Racial & Identity Profiling Advisory Board
2022 Annual Report

A Critical Analysis
by Brian L. Withrow, Ph.D.
January 2, 2023

Executive Summary

Racial bias exists. It exists in healthcare, housing, and consumer finance. It exists in technology, higher education, and in the way we divvy up our state's natural resources. Racial bias is unfortunately a factor in every aspect of American society, and law enforcement is no different.

California is experiencing a public safety crisis. Law enforcement departments are understaffed, underfunded, and underappreciated. While criminals get slaps on the wrist, officers and officer behavior is being legislated, scrutinized, and even demonized. Peace officers hold a unique position of authority within our society, and increased scrutiny is warranted, but only insofar as that scrutiny is based on an accurate assessment of officer behavior.

The normalization of anti-police sentiments and rising crime on our streets have inevitably damaged relationships between police and the communities they serve. That is why it is so important that Californians receive a clear, unbiased, and transparent look into how officers interact with residents and communities of color.

Unfortunately, California's Racial and Identity Profiling Advisory Board (RIPA) has pursued an inherently flawed approach to assessing police stop data that both misrepresents the data itself and misleads the public to believe things that simply are not true. Californians deserve appropriate scrutiny of officer behavior, but they also deserve the truth. To do otherwise would only sow further division between law enforcement and the communities they risk their lives every day to serve – limiting recruitment potential, negatively impacting officer morale, and likely decreasing positive public safety outcomes.

To ensure the public is afforded the opportunity to place RIPA's work in the appropriate context, the Peace Officers Research Association of California (PORAC) engaged the assistance of Dr. Brian L. Withrow to assist in its review of the 2022 Annual RIPA Report. The report includes information on nearly 3,000,000 traffic and pedestrian stops occurring in California during calendar year 2020. Statewide, only 18 law enforcement agencies provided data for this report. According to the U.S. Bureau of

Justice Statistics' 2018 Census of State and Local Law Enforcement Agencies, California has a total of 531 law enforcement agencies.

Dr. Withrow is one of the nation's leading experts on racial profiling. He has authored three books and numerous articles and reports on the subject over the past 22 years. Dr. Withrow's research methods textbook (Research Methods in Crime and Justice, Second Edition (2016), Routledge Publishing) is used extensively throughout the United States and in other countries. In 2018, he coauthored a popular textbook on ethics (Police Ethics: The Corruption of Noble Cause, Fourth Edition (2018), Routledge Publishing) with Michael A. Caldero and Jeffrey C. Dailey.

While information on police stops does provide some insight into police officer decision making, Dr. Withrow found that the data RIPA used for the 2022 report is incomplete, and their methodology does not hold up to statistical rigor. This severely limits the conclusions we can make from the data and calls into question the value of the data and conclusions drawn by the RIPA Board.

Upon his review of the 2022 RIPA Report and data, Dr. Withrow has drawn the following conclusions, the rationale and justification for which are contained in the body of this analysis:

- **RIPA Board Chose Not to Include Data on 57.7% of All Stops**: California Highway Patrol data, which accounts for 57.7% of all stops in the state, was not included uniformly throughout all analyses in the report – meaning that the analysis only considered 42.3% of possible stops yet made sweeping generalizations.
- **RIPA Board Distorts Definition of Racial Profiling:** California defines racial profiling as requiring an officer to initiate a stop based on race – i.e., the officer must know the race or ethnicity of the driver ***prior*** to the stop. However, the data collected in this report was collected ***after*** the stop already occurred.
- **RIPA Board Employs Fundamentally Flawed Statistical Analysis**: The study includes only one independent variable – the driver's race or ethnicity – meaning that race or ethnicity is ***the only*** attributable explanation for why an officer initiated a stop. This fails to consider vitally important contextual information and a myriad of other potential motivations or justifications for the enforcement action.
- **RIPA Fails to Conduct a Thorough Literature Review**: The report cherry picks print and TV news articles as "sources" to support RIPA's stance – neglecting to provide a literature review of pre-existing studies.

Introduction

Routine police stops are likely the most ubiquitous and visible manifestation of government authority. A report from the U.S. Department of Justice in 2020 informs us that 24 percent of U.S. residents that are sixteen years of age or older had at least one contact with the police in 2018. This is an increase from 21 percent reported for 2015. About 11 percent of these stops are police-initiated. The most common type of stop is when residents are pulled over during a traffic stop (Harrell & Davis, 2020).

Prior to 1994, if a police chief were asked even basic information about police stops, it is not likely such information would be readily available. Outside of a few attempts at gathering data on police stops as measures of productivity, basic information about who gets stopped, why they are stopped, what happens to them during stops, how long stops last, etc. was largely unknown to most police administrators. Then came the New Jersey Turnpike Study. In 1994, Professor John Lamberth published his statistical analysis of traffic stops occurring on the New Jersey Turnpike (Lamberth, 1994). Despite its methodological problems, this report ushered in what we now know as the field of racial profiling statistical analysis. Perhaps the most important contribution of the Lamberth study is the realization that not all people have the same experiences during a police stop, or come away from them merely annoyed because they were issued a citation.

Nearly three decades later, the amount of information available on police stops is staggering. Hundreds, perhaps thousands, of scholarly studies, reports, data collections, litigations, court rulings, etc. have produced an immense body of knowledge on millions of police stops. The findings within this body of knowledge are consistent. With rare exception, we know that racial and ethnic minorities appear to be overrepresented in police stops when compared to estimates of their proportional representation of drivers either available or at risk of being stopped. Racial and ethnic minorities appear to be searched more frequently, arrested more frequently, and engage in more physical confrontations with police officers than other residents. Unfortunately, we really aren't sure that our estimates of the driving population are accurate. We really do not know *why* racial and ethnic minorities are searched and arrested more frequently. And, we are *not really certain* about the origin of physical confrontations between the police and members of the public.

The purpose of this report is to offer a critical analysis of the Racial & Identity Profiling Advisory Board's 2022 Annual Report (the RIPA Report 2022). The RIPA Report 2022 contains detailed analyses on 2,937,662 police stops (pedestrian and vehicle) occurring within the State of California from January 1, 2020, through December 31, 2020, from 18 law enforcement agencies. The agencies providing information on their stops represent most of the largest law enforcement agencies in California. Most notably, more than half of the stops (57.7 percent) were conducted by the California Highway Patrol.

This critical analysis begins with a qualitative overview of the common methodological and analytical challenges that threaten the racial profiling research agenda's ability to find explanations for what appears to be the over representation or disparate treatment of historically marginalized individuals. This is followed by a more detailed criticism of the methodological and analytical strategies used by the RIPA analysts, as well as the conclusions drawn from these strategies. This report's conclusions offer some suggestions on how the data collection effort may be improved to enhance our understanding of police stops, as well as some additional insight into police officer motivation.

Qualitative Overview of Methodological and Analytical Challenges in the Racial Profiling Research Agenda

On the surface, racial profiling research appears relatively straightforward. It would appear to only require a measurement of what is supposed to happen with what actually happened. If those measures are not equal, then something might be wrong, or more specifically, somebody is to blame. Unfortunately, the practice of policing is far more complicated than what can be explained by current data sets. We almost never know what should happen because important explanatory factors exist within the context of every stop that are not measured. We are a little better at measuring what actually happened, but seldom do we collect enough data to provide useful insight into an officer's motivation for what actually happened. Here are some of the common methodological and analytical challenges in the racial profiling research agenda. They are offered here because they are relevant to the more detailed critical analysis in the section that follows.

<u>Estimating the driving population</u>

It is surprisingly difficult to measure the demographic features of a driving population. Traffic patterns and the drivers within them change constantly. Historically, racial profiling researchers have utilized three benchmarking strategies.

Population-based benchmarks are by far the most commonly used. Most are based on the racial/ethnic proportions within the residential population as reported or estimated by the United States Bureau of the Census or other authoritative sources. If the police stop data contain information on all individuals (including juveniles) who may come in contact with the police, then the researcher should use the entire population of the community (Withrow 2002a, 2006). Some researchers only use the estimated population of licensed drivers (Smith and Petrocelli 2001). And a few researchers adjust the population using weighting factors like vehicle ownership (Rojek, Rosenfeld, and Decker 2004) or population centers within metropolitan area (Novak 2004).

Field observation-based benchmarks are developed by systematically observing drivers in traffic at and during randomly selected locations and times. Most studies attempt to record the race, ethnicity, gender, and age of the observed drivers (Police Foundation 2003). In addition, some researchers attempt to identify the drivers (also by race/ethnicity, age, and gender) who are observed violating the traffic law, and therefore more likely to be stopped (Lamberth 1994). One research team even used digital cameras and speed detection technology to record images of drivers and their speed (Lange, Blackman, and Johnson 2001). Field observation benchmarks are then compared to the police stop data that is collected at or near where the benchmarks are collected.

