The California Highway Patrol: An Evaluation of Public Contacts in Stop Data

EMILY OWENS and JACLYN ROSENQUIST
The California Policy Lab builds better lives through data-driven policy. We are a project of the University of California, with sites at the Berkeley and Los Angeles campuses.

We would like to thank our partners for their important contributions. At the California Highway Patrol (CHP), we would like to thank Nicholas Mosley for providing extensive knowledge on the inner workings of the CHP. We would also like to thank Janey Rountree, Sean Coffey and Nathan Hess at the California Policy Lab, whose insightful suggestions were invaluable to final production of this report.

This research reflects the views of the authors and not necessarily the views of our funders, our staff, our advisory board, the California Highway Patrol, or the Regents of the University of California. All errors should be attributed to the authors.

CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Benchmarking</td>
<td>10</td>
</tr>
<tr>
<td>Non-Enforcement vs Standard Benchmarks</td>
<td>11</td>
</tr>
<tr>
<td>Division Level Benchmarking</td>
<td>15</td>
</tr>
<tr>
<td>What Determines Stop Rates?</td>
<td>18</td>
</tr>
<tr>
<td>Hit Rate Analysis</td>
<td>23</td>
</tr>
<tr>
<td>Are Stop Outcomes Related to Disparate Stop Rates?</td>
<td>25</td>
</tr>
<tr>
<td>Conclusion</td>
<td>29</td>
</tr>
<tr>
<td>Works Cited</td>
<td>31</td>
</tr>
<tr>
<td>Appendix A: Auxiliary Tables and Figures</td>
<td>32</td>
</tr>
<tr>
<td>Appendix B: Division by Area Office Figures</td>
<td>35</td>
</tr>
<tr>
<td>Appendix C: Veil of Darkness Tests</td>
<td>59</td>
</tr>
</tbody>
</table>

The California Policy Lab builds better lives through data-driven policy. We are a project of the University of California, with sites at the Berkeley and Los Angeles campuses.

We would like to thank our partners for their important contributions. At the California Highway Patrol (CHP), we would like to thank Nicholas Mosley for providing extensive knowledge on the inner workings of the CHP. We would also like to thank Janey Rountree, Sean Coffey and Nathan Hess at the California Policy Lab, whose insightful suggestions were invaluable to final production of this report.

This research reflects the views of the authors and not necessarily the views of our funders, our staff, our advisory board, the California Highway Patrol, or the Regents of the University of California. All errors should be attributed to the authors.
EXECUTIVE SUMMARY

In order to better understand the role that race or ethnicity may play in who is stopped by their officers, the California Highway Patrol (CHP) provided the California Policy Lab (CPL) with a data set of 2,141,817 enforcement stops made by the CHP from January to December of 2019. The data was collected pursuant to California’s Racial and Identity Profiling Act of 2015 (RIPA). RIPA requires that California law enforcement agencies employing peace officers collect data on the perceived demographics of individuals involved in all enforcement stops (e.g. speeding or driving under the influence), the reason for making a stop, the location of stop, as well as the outcome of the stop. As one of the largest law enforcement agencies in California, CHP has been collecting the data required by RIPA since 2018. In addition to enforcement, or “RIPA” stops, the CHP also collects the same data each time an officer interacts with civilians for non-enforcement reasons, including aiding a disabled motorist, responding to a crash, or other motorist services. In order to extend the statistical analysis presented in the 2021 Annual RIPA Board Report¹, we evaluated enforcement stops in combination with non-enforcement stops using two generally accepted approaches to measure racially disparate policing: benchmarking and a hit rate analysis.

Benchmarking involves comparing the composition of people subject to enforcement actions to the composition of people “at risk” - those who could have enforcement actions taken against them. This “at-risk” population is called the “benchmark” population. If members of a particular racial or ethnic group are stopped more often than the benchmark population, then that may suggest evidence of potential bias in who officers are deciding to stop. In this study, we focus on a unique benchmark made possible by CHP’s additional data collection of non-enforcement contacts, and examine how measurements of racial disparities using this benchmark diverge from disparities estimated with more traditional benchmarks, specifically residential or commuting populations.

Hit rate analyses focus on potential bias in an officer’s decision to engage in a particular behavior, such as a search. In its most common application, the “hit rate” is defined as the number of times contraband is found divided by the number of searches conducted. In this study, we conduct a modified hit rate analysis, which examines the severity of an enforcement outcome (e.g. if a stop resulted in a verbal warning, written warning, or a more severe citation).

We focused on the RIPA stop rates of people in four racial and ethnic groups, as perceived by CHP officers: Asian, Black, Hispanic, and White.²

**We Found the Following:**

When residential population data is used as a benchmark, Black people are stopped by CHP in enforcement actions 74% more often than White people. When using non-enforcement stops (e.g. disabled motorist or collision responses, which CHP officers are required to respond to) as a benchmark, Black people are stopped by CHP officers for RIPA-reportable reasons 15% more often than White people. In other words, comparing the people stopped for enforcement reasons to the composition of people CHP encounters for non-enforcement reasons, rather than the composition of the residential population, reduces estimated Black-White disparities in RIPA stops by 79%. Determining which benchmark is the most accurate description of CHP officer decisions depends on how closely the benchmark population mirrors the group of drivers who CHP might potentially stop. Local residential population might not be the most accurate benchmark against which to evaluate CHP officer decisions for a variety of reasons, including the fact that many non-residents utilize California highways where CHP has jurisdiction, and the possibility that driving behavior could vary across identity groups.³ The non-enforcement benchmark may understate racial disparities of CHP enforcement stops if Black, Hispanic, or Asian drivers are, on average, more likely to drive older, higher-mileage cars that require more assistance, than White drivers. These examples highlight a limitation of benchmarking exercises in a context where broader social disparities may interact with police officer decisions.

The CHP has jurisdiction across California, and the state is divided into eight geographic regions, called “Divisions,” each lead by a Division Chief. Using the non-enforcement stop benchmark, we found that officers in different CHP Divisions conduct RIPA and non-enforcement stops at different rates. Black motorists are stopped more than White motorists in every Division, but the disparities ranged from 5% to 41% more Black drivers stopped than White drivers. Across and within Divisions, we found larger racial disparities in areas with smaller residential populations.

A hit rate analysis reveals that as the benchmarked stop rates for Black and Asian drivers increases relative to White benchmarked stop rates, CHP officers are more likely to resolve stops of Black and Asian drivers with a (less severe)

---

² Other racial/ethnic groups are excluded from the analysis due to small number of stops, meaning that reporting descriptive statistics could potentially compromise the stopped individual’s identity: no race provided (0.15% of all stops), multiracial (.78%), Middle eastern/South Asian (5.64%), Native American (.23%), and Pacific Islander (0.54%).

³ We attempt to adjust for any differences in driving behavior using accident data from the CA Statewide Integrated Traffic Records System (SWITRS) database. Across California, Asian and Black drivers are each at fault in about 9% of accidents. This is slightly more accidents for Black drivers and slightly fewer accidents for Asian drivers than their share of the population. White and Hispanic drivers are at fault in 36% and 38% of accidents, which is very close to their population share.
verbal warning and less likely to resolve a stop with a (more severe) citation. While not conclusive evidence of racial bias in stop decisions, this is consistent with CHP officers being more likely to stop Black and Asian drivers than a White driver engaged in the same behavior. We do not observe a comparable relationship for Hispanic drivers, meaning that the hit rate analysis is not informative about the role of race in the decision to stop Hispanic motorists.

