Statistical Transparency of Policing Report

Per House Bill 2355 (2017)

December 1, 2022

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

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, so the Commission 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 Commission, was created along with the Oregon State Police and the Oregon Department of Public Safety Standards and Training (DPSST). This is the fourth annual report to the Oregon Legislature by the STOP Program examining data received pursuant to HB 2355.

Table E.1. Descriptive Statistics for Aggregate Year 4 Stop Data

Variable	Tier 1	Tier 2	Tier 3
Traffic Stop	98.1%	97.6%	94.4%
Race/Ethnicity			
Asian/PI	3.4%	2.7%	2.0%
Black	5.2%	3.2%	2.0%
Latinx	15.6%	14.1%	12.8%
Middle Eastern	1.4%	0.9%	0.8%
Native American	0.5%	0.3%	0.5%
White	73.3%	78.7%	80.1%
Gender			
Male	66.9%	63.6%	59.0%
Female	32.9%	36.2%	31.4%
Nonbinary	0.1%	0.2%	8.8%
Age			
Under 21	10.9%	12.8%	12.3%
21-29	24.2%	22.6%	21.8%
30-39	25.8%	24.8%	23.7%
40-49	16.9%	17.4%	17.1%
50 and Older	22.2%	22.4%	25.1%
Stop Disposition			
None	2.8%	7.6%	5.0%
Warning	56.6%	59.8%	69.4%
Citation	37.8%	30.1%	23.9%
Juvenile Summons	0.0%	0.0%	0.0%
Arrest	2.7%	1.9%	1.3%
Search Conducted	2.2%	1.5%	0.9%

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 December 2022, the STOP Program has received at least two full years of data from 148 law enforcement agencies in the state and analyses using those data are presented in this report. This is the first STOP report to include two years of data from all Tier 1, Tier 2, and Tier 3 agencies.

Table E.1. reports descriptive statistics for the combined Tier 1, Tier 2, and Tier 3 data, which represents stops made from July 1, 2021 through June 30, 2022. Across all agencies, the vast majority of the reported data were for traffic stops, although the share of pedestrian stops made by Tier 3 agencies was higher than that for their larger counterparts. 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 Oregonians. Once the stop had been initiated, stopped individuals either were subject to no further action or merely given a warning in a little under 60 percent of stops for Tier 1, a little under 70 percent of stops for Tier 2, and a little under 75 percent of stops for Tier 3. Tier 3 made more stops of nonbinary individuals than Tier 1 and Tier 2, which is attributable to five Tier 3 agencies that reported a high proportion of stops for nonbinary individuals that stems from a data entry issue.

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.

No agency was identified as having a statistically significant disparity in two or more tests performed on the STOP data this year. Therefore, no agency is referred to receive technical assistance from DPSST in this report. However, that does not mean that the results for any agencies should be ignored or are not close to the threshold of identification. Regardless of whether an agency is officially referred to DPSST, the CJC urges each agency to scrutinize their full set of results.

Tables and Figures

1. Background

This is the fourth annual report from the Statistical Transparency of Policing (STOP) Program. 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 the treatment of Oregonians by law enforcement. To implement this mandate, the Legislature required the largest agencies to collect data first, followed by medium and smaller agencies in the intervening years. In December 2019, the Criminal Justice Commission published its first annual STOP report, which contained data and analyses for the 12 largest law enforcement agencies in the state. In December 2020, the Criminal Justice Commission published its second report, which included an additional 39 mid-sized police agencies. The third annual STOP report, released in December 2021, included all agencies in the state. This report builds on the first three by including analyses that incorporate two years of data for all sized agencies. The inclusion of two full years of data means that this report analyzes stops from 148 law enforcement agencies in the state[1].

1.1. 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 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 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 Oregon Statistical Transparency of Policing (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[2] and pedestrian[3] 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 Oregon Criminal Justice Commission, in partnership with the Oregon State Police and the Department of Justice, worked to develop a standardized method for collecting the data elements required

[1] For a full list of agencies see Appendix A, and for reporting rates by agency see Appendix B – Data Audit.
[2] 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).
[3] 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).

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 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

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 solution 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 in 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.

2. Methodological Approach

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 (1993) 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[4]. 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.

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-citizen 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[5] 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[6]:

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 versus 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

---

[4] 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)](oregon.gov).

[5] 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.

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

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

2.3. Analytical Sample

2.3.1. Data Elements

A total of 453,476 records were submitted by 148 Tier 1, Tier 2, and Tier 3 agencies during the fourth 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[7].

Perceived Race/Ethnicity of Subject. Law enforcement officers are required by HB 2355 to record their perception of a subject's race/ethnicity (for traffic stops, only the perceived race/ethnicity of the driver is reported). 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[8], and doesn't account for the additional nuance of multiple races and individuals who are not white and Latinx. However, in an effort to simplify the data collection process

---

[7] For instance, 130 Year 4 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.

[8] https://www.census.gov/topics/population/race/about.html and https://www.census.gov/topics/population/hispanic-origin/about.html

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 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 is reported).

Legal Basis for the Stop. The legal basis for each stop is reported to the STOP Program. This includes violations of an Oregon statute, a municipal traffic code, a municipal criminal code, a county code, tri-met 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)[9].

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 4 data included only 121 juvenile summons (0.03 percent of all dispositions).

Whether a Search was Conducted. Law enforcement officers utilize a binary variable to report whether a search was conducted to 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.

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

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 X,Y coordinates, while others submit textual descriptions of the location (e.g., 123 Main Street, intersection of Main and Maple Streets).

The STOP Program created four of its own variables for use in its analyses. Following best practices, variables representing both the daily agency stop volume and daily county stop volume were created. For agency stop volume, the aggregate number of stops for a single date are divided by the maximum number of daily stops for the agency unit in question. Thus, if an agency stopped 1,000 drivers on its busiest day, this would be the denominator against which all other days would be compared. A measure of the county stop volume would be calculated the same way, although all stops made by agencies within a single county would be included together. Additionally, variables representing sunrise time and sunset time were made for use in the Decision to Stop and Stop Outcomes analyses[10]. 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.

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.

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, seven Tier 3 agencies made fewer than 100 stops in Year 4. In cases where the sample size is too small, statistical analyses cannot be conducted.

---

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

Table 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[11]
Stop Outcomes	Model convergence[12]
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 2.3.2.1.

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 2.3.2.2.a., Table 2.3.2.2.b., 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, 2021, through June 30, 2022, 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 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.

Table 2.3.2.2.a. Race/Ethnicity Reporting for Tier 1 Agencies for All Reported Stops

Agency Name	Asian/PI	Black	Latinx	Middle Eastern	Native American	White	Total
Beaverton PD	724	1,156	2,466	341	54	8,265	13,006
Clackamas CO SO	721	958	2,354	252	178	13,552	18,015
Eugene PD	278	758	919	0	0	9,609	11,564
Gresham PD	94	312	450	30	7	1,358	2,251
Hillsboro PD	365	366	1,748	160	25	3,698	6,362
Marion CO SO	421	401	2,616	146	12	10,142	13,738
Medford PD	55	154	696	17	5	3,284	4,211
Multnomah CO SO	236	762	1,016	94	27	4,464	6,599
Oregon State Police	3,316	4,129	18,006	1,669	746	100,782	128,648
Portland PB	674	2,324	1,564	186	74	8,611	13,433
Salem PD	123	191	1,272	29	22	3,123	4,760
Washington CO SO	1,304	1,163	5,000	618	136	12,578	20,799
Total Tier 1	8,311	12,674	38,107	3,542	1,286	179,466	243,386

[11] 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.
[12] All possible racial group and stop outcome models are estimated in Stata. Models that did not converge are not included in the results.

Table 2.3.2.2.b. Race/Ethnicity Reporting for Tier 2 Agencies for All Reported Stops

Agency Name	Asian/PI	Black	Latinx	Middle Eastern	Native American	White	Total
Albany PD	65	96	424	16	8	3,154	3,763
Ashland PD	48	48	82	13	0	873	1,064
Bend PD	70	92	371	27	11	4,175	4,746
Benton CO SO	189	125	374	35	11	3,912	4,646
Canby PD	42	34	538	15	4	1,865	2,498
Central Point PD	34	55	245	8	2	1,734	2,078
Corvallis PD	331	290	581	150	13	5,260	6,625
Deschutes CO SO	38	25	185	8	0	1,688	1,944
Douglas CO SO	15	17	47	6	0	741	826
Forest Grove PD	139	121	1,414	39	4	3,152	4,849
Grants Pass PD	43	48	177	10	2	2,222	2,502
Hermiston PD	53	105	2,293	15	44	3,046	5,556
Hood River CO SO	45	23	416	14	0	1,024	1,522
Jackson CO SO	68	110	586	19	0	3,377	4,160
Keizer PD	53	68	557	9	0	1,396	2,083
Klamath CO SO	13	6	39	4	2	254	318
Klamath Falls PD	162	64	383	11	7	2,098	2,725
Lake Oswego PD	243	248	444	106	55	4,450	5,546
Lane CO SO	59	92	188	21	6	2,577	2,943
Lebanon PD	14	18	42	1	0	918	993
Lincoln City PD	60	45	175	18	0	1,148	1,446
Lincoln CO SO	70	42	197	11	6	1,687	2,013
Linn CO SO	52	71	328	16	10	3,906	4,383
McMinnville PD	32	39	385	5	1	1,560	2,022
Milwaukie PD	109	283	281	44	6	2,612	3,335
Newberg-Dundee PD	126	82	625	35	0	3,804	4,672
OHSU PD	5	13	5	2	0	43	68
Oregon City PD	108	186	515	46	37	4,272	5,164
Polk CO SO	121	119	714	32	11	2,553	3,550
Port of Portland PD	63	146	113	29	8	785	1,144
Redmond PD	38	28	248	4	0	1,462	1,780
Roseburg PD	25	58	210	18	12	3,681	4,004
Springfield PD	57	298	491	2	0	5,176	6,024
Tigard PD	232	247	594	125	25	2,713	3,936
Tualatin PD	194	194	782	64	7	3,436	4,677
UO PD	8	16	13	4	1	215	257
West Linn PD	81	101	252	46	32	1,749	2,261
Woodburn PD	3	7	462	0	1	261	734
Yamhill CO SO	129	105	927	41	14	3,988	5,204
Total Tier 2	**3,237**	**3,765**	**16,703**	**1,069**	**340**	**92,967**	**118,081**

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.

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 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 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.

Table 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[13], 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[14]. 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.

2.5. 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

[13] The Connecticut Racial Profiling Prohibition Project is located at http://www.ctrp3.org/.
[14] 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.

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. Characteristics of Year 4 Stop Data

3.1. General Characteristics

Table 3.1.1.a. Number and Percent of Tier 1 Agency Stops by Stop Type Traffic vs. Pedestrian

Agency Name	Traffic		Pedestrian		Total
Beaverton PD	12,557	96.5%	449	3.5%	13,006
Clackamas CO SO	17,457	96.9%	558	3.1%	18,015
Eugene PD	10,095	86.6%	1,563	13.4%	11,658
Gresham PD	2,248	99.9%	3	0.1%	2,251
Hillsboro PD	6,193	97.3%	169	2.7%	6,362
Marion CO SO	13,668	99.5%	72	0.5%	13,740
Medford PD	3,583	85.1%	628	14.9%	4,211
Multnomah CO SO	6,409	97.1%	190	2.9%	6,599
Oregon State Police	129,615	99.6%	501	0.4%	130,116
Portland PB	13,365	99.5%	68	0.5%	13,433
Salem PD	4,512	94.8%	248	5.2%	4,760
Washington CO SO	20,706	99.6%	93	0.4%	20,799
Tier 1 Total	**240,408**	**98.1%**	**4,542**	**1.9%**	**244,950**

While the analyses contained in Sections 4., 5., and 6. 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, 2021, through June 30, 2022. In total, 453,476[15] stops were submitted to the STOP Program by 148 Tier 1, Tier 2, and Tier 3 agencies during Year 4. The number of stops reported by each agency is displayed in Table 3.1.1.a., and Table 3.1.1.b., 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 in terms of officers employed. The Oregon State Police, which is the state's largest law enforcement agency, made 130,116 stops in Year 4, the largest number reported by any one agency and accounting for over a quarter of all stops in the state. At the other end of the continuum, Merrill PD made the fewest stops, totaling 34, accounting for less than 0.008 percent of the reported stops in Year 4.

