STOP REPORT 2024

Per House Bill 2355 (2017)

December 1, 2024

Oregon Criminal Justice Commission

Ken Sanchagrin
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.

Table of Contents

Tables and Figures

Executive Summary

House Bill 2355 (2017) mandates all Oregon law enforcement agencies to submit data regarding officer-initiated traffic and pedestrian stops to the Oregon Criminal Justice Commission (CJC), so the CJC can analyze the submitted data for evidence of racial or ethnic disparities on an annual basis. The Oregon Statistical Transparency of Policing (STOP) Program, housed at the CJC, was created with assistance from the Oregon State Police and the Oregon Department of Public Safety Standards and Training (DPSST). This is the sixth annual report to the Oregon Legislature by the STOP Program examining data received pursuant to HB 2355.

Table 0.1. Descriptive Statistics for Aggregate Year 6 Stop Data

Variable	Tier 1	Tier 2	Tier 3
Traffic Stop	99.1%	99.2%	99.6%
Race/Ethnicity			
Asian/PI	3.6%	2.7%	2.3%
Black	5.3%	3.3%	2.1%
Latinx	17.2%	14.7%	14.5%
Middle Eastern	1.7%	1.1%	0.8%
Native American	0.6%	0.2%	0.3%
White	71.7%	78.0%	80.0%
Gender			
Male	67.3%	64.3%	63.2%
Female	32.5%	35.5%	34.6%
Nonbinary	0.3%	0.2%	2.2%
Age			
Under 21	10.4%	12.0%	12.3%
21-29	22.4%	20.8%	19.9%
30-39	24.7%	23.9%	22.2%
40-49	18.2%	19.0%	17.9%
50 and Older	24.3%	24.2%	27.7%
Stop Disposition			
None	1.6%	5.9%	3.7%
Warning	62.6%	63.8%	72.5%
Citation	34.3%	28.7%	22.8%
Juvenile Summons	0.0%	0.0%	0.0%
Arrest	1.5%	1.6%	0.9%
Search Conducted	1.5%	1.1%	0.5%

Table 0.1 reports descriptive statistics for the combined Tier 1, Tier 2, and Tier 3 data, which represents stops made from July 1, 2023 through June 30, 2024. The majority of stops in Oregon involved white individuals, which, in and of itself, is not surprising given the demographic makeup of Oregon as a whole. Overall, a little over one-quarter of Tier 1 stops and close to one-fifth of Tier 2 and Tier 3 stops involved Asian or Pacific Islander, Black, Latinx, Middle Eastern, or Native American individuals in Oregon. Once the stop had been initiated, stopped individuals either were subject to no further action or merely given a warning in 64 percent of stops for Tier 1, 70 percent of stops for Tier 2, and 76 percent of stops for Tier 3.

To examine the traffic and pedestrian stop data acquired by the STOP Program for racial/ethnic disparities, STOP Program researchers utilized three methods. The first method, which is used to examine the initial decision to stop an individual, was the Decision to Stop analysis. This analysis takes advantage of natural variations in daylight and darkness throughout the year, and is based on the assumption that it is easier for an officer to discern the race/ethnicity of an individual during the day when it is light versus the night when it is dark. Accordingly, the 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 would be evidence of a disparity.

The second analytical method employed by the STOP Program is the Stop Outcomes analysis, which 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). If, after matching on all available data points in the stop data (e.g., time of day and day of the week the stop was made, reason for the stop, gender, age), minority individuals are either cited, searched, or arrested more often than similarly situated white individuals, then there would be evidence of a disparity.

Finally, the STOP Program utilized the Search Findings analysis, which compares relative rates of successful searches (i.e., those resulting in the seizure of contraband) across racial/ethnic groups. It is based on the assumption that if search decisions by officers are made based on race/ethnicity neutral criteria, then 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 would be evidence of a disparity.

To determine if disparities 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, an estimated disparity 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 disparity 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, regardless of the number of analyses that are statistically significant.

Using the above-mentioned analyses and thresholds, the STOP Program identified three agencies that had statistically significant results across two of the tests performed on the data. Canby PD, Deschutes CO SO, and Oregon State Police show statistically significant disparities in two of the three analytical tests described in this report. These agencies 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[1] and engage with DPSST on any results that show a statistically significant disparity.

[1] 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 were to collect data concerning all officer-initiated traffic and pedestrian stops. The mandate also required that the Oregon Criminal Justice Commission (CJC) analyze the collected data to determine whether racial disparities exist in stops and stop outcomes. Since the passage of HB 2355, the STOP Program has developed a standardized method for data collection as well as data collection software offered free of charge to all state law enforcement agencies. As of September 2024, the STOP Program received data from 142 law enforcement agencies in the state and analyses using those data are presented in this report[2].

2. Characteristics of Year 6 Stop Data

2.1. General Characteristics

While the analyses contained in Sections 3., 4., and 5. utilize two years of submitted data, this section analyzes data collected by the STOP Program for officer-initiated traffic and pedestrian stops solely for the most recent year, which includes stops made between July 1, 2023, through June 30, 2024. In total, 582,500 stops were submitted to the STOP Program by 142 Tier 1, Tier 2, and Tier 3 agencies during Year 6. The number of stops reported by each agency is displayed in Table 2.1.1., Table 2.1.2., and Table C.2. in Appendix C. There was significant variation in the frequency with which Tier 1, Tier 2, and Tier 3 agencies stopped individuals. Tier 1 agencies generally made more stops than Tier 2 agencies, which in turn made more stops than Tier 3 agencies, which is consistent with size differences between Tiers in terms of officers employed. The Oregon State Police, which is the state's largest law enforcement agency, made 213,737 stops in Year 6, the largest number reported by any one agency and accounting for over one-third of all stops in the state. At the other end of the continuum, PSU CPS made the fewest stops, with only one stop reported for Year 6.

Table 2.1.1. Tier 1 Agency Stops by Stop Type

Agency Name	Traffic		Pedestrian		Total
Beaverton PD	15,024	96.7%	518	3.3%	15,542
Clackamas CO SO	16,633	97.5%	426	2.5%	17,059
Eugene PD	10,909	100.0%	0	0.0%	10,909
Gresham PD	3,735	99.5%	18	0.5%	3,753
Hillsboro PD	9,779	99.4%	61	0.6%	9,840
Marion CO SO	12,760	98.7%	165	1.3%	12,925
Medford PD	3,747	99.6%	17	0.5%	3,764
Multnomah CO SO	10,426	98.4%	172	1.6%	10,598
Oregon State Police	212,529	99.4%	1,208	0.6%	213,737
Portland PB	21,490	99.5%	110	0.5%	21,600
Salem PD	3,448	93.7%	231	6.3%	3,680
Washington CO SO	23,727	99.5%	113	0.5%	23,840
Tier 1 Total	344,207	**99.1%**	3,039	**0.9%**	347,247

[2] One agency, Union Pacific Railroad, submitted data after September 1, 2024; therefore, their data is excluded from the tables and analyses presented in this report.

Table 2.1.2. **Tier 2 Agency Stops by Stop Type**

Agency Name	Traffic		Pedestrian		Total
Albany PD	7,275	99.7%	25	0.3%	7,300
Ashland PD	2,146	99.6%	9	0.4%	2,155
Bend PD	3,394	99.7%	10	0.3%	3,404
Benton CO SO	6,692	99.9%	9	0.1%	6,701
Canby PD	4,325	99.1%	38	0.9%	4,363
Central Point PD	1,638	99.1%	15	0.9%	1,653
Corvallis PD	7,365	99.2%	60	0.8%	7,425
Deschutes CO SO	6,462	99.7%	20	0.3%	6,482
Douglas CO SO	1,744	100.0%	0	0.0%	1,744
Forest Grove PD	5,162	99.5%	24	0.5%	5,186
Grants Pass PD	1,856	95.8%	81	4.2%	1,937
Hermiston PD	3,854	99.1%	35	0.9%	3,889
Hood River CO SO	1,150	99.5%	6	0.5%	1,156
Jackson CO SO	7,360	99.4%	43	0.6%	7,403
Keizer PD	1,304	100.0%	0	0.0%	1,304
Klamath CO SO	530	100.0%	0	0.0%	530
Klamath Falls PD	1,947	100.0%	0	0.0%	1,947
Lake Oswego PD	5,144	99.7%	18	0.4%	5,162
Lane CO SO	6,555	99.5%	34	0.5%	6,589
Lebanon PD	1047	99.7%	3	0.3%	1,050
Lincoln CO SO	2,306	100.0%	0	0.0%	2,306
Lincoln City PD	1,853	99.4%	11	0.6%	1,864
Linn CO SO	6,736	99.9%	8	0.1%	6,744
McMinnville PD	2,303	99.5%	12	0.5%	2,315
Milwaukie PD	6,503	98.2%	118	1.8%	6,621
Newberg-Dundee PD	4,296	99.3%	31	0.7%	4,327
OHSU PD	59	96.7%	2	3.3%	61
Oregon City PD	6,033	97.3%	166	2.7%	6,199
Polk CO SO	3,366	99.8%	6	0.2%	3,372
Port of Portland PD	1,676	99.6%	7	0.4%	1,683
Redmond PD	5,504	99.9%	5	0.1%	5,509
Roseburg PD	2,782	94.3%	169	5.7%	2,951
Springfield PD	7,446	100.0%	2	0.0%	7,448
Tigard PD	5,484	97.9%	119	2.1%	5,603
Tualatin PD	3,572	99.5%	17	0.5%	3,589
UO PD	341	96.6%	12	3.4%	353
West Linn PD	2,385	99.9%	2	0.1%	2,387
Woodburn PD	1,313	100.0%	0	0.0%	1,313
Yamhill CO SO	5,294	99.9%	7	0.1%	5,301
Tier 2 Total	**146,202**	**99.2%**	**1,124**	**0.8%**	**147,326**

Tables 2.1.1. and 2.1.2. above and Table C.2. in Appendix C report the number and percentage of stops by agency broken down by stop type—traffic or pedestrian—and separated by Tier. Stop type has been

adjusted as described in Section B.2.3.1. By agency and within Tier, the frequency with which pedestrian stops were made, as well as the degree to which those stops affected a department's overall stop profile, varied significantly. Across all tiers, Tier 1 agencies had the highest proportion of pedestrian stops with 0.9 percent compared to Tier 2 and Tier 3's 0.8 percent and 0.4 percent, respectively. Of all Tier 1 agencies, Salem PD made the highest proportion of pedestrian stops, followed by Beaverton PD, which is in line with findings from last year's reporting. Of Tier 2 agencies, Roseburg PD had the highest proportion of pedestrian stops. Of Tier 3 agencies who submitted their stops prior to the reporting deadline, Silverton PD had the highest proportion of pedestrian stops.

The demographic breakdowns for traffic and pedestrian stops are reported in Table 2.1.3. For all agencies contained in this report, the majority of stops were of white drivers/pedestrians, with Latinx and Black individuals being the two most frequently stopped minority groups overall. This pattern held when broken down by traffic versus pedestrian stops, although white individuals made up a higher proportion of pedestrians across all Tiers. With regard to gender, more males were stopped than females. This gender difference is more pronounced in pedestrian stops. Most traffic and pedestrian stops are of individuals perceived to be aged in their thirties, slightly more so for pedestrians, across all Tiers. This echoes previous years' data.