**Recommendations:**

Non-enforcement stops represent an opportunity for CHP officers to improve the public’s perception of the CHP. Ensuring that officers are following best practices with regards to professionalism and procedural justice during non-enforcement stops can be thought of as a targeted form of community policing, to the extent that these stops are non-investigatory encounters with people in the same racial and ethnic groups as people who the CHP encounters in an enforcement context.

Our hit rate analysis suggests that in places with larger disparities in benchmarked stop rates, CHP officers may be able to stop fewer Black and Asian drivers without reducing the amount of criminal behavior detected. This may be particularly impactful in low population jurisdictions, which tend to have larger group disparities in how frequently people are stopped.
The California Highway Patrol: An Evaluation of Public Contacts in Stop Data

In 2015, the California legislature passed Assembly Bill 953 - the Racial and Identity Profiling Act (RIPA). Under RIPA, almost all law enforcement agencies in California will be required to report information on almost every enforcement stop made by police officers by 2023. The purpose of RIPA was to allow the California Office of the Attorney General to identify, and take steps to reduce, racial and other identity group disparities (e.g. age, gender expression, sexual orientation, or religion) in law enforcement contacts.

Concerns about racially disparate or “biased” policing likely play an important role in the gap between how satisfied Black and White Californians are with law enforcement; there is a 24-percentage point gap between the percent of Black (50%) and White (74%) California residents who are satisfied with the police (DiCamillo 2020). Like every aspect of policing, bias in policing is a multilayered phenomenon that is more easily understood intuitively than it is measured in administrative data in a precise and quantifiable way.

Within the academic community, two generally accepted approaches to measuring racially disparate policing are known as “Benchmarking” and “Hit Rate” tests. Each of these tests takes a different approach to identify racially disparate policing. Each is good at identifying some aspects of what people experience as - and what the broader public considers to be - racial bias, but does not identify other aspects of this phenomenon, like the way in which officers communicate with civilians, or how different cultural norms may affect how civilians respond to an officer’s commands. While these approaches are not perfect, they nonetheless help communities, researchers, and law enforcement agencies detect where potential racial bias may exist in policing, and work towards solutions that increase both public safety and equity.4

What is Benchmarking?

Benchmarking involves comparing the composition of people subject to enforcement actions to the composition of people who could have enforcement actions taken against them. The later population is called the “benchmark” population. Benchmarking CHP stops reported under RIPA estimates how likely a person in the “at-risk” population is to have a CHP encounter. If the benchmarked estimates vary across groups, it may be the case that CHP officers are targeting people because of their racial or ethnic identity.

4In Appendix C, we describe a third generally accepted statistical test of bias, the Veil of Darkness or “VOD” test. While the state RIPA report conducts a VOD test of CHP data, we do not in this report because we believe the number of people affected by the type of bias that the VOD test examines is almost certainly too small to draw firm statistical conclusions.
A strength of benchmarking is that it can identify actual, group-level, disparities in police stops. The weakness of benchmarking is that it can be difficult, if not impossible, to discern the source of an observed disparity across identity groups. The disparity could be due to CHP actions, or it could be due to differences in the behavior of civilians in different groups.

Since benchmarking does not identify why a stop disparity exists, it can be difficult to draw conclusions that point directly to specific policy changes. The most policy-informative benchmark for law enforcement is one where the proportion of people in a particular racial group is identical to the proportion of people who would be at risk of being stopped by a bias-free officer conducting an objective assessment of their behavior. Carefully selecting an appropriate population at risk of being stopped can make benchmarking more informative about the source of disparities, but incorrect benchmark populations can understate or overstate the amount of disparity directly caused by CHP officer actions.

Differences in rates of criminal behavior may be generated by inequality in opportunities that exist outside of police involvement. For example, suppose RIPA stops were benchmarked against the California residential population. Access to well-resourced schools as a child has been shown to substantially reduce criminal and otherwise risky behavior (Deming 2012). To the extent that people in different identity groups have different access to high quality schools, this can lead to differences in criminal and otherwise risky behavior across groups, and thus differences in the number of people at risk of being stopped. Even if all CHP stops were made in a bias-free way, the number of stops across groups would vary because inequality in education policy leads to differences in average behavior across groups. Population-based benchmarks would, accurately, identify a population-level disparity in police stops. However, these hypothetical disparities would not be due to CHP actions, but rather social inequities that exist independently of CHP decisions. In this case, a population-based benchmark identifies disparities, but is not informative about whether changes in educational or policing policy would be more effective in reducing those disparities in a meaningful way.

At another extreme, benchmarking can understate the extent of racial disparities in policing. This could occur if the benchmarked population itself is affected by biased law enforcement decisions. For example, suppose the population of people arrested by law enforcement was used as a benchmark against which to compare the population of people stopped. If arrest decisions were also biased, then that bias in the benchmark would effectively cancel out any bias in the stop decision, making the benchmarking estimates appear more equal across groups than stop decisions actually are. To consider an extreme example, suppose that officers were more likely to stop a Black driver than a White driver, and also checked for outstanding warrants on every Black driver
they stopped, but never checked for outstanding warrants on White drivers. Since finding an outstanding warrant will lead to an arrest, this racially-motivated scrutiny would increase the number of Black people relative to the number of White people in the benchmark population, leading a conclusion that an “unbiased” set of RIP stops would contain more Black drivers than White drivers.

In this study we conduct a benchmarking analysis focusing on three potential benchmarks: residential population, the residential population that commutes to work, and the population of drivers involved in non-enforcement contacts with CHP officers. While we do not have information on a truly random sample of the population driving on California’s highways, this range of estimates will allow us to examine the importance of different assumptions about the driving population in identifying how CHP officers decide to stop drivers from different racial or ethnic groups.

What is a Hit Rate Test?

Hit rate tests focus on potential bias in an officer’s decision to conduct a search. In its most common application, the hit rate is defined as the number of times contraband is found divided by the number of searches conducted. Hit rate tests incorporate potential behavioral differences across groups and estimate whether CHP actions taken against members of different groups, on average, have similar outcomes.

The basic idea underlying a hit rate test is that officers decide whether to conduct a search based on the intensity of their suspicion that a civilian has violated a law. The hit rate test assumes that the more suspicious an officer judges a civilian to be, the more likely it is that a search will reveal contraband. However, if officers are less likely to find contraband when they search members of a particular racial or ethnic group, then there is a lower hit rate for that group. If there are lower hit rates for a particular group, then it also suggests that officers may be making biased decisions about who to search, and specifically using both a civilian’s race and their behavior in deciding what law enforcement actions to take.

Hit rate tests specifically attempt to estimate whether the threshold of suspicion that triggers a stop or search is higher or lower for different groups. Note that this is a specific type of potential disparate treatment by CHP officers: Would behavior that led an officer to justify a stop or search for someone in one racial group lead that same officer to justify a stop or search for someone in a different racial group?

Like benchmarking, hit rate tests have limitations. If stops of some groups are less likely to result in a finding of criminal behavior than others, hit rate tests suggest that officers may be applying different standards to people in different
groups. However, officers could be using different standards, and still generate equal hit rates across groups, meaning that in some circumstances hit rate tests can fail to detect biased decision making. This specific problem is referred to as the “inframarginality problem.”

