Statistical Transparency of Policing Report
Per House Bill 2355 (2017)

December 1, 2020

Revised December 8, 2020

Oregon Criminal Justice Commission

Kenneth Sanchagrin

Interim Executive Director

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

Authors and Contributors

Kelly Officer

Siobhan McAlister

Michael Weinerman

Courtney Rau

Katherine Tallan

Oregon Criminal Justice Commission Researchers

Executive Summary

House Bill 2355 (2017) mandated that by 2021, all Oregon law enforcement agencies must submit data regarding officer initiated traffic and pedestrian stops to the Oregon Criminal Justice Commission, so the Commission could analyze the submitted data for evidence of racial or ethnic disparities on an annual basis. To accomplish these ends, the Commission, along with the Oregon State Police and the Oregon Department of Public Safety Standards and Training (DPSST), created the Oregon Statistical Transparency of Policing (STOP) Program. This is the second annual report to the Oregon Legislature by the STOP Program examining data received pursuant to HB 2355.

Since the passage of HB 2355, the STOP Program 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 this time, the STOP Program has received at least one full year of data from the fifty-one largest law enforcement agencies in the state and analyses using those data are presented in this report. In 2021, the STOP Program will report on all Oregon police departments and sheriffs' offices, as required by the Bill.

Table E1 reports descriptive statistics for the combined Tier 1 and Tier 2 data contained in this report, which represents stops made from July 1, 2019 through June 30, 2020. Across all agencies, the vast majority of the reported data were for traffic stops, although the share of traffic stops made by Tier 2 agencies was slightly lower than 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 one-fifth of Tier 2 stops involved Asian/PI, 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 over 60 percent of stops. Other outcomes, including receiving a citation or being arrested, varied widely across agencies and are discussed in detail in the main body of the report. Finally, with regard to searches, approximately 2.7percent of all stops resulted in a search of some type.

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 Veil of Darkness Analysis (VOD). The VOD 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

Table E1.
Descriptive Statistics for Aggregate Year 2 Stop Data

Variable	Tier 1	Tier 2
Traffic Stop	97.8%	95.6%
Race/Ethnicity		
Asian/PI	3.4%	2.5%
Black	5.6%	3.1%
Latinx	12.9%	11.3%
Middle Eastern	1.4%	0.9%
Native American	0.6%	0.4%
White	76.2%	81.9%
Gender		
Male	66.9%	64.1%
Female	32.7%	35.7%
Non-Binary	0.4%	0.1%
Age		
Under 21	10.3%	11.6%
21 – 29	24.0%	22.2%
30 – 39	25.1%	24.5%
40 – 49	17.1%	17.2%
50 and Older	23.5%	24.5%
Stop Disposition		
None	2.6%	7.5%
Warning	57.1%	57.2%
Citation	37.6%	32.9%
Juvenile Summons	0.0%	0.0%
Arrest	2.7%	2.4%
Search Conducted	2.6%	2.8%

officer to discern race/ethnicity during the day when it is light versus the night when it is dark. Accordingly, the VOD 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 VOD 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 Predicted Disposition 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 KPT Hit-Rate 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 Oregon Department of Public Safety Standards and Training (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 two of the three analytical tests performed on the STOP data.

Using the above mentioned analyses and thresholds, the STOP Program identified one agency that had statistically significant results across two of the tests performed on the data: the Clackamas County Sheriff's Office. Specifically, results indicated that the Clackamas County Sheriff's Office had disparities in the Predicted Disposition Analysis with regard to citation patterns involving Black and Latinx individuals and in the KPT Hit-Rate with regard to searches of Latinx individuals. Thus, it is recommended that the Clackamas County Sheriff's Office be examined in greater detail by STOP Program researchers and receive technical assistance from DPSST.

Table of Contents

Table of Contents

Appendices

Table of Contents

List of Tables

Table of Contents

Table of Contents

Table of Contents

List of Figures

1. Background

This is the second annual report from the Statistical Transparency of Policing (STOP) Program. In 2017, the Oregon Legislature mandated that by July 2021 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 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 of 2019, the Criminal Justice Commission published its first annual STOP report, which contained data and analyses for the twelve largest law enforcement agencies in the state. This report builds on the first by including an additional thirty-nine agencies. The inclusion of these "Tier 2" agencies means that this report analyzes stops from the fifty-one largest law enforcement agencies in the state (for a full list, see Appendix A). Analyzed stops occurred from July 1, 2019 through June 30, 2020.

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), provided for the voluntary collection of stop data by agencies, and provided for analysis of collected data by the LECC.

With the exception of 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 the state.

HB 2355, which is codified in ORS 131.930 et seq., created a statewide data collection effort for all officer initiated traffic[1] and pedestrian[2] 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 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.

1 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).

2 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).

1. Background

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.

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

Tier	Number of Officers	Data Collection Begins	Reporting Begins
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

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, and similar to the previous method, the STOP Program also provides mobile applications free of charge for both iPhones and Android phones, through which officers can submit stop data for qualifying police-citizen interactions.

2. Methodological Approach

2.1. Background

The formal examination of police traffic and pedestrian stop data began 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.[3] 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 the residential demographics and those of the driving population.[4]

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, the nature of traffic and pedestrian stops necessitates the use of multiple tests. 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. Within the time it takes to execute and conclude a single stop, however, there are numerous opportunities where racially disparate treatment may be present. 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.

3 See STOP Program Research Brief: Analytical Approaches to Studying Stops Data (October 2018), which can be found at www.oregon.gov/cjc.
4 Using 2017 Census data via https://onthemap.ces.census.gov, it is possible to view the impact that work commuting has on Oregon cities and thus to understand the possible scope of the disjuncture between the driving population and residential census population of a given area. In Portland, for instance, the Census estimates that over 240,000 individuals commute into the city for work each day (about 60 percent of the city's workforce). In Beaverton, this pattern is even more pronounced, as over 85 percent of individuals working in Beaverton commute in from outside the city. Notably, commuting patterns do not just affect the Portland metro area, as Eugene, for example, displays a similar pattern. Specifically, it is estimated that 65 percent of individuals working in Eugene, approximately 91,000 people, commute into the city for work each day.

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]

Veil of Darkness Analysis. The Veil of Darkness test 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 Veil of Darkness test, 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.

Predicted Disposition Analysis. The Predicted Disposition test 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, whether the stop was made in the daylight, and agency and county stop volumes to determine whether one group is cited more often, searched more often, or arrested more often. If, after matching on all of the factors listed above, minority individuals are either cited, searched, or arrested more often than similarly situated white individuals, then there is evidence of a disparity.

Hit-Rate Analysis. The Hit-Rate test 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 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 501,064 records were submitted by the fifty-one Tier 1 and Tier 2 agencies during the second 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 data elements contained in the statute.

Date and Time the Stop Occurred. Law enforcement personnel are required to record the date (month/day/year)

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, technical descriptions of these analyses can be found in Appendices C, D, and E.

and time that the stop occurred. 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 small number of cases, some 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.

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 non-binary.

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 Oregon statute, which represent over 84 percent of all stops, law enforcement provide 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).[8]

Disposition of the Stop. The most serious 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 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 data included only 158 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.

7 For instance, approximately 5,000 stops labeled as "pedestrian" were for moving or speeding violations. Alternatively, 473 "traffic" stops were for ORS 814.020, which is specifically when a pedestrian fails to obey a traffic control device. For a full list of these inconsistent stops, please contact the STOP program..

8 Details on the offenses falling into each category are available upon request.

Justification for the Search. Law enforcement officers can provide several bases for a search using the following categories: consent search, consent search denied, or "other" search. The "other" search category includes frisks, probable cause searches, and other administrative searches. Multiple data points are allowed so that the data can include several search justifications. For example, if an officer initially requests to search an individual but consent is not given, an officer may then perform a search based on probable cause. In this example, the officer could record both "consent search denied" as well as "other search" into the database.

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

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

Finally, the STOP Program created two 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.

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. This issue was particularly acute in the newly added Tier 2 agencies, both because some agencies stopped a relatively low number of individuals and also because of changes in stop patterns due to the COVID-19 pandemic. In cases where the sample size is too small, statistical analyses cannot be conducted.

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 following sample size thresholds:

Table 2.3.1.
Sample Size Thresholds for Conducting Statistical Analyses

Statistical Test	Sample Size Threshold
Veil of Darkness	Minimum of 100 observations for an individual racial/ethnic group
Predicted Disposition	Minimum of 100 observations for an individual racial/ethnic group
Hit-Rate	Minimum 30 observations per racial/ethnic group analyzed; no cell with less than 5 observations

2. Methodological Approach

The sample size issue identified above had a significant impact on the STOP Program research team's ability to conduct analyses on all of the racial/ethnic groups found in the stop database. **Table 2.3.2a** and **Table 2.3.2b** report the breakdown by race/ethnicity and agency for all Tier 1 and Tier 2 agencies for stops occurring from July 1, 2019 through June 30, 2020. In several cases the total number of stopped individuals for certain racial/ethnic groups falls under the thresholds defined in **Table 2.3.1**. 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 Veil of Darkness), many of these counts fell under the requisite thresholds.

Table 2.3.2a.
Race/Ethnicity Reporting by Tier 1 Agency for All Reported Stops

Agency Name	Asian/PI	Black	Latinx	Middle Eastern	Native American	White
Beaverton PD	1,053	1,410	3,049	480	85	11,858
Clackamas Co SO	957	1,227	2,802	331	218	20,261
Eugene PD	371	880	921	0	74	14,261
Gresham PD	292	878	1,015	62	32	3,884
Hillsboro PD	620	547	2,349	320	59	6,622
Marion Co SO	421	418	2,956	148	21	10,538
Medford PD	50	143	472	16	3	2,493
Multnomah Co SO	364	850	1,207	117	42	7,211
Oregon State Police	3,480	4,497	17,903	2,004	932	132,089
Portland PB	1,984	6,227	3,736	480	134	23,956
Salem PD	227	269	1,785	31	32	5,617
Washington Co SO	1,483	1,322	4,850	682	205	16,308
Tier 1 Total	**11,302**	**18,668**	**43,045**	**4,671**	**1,837**	**255,098**

STOP Program researchers faced similar sample size issues with pedestrian stops. Across all Tier 1 and Tier 2 agencies, only 2.9 percent of stops, which represents 14,489 individual encounters, were pedestrian stops in the second year of data collection. In nearly all instances, models could not be estimated for pedestrian stops on their own. Further, when agency-level pedestrian stops are disaggregated by race/ethnicity, the problem becomes more acute. For instance, only two agencies stopped at least 100 Black pedestrians or 100 Latinx pedestrians in the second year of data collection. With regard to Asian/PI, Native American, or Middle Eastern pedestrians, no agency reported more than 50 stops. Due to these sample size issues, pedestrian and traffic stops were analyzed together in this report for all post-stop outcomes.[11]

9 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, which is certainly the case in many of the STOP research team's analyses, larger sample sizes are needed. For the analyses 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 observation for an individual racial/ethnic group at 100.

10 In some instances, despite having the minimum number of observations required to run a model, the models did not converge when estimated in Stata.

11 As the STOP Program database grows, it is likely that robust samples for pedestrian stops will be obtained. Once thresholds are met, these stops will be analyzed separately from traffic stops in future reports.

Table 2.3.2b.
Race/Ethnicity Reporting by Tier 2 Agency for All Reported Stops

Agency Name	Asian/PI	Black	Latinx	Middle Eastern	Native American	White
Albany PD	92	190	759	39	21	6,973
Ashland PD	99	132	241	39	5	3,193
Bend PD	105	127	600	30	21	7,086
Benton Co SO	174	183	408	75	24	4,597
Canby PD	33	45	438	12	7	1,840
Central Point PD	35	68	230	14	0	1,582
Corvallis PD	275	160	334	100	19	3,937
Deschutes Co SO	78	76	404	15	7	4,868
Douglas Co SO	63	51	175	33	2	3,252
Forest Grove PD	104	85	913	23	8	3,047
Grants Pass DPS	29	48	183	7	2	2,763
Hermiston PD	39	96	2,350	23	56	3,384
Hood River Co SO	45	16	298	6	3	934
Jackson Co SO	89	120	783	16	8	4,633
Keizer PD	37	53	485	8	2	1,505
Klamath Co SO	15	9	44	4	6	365
Klamath Falls PD	179	130	479	25	28	3,158
Lake Oswego PD	166	174	346	122	77	4,542
Lane Co SO	49	75	167	24	2	2,148
Lebanon PD	8	14	32	2	0	675
Lincoln City PD	101	91	253	20	3	2,034
Lincoln Co SO	114	51	235	25	17	2,600
Linn Co SO	85	105	395	43	35	5,203
McMinnville PD	59	55	613	17	1	2,622
Milwaukie PD	158	343	461	92	18	4,633
Newberg-Dundee PD	138	138	779	32	1	4,661
Oregon City PD	119	183	509	57	34	5,412
OHSU PD	45	69	53	36	0	505
Polk Co SO	121	94	637	26	13	2,639
Port of Portland PD	240	498	349	87	6	2,249
Redmond PD	41	30	225	7	4	2,069
Roseburg PD	34	133	296	24	8	5,213
Springfield PD	121	514	672	24	30	10,949
Tigard PD	355	342	751	125	36	4,853
Tualatin PD	222	224	793	63	13	4,344
U of O PD	31	39	42	15	0	673
West Linn PD	267	234	522	144	37	5,028
Woodburn PD	10	13	542	1	0	414
Yamhill Co SO	64	49	510	14	8	2,315
Tier 2 Total	**4,039**	**5,057**	**18,306**	**1,469**	**562**	**132,898**

Further, due to these sample size limitations, statistical models aggregated at STOP Tier level are conducted for all racial/ethnic groups where possible, while models for agency-specific analyses are limited to comparisons between white and Black individuals and white and Latinx individuals. In the third annual STOP report, the STOP Program will conduct agency-specific analyses of Asian/PI, Native American, and Middle Eastern individuals when possible by combining several years of data.

A final concern is the prevalence of missing data. 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. Although missing data attributable to both of these sources were found, the total amount of missing data in the STOP database is low compared to the overall sample size and is not a threat to the validity of the statistical models from a methodological point of view. [12]

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 Oregon Department of Public Safety Standards and Training (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.[13] 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.

Multiple comparisons were made for each agency- or tier-level test (i.e., Black-white and Latinx-white each with multiple outcomes). 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 or tier 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 Black individuals or Latinx individuals to run tests for those particular groups. In theory, this should cause the Bonferroni adjustment to differ by agency, conditional on the number of tests that were possible to run for each agency. In this report we do not vary the Bonferroni correction, thereby holding all agencies to the same standard of significance. In future iterations of this report, we may apply varying Bonferroni corrections depending on agency-level data availability.

Table 2.4.1.
Bonferroni Adjustment by Analysis

Statistical Test	Number of Comparisons
Veil of Darkness	Tier-level: 5 comparisons Agency-specific: 2 comparisons
Predicted Disposition	Tier-level: 20 comparisons Agency-specific: 8 comparisons
Hit-Rate	Tier-level: 5 comparisons Agency-specific: 2 comparisons

Adjusting the Bonferroni correction by agency would have implications for which agencies have disparities indicated and which agencies are referred for additional training. Variations in the appropriate Bonferroni thresholds will continue to be present in next year's report, which will include the smallest police agencies in the state (Tier three agencies), where there are fewer overall stops and may be greater variations in stops by race. Due to these concerns, and

12 See Appendix B for more details.

13 Given that multiple comparisons were made for each test (e.g., Black-white and Latinx-white across various outcome variables), Bonferroni adjustments were utilized to adjust for the likelihood of a given test resulting in a false positive. The Bonferroni adjustment differed for each test, and is described in greater detail in Sections 4, 5, and 6, as well as in the corresponding technical appendices.

also low statistical power due to small sample sizes, in the Year 3 STOP report and in reports following, two years of data will be used for each agency-level analysis, where available. The CJC anticipates that Tier three agency data will be sparse for the first year, but has not yet received those data and so can neither describe the capacity to run analyses for that first year nor describe any potential Bonferroni implications for that first year of Tier 3 data.

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,[14] 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.[15] 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 are, in many instances, 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 the vast majority of 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.

In spite of its promise as a means for systematically analyzing statewide data concerning police-citizen interactions, the STOP Program and its associated data and analyses have limitations. First, the statistical analyses can only identify *disparities* in police/citizen interactions. 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 department 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.

An additional limitation to the current report is the effect of the COVID-19 pandemic. In March 2020, many police agencies reduced the number of individuals stopped and some curtailed their discretionary stops entirely. The effects of this unexpected development are discussed in more detail in Section 3.2, although it is important to mention that this decrease in stops further exacerbated sample size limitations in both Tier 1 and Tier 2 analyses.

In spite of these limitations, the statistics 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 this report and believes that the results presented herein will contribute to the dialogue between law enforcement and Oregonians.

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

15 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 out of three threshold. Importantly, even if the Connecticut threshold was used for the STOP Program analyses, the referral results would be the same.

3. Characteristics of the Year 2 Stop Data

3.1. General Characteristics

This report analyzes data collected by the STOP Program for officer-initiated traffic and pedestrian stops occurring in Oregon from July 1, 2019 through June 30, 2020. In total, 501,064[16] stops were submitted to the STOP Program by the fifty-one Tier 1 and Tier 2 agencies. The number of stops reported by each agency are displayed in **Table 3.1.1a** and **Table 3.1.1b**. There was significant variation in the frequency with which Tier 1 and Tier 2 agencies stopped individuals. Tier 1 agencies generally made more stops than Tier 2 agencies, which is consistent with size differences in terms of officers employed and area patrolled. The Oregon State Police made 164,752 stops in Year 2, the largest number reported by any one of the fifty-five agencies, and one third of all stops in the state. Klamath Co. SO made the fewest stops, 447 stops, accounting for less than 0.1 percent of the reported stops in Year 2.

Table 3.1.1a and **Table 3.1.1b** report the number and percentage of stops by agency broken down by stop type—traffic or pedestrian. 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 of the Tier 1 agencies, only 2.2 percent of stops were of pedestrians. In year 2, Tier 1 agencies made predominately traffic stops. Only one Tier 1 agency, Medford PD, made less than 90 percent traffic stops. Across all Tier 2 agencies, a higher percentage, 4.4 percent, of stops were of pedestrians, however Tier 2 agencies exhibit more variability in stop type by agency. At one end of the continuum, 35 percent of University of Oregon PD stops, 24.3 percent of Port of Portland PD stops, and 13.5 percent of Roseburg PD stops were pedestrian stops. At the other end, however, two Tier 2 agencies, Keizer PD and Hood River Co SO, each made only one pedestrian stop in Year 2.

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

Agency	Traffic Stops		Pedestrian Stops	
Beaverton PD	16,972	94.6%	963	5.4%
Clackamas Co SO	24,362	94.4%	1,434	5.6%
Eugene PD	15,094	91.4%	1,424	8.6%
Gresham PD	6,089	98.8%	74	1.2%
Hillsboro PD	10,075	95.8%	442	4.2%
Marion Co SO	14,467	99.7%	37	0.3%
Medford PD	2,540	79.8%	643	20.2%
Multnomah Co SO	9,310	95.1%	481	4.9%
Oregon State Police	163,942	99.5%	810	0.5%
Portland PB	36,029	98.7%	488	1.3%
Salem PD	7,856	97.3%	217	2.7%
Washington Co SO	24,503	98.6%	347	1.4%
Tier 1 Total	**331,239**	**97.8%**	**7,360**	**2.2%**

16 Eight of these 501,064 stops were not definitively identified as either a pedestrian or traffic stop, so stop totals in Table 3.1.1. do not add up exactly to 501,064. All eight stops were made by Hermiston PD, representing about 1.3 percent of Hermiston's total stops.

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

Agency	Traffic Stops		Pedestrian Stops	
Albany PD	7,358	91.1%	716	8.9%
Ashland PD	3,437	92.7%	272	7.3%
Bend PD	7,909	99.3%	60	0.8%
Benton Co SO	5,435	99.5%	26	0.5%
Canby PD	2,345	98.7%	30	1.3%
Central Point PD	1,866	96.7%	63	3.3%
Corvallis PD	4,767	99.8%	58	1.2%
Deschutes Co SO	5,428	99.6%	20	0.4%
Douglas Co SO	3,412	94.8%	187	5.2%
Forest Grove PD	4,069	97.3%	111	2.7%
Grants Pass DPS	2,885	94.9%	156	5.1%
Hermiston PD	5,818	97.6%	141	2.4%
Hood River Co SO	1,302	99.9%	1	0.1%
Jackson Co SO	5,508	97.5%	141	2.5%
Keizer PD	2,089	99.9%	1	0.1%
Klamath Co SO	408	91.3%	39	8.7%
Klamath Falls PD	3,848	96.2%	151	3.8%
Lake Oswego PD	5,362	98.8%	67	1.2%
Lane Co SO	2,370	95.7%	107	4.3%
Lebanon PD	721	98.5%	11	1.5%
Lincoln City PD	2,474	98.9%	28	1.1%
Lincoln Co SO	3,052	99.6%	13	0.4
Linn Co SO	5,788	98.7%	78	1.3%
McMinnville PD	3,317	98.4%	54	1.6%
Milwaukie PD	5,276	92.5	429	7.5%
Newberg-Dundee PD	5,621	97.8	128	2.2%
Oregon City PD	5,827	92.3%	487	7.7%
OHSU PD	683	95.4%	33	4.6%
Polk Co SO	3,449	97.7%	81	2.3%
Port of Portland PD	2,601	75.7%	834	24.3%
Redmond PD	2,253	94.8%	123	5.2%
Roseburg PD	4,939	86.5%	769	13.5%
Springfield PD	11,473	93.2%	837	6.8%
Tigard PD	6,153	95.2%	309	4.8%
Tualatin PD	5,549	98.1%	110	1.9%
U of O PD	520	65.0%	280	35.0%
West Linn PD	6,130	98.4%	102	1.6%
Woodburn PD	964	98.4%	16	1.6%
Yamhill Co SO	2,922	98.0%	60	2.0%
Tier 2 Total	**155,328**	**95.6%**	**7,129**	**4.4%**

3. Characteristics of the Year 2 Stop Data

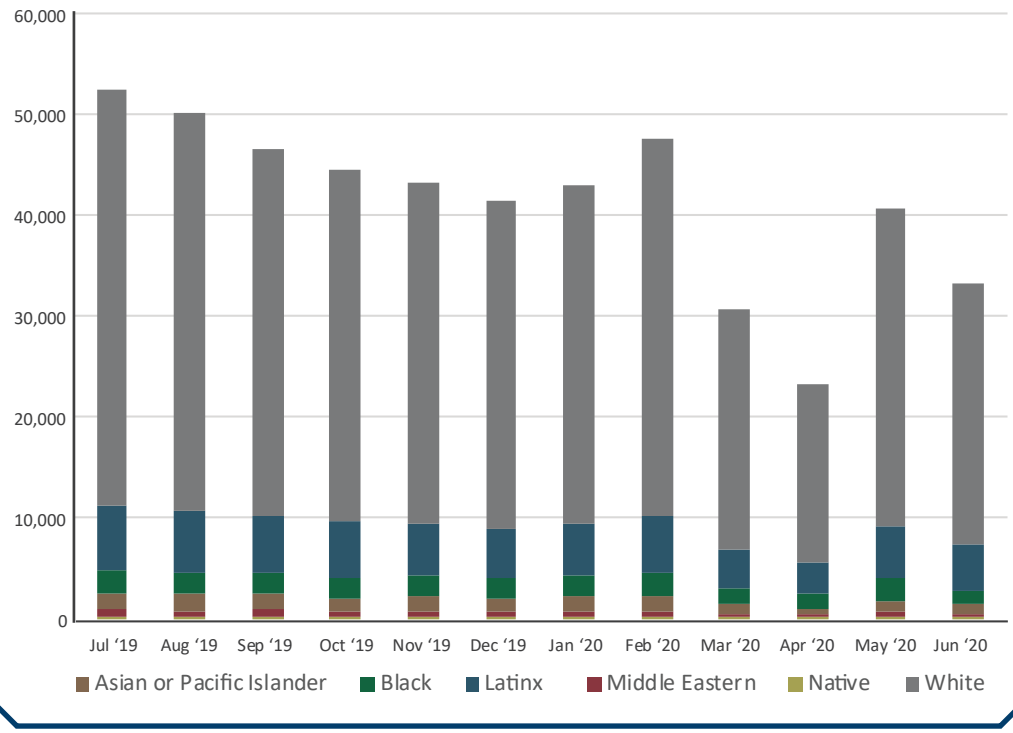
The demographic breakdowns for traffic and pedestrian stops are reported in **Table 3.1.2**. For both the Tier 1 and Tier 2 agencies contained in this report, the vast majority of stops were of white drivers/pedestrians, with Latinx and Black individuals being the two most frequently stopped minority groups. This pattern held when broken down by traffic versus pedestrian stops, although white individuals made up an even higher proportion of pedestrians. With regard to gender, more males were stopped than females. This gender difference is slightly more pronounced in pedestrian stops. Less than 1 percent of either traffic or pedestrian stops were of an individual perceived to be non-binary, with proportions similar across stop type and tier. Most traffic and pedestrian stops of are of individuals perceived to be aged in their thirties, slightly more so for pedestrians.