Table 2.1.3. Demographics by Tier and Stop Type

	Tier 1			Tier 2			Tier 3		
	Traffic	**Ped.**	**Total**	**Traffic**	**Ped.**	**Total**	**Traffic**	**Ped.**	**Total**
Race/Ethnicity									
Asian/PI	3.6%	1.7%	3.6%	2.8%	1.3%	2.7%	2.4%	1.0%	2.3%
Black	5.2%	8.6%	5.3%	3.3%	4.4%	3.3%	2.1%	2.2%	2.1%
Latinx	17.2%	11.7%	17.2%	14.7%	10.4%	14.7%	14.5%	11.4%	14.5%
Mid. East.	1.7%	0.7%	1.7%	1.1%	0.5%	1.1%	0.8%	0.0%	0.8%
Native	0.6%	0.6%	0.6%	0.2%	0.3%	0.2%	0.3%	1.6%	0.3%
White	71.7%	76.7%	71.7%	77.9%	83.1%	78.0%	80.0%	83.9%	80.0%
Gender									
Female	32.6%	19.9%	32.5%	35.6%	23.1%	35.5%	34.7%	25.1%	34.7%
Male	67.2%	79.8%	67.3%	64.2%	76.6%	64.3%	63.2%	74.6%	63.2%
Nonbinary	0.3%	0.3%	0.3%	0.2%	0.3%	0.2%	2.2%	0.3%	2.2%
Age									
Under 21	10.4%	4.7%	10.4%	12.0%	8.8%	12.0%	12.3%	12.9%	12.3%
21-29	22.5%	16.4%	22.4%	20.9%	13.4%	20.8%	19.9%	17.9%	19.9%
30-39	24.6%	38.3%	24.7%	23.9%	27.9%	23.9%	22.2%	27.7%	22.2%
40-49	18.2%	22.1%	18.2%	19.0%	27.1%	19.0%	17.9%	20.4%	17.9%
50+	24.3%	18.5%	24.3%	24.3%	22.8%	24.2%	27.7%	21.1%	27.7%

Table 2.1.4., Table 2.1.5., and Table C.1. in Appendix C further break down stops by race/ethnicity and agency for all Tier 1, Tier 2, and Tier 3 agencies, respectively, for stops occurring from July 1, 2023, through June 30, 2024, the most recent year of data collection.

Table 2.1.4. **Race/Ethnicity Reporting for Tier 1 Agencies– Year 6**

Agency	Asian/PI	Black	Latinx	Middle Eastern	Native American	White	Total
Beaverton PD	821	1,394	3,843	507	112	8,865	15,542
Clackamas CO SO	914	1,023	2,637	272	64	12,149	17,059
Eugene PD	307	689	997	0	0	8,841	10,834
Gresham PD	180	548	929	57	25	2,014	3,753
Hillsboro PD	649	531	2,863	306	54	5,437	9,840
Marion CO SO	294	370	3,188	145	9	8,919	12,925
Medford PD	63	161	848	24	7	2,661	3,764
Multnomah CO SO	455	1,348	2,070	154	60	6,511	10,598
Oregon State Police	5,561	6,721	31,506	3,070	1,490	163,489	211,837
Portland PB	1,386	3,914	3,228	432	114	12,526	21,600
Salem PD	85	157	1,090	35	17	2,293	3,677
Washington CO SO	1,634	1,323	6,109	801	74	13,899	23,840
Total Tier 1	**12,349**	**18,179**	**59,308**	**5,803**	**2,026**	**247,604**	**345,269**

Table 2.1.5. **Race/Ethnicity Reporting for Tier 2 Agencies– Year 6**

Agency	Asian/PI	Black	Latinx	Middle Eastern	Native	White	Total
Albany PD	122	207	1,005	35	14	5,917	7,300
Ashland PD	63	88	165	20	2	1,817	2,155
Bend PD	41	66	233	8	3	3,053	3,404
Benton CO SO	207	242	748	144	14	5,346	6,701
Canby PD	97	91	1,042	23	6	3,104	4,363
Central Point PD	33	40	294	4	0	1,282	1,653
Corvallis PD	415	308	702	173	37	5,790	7,425
Deschutes CO SO	107	100	733	36	20	5,486	6,482
Douglas CO SO	23	37	104	14	0	1,566	1,744
Forest Grove PD	144	159	1,628	59	14	3,182	5,186
Grants Pass PD	25	37	129	4	0	1,742	1,937
Hermiston PD	39	79	1,716	8	30	1,999	3,871
Hood River CO SO	51	10	249	16	0	830	1,156
Jackson CO SO	121	159	1,106	43	3	5,971	7,403
Keizer PD	26	44	417	8	0	809	1,304
Klamath CO SO	24	17	67	5	1	416	530
Klamath Falls PD	100	59	242	13	6	1,527	1,947
Lake Oswego PD	295	236	457	124	41	4,009	5,162
Lane CO SO	89	220	425	32	1	5,819	6,586
Lebanon PD	8	21	54	3	0	964	1050
Lincoln CO SO	104	32	231	21	8	1,795	2,191
Lincoln City PD	74	35	302	10	0	1,443	1,864
Linn CO SO	70	106	529	25	16	5,998	6,744
McMinnville PD	44	32	536	9	1	1,692	2,314

(Table 2.1.5. continued on next page)

Milwaukie PD	259	472	759	100	28	5,003	6,621
Newberg-Dundee PD	109	93	743	26	0	3,356	4,327
OHSU PD	5	7	7	6	0	36	61
Oregon City PD	161	274	646	54	23	5,041	6,199
Polk CO SO	101	104	727	32	7	2,401	3,372
Port of Portland PD	116	227	207	49	2	1,082	1,683
Redmond PD	81	59	762	21	0	4,586	5,509
Roseburg PD	29	43	158	18	2	2,701	2,951
Springfield PD	99	311	625	2	0	6,290	7,327
Tigard PD	337	437	1,105	231	37	3,456	5,603
Tualatin PD	153	135	628	70	10	2,593	3,589
UO PD	12	30	29	8	1	273	353
West Linn PD	118	63	205	61	12	1,928	2,387
Woodburn PD	23	14	785	4	2	485	1,313
Yamhill CO SO	107	135	1,118	44	12	3,885	5,301
Total Tier 2	**4,032**	**4,829**	**21,618**	**1,563**	**353**	**114,673**	**147,068**

Table 2.1.6. displays the most serious dispositions reported by law enforcement for the most recent year of data collection. Most police stops did not result in further action taken against the stopped individual. The most common outcome of a stop regardless of type or Tier was a warning[3]. Over 75 percent of stops by Tier 3 agencies end in no action or a warning, which is a higher proportion than Tier 1 and Tier 2 agencies. Juvenile summons remains a rare outcome as in past reports.

Table 2.1.6. Disposition by Stop Type and Tier

	Tier 1			**Tier 2**			**Tier 3**		
Disposition	**Traffic**	**Ped.**	**Total**	**Traffic**	**Ped.**	**Total**	**Traffic**	**Ped.**	**Total**
None	1.6%	2.6%	1.6%	5.8%	13.7%	5.9%	3.7%	8.5%	3.7%
Warning	62.8%	47.8%	62.6%	63.9%	60.4%	63.8%	72.4%	73.6%	72.5%
Citation	34.3%	35.9%	34.3%	28.8%	17.0%	28.7%	22.9%	5.4%	22.8%
Juv. Summons	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Arrest	1.4%	13.7%	1.5%	1.5%	8.9%	1.6%	0.9%	12.6%	0.9%

Table 2.1.7. provides Year 6 search data, stratified by Tier. Tier 1 agencies conducted searches in 1.5 percent of stops, a higher percentage than Tier 2 and Tier 3 agencies. Pedestrians were searched more often than drivers for all Tiers, but the rate of successful searches varied by Tier. For Tier 1 agencies, 27.3 percent of all searches were consent searches, while 42.0 percent of all Tier 2 searches were consent searches. Tier 3 agencies had the least consent searches of the Tiers, at just under 20 percent of all searches. Echoing previous STOP reports, drugs were the most common form of contraband found in Tier 1 and Tier 2 searches. Conversely, Tier 3 agencies found alcohol most often (38.5 percent) during a search and found alcohol more often than Tier 2 (31.1 percent) or Tier 1 agencies (20.5 percent).

[3] It is the policy of many agencies to give a warning to everyone who is stopped.

Table 2.1.7. **Search Results by Stop Type and Tier**

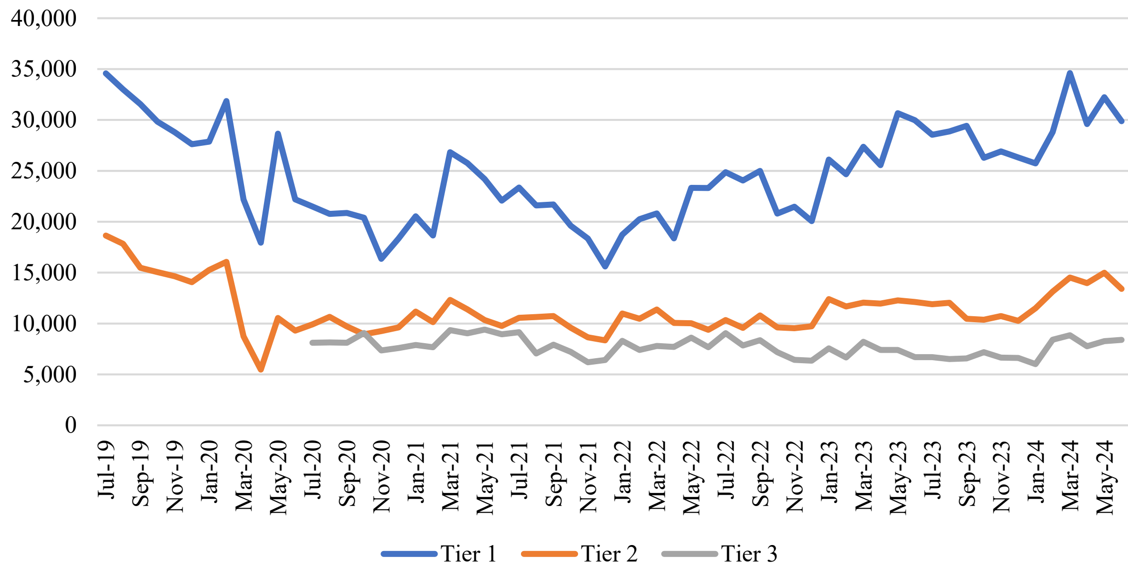
	Tier 1			Tier 2			Tier 3		
	Traf.	**Ped.**	**Total**	**Traf.**	**Ped.**	**Total**	**Traf.**	**Ped.**	**Total**
Search Conducted	1.2%	29.0%	1.5%	1.0%	7.2%	1.1%	0.5%	16.3%	0.5%
Reason*									
Consent Search	30.0%	14.5%	27.3%	41.5%	50.6%	42.0%	21.3%	3.8%	19.3%
Consent Search Denied	0.6%	0.6%	0.6%	0.1%	0.0%	0.1%	0.7%	0.0%	0.7%
'Other' Search	70.8%	85.6%	73.3%	61.9%	51.9%	61.4%	81.9%	96.2%	83.5%
Percent Successful	40.5%	72.9%	46.0%	47.1%	43.2%	46.9%	50.4%	21.6%	47.1%
Item Seized**									
Alcohol	27.5%	1.7%	20.5%	32.1%	11.4%	31.1%	40.1%	9.1%	38.5%
Drugs	44.3%	90.7%	56.9%	38.8%	60.0%	39.8%	36.1%	54.5%	37.1%
Weapons	18.7%	3.1%	14.5%	14.0%	14.3%	14.0%	12.4%	18.2%	12.7%
Stolen Property	4.0%	3.2%	3.8%	2.4%	8.6%	2.7%	0.0%	27.3%	2.3%
Other Evidence	14.7%	5.0%	12.1%	10.4%	2.9%	10.1%	13.9%	0.0%	13.1%
Other Non-Evidence	8.1%	2.8%	6.7%	19.8%	22.9%	20.0%	11.4%	18.2%	11.7%

*Officers can designate multiple search types on a stop, therefore in some cases, percentages for Search Reason may add up to more than 100 percent.
** Items Seized includes only incidents where an item was seized during a search. Multiple types of items may be seized during a search, therefore Item Seized totals may equal more than 100 percent.