Accident records are used extensively by traffic engineers and automobile insurance companies to develop risk factors among drivers (Alpert, Smith, and Dunham 2003). This is one reason why young drivers pay high automobile insurance rates. In racial profiling research, the focus is on the not-at-fault drivers in two vehicle accidents. The Washington State Patrol (2001) introduced this technique to the racial profiling research agenda. The logic of this benchmarking strategy is based on two factors. First, from the not-at-fault drivers' perspectives, an accident is a random event. An appropriate sample of these events provides insight into the race, ethnicity, gender, and age of the driving population. Second, individuals who drive more often are more at risk of being involved in an accident. They are also more at risk of being observed by a police officer. So, accident records-based benchmarks naturally include a weighting factor for driving frequency.

The RIPA Report 2022 utilizes a relatively new strategy called the Veil of Darkness analysis. This strategy assumes that during the intertwilight time (i.e. at night) police officers are less able to see the race or ethnicity of a driver prior to the stop and are therefore less able to initiate a stop based on *racial animus*. This analytical technique was first developed in the late 1960's (Black, 1971; Black & Reiss, 1970; LaFave, 1965, Reiss, 1971). It was later developed for use in racial profiling research by Engel & Calnon, 2004, Grogger and Ridgeway, 2006, Smith & Petrocelli, 2001, Smith et al, 2004, Smith et al, 2017, and Withrow, 2006). While it would seem this strategy would overcome some of the measurement challenges of previous benchmarking strategies, it is not at all clear that it is as effective as proposed (Stacey & Bonner, 2020).

Taniguchi, et al, 2017 finds when omitting stops between sunset and the end of civil twilight that the officers' ability to identify race did not influence who was being stopped (see also Oakland Police Department, 2004). Other researchers express concerns about the inability of this strategy to account for seasonal differences in the length of the intertwilight time (Ridgeway, 2009, Worden, et al, 2012, Taniguchi et al, 2017). Other researchers are critical of this strategy because it does not account for variations in ambient lighting (Horrace & Rohlin, 2016).

A final strategy that deserves some mention here, because it may provide more insight within this particular research context, is called an internal benchmark. This strategy compares the stops among similarly situated police officers in an attempt to identify specific officers that appear to be exhibiting troublesome behaviors. This method was first proposed as a means to identify officers with potential problematic behaviors (Walker, et al 2001). Later, Withrow, Dailey, and Jackson (2008) demonstrated the effectiveness of this strategy within a racial profiling context.

The misapplication of classic probability

Racial profiling research reports often contain various phrases suggesting a probabilistic outcome. Phrases like, "Asian drivers are underrepresented in stops", Black drivers are overrepresented in stops", "Hispanic residents are searched at a higher rate than expected or anticipated", and "Black drivers are more than twice as likely to be involved in physical confrontation with the police". These statements likely come from a common misunderstanding of a popular statistical technique called the Classic Probability Model. This model proposes that the probability of any outcome can be calculated by dividing its possible frequency of occurrence by the total number of possible occurrences. Probability is always expressed in decimal form.

Number of desired outcomes/Number of total possible outcomes

Using this model, one can estimate the probability of rolling a '2' on a single die. There is only one '2' on a single die and there are a total of six possible outcomes.

Number of 2's on a single die/Total number of possible outcomes on a single die

(1/6) = 0.167

All statistical models, including this one, have associated assumptions (i.e. rules) that regulate when they can be used, as well as how their results should be interpreted. Most importantly, one of the key assumptions for this model is random selection. In order for this model to work, all possible outcomes must be equally possible. In the example above, all of the six outcomes of a roll of a single die must be equally possible. No loaded dice. Formally, all possible outcomes must have an equal and nonzero chance of happening.

For the Classic Probability Model to work as intended in a study of police stops, one must assume that all individuals within the population (however defined) have an equal and nonzero chance of being stopped. We know this is not the case. Nothing happens randomly (in the scientific sense of the word) in police operations. For example, people that live in neighborhoods which emanate high numbers of calls for service are inadvertently subjected to higher levels of routine police supervision. As a result, these individuals are at a higher risk of being stopped. Other factors, like age, gender, driving performance, etc. play a role in reducing the randomness within the process by which individuals are 'selected' into the 'sample' of drivers that are stopped.

<u>The rules of causality</u>

An allegation that something is the cause of something else is easily adjudicated in the physical sciences where concepts and change can be measured empirically. For example, if we hypothesize that a baseball hit by a wooden bat swung at a particular angle and with a particular amount of force is likely to result in a home run at Fenway Park in Boston, then we can actually test this. Determining causal relationships in social science research is a little more difficult. Our concepts often do not lend themselves well to empirical measure. Luckily, social researchers agree on a general rule that defines a causal relationship.

In order to confirm that any factor is the cause of any other factor, a researcher must establish three things. The first is temporal order. This simply means that the cause must precede the effect. The second condition is correlation. The relationship between an alleged cause and effect must be consistent and demonstrable. Third, the researcher must remove all plausible alternative explanations. If other plausible alternative explanations remain, then the original allegation of a causal relationship is determined spurious (i.e. false).

This series of causal rules is a generally agreed upon standard in social research (Withrow, 2016), appropriate in racial profiling research, and it is recognized as legitimate in court rulings (see for example, U.S. v. Alcaraz-Arellano, 2006). An allegation of racial profiling alleges that a police officer initiates a stop based a *racial animus* towards an individual because of that individual's race or ethnicity. In order to establish that race is the actual cause (or even a motivation) for a stop, we must demonstrate that the officer knew the driver's race *prior* to the stop and was motivated by *racial animus*. Racial profiling data sets (including the RIPA Report 2022) typically do not measure whether the officers knew the race of drivers prior to the stop. Racial profiling researchers (including those that analyzed the data for the RIPA Report 2022) did not establish a correlation between the races of individuals stopped and the races of individuals available to be stopped. And finally, racial profiling research nearly never attempts to eliminate plausible alternative explanations for an alleged disparity with respect to race, ethnicity, or other identity-related factors. The RIPA Report 2022 is no exception.

The importance of context

There is a common trick played by interview boards during the process of evaluating candidates for police officer positions. It usually starts when a member of the board asks the candidate a question like, "If you were a police officer today, how would you respond if you saw a ten-year-old girl walking alone wearing a backpack in a residential neighborhood?" Normally, the candidate (in an attempt to impress the board) will respond, "I would exit my patrol vehicle and question the little girl to make sure she is safe". The interviewer then asks, "Really? What if it is a school day and she appears to be walking toward a school bus stop?" The candidate will usually change his answer. To which the board member will suggest the possibility of the candidate's lack of decisiveness. It is a game. It is done on purpose to determine how the candidates incorporate changes of context in their decision making. It also demonstrates a reality in policing: that the legitimacy of a decision in a large part depends on the context in which an event occurs.

Context matters in racial profiling research because knowing the context of an event provides a more informed means of evaluating the legitimacy of an officer's behavior. Measuring the contexts of stops is incredibly challenging, given the near limitless number of possible factors that could happen and do influence a police officer's decisions. The data set used by the RIPA Board contains several variables that attempt to capture detailed information about the context of stops. The attributes for these variables appear on their face to be exhaustive. Unfortunately, they are not. For example, the variable describing the reason for the stop contains only eight attributes. Most stops (86.1 percent) are predicated on a traffic violation. These stops can further be described by another variable called type of traffic violation. This variable has four attributes. Most stops for a traffic violation (73.2 percent) are defined as a moving violation. These stops can be further defined by a statutory citation from the traffic code. This level of detail is remarkable. Unfortunately, none of these variables actually measure the severity or flagrance of the alleged offense, which is strongly predictive of a stop and known to influence an officer's desire to initiate a stop as well as the outcome of a stop.

Defining racial profiling

How racial profiling is defined is important to our analysis and not unlike our use of the criminal code that defines the elements of a crime. For example, in most cases a charge of driving while intoxicated requires that the driver be intoxicated, operating a motor vehicle on a public road. Some states allow DWI charges to go forward if the driver is operating a motor vehicle in a public place, like a parking lot. The point is that the state must prove all elements of the crime beyond a reasonable doubt before a verdict of guilty can be reached.

There are two competing definitions of racial profiling. The first commonly used definition was offered by Ramirez, McDevitt, & Farrell, 2000:3.

> …any police-initiated action that relies on the race, ethnicity, or national origin rather than the behavior of an individual or information that leads the police to a particular individual who has been identified as being, or having been, engaged in criminal activity.

This is a conceptual definition of racial profiling. To sustain an allegation of racial profiling the accuser must be able to prove that:

- The police officers were aware of an individual's race or ethnicity prior to initiating the stops, and
- The police officers used this information as a reason for initiating a stop.

The other type of definition was originally offered by Lamberth (1994:4).