[15] 1038, or 0.23% of these 453,476 stops were not definitively identified as either a pedestrian or traffic stop, and were therefore excluded from Table 3.1.1.a, Table 3.1.1.b., and Table C.2. therefore, stop totals in these tables do not add up to exactly 453,476.

Tables 3.1.1.a. above and 3.1.1.b. below 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 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 3 agencies had the highest proportion of pedestrian stops, 4.5 percent, compared to Tier 1's 1.9 percent, and Tier 2's 2.4 percent. This is likely due to the presence of agencies which are small and do not patrol highways or streets. For instance, one Tier 3 agency, Union Pacific Railroad, reported 100 percent pedestrian stops. Of Tier 1 agencies, Eugene PD and Medford PD made the highest proportion of pedestrian stops, echoing past reports. Of Tier 2 agencies, UO PD had the highest proportion of pedestrian stops; just under half of their stops were of pedestrians.

Table 3.1.1.b. Number and Percent of Tier 2 Agency Stops by Stop Type Traffic vs. Pedestrian

Agency Name	Traffic		Pedestrian		Total
Albany PD	3,684	97.9%	79	2.1%	3,763
Ashland PD	960	90.2%	104	9.8%	1,064
Bend PD	4,721	99.5%	25	0.5%	4,746
Benton CO SO	4,627	99.6%	19	0.4%	4,646
Canby PD	2,488	99.6%	10	0.4%	2,498
Central Point PD	2,025	97.4%	53	2.6%	2,078
Corvallis PD	6,545	98.8%	80	1.2%	6,625
Deschutes CO SO	1,944	100.0%	0	0.0%	1,944
Douglas CO SO	804	96.9%	26	3.1%	830
Forest Grove PD	4,839	99.4%	30	0.6%	4,869
Grants Pass PD	2,264	90.5%	238	9.5%	2,502
Hermiston PD	5,461	98.3%	95	1.7%	5,556
Hood River CO SO	1,522	100.0%	0	0.0%	1,522
Jackson CO SO	4,160	100.0%	0	0.0%	4,160
Keizer PD	2,083	100.0%	0	0.0%	2,083
Klamath CO SO	309	97.2%	9	2.8%	318
Klamath Falls PD	2,725	100.0%	0	0.0%	2,725
Lake Oswego PD	5,508	99.3%	38	0.7%	5,546
Lane CO SO	2,899	98.0%	60	2.0%	2,959
Lebanon PD	993	100.0%	0	0.0%	993
Lincoln City PD	1,423	98.4%	23	1.6%	1,446
Lincoln CO SO	1,998	99.8%	5	0.2%	2,003
Linn CO SO	4,280	97.7%	103	2.4%	4,383
McMinnville PD	1,860	92.0%	162	8.0%	2,022
Milwaukie PD	3,199	95.9%	136	4.1%	3,335
Newberg-Dundee PD	4,512	96.6%	160	3.4%	4,672
OHSU PD	68	100.0%	0	0.0%	68
Oregon City PD	4,818	93.3%	346	6.7%	5,164
Polk CO SO	3,525	99.3%	25	0.7%	3,550
Port of Portland PD	992	86.6%	153	13.4%	1,145
Redmond PD	1,764	99.1%	16	0.9%	1,780
Roseburg PD	3,823	95.5%	181	4.5%	4,004
Springfield PD	5,827	95.8%	258	4.2%	6,085
Tigard PD	3,747	95.2%	189	4.8%	3,936
Tualatin PD	4,604	98.4%	73	1.6%	4,677
UO PD	134	52.1%	123	47.9%	257
West Linn PD	2,247	99.4%	14	0.6%	2,261
Woodburn PD	728	99.2%	6	0.8%	734
Yamhill CO SO	5,199	99.9%	5	0.1%	5,204
Total Tier 2	**115,309**	**97.6%**	**2,844**	**2.4%**	**118,153**

The demographic breakdowns for traffic and pedestrian stops are reported in Table 3.1.2. 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. Tier 3 agencies stopped a higher proportion of older individuals than other Tiers.

Table 3.1.2. Aggregate 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	3.4%	1.7%	3.4%	2.8%	1.3%	2.7%	2.1%	1.2%	2.0%
Black	5.2%	4.8%	5.2%	3.2%	3.8%	3.2%	2.0%	1.6%	2.0%
Latinx	15.7%	9.8%	15.6%	14.3%	7.6%	14.1%	13.1%	6.8%	12.8%
Middle Eastern	1.5%	0.4%	1.4%	0.9%	0.5%	0.9%	0.8%	0.3%	0.8%
Native American	0.5%	0.5%	0.5%	0.3%	0.4%	0.3%	0.5%	1.1%	0.5%
White	73.1%	82.4%	73.3%	78.5%	86.4%	78.7%	79.8%	87.7%	80.1%
Gender									
Male	66.6%	83.0%	66.9%	63.3%	78.8%	63.6%	59.0%	61.3%	59.0%
Female	33.2%	16.7%	32.9%	36.5%	21.1%	36.2%	31.4%	31.7%	31.4%
Nonbinary*	0.1%	0.1%	0.1%	0.2%	0.1%	0.2%	9.0%	6.8%	8.8%
Age									
Under 21	11.0%	7.5%	10.9%	12.9%	8.1%	12.8%	22.0%	18.8%	12.3%
21-29	24.3%	20.7%	24.2%	22.8%	17.2%	22.6%	23.6%	24.5%	21.8%
30-39	25.7%	32.5%	25.8%	24.6%	28.9%	24.8%	17.0%	17.5%	23.7%
40-49	16.9%	21.2%	16.9%	17.3%	21.8%	17.4%	25.1%	27.3%	17.1%
50 and Older	22.2%	18.2%	22.2%	22.4%	24.0%	22.4%	12.3%	11.9%	25.1%

*The higher percentage of Nonbinary stops by Tier 3 agencies is largely attributable to a group of northern coastal agencies including Astoria PD, Cannon Beach PD, Clatsop CO SO, Seaside PD, and Warrenton PD which stems from a data entry issue.

Table 3.1.3. displays the most serious dispositions reported by law enforcement. 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[16]. About 69 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 3.1.3. Stop Disposition by Stop Type and Tier

Disposition	Tier 1			Tier 2			Tier 3		
	Traffic	**Ped.**	**Total**	**Traffic**	**Ped.**	**Total**	**Traffic**	**Ped.**	**Total**
None	2.6%	13.7%	2.8%	7.3%	18.0%	7.6%	5.0%	6.4%	5.0%
Warning	56.7%	50.2%	56.6%	60.0%	55.0%	59.8%	69.0%	77.3%	69.4%
Citation	38.2%	16.8%	37.8%	30.4%	15.9%	30.1%	24.8%	11.7%	23.9%
Juv. Summons	0.0%	0.1%	0.0%	0.0%	0.2%	0.0%	0.0%	0.1%	0.0%
Arrest	2.4%	18.3%	2.7%	1.7%	10.4%	1.9%	1.1%	4.5%	1.3%

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

Table 3.1.4. provides Year 4 search data, stratified by Tier. Tier 1 agencies conduct searches in 2.2 percent of stops, a higher percentage than Tier 2 and Tier 3. Pedestrians were searched more often than drivers, but searches were less successful. For Tier 1 and Tier 2 agencies, about 40 percent of all searches were consent searches. For Tier 3 agencies, consent searches made up less, at about a quarter of all searches. Echoing previous STOP reports, drugs were the most common form of contraband found in searches. Tier 3 agencies found alcohol more often (38.2 percent) during a search than Tier 2 (16.7 percent) or Tier 1 agencies (27.7 percent).

Table 3.1.4. Search Results by Stop Type and Tier

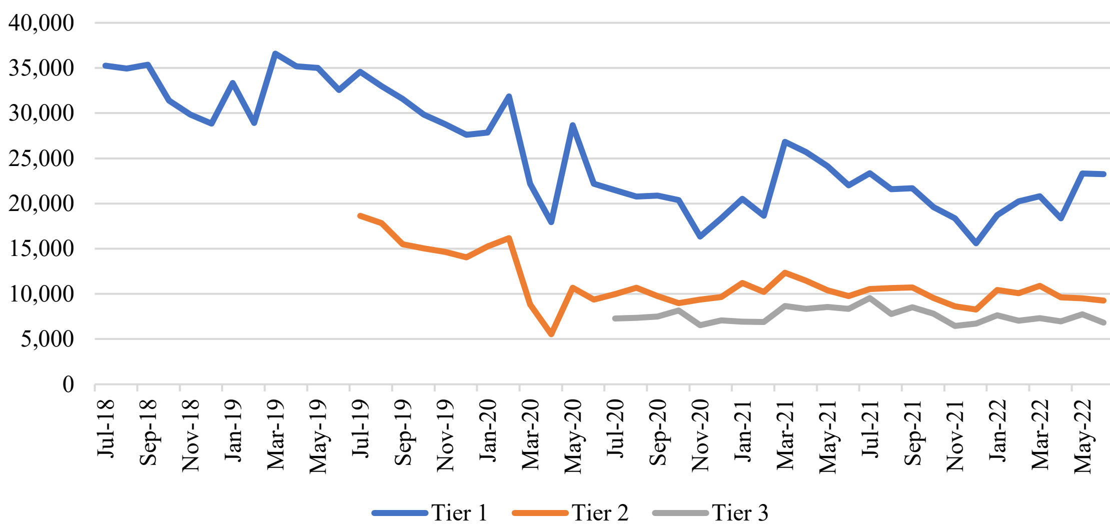
Variable	Tier 1			Tier 2			Tier 3		
	Traf.	Ped.	Total	Traf.	Ped.	Total	Traf.	Ped.	Total
Search Conducted	2.0%	14.0%	2.2%	1.4%	8.5%	1.5%	0.7%	5.1%	0.9%
Reason									
Consent Search	40.2%	46.2%	41.0%	40.8%	46.4%	41.5%	27.9%	19.5%	25.6%
Consent Search Denied	0.8%	0.8%	0.8%	0.3%	0.0%	0.3%	0.0%	0.0%	0.0%
“Other” Search	59.0%	53.0%	58.2%	58.9%	53.6%	58.2%	72.1%	80.5%	74.4%
Percent Successful	50.9%	43.4%	49.9%	40.0%	32.3%	39.0%	50.5%	36.2%	46.5%
Item Seized*									
Alcohol Found	30.0%	8.3%	27.7%	17.6%	9.2%	16.7%	45.5%	11.8%	38.2%
Drugs Found	47.6%	46.2%	47.4%	51.2%	50.0%	51.1%	39.6%	52.6%	42.5%
Weapons Found	16.8%	24.0%	17.6%	18.8%	18.4%	18.7%	10.9%	18.4%	12.5%
Stolen Property Found	6.0%	9.0%	6.3%	5.7%	2.6%	5.4%	4.7%	2.6%	4.3%
Other Evidence Found	16.5%	27.1%	17.6%	15.7%	11.8%	15.2%	14.2%	11.8%	13.7%
Other Non-Evidence Found	4.9%	10.4%	5.5%	17.5%	21.1%	17.9%	6.2%	9.2%	6.8%

*Percentages are the percent of *searches* in which the type of property was seized, not the percentage of *stops*. Multiple items can be seized in one search, so percentages may add to more than 100%.

