	100.0%	0	0.0%	626
Malin PD	113	100.0%	0	0.0%	113
Manzanita PD	284	100.0%	0	0.0%	284
Milton-Freewater PD	1,040	99.4%	6	0.6%	1,046
Molalla PD	1,931	98.0%	39	2.0%	1,970
Monmouth PD	1,120	99.7%	3	0.3%	1,123
Morrow CO SO	2,060	99.9%	3	0.1%	2,063
Mt. Angel PD	337	98.5%	5	1.5%	342
Myrtle Creek PD	1,532	99.8%	3	0.2%	1,535
Myrtle Point PD	163	100.0%	0	0.0%	163
Newport PD	802	99.4%	5	0.6%	807
North Bend PD	4,443	99.9%	3	0.1%	4,446
OSU PD	1,090	99.4%	7	0.6%	1,097
Oakridge PD	260	99.6%	1	0.4%	261
Pendleton PD	1,338	94.1%	84	5.9%	1,422
Philomath PD	1,608	99.6%	6	0.4%	1,614
Phoenix PD	1,276	100.0%	0	0.0%	1,276
Pilot Rock PD	325	100.0%	0	0.0%	325
Port Orford PD	140	100.0%	0	0.0%	140
Prineville PD	635	98.0%	13	2.0%	648
Rainier PD	1,333	99.9%	1	0.1%	1,334
Reedsport PD	17	100.0%	0	0.0%	17
Rogue River PD	347	99.7%	1	0.3%	348
Sandy PD	3,223	99.4%	18	0.6%	3,241
Seaside PD	2,486	100.0%	1	0.0%	2,487
Sherman CO SO	794	100.0%	0	0.0%	794
Sherwood PD	5,565	99.4%	31	0.6%	5,596
Silverton PD	2,703	97.6%	66	2.4%	2,769
Stanfield PD	1,404	99.7%	4	0.3%	1,408
Stayton PD	902	100.0%	0	0.0%	902

Continue on next page...

...Continued from previous page

	Traffic		Pedestrian		
Agency	**Count**	**Pct**	**Count**	**Pct**	**Total**
Sunriver PD	1,581	100.0%	0	0.0%	1,581
Sutherlin PD	1,109	100.0%	0	0.0%	1,109
Sweet Home PD	379	100.0%	0	0.0%	379
Talent PD	1,603	97.9%	34	2.1%	1,637
The Dalles PD	1,207	99.3%	9	0.7%	1,216
Tillamook CO SO	1,361	100.0%	0	0.0%	1,361
Tillamook PD	998	99.5%	5	0.5%	1,003
Toledo PD	360	100.0%	0	0.0%	360
Turner PD	231	100.0%	0	0.0%	231
Umatilla CO SO	1,331	99.2%	11	0.8%	1,342
Umatilla PD	2,537	100.0%	1	0.0%	2,538
Union CO SO	372	100.0%	0	0.0%	372
Vernonia PD	194	93.3%	14	6.7%	208
Wallowa CO SO	160	100.0%	0	0.0%	160
Warrenton PD	1,526	100.0%	0	0.0%	1,526
Wasco CO SO	542	99.4%	3	0.6%	545
Wheeler CO SO	139	100.0%	0	0.0%	139
Winston PD	608	100.0%	0	0.0%	608
Yamhill PD	1,061	99.7%	3	0.3%	1,064
Tier 3 Total	**104,041**	**99.4%**	**592**	**0.6%**	**104,633**

Appendix D Decision to Stop Technical Appendix

The Decision to Stop (DTS) analysis, first developed by Grogger and Ridgeway (2006) as the Veil of Darkness analysis, analyzes stop data for differences in stop decisions by racial/ethnic characteristics of the driver. DTS is based on the basic assumption that officers can better detect a driver's race during daylight hours as compared to darkness. Specifically, relying on variations in daylight throughout the year, the DTS test compares the racial composition of stops in daylight to those in darkness during a combined inter-twilight window, which occurs during morning and evening commute times. The primary advantage of the test is that it does not rely on a benchmark comparison of either the estimated driving population or the residential population. Further, it is a widely accepted technique that does not suffer from benchmarking issues, and when deployed via a multivariate analysis, provides a strong test of racial disparities (Fazzalaro and Barone 2014).