2.2. Longitudinal STOP Data Trends

While the analyses contained in Sections 3, 4, and 5 utilize two years of submitted data, this section analyzes data collected by the STOP Program for officer-initiated traffic and pedestrian stops since Tier 2 agencies began reporting in 2019. Figure 2.2.1. displays stops made by Oregon law enforcement agencies from July 2019 through June 2024, stratified by Tier. While Tier 1 and Tier 2 agencies began reporting in 2018 and 2019 respectively, Tier 3 agencies were not required to submit data until July 2020. From February to April 2020, when COVID-19 mitigation efforts were first put in place, Tier 1 stop volume dropped by 44 percent and Tier 2 stop volume dropped by a greater percentage, 66 percent.

Figure 2.2.1. Monthly Stops by Tier



As COVID-19 vaccines became more widely available, stop volume increased and generally peaked in March 2021. From March to December 2021, stop volume shows an overall decline, likely due to subsequent COVID-19 waves, case counts, and other resource challenges including staffing shortages. As case counts declined and the pandemic abated, stop volume increased by varying levels across tiers. From December 2022 to June 2024, Tier 1 agencies show a 49 percent increase in stop volume, while Tier 2 agencies increased 34 percent, and Tier 3 agencies show a 32 percent increase.

In March 2022, the Oregon Legislature passed SB 1510[4], which includes several public safety law changes. Sections 1 through 8 specifically address law enforcement officer stops of individuals. Sections 1 and 2 require officers to inform a person that they have the right to refuse a consent search request. Section 6 modifies vehicle lighting violations such that an officer may not initiate a traffic stop if certain criteria are met. While these changes were effective January 1, 2023, many agencies implemented them when the bill passed. Table 2.2.1. shows search rates by Tier and Year and includes searches from July 2018 to June 2024. Overall search rates have dropped, with Tier 1 agencies showing a search rate of 2.9 percent in Year 1 and dropping to 1.5 percent in Years 5 and 6. Tier 2 agencies drop from 2.8 percent in Year 2 to 1.1 percent in Year 6. Finally, Tier 3 agencies show a search rate of half a percent in Year 6.

Table 2.2.1. **Search Rates by Year and Tier**

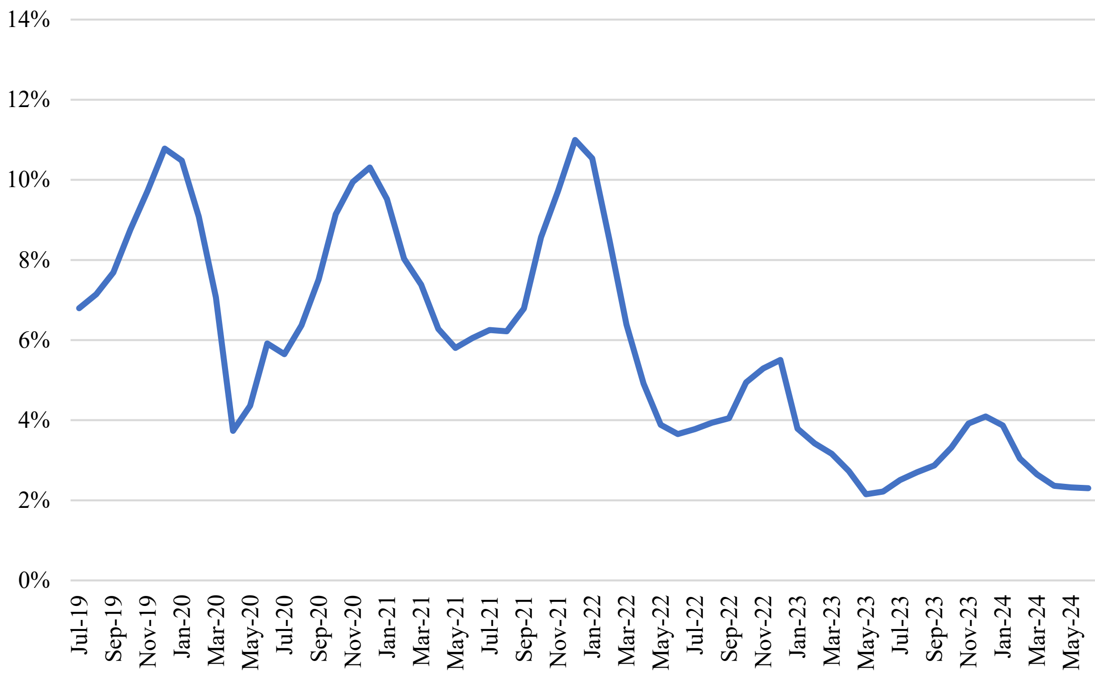
Year	Tier 1	Tier 2	Tier 3
Year 1 (18-19)	2.9%	N/A	N/A
Year 2 (19-20)	2.6%	2.8%	N/A
Year 3 (20-21)	2.5%	1.9%	1.4%
Year 4 (21-22)	2.2%	1.6%	0.9%
Year 5 (22-23)	1.5%	1.3%	0.7%
Year 6 (23-24)	1.5%	1.1%	0.5%

Figure 2.2.2. shows the percent of stops for lighting violations from July 2019 to June 2024. The lighting

[4] https://olis.oregonlegislature.gov/liz/2022R1/Downloads/MeasureDocument/SB1510/Enrolled

violations include stops for ORS 811.520, Unlawful Use or Failure to Use Lights, and ORS 816.330, Operation Without Required Lighting Equipment. The historic trend shows a seasonal increase in the percentage of stops in the winter months, as expected with more hours of darkness. For the seasonal peak in December 2021, lighting violations accounted for 11 percent of stops. The percentage of stops decreased with the passage and implementation of SB 1510 in March 2022 and January 2023 respectively. The seasonal peak in December 2023 is less than half the rate of the peak year at 4.1 percent. The percent of stops in June 2023 and June 2024 show historic lows of just over 2 percent.

Figure 2.2.2. Percent of Monthly Stops for Lighting Violations



2.3. Limitations

The data collected by the STOP Program for the State of Oregon represent one of the most robust stop data collection efforts in the United States. While data are collected by some jurisdictions in most states, few states can boast a statewide, statutorily mandated data collection effort like Oregon's. This robust database and the statistical evaluation of stop data can form the foundation of a transparent dialogue between state leaders, government agencies, law enforcement, and the communities law enforcement agencies serve.

Despite its promise as a means for systematically analyzing statewide data concerning police-citizen interactions, the STOP Program and its associated data and analyses have limitations. First, the statistical analyses can only identify disparities in police/citizen interactions during discretionary stops. This means that the analyses contained in this report cannot be used either as absolute proof that a law enforcement agency engaged in racially biased conduct or as disproof of racially biased conduct. Further, the results in this report are conducted at the police agency level because HB 2355 expressly forbids the collection of data that identify either stopped individuals or officers. These analyses, therefore, can only identify systematic disparities across a law enforcement agency or at a larger level of aggregation. As such, regardless of whether a department is reported to have an identified disparity or not, this report cannot and does not discount or speak to the personal experiences of individuals who have been subjected to biased treatment.

Despite these limitations, the statistical results presented in the following sections demonstrate that after the application of rigorous standards, if multiple disparities are identified for an agency, then there is cause for concern, further investigation, and technical assistance. STOP Program researchers have selected highly respected, thoroughly vetted and peer reviewed, cutting-edge analyses. The STOP Program stands behind the significant amount of work that went into the analyses and crafting of this report and believes that the results presented herein will contribute to the dialogue between law enforcement and Oregonians.

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[5], 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 to test for disparities when officers can more easily perceive the race/ethnicity of drivers. 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. The chief advantage of this approach is that the analysis does not rely on a benchmark comparison with the estimated driving or residential population to the population of stopped individuals. Rather, the DTS analysis takes advantage of natural variations in daylight over the course of the year to compare minority stops made in daylight to those made in darkness at similar times of the day when commuting patterns should be relatively consistent.

More specifically, the DTS analysis relies on comparing the racial composition of individuals 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. Given 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.

Beyond this central assumption underlying the DTS approach, the analytical test also assumes that driving behavior does not change throughout the year or between daylight and darkness, and that driving patterns have little seasonal variation during the morning and evening commute times. While this assumption is likely too strong and not reflective of actual driving patterns, it can be accounted for statistically by including additional control variables available in the STOP Program database, such as: age, gender, reason for stop, day of week, time of day, quarter or season, county stop volume, and agency stop volume.

To accomplish the analysis described above, the DTS approach tests whether the odds of non-white traffic stops during daylight are significantly different from the odds of non-white traffic stops during darkness. In the table that follows in the next subsection, this difference in odds is presented as an odds ratio, which displays the change in odds for non-white stops during daylight compared to darkness. 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

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

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. In sum, following best practices, the STOP Program identifies all agencies with disparities above 1.0 that are statistically significant at the 95 percent confidence level in any minority group at the agency level.

3.1. Agency-Level Decision to Stop Analysis

The following analyses utilized two years of data for Tier 1, Tier 2, and Tier 3 agencies. At the agency level, therefore, it is possible to estimate DTS models for many of the non-white groups reported in the stop database given a sufficient sample size. As described in Appendix B.2.3.2, the sample size requirement for the DTS model was at least 100 stops in each racial/ethnic group within the inter-twilight windows for the two years of data provided. Twenty-two agencies had sufficient sample sizes to run models for Black drivers. Models for Latinx drivers were run for 80 agencies. Models for Asian/PI drivers were run for 14 agencies. Models for Middle Eastern drivers were run for seven agencies. Only one agency met the sample size criteria to run a model for Native American drivers. While a number of agencies have odds ratios above 1.0, five agencies showed statistically significant differences in the rate of stopped drivers in daylight compared to darkness[6]. Table 3.1.1. displays the odds ratios for the DTS models for non-white stopped drivers with statistically significant results. Results for all models are shown in Appendix D.

Table 3.1.1. Decision to Stop Analyses

Agency	Race/Ethnicity	Odds Ratio
Corvallis PD	Latinx	1.81**
Canby PD	Latinx	1.91**
Deschutes CO SO	Latinx	1.48*
Milwaukie PD	Black	2.17*
Sandy PD	Latinx	1.75*

* p<0.05, ** p<0.01, *** p<0.001

As shown in Table 3.1.1, five agencies showed statistically significant differences in the odds of minority stops in daylight compared to darkness. For Corvallis Police Department (PD), the odds of stops for Latinx drivers in daylight were 1.81 times the odds for white drivers. For Canby PD, the odds for Latinx drivers being stopped in the daylight were 1.91 times the odds of white drivers. For Deschutes County Sheriff's Office, the odds for Latinx drivers being stopped in daylight were 1.48 times the odds for white drivers. For Milwaukie PD, the odds of stops for Black drivers in daylight were 2.17 times the odds of white drivers. For Sandy PD, the odds of stops for Latinx drivers in daylight were 1.75 times the odds for white drivers. Analyses for these five agencies indicated statistically significant differences, evidencing disparities in the rate of stopped drivers in daylight compared to darkness. These five agencies have initiated additional analysis of the STOP data.

4. Stop Outcomes Analysis

This report presents results from two analyses assessing outcomes occurring after the initial stop decision has been made and an individual has been stopped by law enforcement. The first of these two approaches, the Stop Outcomes Analysis, is presented in this section. The Stop Outcomes Analysis focuses on the

[6] The odds ratio for Oregon State Police for Native American drivers (1.29) shows a p-value of 0.045. The odds ratio for Middle Eastern drivers (1.22) shows a p-value of 0.018. With the Bonferroni adjustment with four comparisons, these are not statically significant. However, without the adjustment, the p-values are below the 0.05 threshold.

outcomes of stops, including whether stopped individuals were cited, searched, and/or arrested during their encounter with law enforcement. The analysis estimates whether each race/ethnic group is more likely than the white group to have a stop end in the disposition in question 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[7]. 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.