3.2. Longitudinal STOP Data Trends

Figure 3.2.1. displays stops made by Oregon law enforcement agencies from July 2019 through June 2022, stratified by Tier. From July 2019 through June 2020, only Tier 1 and Tier 2 agencies reported stops. In July 2020, Tier 3 agencies started reporting. From February to March 2020, when COVID-19 mitigation efforts were first put in place, Tier 1 stop volume dropped by 30 percent and Tier 2 stop volume dropped by a greater percentage, 45 percent.

Figure 3.2.1. Stops by Month of Year, by Tier



Overall stop volume dropped a further 24 percent in April 2020, before rebounding in May 2020 to 82 percent of the pre-pandemic stop volume. In November 2020, a two-week statewide freeze was implemented to prevent the spread of COVID-19, and stop volume dropped particularly for Tier 1 agencies. 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. Tier 1 agencies show a 42 percent drop in stop volume from March to December 2021, while Tier 2 agencies dropped 33 percent and Tier 3 agencies show a 22 percent decrease.

Figure 3.2.2. Stops by Race/Ethnicity, July 2019 - June 2022

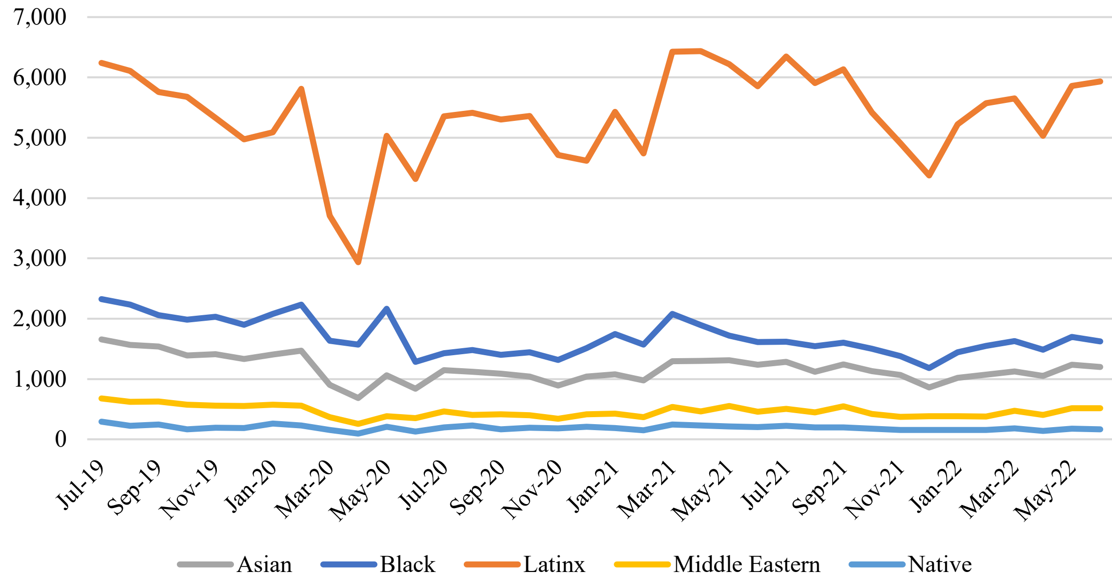


Figure 3.2.2. shows all reported police stops by racial/ethnic group—excluding stops of white individuals—from July 2019 through June 2022[17]. In the December 2020 report, STOP researchers noted that, although stops dropped for all racial groups in March and April 2020, stops of Black individuals did not fall as much as other racial/ethnic groups, potentially because white workers were more likely to work from home[18]. This result is partially echoed in later months as the pandemic continued to influence stop volumes. In November 2020, when overall stop volume dropped 14 percent from October, stops of white individuals dropped 15 percent, Latinx individuals dropped 12 percent, Asian/PI individuals dropped 17 percent, Middle Eastern individuals dropped 14 percent, and stops of Black and Native individuals dropped only 9 and 6 percent, respectively. While stop volume increased in March 2021, the general decrease to the end of the year was not uniform across racial/ethnic groups. From March to December 2021, stops of white individuals dropped 36 percent, while stops of Black individuals decreased 43 percent and stops of Native American individuals dropped 37 percent. The drop in stops for other racial groups was somewhat more muted with a 32 percent drop for Latinx individuals, 34 percent decrease for Asian individuals, and a 29 percent drop for Middle Eastern individuals stopped.

---

[17] White stops make up 77 percent of monthly stops, on average, and largely echo the monthly variation shown in Figure 3.2.1.

[18] From June 2020 article, *Ability to Work From Home: Evidence From Two Surveys and Implications for the Labor Market in the COVID-19 Pandemic*.https://www.bls.gov/opub/mlr/2020/article/ability-to-work-from-home.htm

Table 3.2.1. Search Rates by Year and Tier

Year	Tier 1	Tier 2	Tier 3
Year 2 (19-20)	2.6%	2.7%	N/A
Year 3 (20-21)	2.5%	1.9%	1.4%
Year 4 (21-22)	2.2%	1.5%	0.9%

In March 2022, the Oregon Legislature passed SB 1510[19], 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 are not effective until January 1, 2023, many agencies have already implemented them. Table 3.2.1. shows search rates by Tier and Year, includes searches from July 2019 to June 2022. Overall search rates have dropped, with Tier 1 agencies showing a search rate of 2.6 percent in Year 2 and dropping to 2.2 percent in Year 4. Tier 2 agencies show the largest drop from 2.7 percent in Year 2 to 1.5 percent in Year 4. Finally, Tier 3 agencies show a search rate of just under one percent in Year 4.

4. Decision to Stop Analysis

Often referred to as the "gold standard" of statistical analyses examining the initial law enforcement decision to stop an individual[20], 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 to 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 including: 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 tables that follow in the next subsection, this difference in odds is presented as an odds ratio, which

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

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

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 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.

4.1. 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. First, Table 4.1.1. 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. As described in Section 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. For the full Tier 1 and Tier 2 models in Table 4.1.1., all comparisons show no statistically significant differences in the odds of minority stops in daylight compared to darkness[21].

Table 4.1.1. Logistic Regression of Minority Status on Daylight by Tier 1 or Tier 2 Agency

Agency	Asian/PI	Black	Latinx	Middle Eastern	Native American
Beaverton PD	0.81	0.71	0.86	0.62	--
Clackamas CO SO	1.13	0.99	1.19	0.68	1.20
Corvallis PD	0.60	1.04	0.82	--	--
Eugene PD	1.15	1.34	1.00	--	--
Gresham PD	--	1.60	1.23	--	--
Hillsboro PD	1.08	0.86	1.06	0.97	--
Lake Oswego PD	1.81	0.84	1.09	--	--
Marion CO SO	1.24	1.72	1.03	--	--
Milwaukie PD	--	1.11	1.16	--	--
Multnomah CO SO	1.83	0.95	1.25	--	--
Oregon State Police	0.93	1.11	1.03	1.14	1.21
Portland PB	1.08	0.99	1.17	0.90	--
Springfield PD	--	0.71	0.73	--	--
Tigard PD	0.59	1.07	1.39	--	--
Tualatin PD	2.51	1.11	0.77	--	--
Washington CO SO	0.85	0.89	0.89	0.63	--

Notes: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction based on number of comparisons).
Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, year, county stop volume, and agency stop volume.

[21] The odds ratio for Multnomah COSO for Asian/PI drivers (1.83) shows a p-value of 0.03. With the Bonferroni adjustment with three comparisons this is not significant, however without the adjustment the p-value is below the 0.05 threshold. Multnomah SO is not identified in the Stop Outcomes or Search Findings analysis. The odds ratio for Tualatin PD for Asian/PI drivers (2.51) shows a p-value of 0.03. With the Bonferroni adjustment with three comparisons this is not significant, however without the adjustment the p-value is below the 0.05 threshold. Tualatin PD is not identified in Search Findings analysis but is identified in the Stop Outcomes analysis.

Table 4.1.2.a. reports the Tier 1 and Tier 2 agency specific model results for Latinx drivers compared to white drivers for agencies not displayed above. While a number of agencies have odds ratios above 1.0, all agencies show no statistically significant difference in the rate of stopped Latinx drivers in daylight compared to darkness.

Table 4.1.2.a. Logistic Regression of Latinx Drivers on Daylight by Tier 1 or Tier 2 Agency

Agency	Latinx	Agency	Latinx
Albany PD	1.04	Lincoln City PD	1.12
Bend PD	0.89	Lincoln CO SO	1.32
Benton CO SO	0.95	Linn CO SO	0.80
Canby PD	1.09	Medford PD	0.71
Central Point PD	1.00	McMinnville PD	0.55
Deschutes CO SO	0.74	Newberg-Dundee PD	0.86
Forest Grove PD	1.07	Oregon City PD	0.99
Grants Pass PD	1.09	Polk CO SO	1.11
Hermiston PD	0.80	Redmond PD	0.84
Hood River CO SO	0.57	Roseburg PD	0.88
Jackson CO SO	1.36	Salem PD	1.09
Keizer PD	1.09	West Linn PD	0.97
Klamath Falls PD	0.80	Woodburn PD	1.14
Lane CO SO	0.82	Yamhill CO SO	1.03

Notes: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction based on number of comparisons).
Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, year, county stop volume, and agency stop volume.

Table 4.1.2.b. reports the Tier 3 agency specific model results for Latinx drivers compared to white drivers for agencies with sufficient sample size. Similar to Tier 1 and Tier 2 agencies, most agencies show no statistically significant difference in the rate of stopped Latinx drivers in daylight compared to darkness. For Sandy PD, however, the odds of stops for Latinx drivers in daylight was 2.1 times the odds for white drivers, indicating a statistically significant difference evidencing a disparity in the rate of stopped drivers in daylight compared to darkness.

Table 4.1.2.b. Logistic Regression of Latinx Drivers on Daylight by Tier 3 Agency

Agency	Latinx	Agency	Latinx
Astoria PD	1.71	Morrow CO SO	0.96
Brookings PD	1.82	Mt. Angel PD	1.08
Cannon Beach PD	0.74	Newport PD	1.19
Crook CO SO	0.84	Pendleton PD	0.44
Dallas PD	0.79	Phoenix PD	2.46
Eagle Point PD	1.19	Prineville PD	0.79
Gervais PD	3.42	Sandy PD	2.12*
Gilliam CO SO	0.83	Seaside PD	0.78
Gladstone PD	1.68	Sherman CO SO	1.75
Hood River PD	0.74	Sherwood PD	0.86
Hubbard PD	0.84	Silverton PD	1.42
Jefferson CO SO	1.16	Stanfield PD	1.15
Josephine CO SO	0.89	Tillamook CO SO	1.19
Madras PD	0.52	Umatilla CO SO	1.07
Milton-Freewater PD	0.81	Umatilla PD	0.82
Monmouth PD	0.76		

Notes: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction based on number of comparisons).
Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, year, county stop volume, and agency stop volume.

5. 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 outcomes of stops, including whether stopped individuals were cited, searched, and/or arrested during their encounter with law enforcement.

HB 2355 required all law enforcement agencies to collect data regarding the disposition of stops. Because stops can have multiple dispositions (e.g., 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[22]. This means, therefore, 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.

5.1. Description of Stop Outcomes Analysis

Variation in enforcement outcomes could be due to time of day, day of the week, the offense that led to the stop, or one of 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, gender, or season. 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. Here, STOP Program researchers use these methods to answer the question, "holding other factors constant, do we find 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 us 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 in the statistical literature, 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[23].

Table 5.1.1. Analyses Completed for Each Agency

Disposition of Interest	Comparison Dispositions	Analysis Groups				
Citation	None or Warning	Asian/PI	Black	Latinx	Mid. Eastern	Native
Search	None, Warning, or Citation	Asian/PI	Black	Latinx	Mid. Eastern	Native
Arrest	None, Warning, Citation, or Search	Asian/PI	Black	Latinx	Mid. Eastern	Native
Citation, Search, or Arrest	None or Warning	Asian/PI	Black	Latinx	Mid. Eastern	Native

[22] See Appendix F for more details on how the STOP research team determines the most serious disposition and the appropriate comparison outcomes for each type of disposition.
[23] 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 F.