The DTS analysis relies on two primary assumptions. The first is that in darkness, it is more difficult for officers to determine the race/ethnicity of an individual they intend to stop. Second, the analysis also assumes that driving population is consistent throughout the year, between daylight and darkness, and between the morning and evening commutes. If these assumptions hold, it is possible to model the differences in stops between light and dark using a logistic regression that takes the following form:

$$ln\left(\frac{P(m|\delta)}{1-P(m|\delta)}\right) = \alpha + \delta + \gamma + \omega + \varepsilon$$

where m represents the treatment of a minority group relative to the white majority group, δ is a binary indicator representing daylight, γ is a vector of coefficients, including controls for time of day, day of the week, season, agency stop volume, and county stop volume, and ω is a vector of coefficients representing the demographic characteristics of the stopped individual as well as the reason for the stop.[25] Importantly, the inclusion of controls for time of day, day of the week, and season ensure that the model meets the second assumption regarding the consistency of the driving population throughout the year.

A key factor in the specification of the DTS model is identifying the appropriate periods of daylight and darkness for the analysis. Following Grogger and Ridgeway (2006), the STOP Program analyzes stops that occur within the combined inter-twilight window. The

[25] The covariates included in the models were age, gender, reason for the stop, day of week, time of day, quarter or season, stop year, county stop volume, and agency stop volume. Time of day is modeled as a control variable for morning and evening stops, as well as a restricted cubic spline with three degrees of freedom within each twilight window. Alternative time of day controls were tested and did not change the results.

combined inter-twilight window is created from the Oregon traffic stop data from July 1, 2023, to June 30, 2025. Every traffic stop is defined to have occurred in daylight or darkness based on the date, time, and location of the stop. Astronomical data from the United States Naval Observatory (USNO) is used to determine the sunrise, sunset, and start and end of civil twilight. If the location of the stop has been geo-coded, then those coordinates are used to determine the sunrise, sunset, and civil twilight window for that exact location. If the stop has not been geo-coded due to limitations with location data, the centroid of the city is used. If the city information is unavailable, then the centroid of the county is used.

The dawn inter-twilight period is defined as the earliest start of civil twilight to the latest sunrise. The earliest start of civil twilight is 4:21am in Wallowa County, and the latest sunrise is 7:59am in Clatsop County. Stops that occur in the daily morning twilight window (approximately 30 minutes between the start of civil twilight and the sunrise) are removed since it is neither light nor dark during this time period. Conversely, the dusk twilight window is defined as the earliest sunset to the latest end of civil twilight. The earliest sunset is 4:05pm in Wallowa County, and the latest end of civil twilight is 9:48pm in Clatsop County. Stops that occur in the daily evening twilight window (approximately 30 minutes between sunset and the end of civil twilight) are similarly removed since it is neither light nor dark during this time period. Adjustments have been made to account for daylight savings time (DST) in November and March. In addition, while most of Oregon is on Pacific Standard Time (PST), most of Malheur County is on Mountain Standard Time (MST). The stops in Malheur County have been adjusted to account for this time zone.

The log odds that result from the DTS logistic regression model were then converted to odds ratios. Thus, the model tests whether the odds of non-white traffic stops during daylight are significantly different from the odds of non-white traffic stops during darkness. The DTS approach tests whether the odds ratio is statistically significantly different from 1.0. If the odds ratio is not statistically different from 1.0, then the test finds no difference in stops made during daylight and darkness. If the odds ratio is greater than 1.0 and statistically significant, however, the test concludes the odds of non-white drivers being stopped in daylight is significantly higher than in darkness, which is taken as evidence of a racial disparity in stops, after accounting for additional control variables that are available in the stop data. Conversely, if the odds ratio is less than 1.0 and statistically significant, the odds of a non-white driver being stopped in daylight is significantly lower than in darkness. The logistic regression modeling was compiled using Stata software and utilizing the logistic regression function.