Figure 3.1.1 displays the number of traffic and pedestrian stops by the month of the year and racial/ethnic category. More stops occurred in July

Table 3.1.2.
Aggregate Tier 1 Demographics by Stop Type

	Tier 1			Tier 2		
	Traffic	Pedestrian	Total	Traffic	Pedestrian	Total
Race/Ethnicity						
Asian/PI	3.4%	1.7%	3.4%	2.5%	1.9%	2.5%
Black	5.6%	6.2%	5.6%	3.0%	4.9%	3.1%
Latinx	13.0%	8.2%	12.9%	11.5%	6.6%	11.3%
Middle Eastern	1.4%	0.4%	1.4%	0.9%	0.4%	0.9%
Nat. American	0.6%	0.6%	0.6%	0.3%	0.5%	0.4%
White	76.1%	83.1%	76.2%	81.7%	85.6%	81.9%
Gender						
Male	66.5%	81.8%	66.9%	63.5%	78.1%	64.1%
Female	33.1%	17.6%	32.7%	36.4%	21.7%	35.8%
Non-Binary	0.4%	0.6%	0.4%	0.1%	0.1%	0.1%
Age						
Under 21	10.4%	7.3%	10.3%	11.7%	9.4%	11.6%
21 – 29	24.2%	18.3%	24.0%	22.3%	18.9%	22.2%
30 – 39	25.0%	31.8%	25.1%	24.3%	28.1%	24.5%
40 – 49	17.0%	21.9%	17.1%	17.1%	20.5%	17.2%
50 and Older	23.5%	20.9%	23.5%	24.6%	23.1%	24.5%

Figure 3.1.1.
Traffic and Pedestrian Stops by Month of Year



and August than in other times of the year. There were fewer stops made in March and April than in other months. March and April 2020 were the months in Year 2 when the most extreme COVID-19 restrictions were implemented across the state. More details on the effects of COVID-19 on the STOP data are detailed in Section 3.2.

Figure 3.1.2 depicts traffic and pedestrian stops broken out by time of day. Few stops occured from three am through five am while about six times as many stops are reported from two to three pm. Demographic trends across time of day are generally consistent except during morning commuting times where a higher proportion of minority stops were reported from six am to nine am. Compared to Year 1 data, Year 2 exhibits relatively fewer stops during prime commuting times, five pm through six pm, which is due to the addition of Tier 2 agencies who made relatively fewer stops than Tier 1 agencies during evening commuting hours. **Figure 3.1.2a** shows stops broken out by time of day and tier.

Figure 3.1.2.
Traffic and Pedestrian Stops by Time of Day

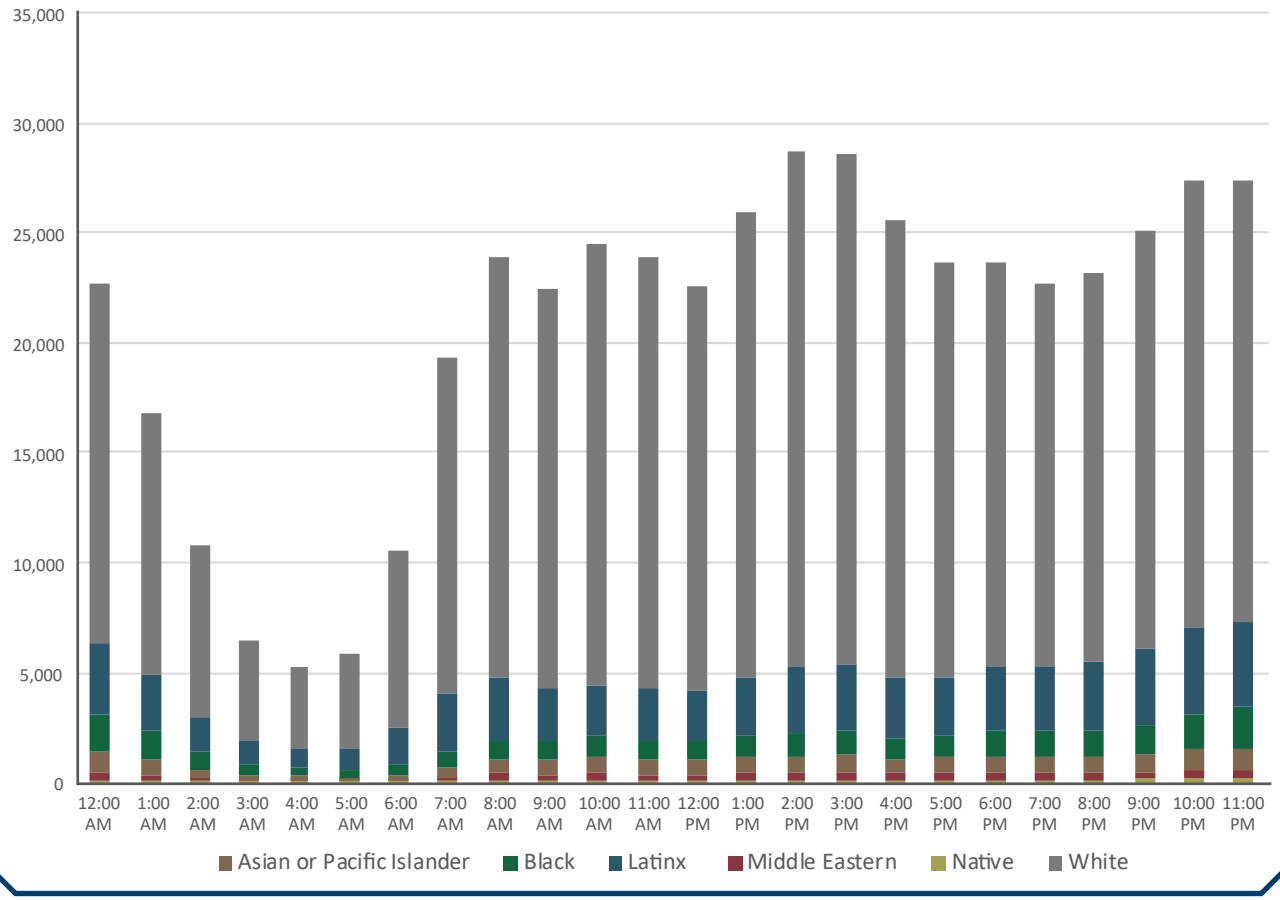


Figure 3.1.2a.
Hourly Stops by Tier

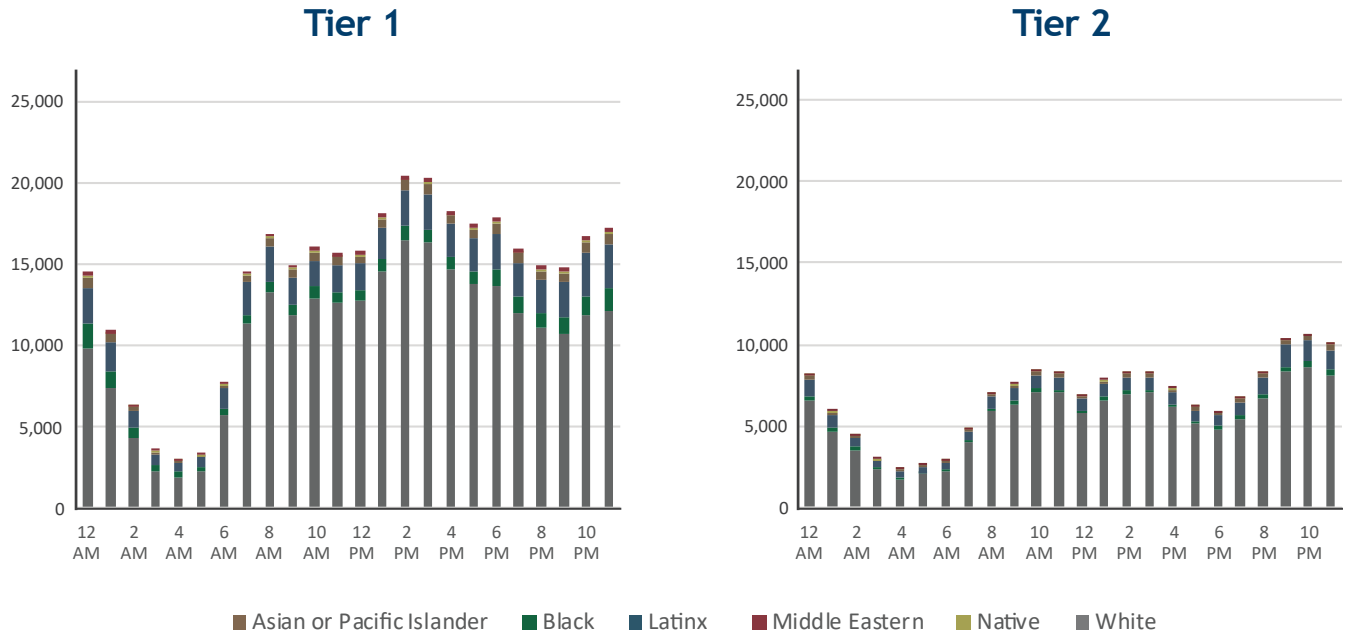


Table 3.1.3 displays the most serious dispositions reported by law enforcement. In all, most police stops did not result in further action taken against the stopped individual. The most common outcome of a police stop regardless of type or Tier was a warning.[17] There were differences, however, in sanction type by stop. Pedestrian stops were more likely to end in an arrest or juvenile summons than traffic stops, though arrests and juvenile sanctions were still relatively rare outcomes. In contrast, traffic stops were more likely to end in a citation than pedestrian stops.

Table 3.1.3.
Aggregate Tier 1 STOP Disposition by Stop Type

	Tier 1			Tier 2		
Disposition	**Traffic**	**Pedestrian**	**Total**	**Traffic**	**Pedestrian**	**Total**
None	2.4%	11.8%	2.6%	7.0%	18.0%	7.5%
Warning	57.0%	58.9%	57.1%	57.2%	57.4%	57.2%
Citation	38.2%	13.8%	37.6%	33.8%	12.9%	32.9%
Juvenile Summons	0.0%	0.2%	0.0%	0.0%	0.2%	0.0%
Arrest	2.4%	15.3%	2.7%	2.0%	11.6%	2.4%

Table 3.1.4 provides Year 2 data concerning searches, stratified by Tier. Overall, nearly 3 percent of stops resulted in a search of some type, with Tier 2 agencies reporting searches slightly more often than Tier 1 agencies. Pedestrians were searched more often by both Tier 1 and Tier 2 agencies, as 14.5 percent of pedestrians stopped by Tier 1 agencies were searched while drivers were only searched in 2.4 percent of stops.

17 It is the policy of many agencies to give a warning to everyone who is stopped.

Similarly, Tier 2 agencies searched 12.1 percent of pedestrians and 2.3 percent of drivers. With regard to the search justification, overall there was nearly an even split between consent searches and "other" searches, which include frisks, searches based on probable cause, and the like. Search justification is similar overall by Tier; 48.3 percent of Tier 1 searches and 48.7 percent of Tier 2 searches are consent searches. However, when stratified by stop type, a greater proportion of Tier 2 pedestrian searches (52.4 percent) are consent searches than Tier 1 pedestrian searches (41.5 percent). Tier 1 agencies conducted more searches that were successful, as 50.1 percent of Tier 1 searches were successful compared to a success rate of 35.7 percent for Tier 2. Drugs were the most common form of contraband found in searches, followed by alcohol. Tier 1 agencies found alcohol more often (13.6 percent) during a search than Tier 2 agencies (5.8 percent).

Table 3.1.4.
Aggregate STOP Search Data, Year 2 Stops

	Tier 1			Tier 2		
Variable	**Traffic**	**Pedestrian**	**Total**	**Traffic**	**Pedestrian**	**Total**
Search Conducted	2.4%	14.5%	2.6%	2.3%	12.1%	2.8%
Search Justification						
Consent Search	49.3%	41.5%	48.3%	47.8%	52.4%	48.7%
"Other" Search	50.7%	58.5%	51.7%	52.2%	47.6%	51.3%
Successful in Total	50.6%	46.5%	50.1%	37.7%	27.3%	35.7%
Alcohol Found	14.2%	9.3%	13.6%	6.6%	2.3%	5.8%
Drugs Found	30.1%	26.6%	29.7%	25.0%	18.7%	23.8%
Weapons Found	6.0%	6.6%	6.1%	5.2%	5.1%	5.2%
Stolen Property Found	2.1%	6.1%	2.6%	1.4%	1.5%	1.4%
Other Evidence Found	9.5%	10.7%	9.6%	6.1%	3.6%	5.6%
Other Non-Evidence Found	2.8%	3.7%	2.9%	3.6%	4.1%	3.7%

3.2 COVID-19 Data Trends

Beginning in January of 2020, the U.S. response to COVID-19 had a significant impact on policing in both the U.S. and in Oregon. In mid-March, when the scale of the pandemic began to become clear, many law enforcement agencies began to curtail their discretionary stops. The Oregon State Police, for instance, provided a directive on March 13 to its troopers to stop violation-based traffic stops absent criminal level behavior. Many other agencies followed OSP's lead and began to curtail their stops as well, as displayed in **Figure 3.1.1**. The effects of the pandemic, though not universal across police agencies, can be seen most dramatically in March, April, and May 2020 stops.

Since Tier 1 agencies began collecting data in July 2019, there is more data on Tier 1 stops pre-COVID-19 than for Tier 2 agencies. In particular, comparisons of stops made in March, April, and May 2020, to potentially comparable months absent COVID-19—March, April, and May 2019—are possible for these twelve agencies. **Table 3.2.1** compares Tier 1 stop volume in March, April, and May 2019, respectively, to March, April and May 2020. All Tier 1 agencies except for Portland Police Bureau and Eugene PD stopped fewer individuals in March, April or May 2020 than in 2019. In April 2020, most Tier 1 agencies stopped at least 22 percent fewer

individuals than in April 2019. Gresham PD had the greatest percent reduction in stops of all Tier 1 agencies, as Gresham officers stopped 92 percent fewer individuals in April 2020 than in April 2019, and 87 percent fewer in May 2020 than in May 2019.

Table 3.2.1.
Stop Volume for Tier 1 Agencies in March, April, and May 2019 and 2020

Agency	March 2019 Stops	March 2020 Stops	Percent Change, From March 2019	April 2019 Stops	April 2020 Stops	Percent Change, From April 2019	May 2019 Stops	May 2020 Stops	Percent Change, From May 2019
Beaverton PD	2,128	954	-55.2%	1,711	543	-68.3%	1,726	938	-45.7%
Clackamas Co SO	2,846	1,449	-49.1%	2,104	1,005	-52.2%	2,154	3,185	47.9%
Eugene PD	991	1,887	90.4%	1,007	1,397	38.7%	902	1,638	81.6%
Gresham PD	667	411	-38.4%	956	80	-91.6%	794	106	-86.6%
Hillsboro PD	1,111	723	-34.9%	1,108	236	-78.7%	989	375	-62.1%
Marion Co SO	1,236	874	-29.3%	1,473	1,147	-22.1%	1,500	1,371	-8.6%
Medford PD	325	154	-52.6%	347	138	-60.2%	286	324	13.3%
Multnomah Co SO	773	681	-11.9%	1,077	467	-56.6%	935	754	-19.4%
Oregon State Police	20,147	10,098	-49.9%	19,239	6,720	-65.1%	19,796	14,218	-28.2%
Portland PB	3,074	3,079	0.2%	2,849	4,452	56.3%	3,144	4,561	45.1%
Salem PD	706	427	-39.5%	556	330	-40.6%	702	614	-12.5%
Washington Co SO	2,714	1,464	-46.1%	2,768	1,462	-47.2%	2,330	1,993	-14.5%

Like Tier 1 agencies, Tier 2 agencies did not respond uniformly to COVID-19. Most, but not all, Tier 2 agencies saw a reduction in stops during March or April and increased stops in May. For example, **Table 3.2.2** displays the five Tier 2 agencies with the greatest drop in March 2020 stops compared to the immediately preceding month, February 2020. Polk Co SO reduced stops by 71 percent in March, and stop volume dropped a further 80 percent from March to April. Stops then tripled in May 2020 and continued to climb into June. Tualatin PD stop volume also swiftly rebounded in May 2020, up from 45 stops in April to 392 in May. Lincoln Co SO, Lebanon PD, and Jackson Co SO's stop volume, in contrast, rebounded more slowly. Stop volume grew in May 2020, and then decreased slightly in June. Generally, agencies reduced discretionary stops in March, reduced further in April, and rebounded somewhat in May.

Table 3.2.2.
Tier 2 Agencies With the Greatest Drop in March 2020 Stops, Compared to February 2020

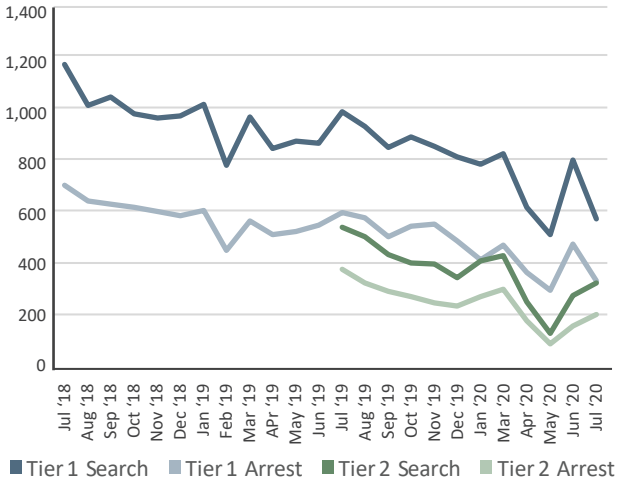
Agency	Month	Stops	Percent Change in Stops Compared to Preceding Month
Jackson Co SO	March-20	194	-66.0%
	April-20	238	22.7%
	May-20	492	106.7%
	Jun-20	299	-39.2%
Lebanon PD	March-20	30	-65.1%
	April-20	22	-26.7%
	May-20	46	109.1%
	Jun-20	25	-45.7%
Lincoln Co SO	March-20	111	-67.2%
	April-20	66	-40.5%
	May-20	107	62.1%
	Jun-20	92	-14.0%
Polk Co SO	March-20	149	-70.8%
	April-20	30	-79.9%
	May-20	133	343.3%
	Jun-20	169	27.1%
Tualatin PD	March-20	265	-62.0%
	April-20	45	-83.0%
	May-20	392	771.1%
	Jun-20	594	51.5%

While stop volume decreased, in general, citations also decreased, though not to the same extent. **Figure 3.2.1** displays stops and citations during Year 1 and Year 2 of data collection. In the Year 2 data, a higher proportion of Tier 1 agency stops ended in a citation than Tier 2 agency stops. Tier 1 and Tier 2 agencies both reduced stop and citation volume in March 2020. However, March and April stop volume dropped by a greater percentage than citations did. When stop volume reduced, the gap between stops and citations also narrowed.

Similarly, search and arrest volume also changed during the pandemic, as displayed in **Figure 3.2.2**. Tier 2 Arrests dropped 42 percent from February to March 2020, and dropped a further 49 percent in April. In contrast, Tier 1 arrest dropped only 25 percent from

Figure 3.2.1.
Stops and Citations

Figure 3.2.2. Searches and Arrests

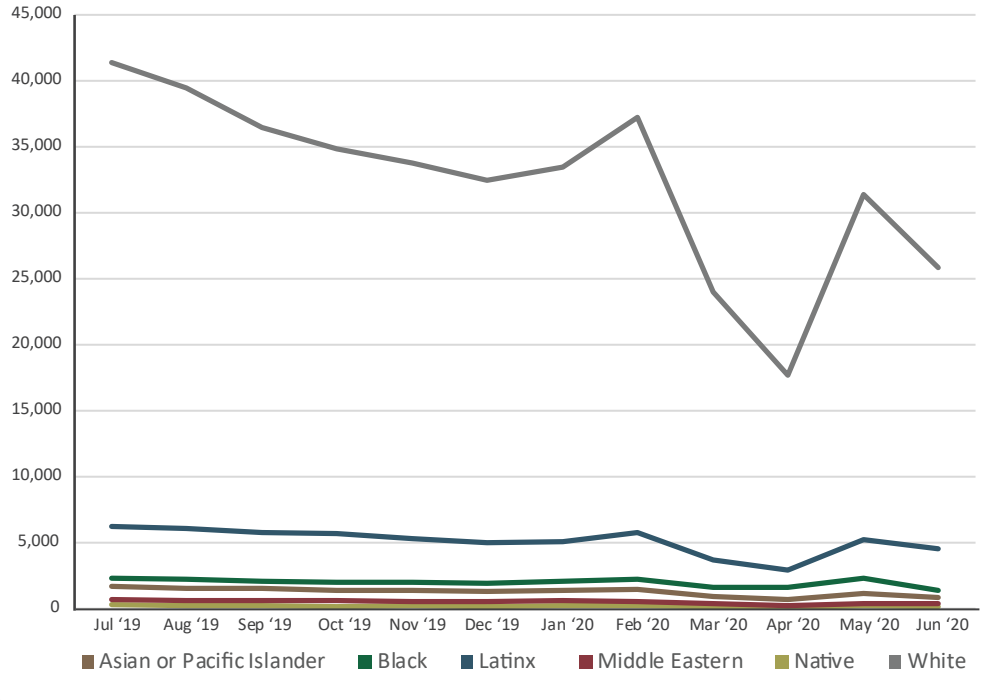


February to March 2020, and a further 17 percent in April. Similarly, Tier 2 searches fell 41 percent in March while Tier 1 searches fell 23 percent; Tier 2 searches then fell a further 51 percent in April, while Tier 1 searches dropped by an additional 19 percent. Searches and arrests dropped during March and April for most agencies, but dropped by a greater percentage for Tier 2 agencies compared to Tier 1 agencies.

