PROMOTING COOPERATIVE STRATEGIES TO REDUCE RACIAL PROFILING

A TECHNICAL GUIDE

OAKLAND POLICE DEPARTMENT

April 20, 2004

RICHARD L. WORD
CHIEF OF POLICE

1 INTRODUCTION

At the heart of effective law enforcement is community trust and confidence. Racial profiling, whether real or perceived, deteriorates the public's trust and confidence in the police and strains police and community relations, especially within communities of color. Attention on the issue of racial profiling has only increased in the national spotlight following the terrorist attacks on September 11th and continues to stimulate intense debate with regard to race and the criminal justice system.

In the mid-1990s the discussion on racial profiling focused primarily on its existence. At that time, most minorities believed racial profiling existed while many non-minorities believed it was only a perception of minorities. For many, these perceptions changed in 1999 when a white state trooper stopped and subsequently shot four unarmed black men driving on the New Jersey turnpike. This incident and high profile racial profiling studies of the I-95 in Maryland (Lamberth, 1996) and of stop and frisk practices in New York City (New York Attorney General, 1999), which revealed significant disparities in stops of minorities, transformed racial profiling from a minority-community perception to a national reality. How we end racial profiling is now the largest challenge we face and where most of the disagreement exists.

In response to racial profiling, many law enforcement agencies in the United States have implemented some form of traffic stop-data collection. Fourteen states have passed racial profiling legislation that require law enforcement agencies adopt racial profiling policies, provide some form of anti-racial profiling training, and implement data collection and analysis programs. California passed similar legislation in 2001, however, this law does not require police departments to collect stop-data.

On the federal level, United States Congressman John Conyers introduced the End Racial Profiling Act of 2001, which would mandate data collection for all law enforcement agencies receiving federal funds. The pending legislation will be reintroduced in early 2004.

Many people believe data collection is necessary to end racial profiling. Others believe data collection offers no practical value and simply validates what is already known. The debate surrounding stop-data collection remains extremely controversial and many questions remain unanswered: Is data collection a practical and critical step necessary to end racial profiling or is it merely symbolic, a necessary step to appease minority communities in hope of instilling public trust?

On one hand, data collection may prove to be practical. Proper data collection coupled with proper analysis utilizing credible benchmarks not only provides an organizational "snap shot," a look at the organization at a specific point in time, but it also assists administrators in identifying institutional and systemic problems. Data collection also serves as a gesture of openness to the community and a commitment to equality. It represents the willingness of law enforcement to take an introspective look to prevent disparate treatment. It also demonstrates law enforcement's true commitment to responding to community needs and concerns.

On the other hand, the absence of appropriate methods to analyze the data and establish credible benchmarks incorporating the complexities of policing has resulted in error prone analyses and fueled negative perceptions in many communities. When it comes to data collection and

analysis, the police and the community are quite often in direct opposition. Many in the community believe that the data will be able to determine conclusively whether officers engage in racial profiling. Conversely, many law enforcement officers will reject data collection outright and challenge its credibility.

It is against this backdrop and amidst this debate that the Oakland Police Department decided to participate in the COPS Promoting Cooperative Strategies to Reduce Racial Profiling Program and select the ever-challenging data collection and analysis strategy.

As one of only a few agencies in California to voluntarily collect traffic stop data, we believe data collection will prove beneficial in several ways. It will: 1) identify whether our operational practices are resulting in racial and ethnic disparities, 2) provide management a tool to discern whether stop disparities are societal-based or a result of police efforts, 3) serve as an additional performance measure to assess operational effectiveness and cost/benefit tradeoffs, 4) assist agencies in developing strategies to reduce disparate enforcement and improve police-community relations, and 5) enhance public trust and confidence in the Department's ability to establish accountability.

1.1 BACKGROUND

In response to community concerns regarding racial profiling, the Oakland Police Department began voluntarily collecting traffic stop-data in March 2000. At that time, the Department convened a racial profiling task force, which involved stakeholders in the community, to identify what data should be collected and develop data collection methodology. Over the subsequent nine months the Department captured over 22,000 stops. Due to a lack of funds, however, the Department was unable to partner with a research team to conduct a comprehensive analysis of the collected data. Consequently, the data were deemed inconclusive and the Department was unable to make any use of it.