Appendix E Stop Outcomes Technical Appendix

Propensity score methods are a family of statistical methods for drawing causal inference about treatment effects in situations where randomized control trials are not feasible. Randomized control trials ensure that treatment assignment is independent of all covariates. Without this randomization, confounders may bias the estimated treatment effects. Confounding variables are a major hurdle to estimating effects in real-world settings and balancing based on the propensity to receive treatment (i.e., propensity score) is one way to mitigate this bias in non-experimental settings. In general, propensity score techniques aim to balance the characteristics (or confounding variables) of the treatment and control groups. This allows an unbiased comparison between those two groups for the outcome variable of interest, as there are no observed differences between the two groups. These methods are frequently employed in the analysis of disparities in criminal justice settings (Higgins et al. 2011; 2013; Ridgeway 2006; Stringer and Holland 2016; Vito, Grossi, and Higgins 2017).

Propensity score methods measure the characteristics of the "treatment" and "control" groups and then weight one or both groups based on measured characteristics so that the two groups look as similar as possible. The resulting groups are said to be "balanced" if they are statistically similar across measured confounding variables following the balancing procedure. If all confounding variables are measured and balanced, then the difference in the average outcomes between the treatment and control groups is an unbiased measure of the average treatment effect. Similarly, if unmeasured confounding variables are closely correlated with the balanced confounding variables and thus are also likely to be balanced, then the average treatment effect is balanced. Some methods, as employed in the current analysis, go a step further and incorporate regression analysis as an additional controlling method after the balancing process.

There are several different forms of propensity score estimators. Here, the researchers employ Inverse Probability Weighted Regression Adjustment (IPWRA) using the Stata statistical package, version 16.1. The method has the following steps:

1. The treatment equation is estimated including potentially confounding variables. The dependent variable is a binary treatment variable and a probit model is estimated.
2. The predicted treatment values from the estimates in step 1 are stored.

3. Inverse probability weights (IPW) are created for each observation using these values.[26]
    a. For treated observations, IPW $= 1$
    b. For control observations, IPW $= \frac{(propensity\ score)}{1-(propensity\ score)}$
4. The outcome equation is estimated using the weights created in step 3 in a regression analysis, including all covariates that are theoretically relevant predictors of the outcome variable.

One advantage of the IPWRA estimator relative to other propensity score estimators is that it benefits from the Double Robust property by estimating the regression equation after the balancing procedure: If *either* the treatment equation *or* the outcome equation is correctly specified then the estimator is unbiased. Put alternatively, the estimates from IPWRA estimation are robust to misspecification errors in either the treatment or outcome equation. Two-stage propensity score estimators such as IPWRA balance for important covariates at both the treatment selection and outcome stages of estimation.[27]

Assumptions

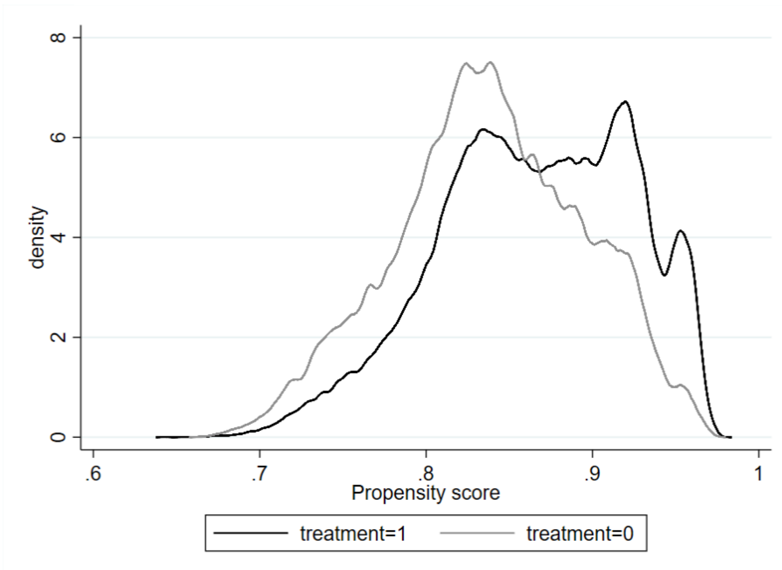
There are a few assumptions that must hold for propensity score estimators to be unbiased. The first is the conditional independence assumption[28], which states that the outcome variable is conditionally independent of the treatment. This means that if researchers include all relevant confounding variables in estimating the treatment equation, i.e., the treatment equation is properly specified, and these variables are balanced across the two groups following match selection, then the outcomes are conditionally independent of the treatment. For this assumption to hold, changes in any unobserved variables that have an effect on the outcome variable must not also influence the treatment variable. This assumption is a theoretical consideration that is not possible to directly test, as a variable may be correlated with both treatment and outcome but may be a spurious correlation. The analyst may, however, ensure that all the measured confounding variables are equally represented in both the treatment and control groups and thus that the confounding variables are not the drivers of remaining variance in treatments and outcomes.