Mathematically, typical hit rate tests compare the number of searches that yield contraband to the total number of searches conducted across different groups. Based on that, researchers make an inference about how likely an officer thought that the least suspicious (or “marginal”) person searched in each group had violated the law. However, only rarely are researchers able to distinguish between searches of marginal people and searches of “inframarginal” people – those whose behavior would always lead to a search, by any officer.

To see how this is a problem for interpreting hit rate tests, suppose that there is a set of “very suspicious” people who would always be searched regardless of their identity, and who almost always have violated a law. Further, suppose there are relatively more Hispanic people in this “very suspicious” group than White people. This will increase the average hit rate for Hispanic people overall, because more Hispanic people are carrying contraband. However, it also could be the case that officers are biased against Hispanic people, and the least suspicious, “marginal,” Hispanic person searched would not be searched if they were White. These marginal people are searched simply because they are Hispanic, which is exactly the sort of bias hit rate tests are intended to detect. However, if there happen to be more “very suspicious” Hispanic people than “marginal” Hispanic people, then a hit rate test, which combines searches of suspicious and marginal people, may not reveal the biased policing of “marginal” Hispanic people.

For this reason, the National Academy of Sciences concluded that hit rate tests can possibly “rule in” potentially problematic police officer behavior, but not necessarily rule out the existence of bias (National Academy of Sciences, 2018). In other words, hit rate tests should be interpreted as uninformative when they show no racial disparities, rather than a confirmation that there is no bias. Hit rate tests can be a useful tool to point out potentially problematic behavior, particularly when they suggest disparities against historically disadvantaged groups.

In this study, we conduct a modified version of a basic hit rate test by looking at how the severity of a stop outcome changes as enforcement intensity – specifically, the relative benchmarked RIPA stop rate - increases. We measure the severity of a stop outcome from least severe to most severe as follows: 1) verbal warning, 2) written warning, 3) citation, and 4) arrest.
Benchmarking the CHP

The central contribution of this report is to generate new benchmarking estimates of racial disparities in stops made by the California Highway Patrol (CHP). The CHP collects information on a broader set of officer-civilian interactions than is required by RIPA, specifically non-enforcement stops that CHP officers are expected to make when they observe a motorist in distress. To the best of our knowledge, within California this collection of non-enforcement data is unique to the CHP.

CPL was provided with data on 3,179,857 stops made by the CHP from January to December of 2019. Of those, 2,141,817 were initiated for enforcement reasons (RIPA stops) and 908,391 were initiated for non-enforcement reasons. CHP officers initiated non-enforcement encounters for the following reasons: to aid a disabled motorist (275,782), for a crash investigation (110,339) or crash report (222,135), or for motorist services (249,550). A total of 28,974 stops were made for other miscellaneous reasons.

159,133 stops (5% of all stops) were excluded because they were stops of commercial vehicles which took place at mandatory checkpoints. 129,649 stops (4% of all stops) were excluded because of missing or conflicting information on the primary activity (enforcement or non-enforcement) that occurred during a stop. Our final sample consists of 2,047,970 RIPA and 886,780 non-enforcement stops.

The composition of motorists involved in non-enforcement stops can be used as an informative benchmark for RIPA stops if the composition of people CHP encounters in non-enforcement stops is similar to the composition of people who the CHP would also consider stopping for enforcement, RIPA reportable, reasons. In Figure 1, we plot when non-enforcement, RIPA stops, and arrests are made. Non-enforcement and RIPA stops tend to occur at similar times of day, with large peaks during morning and evening rush hours, and a small local peak around midnight. This is consistent with the number of cars on the road being a central determinant of both types of stops. At the same time, the number of non-enforcement stops made by CHP officers is relatively constant over the course of the year (about 74,000 a month), while RIPA stops and arrests (made by any law enforcement agency in California) have a stronger cyclical pattern.

Based on Figure 1, we conclude that non-enforcement stops likely track traffic volume that varies by time of day, but may not track variation in the fraction of drivers engaged in risky driving or criminal behavior, which appears to be higher

---

5 Arrest information is from the FBI’s Uniform Crime Reports (UCR) data. All participating law enforcement agencies in the state of California are used to calculate the average number of arrests per law enforcement agency in each month of 2019.
in the spring and summer. While we cannot directly assess whether non-enforcement stops are a truly random sample of the driving population, because the stops seem to track the number of motorists patrolled by the CHP, we conclude that these stops create a compelling benchmark against which to compare the identity of people initially stopped by the CHP for enforcement reasons.

At the same time, it is important to acknowledge limitations of using non-enforcement stops as a benchmark. Thirty-one percent of the non-enforcement stops in our final sample are made in response to a vehicle breaking down. If breakdowns are more common for older, higher mileage vehicles that have undergone less routine maintenance, we might expect lower-income people to be more likely to be stopped for these reasons. To the extent that income is higher for White people than non-White people, this benchmark might overstate the number of non-White people “at risk” of being stopped for enforcement reasons. This would tend to understate bias in enforcement stop decisions. Without access to data on a true random sample of who is driving on California’s highways, the magnitude of this measurement error is unknown.

**How do Non-Enforcement Stops Compare to Standard Benchmarks?**

Across California, there are 2.04 RIPA stops of Asian drivers per non-enforcement stop, 2.58 RIPA stops of Black drivers per non-enforcement stop, 2.33 RIPA stops of Hispanic drivers per non-enforcement stop, and 2.23 RIPA stops of White drivers per non-enforcement stop (Figure 2A). If statewide residential population is used as the benchmark, then for every 100 people in the residential population, there are 2.5 RIPA stops of Asian drivers, 10 stops of
Black drivers, 5 stops of Hispanic drivers, and 6 stops of White drivers. This section focuses on whether benchmarking reveals racial or ethnic disparities in stop rates. As discussed previously, benchmarking exercises are not good at identifying the source of observed disparities. We discuss the role of other factors, such as risky driving behavior, in a later section on what determines stop rates.

To highlight the magnitude of disparities, in Figure 2B we plot the “relative stop rate” for a Division or racial group, which is that group’s benchmarked stop rate divided by the benchmarked stop rate for White people. This relative measure reflects how over, or under, represented a particular group is in the benchmarked RIPA stop rate relative to White people. If the relative stop rate for Black and White people is equal to 1, this means the number of RIPA stops of Black and White people matches the rate at which CHP encounters people in each group for non-enforcement reasons (like providing roadside assistance). A relative benchmarked stop rate of 3 means that the rate at which Black people are stopped is three times the rate of White people, or 200% more than what we would expect to observe if stops were not related to race, given the composition of the benchmarked population.