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, if heavy traffic flows prevent drivers from exceeding the speed limit then the likelihood of receiving a citation for speeding would be reduced at that time. Variation could also be attributed to other factors, including age of the driver, gender of the driver, or season of the stop. Propensity score analysis is employed here to account for as many of these differences as possible and isolate the effect, if any, that the race of the stopped individual has on the disposition of the stop.

Propensity score methods have a long and well-established history in applied statistics. STOP Program researchers use these methods to determine, when other factors are held constant, whether there are different dispositional outcomes across racial/ethnic groups. 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. This approach enables STOP Program researchers to make the white comparison group look 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, stop volume) are identical across the two groups then the remaining difference in outcomes is evidence of a disparity due to racial/ethnic differences (Ridgeway, 2006).

Many different propensity score methods have been developed, but they all have a similar goal of making two groups comparable to one another. The best of these methods to employ for a given research program depends on available data, sample size, data completeness, and other factors; there is no one-size-fits-all approach. Here, the STOP Program employed Inverse Probability Weighted Regression Adjustment[8].

The current analysis includes twenty sub-analyses for each agency: each outcome of citation, search, arrest, or any non-warning disposition across each racial/ethnic group of Asian/PI, Black, Latinx, Middle Eastern, and Native American individuals. The comparison group is drawn from the group of white stops for the agency in question. Bonferroni adjustments are applied at the agency level based on the number of

[7] 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.
[8] Inverse Probability Weighted Regression Adjustment weights the groups based on the propensity score and then uses these weighted data to estimate the effect of race/ethnicity on dispositional outcomes through regression analysis. For a thorough discussion of this methodology see Appendix E.

analyses completed.[9]

4.2. Stop Outcomes Results

As with the Decision to Stop analysis in the previous section, the analyses conducted here includes two years of data for all agencies. Table 4.2.1. reports agency-level results for agencies where a statistically significant disparity is found for a search or arrest outcome, sometimes in addition to citation or any outcome. For seven agencies, Boardman PD, Gilliam CO SO, Hermiston PD, Marion CO SO, Oregon State Police[10], Portland PB, and Washington CO SO, disparities were reported for either searches and/or arrests of Latinx individuals, sometimes in addition to citations. Oregon State Police also showed a disparity for arrests of Native American individuals and searches of Black individuals. Portland PB showed a disparity in searches and arrests of Black individuals.

Table 4.2.1. **Stop Outcome Analyses**

Agency	Race/ Ethnicity	Citation		Search		Arrest		Any Outcome	
		Actual	Pred.	Actual	Pred.	Actual	Pred.	Actual	Pred.
Boardman PD	Latinx	26.3%	17.5%	--	--	11.6%	3.8%	34.6%	21.4%
Gilliam CO SO	Latinx	70.9%	59.2%	4.1%	2.0%	4.8%	2.5%	72.3%	60.2%
Hermiston PD	Latinx	24.6%	17.6%	--	--	3.1%	1.7%	26.9%	19.0%
Hermiston PD	Native	32.8%	14.2%	--	--	--	--	--	--
Marion CO SO	Latinx	--	--	3.9%	2.9%	4.5%	2.9%	--	--
Oregon State Police	Asian	34.9%	33.3%	--	--	--	--	35.4%	34.0%
Oregon State Police	Black	39.7%	34.8%	1.8%	1.2%	--	--	40.8%	35.7%
Oregon State Police	Latinx	41.2%	34.3%	1.3%	1.0%	1.1%	0.9%	42.3%	35.1%
Oregon State Police	Mideast	38.7%	33.9%	--	--	--	--	38.9%	34.5%
Oregon State Police	Native	41.2%	33.8%	--	--	2.4%	1.2%	43.0%	34.9%
Portland PB	Black	--	--	7.0%	4.4%	6.2%	4.5%	--	--
Portland PB	Latinx	--	--	6.4%	4.1%	--	--	--	--
Washington CO SO	Latinx	24.3%	19.9%	--	--	3.0%	2.3%	26.6%	21.7%

*Only statistically significant results following Bonferroni corrections are displayed in this table. The Bonferroni correction is described in Appendix B. For full results by agency, please see the STOP dashboards.

Where disparities were found, the average gap in the predicted versus the actual disposition rate varied by agency and type of disposition. These differences may be especially apparent between large and small agencies. Larger agencies make more stops and thus have a greater sample size, which leads to more precise statistical tests and a lower threshold for identifying statistically important differences. Agencies where a statistically significant disparity was found for only either a citation or for the combined measure of all dispositions (i.e., citation, search, or arrest; referenced as "Any Outcome") are displayed in Table 4.2.2. For four Tier 1 agencies, Beaverton PD, Clackamas CO SO, Hillsboro PD, and Salem PD disparities were detected only for citations and/or for the combined measure of all dispositions (i.e., citation, search, or arrest). This indicates that, for these agencies, it is likely that the only relevant

[9] 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.
[10] In partnership with Portland Central City Task Force, Oregon State Police (OSP) began focused enforcement in October 2023 in Portland on specific offenses related to fentanyl possession and delivery. This has resulted in increased stop volume for OSP in Multnomah County, and in particular an increase in pedestrian stops. The pedestrian stops have a higher search rate than previous trends. https://www.portlandcentralcitytaskforce.com/

disparity is for citations and not the other outcomes. Tier 2 agencies tend to have far fewer stops than Tier 1 agencies. Combined with the already relatively low minority populations in the state, and especially outside of major metro areas, many of the Stop Outcome analyses for the Tier 2 agencies did not have sufficient sample sizes to complete the analysis. That said, of the analyses that were completed, Canby PD, Deschutes CO SO, Forest Grove PD, Linn CO SO, Oregon City PD, Polk CO SO, Tigard PD, Tualatin PD, West Linn PD, Woodburn PD, and Yamhill CO SO had statistically significant disparities indicated for at least one of citations and any outcome.

Table 4.2.2. Predicted Stop Outcome – Citation & Any Outcomes

Agency	Race/Ethnicity	Citation		Any Outcome	
		Actual	Pred.	Actual	Pred.
Astoria PD	Latinx	35.9%	22.2%	44.7%	34.5%
Beaverton PD	Latinx	34.0%	31.7%	38.2%	35.6%
Canby PD	Latinx	40.3%	33.2%	41.8%	34.7%
Cannon Beach PD	Latinx	23.7%	14.6%	27.1%	18.0%
Clackamas CO SO	Latinx	31.6%	29.3%	33.8%	31.3%
Clackamas CO SO	Native	39.1%	30.0%	--	--
Cottage Grove PD	Latinx	43.2%	22.8%	43.2%	23.5%
Deschutes CO SO	Latinx	14.5%	10.9%	17.1%	13.0%
Forest Grove PD	Latinx	31.2%	23.0%	33.0%	24.7%
Hillsboro PD	Latinx	29.7%	25.8%	32.0%	27.7%
Hubbard PD	Latinx	33.1%	25.1%	35.5%	26.9%
Linn CO SO	Latinx	39.9%	34.2%	39.1%	33.4%
Morrow CO SO	Latinx	25.4%	18.5%	26.6%	19.0%
Nyssa PD	Latinx	48.9%	21.6%	67.6%	52.8%
Oregon City PD	Latinx	33.5%	28.5%	35.3%	30.5%
Phoenix PD	Latinx	37.9%	29.4%	38.3%	29.9%
Polk CO SO	Latinx	26.8%	20.6%	28.5%	21.8%
Salem PD	Latinx	62.9%	59.9%	65.2%	61.9%
Sherwood PD	Latinx	--	--	31.5%	27.0%
Sutherlin PD	Latinx	50.2%	38.3%	50.7%	38.5%
Tigard PD	Latinx	38.4%	30.9%	40.4%	33.7%
Tualatin PD	Latinx	51.9%	44.2%	52.9%	45.1%
Turner PD	Latinx	--	--	37.1%	24.1%
Umatilla CO SO	Latinx	24.3%	15.7%	27.7%	18.1%
Umatilla PD	Latinx	32.0%	24.4%	33.2%	25.3%
Woodburn PD	Latinx	--	--	12.3%	8.2%
Yamhill CO SO	Latinx	--	--	27.2%	24.1%

*Only statistically significant results following Bonferroni corrections are displayed in this table. The Bonferroni correction is described in Appendix B. For full results by agency, please see the STOP dashboards.

Sample size issues were even more pronounced for Tier 3 agencies. However, the following Tier 3 agencies were identified as having significant disparities in only citations and/or any disposition for one of the analysis groups: Astoria PD, Cannon Beach PD, Cottage Grove PD, Hubbard PD, Morrow CO SO, Nyssa PD, Phoenix PD, Sherwood PD, Sutherlin PD, Turner PD, Umatilla CO SO, and Umatilla PD.

4.3. Stop Outcomes Analysis including the Reason for the Stop Outcome

In February 2021 law enforcement agencies started submitting additional data to the STOP Program on the reason for the most serious disposition of each stop. Previously, for example, if an officer stopped someone for a moving violation but the stop ended in an arrest because of an outstanding warrant, analysts would only be able to see a moving violation ending in arrest, which is inaccurate. This additional data point allows the STOP program analysts to more accurately account for the reason for the stop outcome in addition to the reason for the stop. These additional data points are submitted voluntarily by STOP agencies and are not statutorily required data elements. The quality and completeness of these data submitted to CJC is, thus, inconsistent. For Years 5 and 6 of data collection (July 2022-June 2024), 43.6 percent of stops with a citation, search, or arrest outcome had a missing most serious disposition code value.[11,12] The CJC uses this data element to run an adjusted Stop Outcomes analysis to provide additional context to the baseline results.

Most serious disposition reasons are inconsistently reported across agencies, with some agencies reporting little or no additional data. Of agencies identified by the baseline stop outcome analysis in the previous section, six agencies either did not submit any additional data or submitted a small amount of data that is insufficient to conduct the additional analysis. As seen in Table 4.3.1, the six agencies are: Gilliam CO SO, Linn CO SO, Marion CO SO, Nyssa PD, Portland PB, and Turner PD. Submission of the additional data is not required but allows the CJC to provide the adjusted Stop Outcomes analysis displayed is this section. Other agencies had relatively low reporting rates but submitted sufficient data to run an analysis that differed from the baseline analysis.

Beyond agencies that reported insufficient data, reporting practices varied widely. Some agencies submitted outcome reason information only when the outcome reason differed from the stop reason. Other agencies submitted the outcome reason on all or close to all stops, regardless of whether the outcome reason differed from the stop reason. In either case, if a sufficient volume of additional data was submitted, the additional analysis could be conducted. The percent of stops with a non-warning outcome that are missing the outcome reason are reported in Table 4.3.1.

[11] This subset of outcomes is useful for determining whether additional analysis is possible. When including all stops and counting Warning or None outcomes as non-missing, 14.6% of observations were missing outcome reason information.

[12] This is an overall improvement in reporting rates from last years STOP report, where 47.9% of these stops were missing a most serious disposition value.

Table 4.3.1. **Observations Missing Stop Outcome Reason**

	Agency	Percent of citation, search, and arrest outcomes with missing outcome reason
Agencies who did not submit enough data for additional analysis.	Gilliam CO SO	100.0%
	Linn CO SO	100.0%
	Marion CO SO	100.0%
	Nyssa PD	100.0%
	Portland PB	100.0%
	Turner PD	100.0%
Agencies that submitted enough data for additional analysis.	Astoria PD	57.8%
	Beaverton PD	3.7%
	Boardman PD	0.8%
	Canby PD	62.6%
	Cannon Beach PD	53.1%
	Clackamas CO SO	55.1%
	Cottage Grove PD	8.3%
	Deschutes CO SO	59.4%
	Forest Grove PD	1.5%
	Hermiston PD	1.5%
	Hillsboro PD	19.0%
	Hubbard PD	55.8%
	Morrow CO SO	0.7%
	Oregon City PD	55.2%
	Oregon State Police[1]	1.4%
	Phoenix PD	83.5%
	Polk CO SO	65.6%
	Salem PD	13.7%
	Sherwood PD	72.1%
	Sutherlin PD	33.3%
	Tigard PD	5.8%
	Tualatin PD	91.1%
	Umatilla CO SO	0.0%
	Umatilla PD	0.4%
	Washington CO SO	14.6%
	Woodburn PD	81.5%
	Yamhill CO SO	73.8%

[1] Oregon State Police submitted sufficient data for the additional citation analysis, but not for arrest or search outcomes.