The current analysis included 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 was drawn from the group of white stops for the agency in question. Each row of Table 5.1.1. describes the tests conducted for each agency. In row 1, STOP Program researchers tested whether there was a disparity in issuing citations between each of the racial groups shown in the analysis groups column and a matched white group.[24] Row 2 does the same for searches, row 3 for arrests, and row 4 describes tests for any Citation, Search, or Arrest disposition.

5.2. Stop Outcomes Results

As with the Decision to Stop analysis in the previous section, the analyses conducted in this section include two years of data for all agencies. Table 5.2.1. reports agency-level results for agencies where a statistically significant disparity was found for a search or arrest outcome, in addition to citation or any outcome. For six agencies, Beaverton PD, Jefferson CO SO, Marion CO SO, Oregon State Police, Salem PD, and Washington CO SO, disparities were reported for either searches and/or arrests of Latinx individuals, sometimes in addition to citations. Oregon State Police and Pendleton PD show disparities in searches and arrests for Native individuals, while Portland PB shows a disparity in searches of Black individuals.

Table 5.2.1. Predicted Disparity by Agency and Disposition (only statistically significant results displayed)

Agency	Race/ Ethnicity	Citation		Search		Arrest		Any Outcome	
		Actual	Pred.	Actual	Pred.	Actual	Pred.	Actual	Pred.
Beaverton PD	Asian	53.8%	49.8%	--	--	--	--	--	--
Beaverton PD	Latinx	--	--	4.6%	3.3%	7.8%	6.5%	--	--
Jefferson CO SO	Latinx	--	--	--	--	2.9%	0.6%	--	--
Marion CO SO	Asian	--	--	--	--	--	--	88.2%	85.1%
Marion CO SO	Latinx	84.0%	81.5%	3.9%	2.6%	3.7%	2.6%	84.7%	81.9%
Oregon State Police	Asian	40.0%	37.5%	--	--	--	--	40.7%	38.6%
Oregon State Police	Black	42.6%	38.8%	--	--	--	--	43.9%	40.0%
Oregon State Police	Latinx	44.4%	38.5%	2.5%	1.7%	1.9%	1.7%	46.3%	39.9%
Oregon State Police	Mideast	40.0%	37.7%	--	--	--	--	--	--
Oregon State Police	Native	43.8%	37.8%	4.2%	2.3%	4.5%	2.4%	47.1%	39.7%
Pendleton PD	Native	37.0%	20.8%	23.6%	15.2%	24.2%	14.8%	52.8%	33.0%
Portland PB	Black	--	--	6.5%	4.6%	--	--	--	--
Salem PD	Latinx	--	--	10.1%	7.9%	7.8%	5.1%	65.6%	62.3%
Washington CO SO	Latinx	22.6%	19.9%	1.8%	1.3%	3.5%	2.7%	25.6%	22.3%

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 a citation or any outcome are displayed in Table 5.2.2. For three Tier 1 agencies, Clackamas CO SO, Eugene PD, and Hillsboro PD, disparities were detected only for citations and/or for the combined measure of all dispositions (i.e., citation or search or arrest). This indicates that for these agencies, it is likely that the only relevant disparity is for citations and not the other outcomes. As described in Section 3, Tier 2 agencies 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

[24] Each matched white group will differ from the next, since the characteristics of the stops of the group being matched differ.

sufficient sample sizes to complete the analysis. That said, of the analyses that were completed, Forest Grove PD, Hermiston PD, Lake Oswego PD[25], Oregon City PD, Polk CO SO, Tigard PD, Tualatin PD, Woodburn PD, and Yamhill CO SO had statistically significant disparities indicated for one or more of the analysis groups for citations and any outcome.

Table 5.2.2. Predicted Outcome by Agency and Disposition (only statistically significant results displayed)

Agency	Race/ Ethnicity	Citation		Any Outcome	
		Actual	Pred.	Actual	Pred.
Astoria PD	Latinx	27.4%	16.4%	--	--
Brookings PD	Latinx	18.3%	11.6%	18.3%	11.6%
Cannon Beach PD	Latinx	27.7%	17.5%	--	--
Clackamas CO SO	Asian	35.4%	31.6%	--	--
Coburg PD	Latinx	57.8%	46.1%	57.8%	46.8%
Cottage Grove PD	Latinx	--	--	60.7%	26.2%
Eugene PD	Latinx	36.5%	32.6%	--	--
Forest Grove PD	Latinx	37.4%	30.1%	39.4%	32.2%
Gervais PD	Latinx	78.2%	67.8%	78.7%	68.5%
Gilliam CO SO	Latinx	70.4%	61.4%	71.3%	61.5%
Hermiston PD	Latinx	29.5%	22.3%	30.7%	23.2%
Hillsboro PD	Latinx	32.1%	27.1%	34.6%	29.2%
Hubbard PD	Latinx	25.2%	18.4%	27.8%	20.7%
Independence PD	Latinx	28.8%	20.8%	28.6%	21.2%
Lake Oswego PD	Latinx	49.7%	42.7%	50.1%	43.2%
Madras PD	Native	--	--	48.8%	29.0%
McMinnville PD	Latinx	37.8%	30.0%	38.5%	30.8%
Morrow CO SO	Latinx	33.3%	26.5%	33.9%	27.2%
Newport PD	Latinx	32.6%	20.4%	34.9%	21.8%
Oregon City PD	Asian	37.8%	27.7%	--	--
Polk CO SO	Latinx	26.9%	21.2%	28.9%	23.5%
Sherwood PD	Latinx	25.7%	19.7%	27.7%	21.3%
Stanfield PD	Latinx	--	--	29.2%	23.6%
Tigard PD	Latinx	44.9%	32.1%	46.9%	34.9%
Tillamook CO SO	Latinx	46.4%	38.0%	--	--
Tualatin PD	Latinx	49.5%	44.0%	51.1%	45.8%
Umatilla PD	Latinx	26.1%	21.0%	27.5%	21.8%
Woodburn PD	Latinx	45.7%	39.1%	47.6%	40.7%
Yamhill CO SO	Latinx	24.4%	20.9%	--	--

Sample size issues were even more pronounced for Tier 3 agencies. However, the following sixteen Tier 3 agencies were identified as having significant disparities in only citations and/or any disposition for one of the analysis groups: Astoria PD, Brookings PD, Cannon Beach PD, Coburg PD, Cottage Grove PD, Gervais PD, Gilliam CO SO, Hubbard PD, Independence PD, Madras PD, Morrow CO SO, Newport PD, Sherwood PD, Stanfield PD, Tillamook CO SO, and Umatilla PD. Pendleton PD showed statistically significant disparities in citations, searches, and arrests of Native Americans as compared to whites. Jefferson CO SO showed a statistically significant disparity in the arrests of Latinx individuals.

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

[25] Lake Oswego PD has identified a potential technology issue that could impact the stop data submission. CJC will re-run the analyses for Lake Oswego PD if needed once the issue is resolved.

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.

6. 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, to name a few) 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.[26] 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 G for more detailed technical information about the KPT Hit-Rate model and statistical tests.

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

6.1. 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) depending upon sample size. Results for these analyses are presented in Table 6.1.1. below.

Table 6.1.1. Hit-Rates and Significance by Agency and Race/Ethnicity

Agency	Race/Ethnicity	Minority Hit-Rate	White Hit-Rate	Significance
Albany PD	Latinx	40.9%	42.9%	
Beaverton PD	Black	60.8%	61.4%	
	Latinx	65.4%	61.4%	
Bend PD	Latinx	14.0%	8.8%	
Clackamas CO SO	Latinx	48.8%	46.9%	
Eugene PD	Black	32.0%	39.1%	
	Latinx	34.7%	39.1%	
Gresham PD	Latinx	60.6%	39.7%	
Hermiston PD	Latinx	44.4%	37.1%	
Hillsboro PD	Latinx	48.9%	52.5%	
Hubbard PD	Latinx	42.9%	38.3%	
Marion CO SO	Latinx	9.8%	12.3%	
Medford PD	Black	46.7%	31.3%	
	Latinx	46.6%	31.3%	
Multnomah CO SO	Black	56.9%	57.5%	
	Latinx	52.6%	57.5%	
OSP	Black	70.4%	65.3%	
	Latinx	67.2%	65.3%	
	Asian/PI	75.3%	65.3%	
	Native	69.4%	65.3%	
Pendleton PD	Latinx	42.9%	34.1%	
	Native	41.8%	34.1%	
Polk CO SO	Latinx	72.7%	68.8%	
Portland PB	Black	51.9%	49.2%	
	Latinx	50.3%	49.2%	
	Asian/PI	59.6%	49.2%	
Salem PD	Latinx	44.2%	38.0%	
Springfield PD	Latinx	44.2%	38.8%	
Washington CO SO	Latinx	71.5%	64.0%	

Notes: * $p<0.05$, ** $p<0.01$, *** $p<0.001$ (Statistical Significance includes a Bonferroni Correction by agency with the number of comparisons shown)

As shown in Table 6.1.1., all agencies have differences in search success rates between white individuals and the comparison groups. These differences in nearly all cases were relatively small, and none of the differences reported were statistically significant. The lack of statistical significance could be attributed to the relatively small sample sizes found across agencies (particularly for Tier 2 and Tier 3 agencies), but it is also important to note that 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. No search findings comparisons made in this report were found to be statistically significant. This means that no agency was identified as having a statistically significant disparity for the Search Findings analysis.

7. Findings from 2022 Analysis

7.1. Aggregate Findings

Similar to previous STOP Reports, in all, the STOP data demonstrates that the vast majority of discretionary police-citizen interactions in Oregon are traffic stops. The breakdown between traffic and pedestrian stops varies by both agency as well as Tier, however, as some law enforcement agencies engage in more pedestrian stops than others, and Tier 3 agencies, on average, logged more pedestrian stops proportionally than Tier 1 and Tier 2 agencies.

With regard to the demographic characteristics of stopped individuals, the aggregate data continue to indicate that the majority of stops in Oregon were of white drivers or pedestrians. This is not surprising given the demographic makeup of Oregon as a whole. When disaggregated by traffic versus pedestrian stops, the data indicate that minorities made up a larger share of individuals stopped for traffic violations compared to those stopped as pedestrians. With regard to gender, males were stopped more often than females and nonbinary individuals, and this split was greater for pedestrian stops versus traffic stops.

Law enforcement agencies reported that stopped individuals either were subject to no further action or merely given a warning in a little under 60 percent of stops for Tier 1, a little under 70 percent of stops for Tier 2, and a little under 75 percent of stops for Tier 3 agencies. Other outcomes, including receiving a citation or being arrested, varied widely across traffic and pedestrian stops, as pedestrian stops were more likely to end in an arrest, and traffic stops were more likely to end in a citation, regardless of Tier.

7.2. Decision to Stop Analysis Results 2022

One of the few consistent findings reported across the academic and professional literature examining police stop data is that comparisons between stops initiated by law enforcement and residential Census data often leads to invalid, biased results. To examine the decision to stop a driver in a manner that does not rely on benchmarks, STOP Program researchers again utilized the Decision to Stop analysis, which examines stops made in daylight versus darkness surrounding sunrise and sunset. The threshold for identifying disparities was a resulting odds ratio above 1.0 that was statistically significant at the 95 percent confidence level for any minority group at the agency level.

In this analysis, one Tier 3 agency was found to have a disparity in the rate of stopped minority drivers in daylight versus darkness compared to white drivers. Sandy PD shows the odds of stops for Latinx drivers in daylight was 2.1 times the odds of white drivers.