In 2001, the Department received a $200,000 grant from the United States Department of Justice, Community Oriented Policing Services (COPS) to further our efforts and address this critical issue. As a part of the COPS project, the Department reconvened its racial profiling task force – making some changes to its composition – and contracted with the RAND Corporation as its research partner. The overarching goals of the new group were to:

- Initiate an effective data collection program,
- Establishing baseline comparison data,
- Establishing credible benchmarks that incorporate relevant local variables, and
- Develop a comprehensive process to analyze the data to be meaningful to the agency and the community.

1.2 PROJECT OVERVIEW

The purpose of this report is to provide police departments, policymakers, and researchers with a summary of Oakland's racial profiling project. The report outlines our project goals and objectives and highlights the Department's efforts and accomplishments, as well as our shortcomings and the many lessons learned. Oakland's new racial profiling policy is one of the successful products of the project's efforts, produced in collaboration with all of the members of the project. We describe in the development of this policy to serve as a guide for other departments needing to craft a policy of their own. This report also includes a comprehensive analysis of Oakland's vehicle stop data. From research questions formulated during task force meetings, the RAND team developed methods for addressing these questions culminating in the analysis presented in Section 9 of this report. For police agencies considering data collection or exploring methods for analyzing their vehicle stop data, that section can serve as an example of some available methods and the type of findings a racial profiling study might produce. This report also provides in-depth technical assistance with regard to data collection technology. Section 8 describes the technology used in Oakland for data collection with the assistance of SCANTRON Corporation.

Accordingly, the report will briefly discuss our activities and accomplishments, and provide recommendations with regard to each of the identified ten objectives of the project, which are as follows:

1) Form a local racial profiling task force involving all stakeholders;

2) Assess community perceptions on racial profiling, data collection and analysis;

3) Identify what data should be collected;

4) Develop data collection methodology;

5) Identify process to determine baseline comparison data;

6) Identify local-based variables in establishing credible benchmarks;

7) Develop methods for analyzing the vehicle stop data and complete an analysis of Oakland's data;

8) Define how data will be useful to the agency;

9) Develop marketing strategy to garner public support and instill community trust; and

10) Identify to what extent the data collection and analysis program influenced public perception.

2 FORMING A LOCAL RACIAL PROFILING TASK FORCE

2.1 INVOLVING STAKEHOLDERS

The Department believes that the most effective way to develop a comprehensive stop-data

collection and analysis program is to involve a cross-section of stakeholders. The challenge however, is deciding which stakeholders should be involved.

With a city as diverse as Oakland, there are many representative organizations and interest groups. While we wanted to ensure the largest cross-section possible, we also recognized that too large a group would be impossible to manage, and the discussions and debates that would inevitably take place – and needed to resolve problems – would be impossible to facilitate. We therefore decided to limit the size of the task force to no more than 15 participants. In order to compensate for such a small representative group, we decided to hold public venues to seek input and solicit feedback from those groups not involved on the task force. Community input and feedback will be discussed in more detail in Section 4 of this report.

The Department was still faced with a tough question: How do you decide what groups are key stakeholder groups? In making this decision, the Department considered the following factors:

- Racial and ethnic make-up of the City of Oakland. In other words, to the extent possible, we wanted to make sure that the task force was as diverse as our city. Although the task force did not perfectly match the demographics of the City, the group was extremely representative and diverse: there were 5 men and 6 women; 4 whites, 5 blacks, one Asian, and one Hispanic.

- Constituency of the interested group. We tried to select representative groups that were established and served large constituencies. This would prove necessary to receive input from the community, and to effectively market our efforts and promote our successes.

- Prior work in the community. The best predictor of future behavior is past performance. We recognize that the project required extensive work and time commitment. We wanted to ensure that the representative groups had been successful in prior projects, and that they were willing to commit the time and resources necessary for the project.

- Ability to be both fair and objective. This area was probably the most important. Because of the nature of racial profiling, many people of predisposed and lack objectivity. For the project to be successful, representatives must come with an open mind, ready to learn new ideas and methods. We wanted to prevent the task force from being used as a forum to air grievances.

- National exposure. Although we recruited local organizations, we looked for those organizations that were affiliated in some way with a national organization or involved in national projects. We felt this was necessary considering the importance of this grant to the industry.