For the additional analysis, the CJC creates a variable indicating whether the stop outcome was a "low-discretion" offense. Oregon State Police policy identifies offenses that allow the stopping officer relatively low discretion in their decision to cite, search, and/or arrest the stopped individual. Discussions with police agencies identified the reasons for these policies as a combination of limiting liability[13] and state-level policies. Exact policies vary by agency, but discussions with agencies suggest that the Oregon State Police policy represents a norm across the state and the best basis for a consistent variable across agencies.[14] When the officer reports the reason for the most serious disposition on the stop as one of these

[13] For example, if a police officer stops an individual who is unlicensed that officer may be held liable if they allow that person to drive after that stop.

[14] The policy indicates that stops where the following were present require additional officer actions: Reckless Endangering Another Person (ORS 163.195), Aggravated Driving while Suspended or Revoked (ORS 163.196), Driving Uninsured (ORS 806.01), licensing violations (ORS 807.010, ORS 807.570), Failure to Yield to Pedestrian (ORS 811.025), Reckless Driving (ORS 811.140), Driving while Suspended or Revoked (ORS 811.175), Criminal

statutes it severely limits the officer's discretion in allowing the driver to continue driving with just a warning. This variable is added to the baseline outcome analysis, which controls for the fact that some groups may have a higher or lower propensity to be cited, searched, and/or arrested for one of these low-discretion offenses.

In Tables 4.3.2 through 4.3.5, lightly shaded predicted values indicate a result that is not statistically different (i.e., insignificant) than the actual outcome rate. Conversely, results that are the standard shade indicate a result that is statistically different (i.e., significant) than the actual outcome value. So, for example, when a result in the Original analysis column is the standard shade and the result in the Low-Discretion column is lightly shaded this indicates that the inclusion of the low-discretion variable caused the originally significant estimate to become insignificant.

Table 4.3.2 presents the baseline and additional analysis results for agencies where citation outcomes were significant in the baseline analysis. For most of these agencies the inclusion of the low-discretion variable in the analysis changes statistically significant differences between the actual and predicted citation rates to become statistically insignificant. Where statistically significant results remained, the difference between the actual and predicted rates typically closed. This suggests that, systematically across police agencies in Oregon, low-discretion policies *tend* to increase perceptible racial disparities in citations for some groups.

For one agency, Tigard PD, the addition of the low-discretion variable led to statistically significant results for the Middle Eastern group where these results were insignificant in the baseline analysis. This suggests that drivers perceived as Middle Eastern have a relatively small proportion of stops resulting in low-discretion citations for, at least, this agency.

Driving while Suspended or Revoked (ORS 811.182), Reckless Endangerment of Highway Workers (ORS 811.231), Fleeing or Attempting to Elude (ORS 811.540), Failure to Perform Duties of a Driver (ORS 811.700, ORS 811.705), Driving Under the Influence of Intoxicants (ORS 813.010), Fleeing (811.540), Sanctions (ORS 33.045), Warrants (ORS 135.280), Failure to Appear in the First Degree (162.205), and controlled substance violations (ORS 475.752)

Table 4.3.2. **Predicted Citation Outcome – Baseline v. Low Discretion Analysis**

Agency	Race/ Ethnicity	Actual	Predictions by Analysis	
			Original	Low-Discretion
Astoria PD	Latinx	35.9%	22.2%	22.4%
Beaverton PD	Latinx	34.0%	31.7%	33.4%[1]
Boardman PD	Latinx	26.3%	17.5%	27.5%[1]
Canby PD	Latinx	40.3%	33.2%	35.4%
Cannon Beach PD	Latinx	23.7%	14.6%	--[1]
Clackamas CO SO	Latinx	31.6%	29.3%	31.3%[1]
Clackamas CO SO	Native	39.1%	30.0%	29.1%
Cottage Grove PD	Latinx	43.2%	22.8%	31.3%
Deschutes CO SO	Latinx	14.5%	10.9%	11.8%[1]
Forest Grove PD	Latinx	31.2%	23.0%	28.5%
Hermiston PD	Latinx	24.6%	17.6%	20.3%
Hermiston PD	Native	32.8%	14.2%	14.4%
Hillsboro PD	Latinx	29.7%	25.8%	28.7%[1]
Hubbard PD	Latinx	33.1%	25.1%	30.4%[1]
Morrow CO SO	Latinx	25.4%	18.5%	22.9%[1]
Oregon City PD	Latinx	33.5%	28.5%	30.3%[1]
Oregon State Police	Asian	34.9%	33.3%	30.1%
Oregon State Police	Black	39.7%	34.8%	38.0%
Oregon State Police	Latinx	41.2%	34.3%	38.3%
Oregon State Police	Mideast	38.7%	33.9%	29.8%
Oregon State Police	Native	41.2%	33.8%	42.7%[1]
Phoenix PD	Latinx	37.9%	29.4%	31.8%
Polk CO SO	Latinx	26.8%	20.6%	23.0%
Salem PD	Latinx	62.9%	59.9%	60.8%[1]
Sutherlin PD	Latinx	50.2%	38.3%	38.2%
Tigard PD	Latinx	38.4%	30.9%	34.6%[1]
Tigard PD	Mideast	36.3%	30.4%	28.4%[2]
Tualatin PD	Latinx	51.9%	44.2%	--[1]
Umatilla CO SO	Latinx	24.3%	15.7%	18.9%
Umatilla PD	Latinx	32.0%	24.4%	13.9%
Washington CO SO	Latinx	24.3%	19.9%	23.1%[1]

Unless otherwise indicated, results remained statistically significant.
1 Indicates a result that went from statistically significant to statistically insignificant.
2 Indicates a result that went from statistically insignificant to statistically significant.

Of agencies that had statistically significant search outcome results, no agencies submitted sufficient data to complete the low-discretion analysis.

Table 4.3.3. Predicted Search Outcome – Baseline v. Low Discretion Analysis

*No results to report for this year's analysis

Of agencies that had statistically significant arrest outcome results, three agencies submitted sufficient outcome reason data to conduct the additional analysis. One agency's results became insignificant with the inclusion of the low-discretion variable, as seen in Table 4.3.4, but the other two agency's gap between actual and predicted grew smaller.

Table 4.3.4. Predicted Arrest Outcome – Baseline v. Low Discretion Analysis

Agency	Race/ Ethnicity	Actual	Predictions by Analysis	
			Original	Low-Discretion
Boardman PD	Latinx	11.6%	3.8%	5.8%
Hermiston PD	Latinx	3.1%	1.7%	1.9%
Washington CO SO	Latinx	3.0%	2.3%	3.2%[1]

Unless otherwise indicated, results remained statistically significant.
[1] Indicates a result that went from statistically significant to statistically insignificant.

Similar to the citation outcome results, most agencies that had a statistically significant result for the any outcome had the estimated gaps grow smaller, as seen in Table 4.3.5 below. In many of these cases, the results became statistically insignificant in the low-discretion analysis.

Table 4.3.5. Predicted Any Outcome – Baseline vs. Low Discretion Analysis

Agency	Race/ Ethnicity	Actual	Predictions by Analysis	
			Original	Low-Discretion
Astoria PD	Latinx	44.7%	34.5%	34.7%
Beaverton PD	Latinx	38.2%	35.6%	37.9%
Boardman PD	Latinx	34.6%	21.4%	35.2%
Canby PD	Latinx	41.8%	34.7%	37.2%
Cannon Beach PD	Latinx	27.1%	18.0%	--[1]
Clackamas CO SO	Latinx	33.8%	31.3%	31.2%
Cottage Grove PD	Latinx	43.2%	23.5%	31.8%
Deschutes CO SO	Latinx	17.1%	13.0%	14.0%
Forest Grove PD	Latinx	33.0%	24.7%	30.5%[1]
Hermiston PD	Latinx	26.9%	19.0%	22.1%
Hillsboro PD	Latinx	32.0%	27.7%	31.1%[1]
Hubbard PD	Latinx	35.5%	26.9%	32.7%[1]
Morrow CO SO	Latinx	26.6%	19.0%	24.0%[1]
Oregon City PD	Latinx	35.3%	30.5%	32.5%[1]
Oregon State Police	Asian	35.4%	34.0%	30.8%
Oregon State Police	Black	40.8%	35.7%	38.7%
Oregon State Police	Latinx	42.3%	35.1%	39.0%
Oregon State Police	Mideast	38.9%	34.5%	30.4%
Oregon State Police	Native	43.0%	34.9%	43.6%[1]
Phoenix PD	Latinx	38.3%	29.9%	24.0%[1]
Polk CO SO	Latinx	28.5%	21.8%	24.6%
Salem PD	Latinx	65.2%	61.9%	62.7%[1]
Sherwood PD	Latinx	31.5%	27.0%	30.1%[1]
Sutherlin PD	Latinx	50.7%	38.5%	38.6%
Tigard PD	Latinx	40.4%	33.7%	37.4%[1]
Tualatin PD	Latinx	52.9%	45.1%	45.9%
Umatilla CO SO	Latinx	27.7%	18.1%	21.9%
Umatilla PD	Latinx	33.2%	25.3%	30.1%[1]
Woodburn PD	Latinx	12.3%	8.2%	8.1%
Washington CO SO	Latinx	26.6%	21.7%	25.4%[1]
Yamhill CO SO	Latinx	27.2%	24.1%	25.2%[1]

Unless otherwise indicated, results remained statistically significant.
[1] Indicates a result that went from statistically significant to statistically insignificant.
[2] Indicates a result that went from statistically insignificant to statistically significant.

5. Search Findings Analysis

The second analysis conducted examining post-stop outcomes is the Search Findings analysis. Originally developed in the context of economics, various hit-rate models use outcomes as indicators of economic discrimination in areas such as mortgage loan decision making (Becker 1957, Becker 1993). In the past few decades, this approach to examining outcomes to identify discrimination has been adapted extensively in analyses of policing. The most widely used model is known as the KPT Hit-Rate model developed by Knowles, Persico, and Todd (2001). Throughout this report, this will be referred to as the Search Findings analysis.

The Search Findings analysis examines whether the likelihood of a "successful" police search differs across racial/ethnic groups, where success is defined as finding contraband. The model assumes that officers make the decision to search a person based on visual and other contextual evidence that they are carrying contraband (e.g., location, furtive movements, or odors associated with drugs) in order to maximize search success rates. The model also assumes that motorists adjust their decision to carry contraband based on their likelihood of being searched. In the case that a certain group is more likely to carry contraband, officers will search this group more often in order to maximize their hit-rate, and the group, as a whole, will adjust their likelihood to carry contraband downward. Eventually an equilibrium is reached at which search success rates (or hit-rates) are the same across all groups. However, if officers are subjecting a group to more frequent searches based on racial or ethnic bias, then their hit-rate for that group will decrease. If a minority group's hit-rate is less than the white hit-rate, this indicates that the minority group is "over-searched," which is evidence of a disparity. Put simply, if search decisions are based on race/ethnicity-neutral factors, then hit-rates across all racial/ethnic groups should be similar. If they are substantially dissimilar, then a disparity is identified.