The second main assumption is the overlap assumption, whereby the range of estimated propensity scores for the treated group must overlap with those of control group observations. If an observation is not within this range, then it is omitted from the sample as it is impossible to form a valid match from the comparison group. This idea is best

[26] These differ whether the estimate is the Average Treatment Effect (ATE) or the Average Treatment Effect on the Treated (ATET). Here we are estimating the ATET. *See* Austin and Stuart 2015.
[27] For a thorough discussion of IPWRA methods, *see* Wooldridge 2010, Chapter 21.3.4.
[28] This assumption is also referred to as the unconfoundedness assumption.

represented with a pre-balance propensity score distribution graph, as seen in the examples below. Figure E.1 shows that for most values of the propensity score (horizontal axis) there is an observation for both the treated (treatment=1) and untreated (=0) groups, but also that at the upper and lower ends there are treated observations that do not have a comparable observation in the untreated group. To satisfy this assumption, for this example these observations with extreme propensity scores would be dropped.

Figure E.1. Overlap Example



With a limited range of covariates, including mostly categorical variables, and the large sample sizes with this set of tier 1 agencies, each analysis completed here had no omitted observations because of a violation of the overlap assumption.[29]

Finally, the Stable Unit Treatment Value Assumption (SUTVA) is similar in concept to the independent and identically distributed (i.i.d.) assumption, but specific to the treatment assignment setting. SUTVA requires that any given unit's treatment assignment does not have a causal relationship with another observation's treatment assignment. This assumption would be violated in this case if, for example, the stop of a Hispanic individual causes another Hispanic individual to be stopped. There may be clustering of stops by race/ethnicity group based on policing strategies, but this assumption is not likely to be

[29] Omitted treatment variables per analysis are not included in this report due to the high number of analyses conducted.

violated in this case as the race of a stopped individual does not plausibly impact the race of subsequently stopped individuals.[30]

Estimation

If the above assumptions hold then estimation may proceed. The teffects ipwra command is used in Stata to estimate these models. First the "treatment" equation is estimated. The treatment variables in this case are indicator variables for each of:

1. Officer perception of race/ethnicity: = 1 if Asian/PI, = 0 if white
2. Officer perception of race/ethnicity: = 1 if Black, = 0 if white
3. Officer perception of race/ethnicity: = 1 if Hispanic, = 0 if white
4. Officer perception of race/ethnicity: = 1 if Middle Eastern, = 0 if white
5. Officer perception of race/ethnicity: = 1 if Native American, = 0 if white

The standard language of treatment/control used with the IPWRA methodology is ill-suited to this STOP analysis. The current analysis balances the two groups under each sub-analysis across all observed covariates, rather than giving one group a treatment, but not the other. This method makes it so that the only perceptible difference between the two groups is the race/ethnicity of those two groups, but race/ethnicity does not conform to this "treatment" description. This language is preserved simply to remain consistent with the relevant literature.