7.3. Stop Outcomes Analysis Results 2022

The Stop Outcomes analysis, which relies on balancing samples across racial/ethnic groups to compare similarly situated individuals, was the first of two models used to examine stop outcomes after the decision to stop a driver has been made. For this analysis, STOP Program researchers identified all agencies with statistically significant disparities in their predicted versus actual dispositional outcomes for Asian/PI, Black, Latinx, Middle Eastern, and Native American groups, respectively.

In total, eight Tier 1 agencies, nine Tier 2 agencies, and eighteen Tier 3 agencies were identified as meeting this threshold. For Tier 1 agencies this included: Beaverton PD, Clackamas CO SO, Eugene PD, Hillsboro PD, Marion CO SO, Oregon State Police, Salem PD, and Washington CO SO. Among Tier 2 agencies, Forest Grove PD, Hermiston PD, Lake Oswego PD, Oregon City PD, Polk CO SO, Tigard PD, Tualatin PD, Woodburn PD, and Yamhill CO SO were identified for disparities. For Tier 3 agencies, Astoria PD, Brookings PD, Cannon Beach PD, Coburg PD, Cottage Grove PD, Gervais PD, Gilliam CO

SO, Hubbard PD, Independence PD, Jefferson CO SO, Madras PD, Morrow CO SO, Newport PD, Pendleton PD, Sherwood PD, Stanfield PD, Tillamook CO SO, and Umatilla PD were identified.

The most common dispositional outcome identified with disparate outcomes was citations, which, in general, is a much more common outcome than searches and arrests. Similarly, the group most often identified for disparate outcomes was Latinx, for which there were generally more stops relative to other non-white groups. There were, however, additional findings regarding searches and arrests. Beaverton PD, Marion CO SO, Oregon State Police, Salem PD, and Washington CO SO were identified for searches of Latinx individuals. Beaverton PD, Jefferson CO SO, Marion CO SO, Oregon State Police, Salem PD, and Washington CO SO were indicated for arrests of Latinx individuals. Oregon State Police and Pendleton PD were identified for searches and arrests of Native American individuals. And Portland PB was identified for searches of Black individuals. Notably, many analyses for several agencies could not be estimated due to low sample sizes, especially for smaller agencies. In these situations, we cannot detect the presence of a disparity with current data limitations.[27]

The findings of the Stop Outcomes analysis are likely influenced, at least in part, by departmental policies and the reason for the stop disposition. CJC has conducted an additional analysis for agencies that submitted optional data elements to account for the reason for disposition in the Stop Outcomes analysis[28]. While CJC has identified missing data and other technology challenges with the optional data elements, this additional analysis should further contribute to conversations about disparities in stop outcomes.

7.4. Search Findings Results 2022

The second of two analyses examining post stop outcomes was the Search Findings analysis, which compared the percentages of successful searches across different racial/ethnic groups. As discussed in detail in Section 6., the theoretical idea at the foundation of this test is that if law enforcement personnel apply search criteria or standards equally across race/ethnicity, then similar success rates should be found for all racial/ethnic groups. For this analysis, STOP Program researchers identified all agencies with disparities in their hit-rates where those differences were statistically significant at the 95 percent confidence level in any minority group at the agency level. In this year's analysis, there were no agencies identified as having statistically significant results. This means that no agency was identified as having a statistically significant disparity for the Search Findings analysis.

7.5. Conclusions

The data contained in this report are intended to be used as a tool for law enforcement, citizens and 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 in the search 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

[27] Full results, including for tests that could not be completed, are available upon request.
[28] See Appendix D Stop Outcomes Additional Analysis

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.

No agency was identified as having a statistically significant disparity in two or more tests performed on the STOP data this year. Therefore, no agency is referred to receive technical assistance from DPSST in this report. However, that does not mean that the results for any agencies should be ignored or are not close to the threshold of identification. Regardless of whether an agency is officially referred to DPSST, the CJC urges each agency to scrutinize their full set of results.

7.6. Next Steps and Future Work

The fourth annual STOP Program report includes data from 148 Oregon law enforcement agencies and is the first report to include two years of data for all Tier 1, Tier 2, and Tier 3 agencies. This achievement culminates a multi-year effort to create a statewide data collection system for all officer-initiated traffic and pedestrian stops that are not associated with calls for service. While this is a significant milestone, the STOP program has encountered several challenges, including the impacts of the ongoing COVID-19 pandemic, resource limitations for some agencies, and the ongoing challenge of sample size limitations to conduct the rigorous statistical analyses necessary to identify disparities. Despite these challenges, the full statewide STOP Program report includes the results of rigorous analyses for nearly 100 agencies across the state and contributes to dialogues between law enforcement agencies and the communities they serve.

8. Oregon Law Enforcement Contacts and Data Review Committee Report

8.1. 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 Senate Bill 415. In 2015, HB 2002 created a standard definition of profiling[29], 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, after which it was transferred to the Criminal Justice Commission (CJC) by order of House Bill 5050, Section 13. This report summarizes the information found in the profiling complaints the LECC received from Oregon law enforcement agencies in calendar years 2020 and 2021. In previous years, this report was published independently; however, in 2022 it was added as an additional section to the existing STOP Report.

8.2. Summary of 2020 and 2021 Reports

Table 8.2.1. summarizes law enforcement agency reporting for 2020 and 2021. During both reporting years, 113 law enforcement agencies reported the number of profiling complaints they received. This represents a reporting compliance rate of 66 percent for both years. Of those agencies that reported in

[29] 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."

2020, 17 percent (N=19) reported at least one complaint, and across those 19 agencies there were a total of 74 complaints. In 2021, 24 percent (N=27) of agencies that reported had at least one complaint and across those agencies, 85 total complaints were received.

Table 8.2.1. Law Enforcement Annual Reporting Compliance, 2020 and 2021

	2020	2021
Agencies Reporting	113	113
Total Reported Complaints	74	85
Agencies Reporting No Complaints	94	86
Agencies Reporting 1+ Complaints	19	27

Table 8.2.2. shows the number of complaints reported by agency in 2020 and 2021. Across those two years, the largest law enforcement agency in the state, the Oregon State Police, had the highest complaint volume with 25 complaints. The agencies with the next highest report volume over that period were Multnomah CO SO (N=20), Clackamas CO SO (N=18), Portland PB (N=15), and Eugene PD (N=11), which are all tier 1 agencies.

Table 8.2.2. Reported Incidents by Agency, 2020 and 2021

Department	2020	2021
Albany PD	0	1
Ashland PD	2	3
Astoria PD	4	0
Beaverton PD	0	3
Bend PD	0	6
Central Point PD	1	0
Clackamas CO SO	9	9
Clatsop CO SO	0	1
Corvallis PD	0	1
Eugene PD	5	6
Forest Grove PD	0	1
Keizer PD	0	2
Klamath CO SO	0	1
Lake Oswego PD	2	1
Lane CO SO	1	1
Marion CO SO	2	1
Medford PD	2	4
Milwaukie PD	3	2
Multnomah CO SO	8	12
North Bend PD	1	0
Oregon City PD	5	1
Oregon State Police	14	11
Pendleton PD	0	1
Polk CO SO	0	1
Portland PB	8	7
Springfield PD	1	4
St. Helen PD	0	1
Talent PD	1	0
The Dalles PD	1	0
Tigard PD	3	1
Tillamook PD	1	0
Washington CO SO	0	3
Total	74	85

Table 8.2.3. shows the dispositions of those complaints that were reported in 2020 and 2021[30]. Of the 159 complaints in those years where copies were sent to the CJC, not a single complaint received a disposition of sustained. For comparison purposes, a report by the California Racial and Identity Profiling Advisory Board that analyzed data on 10,044 racial profiling complaints in California in 2018 found that 10.8% of all reports were sustained[31]. The most common disposition in both years was "unfounded" followed by "not sustained."

Table 8.2.3. Reported Profiling Complaints by Disposition

Disposition	2020	2021
Exonerated	5	5
Not Sustained	12	25
Unfounded	29	34
Administrative Closure	3	0
No Basis for Further Investigation	12	5
Other	5	11

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. For example, a recent report by the CJC[32] found that in 2020, 149 individuals reported to the Oregon Department of Justices' (DOJ) Bias Response Hotline (BRH) that they had experienced a bias crime or incident, and that the reported individual committing the bias was a law enforcement agent. In 76 of those reports, the reporter indicated reporting the incident to law enforcement, while in 33 reports, the reporter indicated that they attempted to report the incident to law enforcement.

8.3. LECC and Bias Response Hotline Report Demographics

Table 8.3.1. shows the reports submitted to the BRH where the perpetrator was a law enforcement agent, along with the reports submitted by law enforcement agencies to the LECC, broken out by gender and race. As these tables show, the demographics of the victims in reports submitted to the BRH differed from those in the reports submitted by law enforcement agencies through the LECC. Specifically, reports from the hotline were much more likely to come from female or gender nonbinary individuals than reports provided to the CJC by law enforcement. This may indicate that individuals in those groups are less likely than males to report instances of profiling directly to the law enforcement agency involved.

Table 8.3.1. Reported Complaints in 2021 by Complainant Race and Gender

Demographics	LECC	Percent	Bias Hotline	Percent
Gender				
Male	40	47%	43	40%
Female	18	21%	36	34%
Nonbinary	1	1%	8	7%
Unknown	26	31%	20	19%
Race				
Asian	1	1%	1	1%
Bi-Racial	0	0%	5	5%
Black	29	34%	33	33%
Latinx	7	8%	12	12%
Native American	2	2%	4	4%
White	10	12%	11	11%
Unknown	36	42%	34	34%

[30] Note that the totals here and in other tables that break complaints down by information contained in the complaints do not equal the total for complaints reported, as copies of some of the reported complaints were not sent to the LECC.
[31] See https://oag.ca.gov/sites/all/files/agweb/pdfs/ripa/ripa-board-report-2020.pdf
[32] https://www.oregon.gov/cjc/CJC%20Document%20Library/SB577ReportJuly2021.pdf

8.4. Conclusion

This report provides an overview of the profiling complaints the LECC received from law enforcement agencies across Oregon in 2020 and 2021. During that period, 140 Oregon law enforcement agencies reported receiving a total of 159 reports. Of those reports, nearly half were disposed as “unfounded” or “not sustained.” A comparison of these reports to reports involving law enforcement made to the DOJ’s BRH suggest that certain demographic groups, specifically those who identify as female or nonbinary, may be less likely to report instances of perceived profiling to law enforcement. This comparison also suggests that the BRH may be a useful way of collecting information on incidents of perceived profiling that are not reported directly to law enforcement, especially for certain demographic groups.

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.

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.

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.

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... "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).

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.

"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, Kareem L. Jordan, and Shaun L. Gabbidon. 2011. "Racial Profiling in Decisions to Search: A Preliminary Analysis Using Propensity-Score Matching." *International Journal of Police Science & Management* 13 (4): 336–47.

Higgins, George E., Melissa L. Ricketts, James D. Griffith, and Stephanie A. Jirard. "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).

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.

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.

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.” 2013. StataCorp. https://www.stata.com/manuals13/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.

“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.

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

Appendix B – Data Audit

This STOP report uses data with a frequency of missingness displayed in Table B.1. This missingness manifests in overall rates of missingness within variables (displayed in Table B.1.) and as stops missing altogether. Additionally, some data used in this report look atypical compared to overall rates. The STOP Program team has worked to the extent possible to correct for incorrect and improbable data before the release of this report, but some challenges remain.