> …implied as when minorities are stopped at disproportionately higher rates than they are represented within the benchmark that indicates the proportional racial representation of actual roadway users.

This is often called an operational definition of racial profiling. Using this definition to prove racial profiling requires only that the percentage of stops involving members from a particular race or ethnic group is higher than they are represented in a measure of the driving population.

There appear to be two legal definitions of racial profiling in California law. The first comes from Chapter 684 and defines racial profiling as "the practice of detaining a suspect based on a broad set of criteria which casts suspicion on an entire class of people without any individualized suspicion of the particular person being stopped." The second comes from Assembly Bill 953, Section 4(e) as follows.

> (e) "Racial or identity profiling," for purposes of this section, is the consideration of, or reliance on, to any degree, actual or perceived race, color, ethnicity, national origin, age, religion, gender identity or expression, sexual orientation, or mental or physical disability in deciding which persons to subject to a stop or in deciding upon the scope or substance of law enforcement activities following a stop, except that an officer may consider or rely on characteristics listed in a specific suspect description. The activities include, but are not limited to, traffic or pedestrian stops, or actions during a stop, such as asking questions, frisks, consensual and nonconsensual searches of a person or any property, seizing any property, removing vehicle occupants during a traffic stop, issuing a citation, and making an arrest.

It would appear both of these definitions are conceptual, meaning that to sustain an allegation of racial profiling the accuser must *at a minimum* prove that the officers knew the race, ethnicity, gender status, disability, etc. of drivers prior to initiating the stop and used that information as a justification for the stop. The data set used as the basis for the RIPA Report 2022 does not measure what the officers knew about the people they stopped prior to initiating the stop. In other words, under California's own legal definition, the current data set cannot establish an accusation of racial or identity profiling.

Critical Analysis of the RIPA 2022 Report

The RIPA Report 2022 contains information about 2,937,662 stops conducted by eighteen agencies throughout California occurring from January 1 through December 31, 2020. This is the third year that the RIPA Board has published this report. The number of agencies that provide data for this report increased from 15 to 18 agencies in 2020. Despite this increase of agencies, the number of actual stops decreased from 2019 when 3,995,686 stops were reported. The RIPA Report 2022 explains this reduction was caused by the travel restrictions imposed on Californians due to the COVID-19 pandemic.

The following critical analysis focuses on the key parts of the RIPA Report 2022 that attempt to provide evidence of disparity with respect to race, ethnicity, gender identity, and disability. The conclusions drawn in this critical analysis are limited to a review of the RIPA Report 2022, its associated appendices, and a review of the data set available.

<u>Measuring who gets stopped</u>

Perhaps the most fundamental statistic captured in a racial profiling analysis is a measure of the race or ethnicity of the individuals stopped. The measure of who gets stopped by race or ethnicity in the RIPA Report 2022 is unnecessarily complicated, and because of this it is collected two times.

There is a single variable (RAC_FULL) that describes an individual's race or ethnicity. This variable has eight attributes (Asian, Black/African American, Hispanic/Latino, Middle Eastern/South Asian, Native American, Pacific Islander, White, and Multiracial). In addition, the race or ethnicity of a driver may be recorded in a series of other single variables (e.g. RAE_ASIAN, RAE_BLACK_ AFTICAN_AMERICAN, etc). For these variables the attributes are either 'NO' or 'YES'. The variables that measure an individual's gender/gender status and disability are similarly formatted. It is not abundantly clear why the RIPA Board collects this information in two places, but the confusion is evident. Furthermore, due to the fundamental importance of these particular measures, the confusion is confounded in all analyses throughout the RIPA Report 2022.

The RIPA Board (p. 29) reports that in 2020;

> Officers perceived the highest proportion of individuals they stopped to be Hispanic (40.4%; 1,187,728), followed by White (31.7%; 929,776), Black (16.5%; 484,364), Asian (5.2%; 151,813), Middle Eastern/South Asian (4.7%; 136,806), Multiracial (0.9%; 25,777), Pacific Islander (0.5%; 15,292), and Native American (0.2%; 6,105).

The footnote associated with this statement (#56) is;

> Officers may select multiple racial/ethnic categories per individual when recording stop data. To avoid counting the same stopped individual in multiple racial/ethnic groups, all stopped individuals whom officers perceived to be part of multiple racial/ethnic groups were categorized as Multiracial. The distribution of the race/ethnicity categories that officers selected when they selected more than one category was as follows: Asian (21.9%), Black (31.6%), Hispanic (71.5%), Middle Eastern/South Asian (27.8%), Native American (15.4%), Pacific Islander (17.0%), and White (65.8%).

It is not clear which of the two variables that are available to the officers to report the race or ethnicity of the drivers they stopped were actually used in the RIPA Report 2022. Interestingly, the percentages reflected in footnote 56 total to 252.0 percent, indicating that the officers perceived a great deal of racial and ethnic diversity in the individuals they stopped. There is no percentage associated with 'Multiracial' in footnote 56, even though it appears the officers were allowed to select this attribute in both variables. The RIPA Report 2022 explains that cases containing more than one race or ethnicity were redefined as Multiracial, but provides little documentation as to how this was done, and most importantly, the effect of this on the distribution of individuals by race or ethnicity. For example, a review of the data set indicates nearly 100 different races and combinations of races reported by police officers.

Because of this, there is potential for a reactive effect in the data collection. Reactivity is a condition whereby research subjects change behavior when they are aware that they are being observed. In this case, the concern is that this reporting protocol may mask the actual number and percentage of stopped individuals by race or ethnicity.

A similar confusion exists in many of the tables listed in the appendices. For example, Table A.4. Stops by Identity Group and Calls for Service reports that of the stops involving Hispanic individuals, 1,128,563 (95.0 percent) were officer-initiated stops and

59,165 (5.0 percent) are calls for service stops. The total for this row is reported as 1,186,470 (100 percent). The actual sum of these two types of stops involving Hispanic residents is 1,187,728, an additional 1,258 stops. Overall, there are 2,910 more stops in the actual total than in the reported total.

Similar incongruences exist in the tables associated with stops by other identity categories. The differences are admittedly slight and do not change the overall findings of their associated analyses. What is important here, at least from a data quality perspective, is to account for (and report) cases wherein data are missing. This information provides important indicators of reporting inconsistencies and potential training needs.

When (during the stop) identity is perceived

Throughout the RIPA Report 2022 the phrase "perceived to be…" followed by one of the six identity-related categories appears multiple times. Of course, we cannot actually know whether this is accurate because officers are not permitted to ask individuals to self-identify for data collection purposes (p. 29). It is likely some of this information may be imperceptible by the officer.

More importantly, the point at which the identity-related category is 'perceived' by the officer is important for sustaining an allegation of profiling. Racial profiling in California is defined conceptually. In order to sustain an allegation of racial profiling, or any other related misbehavior, what the officer knew about the individual stopped *prior* to initiating the stop is an essential condition, the other being some evidence of discriminatory intent. There is no measure of what the officer knew about the driver prior to initiating the stop, much less whether the officer acted with an intention to discriminate. As a result, it is not possible from the RIPA data to allege that individuals were stopped on the basis of their identity-related category.

Evaluating the reason for the stop

The reason for the stop is an important evaluation within the racial profiling research agenda. Associated with this is a potential for a disparate impact. Historically, there has been some concern that racial and ethnic minorities are stopped for less serious violations of the law, thereby producing the probable cause necessary to legally justify a stop.

The RIPA Report 2022 reports that most stops are predicated on a traffic violation (86.7 percent), and most of those are moving violations (73.2 percent). The next most frequently reported reason for a stop is reasonable suspicion (11.5 percent). There are a total of eight attributes in the reason for the stop variable (REASON_FOR_STOP) and a total of three attributes that may further describe a moving violation (RFS_TRAFFIC_VIOLATION_TYPE). In addition, the officer may report the actual traffic code violation to further describe the reason for traffic violations (RFS_TRAFFIC=VIOLATION_CODE). This final qualifier has the potential for measuring the relative severity or flagrance of an alleged traffic violation. Unfortunately, this information is not considered in the RIPA Report 2022.

The analyses presented within the RIPA Report 2022 comparing the alleged violation by traffic code and race or ethnicity of the driver (Tables 5-8, pp. 137-140) indicate a high degree of correlation in actual alleged traffic code violations between racial and ethnic groups. The same patterns exist with respect to other identity-related categories. This indicates that, for the most part, identity-related category does not play a role in the reason for the stop.

It is important to note that data on stops from the California Highway Patrol (57.7 percent of the data set) are excluded from some of these analyses. The reason for this is not clear, especially when this portion of the data is included, and the overall findings do not change.