As a general rule, using residential population as benchmark reflects the actual probability that a resident in each identity group will have contact with the police. However, in the presence of broader societal disparities this benchmark will overestimate the amount of disparity that is due to police actions. In the specific case of the CHP, population benchmarks are potentially even less informative. CHP has primarily jurisdiction on California highways, and at any given point in time, the local population on a stretch of California highway is composed of local residents, residents from other parts of California, and out-of-state residents. This means that not only does a population benchmark not isolate the role of the CHP, the residential population may not actually reflect the population patrolled by CHP officers. Using the residential population of each Division, estimated using the 2015-2019 American Community Survey (ACS) as the benchmark, suggests that Black motorists are stopped 74% more often than White motorists. Both Asian and Hispanic drivers are stopped less frequently than White motorists, with the benchmarked stop rate for Hispanic motorists being very close to that of White drivers (9% lower).
Panel A of Figure 2 displays the stop rates by race, using the three primary benchmarks in this report. Within each race, the darkest shade of blue shows the relative stop rate using population as the benchmark, the middle shade of blue uses the commuting population as the benchmark, and the lightest shade of blue uses non-enforcement contacts as the benchmark. In panel B of Figure 2, relative stop rates are presented so that disparities are comparable across benchmarks. The adjustment scales the stop rate of a particular racial or ethnic group relative to the stop rate for White drivers, such that the relative stop rates equal 1 (line of equality) when stop rates between a group and White drivers are equal (i.e. when there are no disparities).
The second benchmark we use is the number of residents in each place who commute to work in a private vehicle - either driving alone or in a carpool. We estimate this using the 2015-2019 ACS. Using commuters as the benchmark increases the disparity for Black drivers, who are stopped 210% more frequently than White drivers. The commuter benchmark also shows disparities for Hispanic drivers, who are stopped 52% more often than White motorists.

The observed increase in disparities in the commuting benchmark highlights the previously discussed limitations of benchmarking. Based on the 2015-2019 ACS estimates we utilize in this report, the average household income in our sample of commuters is $132,404 and 60% identify themselves as White, non-Hispanic. In contrast, in the residential population, the average household income is $115,954 and 36% of people identify themselves as White, non-Hispanic. To the extent that people with higher income may be less likely to engage in criminal or risky behavior, (Duncan and Brooks-Gunn 1997), and that in our benchmarks – on average - commuters are wealthier than the residential population, it is plausible that the commuting population may be a particularly biased estimate of the population of people engaged in behavior that puts them at risk of being stopped than the CHP.

To be more specific, to the extent that the commuting population is more likely to be wealthier and White than the population as a whole, this would increase the number of White people in the benchmark, meaning that a “no disparities” estimate would have even more stops of White people than suggested by the population benchmark. Conceptually, the commuting population benchmark may incorrectly attribute the disparities in education and labor market outcomes, which create disparities in car ownership, to the decisions made by CHP officers. The observed differences between the residential and privately commuting population suggests that the large commuter-benchmarked disparities we find may be caused by social inequities in systems that are outside of CHP control.

Finally, we present the non-enforcement benchmarked RIPA stop disparities. Overall, the non-enforcement benchmark results in more equal stop rates for all three groups relative to the White stop rates. Across all three non-White groups, the non-enforcement-based benchmark suggests the most equal treatment of drivers by CHP officers relative to the other benchmarks. However, it is important to note that this benchmark does not completely eliminate disparities - Black motorists are still 15% more likely to be stopped than White drivers. For Hispanic and Asian drivers, the number of RIPA stops per non-enforcement stops is even closer to the rate for White drivers (5% more likely and 9% less likely, respectively). In addition, we are unable to assess the extent to which drivers who are Black, Asian or Hispanic are more likely than White drivers to be driving older, higher-mileage cars that break down...
more frequently or create other reasons for non-enforcement assistance by CHP.

**Benchmarked Stop Rates Across Divisions**

At the state level, the ratio of RIPA to non-enforcement stops is relatively similar across racial groups. However, state level measures may not reflect the practices of CHP officers assigned to different jurisdictions, and potentially working in very different contexts. The CHP is organized into eight distinct geographic Divisions (see Appendix A Figure A1), each under the supervision of a separate Division Chief. To account for local variation in officer and driver behavior, we replicated our state level analysis at the Division-level.

When averaged across Divisions, CHP officers make 2.59 RIPA stops of Asian drivers per non-enforcement stop, 2.88 RIPA stops of Black drivers per non-enforcement stop, 2.60 RIPA stops of Hispanic drivers per non-enforcement stop, and 2.28 RIPA stops of White drivers per non-enforcement stop. Note that the ratio of enforcement to non-enforcement stops varies by CHP Division. There are many reasons for this variation across Divisions. An important one is the jurisdictional authority of the CHP: we estimate that a fully unincorporated jurisdiction, where the CHP will be the primary law enforcement agency, will conduct 1.34 additional enforcement stops per non-enforcement stop than jurisdictions that are fully incorporated (meaning local law enforcement will respond to most criminal activity).

In Figure 3, we present the relative benchmarked stop rate which directly compares stops of drivers in non-White groups to stops of White drivers, using non-enforcement stops as a benchmark, for each of the eight CHP Divisions. Black motorists are stopped more than White motorists in every Division, but the highest disparities exist in Division 1 (the Black stop rate is 41% higher than the White stop rate), Division 3 (Black drivers are stopped 34% more often) and Division 4 (Black drivers are stopped 37% more often). Division 8 has the least disparate stop rates overall, with 5% more Black drivers, 13% fewer Hispanic drivers, and 3% more Asian drivers stopped than White drivers.

---

6 Note that the CHP headquarters are a separate Division (Division 0) that is physically located within Division 2, but preside over the entire state. Stops made by Division 0 are excluded from calculations for other Divisions.

7 When calculated at the state level, the number of RIPA stops per non enforcement stops in all 8 CHP Divisions is equal to \( \frac{RIPA_1 + RIPA_2 + RIPA_3 + RIPA_4 + RIPA_5 + RIPA_6 + RIPA_7 + RIPA_8}{NonEnf_1 + NonEnf_2 + NonEnf_3 + NonEnf_4 + NonEnf_5 + NonEnf_6 + NonEnf_7 + NonEnf_8} \). When averaged across the 8 CHP Divisions, the average number of RIPA stops per non enforcement stop is \( \frac{1}{8} \sum_{i=1}^{8} \frac{RIPA_i}{NonEnf_i} \).

8 From a linear probability model estimating the effect of the share of a division that is unincorporated on the share of total stops made for enforcement reasons.

9 Relative to White drivers, RIPA stop rates of Asian drivers are lower, but qualitatively identical to Division 8 in Divisions 2, 3 and 6.
Each Division of Figure 3 depicts RIPA stops per non-enforcement stops of the same race, relative to White RIPA stops per White non-enforcement stops: e.g. (Black stop rate) / (White stop rate). Divisions are colored white when stop rates are equal. Darker colors represent more disparate stop rates. Yellow tones mean the stop rate of the racial/ethnic group is lower than the White stop rate and blue tones mean the stop rate of the group is higher than the White stop rate.

For sake of comparison, we present similar Division-level maps using residential population and commuters as a benchmark in Figure 4. In general, using Census-based numbers as a benchmark increases the differences in RIPA stop rates across identity groups, with population-benchmarked stop rates of Black people increasing relative to White stop rates, and the population-benchmarked stop rates for Hispanic and Asian people decreasing or increasing relative to White stop rates, depending on the Division.
FIGURE 4. Relative Stop Rates using Traditional Benchmarks

A. Relative Enforcement Per Population Stop Rate

B. Relative Enforcement Per Commuter Stop Rate

Each Division of Figure 4 depicts RIPA stops per population or commuter of the same race, relative to White RIPA stops per White population or commuter. Example: (Black stop rate) / (White stop rate). Figures are white when stop rates are equal. Darker colors represent more disparate stop rates. Yellow tones mean the stop rate of the racial/ethnic group is lower than the White stop rate and blue tones mean the stop rate of the group is higher than the White stop rate.