Table B.1. Missing Data for STOP Program Variables used in Year 4 Report Analyses

Variable	Description	% Missing
Age	Age as perceived by officer	0.6%
Agency	Stopping agency	0.0%
Arrest	Physical custody arrest (yes/no)	0.1%
Stop Reason*	Category of stop reason (Move/Spd, Ser Move/Spd, Very Ser Move/Sp, Equip Vio/Cell/Seatbelt, Reg/License, Other)	1.4%
county	County in which stop occurred	0.0%
disposition	Most severe disposition of stop (none, warning, citation, search, arrest)	0.3%
gender	Gender perceived by officer (male, female, nonbinary)	0.2%
race	Race/ethnicity perceived by officer (Asian/PI, Black, Latinx, Middle Eastern, Native American, white)	0.7%
sdate	Date of stop. Converted into day of the week, season, and time sun rises and sets of the day of the stop.	0.0%
search	Whether a discretionary search occurred (yes/no)	0.0%
search_f1**	What was found if a search occurred (Nothing, Alcohol, Drugs, Stolen Property, Weapons, Other Evidence, Other non-Evidence)	0.1%
search_t1**	Search type	0.0%
stime	Time of stop. Converted into time categories (12-5 am, 5-10 am, 10 am-3 pm, 3-8 pm, and 8 pm-12 am)	0.0%
stop_type	Type of stop (traffic, pedestrian)	0.2%

*Stop Reason is a condensed variable created from the original variables that contain the code and text, which denote the ORS code and text description, respectively, for the stop reason.
**These missing percentage reflects the percent of missing when an entry is likely expected. In the case that Search= "no," there is not an entry expected, so these are not included in the missing percentage in this table.

Table B.1. displays the overall rates of missingness for variables used in STOP analyses for Year 4 data, however these rates vary widely between agencies. In addition, some agencies show atypical patterns for submitted data. For example, five Tier 3 agencies (Astoria PD, Cannon Beach PD, Clatsop CO SO, Seaside PD, and Warrenton PD) show that over 80 percent of stops are for nonbinary individuals which stems from a data entry issue. In addition, eleven Tier 3 agencies show an arrest rate over 75 percent. CJC will continue to work with agencies on trouble shooting the stop data submission process.

For Year 4, CJC did not receive data from nine Tier 3 agencies. All Tier 3 agencies were required to begin reporting in July 2020. Five of these agencies were considered active agencies, while four were considered presently inactive for the purposes of this report. These agencies are listed in Table B.2. and Table B.3. respectively. No analysis was done for these agencies.

Table B.2. Active Agencies Which Did Not Submit Year 4 Data
Burns PD
Harney CO SO
Hines PD
Lake CO SO
PSU CPS

Table B.3. Inactive Agencies Which Did Not Submit Year 4 Data
Butte Falls PD
Gearhart PD
Powers PD
Rockaway Beach PD

Appendix C – Stop Characteristics for Tier 3 Agencies

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

Agency	Asian/PI	Black	Latinx	Middle Eastern	Native American	White	Total
Astoria PD	26	33	111	5	2	1,673	1,850
Aumsville PD	5	4	76	4	0	411	500
Baker City PD	6	8	40	1	8	550	613
Baker CO SO	19	17	87	6	0	1,141	1,270
Bandon PD	34	13	45	11	0	670	773
Black Butte Ranch PD	10	7	25	5	0	301	348
Boardman PD	0	1	52	0	0	128	181
Brookings PD	30	24	124	9	3	994	1,184
Cannon Beach PD	54	45	166	29	3	1,688	1,985
Carlton PD	9	2	58	1	0	384	454
Clatsop CO SO	17	31	83	15	0	1,114	1,260
Coburg PD	36	53	118	34	0	873	1,114
Columbia City PD	2	5	14	2	0	138	161
Columbia CO SO	26	29	66	12	6	1,381	1,520
Coos Bay PD	10	23	75	7	6	2,212	2,333
Coos CO SO	12	5	55	6	3	793	874
Coquille PD	5	12	40	3	1	674	735
Cottage Grove PD	5	3	15	0	0	207	230
Crook CO SO	10	21	114	4	2	1,153	1,304
Curry CO SO	17	11	15	4	1	325	373
Dallas PD	20	26	107	4	0	1,014	1,171
Eagle Point PD	18	17	173	4	1	1,135	1,348
Enterprise PD	4	1	9	0	0	106	120
Florence PD	10	8	27	1	0	421	467
Gervais PD	4	16	152	2	0	466	640
Gilliam CO SO	18	18	44	5	0	307	392
Gladstone PD	63	153	353	44	19	2,329	2,961
Gold Beach PD	20	16	50	18	1	401	506
Grant CO SO	1	0	4	1	1	79	86
Hood River PD	20	21	261	12	14	866	1,194
Hubbard PD	24	27	742	3	0	817	1,613
Independence PD	15	19	231	6	1	541	813
Jacksonville PD	0	0	3	0	0	35	38
Jefferson CO SO	60	27	404	20	30	1,686	2,227
John Day PD	0	0	2	0	0	53	55
Josephine CO SO	36	49	200	17	2	1,658	1,962
Junction City PD	3	9	27	2	3	241	285
La Grande PD	45	30	48	5	0	1,016	1,144
Madras PD	17	11	148	6	42	339	563
Malheur CO SO	2	3	53	2	2	197	259
Malin PD	3	1	29	0	0	83	116
Manzanita DPS	28	8	45	4	0	319	404
Merrill PD	4	1	5	3	0	21	34
Milton-Freewater PD	10	14	241	1	0	461	727
Molalla PD	14	11	103	13	3	1,067	1,211
Monmouth PD	43	53	231	9	0	850	1,186
Morrow CO SO	14	17	587	5	12	1,392	2,027
Mt. Angel PD	13	10	182	2	0	396	603
Myrtle Creek PD	8	17	30	2	0	862	919
Myrtle Point PD	0	4	8	1	0	91	104

(Table C.1. continued on next page)

Newport PD	34	25	170	5	6	876	1,116
North Bend PD	6	7	18	2	1	279	313
Nyssa PD	0	2	124	0	0	207	333
Oakridge PD	24	5	7	1	0	168	205
Ontario PD	4	12	67	0	0	295	378
OSU PD	19	9	17	0	1	101	147
Pendleton PD	57	60	179	6	120	1,599	2,021
Philomath PD	69	62	131	26	3	1,331	1,622
Phoenix PD	13	23	142	9	0	560	747
Pilot Rock PD	4	1	5	0	0	120	130
Port Orford PD	74	9	49	42	1	632	807
Prineville PD	32	43	342	12	0	3,447	3,876
Rainier PD	4	4	24	2	0	259	293
Reedsport PD	10	6	7	4	0	170	197
Rogue River PD	7	6	56	1	0	264	334
Sandy PD	93	62	235	29	37	2,152	2,608
Scappoose PD	21	22	39	9	2	740	833
Seaside PD	73	51	220	26	4	1,920	2,294
Sherman CO SO	45	18	172	18	3	751	1,007
Sherwood PD	111	85	388	37	10	2,699	3,330
Silverton PD	19	21	222	9	0	1,112	1,383
St. Helens PD	1	5	9	2	0	502	519
Stanfield PD	18	37	384	20	19	1,074	1,552
Stayton PD	18	13	128	1	0	856	1,016
Sunriver PD	27	21	104	3	1	1,106	1,262
Sutherlin PD	15	17	69	8	0	1,109	1,218
Sweet Home PD	1	1	3	0	1	167	173
Talent PD	24	43	142	9	0	1,200	1,418
The Dalles PD	2	4	77	4	3	263	353
Tillamook CO SO	23	19	121	14	0	811	988
Tillamook PD	23	18	105	16	3	678	843
Toledo PD	18	27	135	6	27	2,193	2,406
Turner PD	1	5	21	0	1	220	248
Umatilla CO SO	6	16	272	7	5	682	988
Umatilla PD	16	44	1204	3	13	1,418	2,698
Union CO SO	24	27	42	11	0	526	630
Union Pacific Railroad PD	0	2	4	0	0	51	57
Vernonia PD	1	0	2	0	0	94	97
Wallowa CO SO	2	1	1	0	1	93	98
Warrenton PD	18	14	70	5	4	1,318	1,429
Wasco CO SO	5	3	49	4	5	278	344
Wheeler CO SO	17	7	38	11	6	742	821
Winston PD	5	14	36	1	1	1,096	1,153
Yamhill PD	14	5	46	5	0	202	272
Grand Total	1,848	1,810	11,554	718	444	72,420	88,794

Table C.2. Percent and Number of Tier 3 Agency Stops by Stop Type Traffic vs. Pedestrian

Agency	Traffic		Pedestrian		Total
Astoria PD	1,850	100.0%	0	0.0%	1,850
Aumsville PD	499	99.8%	1	0.2%	500
Baker City PD	613	100.0%	0	0.0%	613
Baker CO SO	1,270	100.0%	0	0.0%	1,270
Bandon PD	773	100.0%	0	0.0%	773
Black Butte Ranch PD	348	100.0%	0	0.0%	348
Boardman PD	190	99.5%	1	0.5%	191
Brookings PD	1,184	100.0%	0	0.0%	1,184
Cannon Beach PD	1,838	92.6%	147	7.4%	1,985
Carlton PD	437	96.3%	17	3.7%	454
Clatsop CO SO	1,259	99.9%	1	0.1%	1,260
Coburg PD	1,099	98.7%	15	1.3%	1,114
Columbia City PD	161	100.0%	0	0.0%	161
Columbia CO SO*	1,515	99.7%	5	0.3%	1,521
Coos Bay PD	732	31.4%	1,601	68.6%	2,333
Coos CO SO	863	98.7%	11	1.3%	874
Coquille PD	184	25.0%	551	75.0%	735
Cottage Grove PD	230	100.0%	0	0.0%	230
Crook CO SO	1,274	97.7%	30	2.3%	1,304
Curry CO SO	367	98.4%	6	1.6%	373
Dallas PD	1,184	100.0%	0	0.0%	1,184
Eagle Point PD	1,323	98.1%	25	1.9%	1,348
Enterprise PD	119	99.1%	1	0.8%	120
Florence PD	466	99.8%	1	0.2%	467
Gervais PD	645	99.8%	1	0.2%	646
Gilliam CO SO	392	100.0%	0	0.0%	392
Gladstone PD	2,930	99.0%	31	1.0%	2,961
Gold Beach PD	504	99.6%	2	0.4%	506
Grant CO SO	86	100.0%	0	0.0%	86
Hood River PD	1,188	99.5%	6	0.5%	1,194
Hubbard PD	1,514	93.9%	99	6.1%	1,613
Independence PD	806	99.1%	7	0.9%	813
Jacksonville PD	38	100.0%	0	0.0%	38
Jefferson CO SO	2,226	100.0%	1	0.0%	2,227
John Day PD	55	100.0%	0	0.0%	55
Josephine CO SO	1,890	96.3%	72	3.7%	1,962
Junction City PD	285	100.0%	0	0.0%	285
La Grande PD	1,144	100.0%	0	0.0%	1,144
Madras PD	563	100.0%	0	0.0%	563
Malheur CO SO	252	97.3%	7	2.7%	259
Malin PD	116	100.0%	0	0.0%	116
Manzanita DPS	404	100.0%	0	0.0%	404
Merrill PD	34	100.0%	0	0.0%	34
Milton-Freewater PD	725	99.7%	2	0.3%	727
Molalla PD	1,048	86.5%	163	13.5%	1,211
Monmouth PD	1,185	99.9%	1	0.1%	1,186
Morrow CO SO	2,020	99.1%	19	0.9%	2,039
Mt. Angel PD	603	100.0%	0	0.0%	603
Myrtle Creek PD	878	95.5%	41	4.5%	919
Myrtle Point PD	104	100.0%	0	0.0%	104
Newport PD*	1,359	93.1%	92	6.3%	1,460
North Bend PD	293	93.6%	20	6.4%	313
Nyssa PD	333	100.0%	0	0.0%	333
Oakridge PD	205	100.0%	0	0.0%	205

(Table C.2. continued on next page)