Strikingly absent in any of these data or their analyses is any mention of how the relative flagrance or severity of an alleged traffic violation (moving or non-moving) affects anything associated with the stop (e.g. outcome). As previously mentioned, an understanding of the context within which stops occur is critically important. For example, a speeding violation could be 80 miles per hour in a 70 miles per hour on a rural interstate highway or 30 miles per hour in a 20 miles per hour in an active school zone. Both are speeding violations, but there is a substantial difference between them with respect to the potential threat to public safety. Overlooking this nuance within the data collection protocol eliminates the possibility of more meaningful inquiry into officer decision making.

It appears there is some confusion with respect to the reporting of stops predicated on reasonable suspicion. The RIPA Report 2022 explains this in footnote 74 appearing on the bottom of page 35, "Reasonable suspicion may be reported as a reason for the stop when an officer suspects criminal activity or when officers initiate a contact for "community caretaking purposes". The result of this is an unreliable understanding of the context in which officers are reporting this reason. The RIPA Board has proposed changes that will clarify their reporting protocols in this regard. Until then, the reasonable suspicion attribute within the reason for the stop variable is unreliable (meaning inconsistently reported).

Non-moving violations represent a small portion of the total traffic violations (13.7 percent). About half of the stops (51.4 percent) within this category are related to no registration, display of plates/tags, and failure to comply with commercial vehicle rule. The other half (42.8 percent) are equipment violations. The RIPA Report 2022 largely ignores non-moving violations, even though on page 14 they allege that non-moving violations could be "ripe for pretext if an officer was using minor traffic violations to take further actions". One wonders why this year's report does not include such analyses. The concept of a pretext stop is discussed later in this analysis.

Finally, there is a short analysis of bicycle-related stops, again a very small number of stops. This issue is raised herein because the analytical strategy presented in the RIPA Report 2022 does not appear to consider the demographics (e.g. race or ethnicity) of the individuals that live in the areas where bicycle violations are likely to occur. One would anticipate that bicycle stops would be higher in locations where a bike lane (or similar accommodation for bicycle travel) is not available. These neighborhoods may be populated principally by racial or ethnic minority residents. A simple adjustment in the population-based benchmark (focusing only on the census tract or block) would be necessary for this analysis.

<u>Measuring what happens during stops</u>

There are a series of variables that may be used to indicate actions taken (by the police officer) during the stop (ADS_). The RIPA Report 2022 indicates that there a total of 23 possible actions that an officer can take (p. 37). A count of the variables within this category indicates there are closer to 25. Either way, this number of variables describing the events occurring in each stop could provide substantial insight into officer decision making. Unfortunately, the data variables are neither appropriately conceptualized nor associated with other variables that might explain officer behavior.

The RIPA analysis indicates that in 80.9 percent of the stops the officers took no action. This is astonishing, given the detail available in this series of variables. The result of this is there appears to be no appreciable differences in the number of actions taken by police officers with respect to the identity-related category (see Appendix A.6). Some variation may exist between identity-related categories when all variables are spread across all categories. Unfortunately, the data set does not collect information that would explain these differences. Here are some examples of why this is important.

Handcuffing may or may not be a punitive action. Individuals are handcuffed for a variety of reasons (e.g. pursuant to an arrest, officer security, to protect an individual from self-harm or harming from another person, etc.). The RIPA data set only allows an officer to select 'handcuffed'. No other variables existing within the data set would appear to indicate the policing objective associated with this action.

Curbside and patrol car detentions may appear punitive, however, they may be necessary to protect an individual from further harm, to separate an individual from a suspected assailant, to conduct required searches of vehicles (e.g. inventory or pursuant to an arrest), or for officer protection. Unfortunately, it is not possible to collect this descriptive data that would explain officer motivation within the context of the current data collection protocols.

It is not possible to tell from the data available in the appendices whether or not individual drivers are being double counted. For example, Appendix A.8 All Actions Taken During Stop by Race/Ethnicity, 7.3 percent of all stops involving Black residents included 'Removed from the Vehicle by Order' and 7.5 percent of all stops involving Black residents included 'Patrol Car Detention' as an action taken during the stop. The police have an obligation to protect all residents. A person ordered removed from a vehicle during a stop must be placed in the most secure location possible. During a traffic stop, this is likely a patrol car.

The issue here is that each of the variables that include a measure of what happens during a stop are independently selected, rather than being arranged as attributes of a single variable. This means each 'action' is a separate event when they should actually be considered as a combination of events occurring within the same stop. It is not clear whether the RIPA analyst considered this. The series of variables used to document actions taken during the stop are largely a collection of unconnected discrete descriptors. There is no provision that would allow a police officer to indicate the justification for initiating these actions. Furthermore, there is no provision for reporting whether these actions were initiated solely at the discretion of the police officer or as a result of another individual's behavior. In addition, and perhaps more importantly, the order in which these actions happened is not possible to determine from these data. For example, whether the individual was handcuffed *after* threatening an officer's safety is not possible to determine because of the manner in which the data are collected.

The analyses of these variables are confounding because additional factors that may justify an action are not defined. For example, on page 8 of the RIPA Report 2022 there is a table representing the percentage of stops by race that include searches, curbside/patrol car detention, handcuffed, or ordered vehicle exit. This analysis with respect to searches is problematic because it appears to include all searches. It would appear beneficial to exclude non-discretionary searches (e.g. pursuant to an arrest or vehicle impoundment), searches predicated on a review of the evidence by an independent magistrate (search warrant), or searches predicated on exigent circumstances from this analysis.

Finally, the large percentage of stops resulting in 'no action' suggests that the overall variables describing what happens during a stop are poorly conceptualized. Just because an officer indicates 'no action' does not mean that the officer actually took no action. It only means that a better attribute describing what actually happened is not available.

Assessing the results of stops

The variables that document the results of the stop are organized into thirteen discrete variables that may be further described in four additional variables. Officers are allowed to report multiple results for a single stop. This series of variables includes another possibility for a stop to result in no action. It is not clear how the analysis managed cases that appear to have no action reported as a stop event or as a stop result.

The analysis reveals higher proportions of stops involving Black residents (13.1 percent) and Multiracial residents (7.2 percent) resulting in no action. By comparison, 5.6 percent of all stops involving White residents resulted in no action. Alternatively, because the percentage of stops resulting in no action for Black residents is high, stops involving Black residents had lower percentages of stops resulting in citations, warnings, or arrests. The RIPA analysis provides no explanation for this difference, or whether the difference is statistically significant.

Determining overrepresentation and disparate impact

The RIPA Report 2022 presents four independent analyses used to determine the presence of disparity in police stopping behavior. Two of these analyses focus on who (meaning identity-related category) gets stopped and whether or not any one of these groups of residents are overrepresented in stops. These involve comparing stop data with a benchmark based on the residential population by race or ethnicity and a comparison of stops occurring during the daytime and nighttime. The other two analyses focus on what happens to individuals (by identity-related category) once they are stopped. These analyses focus on searches and use of force.

Comparisons to residential population data

The objective of this analysis is to determine whether individuals of particular identity-related groups are either overrepresented or underrepresented in stops. The principal focus of this analysis is on the identity-related categories for race and ethnicity.

The analytical strategy of this analysis is quite simple. It involves a comparison between the proportions of individuals stopped by race or ethnicity with the proportions of individuals that might be stopped by race or ethnicity. The RIPA data set contains the racial and ethnic identities of the individuals stopped by the police, or at least the officers' post-stop subjective perceptions of these residents. Because there is no specific measure of the actual driving population with respect to the race or ethnicity of the drivers, the RIPA analysts used a proxy measure based on the residential population. Specifically, they used the 2019 edition of the American Community Survey (ACS). The ACS is an estimate of a population, typically done in the years between the decennial censuses. Using the residential population to estimate the driving population is based on an assumption that "who is stopped would be similar to who resides within a comparable geographic region" (RIPA Report 2022:48). Unfortunately, this assumption is false.

Estimates of the driving population based on an associated residential population are not useful for numerous reasons. First, an estimate based on the residential population does not account for transient drivers, i.e. individuals that drive through an area but do not live there. The error associated with this could be substantial in communities that have interstate highways or major state and regional roadways coursing through them. Second, population estimates represent who lives in an area but do not estimate how much they drive. For example, a 35-year-old Hispanic male who works as a delivery driver likely drives all day, while his 89-year-old grandfather drives only occasionally. Both individuals count equally in the census as Hispanic males, but their relative risk of being observed by a police officer are quite different. Third, population-based estimates of the driving population do not account for differential levels of exposure to routine police observation. The error associated with this happens in two ways.

Numerous factors (other than race or ethnicity) affect how much and when residents drive. Age, occupation, socio-economic status, and the availability of public transportation are examples of these factors. The potential for error in this regard is particularly poignant for this year's analysis because the stops occurred during government-imposed lockdown periods caused by the COVID-19 pandemic. It is likely the travel restrictions imposed during the pandemic changed the demographics of the driving population. There is a growing body of evidence that Black and Hispanic individuals are less likely to be able to work at home (U.S. Bureau of Labor Statistics, 2020, Economic Policy Institute, 2020). As a result, it is likely that the proportions of Black and Hispanic drivers in 2020 were higher than the proportions of Black and Hispanic residents estimated by the 2019 ACS.