The spatial patterns of relative RIPA stops calculated using different benchmarks suggests that the gap between the population at risk of being stopped by the CHP and Census-based benchmarks not only changes the size of RIPA stop disparities, but also changes where in California the larger disparities are most likely to exist—places where further investigation into stop actions may be most beneficial for the CHP.
What Determines Stop Rates?

The last step in our benchmarking analysis is to statistically explain some of the differences in benchmarked stops across Divisions and racial groups. Many contextual and unmeasured factors affect a CHP officer’s decision to make a stop. In this section, we focus on estimating the role of two factors that can contribute to group-level disparities. We first look at how disparities vary in low population jurisdictions, as population may be correlated with the scope and nature of CHP officer’s typical daily tasks. We also estimate the role of group and region-specific differences in unsafe driving practices that maybe correlated with a driver’s racial and ethnic identity. Both of these factors may contribute to the overall disparity in CHP RIPA contacts.

First, we explore the potential role of CHP jurisdictional authority – specifically differences across California in the places where CHP is responsible for enforcing California law. In heavily populated areas, like Los Angeles County, CHP officers may be exclusively responsible for patrolling heavily trafficked restricted access highways. In less populated areas, a larger share of CHP patrol areas include surface (non-highway) roads. Across Divisions, RIPA stops make up a larger share of total stops made by CHP officers in jurisdictions with lower populations (Appendix Figures A4, A5).

Each Division in the CHP is divided into operational units known as “area offices.” In Figure 5, we plot, for each area office and Asian, Black, and Hispanic group, the relative number of RIPA to non-enforcement stops against the total residential population of that area office's jurisdiction. There are three observations per area office, one per non-White group, and each benchmarked stop rate is compared to the White benchmarked stop rate. Stop rates greater than 1 suggest a higher stop rate for the non-White group than the White group.

FIGURE 5. Relationship between Stop Rates and Population (continued on next page).
Figure 5 displays the relationship between relative stop rates by race (group specific RIPA stops per non-enforcement stop divided by White stop rate) and the total population, by Division. There are three dots (Hispanic, Asian, Black) per area office within each Division. The yellow line at $y = 1$ is where there is equality between the stop rate of the group and White stop rates for each area office.

To provide further context, we will use Division 1 as an example. If population was not related to disparate stop rates, we would expect the scatter plot to have a flat trend. In Division 1, however, disparities decrease as residential population increases. The largest disparities exist in area offices where the residential population is under 50,000. In some of these area offices, Asian motorists are stopped at more than two times the rate of White motorists.
With the exception of Divisions 2, 3 and 5, the stop rate disparities are largest in low population jurisdictions.

This particular analysis does not identify why lower population jurisdictions have more disparate stop rates, but simply that population is a predictor of larger disparities. Further qualitative investigation into the regular tasks, experience, and training of CHP officers working in low vs high population areas may help CHP better understand why this relationship exists.

We next use multivariate regression to predict the number of RIPA stops per non-enforcement stop that would be made of White, Black, Hispanic, and Asian drivers if all motorists engaged in similar driving behavior, and the racial composition of drivers was the same across California – meaning there was no difference in the racial composition of drivers across Divisions. Mathematically, this involves calculating the number of RIPA stops per non-enforcement stops for each group in each Division, and modeling that as a function of the identity group, residential population size, and a measure of the amount of "risky" driving by group members that would justify a RIPA stop.

We quantify the behavior of drivers in each identity group and CHP Division in two ways. We first calculate the fraction of drivers involved in fatal accidents for each racial group using Fatality Analysis Reporting System (FARS) data. We also calculate the fraction of at-fault drivers involved in collisions for each racial group using data from the CA Statewide Integrated Traffic Records System (SWITRS). Because we are using these accident reports to approximate differences in the underlying driving risk of the driving population as a whole – e.g. How “risky” are White drivers relative to Asian drivers in this area? - we use all accidents reported in both systems, even if CHP officers did not respond to the incident.

Figure 6 presents both baseline and risk-adjusted RIPA estimated stop rates, which adjust for any average difference in the relative number of RIPA and non-enforcement stops across Divisions that is the same for all identity groups (e.g. poor road conditions in one Division leading to more disabled vehicles, and if there were more people in a particular group who lived in that Division). Recall that, on average across Divisions, White drivers are stopped for enforcement reasons 2.3 times per non-enforcement stop. Black drivers are stopped 2.9 times per non-enforcement stop, and Asian and Hispanic drivers are stopped 2.6 times per non-enforcement stop.

If there are differences in driver behavior across racial groups, this could affect the frequency of RIPA stops across groups that are not caused by bias on the part of CHP, and the existence of such behavioral differences is unknown. Statistically accounting for one imperfect proxy for differences in driver behavior reduces the difference in expected stop rates across all groups, most noticeably across White, Black, and Asian drivers. If all drivers in each group
were responsible for the same fraction of FARS accidents, and the racial and ethnic composition of drivers was the same across California, we would observe White drivers stopped 2.5 times per non-enforcement stop, Asian drivers stopped 2.4 times, Black drivers stopped 2.5 times, and Hispanic drivers stopped for enforcement reasons 2.8 times per non-enforcement stop. Using data from the SWITRS, which includes less serious traffic accidents, but may also be affected by differences in the probability that an accident is reported to law enforcement, generates essentially the same results.

There are many possible reasons that RIPA stop rates could be different in high and low population areas, some of which are more or less potentially problematic (e.g. better maintained roads vs. increased CHP bias towards specific groups in more rural places). Whether or not adjusting for population increases the accuracy or decreases the accuracy of our estimate of racial disparities in CHP officer decisions depends on the specific reason why population affects estimates of disparities, which likely requires additional review by the CHP.

**FIGURE 6. Predicted RIPA Stops Per Non-Enforcement Stops**

Figure 6 displays the predicted average stop rates by race, using non-enforcement stops as the benchmark. The first set of columns displays the actual, unadjusted stop rates, while the subsequent columns display the predicted stop rate we would expect to observe, based on our multivariate regression analysis, if there was no geographic variation in the racial composition of drivers and no variation in the racial share of drivers involved in fatal accidents (FARS) or at-fault collisions (SWITRS).

In contrast, if we adjust for SWITRS group-specific accident risk, but do not assume that residential population should affect the benchmarked stop rate (i.e. low population jurisdictions may have different stop rates for potentially problematic reasons), we predict that White and Asian drivers are stopped 2.4 times per non-enforcement stop, and that Black and Hispanic drivers are stopped 2.7 times per non-enforcement stop, which is 13% higher than the White and Asian stop rate. This confirms our finding in Figure 5, that small
population Divisions appear to be important contributors to disparities in stops. For Black-White disparities, understanding exactly why CHP officers in small-population jurisdictions stop Black and White drivers at different rates, potentially through a further audit of a random sample of stops in those places, may help CHP better understand the cause of these disparities.

In Figure 7, we present the results of the same exercise in predicting RIPA stop rates but using residential population as a benchmark. There are two main conclusions to draw from comparing Figures 6 and 7. First, consistent with Figures 3 and 4, disparities are much larger when population is used as a benchmark. Second, adjusting stop rates for driver accident risk reduces disparities more when non-enforcement stops are used as a benchmark than when population is used as a benchmark.