Ontario PD	378	100.0%	0	0.0%	378
OSU PD	124	84.4%	23	15.6%	147
Pendleton PD	1,660	82.1%	361	17.9%	2,021
Philomath PD	1,615	99.6%	7	0.4%	1,622
Phoenix PD	608	81.4%	139	18.6%	747
Pilot Rock PD	130	100.0%	0	0.0%	130
Port Orford PD	806	99.9%	1	0.1%	807
Prineville PD*	3,435	77.6%	14	0.3%	4,425
Rainier PD	290	98.3%	5	1.7%	295
Reedsport PD	194	98.5%	3	1.5%	197
Rogue River PD	312	93.4%	22	6.6%	334
Sandy PD	2,600	99.7%	8	0.3%	2,608
Scappoose PD	824	98.9%	9	1.1%	833
Seaside PD	2,222	96.9%	72	3.1%	2,294
Sherman CO SO	997	99.0%	10	1.0%	1,007
Sherwood PD	3,301	99.1%	29	0.9%	3,330
Silverton PD	1,383	100.0%	0	0.0%	1,383
St. Helens PD	516	99.4%	3	0.6%	519
Stanfield PD	1,562	99.7%	4	0.3%	1,566
Stayton PD	963	94.8%	53	5.2%	1,016
Sunriver PD	1,256	99.5%	6	0.5%	1,262
Sutherlin PD	1,070	87.8%	148	12.2%	1,218
Sweet Home PD	173	100.0%	0	0.0%	173
Talent PD	1,385	97.7%	33	2.3%	1,418
The Dalles PD*	537	95.5%	3	0.5%	562
Tillamook CO SO	976	98.7%	13	1.3%	989
Tillamook PD	843	100.0%	0	0.0%	843
Toledo PD	2,404	99.9%	2	0.1%	2,406
Turner PD	248	100.0%	0	0.0%	248
Umatilla CO SO	1,126	99.3%	8	0.7%	1,134
Umatilla PD	2,737	99.5%	15	0.5%	2,752
Union CO SO	630	100.0%	0	0.0%	630
Union Pacific Railroad PD	0	0.0%	57	100.0%	57
Vernonia PD	97	100.0%	0	0.0%	97
Wallowa CO SO	98	100.0%	0	0.0%	98
Warrenton PD	1,429	100.0%	0	0.0%	1,429
Wasco CO SO*	532	96.7%	0	0.0%	550
Wheeler CO SO	786	95.7%	35	4.3%	821
Winston PD	1,152	99.9%	1	0.1%	1,153
Yamhill PD	271	99.6%	1	0.4%	272
Total Tier 3*	**85,273**	**94.4%**	**4,062**	**4.5%**	**90,361**

*Traffic and Pedestrian percentages do not add to 100% due to missing stop type data.

Appendix D – Stop Outcomes Additional Analysis

Beginning in February 2021, the CJC began collecting an additional data element termed "Most Serious Disposition." Since STOP analysis began, some law enforcement officers and agencies have reported that for some types of infractions they have little discretion in whether they cite, search, and/or arrest an individual. This low-level of officer discretion may contribute to disparate outcomes when a particular racial or ethnic group is more likely to have stops involving those types of low-discretion infractions. Prior auxiliary analyses of OSP have shown that accounting for low-discretion situations sometimes does influence the level of disparity found in outcome analyses.[33] As a result, Most Serious Disposition fields were added as optional data elements for submission in February 2021.

This additional data element allows officers to report the type of infraction for which the most serious disposition was allotted for a stop. If, for example, an individual was stopped for speeding, a search was conducted, an open container of alcohol was found, and the individual was arrested for DUII then the initial stop reason would be speeding but the most serious disposition would be DUII. Accounting for this information is important when comparing groups.

For many stops this additional data element has not been submitted. This may be due to a lack of training regarding what the information is meant to convey, lack of emphasis on the importance of this additional element, or some other factor. In some cases, it remains unclear what a missing value conveys: the officer may have simply neglected to complete the optional field or it may be that the initial stop reason was also the reason for the most serious disposition and thus the officer decided not to complete the seemingly redundant field. Anecdotally, the CJC has heard that both scenarios are present in the data, but it is impossible to tell the extent to which either reason is present, or for which stops each explanation prevails. Additionally, several agencies have identified technical issues to address before the data can be included in this analysis.[34]

For Year 4 of data collection (July 2021-June 2022), 22.9% of all stop observations had a missing most serious disposition code value.[35] Of the 145 agencies in the full Year 4 data set, 117 submitted the most serious disposition for 50% or more of stops, counting outcomes of None and Warning as completed. Of these, 71 agencies had enough remaining observations to complete at least one of the outcome analyses in this section.[36] These agencies are listed in Table D.1.

---

[33] For an example of the effect these policies may have on STOP Program analysis results, please see Appendix E of the 2019 Statistical Transparency of Policing Report (https://www.oregon.gov/cjc/CJC%20Document%20Library/STOP_Report_Final.pdf)

[34] Oregon State Police was able to submit the Most Serious Disposition information for stops that result in a citation only, and are not included in the other outcomes. Salem PD and Lake Oswego PD have identified technical issues to address.

[35] Counting stops with a disposition of None or Warning as non-missing, regardless of the most serious disposition variables' contents.

[36] Some agencies submitted text entries, and only text entries, in the most serious disposition variables. Fully cleaning and organizing text entries is a significant undertaking and was not completed at time of writing. Therefore only observations that included ORS numbers were used for the analysis at this time. Police agencies are strongly encouraged to submit ORS numbers in the future.

Table D.1. Agencies Included in Stop Outcomes Additional Analysis

Aumsville PD	Jackson CO SO	Polk CO SO
Beaverton PD	Jefferson CO SO	Port of Portland PD
Benton CO SO	Josephine CO SO	Prineville PD
Boardman PD	Keizer PD	Redmond PD
Canby PD	Klamath CO SO	Salem PD
Cannon Beach PD	La Grande PD	Sandy PD
Central Point PD	Lake Oswego PD	Seaside PD
Clackamas CO SO	Lane CO SO	Sherwood PD
Coburg PD	Lincoln City PD	Silverton PD
Coos CO SO	Lincoln CO SO	Stanfield PD
Corvallis PD	McMinnville PD	Sunriver PD
Crook CO SO	Milwaukie PD	Sutherlin PD
Deschutes CO SO	Molalla PD	Talent PD
Douglas CO SO	Morrow CO SO	Tigard PD
Forest Grove PD	Mt Angel PD	Tillamook CO SO
Gladstone PD	Multnomah CO SO	Tualatin PD
Grants Pass PD	Newberg-Dundee PD	Umatilla CO SO
Gresham PD	Newport PD	Umatilla PD
Hermiston PD	Nyssa PD	Wasco CO SO
Hillsboro PD	Oregon City PD	Washington CO SO
Hood River CO SO	Oregon State Police	West Linn PD
Hood River PD	Pendleton PD	Woodburn PD
Hubbard PD	Philomath PD	Yamhill CO SO
Independence PD	Phoenix PD	

The CJC has completed this additional Stop Outcomes analysis to further contribute to conversations about disparities in outcomes, as well as the utility of this data field and to encourage agencies to submit the additional data elements in the future. The CJC completed a comparative outcome analysis for the set of agencies that had greater than 50% submissions in this field in Year 4 *and* had enough data in Year 4 for a standalone analysis[37]. For these agencies in Year 4 the CJC completed a baseline analysis where the most serious disposition field is not used in analysis and compared this to the results of an otherwise identical analysis where the most serious information is used.

Agency-level policies regarding officer discretion by infraction type vary by agency. As a test for the current analysis, the CJC created a Reason for Disposition variable based on the polices of the Oregon State Police, which are likely to be a close approximation of the policies of most individual agencies across the state.[38] For these stops, the choices of the stopping officer are highly limited after they observe these statute violations and the stop is likely to result in a citation, search, and/or arrest. In the "Original" outcome analysis this variable is not included in the analysis and in the Most Serious Disposition analysis this indicator variable is included in the balancing equation and in the outcome regression equation.[39]

[37] The data were limited to Year 4 because submission of the Most Serious Disposition variable began part way through Year 3. While data submission began in February 2021, the level of missing data does improve beginning in June 2021.

[38] These ORS codes include 163.195, 163.196, 806.010, 807.010, 807.570, 811.125, 811.140, 811.175, 811.182, 811.231, 811.540, 811.700, 811.705, 813.010, 813.011, 830.035, 33.045, 135.280, 162.205, and any codes in the 475 section.

[39] For more information see Appendix F which provides technical details of the Stop Outcomes analysis.

Of the 71 agencies analyzed, 21 showed statistically significant indications of a disparity for at least one test. Table D.2. displays the results for the citation outcome. In many cases, after the Reason for Disposition variable was added to the analysis the results became statistically insignificant (indicated by ***) or the gap between the predicted outcome and the actual outcome decreased. These patterns, however, are not universal. In some cases, the inclusion of the Reason for Disposition variable caused the gap between the actual outcome rate and the predicted to widen. And for one agency the Reason for Disposition results were statistically significant whereas they were not statistically significant prior to the inclusion of the Reason for Disposition variable.

Table D.2. Additional Stop Outcomes Analysis - Citations (only statistically significant results displayed)

Agency	Race/ Ethnicity	Original		Most Serious Disposition	
		Actual	Pred.	Actual	Pred.
Forest Grove PD	Latinx	35.6%	28.8%	***	***
Hermiston PD	Latinx	32.2%	24.8%	***	***
Hubbard PD	Latinx	27.8%	20.8%	***	***
Independence PD	Latinx	29.4%	19.2%	***	***
Madras PD	Native	44.7%	24.6%	***	***
Morrow CO SO	Latinx	33.9%	25.6%	***	***
Mt Angel PD	Latinx	22.7%	13.5%	***	***
Newport PD	Latinx	30.7%	17.9%	30.7%	17.7%
Oregon State Police	Asian	40.3%	37.3%	40.3%	35.2%
	Black	42.9%	38.5%	42.9%	40.7%
	Latinx	44.1%	38.3%	44.1%	41.8%
	Mideast	***	***	40.0%	33.5%
Polk CO SO	Latinx	27.9%	20.4%	27.9%	22.2%
Redmond PD	Latinx	49.2%	38.6%	49.2%	41.3%
Tigard PD	Latinx	47.8%	37.9%	***	***
Umatilla PD	Latinx	26.9%	20.1%	***	***
Washington CO SO	Latinx	22.1%	19.8%	***	***

*** Indicates a statistically insignificant result.

Table D.3. Additional Stop Outcomes Analysis - Any Outcome (only statistically significant results displayed)

Agency	Race/ Ethnicity	Original		Most Serious Disposition	
		Actual	Pred.	Actual	Pred.
Boardman PD	Latinx	34.6%	13.0%	***	***
Forest Grove PD	Latinx	37.4%	30.9%	***	***
Hermiston PD	Latinx	33.2%	25.7%	***	***
Hillsboro PD	Latinx	31.8%	27.3%	***	***
Hubbard PD	Latinx	30.0%	22.3%	30.0%	23.0%
Independence PD	Latinx	29.3%	19.6%	***	***
Morrow CO SO	Latinx	34.4%	26.1%	***	***
Mt Angel PD	Latinx	23.1%	13.8%	***	***
Newport PD	Latinx	32.4%	19.8%	32.4%	19.8%
Pendleton PD	Native	52.9%	35.3%	52.9%	38.4%
Polk CO SO	Latinx	29.4%	22.2%	29.4%	24.0%
Redmond PD	Latinx	49.6%	39.5%	***	***
Tigard PD	Latinx	49.7%	40.6%	***	***
Umatilla PD	Latinx	28.4%	20.7%	***	***
Washington CO SO	Latinx	25.2%	22.1%	***	***

*** Indicates a result that went from statistically significant to statistically insignificant

The results for any outcome are displayed in Table D.3. Again, many agencies show that after the Reason for Disposition variable was added to the analysis, the results became statistically insignificant or the difference between the predicted and actual outcome decreased. Finally, Table D.4. displays the results for the arrest outcome, and of the three agencies that show a significant difference in the original model, all are statistically insignificant when including the Reason for Disposition variable.