The other factor causing differential rates of routine police observation is related to where police officers are assigned to work. Police resources, specifically patrol, are not assigned evenly throughout a jurisdiction. Areas of a community that experience high levels of crime/victimization or from which numerous calls for service are received are assigned additional patrol officers. This means that the residents of these areas are inadvertently subjected to higher levels of routine police observation. Unfortunately, many of these areas may also be populated principally by racial and ethnic minority residents. The RIPA Report 2022 indicates that the analysts were aware of some of the factors that produce errors in this estimate, but they made no attempt to account for them.

The RIPA Report 2022 reports four principal computations for this analysis.

1. The absolute difference – the difference between the percentage of individuals stopped and the percentage of individuals estimated by the ACS, by race or ethnicity.
2. A relative percentage difference – The absolute difference divided by 100, which reflects a percentage difference. When the result of this calculation is negative, it indicates that individuals from that racial or ethnic group are underrepresented in stops. When the result of this calculation is positive, it indicates that the individuals from that racial or ethnic group are overrepresented in stops.
3. A disparity index – Calculated by dividing the percentage of individuals stopped by the percentage of individuals estimated available for stop, by race or ethnicity. A disparity index below 1.00 indicates individuals from that racial or ethnic group are underrepresented in stops. A disparity index greater than 1 indicates individuals from that racial or ethnic group are overrepresented in stops.
4. A ratio of disparity. This is calculated for each minority group. It represents the disparity index for each minority group of color (m) divided by the disparity index for White individuals (w), or $E(m)/E(w)$.

A consistent pattern emerges from these data. RIPA bases its conclusion primarily on the disparity index. They consistently find that Black residents are stopped in higher proportions than they are estimated in the residential population. The disparity indices for the overall analysis are 2.52 and range from 1.92 (Fresno PD) to 5.54 (Davis PD). The disparity indices for Black residents are consistently well over 1.00, indicating an overrepresentation of Black residents. Specifically, the RIPA Report 2022 concludes:

> Overall, the disparity between the proportion of stops and the proportion of residential population was greatest for Multiracial and Black individuals. Multiracial individuals were stopped 81.6 percent *less frequently than expected*, while Black individuals were stopped 151.5 percent *more frequently than expected*. The proportion of stops corresponding to Hispanic individuals most closely matched estimates from residential population data (4.7% *more frequent than expected*). Compared to White individuals, who were stopped 10 percent *less frequently than expected* based on their share of the residential population, the greatest disparities between stop data and residential population data estimates occurred for Black and Multiracial individuals. The disparity for Black individuals was 2.8

> times as great as the disparity for White individuals. For Multiracial individuals, the disparity was 0.2 times as great as the disparity for White individuals. This indicates that individuals perceived as Black were *substantially more likely* to be stopped compared to White individuals, while individuals perceived as Multiracial were *substantially less likely* to be stopped. After excluding California Highway Patrol records from the analysis, the data continued to show the greatest disparities for the stops of Black and Multiracial individuals; relative disparities compared to those of White individuals were larger than the all-agency disparities for individuals perceived to be Asian, Black, Hispanic, and Pacific Islander (p. 50, italics mine).

The phrases used in this conclusion of particular concern are "less frequently than expected", "more frequently than expected", "substantially less likely", and "substantially more likely". These phrases suggest that there is a probabilistic nature governing who gets stopped. There is not. It is not possible to legitimately calculate probabilities like this because not all drivers have an equal probability of being stopped. Police officers do not stop individuals randomly. The probability of being stopped is much more complicated than race or ethnicity. The RIPA Report 2022 makes no attempt to include additional factors that might play a role in who gets stopped. Adjusting the residential benchmark to include narrower measures of population (i.e. by neighborhood, patrol beat, census tract or block) would likely affect the outcome of these analyses (Withrow, 2002b). In the RIPA Report 2022, the effects of race or ethnicity on the probability of a stop are actually unknown because there is no evidence in the data set that the officers even knew the resident's race or ethnicity prior to the stop.

The calculation for the ratio of disparity is especially misleading. This ratio is computed by dividing the disparity index for the residents of each racial and ethnic minority group by the disparity index for White residents. In most cases, the disparity index for Whites is one (1) or less, meaning that Whites tend to be underrepresented in stops. Because disparity indexes for minority residents tend to be greater than one (1), the effect of this overinflates the disparity for minority residents. Using stops involving White residents (or for any other identity-related category for that matter) as a baseline has no methodological or analytical justification, other than to misrepresent the disparity indexes associated with other races or ethnicities.

Veil of Darkness Analysis

This analysis involves comparing stops occurring at night with stops occurring during the day, with respect to the race or ethnicity of the drivers. The utility of this analysis to racial profiling research is based on an assumption that police officers are less likely to perceive an individual's race or ethnicity at night (when it is dark) than they are during the day (when it is light). Proponents of this analytical strategy argue that a finding that the percentage of stops involving a particular racial group is higher during the day than during the night supports an allegation of racial profiling. Alternatively, this finding could also support the hypothesis that individuals from that particular racial group drive more during the day than at night. The RIPA Report 2022 explains this as follows:

> In other words, to the extent that it is harder to identify someone at night, we would expect darkness to decrease the likelihood that individuals of racial/ethnic groups of color are disproportionately stopped relative to White individuals (p. 56).

This part of the analysis relies on four separate (independent) statistical models. Each attempts to estimate the effect of a resident of color being stopped in darkness versus a White resident being stopped in darkness. The four models compare:

- Stops involving Black residents versus stops involving White residents
- Stops involving Hispanic residents versus stops involving White residents
- Stops involving Asian residents versus stops involving White residents
- Stops involving Other race residents versus stops involving White residents

The Other race category was created by combining stops involving Middle Eastern/South Asian, Multiracial, Native American, and Pacific Islander residents. The results from Table D.3 Veil of Darkness Analysis Table indicate two of the four analyses produced statistically significant results. Statistical significance is a measure of the probability that a statistic is due to random chance. It does not estimate the actual effect of an alleged causal variable on an outcome variable. The coefficient for the Black/White analysis is -0.021. The coefficient for the Hispanic/White model is -0.021. These coefficients are not particularly remarkable (i.e. powerful) with respect to their actual effect on the dependent variable. Depending upon how the models were set up, these coefficients would need to be much greater in absolute value to have a perceptible effect on the outcome variable.

Unfortunately, the results offered do not provide insight into whether other variables included in the model would have more effect. Furthermore, the amount of overall variance explained by these models is comparatively low at 0.358 (about 36 percent for the Black/White model) and 0.230 (about 23 percent for the Hispanic/White model). This suggests that other factors, not included in this analysis, may play a more profound role estimating the effect of race or ethnicity on the probability of a stop at night. To be fair, the RIPA analysts recognized this when on page 58 of the RIPA Report 2022, they write, "These disparities could reflect biased police behavior *or the effect of some factor that is not yet being considered by this test*." (italics mine).

Here again, as in the previous analysis, there is no theoretical, methodological, or analytical justification for using the proportion of stops involving the residents of one racial group (White) as a factor that might affect the probability of stops involving residents from other racial groups (all other racial and ethnic groups). If race is less likely to be perceived at night, then this factor affects all races of residents. Including this variable in the statistical models produces unnecessary complication and, most importantly, generates a statistical misrepresentation.

In addition, the RIPA Report 2022 indicates other "fixed factors" were included in these statistical models. They include time of day, day of week, month, and officer conducting the stop. It is assumed the variables describing when the stop occurred (i.e. time of day, day of week, and month) are necessary for differentiating between stops occurring during the day and stops occurring during the night. Additional information on how the analysts specifically used this information to account for seasonal changes in the sunrise/sunset time, governmentally imposed changes to the time, geography, and other factors to define the intertwilight (dark versus light) periods is necessary to fully evaluate the soundness of this methodology.

The perplexing "fixed factor" in these models is "the officer conducting the stop". It is not reported how (i.e. which variable) was used to describe and/or differentiate between officers. The only insight offered lies in the more detailed description of the methodology appearing on page 61. The analyst reports, "The standard errors were clustered at the officer level to account for unobserved correlations between stops made by the same officers." It is not abundantly clear why this is necessary.

There are potentially other limitations within these models. First, only discretionary stops for traffic violations were included in the analysis. It is not clear why calls for service and other stop reasons were excluded. Second, it appears pedestrian stops were included. It would seem the race or ethnicity of an individual would be much more perceptible during a pedestrian stop. Third, the analysis did not control for traffic violations that can only occur at night (e.g. headlight). Finally, the analysis did not control for urban areas wherein artificial illumination could increase the likelihood that an officer would be able to perceive a resident's race or ethnicity. These issues are discussed in the methodology but do not appear to be controlled in the statistical models.