Hit-rates are calculated by dividing the number of searches in which contraband was found by the total number of searches for each racial/ethnic group. The results for non-white groups are then compared to the outcomes for white individuals to determine whether the success rates are similar. Agency level search data were analyzed for disparities between the white baseline group and individuals identified as Black, Latinx, Asian/PI, Middle Eastern, and Native American. In order to perform these analyses for an agency for a particular racial/ethnic group the agency must have searched at least 30 people of both the minority group and the white group. This protects against statistical anomalies due to low search counts and aligns with best practices.[15] Because of this requirement, the Search Findings analysis was unable to be performed for certain agencies and racial/ethnic groups. Finally, chi-square tests of independence with a Bonferroni adjustment were performed for each comparison to determine if observed differences in hit-rates are statistically significant. Following best practices, the STOP Program identifies all agencies with disparities in the Search Findings analysis. For individual agencies, this includes minority group hit-rates less than the white hit-rate and statistically significant at the 95 percent confidence level. See Appendix F for more detailed technical information about the KPT Hit-Rate model and statistical tests.

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. In this report, the Search Findings analysis was performed for each agency for up to five minority racial/ethnic groups (Black, Latinx, Asian/PI, Middle Eastern, and/or Native American) depending upon sample size. Significant results for these analyses are presented in Table 5.1.1. below. All other results for combinations of agencies and races for which the test was run can be found in Appendix F.

Table 5.1.1. Search Findings Analysis

Agency	White	Black	Latinx	Asian/PI	Native	Middle Eastern
Oregon State Police	61.0%	--	54.9%**	--	--	---
* p<0.05, ** p<0.01, *** p<0.001						
Full detailed results can be found in Appendix F.						

All agencies have differences in search success rates between white individuals and the comparison

[15] Connecticut Racial Profiling Prohibition Project (2019).

groups. These differences in nearly all cases were relatively small, and only one of the differences reported was statistically significant. Small, statistically insignificant differences in search success rates are likely to occur due to random chance even in the absence of policies or practices that could lead to disparate treatment of different groups. One search findings comparison made in this report was found to be statistically significant. Oregon State Police[16] was identified as having a statistically significant disparity for the Latinx group.

6. Conclusions

The data contained in this report are intended to be used as a tool for law enforcement, community members, researchers, Legislators and policy makers, and other interested parties to focus training and technical assistance on agencies found to have disparities in outcomes for minority groups. As described previously, STOP Program researchers utilized three rigorous statistical analyses, consistent with best practices, to identify disparities in Oregon. The use of these three tests allows the STOP Program researchers to evaluate numerous decision points before and during a stop, while also providing numerous points of analysis for disparate outcomes.

To determine if identified disparities 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) An estimated disparity in an individual analysis must have met the 95 percent confidence level for it to be statistically significant. This means STOP Program researcher must be at least 95 percent confident that differences or disparities identified by the analyses were not due to random chance: (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 disparity 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, the STOP Program identified three agencies that had statistically significant results across two of the tests performed on the data. Canby PD, Deschutes CO SO, and Oregon State Police show a statistically significant disparity in two of the analytical tests described in the report. These agencies, as well as several other agencies with a disparity identified 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[17] and engage with DPSST on any results that show a statistically significant disparity.

[16] In partnership with Portland Central City Task Force, Oregon State Police (OSP) began focused enforcement in October 2023 in Portland on specific offenses related to fentanyl possession and delivery. This has resulted in increased stop volume for OSP in Multnomah County, and in particular an increase in pedestrian stops. The pedestrian stops have a higher search rate than previous trends. https://www.portlandcentralcitytaskforce.com/

[17] 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[18], 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 2022 and 2023. Prior to 2022, this section was published as a separate report. Since 2022, this information has been included as an additional section within the existing STOP report. 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 2022 and 2023 Reports

Table 7.2.1. summarizes law enforcement agency reporting for 2022 and 2023. In 2022, 127 of 154 (82.5 percent) law enforcement agencies reported the number of profiling complaints they received and in 2023, 117 of 154 (76.0 percent) law enforcement agencies reported the number of profiling complaints they received for each respective calendar year. Of those agencies that reported in 2022, 23 (18.1 percent) reported at least one complaint, and across those 23 agencies there were a total of 62 complaints. In 2023, 22 (18.8 percent) agencies that reported had at least one complaint and across those agencies, 75 total complaints were received.

Table 7.2.1. Law Enforcement Annual Reporting Compliance, 2022 and 2023

	2022	2023
Agencies Reporting	127	117
Total Reported Complaints	62	75
Agencies Reporting No Complaints	104	94
Agencies Reporting 1+ Complaints	23	22

Table 7.2.2. shows the number of complaints reported by agency in 2022 and 2023. Across those two years, Oregon State Police had the highest complaint volume with 21 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 Clackamas County SO and Portland PB with 15 reported complaints apiece.

[18] 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. Reported Profiling Complaints by Agency, 2022 and 2023

Department	2022	2023
Albany PD	0	1
Beaverton PD	2	5
Bend PD	2	7
Clackamas CO SO	9	6
Corvallis PD	2	2
Dallas PD	1	0
Deschutes CO SO	0	1
Eagle Point PD	0	1
Eugene PD	7	3
Gresham PD	1	0
Hillsboro PD	1	4
Independence PD	1	0
Jackson CO SO	0	2
Keizer PD	3	1
La Grande PD	1	0
Lake Oswego PD	2	2
Lane CO SO	1	6
Marion CO SO	1	1
Medford PD	3	2
Multnomah CO SO	0	2
Oregon City PD	1	0
Oregon State Police	10	11
OSU PD	2	0
Portland PB	5	10
Redmond PD	0	1
Springfield PD	3	4
The Dalles PD	1	0
Tigard PD	1	1
Washington CO SO	2	2
Total	62	75

Table 7.2.3. shows the dispositions of complaints that were reported in 2022 and 2023. The most common disposition in both years was "unfounded", followed by "not sustained" in 2022 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.

For comparison purposes, a report by the California Racial and Identity Profiling Advisory Board that analyzed data on 10,648 civilian complaints in California in 2020 found that 9.4 percent of all reports were sustained, with the most common disposition for that year being "unfounded" followed by "exonerated."[19]

[19] See https://oag.ca.gov/system/files/media/ripa-board-report-2022.pdf

Table 7.2.3. Reported Profiling Complaints by Disposition, 2022 and 2023

Disposition	2022	2023
Exonerated	2	8
Not Sustained	10	4
Unfounded	33	43
Administrative Closure	3	7
No Basis for Further Investigation	9	11
Other	3	4

The reports received by law enforcement agencies varied greatly in terms of providing details about the incidents being reported on, which made it difficult for CJC researchers to identify trends in the nature of these incidents. This indicates that law enforcement agencies may need further guidance on filling out these forms. In addition, it is difficult to determine what proportion of actual incidents of racial profiling in Oregon these reports represent.

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 County SO	Marion County SO	Portland PB
Eugene PD	Medford PD	Salem PD
Gresham PD	Multnomah County SO	Washington County SO

Table A.2. Tier 2 Agencies

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

Table A.3. **Tier 3 Agencies**

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

*Inactive Agencies

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[20] and pedestrian[21] 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.).

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 1.1. 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.1.1. Three-Tier Reporting Approach in HB 2355 (2017)

Tier	Number of 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

[20] 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).

[21] 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).

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

B.2.1. 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[22]. 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.

---

[22] See STOP Program Research Brief: Analytical Approaches to Studying Stops Data (October 2018), which can be found at Traffic_Stop_Research_Memo_Final_Draft-10-16-18.pdf (oregon.gov).

B.2.2. 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 disparities 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 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[23] 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[24]:

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 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

[23] 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.
[24] More detailed and technical descriptions of these analyses can be found in Appendices E, F, and G.

than those reported for white individuals, then there is evidence of a disparity.

B.2.3. Analytical Sample

B.2.3.1. Data Elements

A total of 582,500 records were submitted by 142 Tier 1, Tier 2, and Tier 3 agencies during the sixth 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[25].

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, Latinx, 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[26], and doesn't account for the additional nuance of multiple races and individuals who are not white and Latinx. 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

---

[25] 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.

[26] 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

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)[27].

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" 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

---

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

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[28]. 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 race of the driver 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 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. Appendix D includes an additional analysis for the Stop Outcomes analysis for agencies that submitted the additional optional variables.

B.2.3.2. Sample

While the overall number of records was substantial, the STOP Program team faced challenges with regard to sample size when the data were broken down into subsamples based on race/ethnicity and agency. Tier 3 agencies have fewer officers than Tier 1 and Tier 2 agencies, and therefore submit a relatively low number of police stops. For example, four Tier 3 agencies made fewer than 100 stops in Year 5. In cases where the sample size is too small, statistical analyses cannot be conducted.

Table B.2.3.2.1. 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[29]
Stop Outcomes	Model convergence[30]
Search Findings	Minimum 30 observations per racial/ethnic group analyzed; no cell with less than 5 observations

To determine appropriate thresholds for sample size, the STOP Program relied on established criteria set in the academic and professional literature. Drawing on standards described by Wilson, Voorhis, and Morgan (2007), the STOP Program used the sample size thresholds in Table B.2.3.2.1.

[28] 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.
[29] 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.
[30] 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.

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. Table 2.1.4., Table 2.1.5., and Table C.1. in Appendix C report the breakdown by race/ethnicity and agency for all Tier 1, Tier 2, and Tier 3 agencies, respectively, for stops occurring from July 1, 2023, through June 30, 2024, the most recent year of data collection.

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 in Table B.2.3.2.1. 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 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 is 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 of these sources were found.

B.2.4. Threshold for Statistical Significance

To determine if disparities 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, an estimated disparity must meet the 95 percent confidence level for it to be statistically significant. This means that the STOP Program research team must be at least 95 percent confident that differences or disparities identified by the analyses were not due to random variation in statistical estimates. In some cases, confidence in the reported results exceeded the 95 percent confidence threshold.

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[31] 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 Table B.2.4.1. Some agencies had too few stops of Asian/PI, Black, Latinx, 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.

[31] 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.2.4.1. **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

Beyond the 95 percent confidence threshold for each individual analysis, STOP Program researchers also established a threshold at which identified disparities 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[32], 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 disparity in at least two of the three analytical tests performed on the STOP data[33]. 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.

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

[33] The State of Connecticut applies a sliding scale in its analyses, whereby a disparity identified via the Veil of Darkness analysis alone results in an agency being identified for further analysis. For its other analyses, two or more identified disparities results 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 – Year 6**

Agency	Asian/PI	Black	Latinx	Middle Eastern	Native	White	Total
Astoria PD	37	39	217	3	0	1,909	2,205
Aumsville PD	11	12	171	6	0	640	840
Baker CO SO	12	11	58	7	0	946	1,034
Baker City PD	3	6	30	4	1	496	540
Bandon PD	1	1	0	0	0	21	23
Black Butte Ranch PD	8	5	26	4	0	253	296
Boardman PD	2	9	326	0	2	176	515
Brookings PD	54	21	186	11	7	1,557	1,836
Burns PD	15	20	40	8	0	399	482
Cannon Beach PD	69	30	206	29	1	1,591	1,926
Carlton PD	8	5	38	3	0	240	294
Clatsop CO SO	56	49	242	14	1	2,272	2,634
Coburg PD	18	24	100	18	0	609	769
Columbia CO SO	30	34	107	11	1	1,711	1,894
Columbia City PD	3	4	18	3	1	169	198
Coos Bay PD	40	44	181	7	5	3,699	3,976
Coos CO SO	7	1	37	1	2	365	413
Coquille PD	3	4	24	1	2	409	443
Cottage Grove PD	18	7	61	1	0	530	617
Crook CO SO	24	19	167	6	1	1,523	1,740
Curry CO SO	15	6	21	6	0	366	414
Dallas PD	32	24	160	8	0	996	1,220
Eagle Point PD	29	25	231	7	0	1,602	1,894
Enterprise PD	0	1	1	1	0	7	10
Florence PD	2	6	12	1	0	475	496
Gearhart PD	6	4	35	3	0	233	281
Gervais PD	1	1	3	0	0	12	17
Gilliam CO SO	50	79	368	11	5	2,108	2,621
Gladstone PD	93	118	310	42	9	1,852	2,424
Gold Beach PD	11	6	24	18	1	331	391
Harney CO SO	8	7	20	3	2	165	205
Hood River PD	39	29	399	13	7	1,024	1,511
Hubbard PD	11	8	383	1	0	281	684
Independence PD	38	50	442	10	3	989	1,532
Jefferson CO SO	43	15	181	5	2	967	1,213
Josephine CO SO	13	24	99	4	0	959	1,099
Junction City PD	8	10	21	4	0	240	283
La Grande PD	37	10	26	4	0	454	531
Madras PD	8	5	26	3	2	70	114
Malheur CO SO	0	4	48	0	0	113	165