Finally, in addition to the effect that COVID-19 has had on stop volume, the pandemic has also had an impact on the driving population and therefore the rates at which different individuals and groups were stopped during 2020. Research reports that white workers are more likely to be able to work remotely and that non-white individuals are overrepresented in "essential" occupations. According to the Bureau of Labor Statistics, for example, white workers are 50 percent more likely to have the privilege of working from home compared to Black workers.[18] News reports in Oregon suggest similar patterns, as Black and Latinx Oregonians are overrepresented in essential industries and occupations.

While the overrepresentation of people of color in essential industries and occupations poses a significant risk for their health and wellbeing, it also means that it is possible that their share of the driving population also increased during the pandemic relative to white Oregonians. Unfortunately, data concerning driving populations does not exist in the best of times and certainly does not exist at a systematic, statewide level during the COVID-19 pandemic. Thus, the best proxy the STOP Research Team can use is to examine stop rates immediately preceding and following the beginning of the pandemic based on the assumption that patterns during this short period of time would reflect changes in driving populations, not coincidental changes in enforcement practices.

Figure 3.2.3. Stop Volume by Race



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

Figure 3.2.3 displays stops from July 2019 through June 2020, by race. Stop volume dropped for all demographic groups in March and April 2020, though not uniformly. In February 2020, Tier 1 and Tier 2 agencies stopped 37,252 white individuals. In March 2020, agencies stopped 23,984 white individuals, a 35.6 percent drop. Similarly, Latinx stops dropped 36 percent, Asian/PI stops dropped 38.5 percent, Middle Eastern stops dropped by 35.3 percent, and Native American stops dropped 32.3 percent. However, over the same time period, stops for Black individuals fell by only 27 percent, from 2,242 to 1,637 stops. For all demographic groups, stops dropped further in April. Stops for white individuals dropped by 26.3 percent from March to April, but dropped only a further 3.7 percent for Black individuals. Though stop volume reduced overall during the early days of the COVID-19 pandemic, stops of Black individuals did not reduce as much as other racial groups.

In all, the COVID-19 pandemic and the response to the dangers posed by the disease by law enforcement had a significant impact on the STOP data collection for the 2019-2020 reporting year. As discussed in this section, overall, stops were down substantially, although this effect was not universal. Due to the significant changes in policy ushered in by the pandemic, as well as the likely changes in the driving population due to the unequal distribution of telework opportunities across different racial and ethnic groups, the STOP Program would strongly caution readers of this report against making direct comparisons of year-to-year results in the models that follow.

4. Veil of Darkness Analysis

Often referred to as the "gold standard" of statistical analyses examining the initial law enforcement decision to stop an individual,[19] the Veil of Darkness (VOD) 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 VOD 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 VOD 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 VOD 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 VOD analysis can compare stops made on January 10, 2020 when it was dark at 5:00pm to stops made two months later at the same time on March 10, 2020, 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 VOD 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 VOD 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 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 aggregate, Tier level, or for the Black or Latinx alone groups at the agency level.

4.1.Aggregate Veil of Darkness Analysis for All Tier 1 Agencies

At the statewide level, it is possible to estimate VOD models for all of the non-white groups reported in the stop database. First, **Table 4.1.1** displays the odds ratios for the aggregate Tier 1 and Tier 2 VOD models for all non-white stopped drivers, including those perceived as Black, Latinx, Asian/PI, Middle Eastern, and Native

19 See Barone et al. (2018).

American, compared to white stopped drivers. For the full Tier 1 and Tier 2 models, all comparisons show no statistically significant differences in the odds of minority stops in daylight compared to darkness. Further, the coefficients for nearly all comparisons are less than or close to 1.0. Taken together, this indicates that disparities were not detected at a statewide level.

Table 4.1.1.
Logistic Regression of Minority Status on Daylight for All Tier 1 & 2 Agencies

	Asian/PI	Black	Latinx	Middle Eastern	Native American
Tier 1	0.91	1.04	0.97	0.90	0.71
Tier 1 (no OSP)	0.88	1.01	1.02	0.86	0.73
Tier 2	0.99	1.07	1.01	0.95	1.75

NOTES: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 5 Comparisons) Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, county stop volume, agency stop volume, and agency fixed effects.

To check the robustness of the statewide finding reported in the first row of **Table 4.1.1**, an additional analysis was conducted. In the data reported by law enforcement for the second year of data collection, stops made by the Oregon State Police accounted for over half of all traffic stops for Tier 1 agencies. Due to this imbalance, it is possible that patterns in Oregon State Police stop rates could mask disparities occurring in the other eleven agencies when aggregated to the state level. Due to this concern, a VOD model was estimated for all stops other than those made by the Oregon State Police. The results of this robustness check, reported in the second row of **Table 4.1.1**, demonstrate that even with the Oregon State Police removed from the analysis, there is no evidence of a racial disparity at the state level for non-white drivers.

Table 4.1.2
Logistic Regression of Minority Status on Daylight for All Tier 1 & 2 Agencies Restricted to DST Window

	Asian/PI	Black	Latinx	Middle Eastern	Native American
Tier 1	0.85	1.03	0.95	0.98	0.62
Tier 1 (no OSP)	0.76	0.96	0.86	0.88	--
Tier 2	0.85	1.29	0.98	0.81	--

NOTES: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 5 Comparisons) Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, county stop volume, agency stop volume, and agency fixed effects.

An additional robustness check for the VOD analysis is possible when a sufficient number of data points are found in a given dataset. As discussed above, one assumption of the VOD model is that driving behavior and patterns does not vary throughout the year. It is possible, however, that patterns in the summer, for instance, are different than those found in the winter. To account for this possibility, the VOD analysis can be restricted to stops made during the 30-day windows before and after the fall and spring Daylight Savings Time (DST) dates. This ensures that comparisons of stops in daylight and darkness are made between dates that are relatively close to one another, which narrows the

possibility that variation in driving behavior and driving patterns could impact the results of the statistical test. The results of these models are reported in **Table 4.1.2** for Black, Latinx, Asian/PI, Middle Eastern, and Native American drivers.[20] Similar to the results presented in **Table 4.1.1**, above, the results for the restricted sample show no evidence of a statistically significant difference in stops during the daylight compared to those at night.

4.2. Agency-level Veil of Darkness Analysis

While the aggregate, state level results indicate that there is no evidence of a statistically significant disparity in outcomes for any minority group compared to white individuals, it is still possible that analyses of individual law enforcement agencies could detect disparities at a disaggregated level. As such, VOD models were estimated for Tier 1 agencies for Black and Latinx drivers. The sample sizes were insufficient to estimate the VOD models for Black drivers for Marion Co. SO, Medford PD, and Salem PD. In addition, VOD models were estimated for 22 Tier 2 agencies (from a total of 39 agencies) for Latinx drivers. Springfield PD was the only Tier 2 agency with a large enough sample size to estimate models for both Black and Latinx drivers. As described in Section 2, the sample size requirement for the VOD model was at least 100 stops in each race/ethnicity group within the inter-twilight windows.

Table 4.2.1
Logistic Regression of Minority Status on Daylight by Tier 1 Agency

Agency	Black	Latinx
Beaverton PD	1.28	0.91
Clackamas Co SO	1.21	1.05
Eugene PD	0.96	1.08
Gresham PD	1.23	1.02
Hillsboro PD	0.92	1.10
Marion Co SO	--	0.82
Medford PD	--	1.01
Multnomah Co SO	1.05	1.43
Oregon State Police	1.08	0.93
Portland PB	0.99	1.09
Salem PD	--	0.88
Washington Co SO	0.65	1.00

NOTES: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 2 Comparisons) Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, county stop volume, and agency stop volume.

Table 4.2.2
Logistic Regression of Minority Status on Daylight by Tier 2 Agency

Agency	Black	Latinx	Agency	Black	Latinx	Agency	Black	Latinx
Albany PD	--	1.36	Jackson Co SO	--	0.92	Polk Co SO	--	0.97
Bend PD	--	1.05	Keizer PD	--	1.86	Springfield PD	0.94	0.57
Benton Co SO	--	1.14	Klamath Falls PD	--	1.05	Tigard PD	--	1.14
Canby PD	--	0.66	Lincoln City PD	--	0.91	Tualatin PD	--	0.66
Deschutes Co SO	--	1.32	Linn Co SO	--	0.82	West Linn PD	--	1.36
Forest Grove PD	--	0.89	McMinnville PD	--	0.93	Woodburn PD	--	0.77
Hermiston PD	--	0.80	Milwaukie PD	--	1.66	Yamhill Co SO	--	1.23
Hood River Co SO	--	1.29	Newberg-Dundee PD	--	0.81			

NOTES: * p<0.05, ** p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 2 Comparisons) Logistic regression results include controls for age, gender, reason for stop, day of week, time of day, quarter or season, county stop volume, and agency stop volume.

20 To conduct a VOD analysis limited to the 30-day windows around daylight savings time changes, the sample size must be sufficiently large. Unfortunately, at this time, sample sizes were not sufficient to make comparisons of Native American drivers to white drivers.

Table 4.2.1 reports the Tier 1 agency specific model results for Black or Latinx drivers compared to white drivers. While a number of agencies have odds ratios above 1.0, no agency had an odds ratio that was above 1.0 and statistically significant. **Table 4.2.2** reports the Tier 2 agency specific model results for Black or Latinx drivers compared to white drivers. The results are similar and no agency had an odds ratio that was above 1.0 and statistically significant. While some Tier 2 agencies show a higher odds ratio, the estimate is not statistically significantly different from 1.0 and does not indicate a disparity at this time. Thus, taken together, the STOP Program did not find evidence of racial/ethnic disparities in the Veil of Darkness analyses.[21]

4.3. Results for Stops in the Inter-Twilight Window

The VOD models described in the previous section reported estimates for agencies with at least 100 stops in each race/ethnicity group within their inter-twilight windows. Several agencies had an insufficient sample size to estimate the full VOD models. This section displays results for a bivariate comparison of stops in daylight vs. darkness compared to the race of the driver. These analyses are not as robust or rigorous as the full VOD models, but provide results to compare the percentage of stops in daylight vs. darkness in the combined inter-twilight window. Agencies with at least 30 stops for each race/ethnicity group have results displayed in the following tables.

Table 4.3.1 reports results for the three Tier 1 agencies that had an insufficient sample size for the full VOD models for Black drivers. Consistent with the findings reported across agencies in the VOD analyses, the share of Black drivers was

21 The p-values for the agency-level Veil of Darkness models were examined to ensure that the rigorous threshold set by the inclusion of the Bonferroni correction did not result in the under identification of agencies. In no instance does the Bonferroni correction lead to an agency not being identified that otherwise would have had a statistically significant finding at 95 percent confidence level absent the Bonferroni correction.

Table 4.3.1.
Results for Stops of Black Individuals - Tier 1

Agency	% Dark	% Light
Marion Co SO	3.5%	3.0%
Medford PD	5.1%	5.5%
Salem PD	5.3%	3.2%

Table 4.3.2
Results for Stops of Black Individuals - Tier 2

Agency	% Dark	% Light	Agency	% Dark	% Light	Agency	% Dark	% Light
Albany PD	1.9%	2.2%	Jackson Co SO	--	--	OHSU PD	--	--
Ashland PD	--	--	Keizer PD	--	--	Oregon City PD	3.6%	3.3%
Bend PD	--	--	Klamath Co SO	--	--	Polk Co SO	3.5%	3.8%
Benton Co SO	2.8%	3.6%	Klamath Falls PD	4.8%	4.5%	Port of Portland PD	14.1%	12.4%
Canby PD	--	--	Lake Oswego PD	5.0%	4.1%	Redmond PD	--	--
Central Point PD	--	--	Lane Co SO	--	--	Roseburg PD	4.5%	2.7%
Corvallis PD	--	--	Lebanon PD	--	--	Tigard PD	7.5%	5.2%
Deschutes Co SO	--	--	Lincoln City PD	--	--	Tualatin PD	5.2%	3.3%
Douglas Co SO	--	--	Lincoln Co SO	--	--	U of O PD	--	--
Forest Grove PD	2.5%	2.8%	Linn Co SO	--	--	West Linn PD	3.1%	4.8%
Grants Pass DPS	--	--	McMinnville PD	--	--	Woodburn PD	--	--
Hermiston PD	--	--	Milwaukie PD	4.9%	6.6%	Yamhill Co SO	--	--
Hood River Co SO	--	--	Newberg-Dundee PD	2.7%	3.1%			

higher in the dark than during daylight for Marion Co. SO and Salem PD. For Medford PD, more Black drivers were stopped during the daylight, but this difference was not statistically significant.

Table 4.3.2 displays similar bivariate results for Tier 2 agencies for Black drivers, and **Table 4.3.3** displays results for Tier 2 agencies for Latinx drivers. While these analyses are not as robust as the full VOD models in the previous section, no agency shows a significant difference in the percentage of stops in daylight compared to darkness.

Table 4.3.3.
Results for Stops of Latinx Individuals - Tier 2

Agency	% Dark	% Light	Agency	% Dark	% Light
Ashland PD	6.0%	7.0%	Lebanon PD	--	--
Central Point PD	12.3%	11.2%	Lincoln Co SO	7.3%	7.5%
Corvallis PD	9.6%	8.2%	OHSU PD	--	--
Douglas Co SO	4.2%	4.8%	Oregon City PD	8.8%	6.3%
Grants Pass DPS	6.2%	5.4%	Port of Portland PD	16.2%	13.8%
Klamath Co SO	--	--	Redmond PD	8.9%	10.1%
Lake Oswego PD	6.3%	6.4%	Roseburg PD	5.4%	5.7%
Lane Co SO	8.9%	5.6%	U of O PD	--	--

5. Predicted Disposition 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 Predicted Disposition analysis, is presented in this section and 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. **Table 5.0.1.a and Table 5.0.1.b** report the percentages of dispositions broken down by agency and demonstrates that there is considerable variation across agencies examined in this report with regard to the share of drivers cited, searched, and arrested. While most agencies report similar arrest rates, some agencies use warnings or issue citations more often than other agencies. The Marion Co. SO, for example, cites a higher share of drivers than all other Tier 1 agencies (80.8 percent), while the next closest rate is 16 percentage points lower with Bend PD (64.2 percent) and most agencies cite below 50 percent of stopped individuals. Similarly, Springfield PD, Albany PD, and University of Oregon PD all have relatively high arrest rates (above 6.5 percent) and search rates (above 4.5 percent) when compared to other agencies.

Table 5.0.1a
Aggregate Tier 1 and Agency-level Stop Dispositions

Agency	Nothing/ Warning	Citation	Search	Arrest
Beaverton PD	59.1%	34.3%	3.8%	5.9%
Clackamas Co SO	64.6%	32.1%	1.6%	3.1%
Eugene PD	59.0%	35.9%	3.8%	4.7%
Gresham PD	54.9%	41.6%	1.9%	3.0%
Hillsboro PD	69.5%	27.9%	1.3%	2.4%
Marion Co SO	17.7%	80.8%	0.9%	1.4%
Medford PD	63.2%	30.9%	4.3%	5.2%
Multnomah Co SO	77.3%	18.5%	2.4%	3.8%
Oregon State Police	61.9%	36.1%	1.0%	1.7%
Portland PB	47.3%	47.0%	1.8%	4.0%
Salem PD	41.7%	52.9%	4.0%	5.2%
Washington Co SO	72.7%	23.6%	2.0%	3.2%
Tier 1 - All	59.2%	37.5%	1.6%	2.8%
Tier 1 (No OSP)	56.7%	38.9%	2.3%	4.0%

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 as well. 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. 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 that 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 all else 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

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

Table 5.0.1b
Aggregate Tier 1 and Agency-level Stop Dispositions

Agency	Nothing/ Warning	Citation	Search	Arrest	Agency	Nothing/ Warning	Citation	Search	Arrest
Albany PD	60.5%	32.0%	5.6%	7.0%	Lincoln City PD	76.3%	22.3%	0.8%	1.3%
Ashland PD	66.8%	31.5%	1.2%	1.7%	Lincoln CO SO	80.2%	18.7%	0.7%	1.0%
Bend PD	31.4%	64.2%	3.0%	4.1%	Linn CO SO	62.7%	36.0%	0.9%	1.2%
Benton CO SO	76.6%	22.0%	0.8%	1.0%	McMinnville PD	75.2%	23.0%	1.1%	1.4%
Canby PD	81.0%	15.2%	2.5%	3.4%	Milwaukie PD	52.9%	46.4%	0.4%	0.5%
Central Point PD	83.8%	9.2%	3.2%	4.7%	Newberg-Dundee PD	75.2%	24.0%	0.4%	0.7%
Corvallis PD	74.5%	24.7%	0.5%	0.6%	Oregon City PD	56.2%	41.1%	1.6%	2.6%
Deschutes CO SO	79.4%	17.6%	1.0%	1.6%	OHSU PD	93.2%	4.5%	1.4%	1.4%
Douglas CO SO	69.1%	26.0%	2.9%	4.6%	Polk CO SO	76.3%	20.7%	2.2%	2.6%
Forest Grove PD	69.9%	27.3%	0.9%	2.7%	Port of Portland PD	81.7%	15.6%	1.8%	2.4%
Grants Pass DPS	66.1%	29.7%	3.6%	3.9%	Redmond PD	75.0%	23.3%	0.6%	1.4%
Hermiston PD	65.1%	32.8%	1.1%	2.0%	Roseburg PD	39.9%	52.1%	2.9%	7.5%
Hood River CO SO	80.1%	18.4%	0.9%	1.4%	Springfield PD	64.4%	28.4%	4.5%	6.5%
Jackson CO SO	45.4%	52.6%	1.2%	1.9%	Tigard PD	71.8%	26.0%	1.6%	2.1%
Keizer PD	69.2%	30.2%	0.1%	0.5%	Tualatin PD	52.3%	45.9%	0.8%	1.7%
Klamath CO SO	72.9%	24.8%	0.9%	1.8%	U of O PD	81.9%	10.0%	5.9%	6.9%
Klamath Falls PD	69.1%	27.8%	2.2%	2.5%	Washington CO SO	72.7%	23.6%	2.0%	3.2%
Lake Oswego PD	47.0%	52.3%	0.1%	0.5%	West Linn PD	81.5%	18.1%	0.2%	0.4%
Lane CO SO	67.6%	29.8%	1.5%	2.4%	Woodburn PD	65.3%	29.6%	3.3%	4.5%
Lebanon PD	65.8%	31.7%	1.1%	2.2%	Yamhill CO SO	75.5%	23.2%	0.4%	1.0%

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 the data available, the sample size, the completeness of the data, and other factors; there is no one-size-fits-all approach. Here the STOP Program employed Inverse Probability Weighted Regression Adjustment.[23]

23 IPWRA 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 D.

24 Discussions with some law enforcement agencies indicated that the inclusion of warnings in the same category as "none" did not necessarily match their conceptualization of warnings as an enforcement device. In the opinion of these agencies, a warning constitutes an action, albeit one that is not as serious as a citation.

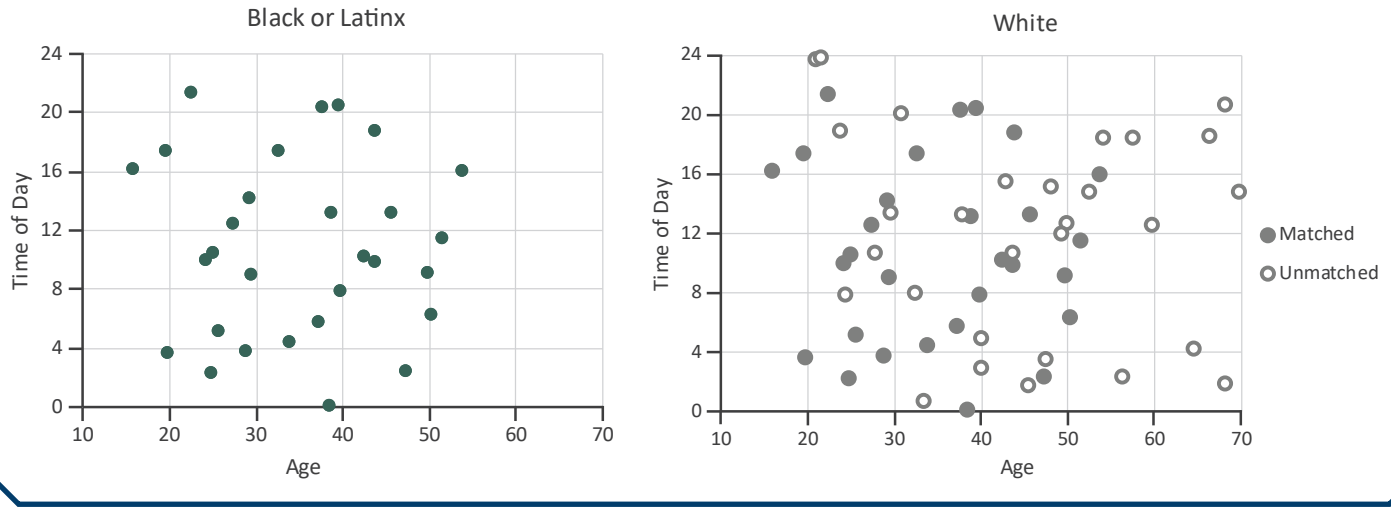
Table 5.0.2
Analyses Completed for Each Agency

Disposition of Interest	Comparison Dispositions[24]	Analysis Groups
Citation	None or Warning	Black Latinx
Search	None, Warning, or Citation	Black Latinx
Arrest	None, Warning, Citation, or Search	Black Latinx
Citation, Search, or Arrest	None or Warning	Black Latinx

The current analysis included eight sub-analyses for each agency: each outcome of citation, search, arrest, or any non-warning disposition across each racial/ethnic group of Black and Latinx individuals (**Table 5.0.2**). With higher sample sizes when aggregating to the tier level, these groups were expanded to include Black, Latinx, Asian/Pacific Islander, 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 5.0.2** describes the two sets of tests conducted for each agency. In row 1, STOP Program researchers tested whether there was a disparity in issuing citations between the Black and Latinx groups, respectively, and a matched white group.[25] Row 2 does the same for searches, row 3 for arrests, and row 4 describes tests for any Citation, Search, or Arrest disposition.

In each of the eight tests for each agency, the results compare the actual, observed rate of the dispositional outcome for the minority group with an estimate of the rate of the disposition for the best possible comparable white group. A stylized example of this process is presented in **Figure 5.0.1**. Here we present a hypothetical group of Black or Latinx individuals on the left and the white group on the right, across age of the stopped individual and time of day of the stop. For the white Group, the solid dots represent the best comparison group across these factors, whereas the hollow observations represent observations that are relatively poor comparisons based on the age and time of stop variables. The STOP Program research team created comparison groups across variables including perceived gender, age, time of day, day of week, reason for stop, season, daylight, agency stop volume, and county stop volume.