This second observation is consistent with our preliminary conclusions regarding Figures 3 and 4. The composition of the residential population is likely to be a poor estimate of the composition of drivers at risk of being stopped by a CHP officer. If population-benchmarked RIPA stops are a less accurate measure of the social phenomenon of interest (CHP officer stop decisions), attempts to model that particular measure will not be as successful as attempts to model a more accurate measure of stop decisions, like RIPA stops benchmarked with non-enforcement stops.

![Figure 7: Predicted RIPA Stops Per 100 Population](image)

**Figure 7** displays the actual and expected average stop rates by race, using population as the benchmark. The first set of columns displays the actual, unadjusted stop rates, while the subsequent columns display the predicted stop rate we would expect to observe, based on our multivariate regression analysis, if there was no geographic variation in the racial composition of drivers and no variation in the racial share of drivers involved in fatal accidents (FARS) or at-fault collisions (SWITRS).
Importance of Non-Enforcement Contacts

One important interpretation of the reduction in disparities that we observe moving from Census-based benchmarks to non-enforcement stop-based benchmarks is that people who are more likely to encounter CHP officers in a potentially adversarial enforcement context are also more likely to encounter CHP officers in a non-enforcement, non-adversarial, context.

While non-enforcement stops may be routine from the perspective of a CHP officer, for the civilian, these are uncommon, highly stressful, and memorable situations. Non-enforcement stops may be one of the few times this individual person ever encounters law enforcement. Borrowing the terminology of police scholars, these non-enforcement stops can be thought of as opportunities for officers to demonstrate the “guardian” aspect of a CHP officer’s role.

Consistent with the recommendation of the Presidential Task Force of 21st Century Policing, CHP officers can use these non-enforcement stops to develop positive connections between civilians and officers, increasing trust in the CHP and public perceptions of integrity and legitimacy of law enforcement.

How are RIPA stops resolved by the California Highway Patrol?

A Hit Rate Test

An officer’s decision to make a stop is different from an officer’s decision about how to resolve that stop. Both decisions contribute to the overall experience that a civilian has with a CHP officer, which may or may not result in a feeling that they were treated fairly.

In the policing context, hit rate tests are commonly used to test for bias in search decisions by comparing the probabilities that searches of different groups of people result in the discovery of contraband. Since CHP officers search civilians at very low rates (approximately 2% of RIPA stops in our data), we propose that searches and contraband discovery occur too infrequently to be a particularly informative descriptor of typical CHP encounters. Instead, we analyze the severity of outcomes after RIPA stops. In this modified hit rate test, we examine the severity of a stop outcome, measured from least severe to most severe as: 1) verbal warning, 2) written warning, 3) citation, and 4) arrest. In this context, the decision to stop a civilian,

10 For sake of comparison, in the 2nd half of 2018, CHP officers searched the person or property of 1.1% of all people stopped, and all other reporting agencies searched the person or property of 21.8% of all people stopped.
or enforcement intensity, plays the role of a search, while the severity of the outcome plays the role of the contraband discovery.

If racial bias plays a role in an officer’s stop decision, specifically that non-White drivers are stopped for behavior that a White driver would not be stopped for, then we might expect to see that, on average, Black drivers stopped by the CHP would end up with less serious consequences after stops. This phenomenon is displayed in the top row of Figure 8.

**FIGURE 8. Bias in Enforcement Severity “Hit Rate”**

![Flow chart of hit rate calculations](chart)

Bias makes this ratio **smaller**.

1) If stop rate and outcome severity \(\uparrow\) at same rate, no change.
2) If stop rate \(\uparrow\) more than outcome severity, smaller.
3) If outcome severity \(\uparrow\) more than stop rate, larger.

Figure 8 displays a flow chart of how hit rate calculations can be affected by different assumptions on how group-based disparities alter officer behavior. In the top row, the outcome severity is held constant, and disparities only alter the stop rate. In the bottom row, disparities alter both the decision to stop as well as the decision to assign more severe outcomes (e.g., a verbal vs a written warning).

Hit rate tests presume that differences in the outcome of a stop (or search) across groups only reflects differences in the actual criminality of the civilian, rather than the outcome also being influenced by racial bias. If CHP officers would stop an Asian driver for actions that a Black driver would not be stopped for, and also would issue a ticket to an Asian driver for behavior that they would verbally warn a Black driver for, the disparate treatment in the outcome would mask disparate treatment in the stop rate. This phenomenon is displayed in the bottom row of Figure 8. To the extent that CHP officers have more discretion in choosing between a verbal warning, written warning, or citation than recording the discovery of contraband, correlated bias in both stop and outcome decisions may be more of a problem in this context than in standard hit rate tests.

With that important caveat in mind, as shown in Figure 9, the distribution of outcomes for White and Black people stopped by the CHP are quite similar. Citations and verbal warnings are the most common stop outcomes for all groups of people. That said, Black people are 1.3 percentage points more likely to be arrested, and 1.4 percentage points less likely to receive a written
warning than White drivers who are stopped. Asian and Hispanic drives are more likely to be cited, and less likely to receive a verbal warning, conditional on being stopped. Substantively, the differences in outcomes across racial groups are small. Since hit rate analysis can be an indicator of potential bias but is not a reliable indicator of the absence of bias, the hit rate analysis results do not support any conclusion about whether, across California, enforcement stops by CHP are racially biased.

FIGURE 9. Hit Rate Analysis: RIPA Stop Outcomes by Group

Figure 9 displays the percent of total RIPA stops for a particular race or ethnicity that result in a verbal warning, written warning, citation or arrest. Note that percentages will not add up to exactly 100% since a small fraction of enforcement stops results in alternative outcomes.

**Are Stop Outcomes Related to Disparate Stop Rates?**

We now use non-enforcement stop data to contextualize the hit rate analysis of stops made by CHP officers. Specifically, we estimate whether hit rate tests tend to suggest more disparate stops in the same CHP Divisions where our benchmarking analysis suggested there may be racially disparate stops occurring.

We examine the impact of disparities in benchmarked RIPA stop rates on the severity of enforcement outcomes using an ordered logit model. The unit of observation is an individual stop, and the relative benchmarked RIPA stop disparity is calculated for Asian, Black, and Hispanic drivers in each Division.

---

11 An ordered logit model is used to study outcomes where the order of the outcome matters. In this case, an arrest is more severe than a citation, and both an arrest and citation are more severe than a written warning. An ordered logit preserves that information.
To the extent that we observe that there is a negative (downward sloping) relationship between the relative RIPA stop rate in an area and the severity of stop outcomes, this means that in Divisions with larger disparities in RIPA stop rates for non-White drivers relative to White drivers, we also observe an increased likelihood that the stops result in less consequential outcomes (like a verbal warning) for these non-White drivers. This would support a conclusion that the driver’s race may have played a role in a CHP officer’s decision to make a stop. In contrast, finding that the benchmarked stop rate was not related to how CHP officers resolved those stops does not allow for any strong conclusions, as this result could occur in the absence of any CHP bias (particularly if there were different numbers of inframarginal members of different groups), or if CHP bias existed in both the decision to stop and the resolution of the stop.