The following confounding variables are balanced across the groups:

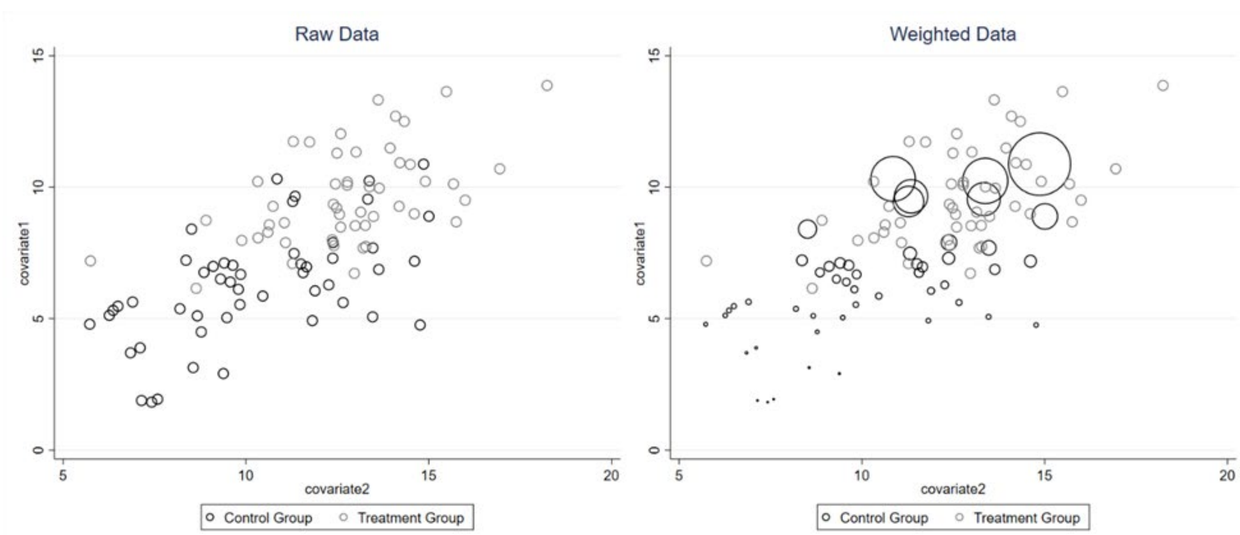
1. Female indicator, 1 = if female, 0 = if any other
2. Age category indicators for each of <21, 21-24, 25-29, 30-39, 40-49, 50+
3. Season indicators for each of Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec
4. Daylight indicator = 1 if stop happened after sunrise and before sunset, = 0 otherwise
5. Time of stop indicators for each of 12am-5am, 5am-10am, 10am-3pm, 3pm-8pm, 8pm-12am
6. Citation category indicators for each of Equipment Violation; Low Speed or Moving Violation; Moving Violation – High; Moving Violation – Medium; Registration/License; Speed Violation – High; Speed Violation – Medium; and Unknown/Other.
7. Day of week indicators
8. Agency stop volume $= \frac{Total\ \#\ of\ stops\ by\ agency\ on\ day\ of\ stop}{Maximum\ \#\ of\ daily\ stops\ by\ \boldsymbol{agency}\ over\ year\ of\ analysis}$
9. County stop volume $= \frac{Total\ \#\ of\ stops\ by\ agency\ on\ day\ of\ stop}{Maximum\ \#\ of\ daily\ stops\ in\ the\ \boldsymbol{county}\ over\ year\ of\ analysis}$
10. If the stop outcome is caused by a low-discretion violation = 1, otherwise = 0

[30] The Stata handbook provides a good description of these assumptions, and the counterfactual model that underlies all matching methods. *See* "Stata Treatment-Effects Reference Manual: Potential Outcomes/Counterfactual Outcomes" 2019.

The first step of the analysis uses a probit model to estimate the propensity of being in the treatment group based on the covariates listed above. Overlap of propensity scores is evaluated and any non-overlapping observations are removed from the sample. Inverse Probability Weights (IPWs) are estimated for each observation based on the propensity scores. For the treatment group in an ATET framework, these weights are equal to 1. For the control group the weight is equal to $p/(1-p)$, where p is the propensity score (see footnote 31). In effect, this process gives more weight to control observations that have a higher propensity score (i.e., are more similar to treated observations).