Figure 5.0.1.
Hypothetical Representation of the Best Possible Comparison White Group



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

5. Predicted Disposition Analysis

5.1. Aggregate Predicted Disposition Results for All Agencies by Tier

Similar to the Veil of Darkness results presented previously, the STOP Program identifies all agencies with disparities in their predicted versus actual outcomes where those differences were statistically significant at the 95 percent confidence level in any minority group at tier level or for the Black or Latinx groups, respectively, at the agency level. **Table 5.1.1** reports the results of models examining the actual and predicted disposition rates for Tier 1 and Tier 2 agencies combined, respectively.[26] For each outcome, which includes receiving a citation, being searched, being arrested, or falling into any one of these categories, the actual versus predicted outcomes are broken down by minority group. Importantly, for this table and for those that follow, only the results that reached statistical significance at the 95 percent confidence level are reported.[27]

As shown in **Table 5.1.1**, there were disparate outcomes for the Latinx group with regard to citations, searches, and arrests. With regard to citations in the "Tier 1 – All" group, for example, the propensity score models predicted that 39.0 percent of Latinx individuals should have been cited based on the comparable white group, while the actual share of Latinx individuals cited was about 6 percentage points higher. Disparities were also found for the Black and Native American groups, but the significance level varied across disposition type and tier group. Among Native Americans in the "Tier 1 – All" analysis group, for instance, the propensity score

26 For all the results tables in this section, only the results that reached statistical significance at the 95 percent confidence level are reported—full results are available upon request.

27 Each of the eight tests is compared to a more stringent Bonferroni adjusted level that is conditional on conducting eight tests at once, at the 95 percent confidence level. See Appendix D for a more thorough discussion of this adjustment and the full table of results.

Table 5.1.1
Predicted Disposition Analysis Results by Tier

Agency	Race/ Ethnicity	Citation		Search		Arrest		Citation, Search, or Arrest	
		Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
Tier 1	Asian/PI	39.6%	37.9%	--	--	--	--	--	--
	Black	--	--	2.6%	2.2%	4.6%	3.8%	--	--
	Latinx	45.0%	39.0%	2.1%	1.8%	3.4%	2.9%	47.3%	41.2%
	Middle Eastern*	--	--	--	--	--	--	--	--
	Nat. American	37.6%	34.2%	--	--	4.5%	3.1%	41.0%	36.6%
Tier 1 No OSP	Asian/PI	--	--	--	--	--	--	--	--
	Black	--	--	3.1%	2.4%	5.4%	4.2%	--	--
	Latinx	44.0%	40.3%	2.9%	2.2%	4.4%	3.6%	47.0%	42.8%
	Middle Eastern	--	--	--	--	--	--	--	--
	Nat. American	--	--	--	--	--	--	--	--
Tier 2	Asian/PI	--	--	--	--	--	--	--	--
	Black	32.5%	30.0%	--	--	--	--	34.9%	32.0%
	Latinx	36.8%	31.2%	--	--	--	--	38.3%	32.8%
	Middle Eastern	--	--	--	--	--	--	--	--
	Nat. American	--	--	--	--	--	--	--	--

* Eugene PD is omitted from the Middle Eastern analysis since this agency does not log a Middle Eastern race/ethnicity

models predicted that arrests should have occurred in 3.1 percent of stops based on the most comparable white group, while the actual share of Native Americans arrested was 4.5 percent. Notably, disparities were identified for the Black and Latinx groups for at least one dispositional outcome for each of the Tier 1, Tier 1 omitting Oregon State Police, and the Tier 2 agency analysis groups.

5.2. Agency-level Predicted Disposition Results

While the aggregate results for the Tier 1 agencies indicate that there is evidence of disparities in dispositional outcomes, it is necessary to examine outcomes individually for each agency, as the policies of individual agencies and the training received by officers following their graduation from DPSST differs from one agency to another. Agency-level results for Tier 1 agencies are presented in **Table 5.2.1**. Ten Tier 1 law enforcement agencies report statistically significant disparities for the Predicted Disposition analysis. For five agencies, Clackamas Co. SO, Eugene PD, Hillsboro PD, Multnomah Co. SO, and Oregon State Police, 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 it is likely for these agencies that the only relevant disparity is for citations and not the other outcomes. For five agencies, Beaverton PD, Marion Co. SO, Portland PB, Salem PD, and the Washington Co. SO, disparities were reported for either searches and/or arrests, sometimes in addition to citations.

Where disparities were found, the average gap in the predicted versus the actual disposition rate varied by the type of disposition. For the disparities in citations, for example, the differences between predicted and actual citation rate averaged 18.7 percent (ranging from 3.3 – 36.6%).[28] Comparatively, the three agencies that showed disparities for searches showed rates that were, on average, 46.7% higher than predicted. From this perspective, the percent difference in rates for searches and arrests represents a larger disparity even though these are lower overall percentages.

Table 5.2.2 presents similar results, but for the Tier 2 agencies reporting their first year of data. As described in Sections 2 and 3, Tier 2 agencies have far fewer stops than Tier 1 agencies. Combined with the already relatively low minority population in the state many of the predicted disposition analyses for the Tier 2 agencies did not have sufficient sample sizes to complete the analysis. That said, Albany PD, Forest Grove PD, Keizer PD, Lincoln Co. SO, McMinnville PD, Newberg Dundee PD, Polk Co. SO, Port of Portland PD, Redmond PD, Roseburg PD, Tigard PD, Tualatin PD, West Linn PD, and Yamhill Co. SO had statistically significant disparities indicated for either the Black and Latinx groups for Citations and Any Disposition. Only Woodburn PD indicated a disparity for searches or arrests among the Tier 2 agencies, indicating that Latinx individuals are about 4 times more likely to be searched than comparable white counterparts, controlling for all covariates.

As indicated elsewhere in this report as well, limited sample sizes were a significant barrier to estimation for many of the Tier 2 agencies: in aggregate across all Tier 2 agencies, we were unable to estimate any models for the Native American group. Regarding the agency-specific models, estimates were unattainable in all of the baseline models for Corvallis PD, Lebanon PD, and University of Oregon PD. Further, the models focused specifically on Black individuals could not be estimated for Ashland PD, Canby PD, Central Point PD, Hood River Co. SO, Keizer PD, Klamath Co. SO, Lincoln Co. SO, Linn Co. SO, Springfield PD, and Woodburn PD. In all these cases, a lack of indicated disparity should not be interpreted as proof-positive that there is no disparity for these groups in these jurisdictions. STOP analysts in these instances were just unable to estimate the models with current data limitations. In future iterations of this report multiple years of data will be combined so that more of these models may be estimated. This will be discussed further in the concluding section of this report.

28 To calculate the percent difference for these comparisons, $\Delta\% = ((Actual\text{-}Predicted)/Predicted)\times 100$.

Table 5.2.1.
Predicted Disparity, Tier 1 Agencies, by Agency and Disposition (Statistically Significant Results)

Agency	Race/ Ethnicity	Citation		Search		Arrest		Citation, Search, or Arrest	
		Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
Beaverton PD	Black	--	--	--	--	--	--	--	--
	Latinx	37.6%	34.7%	--	--	7.2%	5.8%	43.0%	38.9%
Clackamas Co SO	Black	33.9%	29.5%	--	--	--	--	36.9%	32.4%
	Latinx	35.7%	32.0%	--	--	--	--	37.8%	34.4%
Eugene PD	Black	--	--	--	--	--	--	--	--
	Latinx	45.6%	36.5%	--	--	--	--	47.8%	39.7%
Gresham PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Hillsboro PD	Black	--	--	--	--	--	--	--	--
	Latinx	35.5%	26.0%	--	--	--	--	37.7%	28.0%
Marion Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	80.3%	77.7%	--	--	2.9%	1.9%	80.9%	78.2%
Medford PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Multnomah Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	23.1%	18.9%	--	--	--	--	26.0%	22.1%
Oregon State Police	Black	45.5%	37.5%	--	--	--	--	47.0%	39.2%
	Latinx	46.3%	37.2%	--	--	--	--	47.8%	38.7%
Portland PB	Black	--	--	2.9%	1.9%	6.1%	4.3%	--	--
	Latinx	--	--	--	--	--	--	57.4%	55.1%
Salem PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	6.5%	4.6%	7.7%	5.8%	--	--
Washington Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	28.2%	23.4%	3.1%	2.1%	--	--	32.1%	26.6%

Table 5.2.2.
Predicted Disparity, Tier 1 Agencies, by Agency and Disposition (Statistically Significant Results)

Agency	Race/ Ethnicity	Citation		Search		Arrest		Citation, Search, or Arrest	
		Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
Albany PD	Black	--	--	--	--	--	--	--	--
	Latinx	48.1%	40.2%	--	--	--	--	45.5%	39.0%
Ashland PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Bend PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Benton Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Canby PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Central Point PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Corvallis PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Deschutes Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Douglas Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Forrest Grove PD	Black	--	--	--	--	--	--	--	--
	Latinx	35.8%	25.8%	--	--	--	--	37.7%	27.8%
Grants Pass DPS	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Hermiston PD	Black	--	--	--	--	--	--	--	--
	Latinx	37.9%	32.2%	--	--	--	--	39.5%	33.7%
Hood River Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Jackson Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Keizer PD	Black	--	--	--	--	--	--	--	--
	Latinx	36.6%	29.4%	--	--	--	--	37.1%	29.9%
Klamath Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Klamath Falls PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Lake Oswego PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--

5. Predicted Disposition Analysis

Agency	Race/ Ethnicity	Citation		Search		Arrest		Citation, Search, or Arrest	
		Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
Lane Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Lebanon PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Lincoln City PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Lincoln Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	29.6%	19.2%					30.2%	20.2%
Linn Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	48.9%	37.0%
McMinnville PD	Black	--	--	--	--	--	--	--	--
	Latinx	31.8%	22.5%	--	--	--	--	33.0%	23.8%
Milwaukie PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Newberg-Dundee PD	Black	--	--	--	--	--	--	--	--
	Latinx	29.3%	21.8%	--	--	--	--	30.4%	22.5%
OHSU PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Oregon City PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Polk Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	30.5%	21.1%	--	--	--	--	33.6%	23.9%
Port of Portland PD	Black	22.9%	14.6%	--	--	--	--	25.9%	17.5%
	Latinx	--	--	--	--	--	--	--	--
Redmond PD	Black	--	--	--	--	--	--	--	--
	Latinx	36.9%	23.6%	--	--	--	--	37.8%	25.0%
Roseburg PD	Black	77.0%	60.3%	--	--	--	--	78.2%	64.7%
	Latinx	--	--	--	--	--	--	--	--
Springfield PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
Tigard PD	Black	--	--	--	--	--	--	--	--
	Latinx	34.6%	26.4%	--	--	--	--	35.8%	28.1%
Tualatin PD	Black	--	--	--	--	--	--	--	--
	Latinx	49.3%	44.1%	--	--	--	--	50.8%	45.0%
U of O PD	Black	--	--	--	--	--	--	--	--
	Latinx	--	--	--	--	--	--	--	--
West Linn PD	Black	--	--	--	--	--	--	--	--
	Latinx	27.7%	19.7%	--	--	--	--	28.2%	20.1%
Woodburn PD	Black	--	--	--	--	--	--	--	--
	Latinx	34.6%	24.9%	4.5%	1.1%	--	--	38.7%	26.9%
Yamhill Co SO	Black	--	--	--	--	--	--	--	--
	Latinx	31.5%	25.0%	--	--	--	--	32.7%	26.0%

6. KPT Hit-Rate Analysis

The second analysis conducted for this report examining post-stop outcomes is the KPT Hit-Rate test. 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 examining outcomes to identify discrimination has been adapted extensively in analyses of policing, and the most widely used model is the KPT Hit-Rate model developed by Knowles, Persico, and Todd (2001).

The Knowles, Persico, and Todd (KPT) Hit-Rate model examines whether the likelihood of a "successful" police search differs across racial/ethnic groups, where success is defined as finding contraband. The KPT 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 hit-rates are the same across all groups. However, if officers are subjecting a group to more frequent searches based on racial bias, then their hit-rate for that group will decrease. If a minority group's hit-rate is less than the white hit-rate, therefore, 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 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. Statewide search data were analyzed for disparities between the white baseline group and individuals identified as Black, Latinx, Asian/PI, Middle Eastern, and Native American. Due to sample size limitations, agency-specific hit-rates were only calculated for white individuals compared to Black individuals and Latinx individuals. For certain agencies and racial/ethnic groups, the Hit-Rate analysis was unable to be performed, because 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.[29] 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 KPT Hit-Rate analysis. At the aggregate Tier 1 and Tier 2 level, this includes any minority search hit-rate below the white hit-rate and statistically significant at the 95 percent confidence level. For individual agencies, this includes Black or Latinx alone hit-rates less than the white hit-rate and statistically significant at the 95 percent confidence level. See Appendix E for more detailed technical information about the KPT Hit-Rate model and statistical tests.

6.1 Aggregate KPT Hit-Rate Results for all Tier 1 Agencies

Table 6.1.1 presents KPT Hit-Rate results for all Tier 1 agencies combined. As shown in the table, the hit-rate for the Black group was determined to be significantly lower than the white hit-rate at the 99.9% confidence level, indicating a disparity. There was no evidence that statewide Tier 1 hit-rates for other groups, Latinx, Asian/PI, Middle Eastern, or Native American, were statistically significantly different from the white hit-rate.

29 Connecticut Racial Profiling Prohibition Project (2019).

Table 6.1.1.
Statewide Tier 1 Hit-Rates and Significance by Race/Ethnicity

Race/Ethnicity	Minority Hit Rate	White Hit Rate	p-value (Significance)
Asian/PI	0.463	0.536	0.066
Black	0.443	0.536	0.000***
Latinx	0.502	0.536	0.021
Middle Eastern	0.426	0.536	0.172
Nat. American	0.547	0.536	0.933

* p<0.05, **p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 5 Comparisons)

As a robustness check, because searches by Oregon State Police make up 33 percent of all searches across Tier 1 agencies, statewide hit-rates were tested both including (as displayed above) and removing Oregon State Police searches. Removing Oregon State Police searches did not substantively change the results, as shown in **Table 6.1.2**, although the overall hit rate fell slightly across the board, and there were not enough searches to conduct an analysis of Middle Eastern hit-rates.

Table 6.1.2
Statewide Tier 1 (Not including OSP) Hit-Rates and Significance by Race/Ethnicity

Race/Ethnicity	Minority Hit Rate	White Hit Rate	p-value (Significance)
Asian/PI	0.423	0.473	0.273
Black	0.396	0.473	0.000***
Latinx	0.456	0.473	0.340
Middle Eastern	0.379	0.473	(no analysis)
Nat. American	0.421	0.473	0.636

* p<0.05, **p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 5 Comparisons)

Table 6.1.3 presents KPT Hit-Rate results for all Tier 2 agencies combined. While for both the Middle Eastern and Native American groups there were too few searches to perform an analysis, three analyses (for Black, Latinx, and Asian/PI hit-rates) were performed. None of these analyses resulted in statistically significant differences from the statewide white hit-rate for all Tier 2 agencies combined.

Table 6.1.3.
Statewide Tier 2 Hit-Rates and Significance by Race/Ethnicity

Race/Ethnicity	Minority Hit Rate	White Hit Rate	p-value (Significance)
Asian/PI	0.246	0.381	0.031
Black	0.362	0.381	0.650
Latinx	0.388	0.381	0.798
Middle Eastern	0.364	0.381	(no analysis)
Nat. American	0.238	0.381	(no analysis)

* p<0.05, **p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 5 Comparisons)

6.2 Agency-level KPT Hit-Rate Results

While the aggregate results for the Tier 1 agencies indicate that there is evidence of a disparity in outcomes for the Black group, it is necessary to examine outcomes by individual agency, as the policies of individual agencies and the training received by officers likely differs from one agency to another. As discussed above, individual agencies were analyzed only for Black and Latinx groups. Results for these analyses are presented in **Table 6.2.1**.

As shown in **Table 6.2.1**, all agencies have differences in search success rates between white individuals and the two comparison groups. These differences in nearly all cases were relatively small, and in all but one case the differences reported were not statistically significant. The lack of statistical significance could be attributed to the relatively small sample sizes found across agencies (particularly Tier 2 agencies), but it is also important to note that small, statistically insignificant differences in search outcomes 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.

While the vast majority of comparisons present no evidence of disparity in KPT Hit-Rate outcomes and demonstrate only small differences in search outcome

30 All agencies not listed above (Ashland PD, Bend PD, Benton County SO, Canby PD, Central Point PD, Corvallis PD, Deschutes County SO, Douglas County SO, Forest Grove PD, Grants Pass DPS, Hood River County SO, Jackson County SO, Keizer PD, Klamath County SO, Klamath Falls PD, Lake Oswego PD, Lane County SO, Lebanon PD, Lincoln City PD, Lincoln County SO, Linn County SO, McMinnville PD, Medford PD, Milwaukie PD, Newberg-Dundee PD, OHSU PD, Oregon City PD, Port of Portland PD, Redmond PD, Roseburg PD, Tigard PD, Tualatin PD, U of O PD, West Linn PD, Woodburn PD, and Yamhill County SO) did not have enough searches to conduct the hit-rate analysis.

Table 6.2.1
Statewide Hit-Rates and Significance by Agency[30] and Race/Ethnicity

Agency	Race/ Ethnicity	Minority Hit-Rate	White Hit-Rate	p-value (Significance)
Albany PD	Black	0.118	0.364	(no analysis)
	Latinx	0.351	0.364	1.000
Beaverton PD	Black	0.506	0.601	0.129
	Latinx	0.584	0.601	0.752
Clackamas Co SO	Black	0.432	0.481	0.645
	Latinx	0.299	0.481	0.004**
Eugene PD	Black	0.393	0.438	0.592
	Latinx	0.404	0.438	0.741
Gresham PD	Black	0.333	0.518	(no analysis)
	Latinx	0.444	0.518	0.512
Hermiston PD	Black	0.5	0.525	(no analysis)
	Latinx	0.595	0.525	0.676
Hillsboro PD	Black	0.429	0.398	(no analysis)
	Latinx	0.34	0.398	0.595
Marion Co SO	Black	0.182	0.214	(no analysis)
	Latinx	0.143	0.214	0.259
Multnomah Co SO	Black	0.368	0.405	0.802
	Latinx	0.258	0.405	0.167
Oregon State Police	Black	0.77	0.645	0.009†
	Latinx	0.63	0.645	0.618
Polk Co SO	Black	0.714	0.526	(no analysis)
	Latinx	0.394	0.526	0.288
Portland PB	Black	0.353	0.4	0.121
	Latinx	0.38	0.4	0.733
Salem PD	Black	0.333	0.503	(no analysis)
	Latinx	0.398	0.503	0.035
Springfield PD	Black	0.383	0.406	0.879
	Latinx	0.315	0.406	0.245
Washington Co SO	Black	0.689	0.632	0.560
	Latinx	0.703	0.632	0.100

* p<0.05, **p<0.01, *** p<0.001 (Statistical Significance includes a Bonferroni Correction with 2 Comparisons)
†While Oregon State Police has a p-value meeting the threshhold for significance, the hit-rate for white individuals is lower than the hit-rate for Black individuals in this case, and therefore there is no disparity indicated.

percentages, a disparity was found for the Clackamas Co. SO in their Hit-Rate analysis for the Latinx group. More specifically, for the white-Latinx hit-rate comparison, the percentage of successful searches for white individuals was 48.1 percent, while the percentage of successful searches for Latinx individuals was only 29.9 percent. This difference is significant at the 99% confidence level, indicating a disparity.

To aid in contextualizing these comparisons, **Figures 6.2.1** and **6.2.2** present the results reported in **Table 6.2.1** in visual form. In each figure, the white hit-rate occupies the horizontal axis, while the hit-rate for the comparison group is found on the vertical axis. A diagonal line is also included where the hit-rates between the two groups would be exactly equal, and each agency is represented by a numbered dot. The location of the dot represents the relationship between the white and comparison hit-rates. For each agency, it would be expected their dot be close to the diagonal line if a disparity is not present. Alternatively, the likelihood of identifying a disparity increases (dependent upon sample size) as an agency's dot falls further below the diagonal line (the region below the diagonal line is where a comparison group's hit-rate is less than the white hit-rate).