Search discovery rates

Searches and search discovery rates are often used as a means to measure disparity in police/citizen contacts with respect to various identity-related criteria. In an earlier section of this analysis, the importance of understanding context was discussed. Nowhere is this more important than during an analysis of searches conducted by police officers.

The reason for this is that not all searches are 'created' equally. The primary focus of an analysis of searches, within the context of a racial profiling study, must be on the level of discretion that the officers have in conducting the searches.

Some searches are predicated events occurring within stops that require an officer to conduct a search. Nearly all police departments require officers to conduct searches of the individuals they arrest and searches of the vehicles they impound. There are important officer safety and liability reasons for these requirements. Other searches are predicated on some level of proof of nefarious activity. These include searches predicated by a warrant, when exigent circumstances are present, or when evidence of a crime is present during the stop. It is true that officers do have some level of discretion here. They may, for example, merely ignore the evidence that a crime has been committed. Whether they should is an ethical question. Other searches are allowed when the totality of circumstances (i.e. context) around the contact would lead a reasonable officer to conduct a search for his own protection. These are commonly called Terry stops. The consent search is the only type of search that is totally discretionary. Because of this, the consent search deserves most if not all of the attention when evaluating searches for their potential discriminatory impact.

Unfortunately, it is not actually clear how many, or what percent of individuals, by race or ethnicity are actually subjected to a consent search. Table A-8 All Actions Taken During Stops by Race/Ethnicity on page 17 of the appendices reports that 14,752 (3.0 percent) Black residents were asked for consent to search their person and 12,323 (2.5 percent) Black residents were asked for permission to search their property. It is assumed that there is some overlap in these two numbers. On page 99 of the RIPA Report 2022 the analyst concludes:

> Overall, officers asked 2.7 percent of the individuals they stopped for consent to perform a search. The rate at which officers asked for consent to perform a search ranged from 0.7 percent of stopped individuals perceived to be Middle Eastern/South Asian to 4.1 percent of stopped individuals perceived to be Multiracial.

On the next page (p. 100) the RIPA Report 2022 includes an illustration (Figure 39. Stopped Individuals Asked for Consent to Search by Race/Ethnicity) wherein it would appear that Black and Multiracial individuals represent the largest proportions of individual asked for consent to search. The actual numbers or percentages are not reported. Then the RIPA Report 2022 offers an interesting insight into how these analyses were conducted. Just below Figure 39 on page 100 appears the following.

> The results of this analysis reveal a trend in the 2019 and 2020 RIPA data: Black or Multiracial individuals are asked for consent to search at a higher rate *than those who are perceived to be White*. These disparities reported in the RIPA data are consistent with other data around the country demonstrating racial disparities in consent searches (italics mine).

It is not clear why the analysis (again) compares consent searches for each racial category with those involving White residents. As in previous analysis, it would seem inappropriate to consider searches from any one racial group as a baseline for searches involving other racial groups. This results in an inflation of the alleged racial disparity. Furthermore, and perhaps more importantly, the structure of the data set provides only a limited capacity to measure other contextual factors that are known to produce the desire to conduct a consent search.

On page 39, there is another illustration (Figure 9. Actions Taken During Stops by Race/Ethnicity) wherein it appears Black and Multiracial individuals are searched more frequently than White individuals. Here again, the actual numbers and percentages are not reported. In addition, and perhaps more confounding, is that these particular statistics do not differentiate between types of searches.

These analyses are misleading because it is not abundantly clear which types of searches are being considered and searches involving White residents are used as a baseline, thereby inflating the level of disparity in the search discovery rates in searches involving racial and ethnic minority residents. It would seem simple enough for the RIPA analysts to create a table reflecting:

- The number of all stops by all identity-related criteria,
- The number of stops wherein a consent search was performed by all identity-related criteria,
- The number of consent searches that produced something by identity-related criteria, and
- What types of things were discovered during consent searches by identity-related criteria.

Unfortunately, these data do not exist (reliably) within the RIPA Report 2022, its associated appendices, and only to a limited degree in the actual data set. Without such data and its analysis, it is not possible to determine whether the conclusions offered in the RIPA Report 2022 are justified.

Further confounding this critical analysis, the RIPA analysts redefined searches in numerous ways. The data set contains thirteen mutually exclusive bases for a search. On page 53 of the appendices, they propose a completely artificial reclassification of these bases. The effect of this is that searches incident to an arrest, predicated by a warrant or a vehicle inventory, are classified as "Administrative". The methodology argues all other types of searches (consent, officer safety/safety of others, conditions of parole, suspected weapon, visible contraband, odor of contraband, canine detection, evidence of a crime, exigent circumstances, and suspected violation of school property) are discretionary. In reality, discretion is more nuanced than this. Because discretion exists in a continuum, it should be classified at a higher level of measurement than this nominal distinction. In other words, the reconceptualization of searches into 'administrative' and 'discretionary' categories sufficiently limits the capacity for an honest critical analysis and may misrepresent the extent of the alleged disparity. The footnote (#107) at the bottom of page 54 of the RIPA Report 2022 provides some insight into the RIPA analysts' misunderstanding of searches, policing practices and the ethical standards of policing.

> Administrative searches are not instances where the police officer has no discretion at all, but rather where the officer makes an earlier choice that leads to a search, such as a choice to make an arrest that requires a search. Stops where officers perform administrative searches still possess the potential for bias to affect an interaction, either by the officer at points prior to the search, or at a command level when setting policies and priorities.

Finally, a review of the statistical results for this analysis produces a sense that while the coefficients for the independent variables (race/ethnicity) are statistically significant, their actual impact on the dependent variable (discovery rate) is very small. Furthermore, the Adjusted R^2 values (indicating the amount of variation explained by the model) are quite low, ranging from 0.305 (about 31 percent of the variance) to 0.349 (about 35 percent of the variance). It would seem that the breathless nature of the allegations of racial disparity are overstated.

Use of force analysis

The RIPA Report 2022 contains a series of analyses that attempt to identify disparity in use of force with respect to the identity-related categories. Overall, it appears that of the 2,937,662 stops occurring in 2020, only 32,579 (1.1 percent) included an incident of force (use of force) at some level (lethal, less than lethal, or limited force).

The multivariate statistical models used to estimate the effect of race/ethnicity on the probability of use of force are similar to the models used to conduct the Veil of Darkness analysis. There is, however, a notable difference with respect to the independent variables. These analyses include age and gender as independent variables. These two factors are consistently correlated with physical violence and would likely produce a measurable effect on whether or not a use of force incident occurs during a stop. Unfortunately, the actual influence of these factors is not reported in the appendices. The coefficients relevant to the race or ethnicity of the residents are the only ones reported. The overall results are all statistically significant, meaning only that they are not likely the result of random chance. Curiously, the coefficients associated with the race/ethnicity of the resident are abnormally high for a logistic regression. This coupled with the relatively low Adjusted R^2 (a measure of the amount of variation explained by the model) demands a more complete reporting of these statistical models and their diagnostics. These findings warrant additional scrutiny, particularly with respect to the manner in which the statistical model was constructed. A similar model using the same data does not reveal similar findings.

Unfortunately, the RIPA data set is missing critical information. This omission eliminates our ability to completely understand and evaluate use of force incidents. The missing information would describe more about the context existing before, during, and after each use of force incident. For example, there are fundamental differences between a use of force incident initiated by a police officer and a use of force incident initiated by a resident. An unprovoked physical attack by a police officer on a resident would likely be considered assault. Alternatively, an officer who defends himself from an attack perpetrated by a resident would be considered self-defense. From the RIPA data, we cannot measure whether the police are initiating these physical confrontations or perpetuating them.

<u>Measuring discretion</u>

Police officer discretion is, and has always been, controversial. The amount of discretion allowed for officers ebbs and flows depending on the activities in which they are engaged. Discretion may also be limited by law, regulation, policy, procedure or court ruling. For example, in the 1980's, police officers often ignored domestic violence, preferring instead to define it as a civil matter. More than 30 years ago, state legislatures, city councils and other governing bodies began adopting regulations requiring officers to arrest individuals they suspect are guilty of domestic violence, effectively, and likely appropriately, limiting an officer's discretion.

Measuring officer discretion is a critical factor in any evaluation of police practices, and most especially racial profiling analysis. Within the context of the RIPA Report 2022 there are several missed opportunities for assessing the effects of police officer discretion on stops and stop outcomes. The relative severity or flagrance of the violation observed (reason for the stop), the event(s) that initiated a physical confrontation between a police officer and a resident, the level of discretion for a search, and other contextual variables essential to our understanding (and evaluation) of police officer behavior are ignored.