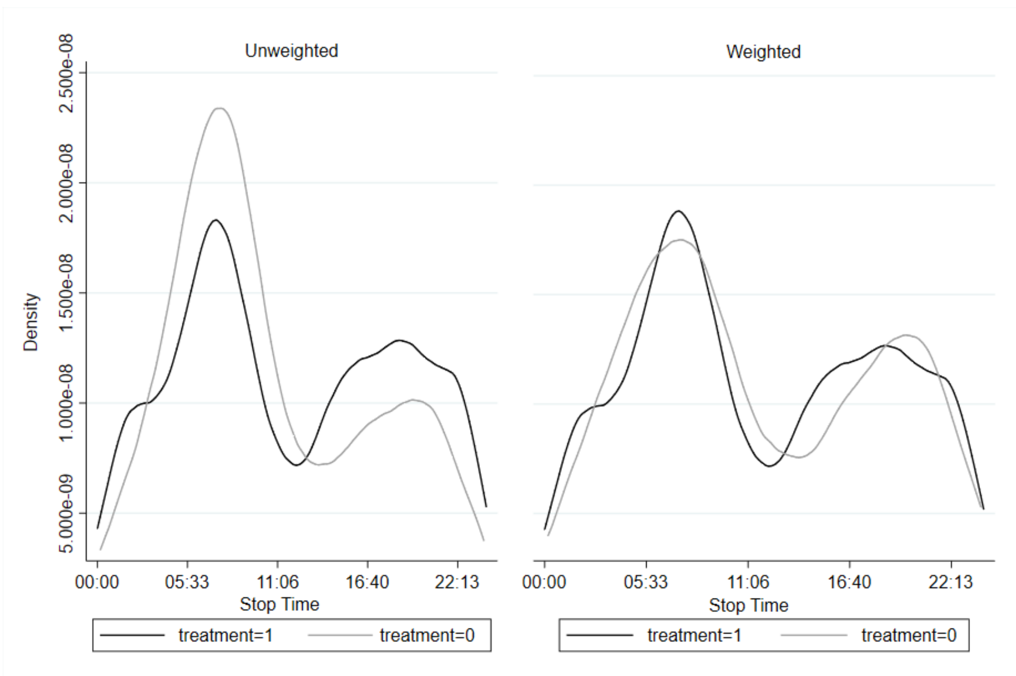
A hypothetical example application of IPWs Figure E.2. The two graphs each represent control and treatment group observations and their respective values for each of two covariates. While there is some overlap between the groups in this example, the treatment (light gray) group tends to have higher values of both variables. In the Raw Data (unweighted) we can see that the two groups are not directly comparable. After calculating IPWs for ATET these weights are applied to the two groups and represented by the size of the circles in the Weighted Data graph. The treatment group remains the same here since the weights = 1, but the importance or weight of control group observations are adjusted. The observations that are closer to the treatment group observations are given a large weight, while those that are not are given a small weight. The weighted control group has observations that are much closer to those of the treatment group than the raw control group.

Figure E.2. Weighting Example



Balance is then measured based on the standardized difference[31] in means and the variance ratio[32] between the treatment and control groups for each of the raw data set and the inverse probability weighted data set. If the resulting standardized difference in the weighted data set is close to zero and the variance ratio is close to 1 for each variable for the weighted data then the sample is said to be balanced. Balance was evaluated in every data subset by agency and strong balance was achieved in every instance, e.g., the standardized differences were always close to zero (usually within .01 of 0, always within 0.05) and the variance ratios were always close to one (usually within .01 of 1, always within 0.05) (Austin 2009a; 2009b). In every case, the data sets were relatively well balanced in the initial, raw data sets, but became more balanced through the weighting process. This balance can also be evaluated graphically for each variable. Figure E.3 is an example of one of these variables for one agency. The Unweighted chart displays the distribution of stop time for each of the treated group and the untreated group. The Weighted chart displays these same distributions with the IPWs applied. The distributions of the two groups more closely resemble each other in the weighted graph than in the unweighted graph, so STOP Program researchers can say that these groups are more balanced when incorporating the IPWs.