In Figure 10, we present our results graphically, identifying the predicted change in the probability that a stop results in any particular outcome as the disparity in RIPA stops increases. Overall, we find a small negative relationship between the enforcement severity and outcome severity. This relationship disappears, however, after controlling for other factors such as driving behavior using the SWITRS data, allowing for a constant average difference in outcomes across race, or allowing for a constant average difference in outcomes across Divisions. After taking these other factors into account, we conclude that knowing the relative benchmarked RIPA stop rate for a specific racial group in a specific Division is not informative about the average stop outcomes for that group. The relationship between the relative stop rate and the probability that a stop ends in any particular outcome is close to zero, without strong evidence of an increase or decrease as the outcome becomes more consequential for the civilian.

When we divide our sample by driver identity, focusing on each non-White group in turn, we do observe distinct patterns. Figures 11-13 break out these results by different racial and ethnic groups.

The probability that a Hispanic driver stopped by the CHP is ultimately warned, cited, or arrested seems to be unrelated to how frequently drivers that look like them are stopped, with or without adjusting for differences in driver risk across Divisions (Figure 11). In other words, officers in Divisions that have larger Hispanic-White disparities in benchmarked RIPA stops are not more likely to let stopped Hispanic drivers go with only a verbal warning, which is the type of bias hit rates tests attempt to find.

We observe different patterns for Black and Asian drivers, both of which suggest that, in areas where these drivers are stopped for enforcement reasons more than their non-enforcement stops would suggest, CHP officers could make fewer RIPA stops while continuing to identify the same amount of illegal behavior (Figures 12 and 13). This is also true when taking the frequency with
which drivers are at fault in accidents in that region, as well as the ability of the driver to effectively communicate with the CHP officer during the stop. Indeed, the difference in outcomes for Asian drivers is most apparent once driver characteristics are accounted for. In contrast to stops of Hispanic drivers, in Divisions with larger disparities in Black-White and Asian-White RIPA stop rates, we also observe an increased likelihood that these stops result in less consequential outcomes for the driver. This is consistent with race playing a role in some CHP stop decisions. Conclusive determination of CHP bias would require additional investigation of specific stop circumstances. These results suggest that such investigations, should CHP wish to undertake them, may be most beneficial in Divisions with the highest disparities in RIPA stops relative to non-enforcement stops.

Figure 10 displays the average marginal effect of enforcement intensity for non-White drivers with 95% confidence intervals. The baseline model has no controls, while the subsequent models control for the racial share of at fault drivers in collisions (SWITRS share), race fixed effects, or Division fixed effects.

---

The ability of the stopped driver to communicate is measured based on the RIPA field on whether the individual is perceived to have limited or no English fluency. Controlling for this measure did not substantively alter the predicted margins relative to the baseline model.
Figure 11 displays the average marginal effect of enforcement intensity for Hispanic people with 95% confidence intervals. The baseline model on the left has no controls, while the model on the right controls for the racial share of at fault drivers in collisions (SWITRS share).

Figure 12 displays the average marginal effect of enforcement intensity for Black people with 95% confidence intervals. The baseline model on the left has no controls, while the model on the right controls for the racial share of at fault drivers in collisions (SWITRS share).
**CONCLUSION**

This report analyzes 2,047,970 enforcement stops conducted by California Highway Patrol officers in 2019, reported to the state under RIPA. For every 100 people in the residential population, there were 3 RIPA stops of Asian drivers, 10 stops of Black drivers, 5 stops of Hispanic drivers, and 6 stops of White drivers. However, comparing the stopped population to the general population, as opposed to the population engaged in behavior that may warrant a stop, will likely overstate the amount of disparity in police contact that is due to police officer decisions.

The collection of non-enforcement stop data by CHP officers provides an alternative estimate of who is driving on California’s Highways. We find that, for each non-enforcement stop made, CHP officers made 2.04 RIPA stops of Asian drivers, 2.58 stops of Black drivers, 2.33 stops of Hispanic drivers, and 2.23 stops of White drivers. Black drivers are 15% more likely to be stopped than White drivers, while Asian and Hispanic drivers are stopped at similar rates as White drivers. However, if on average, non-White people drive vehicles that break down more frequently than vehicles driven by White people (and experienced more non-enforcement stops as a result), then using non-enforcement stops as a benchmark could understate the amount of disparity in police contact that is due to officer decisions.

These alternate benchmarks suggest that CHP officer actions may be responsible for 20% of the increased likelihood that Black Californians are...
stopped by the CHP relative to other groups. Almost 80% of the population-based disparity in Black RIPAs stop rates is potentially due to there being more Black drivers on CHP-patrolled roads than local residential population estimates would suggest.

We also found regional variation in benchmarked stop rates; in particular, officers working in lower population jurisdictions tend to have high ratios of enforcement to non-enforcement stops, and the racial disparities in these stops tends to be larger.

Non-enforcement contacts are important opportunities for positive community outreach on the part of CHP. The non-enforcement nature of these encounters provides both officers and civilians an opportunity to interact in a problem-solving, collaborative way, building the community capital that creates trusting relationships and officer legitimacy. Building community capital during non-enforcement stops could be particularly important because we also find evidence that in places where the benchmarked stop rates of Black and Asian drivers are higher, many of these stops did not result in CHP officers identifying illegal or problematic driver behavior. Conditional on the fraction of accidents where a Black or Asian driver is at fault, officers working in Divisions where the benchmarked stop rate of Black and Asian people is higher relative to the White benchmarked rate are more likely to resolve those stops informally, rather than discovering behavior that leads to a citation or arrest. Additional investigation of stops that are resolved informally could potentially help to reduce stop disparities without sacrificing public safety.

---

13 Recall that the 74% Black-White disparity under the population benchmark is reduced to a 15% Black-White disparity under the non-enforcement benchmark – a 79% reduction.
Works Cited


Appendix A: Auxiliary Tables & Figures

FIGURE A1. CHP Divisions

Figure A1 displays the geographic location of each CHP Division.

FIGURE A2. RIPA Stops Per 100 Population

Each Division of Figure A2 depicts the number of stops per 100 population of the same race, where darker shades of blue represent higher stop rates per capita.
FIGURE A3. RIPA Stops Per 100 Commuters

Each Division of Figure A3 depicts the number of stops per 100 commuters of the same race, where darker shades of blue represent higher stop rates per commuter.

FIGURE A4. RIPA Stops Per Non-Enforcement Stop

Each Division of Figure A4 depicts the RIPA stops per non-enforcement stops of the same race, where darker shades of blue represent higher stop rates.
Figure A5 depicts racial and ethnic group residential population estimates of each CHP Division based on data from the 2015-2019 ACS.