Consent searches

As discussed previously, consent searches are completely discretionary. In fact, when an officer articulates some level of proof justifying a search, then this justification is always subject to a review by an independent magistrate. It is easier for an officer to merely ask "Do you mind if I search your car?" As a result, consent searches have the highest potential for abuse compared to other types of searches and warrant increased scrutiny. The RIPA Report's 2022 principle finding on consent searches with respect to the race/ethnicity of residents is misleading.

> The 2019 and 2020 RIPA data show that Black and Hispanic/Latine(x) (*sic*) individuals are asked for consent to search at higher rates than White individuals. While Black, Hispanic/Latine(x) (*sic*), and Multiracial individuals were searched at higher rates for consent only searches as compared to all other racial/ethnic groups, these consent-only searches resulted in lower rates of discovery of contraband (8.5%, 11.3%, and 13.0% respectively) than searches of all other racial/ethnic groups (p. 11).

The operative comparison is not against individuals from other racial and ethnic categories. The percent of consent searches by racial and ethnic category should be compared against the proportion of individuals stopped by race or ethnicity. This analysis would reveal whether or not disparity changes during a stop with respect to the race or ethnicity of the individual stopped. An additional comparison measuring the level of punitiveness in the results of the stop (arrest versus citation) would further evaluate disparity.

The entire discussion on consent searches in the RIPA Report 2022 misses the point about searches and their justification. Searches are never justified with respect to what they produce. Searches are justified (legally) on the basis of whether or not the police officer adhered to the restrictions of the 4th Amendment to the U.S. Constitution, as defined by legal decision.

Discretionary stops versus calls for service

There appears to be an inconsistency in the RIPA data set with respect to meanings of discretionary and call for service stops. This issue affects the reporting of stops as well as their analyses. The RIPA Report 2022 explains this on page 32 in footnotes #69 and #70, as follows.

> An interaction that occurs when an officer responds to a call for service is only reported if it meets the definition of a "stop" as set forth in section 999.224, subdivision (a)(14) of the RIPA regulations. A call for service is not a reason for stop value under the RIPA regulations. Rather, officers indicate whether or not a stop was made in response to a call for service in addition

to providing a primary reason for stop. The RIPA regulations do not specify whether a stop made after a civilian flags down an officer on the street fits the definition of a call for service; accordingly, data entry for this field may vary across officers and agencies for stops where civilians flagged down officers.

Given that stops for traffic violations constitute a majority of the data, but are less likely to be made in response to a call for service, these analyses were also conducted while excluding data from stops where officers indicated that the primary reason for the stop was a traffic violation.

As a general rule, a police officer has substantially less discretion when dispatched to a call for service than when initiating a stop. The effect of this can be profound. For example, Appendix A.4 Stops by Identity Group and Calls for Service differentiates between officer-initiated (i.e. discretionary) stops and call for service (i.e. less discretionary) stops. A closer analysis of this table indicates that 94.1 percent of all stops are officer-initiated and 5.9 percent of all stops are call for service stops. There are some differences in the percentages of stops with respect to the race or ethnicity of the resident depending on whether the stop is officer-initiated or a call for service. For example, overall Black residents represent:

- 16.5 percent of all stops,
- 15.8 percent of officer-initiated stops, and
- 27.2 percent of all call for service stops.

Could this indicate that Black residents are more likely to seek assistance from the police? Or could it mean that police officers are dispatched to calls that result in a higher proportion of interactions involving Black residents? Interestingly, this is the only racial/ethnic category in which this proportional increase occurs. Alternatively, Hispanic residents represent:

- 40.4 percent of all stops,
- 40.8 percent of officer-initiated stops, and
- 34.4 percent of all call for service stops.

Could this mean that Hispanics are less likely to seek assistance from the police? Or could it mean that police officers are dispatched to stops that result in a lower proportion of stops involving Hispanics? Answering these questions is beyond the scope of this critical analysis. They are offered here to demonstrate the importance that officer discretion may have on officer performance and motivation. At the very least, the RIPA Board should be consistent in its analysis of discretion.

Supervision Stops and Searches

Supervision stops and searches deserve special attention. The population of individuals subjected to a search because of their supervision status (e.g. parole, probation, etc.) is not demographically equivalent to the overall population of residents in California. It is likely this population has a much higher proportion of Black or Hispanic males than the overall population. This difference should be factored into the analysis. In fact, because this practice is so narrowly defined it deserves additional scrutiny. Whether or not these individuals should relinquish their 4th Amendment protections by virtue of their status is a policy question that is beyond the scope of this critical analysis.

<u>The disparate impact experienced by transgendered and disabled residents</u>

The RIPA Report 2022 finds some evidence of an increased disparate impact on transgendered and disabled residents. This type of analysis is relatively new to the racial profiling research agenda. The number of individuals within these identity-related categories is relatively small and this adversely affects the availability of meaningful analytical strategies, or at least the use of higher-level statistical models. The most compelling descriptions of potential disparity come from anecdotal evidence. Interestingly, this is how the racial profiling research agenda began in the early 1990's.

It is likely mental illness, as one category of disability, may impose itself on these analyses substantially. The police response to mental illness is challenging because there are not a lot of options available to officers. It is also likely that a case study approach would produce more meaningful results into this particular potential for disparity than multivariate models or descriptive comparisons.

Conclusions and Recommendations

It is simply not possible (legally or scientifically) to allege racial profiling (as defined by California law) using the data available to the RIPA Board. The only measure of a resident's identity-related criteria happens *after* the stop is initiated. In order to allege individuals are stopped on the basis of their identity-related criteria, what the officer perceives the resident to be *prior* to the stop must be measured, at a minimum.

Related to this is an inappropriate use and interpretation of classic probability. Because individuals are not selected for stops randomly and are subjected to differential levels of routine police supervision, it is not possible to confirm random selection, which is a critical assumption for this statistical technique. As a result, it is not possible to calculate how any identity-related criteria either increases or decreases the probability of a stop, any event occurring during a stop, or any outcome of a stop.

A consistent allegation throughout the history of the racial profiling is that marginalized persons are stopped for less severe violations. Anecdotal evidence tends to support this allegation. Unfortunately, the quantitative data we use to assess this are often lacking any measure of the relative severity or level of flagrance for the alleged violation. The data set used by the RIPA Board is no exception. Closely associated with this aspect of the racial profiling is the notion that reported reasons for a stop may be merely a pretext. It is actually impossible to know this definitively. Throughout the RIPA Report 2022, there are several references to pretext stops. Most of these references suggest that stops for minor violations are more likely to be pretexts. There is really no way to measure this. What may seem a minor violation actually may be important. The enforcement of vehicle registration and license plate laws are important for ensuring sufficient tax revenues are collected and for identifying wanted individuals and stolen vehicles. Likely the best way to limit stops based on registration violations would be to ask the California legislature to remove these violations from the traffic code.

The analyses of stops within the RIPA Report 2022 sometime include stops from the California Highway Patrol (57.7 percent of all stops) and sometimes they do not. The RIPA Report 2022 indicates that data relating to gender identity was missing from the stops reported by the CHP. This is a legitimate reason for excluding CHP stops from these analyses. What is not clear is why these stops were excluded from other analyses, seemingly without legitimate justification. In some cases, the exclusion of CHP stops is peripherally reported in a footnote. Similar alternating and covert patterns exist in the analyses of discretionary/call for service stops and for searches. The portions of the data set that are excluded from an analysis should be reported more substantially and justified more completely.

Perhaps the most disappointing analytical strategy is the RIPA Board's use of stops involving White residents as a baseline for calculating evidence of disparity in stops involving all other racial and ethnic groups. This practice has no analytical or methodological basis. Furthermore, this practice amplifies the level of disparity to the point of misrepresentation. There is nothing in the research literature that would support the notion that stops involving White residents should be a baseline or standard upon which stops of residents from other racial and ethnic groups should be compared. This practice assumes that White residents cannot be victimized by racial profiling. This issue is seldom considered in the academic research, but there is some evidence that White residents may receive additional attention (from police officers) when they are observed in places principally populated by Black or Hispanic individuals (Sansone-Braff, 2014, Withrow, 2002b).

The RIPA Board's analysis of searches and search discovery rates are particularly confounding. The analysts use different portions of the data set and create artificial categories of searches. These artificial categories are not consistent with routine police systems and practices or with professional standards and ethical behavior. This, however, is only one example of a more fundamental problem throughout various parts of the analysis. Variables that could be used to explain the justification for officer behavior or to understand more about the contexts of the stops are not included in the data set. A notable omission is a lack of measurement for the direction of conflict in use of force incidents. The RIPA data cannot measure whether an officer's use of force is initiated by the officer or a reaction to force initiated by a resident.