Figure E.3. Confounding Variable Balance Example



[31] The standardized difference of variable x is: $\delta_x = \frac{\mu_x(t=1)-\mu_x(t=0)}{\sqrt{\frac{\sigma_x^2(t=1)-\sigma_x^2(t=0)}{2}}}$

[32] The variance ratio is simply the variance of the treated group divided by the variance of the control group.

Outcome equations are then estimated for each of the treatment variables across four sets of outcomes:

1. = 0 if a warning/none disposition is observed, = 1 otherwise
2. = 1 if a citation disposition is observed, = 0 if warning/none outcome is observed
3. = 1 if a search disposition is observed, = 0 if a citation or warning/none outcome is observed
4. = 1 if an arrest disposition is observed, = 0 otherwise

In the next step, probit models with the inverse probability weights applied and robust standard errors are estimated for each of the treatment and control groups. Predicted outcomes are stored for each observation and their average yields the potential outcome mean for the control group. The comparison between this mean and the actual average of the treatment group yields the Average Treatment Effect on the Treated (ATET), the main estimate of interest in these models. This estimate is slightly different from the Average Treatment Effect as it focuses specifically on the effect on the treated group rather than the population as a whole. In this case, the estimates may be interpreted as the average difference in predicted probability of the outcome if the treated (minority) group had identical characteristics to the control group, except had a race/ethnicity = white.[33]

Limitations

As with any statistical analysis, there are potential shortcomings of IPWRA analysis that may hinder the validity of the results. In this case, the largest concerns are the data limitations that result in the omission of some confounding variables that may be theoretically relevant. Comparable analyses of bias in police stops in other localities have controlled for additional confounding variables not included here, including police officer identifiers, make/model/year of vehicle, and location of the stop. Other variables may influence officer decision criteria but are rarely included in the comparable analyses in other states due to data availability challenges. These variables include economic characteristics of the driver (i.e., employment status, income, etc.) and information on the driving population from which drivers are stopped. This later variable poses significant estimation challenges as it requires several assumptions regarding directions, populations, and time of travel, as well as frequencies of commuters and tourists at each location in the road system. Without significant preliminary data about these factors any estimation of the driving population is likely to incorporate a significant amount of bias to any statistical estimates built on top of these driving population estimates.

[33] Conversely, the ATE predicts these differences for both the treated group and for the untreated group and averages all these differences. Thus, it estimates the difference in predicted probabilities for both the white group and the minority groups and averages across all observations.

Many of these variables are not described in the statutes establishing Oregon's STOP data tracking system (e.g., make/model). Other variables, such as geographic location of the stop, are highly varied in quality and format across these Oregon agencies. Some Oregon agencies provide precise longitude and latitude of the traffic stop via automatic logging in the cellphone app, other agencies allow officers to enter nearest intersections or mile markers, and others require no location to be entered by their officers. Due to this lack of uniformity in reporting, the STOP research team could not include location information for some agencies with high quality location information while also conducting uniform analyses agencies.

The omission of important confounding variables leads to the low Pseudo-R^2s in the results and also drives the high amount of balance found in the raw data. In each sub-analysis the balancing procedure leads to greater confounder balance than in the raw data, but the groups were not egregiously unbalanced in the raw data. A high number of the confounders are binary indicator variables, which makes it easier to form very close matches and leads to less imbalance in the raw data, but this also shows that these variables may be imprecisely measured.

Appendix F Search Findings Technical Appendix

Model and Assumptions

The Search Findings analyses performed in this report are based on the model presented by Knowles, Persico, and Todd (2001) which details how police and citizens act surrounding searches. In this model, police officers are assumed to make the decision to search someone based on their perception of the likelihood that the person will have contraband in their possession, while also accounting for the economic "cost" of a search. In the case that the cost of searching members of different groups is the same, officers may be expected to search the group that they perceive to be more likely to possess contraband. Similarly, this model assumes that citizens make the decision to carry contraband based on their perception of the likelihood that they will be caught with contraband. If a particular group is more likely to carry contraband, they will be searched more often by police. As a group, they will respond by reducing their likelihood of carrying contraband in order to reduce their risk of being caught. In this way, differences in groups' likelihoods to carry contraband and to be searched by police should tend toward an equilibrium. At equilibrium, STOP Program researchers expect that the hit-rate (the rate at which searches are "successful," or result in finding contraband) should be equal across groups, whereas unequal hit-rates indicate disparate search practices.