Figure 6.2.1.
Black KPT Hit-Rate Comparison

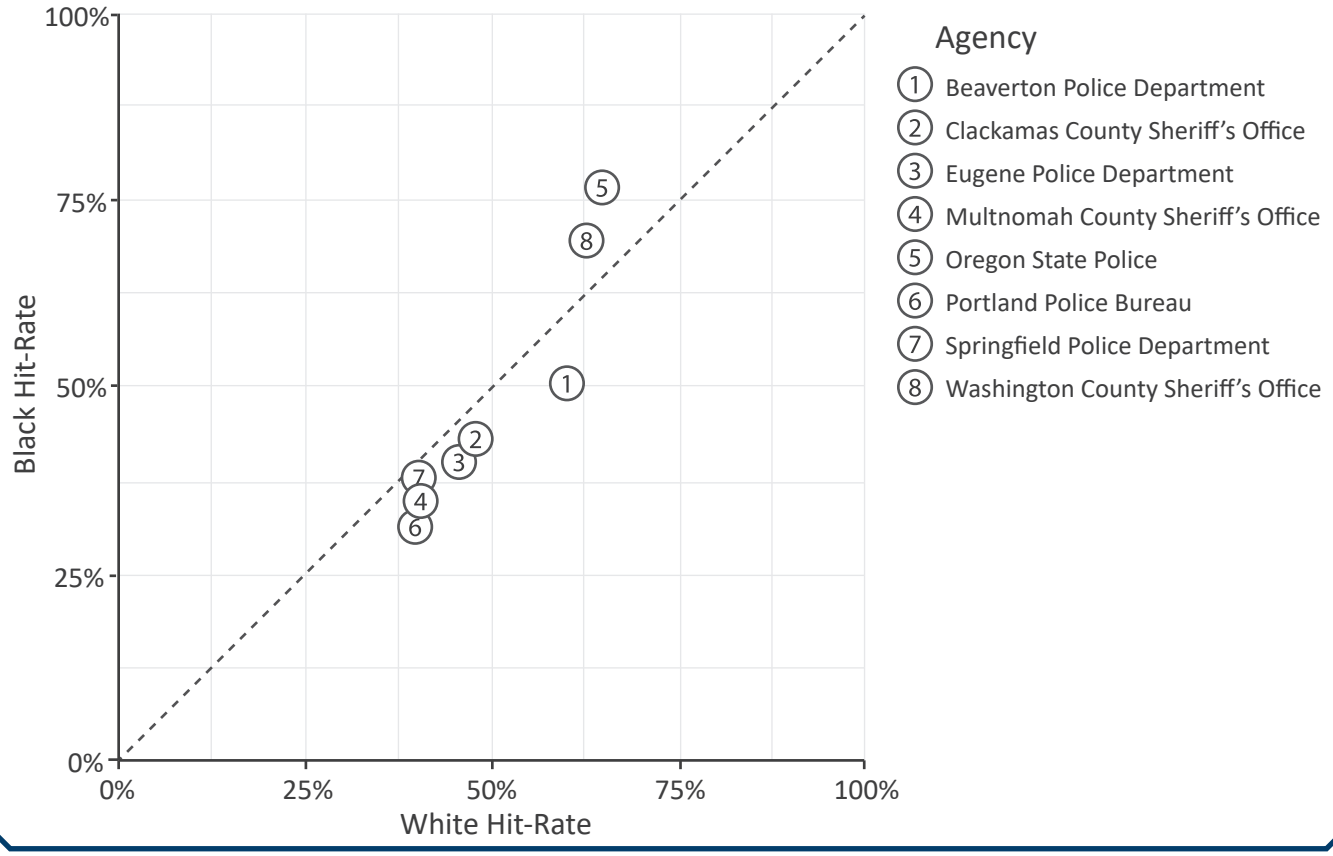
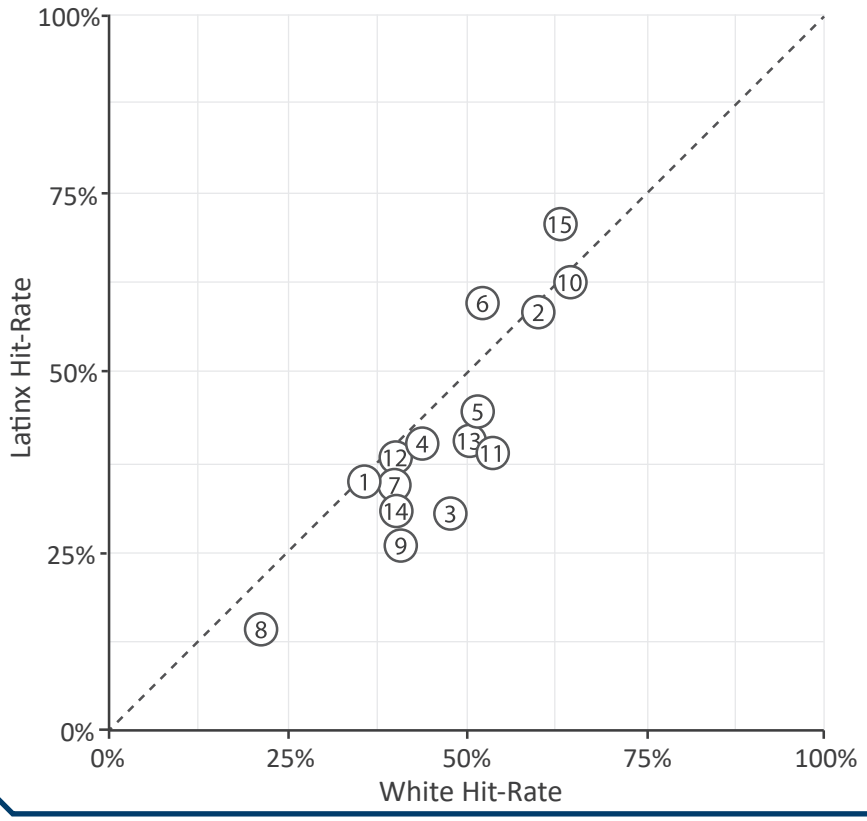


Figure 6.2.2.
Latinx KPT Hit-Rate Comparison



Agency

1. Albany Police Department
2. Beaverton Police Department
3. Clackamas County Sheriff's Office
4. Eugene Police Department
5. Gresham Police Department
6. Hermiston Police Department
7. Hillsboro Police Department
8. Marion County Sheriff's Office
9. Multnomah County Sheriff's Office
10. Oregon State Police
11. Polk County Sheriff's Office
12. Portland Police Bureau
13. Salem Police Department
14. Springfield Police Department
15. Washington County Sheriff's Office

7. Findings from 2019-2020 Tier 1 and Tier 2 Analysis

7.1 Aggregate Findings

Similar to the data reported in the first annual STOP Report, 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 does vary by both agency as well as tier, however, as some law enforcement agencies engage in more pedestrian stops than others and Tier 2 agencies, on average, logged more pedestrian stops than Tier 1 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, in and of itself, 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 non-binary individuals, and this split was greater for pedestrian stops versus traffic stops.

Across both tiers, law enforcement agencies reported that stopped individuals either were subject to no further action or merely given a warning in over 60 percent of stops. Other outcomes, including receiving a citation or being arrested, varied widely across traffic and pedestrian stops, with 30 to 40 percent of traffic stops resulting in a citation versus less than 14 percent of all pedestrian stops. Similarly, while less than 2.5 percent of traffic stops culminated in an arrest in either tier, between 11.6 and 15.3 percent of pedestrian stops resulted in this most serious of outcomes across Tier 2 and Tier 1, respectively.

Finally, with regard to searches, 2.6 to 2.8 percent of all stops resulted in a search of some type. Of those searches, consent was obtained around half of the time, while some other legal basis was reported in the remaining cases. Across Tier 1 agencies, just over half of all searches resulted in the discovery of contraband, with alcohol and drugs being the most commonly reported items. Among Tier 2 agencies, however, the percentage of searches resulting in the seizure of contraband was lower, only around 36 percent. Similar to the data reported in the first annual STOP Report, in all, the STOP data demonstrates that the vast majority of discretionary police-citizen interactions in Oregon are traffic stops.

7.2. Veil of Darkness Findings 2019-2020

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 Veil of Darkness 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 in any minority group at the aggregate Tier 1 or Tier 2 level, or for the Black or Latinx alone groups at the agency level.

Consistent with the first annual STOP Report, the Veil of Darkness Analysis results did not identify any disparities in the decision to stop a driver, either at the state level or for any of the individual Tier 1 or Tier 2 agencies. It is important to note that for many of the Tier 2 agencies, and even a few Tier 1 agencies, STOP Program researchers were unable to estimate the full VOD model. To explore the trends for these agencies, comparisons of the percentages of stops for Black and Latinx individuals during daylight versus darkness were presented. In these cases, no statistically significant differences were reported. Finally, even these pared down analyses could not be conducted on a subset of Tier 2 agencies due to small sample sizes. To combat this issue in the future, the STOP Program will begin to analyze data collected over a two-year period starting with the third annual STOP report when Tier 2 agencies will have submitted their second year of data.

7.3. Predicted Disposition Findings 2019-2020

The Predicted Disposition analysis, which relied 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 disparities in their predicted versus actual dispositional outcomes where those differences were statistically significant at the 95 percent confidence level in any minority group at the aggregate tier level and for the Black group or Latinx group, respectively, at the agency level.

In total, ten Tier 1 agencies were identified as meeting this threshold and sixteen Tier 2 agencies were identified. With regard to citations, Beaverton PD, Clackamas Co. SO, Eugene PD, Hillsboro PD, Marion Co. SO, Multnomah Co. SO, Oregon State Police, and Washington Co. SO were found to have statistically significant differences in the predicted versus actual outcomes for the Tier 1 agencies. Among Tier 2 agencies, Albany PD, Forest Grove PD, Hermiston PD, Keizer PD, Lincoln Co. SO, McMinnville PD, Newberg-Dundee PD, Polk Co. SO, Port of Portland PD, Redmond PD, Roseburg PD, Tigard PD, Tualatin PD, West Linn PD, Woodburn PD, and Yamhill Co. SO were identified for citation disparities.

For searches, Portland PD, Salem PD, and Washington Co. SO among Tier 1 agencies and Woodburn PD were found to have statistically significant disparities. Regarding arrests, Beaverton PD, Marion Co. SO, Portland PB, and Salem PD were identified among Tier 1 agencies, while no Tier 2 agencies were identified. Relatively low occurrences for search and arrest outcomes combined with relatively low occurrences of stops of Black individuals and Latinx individuals are significant barriers to identifying disparities.

Notably, many analyses for several agencies could not be estimated due to low sample sizes, especially for minority groups, search and arrest outcomes, and for smaller Tier 2 agencies. In these situations we cannot detect the presence of a disparity with current data limitations and we must wait until we have at least 2 years of STOP data on which to conduct the analysis. At least one analysis was possible for every agency except Lebanon PD, Corvallis PD, and University of Oregon PD.

7.3.1. Citations

The most common finding of a disparity in this report was for citations, as 24 agencies were reported to have disparities in citation rates for either Latinx and/or Black individuals. Among the agencies where disparities were identified, the average disparity was a difference in approximately 7.7 percentage points between the predicted and actual rates, with a range of 2.5 to 16.7 percentage points.

These findings with regard to citations are likely influenced, at least in part, by departmental policies regarding citations. While the exact extent of this influence is not yet known and it is unlikely that policies of this kind would fully explain away the existence of disparities in all cases, these policies and controlling for them in future analyses represent an important next step in the analysis of stop data in Oregon as well as a potential area for discussion between law enforcement and the communities these agencies serve.[31] Future iterations of the STOP analyses will include additional information regarding citations.

7.3.2. Searches

With regard to searches, four agencies were identified as having statistically significant disparities for at least one comparison groups: Portland PB, Salem PD, Washington Co. SO, and Woodburn PD. Portland PB had a

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

search disparity for Latinx individuals of similar magnitude to Beaverton PD. The analysis for Portland PB identified a disparity with Black individuals, whereas Salem PD, Washington Co. SO, and Woodburn PD had disparities with Latinx individuals. Notably, the expected search rate of the Woodburn PD for Latinx individuals was 1.1 percent, whereas the actual search rate for Latinx individuals is 4.5 percent.

7.3.3. Arrests

With regard to arrests, three agencies were identified as having statistically significant disparities for Black individuals or Latinx individuals: Beaverton PD, Marion Co. SO, and Portland PB. Portland PB was identified for arrests of Black individuals, whereas the other two agencies were identified for arrests of Latinx individuals.

7.4. KPT Hit-Rate Findings 2019-2020

The second of two analyses examining post stop outcomes was the KPT Hit-Rate 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 aggregate Tier 1 or Tier 2 level, or for the Black or Latinx alone groups at the agency level.

In this analysis, only Clackamas Co. SO was found to have a disparity meeting the above criteria. Specifically, Clackamas Co. SO reported successful searches in 48.1 percent of searches involving white individuals but only reported successful searches in 29.9 percent of searches of Latinx individuals.

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 stakeholders to focus training and technical assistance on those 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 allow 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 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.

Based on the above described criteria, it is recommended that the Clackamas Co. SO be examined in greater detail by STOP Program researchers and receive technical assistance from DPSST. Clackamas Co. SO was indicated as having a disparity in the Predicted Disposition analysis with regard to its citations of Black and Latinx individuals, as well as the KPT Hit-Rate analysis with regard to searches of Latinx individuals.

7. Findings from 2019-2020 Tier 1 and Tier 2 Analysis

Regardless of whether an agency is officially referred to DPSST by this report or not, the CJC urges each agency to scrutinize the full set of results for their agency, found in the Appendices. While most agencies are not referred to DPSST in this analysis, that does not necessarily mean that the results for all those agencies should be ignored or are not close to the threshold of identification. All agencies and/or interested stakeholders should contact the CJC should they require technical assistance in interpreting specific statistical results.

7.6. Next Steps and Future Work

The second annual STOP Program report includes data from fifty-one Oregon law enforcement agencies, representing all agencies with at least 25 officers with stop authority. Next year, all Oregon law enforcement agencies, including over one hundred additional Tier 3 agencies, will be examined by the STOP Program. While this will be a significant event, the STOP Program will encounter challenges in analyzing this statewide data, as many Oregon law enforcement agencies do not make enough stops to conduct the necessary analyses. Indeed, as described in detail in this report, even among Tier 1 and Tier 2 agencies the STOP Program often does not have sufficient data points to conduct the rigorous statistical analyses necessary to identify disparities. During the current reporting year, this issue was significantly impacted by the COVID-19 pandemic, but even absent the effects of this unexpected event, it is likely that a number of Oregon's Tier 2 agencies, and possibly all of the Tier 3 agencies, will continue to fail to meet the sample size thresholds necessary for full analysis on an annual basis.

Due to these concerns, in the 2020-2021 report that will be released in December 2021, the STOP Program will begin reporting results using both (i) the one-year annual data reported by agencies, as well as (ii) two-year data covering 2019-2021 for Tier 1 and Tier 2 agencies. In future reports, as the stop rates for Tier 3 agencies become clear, the STOP Program will aggregate stop data as necessary to run its analyses, although it may take several years to build up a sufficient number of data points for these smaller agencies.

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.

Works Cited

"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

Tier 1

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

Tier 2

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

Appendix A – List of Law Enforcement Agencies by Tier

Tier 3

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

Appendix B - Data Audit

Missing data in the STOP Program database can be traced to two sources. First, a value for an individual data point could be truly missing, which means that no information for that particular variable was entered into the STOP database. Second, a data point could be invalid, which means that a value of some type was entered into the STOP database, but it did not conform to the standards of the STOP Program. **Table B1** presents a breakdown of missing data in these two forms for the STOP Program variables used in the analyses contained in this report.

Table B.1
Missing Data for STOP Program Variables used in Year 2 Report Analyses

Variable	Description	Analyses Affected	% Missing
age	Age perceived by officer	Veil of Darkness, Predicted Disposition	3.18%
agency	Stopping agency	Veil of Darkness, Predicted Disposition, Hit-Rate	0.00%
arrest	Physical custody arrest (yes/no)	Predicted Disposition	1.69%
CiteCat*	Category of citation (Move/Spd, Ser Move/ Spd, Very Ser Move/Sp, Equip Vio/Cell/ Seatbelt, Reg/License, Other)	Veil of Darkness	0.00%
cite_type	Citation basis for traffic stop (ORS, Municipal Traffic, Municipal Criminal, County Ordinance)	Veil of Darkness, Predicted Disposition	9.70%
county	County in which stop occurred	Veil of Darkness, Predicted Disposition	0.00%
disposition	Most severe disposition of stop (none, warning, citation, search, arrest)	Predicted Disposition	0.66%
gender	Gender perceived by officer (male, female, non-binary)	Veil of Darkness, Predicted Disposition	0.17%
race	Race/ethnicity perceived by officer (Asian/ PI, Black, Hispanic, Middle Eastern, Native American, white)	Veil of Darkness, Predicted Disposition, Hit-Rate	0.82%
sdate	Date of stop	Veil of Darkness, Predicted Disposition	0.00%
search	Whether a discretionary stop occurred (yes/no)	Predicted Disposition, Hit-Rate	0.00%
search_ f1**	What was found if a search occurred (Nothing, Alcohol, Drugs, Stolen Property, Weapons, Other Evidence, Other non-Evidence)	Hit-Rate	2.25%
stime	Time of stop	Veil of Darkness, Predicted Disposition	0.00%
stop_type	Type of stop (traffic, pedestrian)	Veil of Darkness, Predicted Disposition	0.00%

*cite_cat is a condensed variable created from the original variable cite_code, which denotes the ORS code for the citation. As not every stop has an ORS code basis, some stops are missing cite_code, but are not included in these missing/invalid counts.
**This missing percentage reflects the percent of search_f1 missing when an entry is expected. In the case that Search= “no”, there is not an entry expected for search_f1, so these are not included in the missing percentage in this table.

Appendix B - Data Audit

As shown in **Table B1**, missing data were reported for several, but not all variables used in the STOP Program analyses. On the whole, the data submitted by the Tier 1 and Tier 2 agencies have met the standards of the STOP Program and do not raise significant concerns with the research staff. While it is important to strive for 100 percent reporting in all cases, rarely does reporting meet this goal in the early years of data collection.

Appendix C – Veil of Darkness Technical Appendix and Detailed Results

The Veil of Darkness (VOD) analysis was first developed by Grogger and Ridgeway (2006) for analyzing 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 VOD 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 Veil of Darkness 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:

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

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, and agency 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.[32] Importantly, the inclusion of controls for time of day, day or 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 VOD 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, 2019 to June 30, 2020. Every traffic stop is defined to have occurred in daylight or darkness based on the date, time, and location of the stop. Astronomical data from the United States Naval Observatory (USNO) is used to determine the sunrise, sunset, and start and end of civil twilight. If the location of the stop has been geo-coded, then those coordinates are used to determine the sunrise, sunset, and civil twilight window for that exact location. If the stop has not been geo-coded due to limitations with location data, the centroid of the city is used. If the city information is unavailable, then the centroid of the county is used.

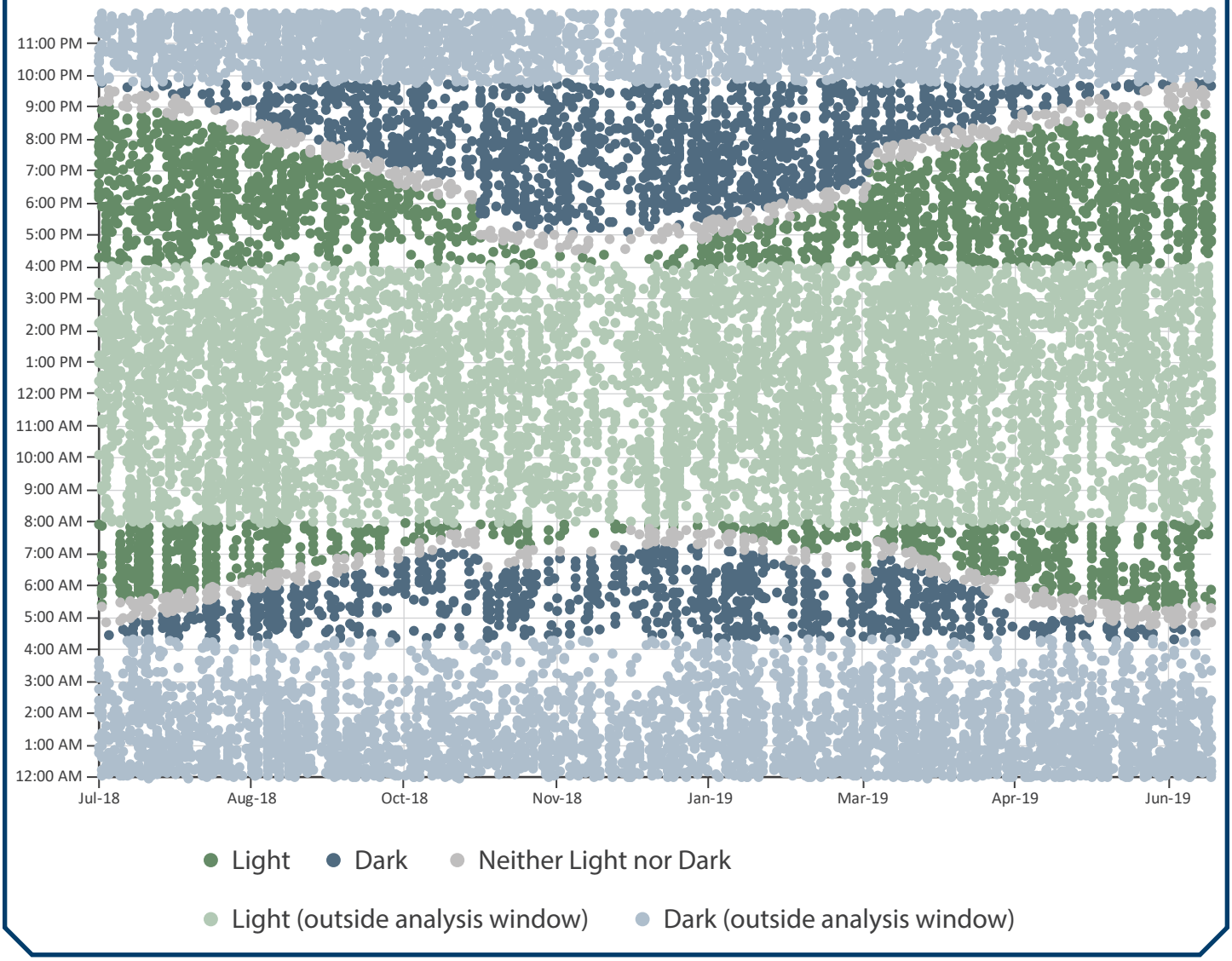
The dawn inter-twilight period is defined as the earliest start of civil twilight to the latest sunrise. The earliest start of civil twilight is 4:21am in Wallowa County, and the latest sunrise is 7:59am in Clatsop County. Stops that occur in the daily morning twilight window (approximately 30 minutes between the start of civil twilight and the sunrise) are removed since it is neither light nor dark during this time period. Conversely, the dusk twilight window is defined as the earliest sunset to the latest end of civil twilight. The earliest sunset is 4:05pm in Wallowa County, and the latest end of civil twilight is 9:48pm in Clatsop County. Stops that occur in the daily evening twilight window (approximately 30 minutes between sunset and the end of civil twilight) are similarly removed since it is neither light nor dark during this time period. Adjustments have been made to account for daylight savings time

32 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, 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. In models examining aggregate Tier 1 outcomes, agency fixed effects were included as well.

(DST) in November 2019 and March 2020. In addition, most of Malheur County is on Mountain Standard Time (MST) and the stops in Malheur County have been adjusted to account for this time zone.

Figure C1 displays the traffic stops for one Tier 1 agency by date and time of day. The gradual variation in the sunrise and sunset times shows the shift in daylight and darkness for traffic stops. The morning and evening inter-twilight windows are displayed, and stops during these times are included in the VOD analysis.

Figure C1.
Sample Statewide Distribution of Traffic Stops by Time and Day



The log odds that result from the Veil of Darkness 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 VOD 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[33].

Detailed Veil of Darkness Results

The following tables display the VOD Results for the statewide and agency specific models. For each model the sample size, coefficient, standard error, p-value, odds ratio, and 95% confidence interval of the odds ratio are displayed. The combined Tier 1 and Tier 2 results in the main report include a Bonferroni correction for 5 comparisons. The agency specific results include a Bonferroni correction for 2 comparisons.