(Table C.1. continued on next page)

Malin PD	4	11	77	2	0	131	225
Manzanita PD	17	6	17	5	0	190	235
Milton-Freewater PD	10	10	230	3	1	531	785
Molalla PD	21	20	189	9	5	1,197	1,441
Monmouth PD	63	58	275	14	1	920	1,331
Morrow CO SO	9	10	566	3	10	1,171	1,769
Mt. Angel PD	8	5	69	2	0	126	210
Myrtle Creek PD	10	15	39	9	0	959	1,032
Myrtle Point PD	1	1	4	1	0	36	43
Newport PD	19	11	146	4	1	706	887
Nyssa PD	0	1	32	0	0	50	83
OSU PD	69	52	83	17	9	502	732
Oakridge PD	14	1	11	7	0	117	150
Ontario PD	3	3	127	1	1	173	308
PSU CPS	0	0	0	0	0	1	1
Pendleton PD	46	50	151	6	47	1,215	1,515
Philomath PD	48	29	157	19	5	1,101	1,359
Phoenix PD	22	55	254	9	0	1,115	1,455
Pilot Rock PD	16	2	25	1	0	284	328
Port Orford PD	14	3	13	1	0	125	156
Powers PD	3	0	1	0	0	56	60
Prineville PD	1	4	46	2	0	588	641
Rainier PD	4	5	17	1	0	207	234
Reedsport PD	0	0	5	0	0	22	27
Rogue River PD	6	3	22	2	0	195	228
Sandy PD	73	75	262	20	23	2,151	2,604
Seaside PD	62	69	333	22	4	2,178	2,668
Sherman CO SO	64	33	183	17	0	665	962
Sherwood PD	192	145	692	68	18	4,092	5,207
Silverton PD	9	28	315	3	1	1,281	1,637
Stanfield PD	8	21	214	8	8	494	753
Stayton PD	13	11	98	7	0	630	759
Sunriver PD	28	11	107	14	0	1,339	1,499
Sutherlin PD	24	29	96	18	0	1,019	1,186
Sweet Home PD	5	7	18	0	2	528	560
Talent PD	32	35	146	7	0	909	1,129
The Dalles PD	9	10	106	5	6	312	448
Tillamook CO SO	30	9	123	12	0	828	1,002
Tillamook PD	70	23	250	20	7	1,502	1,872
Toledo PD	7	10	63	1	7	634	722
Turner PD	7	9	58	3	1	300	378
Umatilla CO SO	7	20	386	13	8	801	1,235
Umatilla PD	18	48	1022	3	24	999	2,114
Union CO SO	13	15	42	11	2	248	331
Vernonia PD	2	2	5	0	0	122	131

(Table C.1. continued on next page)

Wallowa CO SO	1	0	10	1	1	139	152
Warrenton PD	35	35	139	2	0	1,559	1,770
Wasco CO SO	7	12	107	3	11	389	529
Wheeler CO SO	6	3	22	2	0	204	237
Winston PD	14	15	42	3	0	959	1,033
Yamhill PD	23	18	130	12	0	532	715
Total Tier 3	**2,030**	**1,825**	**12,570**	**674**	**258**	**69,291**	**86,648**

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

Agency	Traffic		Pedestrian		Total
Astoria PD	2,202	99.9%	3	0.1%	2,205
Aumsville PD	840	100.0%	0	0.0%	840
Baker CO SO	1,034	100.0%	0	0.0%	1,034
Baker City PD	540	100.0%	0	0.0%	540
Bandon PD	23	100.0%	0	0.0%	23
Black Butte Ranch PD	296	100.0%	0	0.0%	296
Boardman PD	544	99.8%	1	0.2%	545
Brookings PD	1,836	100.0%	0	0.0%	1,836
Burns PD	482	100.0%	0	0.0%	482
Cannon Beach PD	1,924	99.9%	2	0.1%	1,926
Carlton PD	288	98.0%	6	2.0%	294
Clatsop CO SO	2,634	100.0%	0	0.0%	2,634
Coburg PD	768	99.9%	1	0.1%	769
Columbia CO SO	1,887	99.6%	7	0.4%	1,894
Columbia City PD	198	100.0%	0	0.0%	198
Coos Bay PD	3,976	100.0%	0	0.0%	3,976
Coos CO SO	409	99.0%	4	1.0%	413
Coquille PD	443	100.0%	0	0.0%	443
Cottage Grove PD	615	99.7%	2	0.3%	617
Crook CO SO	1,739	99.9%	1	0.1%	1,740
Curry CO SO	413	99.8%	1	0.2%	414
Dallas PD	1,220	100.0%	0	0.0%	1,220
Eagle Point PD	1,892	99.9%	2	0.1%	1,894
Enterprise PD	10	100.0%	0	0.0%	10
Florence PD	496	100.0%	0	0.0%	496
Gearhart PD	280	99.6%	1	0.4%	281
Gervais PD	17	100.0%	0	0.0%	17
Gilliam CO SO	2,989	99.9%	4	0.1%	2,993
Gladstone PD	2,403	99.1%	21	0.9%	2,424
Gold Beach PD	391	100.0%	0	0.0%	391
Harney CO SO	205	100.0%	0	0.0%	205
Hood River PD	1,511	100.0%	0	0.0%	1,511
Hubbard PD	677	99.0%	7	1.0%	684

(Table C.2. continued on next page)

Independence PD	1,529	99.8%	3	0.2%	1,532
Jefferson CO SO	1,210	99.8%	3	0.3%	1,213
Josephine CO SO	1,099	100.0%	0	0.0%	1,099
Junction City PD	283	100.0%	0	0.0%	283
La Grande PD	531	100.0%	0	0.0%	531
Madras PD	114	100.0%	0	0.0%	114
Malheur CO SO	165	100.0%	0	0.0%	165
Malin PD	225	100.0%	0	0.0%	225
Manzanita PD	235	100.0%	0	0.0%	235
Milton-Freewater PD	780	99.4%	5	0.6%	785
Molalla PD	1,435	99.6%	6	0.4%	1,441
Monmouth PD	1,330	99.9%	1	0.1%	1,331
Morrow CO SO	2,007	99.7%	7	0.4%	2,014
Mt. Angel PD	210	100.0%	0	0.0%	210
Myrtle Creek PD	1,027	99.5%	5	0.5%	1,032
Myrtle Point PD	43	100.0%	0	0.0%	43
Newport PD	923	99.9%	1	0.1%	924
Nyssa PD	83	100.0%	0	0.0%	83
OSU PD	709	96.9%	23	3.1%	732
Oakridge PD	150	100.0%	0	0.0%	150
Ontario PD	308	100.0%	0	0.0%	308
PSU CPS	1	100.0%	0	0.0%	1
Pendleton PD	1,467	96.8%	48	3.2%	1,515
Philomath PD	1,357	99.9%	2	0.2%	1,359
Phoenix PD	1,454	99.9%	1	0.1%	1,455
Pilot Rock PD	328	100.0%	0	0.0%	328
Port Orford PD	156	100.0%	0	0.0%	156
Powers PD	60	100.0%	0	0.0%	60
Prineville PD	651	100.0%	0	0.0%	651
Rainier PD	234	100.0%	0	0.0%	234
Reedsport PD	27	100.0%	0	0.0%	27
Rogue River PD	228	100.0%	0	0.0%	228
Sandy PD	2,595	99.7%	9	0.4%	2,604
Seaside PD	2,665	99.9%	3	0.1%	2,668
Sherman CO SO	962	100.0%	0	0.0%	962
Sherwood PD	5,180	99.5%	27	0.5%	5,207
Silverton PD	1,568	95.8%	69	4.2%	1,637
Stanfield PD	886	99.9%	1	0.1%	887
Stayton PD	759	100.0%	0	0.0%	759
Sunriver PD	1,497	99.9%	2	0.1%	1,499
Sutherlin PD	1,185	99.9%	1	0.1%	1,186
Sweet Home PD	560	100.0%	0	0.0%	560
Talent PD	1,114	98.7%	15	1.3%	1,129
The Dalles PD	445	99.3%	3	0.7%	448
Tillamook CO SO	1,002	100.0%	0	0.0%	1,002

(Table C.2. continued on next page)

Tillamook PD	1,872	100.0%	0	0.0%	1,872
Toledo PD	722	100.0%	0	0.0%	722
Turner PD	378	100.0%	0	0.0%	378
Umatilla CO SO	1,262	99.5%	7	0.6%	1,269
Umatilla PD	2,526	99.8%	5	0.2%	2,531
Union CO SO	331	100.0%	0	0.0%	331
Vernonia PD	131	100.0%	0	0.0%	131
Wallowa CO SO	152	100.0%	0	0.0%	152
Warrenton PD	1,770	100.0%	0	0.0%	1,770
Wasco CO SO	521	98.5%	8	1.5%	529
Wheeler CO SO	237	100.0%	0	0.0%	237
Winston PD	1,032	99.9%	1	0.1%	1,033
Yamhill PD	715	100.0%	0	0.0%	715
Tier 3 Total	**87,608**	**99.6%**	**319**	**0.4%**	**87,927**

Appendix D – Decision to Stop Analysis Technical Appendix and Detailed Results

The Decision to Stop (DTS) analysis, first developed by Grogger and Ridgeway (2006) as the Veil of Darkness analysis, analyzes stop data for racial/ethnic disparities and 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 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.[34] 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 combined inter-twilight window is created from the Oregon traffic stop data from July 1, 2021, to June 30, 2023. 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

[34] 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 spline with three degrees of freedom within each twilight window. Alternative time of day controls were tested and did not change the results.

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.

Table D.1 displays the odds ratios for the Tier 1 and Tier 2 DTS models with at least two comparisons for all non-white stopped drivers, including those perceived as Black, Latinx, Asian/PI, Middle Eastern, and Native American, compared to white stopped drivers.

Table D.1 Decision to Stop Analyses by Tier 1 or Tier 2 Agency

Agency	Asian/PI	Black	Latinx	Middle Eastern	Native American
Albany PD	--	1.99	1.15	--	--
Beaverton PD	0.75	1.00	1.08	1.00	--
Benton CO SO	--	0.08	1.09	--	--
Clackamas CO SO	0.89	1.35	1.05	0.80	--
Corvallis PD	0.96	1.17	1.81**	--	--
Eugene PD	0.70	0.86	0.89	--	--
Forest Grove PD	0.69	0.69	0.99	--	--
Gresham PD	--	1.47	1.09	--	--
Hillsboro PD	1.13	0.85	1.03	1.45	--
Lake Oswego PD	0.72	1.14	1.13	--	--
Lane CO SO	--	1.17	0.61	--	--
Marion CO SO	1.19	1.35	1.10	--	--
Medford PD	--	1.02	0.92	--	--
Milwaukie PD	1.06	2.17*	1.30	--	--
Multnomah CO SO	1.49	1.25	1.00	--	--
Oregon City PD	--	0.82	0.78	--	--
Oregon State Police	1.06	1.11	1.00	1.22	1.29
Portland PB	1.23	1.07	1.12	1.12	--
Springfield PD	--	0.81	0.66	--	--
Tigard PD	1.04	1.41	1.08	1.30	--
Washington CO SO	0.91	1.08	1.00	0.98	--
Yamhill CO SO	--	1.48	0.92	--	--

* p<0.05, ** p<0.01, *** p<0.001

Table D.2 reports the Tier 1 and Tier 2 agency specific model results for Latinx drivers compared to white drivers for agencies not displayed above.