Based on these above factors, the Department recruited the following organizations:

- Oakland Public Safety Committee;

- National Association for the Advancement of Colored People (NAACP);
- American Civil Liberties Union (ACLU);
- PUEBLO;
- PolicyLink, Inc.;
- Oakland Citizen Police Review Board;
- Oakland Police Officer Association;
- RAND Corporation; and
- SCANTRON Corporation

Captain Ronald Davis, the Department's racial profiling program manager, chaired the task force. Below is list of the names and positions of the representatives for each of the groups participating in the racial profiling task force. Also provided is a brief description of their organization.

Sarah Chavez, Policy Analyst
Office of Oakland City Councilmember Larry Reid, District 7
Councilmember Reid is the chair of the City's Public Safety Committee and the Vice Mayor of the City of Oakland.

Christopher Swartz-Edmisten
SCANTRON Corporation
SCANTRON Corp. is headquartered in Irvine, Calif., and is a wholly owned subsidiary of Atlanta-based John H. Harland Company (NYSE:JH). Scantron is the acknowledged leader in data collection systems, testing, and assessment and hardware service and repair.

Rashida Grinage
People United for a Better Oakland (PUEBLO)
PUEBLO is a grassroots, non-profit organization which is multi-ethnic, multi-generational, and multi-issue. It's an advocacy organization that campaigns for issues that affect primarily low-income and minority members of the Oakland community. PUEBLO has organized around environmental, health, educational, and criminal justice issues since 1989.

Maya Harris-West, Esq.
Senior Associate, PolicyLink
PolicyLink is a national nonprofit research, communications, capacity building, and advocacy organization, dedicated to advancing policies to achieve economic and social equity based on the wisdom, voice, and experience of local constituencies. Ms. Harris has since left PolicyLink to be the Director of the Racial Justice Project at the American Civil Liberties Union (ACLU) of Northern California.

Jeff Hassna, Executive Board Member
Oakland Police Officer Association
The Oakland Police Officers' Association is the recognized collective bargaining unit for over 700 sworn police officers in the City of Oakland.

Joyce Hicks, Executive Director
Wendy Jan, Senior Policy Analyst
Oakland Citizen Police Review Board
The Citizens' Police Review Board (CPRB) is an advisory board that provides citizen oversight of conduct by sworn police officers and park rangers. The Board consists of nine members and three alternates who are appointed by the Mayor and confirmed by the City Council. One Board member and one alternate must be under twenty-five years old.

Greg Ridgeway, Ph.D., Statistician
RAND Corporation, Public Safety and Justice Unit
RAND conducts research and provides analysis to address challenges that face the United States and the world. Today, RAND emphasizes several areas of research that reflect the changing nature of a global society. Much of this research is carried out on behalf of public and private sponsors and clients.

Shonda Scott, Board Member
Oakland Chapter, National Association for the Advancement of Colored People (NAACP)
The NAACP works at the national, regional, and local level to secure civil rights through advocacy for supportive legislation and by the implementation of our Strategic Initiatives. The NAACP also stands poised to defend civil rights wherever and whenever they are threatened. The Oakland Chapter of the NAACP represents over 14,000 members.

Mark Schlosberg, Esq.
American Civil Liberties Union (ACLU) of Northern California
The ACLU of Northern California, based in San Francisco, is the affiliate that works to protect civil liberties in this region of the country. The ACLU-NC was founded in 1934 during the General Strike to fight against police abuse of striking longshoremen. Today, the ACLU is at the forefront of every civil liberties battle in the state—from the rights of immigrants and reproductive rights to abolition of the death penalty and the rights of lesbians and gay men.

2.2 IDENTIFYING GOALS AND OBJECTIVES

After forming the task force, the next challenge we faced was to identify its specific goals and objectives. As with most projects, the key to success is identifying attainable goals and objects at the outset. In making this decision, the Department took into account current industry challenges and practices with regard to data collection and analysis, and lessons learned from our earlier data collection efforts.

Based on these factors, the task force identified the following overarching goals:

1) Identify what data to collect;

2) Develop data collection processes;

3) Identify local variables that may skew aggregate data;

4) Identify local relevant data fields;

5) Develop data analysis model - incorporating all variables and perspectives;

6) Identify what the data means to local police and community; and

7) Identify how data will be used.

Throughout the year the task force met monthly, in many cases several times monthly, to discuss all pertinent issues and develop solutions. In addition, the task force conducted one town hall meeting at a local high school and received valuable input and feedback from the community.