Table C.1
All Tier 1 Agencies VOD Results

Race/Ethnicity	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Asian/PI	81,184	-0.048	0.031	0.122	0.909	0.776	1.065
Black	83,133	0.017	0.026	0.498	1.035	0.907	1.182
Latinx	91,470	-0.013	0.017	0.450	0.974	0.892	1.064
Middle Eastern	79,039	-0.054	0.048	0.259	0.897	0.700	1.149
Nat. American	78,178	-0.172	0.077	0.026	0.709	0.476	1.056
Asian/PI (no OSP)	37,176	-0.067	0.038	0.074	0.875	0.721	1.061
Black (no OSP)	38,837	0.005	0.030	0.868	1.010	0.865	1.180
Latinx (no OSP)	42,472	0.009	0.023	0.702	1.018	0.902	1.149
Middle Eastern (no OSP)	35,509	-0.078	0.065	0.230	0.856	0.612	1.196
Nat. American (no OSP)	35,005	-0.156	0.111	0.160	0.732	0.413	1.297

Table C.2
All Tier 2 Agencies VOD Results

Race/Ethnicity	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Asian or PI	39,240	-0.004	0.056	0.946	0.992	0.745	1.323
Black	39,323	0.032	0.054	0.562	1.065	0.805	1.409
Latinx	43,772	0.006	0.028	0.836	1.011	0.877	1.166
Middle Eastern	38,493	-0.027	0.093	0.775	0.948	0.588	1.530
Native American	38,244	0.278	0.152	0.066	1.745	0.799	3.810

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

Table C.3
All Tier 1 Agencies Restricted to DST Window VOD Results

Race/Ethnicity	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Asian/PI	26,550	-0.084	0.059	0.155	0.845	0.623	1.147
Black	27,160	0.014	0.051	0.786	1.028	0.792	1.334
Latinx	29,770	-0.028	0.034	0.415	0.946	0.795	1.127
Middle Eastern	25,825	-0.011	0.094	0.911	0.979	0.603	1.590
Nat. American	25,498	-0.243	0.167	0.145	0.615	0.260	1.453
Asian/PI (no OSP)	12,203	-0.139	0.075	0.063	0.758	0.516	1.112
Black (no OSP)	12,739	-0.023	0.061	0.710	0.956	0.700	1.306
Latinx (no OSP)	13,858	-0.074	0.050	0.139	0.863	0.669	1.115
Middle Eastern (no OSP)	11,608	-0.062	0.133	0.638	0.883	0.445	1.749
Nat. American (no OSP)	Insufficient sample size						

Table C.4
All Tier 2 Agencies Restricted to DST Window VOD Results

Race/Ethnicity	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Asian or PI	13,101	-0.081	0.128	0.529	0.851	0.439	1.649
Black	13,150	0.128	0.121	0.292	1.291	0.691	2.410
Latinx	14,654	-0.011	0.060	0.852	0.978	0.717	1.334
Middle Eastern	12,860	-0.106	0.216	0.622	0.808	0.266	2.456
Native American	Insufficient sample size						

Table C.5
Black vs. White Traffic Stops, Agency VOD Results

Agency	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Beaverton PD	3,311	0.123	0.102	0.228	1.280	0.809	2.024
Clackamas Co SO	5,792	0.094	0.107	0.381	1.206	0.747	1.945
Eugene PD	3,630	-0.020	0.140	0.885	0.960	0.512	1.801
Gresham PD	1,399	0.102	0.138	0.462	1.226	0.660	2.277
Hillsboro PD	2,221	-0.040	0.151	0.791	0.923	0.470	1.814
Marion Co SO			Insufficient sample size				
Medford PD			Insufficient sample size				
Multnomah Co SO	2,321	0.026	0.131	0.840	1.054	0.587	1.893
Oregon State Police	44,296	0.039	0.051	0.443	1.081	0.862	1.355
Portland PB	10,638	-0.007	0.044	0.870	0.986	0.811	1.198
Salem PD			Insufficient sample size				
Washington Co SO	5,236	-0.218	0.103	0.035	0.647	0.408	1.027
Springfield PD	2,945	-0.034	0.176	0.848	0.935	0.425	2.055

Table C.6
Latinx vs. White Traffic Stops, Tier 1 Agency VOD Results

Agency	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Beaverton PD	3,777	-0.045	0.071	0.527	0.914	0.666	1.256
Clackamas Co SO	6,369	0.022	0.067	0.740	1.045	0.775	1.409
Eugene PD	3,691	0.040	0.124	0.747	1.083	0.622	1.886
Gresham PD	1,468	0.008	0.128	0.951	1.016	0.571	1.807
Hillsboro PD	2,886	0.049	0.079	0.532	1.104	0.775	1.571
Marion Co SO	2,812	-0.101	0.088	0.251	0.817	0.552	1.211
Medford PD	658	0.006	0.215	0.977	1.013	0.387	2.649
Multnomah Co SO	2,420	0.177	0.115	0.123	1.426	0.851	2.387
Oregon State Police	48,998	-0.038	0.026	0.140	0.927	0.827	1.040
Portland PB	9,872	0.045	0.053	0.394	1.094	0.864	1.387
Salem PD	2,016	-0.066	0.111	0.553	0.877	0.533	1.441
Washington Co SO	6,503	0.001	0.053	0.988	1.002	0.789	1.272

Table C.7
Latinx vs. White Traffic Stops, Tier 2 Agency VOD Results

Agency	Sample Size	Coefficient	Standard Error	P-Value	Odds Ratio	95% CI	
Albany PD	1,935	0.152	0.134	0.256	1.356	0.743	2.475
Bend PD	2,249	0.024	0.143	0.869	1.048	0.552	1.992
Benton Co SO	1,177	0.067	0.207	0.747	1.143	0.452	2.889
Canby PD	684	-0.208	0.183	0.256	0.660	0.291	1.498
Deschutes Co SO	1,732	0.138	0.181	0.444	1.318	0.587	2.960
Forest Grove PD	1,470	-0.061	0.115	0.594	0.885	0.529	1.481
Hermiston PD	1,708	-0.111	0.100	0.267	0.801	0.512	1.254
Hood River Co SO	441	0.128	0.212	0.546	1.292	0.499	3.346
Jackson Co SO	1,582	-0.043	0.142	0.765	0.918	0.486	1.736
Keizer PD	478	0.310	0.285	0.277	1.859	0.518	6.673
Klamath Falls PD	1,171	0.022	0.159	0.890	1.045	0.513	2.130
Lincoln City PD	1,010	-0.047	0.190	0.803	0.910	0.389	2.129
Linn Co SO	1,997	-0.101	0.163	0.536	0.817	0.394	1.696
McMinnville PD	1,408	-0.038	0.130	0.768	0.926	0.517	1.660
Milwaukie PD	1,162	0.252	0.195	0.194	1.657	0.693	3.962
Newberg-Dundee PD	1,611	-0.105	0.125	0.400	0.811	0.464	1.417
Polk Co SO	1,178	-0.016	0.142	0.908	0.968	0.511	1.832
Springfield PD	2,981	-0.283	0.163	0.082	0.568	0.274	1.177
Tigard PD	1,291	0.065	0.157	0.676	1.139	0.567	2.289
Tualatin PD	1,114	-0.208	0.176	0.237	0.660	0.300	1.450
West Linn PD	1,580	0.154	0.188	0.412	1.361	0.587	3.155
Woodburn PD	312	-0.132	0.266	0.621	0.769	0.233	2.534
Yamhill Co SO	1,128	0.102	0.148	0.488	1.227	0.634	2.377

Table C.8
Results for Stops of Black Individuals, Agency Results

Agency	Black		White		Percentage Black		P-Value
	Dark	Light	Dark	Light	Dark	Light	
Marion Co SO	20	49	555	1,578	3.5%	3.0%	0.581
Medford PD	14	16	263	276	5.1%	5.5%	0.821
Salem PD	33	29	591	888	5.3%	3.2%	0.037
Albany PD	19	16	1,004	715	1.9%	2.2%	0.625
Ashland PD	16	12	468	305	--	--	--
Bend PD	21	8	870	1,181	--	--	--
Benton Co SO	17	18	583	488	2.8%	3.6%	0.493
Canby PD	8	7	256	277	--	--	--
Central Point PD	10	2	299	166	--	--	--
Corvallis PD	22	6	628	279	--	--	--
Deschutes Co SO	13	7	972	636	--	--	--
Douglas Co SO	12	8	589	570	--	--	--
Forest Grove PD	13	17	507	582	2.5%	2.8%	0.727
Grants Pass DPS	9	4	303	473	--	--	--
Hermiston PD	9	15	570	527	--	--	--
Hood River Co SO	2	3	120	220	--	--	--
Jackson Co SO	16	12	668	673	--	--	--
Keizer PD	13	5	183	180	--	--	--
Klamath Co SO	0	0	47	41	--	--	--
Klamath Falls PD	27	22	535	472	4.8%	4.5%	0.787
Lake Oswego PD	26	30	491	698	5.0%	4.1%	0.446
Lane Co SO	16	11	359	471	--	--	--
Lebanon PD	2	2	106	117	--	--	--
Lincoln City PD	16	11	587	305	--	--	--
Lincoln Co SO	17	4	787	298	--	--	--
Linn Co SO	10	17	879	965	--	--	--
McMinnville PD	18	9	659	459	--	--	--
Milwaukie PD	24	41	463	581	4.9%	6.6%	0.242
Newberg-Dundee PD	18	22	647	696	2.7%	3.1%	0.692
OHSU PD	17	0	156	39	--	--	--
Oregon City PD	18	23	477	670	3.6%	3.3%	0.768
Polk Co SO	18	17	498	432	3.5%	3.8%	0.805
Port Of Portland PD	41	23	249	162	14.1%	12.4%	0.596

Agency	Black		White		Percentage Black		
	Dark	Light	Dark	Light	Dark	Light	P-Value
Redmond PD	5	2	407	328	--	--	--
Roseburg PD	26	23	556	840	4.5%	2.7%	0.063
Tigard PD	39	34	484	626	7.5%	5.2%	0.102
Tualatin PD	21	19	380	556	5.2%	3.3%	0.134
U Of O PD	3	3	47	27	--	--	--
West Linn PD	27	31	842	620	3.1%	4.8%	0.096
Woodburn PD	2	4	60	48	--	--	--
Yamhill Co SO	11	10	465	443	--	--	--

Table C.9
Results for Stops of Latinx Individuals, Agency Results

Agency	Latinx		White		Percentage Latinx		P-Value
	Dark	Light	Dark	Light	Dark	Light	
Ashland PD	30	23	468	305	6.0%	7.0%	0.571
Central Point PD	42	21	299	166	12.3%	11.2%	0.713
Corvallis PD	67	25	628	279	9.6%	8.2%	0.476
Douglas Co SO	26	29	589	570	4.2%	4.8%	0.607
Grants Pass DPS	20	27	303	473	6.2%	5.4%	0.633
Klamath Co SO	6	4	47	41	--	--	--
Lake Oswego PD	33	48	491	698	6.3%	6.4%	0.922
Lane Co SO	35	28	359	471	8.9%	5.6%	0.058
Lebanon PD	7	4	106	117	--	--	--
Lincoln Co SO	62	24	787	298	7.3%	7.5%	0.930
OHSU PD	12	8	156	39	--	--	--
Oregon City PD	46	45	477	670	8.8%	6.3%	0.096
Port Of Portland PD	48	26	249	162	16.2%	13.8%	0.487
Redmond PD	40	37	407	328	8.9%	10.1%	0.565
Roseburg PD	32	51	556	840	5.4%	5.7%	0.818
U Of O PD	9	5	47	27	--	--	--

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[34]. 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.[35]
    a. For treated observations, IPW =1
    b. For control observations, IPW = $\frac{(propensity\ score)}{1-(propensity\ score)}$

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.[36]

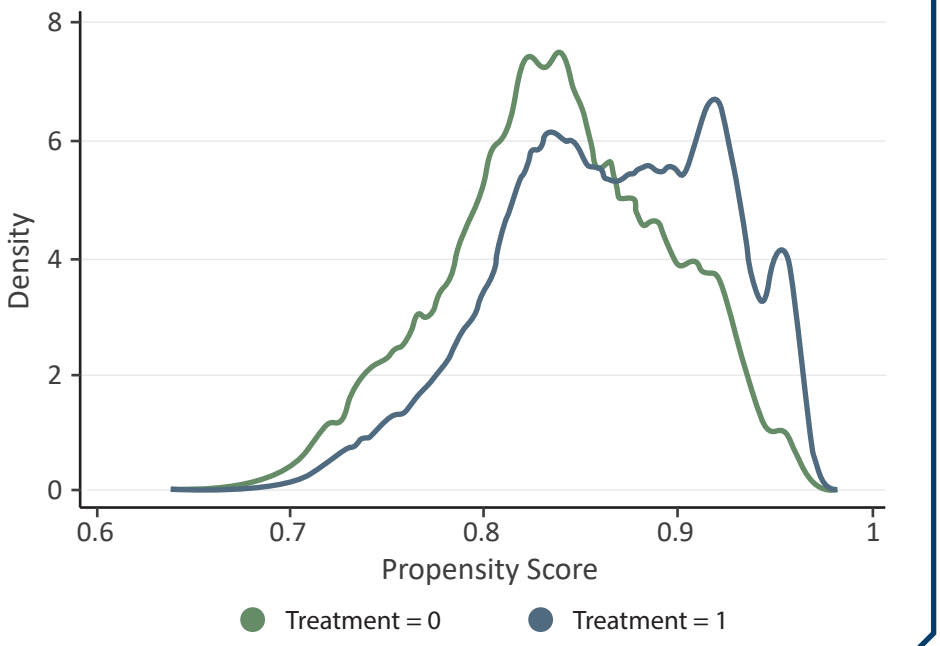
34 StataCorp. 2013. Stata: Release 13. Statistical Software. College Station, TX: StataCorp LP.
35 These differ whether the estimand is the Average Treatment Effect (ATE) or the Average Treatment Effect on the Treated (ATET). Here we are estimating the ATET. (Austin and Stuart 2015)
36 For a thorough discussion of IPWRA methods see (Wooldridge 2010, Chapter 21.3.4)

Assumptions

There are a few assumptions that must hold in order for propensity score estimators to be unbiased. The first is the conditional independence assumption[37], which states that the outcome variable is conditionally independent of the treatment. This means that if researchers include all relevant confounding variables in estimating the treatment equation, i.e., the treatment equation is properly specified, and these variables are balanced across the two groups following match selection, then the outcomes are conditionally independent of the treatment. In order for this assumption to hold, changes in any unobserved variables that have an effect on the outcome variable must not also have an effect on the treatment variable. This assumption is a theoretical consideration that is not possible to directly test, as a variable may be correlated with both treatment and outcome but may be a spurious correlation. The analyst may, however, ensure that all the measured confounding variables are equally represented in both the treatment and control groups and thus that the confounding variables are not the drivers of remaining variance in treatments and outcomes.

The second main assumption is the overlap assumption, whereby the range of estimated propensity scores for the treated group must overlap with those of control group observations. If an observation is not within this range then it is omitted from the sample as it is impossible to form a valid match from the comparison group. This idea is best represented with a pre-balance propensity score distribution graph, as seen in the examples below. **Figure D1** 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 D1.
Overlap Example



With a limited range of covariates, including mostly categorical variables, and the large sample sizes with this set of Tier 1 agencies, each analysis completed here had no omitted observations because of a violation of the overlap assumption.[38]

Finally is 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

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

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.[39]

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 Black, = 0 if white
2. Officer perception of race/ethnicity: = 1 if Latinx, = 0 if white

For the statewide models, a broader set of treatment variables is available because of the higher sample size:

1. Officer perception of race/ethnicity: = 1 if Black, = 0 if white
2. Officer perception of race/ethnicity: = 1 if Latinx, = 0 if white
3. Officer perception of race/ethnicity: = 1 if Asian, = 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 are 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:

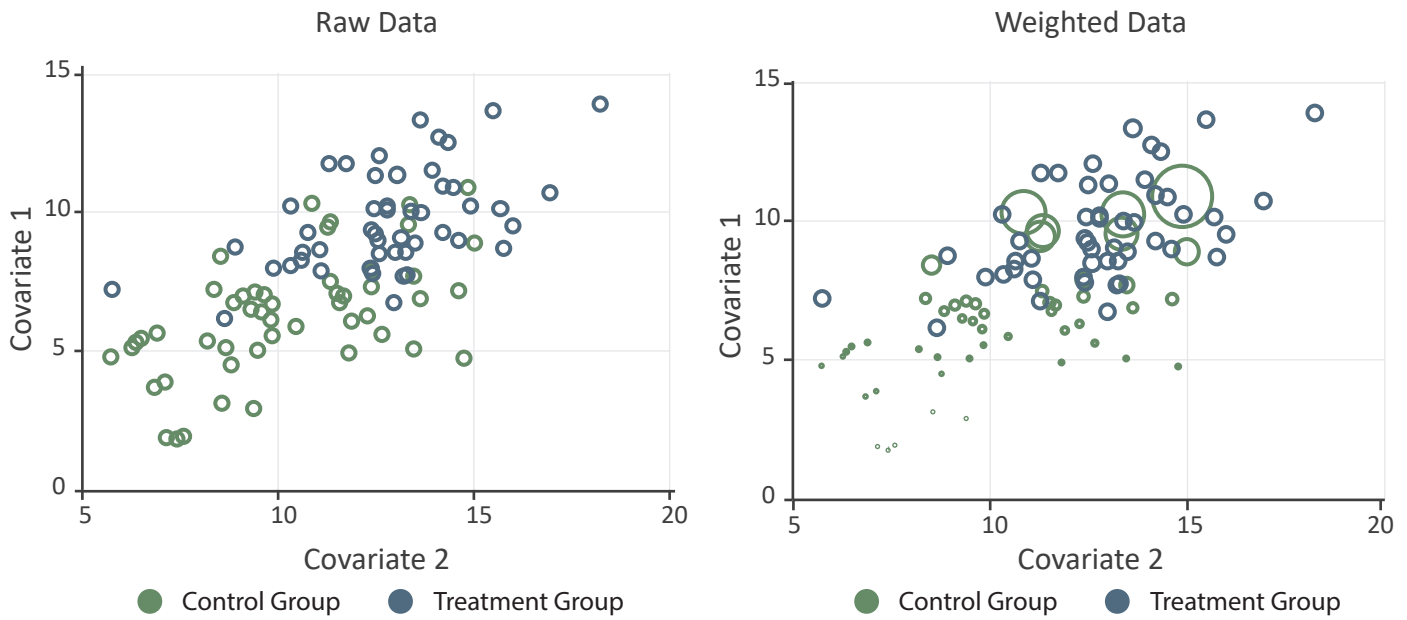
1. Female indicator, 1 = if female, 0 = if any other
2. Age category indicators for each of <21, 21-24, 25-29, 30-39, 40-49, 50+
3. Season indicators for each of Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec
4. Daylight indicator = 1 if stop happened after sunrise and before sunset, = 0 otherwise
5. Time of stop 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 = $\frac{\textit{Total \# of stops by agency on day of stop}}{\textit{Maximum \# of daily stops by } \textbf{\textit{agency}} \textit{ over year of analysis}}$
9. County stop volume = $\frac{\textit{Total \# of stops by agency on day of stop}}{\textit{Maximum \# of daily stops by } \textbf{\textit{county}} \textit{ over year of analysis}}$

39 The Stata handbook provides a good description of these assumptions, and the counterfactual model that underlies all matching methods. ("Stata Treatment-Effects Reference Manual: Potential Outcomes/Counterfactual Outcomes" 2013).

The first step of the analysis uses a probit model to estimate the propensity of being in the treatment group based on the covariates listed above. Overlap of propensity scores is evaluated and any non-overlapping observations are removed from the sample. Inverse Probability Weights (IPWs) are estimated for each observation based on the propensity scores. For the treatment group in an ATET framework these weights are equal to **1**. For the control group the weight is equal to $p/(1-p)$, where p is the propensity score (see **footnote 31**). In effect, this process gives more weight to control observations that have a higher propensity score (i.e., are more similar to treated observations) 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 D2** 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 D2.
Weighting Example



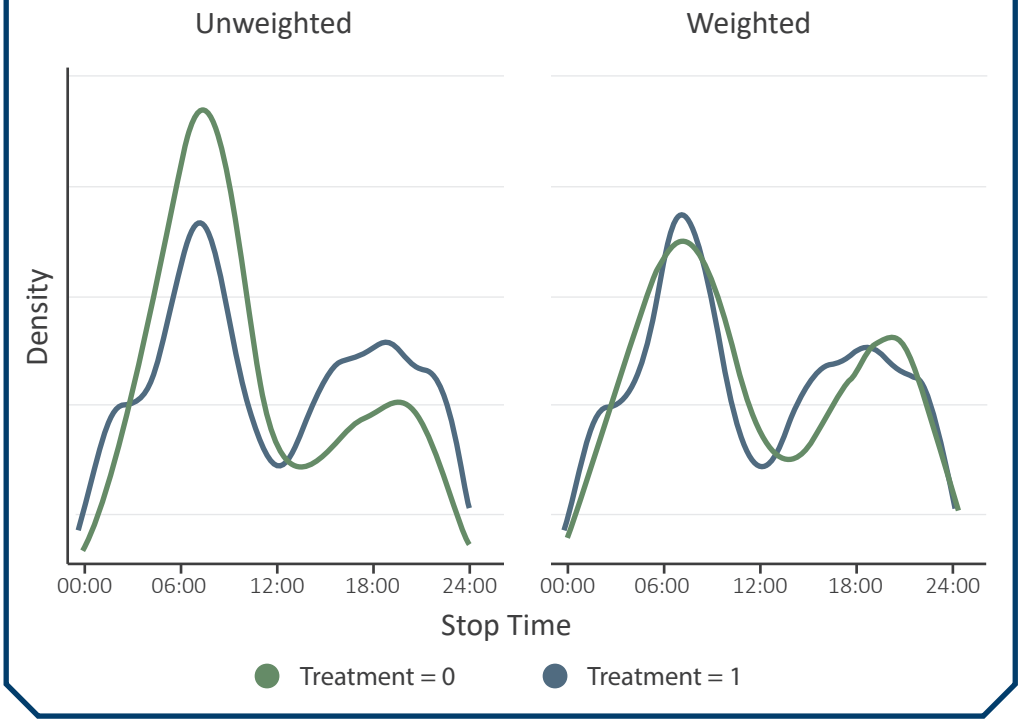
Balance is then measured based on the standardized difference[40] in means and the variance ratio[41] between the treatment and control groups for each of the raw data set and the inverse probability weighted data set. If the resulting standardized difference in the weighted data set is close to zero and the variance ratio is close to 1 for

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

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

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 D3** is an example of one of these variables for one agency. The Unweighted chart displays the distribution of stop time for each of the treated group and the untreated group. The Weighted chart displays these same distributions with the IPWs applied. The distributions of the two groups more closely resemble each other in the weighted graph than in the unweighted graph, so we can say that these groups are more balanced when incorporating the IPWs.

Figure D3.
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 group (Black/Latinx) had identical characteristics to the control group, except had a race/ethnicity = white.[42]

Limitations

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

Results

The threshold for identifying an effect as significant is 95% or a p-value of 0.05 or less. For each agency, however, we are conducting 8 tests (20 for statewide). A Bonferroni adjustment[43] is warranted in this situation to adjust for the likelihood of a given test resulting in a false positive. The appropriate threshold for each agency is thus, p-value $\leq 0.05/8=0.00625$ (statewide p-value $\leq 0.05/20=0.0025$). Statistically significant disparities are indicated by a * next to the average treatment effect on the treated (ATET) in the tables below.

42 Conversely, the ATE is 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 Black/Latinx/Black-Latinx groups and averages across all observations.