Table D.2. Decision to Stop Analyses for Latinx Drivers by Tier 1 or Tier 2 Agency

Agency	Latinx	Agency	Latinx
Bend PD	1.06	Lincoln CO SO	1.16
Canby PD	1.91**	Linn CO SO	0.85
Central Point PD	1.25	McMinnville PD	1.26
Deschutes CO SO	1.48*	Newberg-Dundee PD	0.92
Grants Pass PD	0.75	Polk CO SO	1.07
Hermiston PD	0.87	Redmond PD	0.87
Hood River CO SO	0.55	Roseburg PD	1.25
Jackson CO SO	1.11	Salem PD	0.84
Keizer PD	0.69	Tualatin PD	0.77
Klamath Falls PD	0.87	West Linn PD	0.92
Lincoln City PD	1.44	Woodburn PD	0.72

* p<0.05, ** p<0.01, *** p<0.001

Table D.3 reports the Tier 3 agency specific model results for Latinx drivers compared to white drivers for agencies with sufficient sample size.

Table D.3. Decision to Stop analysis for Latinx Drivers by Tier 3 Agency

Agency	Latinx	Agency	Latinx
Astoria PD	0.87	Molalla PD	0.70
Aumsville PD	1.16	Monmouth PD	1.47
Boardman PD	0.87	Morrow CO SO	1.10
Brookings PD	1.41	Newport PD	0.55
Cannon Beach PD	0.96	Phoenix PD	0.92
Clatsop CO SO	0.78	Sandy PD	1.75*
Coos Bay PD	1.26	Seaside PD	0.90
Crook CO SO	1.39	Sherman CO SO	1.15
Dallas PD	0.83	Sherwood PD	0.85
Eagle Point PD	0.61	Silverton PD	0.86
Gilliam CO SO	0.99	Stanfield PD	1.54
Gladstone PD	1.01	Talent PD	1.44
Hood River PD	0.75	The Dalles PD	0.96
Hubbard PD	1.07	Tillamook CO SO	1.27
Independence PD	0.94	Tillamook PD	1.00
Jefferson CO SO	1.07	Umatilla CO SO	0.79
Josephine CO SO	1.86	Umatilla PD	1.12
Milton-Freewater PD	0.53	Warrenton PD	1.16

* p<0.05, ** p<0.01, *** p<0.001

Appendix E – Stop Outcomes Analysis Technical Appendix and Detailed Results

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.[35]
    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.[36]

[35] 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 (Austin and Stuart 2015).

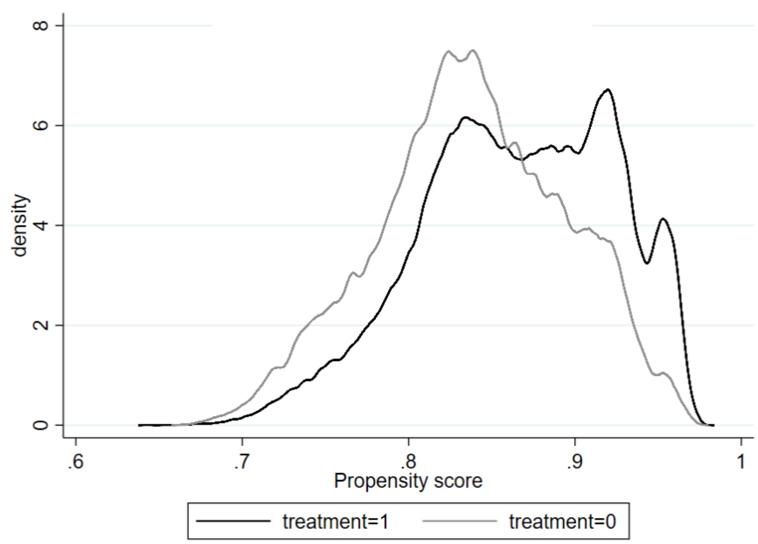
[36] For a thorough discussion of IPWRA methods, see Wooldridge 2010, Chapter 21.3.4.

Assumptions

There are a few assumptions that must hold in order for propensity score estimators to be unbiased. The first is the conditional independence assumption[37], 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. In order for this assumption to hold, changes in any unobserved variables that have an effect on the outcome variable must not also have an effect on 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 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.[38]

[37] This assumption is also referred to as the unconfoundedness assumption.
[38] Omitted treatment variables per analysis are not included in this report due to the high number of analyses conducted.

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 Latinx individual causes another Latinx 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 violated in this case as the race of a stopped individual does not plausibly impact the race of subsequently stopped individuals.[39]

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 Latinx, = 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 weighs 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 $= \frac{Total\ \#\ of\ stops\ by\ agency\ on\ day\ of\ stop}{Maximum\ \#\ of\ daily\ stops\ in\ the\ \boldsymbol{county}\ over\ year\ of\ analysis}$

For the additional analysis, one further variable is included:

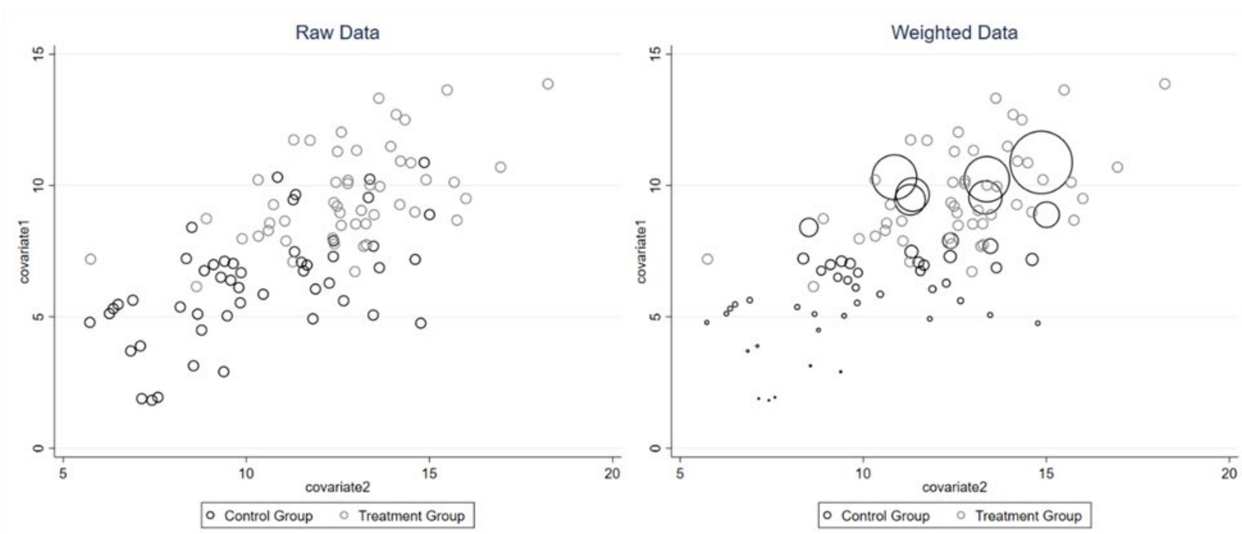
10. If the stop outcome is caused by a low-discretion violation = 1, otherwise = 0

[39] The Stata handbook provides a good description of these assumptions, and the counterfactual model that underlies all matching methods. ("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 is in Figure E.2. below. 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, as a whole, has observations that are much closer to those of the treatment group than the raw control group.

Figure E.2. Weighting Example



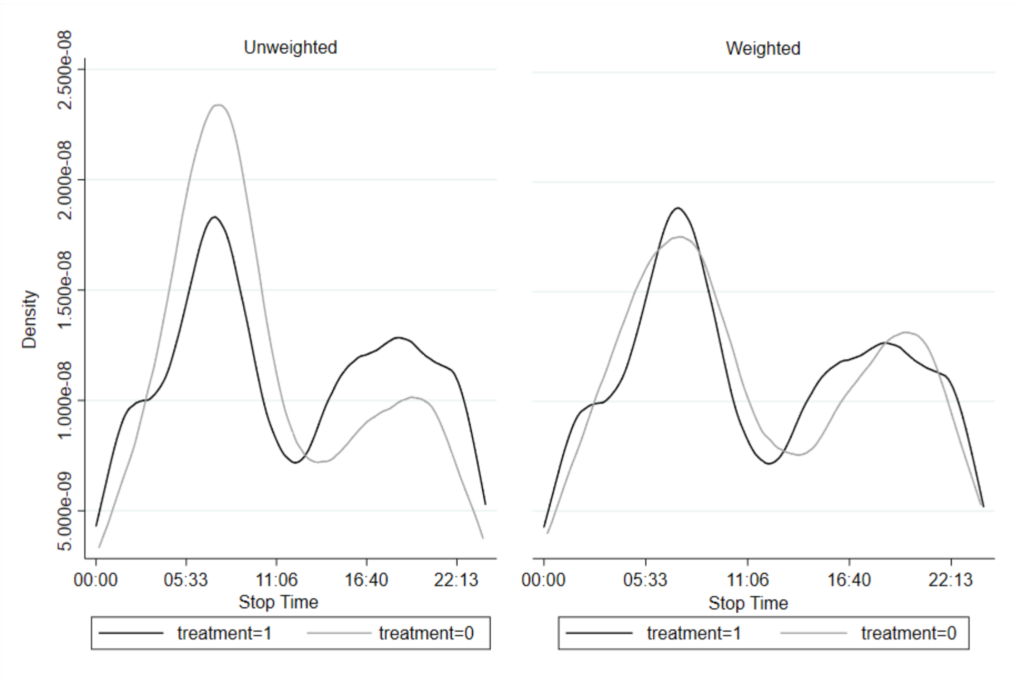
Balance is then measured based on the standardized difference[40] in means and the variance ratio[41] between the treatment and control groups for each of the raw data set and the inverse probability weighted data set. If the resulting standardized difference in the weighted data set is close to zero 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

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

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

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



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.[42]

[42] 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.

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 disparity estimates built on top of these driving population estimates.

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 Analysis 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, the STOP Program researchers expect officers 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 to carry contraband in order to reduce their risk of being caught. In this way, any 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.

The Search Findings analysis assesses whether police are participating in racial/ethnic discrimination by over searching members of a particular group. 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, if a minority racial/ethnic group is "over-searched," then the hit-rate for that group will be lower than that of white individuals, 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, Latinx, 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.

Detailed Results

Table F.1. Search Findings Analysis Detailed Results

Agency	White	Black	Latinx	Asian/PI	Native	Middle Eastern
Albany PD	43.6%	---	45.5%	---	---	---
Beaverton PD	49.7%	45.5%	45.6%	---	---	---
Bend PD	17.3%	---	22.2%	---	---	---
Clackamas CO SO	27.0%	---	48.9%	---	---	---
Eugene PD	35.6%	33.8%	36.6%	---	---	---
Gresham PD	45.5%	---	40.0%	---	---	---
Hillsboro PD	51.2%	---	46.7%	---	---	---
Hubbard PD	41.7%	---	47.5%	---	---	---
Marion CO SO	10.4%	---	6.8%	---	---	---
Multnomah CO SO	37.7%	34.9%	43.6%	---	---	---
Oregon State Police	61.0%	75.4%	54.9%**	64.5%	43.2%	---
Portland PB	39.9%	51.8%	39.3%	41.5%	---	---
Salem PD	33.8%	---	37.8%	---	---	---
Silverton PD	48.2%	---	47.6%	---	---	---
Springfield PD	55.3%	57.1%	52.2%	---	---	---
Washington CO SO	42.2%	---	51.5%	---	---	---

* p<0.05, ** p<0.01, *** p<0.001