Each Division of Figure A6 depicts searches per 100 RIPA stops, where darker shades of blue represent higher search rates.
Appendix B: Division by Area Office

**Figure B1.1** Division 1: Relative Non-Enforcement Stop Rate and Population, Black

<table>
<thead>
<tr>
<th>Relative Benchmark</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x &gt; 1.7$</td>
<td>1M to 2.5M</td>
</tr>
<tr>
<td>$1.5 &lt; x \leq 1.7$</td>
<td>500K to 1M</td>
</tr>
<tr>
<td>$1.3 &lt; x \leq 1.5$</td>
<td>250K to 500K</td>
</tr>
<tr>
<td>$1.1 &lt; x \leq 1.3$</td>
<td>100K to 250K</td>
</tr>
<tr>
<td>$0.9 &lt; x \leq 1.1$</td>
<td>50K to 100K</td>
</tr>
<tr>
<td>$0.7 &lt; x \leq 0.9$</td>
<td>25K to 50K</td>
</tr>
<tr>
<td>$0.5 &lt; x \leq 0.7$</td>
<td>5K to 25K</td>
</tr>
<tr>
<td>$0.3 &lt; x \leq 0.5$</td>
<td>0 to 5K</td>
</tr>
<tr>
<td>$x \leq 0.3$</td>
<td></td>
</tr>
</tbody>
</table>

Each area office is labeled with the area office number. The left column of Figure B1.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B1.1 depicts population estimates of each CHP area office within Division 1 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B1.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B1.2 depicts population estimates of each CHP area office within Division 1 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B1.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B1.3 depicts population estimates of each CHP area office within Division 1 based on data from the 2015-2019 ACS.
FIGURE B2.1 Division 2: Relative Non-Enforcement Stop Rate and Population, Black

Each area office is labeled with the area office number. The left column of Figure B2.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B2.1 depicts population estimates of each CHP area office within Division 2 based on data from the 2015-2019 ACS.
FIGURE B2.2 Division 2: Relative Non-Enforcement Stop Rate and Population, Hispanic

Each area office is labeled with the area office number. The left column of Figure B2.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B2.2 depicts population estimates of each CHP area office within Division 2 based on data from the 2015-2019 ACS.
FIGURE B2.3 Division 2: Relative Non-Enforcement Stop Rate and Population, Asian

Each area office is labeled with the area office number. The left column of Figure B2.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B2.3 depicts population estimates of each CHP area office within Division 2 based on data from the 2015-2019 ACS.
FIGURE B3.1 Division 3: Relative Non-Enforcement Stop Rate and Population, Black

Each area office is labeled with the area office number. The left column of Figure B3.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B3.1 depicts population estimates of each CHP area office within Division 3 based on data from the 2015-2019 ACS.
FIGURE B3.2 Division 3: Relative Non-Enforcement Stop Rate and Population, Hispanic

Each area office is labeled with the area office number. The left column of Figure B3.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B3.2 depicts population estimates of each CHP area office within Division 3 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B3.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B3.3 depicts population estimates of each CHP area office within Division 3 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B4.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B4.1 depicts population estimates of each CHP area office within Division 4 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B4.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B4.2 depicts population estimates of each CHP area office within Division 4 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B4.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B4.3 depicts population estimates of each CHP area office within Division 4 based on data from the 2015-2019 ACS.
FIGURE B5.1 Division 5: Relative Non-Enforcement Stop Rate and Population, Black

Each area office is labeled with the area office number. The left column of Figure B5.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B5.1 depicts population estimates of each CHP area office within Division 5 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B5.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B5.2 depicts population estimates of each CHP area office within Division 5 based on data from the 2015-2019 ACS.
FIGURE B5.3 Division 5: Relative Non-Enforcement Stop Rate and Population, Asian

Each area office is labeled with the area office number. The left column of Figure B5.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B5.3 depicts population estimates of each CHP area office within Division 5 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B6.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B6.1 depicts population estimates of each CHP area office within Division 6 based on data from the 2015-2019 ACS.
FIGURE B6.2 Division 6: Relative Non-Enforcement Stop Rate and Population, Hispanic

Each area office is labeled with the area office number. The left column of Figure B6.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B6.2 depicts population estimates of each CHP area office within Division 6 based on data from the 2015-2019 ACS.
FIGURE B6.3 Division 6: Relative Non-Enforcement Stop Rate and Population, Asian

Each area office is labeled with the area office number. The left column of Figure B6.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B6.3 depicts population estimates of each CHP area office within Division 6 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B7.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B7.1 depicts population estimates of each CHP area office within Division 7 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B7.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B7.2 depicts population estimates of each CHP area office within Division 7 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B7.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B7.3 depicts population estimates of each CHP area office within Division 7 based on data from the 2015-2019 ACS.
Each area office is labeled with the area office number. The left column of Figure B8.1 depicts RIPA stops per non-enforcement stops for Black drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Black stop rate is lower than the White stop rate and blue tones mean the Black stop rate is higher than White stop rate. The right column of Figure B8.1 depicts population estimates of each CHP area office within Division 8 based on data from the 2015-2019 ACS.
FIGURE B8.2 Division 8: Relative Non-Enforcement Stop Rate and Population, Hispanic

Relative Benchmark

Population

Each area office is labeled with the area office number. The left column of Figure B8.2 depicts RIPA stops per non-enforcement stops for Hispanic drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Hispanic stop rate is lower than the White stop rate and blue tones mean the Hispanic stop rate is higher than White stop rate. The right column of Figure B8.2 depicts population estimates of each CHP area office within Division 8 based on data from the 2015-2019 ACS.
FIGURE B8.3 Division 8: Relative Non-Enforcement Stop Rate and Population, Asian

Each area office is labeled with the area office number. The left column of Figure B8.3 depicts RIPA stops per non-enforcement stops for Asian drivers, relative to White stop rates. Lighter colors represent more equal stop rates. Yellow tones mean the Asian stop rate is lower than the White stop rate and blue tones mean the Asian stop rate is higher than White stop rate. The right column of Figure B8.3 depicts population estimates of each CHP area office within Division 8 based on data from the 2015-2019 ACS.
Appendix C: Veil of Darkness Tests

A Veil of Darkness (VOD) test identifies the presence of biased decision making by comparing decisions made in the presence, or absence, of identity information. A classic study of this type of bias was done by Cecilia Rouse and Claudia Goldin, who demonstrated that female musicians were more likely to be selected in “blind” orchestra auditions - where the musician was separated by the judges by an opaque curtain - than in “unblinded” auditions, conducted in full view of the judges (Goldin and Rouse 2000). The VOD test was formalized in the context of policing by Jeffrey Grogger and Greg Ridgeway. Their proposed test essentially compares the racial composition of stops made during the evening rush hour when visibility depends on the time of year (dark in the winter, light in the summer). If a larger fraction of the stopped population is Black during summer evenings than in the winter evenings, this suggests that race is being used in the stop decisions of officers (Grogger and Ridgeway 2006).

The 2021 State RIPA board report did not find evidence of disparities in their VOD analysis of CHP stops. We do not conduct a VOD test in this report, because relative to the other two tests, the VOD evaluates a very narrowly defined pathway through which civilian identity could impact officer behavior, and the number of CHP stops that could be affected by the type of bias measured by the VOD is most likely very small. Because of this, we feel that it can produce less general conclusions about typical encounters between CHP officers and civilians.

Why is the population examined by the VOD very small? VOD tests examine officer responses to motorists whose group identities are visible to officers during the inter-twilight hours of daylight savings time, but not during other times of the year. In some places, this might be a large group of people, in others not very many. We are unaware of credible estimates of how large this population is on California’s highways. Efforts to increase visibility through artificial lighting, and the high speed at which vehicles may be traveling on the highway, will tend to reduce the number of people whose racial or ethnic identities can sometimes be determined, depending on the position of the sun. Note that in the orchestral study by Goldin and Rouse, every musician was a part of this population, as all identities were concealed behind a curtain and none were concealed in unblinded auditions.

Further, the VOD test intends to evaluate officer reactions to a subset of people in this first group whose driving behavior is marginal - where racial identity would influence a biased officer’s stop decision, but not an unbiased officer’s stop decision. This is a further reduction in the number of drivers and officer decisions that are tested in the VOD model, making this type of bias even harder to statistically detect.