43 Weisstein, Eric W. "Bonferroni Correction." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/BonferroniCorrection.html

Table D.1.1
Tier 1 - All

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Asian	Citation	0.3963	0.0184	0.0045	4.10	0.0000	0.3643	0.3915
	Search	0.0103	-0.0052	0.0010	-5.19	0.0000	0.0124	0.0185
	Arrest	0.0166	-0.0103	0.0013	-8.14	0.0000	0.0231	0.0307
	Citation, Search, or Arrest	0.4083	0.0087	0.0045	1.92	0.0542	0.3860	0.4133
Black	Citation	0.3764	-0.0070	0.0038	-1.87	0.0616	0.3721	0.3948
	Search	0.0257	0.0044	0.0012	3.51	0.0005	0.0176	0.0251
	Arrest	0.0459	0.0080	0.0016	4.92	0.0000	0.0033	0.0428
	Citation, Search, or Arrest	0.4203	0.0064	0.0038	1.68	0.0922	0.4025	0.4254
Latinx	Citation	0.4497	0.0619	0.0026	24.06	0.0000	0.3801	0.3956
	Search	0.0209	0.0036	0.0008	4.79	0.0000	0.0149	0.0195
	Arrest	0.0337	0.0047	0.0010	4.98	0.0000	0.0261	0.0318
	Citation, Search, or Arrest	0.4733	0.0638	0.0026	24.87	0.0000	0.4017	0.4172
Middle Eastern*	Citation	0.3471	-0.0046	0.0067	-0.69	0.4920	0.3313	0.3721
	Search	0.0062	-0.0090	0.0012	-7.30	0.0000	0.0115	0.0190
	Arrest	0.0111	-0.0160	0.0016	-9.90	0.0000	0.0222	0.0320
	Citation, Search, or Arrest	0.3565	-0.0182	0.0068	-2.67	0.0076	0.3540	0.3952
Native	Citation	0.3759	0.0364	0.0116	3.13	0.0017	0.3044	0.3746
	Search	0.0294	0.0116	0.0039	2.95	0.0032	0.0059	0.0297
	Arrest	0.0452	0.0149	0.0048	3.13	0.0017	0.0159	0.0447
	Citation, Search, or Arrest	0.4101	0.0473	0.0115	4.11	0.0000	0.3280	0.3976

* Eugene PD is omitted from the Middle Eastern analysis since this agency does not log a Middle Eastern race/ethnicity

Table D.1.2
Tier 1 - No OSP

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Asian	Citation	0.3914	0.0015	0.0052	0.29	0.7702	0.3742	0.4056
	Search	0.0132	-0.0047	0.0013	-3.57	0.0004	0.0139	0.0219
	Arrest	0.0208	-0.0104	0.0017	-6.23	0.0000	0.0262	0.0363
	Citation, Search, or Arrest	0.4063	-0.0086	0.0052	-1.64	0.1016	0.3990	0.4307
Black	Citation	0.3505	-0.0352	0.0042	-8.36	0.0000	0.3730	0.3984
	Search	0.0306	0.0069	0.0015	4.43	0.0000	0.0191	0.0285
	Arrest	0.0539	0.0115	0.0020	5.70	0.0000	0.0363	0.0485
	Citation, Search, or Arrest	0.4047	-0.0158	0.0043	-3.68	0.0002	0.4075	0.4335
Latinx	Citation	0.4402	0.0373	0.0032	11.52	0.0000	0.3931	0.4127
	Search	0.0286	0.0071	0.0011	6.22	0.0000	0.0181	0.0250
	Arrest	0.0435	0.0080	0.0014	5.75	0.0000	0.0313	0.0397
	Citation, Search, or Arrest	0.4700	0.0419	0.0033	12.86	0.0000	0.4182	0.4379
Middle Eastern	Citation	0.3172	-0.0312	0.0082	-3.80	0.0001	0.3236	0.3733
	Search	0.0083	-0.0107	0.0018	-5.82	0.0000	0.0134	0.0246
	Arrest	0.0131	-0.0205	0.0023	-8.82	0.0000	0.0266	0.0406
	Citation, Search, or Arrest	0.3285	-0.0489	0.0083	-5.87	0.0000	0.3522	0.4026
Native	Citation	0.3500	0.0210	0.0156	1.34	0.1794	0.2817	0.3763
	Search	0.0260	0.0051	0.0052	0.99	0.3244	0.0052	0.0366
	Arrest	0.0431	0.0072	0.0065	1.10	0.2729	0.0163	0.0557
	Citation, Search, or Arrest	0.3872	0.0270	0.0156	1.73	0.0842	0.3130	0.4074

* Eugene PD is omitted from the Middle Eastern analysis since this agency does not log a Middle Eastern race/ethnicity

Table D.1.3
Tier 2 - All

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Asian	Citation	0.2896	0.0026	0.0070	0.38	0.7072	0.2660	0.3081
	Search	0.0109	-0.0026	0.0017	-1.54	0.1239	0.0084	0.0185
	Arrest	0.0126	-0.0065	0.0018	-3.64	0.0003	0.0137	0.0245
	Citation, Search, or Arrest	0.2946	-0.0039	0.0069	-0.57	0.5708	0.2776	0.3195
Black	Citation	0.3252	0.0259	0.0065	3.99	0.0001	0.2798	0.3190
	Search	0.0271	0.0058	0.0023	2.50	0.0124	0.0142	0.0283
	Arrest	0.0356	0.0052	0.0026	1.98	0.0477	0.0224	0.0383
	Citation, Search, or Arrest	0.3486	0.0286	0.0065	4.38	0.0000	0.3003	0.3398
Latinx	Citation	0.3682	0.0581	0.0039	14.99	0.0000	0.2985	0.3219
	Search	0.0169	0.0006	0.0011	0.60	0.5512	0.0131	0.0195
	Arrest	0.0251	0.0006	0.0013	0.48	0.6314	0.0206	0.0283
	Citation, Search, or Arrest	0.3830	0.0568	0.0039	14.68	0.0000	0.3145	0.3379
Middle Eastern	Citation	0.2812	-0.0104	0.0113	-0.92	0.3589	0.2574	0.3258
	Search	0.0042	-0.0114	0.0018	-6.26	0.0000	0.0101	0.0211
	Arrest	0.0068	-0.0153	0.0023	-6.78	0.0000	0.0153	0.0290
	Citation, Search, or Arrest	0.2825	-0.0213	0.0112	-1.90	0.0577	0.2698	0.3379
Native	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						

Table D.1.4
Albany PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.4013	0.0132	0.0394	0.34	0.7372	0.2803	0.4958
	Search	0.0765	0.0045	0.0209	0.22	0.8295	0.0149	0.1291
	Arrest	0.0947	0.0183	0.0216	0.84	0.3983	0.0173	0.1357
	Citation, Search, or Arrest	0.4263	0.0395	0.0362	1.09	0.2753	0.2876	0.4860
Latinx	Citation	0.4811	0.0791	0.0203	3.89	0.0001	0.3464	0.4576
	Search	0.0450	-0.0184	0.0087	-2.12	0.0337	0.0397	0.0871
	Arrest	0.0501	-0.0180	0.0084	-2.14	0.0324	0.0450	0.0912
	Citation, Search, or Arrest	0.4545	0.0650	0.0188	3.45	0.0006	0.3381	0.4411

Table D.1.5
Ashland PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.2521	-0.0095	0.0269	-0.35	0.7235	0.1879	0.3354
	Search	0.0124	0.0024	0.0068	0.35	0.7285	-0.0086	0.0287
	Arrest	0.0124	-0.0015	0.0068	-0.22	0.8234	-0.0047	0.0327
	Citation, Search, or Arrest	0.2614	-0.0138	0.0270	-0.51	0.6097	0.2013	0.3491

Table D.1.6
Beaverton PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3407	0.0147	0.0128	1.14	0.2530	0.2909	0.3612
	Search	0.0451	-0.0006	0.0059	-0.10	0.9229	0.0296	0.0618
	Arrest	0.0681	-0.0002	0.0070	-0.02	0.9804	0.0490	0.0875
	Citation, Search, or Arrest	0.3965	0.0143	0.0132	1.08	0.2794	0.3461	0.4183
Latinx	Citation	0.3755	0.0287	0.0093	3.09	0.0020	0.3215	0.3723
	Search	0.0481	0.0103	0.0043	2.41	0.0160	0.0261	0.0495
	Arrest	0.0725	0.0148	0.0051	2.90	0.0037	0.0437	0.0717
	Citation, Search, or Arrest	0.4296	0.0410	0.0096	4.29	0.0000	0.3624	0.4148

Table D.1.7
Bend PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.6271	-0.0305	0.0414	-0.74	0.4611	0.5444	0.7709
	Search	0.0635	0.0144	0.0168	0.86	0.3924	0.0032	0.0951
	Arrest	0.0709	0.0074	0.0178	0.41	0.6788	0.0147	0.1122
	Citation, Search, or Arrest	0.6535	-0.0309	0.0387	-0.80	0.4257	0.5784	0.7904
Latinx	Citation	0.6816	0.0055	0.0185	0.30	0.7672	0.6255	0.7267
	Search	0.0461	0.0113	0.0079	1.43	0.1536	0.0130	0.0565
	Arrest	0.0683	0.0198	0.0093	2.12	0.0343	0.0230	0.0741
	Citation, Search, or Arrest	0.7067	0.0136	0.0176	0.78	0.4383	0.6448	0.7412

Table D.1.8
Benton Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.1534	-0.0516	0.0267	-1.94	0.0529	0.1320	0.2779
	Search	0.0000	-0.0134	0.0025	-5.44	0.0000	0.0066	0.0201
	Arrest	0.0000	-0.0151	0.0025	-6.13	0.0000	0.0083	0.0218
	Citation, Search, or Arrest	0.1694	-0.0435	0.0279	-1.56	0.1185	0.1366	0.2892
Latinx	Citation	0.2296	0.0099	0.0200	0.50	0.6204	0.1651	0.2743
	Search	0.0051	-0.0050	0.0040	-1.26	0.2081	-0.0008	0.0210
	Arrest	0.0049	-0.0084	0.0040	-2.10	0.0360	0.0023	0.0242
	Citation, Search, or Arrest	0.2304	0.0054	0.0198	0.27	0.7858	0.1707	0.2793

Table D.1.9
Canby PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.2071	0.0394	0.0219	1.79	0.0729	0.1077	0.2279
	Search	0.0300	0.0028	0.0092	0.30	0.7613	0.0019	0.0525
	Arrest	0.0411	0.0059	0.0106	0.55	0.5800	0.0062	0.0642
	Citation, Search, or Arrest	0.2443	0.0419	0.0226	1.86	0.0636	0.1405	0.2642

Table D.1.10
Central Point PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.0885	-0.0013	0.0206	-0.06	0.9491	0.0334	0.1463
	Search	0.0131	-0.0138	0.0085	-1.63	0.1039	0.0037	0.0502
	Arrest	0.0174	-0.0226	0.0099	-2.30	0.0216	0.0130	0.0670
	Citation, Search, or Arrest	0.1522	0.0066	0.0261	0.25	0.7988	0.0742	0.2169

Table D.1.11
Clackamas Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3393	0.0438	0.0131	3.34	0.0008	0.2596	0.3313
	Search	0.0191	-0.0001	0.0041	-0.02	0.9830	0.0079	0.0305
	Arrest	0.0391	0.0021	0.0057	0.36	0.7168	0.0215	0.0526
	Citation, Search, or Arrest	0.3692	0.0455	0.0134	3.39	0.0007	0.2869	0.3604
Latinx	Citation	0.3574	0.0371	0.0092	4.03	0.0001	0.2952	0.3455
	Search	0.0167	-0.0001	0.0026	-0.02	0.9836	0.0096	0.0239
	Arrest	0.0310	-0.0017	0.0035	-0.48	0.6289	0.0232	0.0423
	Citation, Search, or Arrest	0.3779	0.0338	0.0093	3.63	0.0003	0.3186	0.3697

Table D.1.12
Corvallis PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI
Black	Citation	Failed to Converge/Not Enough Data for Analysis					
	Search	Failed to Converge/Not Enough Data for Analysis					
	Arrest	Failed to Converge/Not Enough Data for Analysis					
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis					
Latinx	Citation	Failed to Converge/Not Enough Data for Analysis					
	Search	Failed to Converge/Not Enough Data for Analysis					
	Arrest	Failed to Converge/Not Enough Data for Analysis					
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis					

Table D.1.13
Deschutes Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2368	0.0447	0.0449	0.99	0.3201	0.0691	0.3152
	Search	0.0000	-0.0142	0.0044	-3.25	0.0011	0.0023	0.0261
	Arrest	0.0000	-0.0215	0.0051	-4.20	0.0000	0.0075	0.0355
	Citation, Search, or Arrest	0.2500	0.0276	0.0465	0.59	0.5533	0.0951	0.3497
Latinx	Citation	0.2312	0.0493	0.0211	2.34	0.0194	0.1241	0.2396
	Search	0.0050	-0.0095	0.0043	-2.22	0.0265	0.0028	0.0263
	Arrest	0.0149	-0.0071	0.0065	-1.09	0.2757	0.0042	0.0396
	Citation, Search, or Arrest	0.2500	0.0374	0.0216	1.73	0.0839	0.1535	0.2718

Table D.1.14
Douglas Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3878	0.1081	0.0648	1.67	0.0950	0.1024	0.4569
	Search	0.0200	-0.0026	0.0179	-0.14	0.8852	-0.0264	0.0715
	Arrest	0.0392	0.0052	0.0254	0.20	0.8396	-0.0356	0.1037
	Citation, Search, or Arrest	0.4118	0.1019	0.0637	1.60	0.1099	0.1354	0.4843
Latinx	Citation	0.2994	0.0155	0.0350	0.44	0.6580	0.1881	0.3797
	Search	0.0234	-0.0100	0.0111	-0.90	0.3700	0.0029	0.0638
	Arrest	0.0457	-0.0034	0.0156	-0.22	0.8271	0.0064	0.0919
	Citation, Search, or Arrest	0.3314	0.0059	0.0346	0.17	0.8649	0.2309	0.4202

Table D.1.15
Eugene PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3718	0.0167	0.0170	0.99	0.3241	0.3086	0.4015
	Search	0.0556	0.0096	0.0080	1.21	0.2274	0.0241	0.0678
	Arrest	0.0636	0.0085	0.0085	1.01	0.3142	0.0319	0.0783
	Citation, Search, or Arrest	0.4102	0.0168	0.0168	1.00	0.3176	0.3476	0.4394
Latinx	Citation	0.4558	0.0908	0.0168	5.41	0.0000	0.3191	0.4109
	Search	0.0397	-0.0012	0.0067	-0.17	0.8616	0.0226	0.0591
	Arrest	0.0434	-0.0067	0.0069	-0.96	0.3359	0.0311	0.0691
	Citation, Search, or Arrest	0.4777	0.0805	0.0165	4.87	0.0000	0.3520	0.4425

Table D.1.16
Forest Grove PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3253	0.0525	0.0499	1.05	0.2928	0.1362	0.4094
	Search	0.0119	0.0051	0.0120	0.43	0.6681	-0.0260	0.0396
	Arrest	0.0235	-0.0020	0.0170	-0.12	0.9078	-0.0209	0.0719
	Citation, Search, or Arrest	0.3412	0.0491	0.0502	0.98	0.3283	0.1546	0.4295
Latinx	Citation	0.3580	0.0999	0.0178	5.61	0.0000	0.2094	0.3069
	Search	0.0122	0.0053	0.0039	1.36	0.1731	-0.0038	0.0176
	Arrest	0.0241	-0.0017	0.0058	-0.29	0.7741	0.0099	0.0416
	Citation, Search, or Arrest	0.3768	0.0992	0.0179	5.55	0.0000	0.2287	0.3265

Table D.1.17
Grants Pass DPS Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2667	-0.0206	0.0580	-0.35	0.7230	0.1285	0.4459
	Search	0.0625	0.0313	0.0352	0.89	0.3732	-0.0651	0.1275
	Arrest	0.0625	0.0300	0.0352	0.85	0.3934	-0.0638	0.1288
	Citation, Search, or Arrest	0.3125	-0.0109	0.0581	-0.19	0.8516	0.1644	0.4824
Latinx	Citation	0.2614	-0.0196	0.0322	-0.61	0.5421	0.1929	0.3691
	Search	0.0330	-0.0038	0.0134	-0.29	0.7747	0.0001	0.0735
	Arrest	0.0383	-0.0021	0.0141	-0.15	0.8820	0.0018	0.0789
	Citation, Search, or Arrest	0.2951	-0.0207	0.0329	-0.63	0.5289	0.2257	0.4059

Table D.1.18
Gresham PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3993	0.0003	0.0175	0.02	0.9876	0.3512	0.4468
	Search	0.0150	-0.0062	0.0049	-1.27	0.2040	0.0078	0.0347
	Arrest	0.0273	-0.0061	0.0064	-0.96	0.3394	0.0159	0.0509
	Citation, Search, or Arrest	0.4237	-0.0006	0.0177	-0.03	0.9742	0.3758	0.4727
Latinx	Citation	0.4112	0.0001	0.0162	0.00	0.9963	0.3667	0.4556
	Search	0.0239	0.0042	0.0054	0.78	0.4328	0.0050	0.0344
	Arrest	0.0345	0.0046	0.0064	0.72	0.4710	0.0124	0.0474
	Citation, Search, or Arrest	0.4374	0.0045	0.0164	0.27	0.7837	0.3880	0.4779

Table D.1.19
Hermiston PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3043	-0.0008	0.0519	-0.02	0.9873	0.1631	0.4473
	Search	0.0316	0.0157	0.0184	0.86	0.3920	-0.0344	0.0661
	Arrest	0.0417	0.0147	0.0211	0.69	0.4873	-0.0308	0.0848
	Citation, Search, or Arrest	0.3333	0.0102	0.0521	0.20	0.8447	0.1807	0.4656
Latinx	Citation	0.3790	0.0770	0.0138	5.58	0.0000	0.2642	0.3398
	Search	0.0134	0.0006	0.0034	0.17	0.8654	0.0035	0.0222
	Arrest	0.0272	0.0068	0.0047	1.46	0.1453	0.0077	0.0332
	Citation, Search, or Arrest	0.3949	0.0776	0.0139	5.60	0.0000	0.2794	0.3552

Table D.1.20
Hillsboro PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2430	-0.0126	0.0184	-0.68	0.4938	0.2052	0.3061
	Search	0.0037	-0.0130	0.0033	-3.91	0.0001	0.0076	0.0259
	Arrest	0.0146	-0.0134	0.0057	-2.34	0.0195	0.0123	0.0437
	Citation, Search, or Arrest	0.2578	-0.0226	0.0190	-1.19	0.2343	0.2284	0.3322
Latinx	Citation	0.3547	0.0950	0.0112	8.51	0.0000	0.2292	0.2903
	Search	0.0164	0.0025	0.0031	0.81	0.4177	0.0054	0.0224
	Arrest	0.0315	0.0072	0.0041	1.74	0.0820	0.0130	0.0356
	Citation, Search, or Arrest	0.3772	0.0975	0.0113	8.61	0.0000	0.2487	0.3107

Table D.1.21
Hood River Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.1993	0.0386	0.0257	1.50	0.1326	0.0904	0.2310
	Search	0.0202	0.0080	0.0090	0.89	0.3716	-0.0124	0.0368
	Arrest	0.0201	0.0027	0.0089	0.31	0.7584	-0.0071	0.0418
	Citation, Search, or Arrest	0.2181	0.0428	0.0263	1.63	0.1038	0.1033	0.2473

Table D.1.22
Jackson Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.5043	0.0103	0.0401	0.26	0.7973	0.3842	0.6038
	Search	0.0250	0.0063	0.0142	0.44	0.6583	-0.0201	0.0576
	Arrest	0.0250	-0.0019	0.0143	-0.13	0.8927	-0.0122	0.0660
	Citation, Search, or Arrest	0.5167	0.0096	0.0403	0.24	0.8127	0.3967	0.6175
Latinx	Citation	0.5733	0.0247	0.0163	1.52	0.1292	0.5041	0.5931
	Search	0.0052	-0.0069	0.0031	-2.23	0.0256	0.0036	0.0205
	Arrest	0.0153	-0.0036	0.0049	-0.75	0.4547	0.0056	0.0323
	Citation, Search, or Arrest	0.5837	0.0241	0.0163	1.48	0.1397	0.5149	0.6042

Table D.1.23
Keizer PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.3659	0.0714	0.0250	2.86	0.0042	0.2262	0.3628
	Search	0.0021	0.0006	0.0023	0.25	0.8028	-0.0049	0.0079
	Arrest	0.0082	0.0027	0.0049	0.54	0.5872	-0.0079	0.0191
	Citation, Search, or Arrest	0.3711	0.0720	0.0250	2.88	0.0039	0.2307	0.3675

Table D.1.24
Klamath Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.3488	0.0810	0.0712	1.14	0.2552	0.0730	0.4627
	Search	0.0000	-0.0231	0.0150	-1.53	0.1248	-0.0181	0.0642
	Arrest	0.0227	-0.0088	0.0269	-0.33	0.7442	-0.0420	0.1050
	Citation, Search, or Arrest	0.3636	0.0577	0.0770	0.75	0.4536	0.0953	0.5166

Table D.1.25
Klamath Falls PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3415	0.0376	0.0442	0.85	0.3952	0.1830	0.4248
	Search	0.0465	0.0199	0.0157	1.27	0.2042	-0.0164	0.0695
	Arrest	0.0538	0.0232	0.0170	1.36	0.1725	-0.0159	0.0772
	Citation, Search, or Arrest	0.3923	0.0659	0.0435	1.51	0.1301	0.2072	0.4456
Latinx	Citation	0.2714	0.0025	0.0221	0.11	0.9093	0.2084	0.3293
	Search	0.0189	0.0026	0.0066	0.40	0.6885	-0.0019	0.0343
	Arrest	0.0230	0.0048	0.0073	0.67	0.5056	-0.0018	0.0380
	Citation, Search, or Arrest	0.3048	0.0164	0.0226	0.73	0.4684	0.2266	0.3503

Table D.1.26
Lake Oswego PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.4620	-0.0586	0.0389	-1.51	0.1313	0.4143	0.6270
	Search	0.0058	0.0042	0.0058	0.73	0.4681	-0.0144	0.0175
	Arrest	0.0115	0.0038	0.0083	0.45	0.6523	-0.0150	0.0305
	Citation, Search, or Arrest	0.4655	-0.0576	0.0387	-1.49	0.1369	0.4172	0.6290
Latinx	Citation	0.5698	0.0151	0.0255	0.59	0.5542	0.4848	0.6245
	Search	0.0029	0.0015	0.0030	0.52	0.6011	-0.0067	0.0094
	Arrest	0.0058	-0.0009	0.0043	-0.22	0.8291	-0.0052	0.0186
	Citation, Search, or Arrest	0.5723	0.0134	0.0256	0.52	0.6000	0.4888	0.6289

Table D.1.27
Lane Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3099	-0.0052	0.0548	-0.10	0.9237	0.1652	0.4650
	Search	0.0405	0.0219	0.0232	0.95	0.3445	-0.0448	0.0821
	Arrest	0.0533	0.0237	0.0267	0.89	0.3746	-0.0434	0.1027
	Citation, Search, or Arrest	0.3467	0.0074	0.0552	0.13	0.8938	0.1883	0.4903
Latinx	Citation	0.3636	0.0397	0.0387	1.03	0.3041	0.2181	0.4297
	Search	0.0060	-0.0084	0.0071	-1.19	0.2326	-0.0049	0.0338
	Arrest	0.0120	-0.0122	0.0097	-1.25	0.2103	-0.0024	0.0507
	Citation, Search, or Arrest	0.3713	0.0281	0.0387	0.73	0.4676	0.2374	0.4490

Table D.1.28
Lebanon PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI
Black	Citation	Failed to Converge/Not Enough Data for Analysis					
	Search	Failed to Converge/Not Enough Data for Analysis					
	Arrest	Failed to Converge/Not Enough Data for Analysis					
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis					
Latinx	Citation	Failed to Converge/Not Enough Data for Analysis					
	Search	Failed to Converge/Not Enough Data for Analysis					
	Arrest	Failed to Converge/Not Enough Data for Analysis					
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis					

Table D.1.29
Lincoln City PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3444	0.0956	0.0517	1.85	0.0646	0.1073	0.3904
	Search	0.0000	-0.0084	0.0036	-2.37	0.0180	-0.0013	0.0182
	Arrest	0.0110	-0.0043	0.0103	-0.42	0.6764	-0.0129	0.0435
	Citation, Search, or Arrest	0.3516	0.0896	0.0511	1.75	0.0796	0.1222	0.4019
Latinx	Citation	0.2590	0.0296	0.0292	1.01	0.3115	0.1494	0.3094
	Search	0.0040	-0.0047	0.0047	-1.01	0.3139	-0.0041	0.0214
	Arrest	0.0079	-0.0062	0.0063	-0.98	0.3273	-0.0032	0.0314
	Citation, Search, or Arrest	0.2648	0.0240	0.0294	0.82	0.4137	0.1604	0.3212