Table D.4. Additional Stop Outcomes Analysis - Arrest (only statistically significant results displayed)

Agency	Race/ Ethnicity	Original		Most Serious Disposition	
		Actual	Pred.	Actual	Pred.
Beaverton PD	Latinx	7.9%	5.7%	***	***
Boardman PD	Latinx	11.5%	0.0%	***	***
Jefferson CO SO	Latinx	3.2%	0.4%	***	***

*** Indicates a result that went from statistically significant to statistically insignificant

The Reason for Disposition variable, therefore, has an impact on the Stop Outcome analysis results, which leads to some important findings: First, where agencies reliably submitted Most Serious Disposition information, this information often goes some way toward explaining the disparities detected by the Outcome Analysis. Therefore, the CJC strongly encourages all law enforcement agencies to submit this data, which may help explain *why* disparities are detected in a given situation. The CJC would be able to include this additional information in future baseline analyses, should these data become more complete within agencies and more consistent across agencies in the future.

Second, the results suggest that non-white groups, and especially the Latinx group, are generally more likely to violate a low-discretion ORS statute than the white population. Therefore, to address some of the disparities found in the STOP analyses, policymakers should consider how to 1) address the conditions that lead to non-white groups being more likely to be stopped for low-discretion statutes, 2) address the conditions that may lead non-white groups to violate low-discretion ORS statutes, and 3) increase the discretion of patrol officers for these otherwise low-discretion statutes.

Appendix E – 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 (often referred to as the "gold standard"), 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.[40] 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, 2020, to June 30, 2022. 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

[40] The covariates included in the models were age, gender, reason for the stop, day of week, time of day, quarter or season, 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.

County. Stops that occur in the daily evening twilight window (approximately 30 minutes between sunset and the end of civil twilight) are similarly removed since it is neither light nor dark during this time period. Adjustments have been made to account for daylight savings time (DST) in November and March. In addition, while most of Oregon is on Pacific Standard Time (PST), most of Malheur County is on Mountain Standard Time (MST). The stops in Malheur County have been adjusted to account for this time zone.

The log odds that result from the DTS logistic regression model were then converted to odds ratios. Thus, the model tests whether the odds of non-white traffic stops during daylight are significantly different from the odds of non-white traffic stops during darkness. The DTS approach tests whether the odds ratio is statistically significantly different from 1.0. If the odds ratio is not statistically different from 1.0, then the test finds no difference in stops made during daylight and darkness. If the odds ratio is greater than 1.0 and statistically significant, however, the test concludes the odds of non-white drivers being stopped in daylight is significantly higher than in darkness, which is taken as evidence of a racial disparity in stops, after accounting for additional control variables that are available in the stop data. Conversely, if the odds ratio is less than 1.0 and statistically significant, the odds of a non-white driver being stopped in daylight is significantly lower than in darkness. The logistic regression modeling was compiled using SAS software and utilizing the procedure logistic function[41].

[41] SAS software, Version 9.4 of the SAS System for X64_8PRO Windows. Copyright © 2002-2012 SAS Institute Inc., Cary, NC, USA.

Appendix F – 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 of these 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 unbalanced. 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.[42] 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 logistic-type of 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.[43]
    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, 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.[44]

---

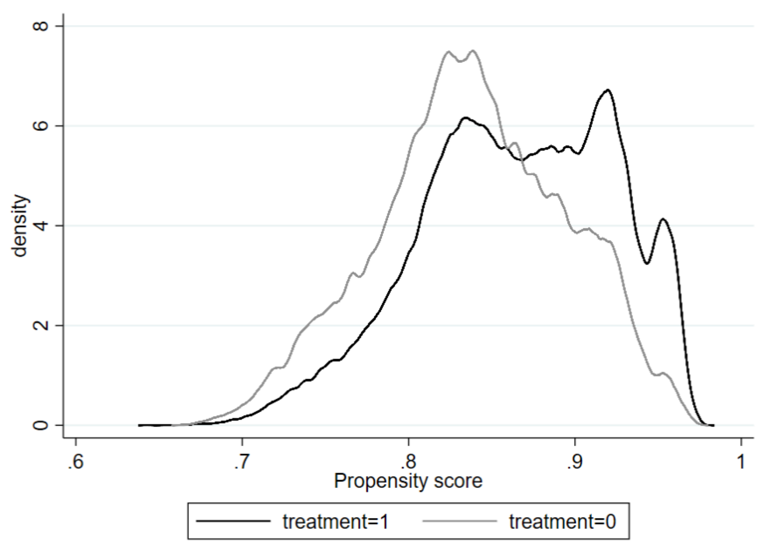
[42] StataCorp. 2013. Stata: Release 13. Statistical Software. College Station, TX: StataCorp LP.
[43] These differ whether the estimand is the Average Treatment Effect (ATE) or the Average Treatment Effect on the Treated (ATET). Here we are estimating the ATET (Austin and Stuart 2015).
[44] For a thorough 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[45], 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 F.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.[46]

[45] This assumption is also referred to as the unconfoundedness assumption.
[46] Omitted treatment variables per analysis are not presented here due to the high number of analyses conducted.

Finally, the Stable Unit Treatment Value Assumption (SUTVA), which 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 directly impact the race of subsequently stopped individuals.[47]

Estimation

If the above assumptions hold then estimation proceeds. 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 Moving/Speeding; Serious or Very Serious Moving/Speeding; Equipment, Cell, or Seatbelt; Registration/License; 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 $= \frac{Total\ \#\ of\ stops\ by\ agency\ on\ day\ of\ stop}{Maximum\ \#\ of\ daily\ stops\ in\ the\ \boldsymbol{county}\ over\ year\ of\ analysis}$

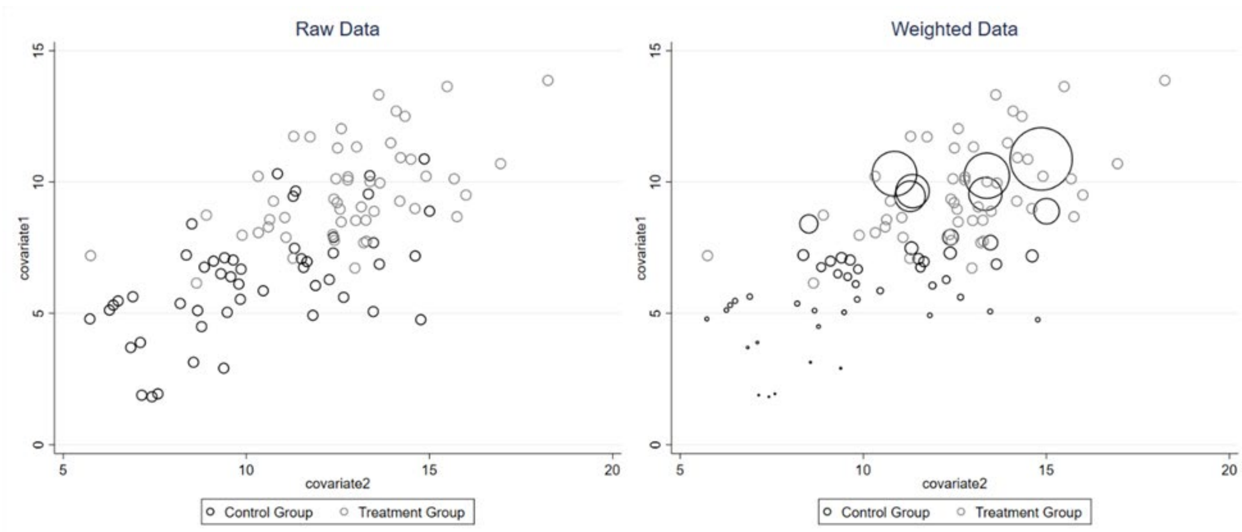
The first step of the analysis uses a probit model to estimate the propensity of being in the treatment group based on the covariates listed above. Overlap of propensity scores is evaluated and any non-overlapping observations are removed from the sample. Inverse Probability Weights (IPWs) are estimated for each observation based on the propensity scores. For the treatment group in an ATET framework,

[47] 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" 2013).

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) and treated observations that have a lower propensity score (i.e., are more similar to control 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 F.2. Weighting Example



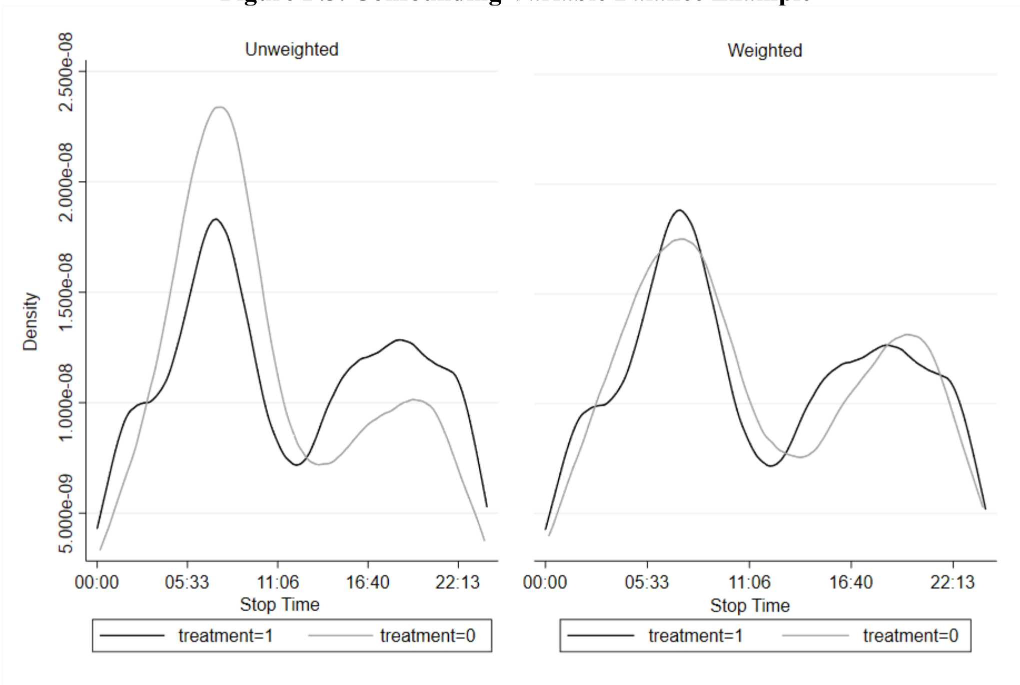
Balance is then measured based on the standardized difference[48] in means and the variance ratio[49] between the treatment and control groups for each of the raw data set and the inverse probability weighted data set. If the resulting standardized difference in the weighted data set is close to zero and the variance ratio is close to 1 for each variable for the weighted data then the sample is said to be balanced. Balance was evaluated in every data subset by agency and strong balance was achieved in every instance, e.g., the standardized differences were always close to zero (usually within .01 of 0, always within 0.05) and the variance ratios were always close to one (usually within .01 of 1, always within 0.05) (Austin 2009a; 2009b). In every case, the data sets were relatively well balanced in the initial, raw data sets, but became more balanced through the weighting process. This balance can also be evaluated graphically for each variable. Figure E.3. is an example of one of these variables for one agency. The Unweighted chart displays the distribution of stop time for each of the treated group and the untreated group. The Weighted chart displays these same distributions with the IPWs applied. The distributions of the two groups more

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

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

closely resemble each other in the weighted graph than in the unweighted graph, so we can say that these groups are more balanced when incorporating the IPWs.

Figure F.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.[50]

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

[50] 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.

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, time of travel, and 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 effect estimates built on top of these 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 across all the Tier 1 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 G – 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, we 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 we 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 our case, if a minority racial/ethnic group is "over-searched," then the hit-rate for that group will be lower than that of whites, 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% 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) 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, we 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% of the time (while all other members of the group carry contraband less than 50% of the time). Suppose also that group B has some portion of members that instead carry contraband 75% of the time (while all other members of the group carry contraband less than 50% of the time). If an officer only searches every person (regardless of group) who has over a 50% 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.