If a group is "over-searched" (searched more often than necessary to maintain the abovementioned equilibrium), then the hit-rate for that group will be lower than that of a baseline group. In the case of this report, the baseline group is white drivers and pedestrians, while the test group are drivers and pedestrians in any one of the racial/ethnic groups present in the data. perhaps indicating what Becker calls "a taste for discrimination" (an economic phrase coined to describe discrimination) in officers conducting searches.

Hit-Rate and Significance Calculation

The hit-rate for a group is simply a proportion. The total number of searches of a group is represented by s and the number of searches of that group which result in finding contraband is represented by f:

$$\text{KPT Hit-Rate } = \frac{f}{s}$$

After calculating hit-rates by agency for each racial/ethnic group, chi-square tests of independence were performed in order to determine whether differences in the hit-rates were statistically significant. Yates's continuity correction for the chi-square test was used to mitigate the test's tendency to produce low p-values due to the discrete nature of the data. However, no substantive difference arose between the results when performed with or without the continuity correction. A confidence level of 95 percent with a Bonferroni correction for multiple testing determined significance. Each agency's white hit-rate was compared to each race group (Black, Hispanic, Asian/PI, Middle Eastern, and Native American) dependent upon sample size, so a Bonferroni corrected p-value of 0.05/5 = 0.01, 0.05/4, 0.05/3, 0.05/2, or 0.05 was used, dependent upon the number of groups for which the analysis was able to be performed. Hit-rate analyses and accompanying statistical tests were performed with the statistical software R.

Limitations

One important assumption of the Search Findings analysis model is that all searches included in the analysis are discretionary. Some searches, such as those made incident to arrest, are non-discretionary, meaning that there is no individual choice (discretion) in the officer's decision to conduct the search. This type of search is not representative of officers' motivations and cannot be used to determine any patterns of behavior. In the STOP Program training that all officers complete prior to submitting data for this study, officers are informed that non-discretionary searches should not be included in the data. This means that when a stop results in an officer arresting someone, although they will always do a "pat-down" to ensure safety at the time of arrest, STOP Program researchers should not always see a search recorded for the stop (as these pat-downs are non-discretionary searches). In some cases, the data seem to show records of searches incident to arrest, however it is not possible to distinguish these "mistakes" from true records of discretionary searches. Accordingly, STOP Program researchers chose to take all data at face value–that is, if a search was recorded, it is included in the KPT Hit-Rate analysis as a discretionary search.

A possible methodological limitation of the hit-rate test is the problem of infra-marginality (Simoiu 2017). Infra-marginality is best explained by example. Suppose that group A has some portion of members that carry contraband 55 percent of the time (while all other members of the group carry contraband less than 50 percent of the time). Suppose also that group B has some portion of members that instead carry contraband 75 percent of the time (while all other members of the group carry contraband less than 50 percent of the time). If an officer only searches every person (regardless of group) who has over a 50 percent chance of carrying contraband, then group A will have a lower hit-rate. In the hit-rate test, this would appear to indicate discrimination against group A, despite the true "group-neutral" manner of the officer's search decisions. While this is one of the widest criticisms of the KPT Hit-Rate test, Persico (of Knowles, Persico, and Todd) independently addressed the criticism of this limitation in a follow up paper. Persico (2009) argues that infra-marginality is alleviated by the allowance in the model for searched groups to respond to search intensity (by lowering their propensity to carry contraband when searched more frequently). This is consistent with KPT's initial assertion that subgroups, as well as larger racial/ethnic groups, should act similarly to larger groups in that they adjust their propensity to carry contraband according to their likelihood of being searched.