Table D.1.30
Lincoln Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.2961	0.1038	0.0311	3.33	0.0009	0.1072	0.2776
	Search	0.0043	-0.0029	0.0048	-0.61	0.5389	-0.0059	0.0203
	Arrest	0.0085	-0.0015	0.0066	-0.23	0.8175	-0.0080	0.0280
	Citation, Search, or Arrest	0.3021	0.1001	0.0312	3.21	0.0013	0.1166	0.2875

Table D.1.31
Linn Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	0.0101	-0.0040	0.0054	-0.73	0.4625	-0.0007	0.0289
	Citation, Search, or Arrest	0.4886	0.1189	0.0260	4.58	0.0000	0.2986	0.4408

Table D.1.32
Marion Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.7764	-0.0421	0.0185	-2.27	0.0230	0.7678	0.8693
	Search	0.0240	0.0121	0.0072	1.69	0.0911	-0.0077	0.0315
	Arrest	0.0263	0.0091	0.0074	1.22	0.2230	-0.0031	0.0376
	Citation, Search, or Arrest	0.7823	-0.0395	0.0184	-2.15	0.0315	0.7715	0.8721
Latinx	Citation	0.8026	0.0264	0.0081	3.24	0.0012	0.7539	0.7985
	Search	0.0202	0.0073	0.0029	2.50	0.0125	0.0049	0.0209
	Arrest	0.0294	0.0104	0.0034	3.05	0.0023	0.0096	0.0284
	Citation, Search, or Arrest	0.8085	0.0274	0.0081	3.40	0.0007	0.7591	0.8032

Table D.1.33
McMinnville PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2778	0.0557	0.0607	0.92	0.3584	0.0559	0.3882
	Search	0.0182	0.0054	0.0186	0.29	0.7710	-0.0382	0.0637
	Arrest	0.0182	0.0030	0.0186	0.16	0.8710	-0.0359	0.0662
	Citation, Search, or Arrest	0.2909	0.0527	0.0610	0.86	0.3872	0.0713	0.4051
Latinx	Citation	0.3185	0.0931	0.0206	4.52	0.0000	0.1690	0.2818
	Search	0.0082	-0.0016	0.0042	-0.38	0.7025	-0.0018	0.0214
	Arrest	0.0114	-0.0011	0.0049	-0.23	0.8207	-0.0010	0.0260
	Citation, Search, or Arrest	0.3295	0.0910	0.0206	4.41	0.0000	0.1820	0.2950

Table D.1.34
Medford PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3308	0.0419	0.0408	1.03	0.3051	0.1772	0.4007
	Search	0.0567	0.0048	0.0196	0.24	0.8081	-0.0018	0.1057
	Arrest	0.0699	0.0067	0.0213	0.31	0.7535	0.0049	0.1216
	Citation, Search, or Arrest	0.3986	0.0604	0.0427	1.41	0.1573	0.2212	0.4551
Latinx	Citation	0.3062	0.0210	0.0228	0.92	0.3575	0.2228	0.3476
	Search	0.0340	-0.0011	0.0091	-0.12	0.9016	0.0102	0.0601
	Arrest	0.0381	-0.0024	0.0097	-0.25	0.8050	0.0140	0.0670
	Citation, Search, or Arrest	0.3411	0.0203	0.0234	0.87	0.3849	0.2567	0.3848

Table D.1.35
Milwaukie PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.4357	-0.0094	0.0228	-0.41	0.6786	0.3827	0.5075
	Search	0.0029	-0.0013	0.0031	-0.40	0.6901	-0.0044	0.0127
	Arrest	0.0029	-0.0028	0.0033	-0.85	0.3927	-0.0032	0.0146
	Citation, Search, or Arrest	0.4402	-0.0099	0.0231	-0.43	0.6674	0.3869	0.5134
Latinx	Citation	0.4913	0.0227	0.0211	1.08	0.2807	0.4108	0.5262
	Search	0.0065	0.0039	0.0038	1.02	0.3069	-0.0078	0.0130
	Arrest	0.0065	0.0029	0.0038	0.76	0.4481	-0.0068	0.0141
	Citation, Search, or Arrest	0.4967	0.0258	0.0213	1.21	0.2262	0.4127	0.5292

Table D.1.36
Multnomah Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.1734	0.0049	0.0140	0.35	0.7265	0.1301	0.2068
	Search	0.0352	0.0091	0.0066	1.38	0.1671	0.0079	0.0441
	Arrest	0.0635	0.0221	0.0086	2.58	0.0099	0.0181	0.0649
	Citation, Search, or Arrest	0.2294	0.0220	0.0153	1.45	0.1484	0.1656	0.2491
Latinx	Citation	0.2307	0.0413	0.0129	3.20	0.0014	0.1541	0.2247
	Search	0.0177	-0.0047	0.0042	-1.12	0.2635	0.0108	0.0340
	Arrest	0.0323	-0.0013	0.0055	-0.24	0.8072	0.0187	0.0486
	Citation, Search, or Arrest	0.2604	0.0396	0.0135	2.94	0.0033	0.1839	0.2576

Table D.1.37
Newberg-Dundee PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3212	0.0868	0.0387	2.24	0.0250	0.1285	0.3403
	Search	0.0072	0.0032	0.0073	0.43	0.6669	-0.0160	0.0242
	Arrest	0.0072	-0.0006	0.0074	-0.07	0.9402	-0.0126	0.0282
	Citation, Search, or Arrest	0.3261	0.0831	0.0385	2.16	0.0309	0.1376	0.3484
Latinx	Citation	0.2931	0.0747	0.0165	4.51	0.0000	0.1732	0.2638
	Search	0.0077	0.0046	0.0032	1.41	0.1582	-0.0057	0.0120
	Arrest	0.0103	0.0037	0.0038	0.96	0.3354	-0.0039	0.0170
	Citation, Search, or Arrest	0.3042	0.0796	0.0167	4.77	0.0000	0.1790	0.2704

Table D.1.38
OHSU PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.0870	0.0523	0.0348	1.50	0.1326	-0.0605	0.1298
	Search	0.0000	-0.0085	0.0059	-1.45	0.1471	-0.0076	0.0246
	Arrest	0.0000	-0.0085	0.0059	-1.45	0.1471	-0.0076	0.0246
	Citation, Search, or Arrest	0.0870	0.0407	0.0355	1.15	0.2519	-0.0508	0.1434
Latinx	Citation	0.1154	0.0729	0.0467	1.56	0.1181	-0.0852	0.1702
	Search	0.0189	-0.0011	0.0175	-0.07	0.9482	-0.0280	0.0680
	Arrest	0.0189	-0.0011	0.0175	-0.07	0.9482	-0.0280	0.0680
	Citation, Search, or Arrest	0.1698	0.0992	0.0530	1.87	0.0614	-0.0745	0.2158

Table D.1.39
Oregon City PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.4091	-0.0034	0.0376	-0.09	0.9290	0.3095	0.5154
	Search	0.0222	0.0067	0.0105	0.63	0.5257	-0.0131	0.0443
	Arrest	0.0383	0.0121	0.0135	0.89	0.3714	-0.0109	0.0632
	Citation, Search, or Arrest	0.4317	0.0050	0.0373	0.13	0.8943	0.3246	0.5289
Latinx	Citation	0.4148	-0.0240	0.0215	-1.12	0.2637	0.3800	0.4976
	Search	0.0119	0.0014	0.0049	0.28	0.7793	-0.0028	0.0238
	Arrest	0.0196	0.0013	0.0063	0.21	0.8367	0.0012	0.0355
	Citation, Search, or Arrest	0.4263	-0.0213	0.0217	-0.98	0.3249	0.3883	0.5070

Table D.1.40
Oregon State Police Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.4549	0.0820	0.0076	10.82	0.0000	0.3522	0.3937
	Search	0.0106	-0.0034	0.0015	-2.29	0.0221	0.0099	0.0181
	Arrest	0.0205	-0.0028	0.0021	-1.34	0.1794	0.0176	0.0290
	Citation, Search, or Arrest	0.4696	0.0803	0.0075	10.72	0.0000	0.3688	0.4098
Latinx	Citation	0.4628	0.0973	0.0041	23.85	0.0000	0.3543	0.3767
	Search	0.0101	-0.0017	0.0008	-2.10	0.0357	0.0096	0.0140
	Arrest	0.0199	-0.0005	0.0011	-0.45	0.6558	0.0174	0.0235
	Citation, Search, or Arrest	0.4780	0.0968	0.0040	23.96	0.0000	0.3702	0.3923

Table D.1.41
Polk Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2955	0.0736	0.0489	1.50	0.1326	0.0881	0.3557
	Search	0.0538	0.0278	0.0223	1.25	0.2124	-0.0350	0.0870
	Arrest	0.0638	0.0343	0.0243	1.41	0.1573	-0.0369	0.0959
	Citation, Search, or Arrest	0.3404	0.0940	0.0492	1.91	0.0558	0.1119	0.3810
Latinx	Citation	0.3046	0.0933	0.0204	4.58	0.0000	0.1555	0.2670
	Search	0.0331	0.0050	0.0080	0.63	0.5317	0.0063	0.0499
	Arrest	0.0345	0.0022	0.0083	0.26	0.7915	0.0097	0.0550
	Citation, Search, or Arrest	0.3359	0.0966	0.0206	4.68	0.0000	0.1828	0.2958

Table D.1.42
Port of Portland PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2292	0.0832	0.0192	4.32	0.0000	0.0933	0.1986
	Search	0.0283	0.0064	0.0083	0.77	0.4416	-0.0008	0.0447
	Arrest	0.0361	0.0072	0.0093	0.78	0.4379	0.0034	0.0544
	Citation, Search, or Arrest	0.2590	0.0839	0.0205	4.09	0.0000	0.1190	0.2313
Latinx	Citation	0.2135	0.0529	0.0234	2.25	0.0241	0.0964	0.2248
	Search	0.0172	0.0025	0.0073	0.35	0.7300	-0.0053	0.0348
	Arrest	0.0201	-0.0010	0.0081	-0.12	0.9027	-0.0011	0.0432
	Citation, Search, or Arrest	0.2292	0.0482	0.0240	2.01	0.0440	0.1154	0.2465

Table D.1.43
Portland PB Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3718	-0.0706	0.0064	-10.98	0.0000	0.4248	0.4600
	Search	0.0293	0.0099	0.0024	4.20	0.0000	0.0129	0.0258
	Arrest	0.0609	0.0176	0.0033	5.29	0.0000	0.0342	0.0524
	Citation, Search, or Arrest	0.4466	-0.0341	0.0067	-5.06	0.0000	0.4623	0.4992
Latinx	Citation	0.5417	0.0190	0.0079	2.41	0.0161	0.5011	0.5443
	Search	0.0165	0.0009	0.0023	0.40	0.6908	0.0094	0.0217
	Arrest	0.0377	0.0028	0.0033	0.83	0.4049	0.0259	0.0441
	Citation, Search, or Arrest	0.5736	0.0224	0.0080	2.82	0.0049	0.5294	0.5730

Table D.1.44
Redmond PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3571	0.1375	0.0790	1.74	0.0817	0.0034	0.4358
	Search	0.0345	0.0282	0.0340	0.83	0.4056	-0.0867	0.0992
	Arrest	0.0667	0.0504	0.0440	1.15	0.2512	-0.1041	0.1366
	Citation, Search, or Arrest	0.4000	0.1674	0.0813	2.06	0.0396	0.0100	0.4552
Latinx	Citation	0.3694	0.1332	0.0327	4.07	0.0000	0.1466	0.3257
	Search	0.0045	-0.0030	0.0051	-0.59	0.5532	-0.0064	0.0213
	Arrest	0.0133	-0.0024	0.0082	-0.29	0.7734	-0.0068	0.0382
	Citation, Search, or Arrest	0.3778	0.1279	0.0330	3.88	0.0001	0.1597	0.3401

Table D.1.45
Roseburg PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.7698	0.1670	0.0299	5.59	0.0000	0.5211	0.6846
	Search	0.0079	-0.0348	0.0101	-3.44	0.0006	0.0150	0.0703
	Arrest	0.0526	-0.0399	0.0210	-1.89	0.0582	0.0349	0.1501
	Citation, Search, or Arrest	0.7820	0.1346	0.0282	4.76	0.0000	0.5701	0.7247
Latinx	Citation	0.6070	0.0530	0.0261	2.03	0.0427	0.4825	0.6256
	Search	0.0138	-0.0251	0.0080	-3.14	0.0017	0.0171	0.0607
	Arrest	0.0372	-0.0515	0.0124	-4.14	0.0000	0.0546	0.1227
	Citation, Search, or Arrest	0.6216	0.0214	0.0250	0.86	0.3916	0.5318	0.6686

Table D.1.46
Salem PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.5000	-0.0492	0.0291	-1.69	0.0912	0.4695	0.6288
	Search	0.0692	0.0265	0.0151	1.75	0.0803	0.0013	0.0842
	Arrest	0.0967	0.0433	0.0176	2.46	0.0138	0.0052	0.1015
	Citation, Search, or Arrest	0.5465	-0.0342	0.0273	-1.25	0.2102	0.5060	0.6553
Latinx	Citation	0.5952	0.0256	0.0123	2.08	0.0372	0.5360	0.6033
	Search	0.0653	0.0188	0.0062	3.01	0.0026	0.0293	0.0635
	Arrest	0.0768	0.0187	0.0067	2.78	0.0055	0.0396	0.0765
	Citation, Search, or Arrest	0.6224	0.0318	0.0118	2.69	0.0071	0.5582	0.6230

Table D.1.47
Springfield PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.3037	-0.0058	0.0179	-0.32	0.7467	0.2606	0.3584
	Search	0.0621	0.0166	0.0095	1.74	0.0814	0.0195	0.0716
	Arrest	0.0789	0.0142	0.0104	1.36	0.1742	0.0361	0.0933
	Citation, Search, or Arrest	0.3676	0.0067	0.0186	0.36	0.7173	0.3100	0.4116

Table D.1.48
Tigard PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2364	0.0153	0.0227	0.67	0.5006	0.1588	0.2833
	Search	0.0294	0.0084	0.0094	0.89	0.3741	-0.0047	0.0468
	Arrest	0.0351	0.0076	0.0103	0.74	0.4622	-0.0007	0.0557
	Citation, Search, or Arrest	0.2632	0.0196	0.0235	0.83	0.4039	0.1792	0.3079
Latinx	Citation	0.3464	0.0824	0.0177	4.65	0.0000	0.2155	0.3125
	Search	0.0120	-0.0034	0.0044	-0.79	0.4294	0.0036	0.0274
	Arrest	0.0160	-0.0045	0.0050	-0.90	0.3696	0.0067	0.0342
	Citation, Search, or Arrest	0.3582	0.0776	0.0178	4.37	0.0000	0.2320	0.3292

Table D.1.49
Tualatin PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.3991	-0.0497	0.0310	-1.60	0.1093	0.3638	0.5337
	Search	0.0000	-0.0071	0.0016	-4.43	0.0000	0.0027	0.0115
	Arrest	0.0268	0.0107	0.0109	0.98	0.3267	-0.0137	0.0459
	Citation, Search, or Arrest	0.4196	-0.0386	0.0315	-1.23	0.2205	0.3720	0.5445
Latinx	Citation	0.4929	0.0519	0.0175	2.96	0.0031	0.3930	0.4889
	Search	0.0141	0.0055	0.0044	1.25	0.2100	-0.0036	0.0206
	Arrest	0.0277	0.0113	0.0061	1.86	0.0636	-0.0003	0.0331
	Citation, Search, or Arrest	0.5082	0.0581	0.0175	3.31	0.0009	0.4020	0.4981

Table D.1.50
U of O PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI
Black	Citation	Failed to Converge/Not Enough Data for Analysis					
	Search	Failed to Converge/Not Enough Data for Analysis					
	Arrest	Failed to Converge/Not Enough Data for Analysis					
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis					
Latinx	Citation	Failed to Converge/Not Enough Data for Analysis					
	Search	Failed to Converge/Not Enough Data for Analysis					
	Arrest	Failed to Converge/Not Enough Data for Analysis					
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis					

Table D.1.51
Washington Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2206	-0.0055	0.0116	-0.47	0.6381	0.1943	0.2579
	Search	0.0261	0.0005	0.0045	0.12	0.9038	0.0132	0.0379
	Arrest	0.0393	-0.0035	0.0055	-0.63	0.5311	0.0277	0.0579
	Citation, Search, or Arrest	0.2587	-0.0071	0.0123	-0.57	0.5655	0.2320	0.2995
Latinx	Citation	0.2816	0.0474	0.0072	6.55	0.0000	0.2144	0.2540
	Search	0.0314	0.0101	0.0027	3.72	0.0002	0.0139	0.0287
	Arrest	0.0443	0.0079	0.0033	2.42	0.0154	0.0275	0.0454
	Citation, Search, or Arrest	0.3212	0.0555	0.0075	7.37	0.0000	0.2451	0.2863

Table D.1.52
West Linn PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2318	0.0375	0.0274	1.37	0.1716	0.1192	0.2693
	Search	0.0043	0.0022	0.0043	0.50	0.6169	-0.0097	0.0139
	Arrest	0.0043	-0.0004	0.0044	-0.09	0.9246	-0.0073	0.0167
	Citation, Search, or Arrest	0.2350	0.0372	0.0275	1.35	0.1772	0.1225	0.2732
Latinx	Citation	0.2775	0.0800	0.0206	3.89	0.0001	0.1411	0.2537
	Search	0.0019	-0.0002	0.0021	-0.10	0.9212	-0.0035	0.0078
	Arrest	0.0057	0.0012	0.0035	0.33	0.7421	-0.0051	0.0143
	Citation, Search, or Arrest	0.2816	0.0804	0.0206	3.90	0.0001	0.1448	0.2577

Table D.1.53
Woodburn PD Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	Failed to Converge/Not Enough Data for Analysis						
	Search	Failed to Converge/Not Enough Data for Analysis						
	Arrest	Failed to Converge/Not Enough Data for Analysis						
	Citation, Search, or Arrest	Failed to Converge/Not Enough Data for Analysis						
Latinx	Citation	0.3464	0.0974	0.0290	3.36	0.0008	0.1696	0.3284
	Search	0.0449	0.0338	0.0101	3.35	0.0008	-0.0165	0.0387
	Arrest	0.0554	0.0311	0.0124	2.51	0.0122	-0.0097	0.0582
	Citation, Search, or Arrest	0.3875	0.1183	0.0300	3.95	0.0001	0.1871	0.3511

Table D.1.54
Yamhill Co SO Results

Race/Ethnicity	Outcome	Group Mean	ATET	Robust S.E.s	z-score	p-value	95% CI	
Black	Citation	0.2857	-0.0021	0.0682	-0.03	0.9750	0.1012	0.4745
	Search	0.0000	-0.0038	0.0022	-1.75	0.0799	-0.0021	0.0097
	Arrest	0.0000	-0.0111	0.0048	-2.32	0.0201	-0.0020	0.0241
	Citation, Search, or Arrest	0.2857	-0.0117	0.0674	-0.17	0.8624	0.1128	0.4820
Latinx	Citation	0.3149	0.0644	0.0209	3.08	0.0021	0.1932	0.3077
	Search	0.0020	-0.0025	0.0026	-0.96	0.3366	-0.0026	0.0114
	Arrest	0.0098	-0.0019	0.0051	-0.38	0.7046	-0.0022	0.0257
	Citation, Search, or Arrest	0.3275	0.0676	0.0213	3.17	0.0015	0.2014	0.3182

Appendix E – KPT Hit-Rate Analysis Technical Appendix

Model and Assumptions

The hit-rate 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 Knowles, Persico, and Todd (KPT) Hit-Rate Model 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 the white group, perhaps indicating what Becker calls "a taste for discrimination" (a phrase coined to describe economic 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 $\boldsymbol{s}$ and the number of searches of that group which result in finding contraband is represented by $\boldsymbol{f}$:

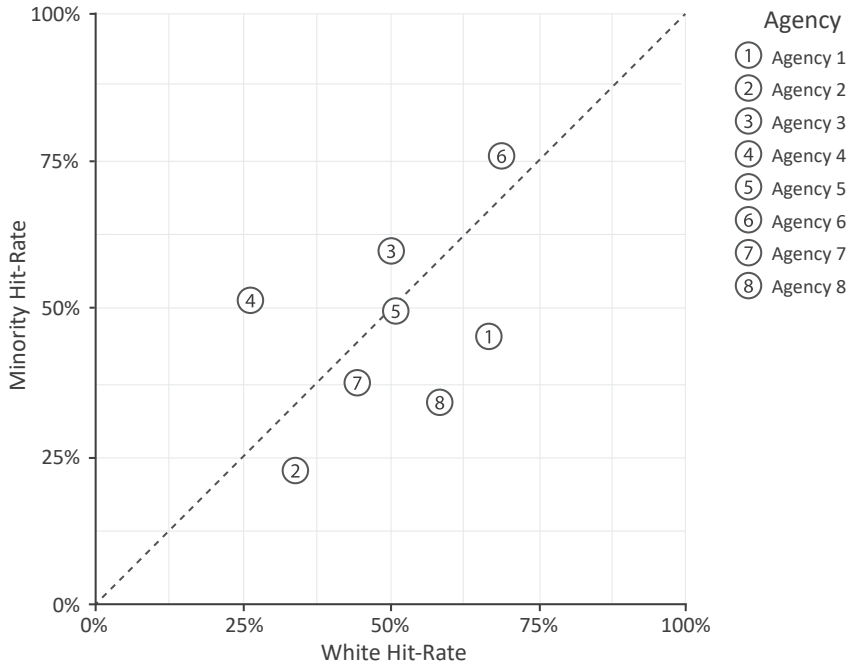
$$\textbf{KPT Hit} - \textbf{Rate} = \frac{\boldsymbol{f}}{\boldsymbol{s}}$$

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-squared test was used to mitigate the test's tendency to produce low p-values due to the discrete nature of the data. A confidence level of 95% with a Bonferroni correction for multiple testing determined significance. Each agency's white hit-rate was compared to each of two groups (Black and Latinx), so a Bonferroni corrected p-value of 0.05/2 = 0.025 or lower was considered indicative of a statistically significant difference between minority and white hit-rates. For statewide results, a Bonferroni corrected p-value of 0.05/5 = 0.01 or lower was considered indicative of a statistically significant difference between minority and white hit-rates. Hit-Rate analyses and accompanying statistical tests were performed with the statistical software R.

Visualization

Figure E1 is a stylized example of the figures that appear in the body of the report and on the accompanying data dashboards found online.[44] Each agency is a single point, with the location of the point indicating the relationship between white and comparison group hit-rates for that agency. Agencies below the diagonal line have a lower minority hit-rate than white hit-rate, which could be indicative of disparate searching practices. An agency falling above or below the line does not, alone, indicate a significant difference between white and minority hit-rates. Significance is noted in the text and tables accompanying figures in the body of the report. It is possible for an agency below the line to have no statistically significant difference between white and minority hit-rates, while an agency "closer" to the line may be found to have a significant difference. The power of the chi-square test to determine significance is dependent upon the sample size (in this case, number of searches). An agency with fewer searches may be "further" below the line than an agency with more searches, yet fail to be statistically significant. Hit-Rate analysis figures are created using the ggplot2 package in R.

Figure E1.
Stylized Example of Hit-Rate Visualization



Limitations

One important assumption of the KPT Hit-Rate 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).

44 Inspiration for these figures came from the display used by the Stanford Open Policing Project, which is located at https://openpolicing.stanford.edu/.

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.