Parsimony is a desirable quality in research methodology and statistical analysis. Generally, it means simpler explanations are preferable. This necessarily extends into data collection protocols and procedures. Unfortunately, the data collection strategies adopted by the RIPA Board are far from parsimonious. For example, race and ethnicity are collected in two distinct formats. This requires the analyst to create additional categories to account for cases wherein officers report multiple racial or ethnic categories. The potential for error in this practice is substantial. Perhaps more importantly, the potential for reactivity is even more threatening. Reactivity occurs when research subjects change behavior when they perceive they are being observed. Police officers completing RIPA reports know they are being observed. Although there is no evidence of this, it is possible for an errant police officer to report race or ethnicity so as to misrepresent a resident's actual race or ethnicity into a broader category defined as Multiracial. The same issues occur when seemingly similar events and outcomes are included in separate variables. "No action taken" exists as both a stop event and as a result of a stop. Perhaps the reporting rules account for this, but the measurement of the same behavior within two separate variables confounds the analysis, as well as its critical analysis.

The appendices associated with the RIPA Report 2022 need substantial correction. Row totals are incorrect and column totals are missing. This is much more than an annoyance or suggestive of a manuscript form violation. The result of these mistakes produces misinformation. Similarly, the tables reporting statistical results (regression models) are incomplete. They do not contain coefficients for other independent or control variables, coefficients for constants, and diagnostic measures that should be used to completely evaluate the soundness of these statistical models. These tables and their methodological descriptions do not contain information on how variables, including the dependent variable, are coded. This information is essential for understanding how the signs of the coefficients for the independent variables specifically contribute to the outcome of the dependent variable. Based on this author's extensive experience as a researcher, it is unlikely incomplete tables like this would survive the rigors of peer review.

It would appear that the reporting of stop data under the regulations adopted by the RIPA Board would be onerous. The amount of data collected per stop is substantially greater than other similar data sets. The average time necessary to complete one report is not known. We do know that in 2020, there were 2,937,662 stop reports submitted. If each report required only one minute to complete, then collectively this represents 2,937,662 minutes, 48,961 hours (2,937,667 minutes/60 minutes per hour), and 23.5 person years (48,961 hours/2,080 hours worked per year). If each report required ten minutes to complete, the amount of personal time required for this endeavor would be equivalent to 235 officers. Given current staff shortages experienced by most police departments, the time commitment associated with completing these reports is potentially substantial.

Perhaps the most important recommendation is to consider revising the reporting criteria, eliminating some variables and adding other variables, and reconsidering the types (and consistency) of the analyses. The objectives of these revisions would be to produce a data set that more completely measures the contextual features of each stop and provides more insight into the factors that might explain (or justify) a police officer's decisions. These sorts of data sets inform policy makers and provide more justification for change. Overall, the goal is to provide a fair and objective analysis of what is happening on the street. This is fundamentally different than using data to confirm a bias.

###

Bibliography

Alpert, G.P., Smith, M.R. and Dunham, R.G. 2003 (March). Toward a Better Benchmark: Assessing the Utility of Not-at-Fault Traffic Crash Data in Racial Profiling Research. Paper presented at Confronting Racial Profiling in the 21st Century: Implications for Racial Justice. Boston, MA: Northeastern University.

Black, D. J. 1971. The social organization of arrest. Stanford Law Review 23:1050-111.

Black, D.J. & Reiss, A.J., Jr. 1970. Police control of juveniles. *American Sociological Review* 35(1): 63-77.

Economic Policy Institute (2020). Not everybody can work from home. Black and Hispanic workers are much less likely to be able to telework. Accessed from the World Wide Web on July 20, 2022 from https://www.epi.org/blog/black-and-hispanic-workers-are-much-less-likely-to-be-able-to-work-from home/#:~:text=Asian%20workers%20are%20the%20most%20likely%20to%20be,Hispanic%20workers%20are%20able%20to%20work%20from%20home

Engel, R. S., Calnon, J. M. (2004). Comparing benchmark methodologies for police–citizen contacts: Traffic stop data collection for the Pennsylvania State Police. Police Quarterly, 7(1), 97–125.

Grogger, J. and Ridgeway, G. (2006). Testing for the racial profiling in traffic stops from behind a veil of darkness. Journal of the American Statistical Association. 101(475), 878 – 887.

Harrell, E. and Davis, E. (2020). Contacts between police and the public, 2018, Statistical tables. US Department of Justice, Bureau of Justice Statistics. Washington, DC

Horrace, W. C., Rohlin, S. M. (2016). How dark is dark? Bright lights, big city, racial profiling. Review of Economics and Statistics, 98(2), 226–232.

LaFave, W. (1965). Arrest: The decision to take a suspect into custody. Little, Brown.

Lamberth, J. 1994. Revised Statistical Analysis of the Incidence of Police Stops and Arrests of Blak Drivers/Travelers on the New Jersey Turnpike Between Exits or Interchanges 1 and 3 from years 1988 through 1999. West Chester, PA: Author. Retrieved July 7, 1999, from www.lamberthconsulting.com/research_articles/asp.

Lange, J.E., Blackman, K.O., and Johnson, M.B. (2001). Speed Violation Survey of the New Jersey Turnpike. Final Report. Calverton, MD: Public Services Research Institute.

Novak, K.J. (2004). Disparity and racial profiling in traffic enforcement. Police Quarterly 7(1): 65-96.

Oakland Police Department (2004). *Promoting cooperative strategies to reduce racial profiling: A technical guide*. Accessed from the World Wide Web on July 15, 2022. https://www.aclunc.org/sites/default/files/040824-oaklandprofilingstudy.pdf

Police Foundation (2003). A Multi-Jurisdictional Assessment of Traffic Enforcement and Data Collection in Kansas. Washington, DC: Author.

Ramirez, D., McDevitt, J. and Farrell, A. (2000). A resource guide on racial profiling data collection systems: Promising practices and lessons learned. US Department of Justice, Bureau of Justice Assistance, Washington DC.

Reiss, A. (1971). The police and the public. Yale University Press.

Ridgeway, G. (2009). Cincinnati Police Department traffic stops: Applying RAND's framework to analyze racial disparities. RAND.

Rojek, J., Rosenfeld, R. and Decker, S. (2004). The influence of drivers' race on traffic stops in Missouri. Police Quarterly 7(1), 126-47.

Sansone-Braff, C. (2014). Racial profiling: It happens to White people, too. Patch Media. Accessed from the World Wide Web on July 2, 2022 from https://patch.com/new-york/easthampton/racial-profiling-it-happens-white-people-too.

Smith, M.R. and Petrocelli, M. (2001). Racial profiling? A multivariate analysis of police traffic stop data. Police Quarterly 4: 4-27.

Smith, W. R., Tomaskovic-Devey, D., Zingraff, M. T., Mason, H. M., Warren, P. Y., Wright, C. P. (2004). *The North Carolina highway traffic study*. Final report to the National Institute of Justice, U.S. Department of Justice.

Smith, M. R., Rojek, J. J., Petrocelli, M., Withrow, B. (2017). Measuring disparities in police activities: A state of the art review. Policing: An International Journal of Police Strategies & Management, 40(2), 166–183.

Stacey, M. and Bonner H.S. (2020). Veil of darkness and investigating disproportionate treatment in policing: When researchers disagree. Police Quarterly, 24(1), 55-73.

Taniguchi, T. A., Hendrix, J. A., Levin-Rector, A., Aagaard, B. P., Strom, K. J., Zimmer, S. A. (2017). Extending the veil of darkness approach: An examination of racial disproportionality in traffic stops in Durham, NC. Police Quarterly, 20(4), 420–448.

US v. Alcaraz-Arellano, 441 F. 3d 1252 – Court of Appeals, 10th Circuit 2006.

US Bureau of Labor Statistics (2020). Ability to work from home: Evidence from two surveys and implications for the labor market in the COVID-19 pandemic. Accessed from the World Wide Web on July 20, 2020 from https://www.bls.gov/opub/mlr/2020/article/ability-to-work-from-home.htm.

Walker, S., Alpert, G.P. and Kenney, D.J. (2001). Early warning systems: Responding to the problem police officer. National Institute of Justice, Research in Brief. US Department of Justice, Washington DC.

Washington State Patrol. 2001. Report to the Legislature on Routine Traffic Stop Data. Olympia: Washington State Patrol.

Withrow, B.L. 2002a. The Wichita Stop Study. Wichita, KS: Wichita State University, Midwest Criminal Justice Institute.

Withrow, B.L. (2002b). The Wichita Stops Study: A Beat Level Analysis, Midwest Criminal Justice Institute.

Withrow, B.L. (2016). Research methods in crime and justice, Second edition. Routledge/Taylor & Frances. Upper Saddle River, NJ.

Withrow, B., Dailey, J.D., and Jackson, H. (2008). Utility of an internal benchmarking strategy in racial profiling surveillance. Justice Research and Policy 2, 19-47

Worden, R. E., McLean, S. J., Wheeler, A. P. (2012). Testing for racial profiling with the veil of-darkness method. Police Quarterly, 15(1), 92–111.