2.3 TRAINING OF TASK FORCE MEMBERS

Because the issue of racial profiling is so emotionally charged, we felt it necessary to provided both informal and formal training to task force members. The training consisted of ride-alongs with Oakland police officers and training courses. We implemented this training to ensure the task force made sound decisions based on experience, expertise, and objective factors, not emotions. When forming a racial profiling task force, departments should not assume that task force members understand racial profiling simply because they are representatives of certain organizations. By providing training to the group, the members gained expertise on the topic and came to understand and respect the varied perspectives that exist, even if they did not agree with them.

Task force members attended formal training courses that the United States Department of Justice, Washington State University, Northwestern University, and Simon Frazier University in Canada developed. Upon their return, the task force members that attended these training courses provided the group an overview of the course and identified how our efforts compared to what they had learned, and then facilitated discussion as to whether the group should make changes in our program or stay the course.

To further the group's understanding of the issue, the task force also conducted a Bay Area law enforcement workshop on racial profiling. Police agencies throughout the Bay Area attended and shared their experiences. This was extremely beneficial in that: 1) the Department was able to provide assistance to agencies in the county that did not have the benefit of our experiences; 2) the Department was able to identify whether certain benchmarking variables were specific to Oakland or extended throughout the county, and 3) task force members were provided the opportunity to interact with other law enforcement agencies and see first hand the shared challenges faced in dealing with this issue.

There was great benefit in having task force members receive both informal and formal training. Members of the task force obtained an expertise in racial profiling, which increased the effectiveness and efficiency of the group, and reduced the level of personal agendas and biases.

2.4 FACILITATING TASK FORCE MEETINGS

Once the formation of the task force is complete, the project manager must effectively facilitate the meetings so that members keep their interest and the group remains on task with its goals and objectives.

The key to facilitating our task force meetings was preparation and planning. Task force members must be given ample notice of meetings and to the extent possible, reminders of the meetings must be provided several days before. This was necessary because most of the representatives are extremely busy individuals. Not only were they representatives of their organizations, in most cases they also had full-time employment and family commitments.

Because time is valuable, program managers must ensure meetings are efficient. Initially, the Department fell woefully short in this area. Meeting agendas were over ambitious with too many items to resolve. Consequently, meetings occasionally went over two hours. As this occurred, participation began to taper off. To reenergize the group the Department began setting meetings with single topic agendas, and altered the times of the meeting to include lunch and dinner meetings.

Another critical aspect of facilitation is the facilitator's ability to listen versus talking. This too was an area the Department fell short in the beginning. Consequently, the Department began to control and monopolize the meeting, which once again had a negative impact on the meeting. As we moved further into the program we adjusted our facilitation style and let others lead the meeting and debate. This proved extremely beneficial and resulted in the Department better understanding the perspectives and concerns of the group.

The most important aspects of facilitation are openness and candor. Because there are apprehensions on all sides, it is necessary for the facilitator to be open to new ideas, and to be candid about what the Department is and is not willing to do or compromise. The worst thing that can happen is for the group to think the department has already made its decision and is simply looking for a stamp of validation.

Facilitation was by far our strongest area. From the beginning, the Department identified the few areas of policy and practice that were non-negotiable. Even then, we opened those areas for discussion and advised the groups that their opinions, even if not adopted, would be represented in our report.

We also opened the Department to the task force by providing task force members with sensitive information about data results and internal shortcomings. In other words, we demonstrated our trust in the group, who in return demonstrated their trust and confidence to the Department. The task force agreed to not discuss data or findings for the duration of the project and, indeed, no data was leaked or released to the press throughout the entire project. By being forthright, open, and candid, the Department was able to establish trust with the task force members and form a true partnership upon which we were able to build consensus upon, and agree to disagree in those few instances in which there was not consensus.

2.5 BUILDING CONSENSUS

Although the Department chose the data collection and analysis strategy, the process used

extended well beyond that goal. One of our greatest achievements was the ability to build consensus on most, if not all, parts of the project. This in and of itself has made the project worthwhile.

After lengthy debates, and in some cases heated arguments, the task force came to unanimous consensus on the following issues:

1) Identifying what data to collect;
2) Developing the data collection form;
3) Selection of the data collection methodology;
4) Defining Racial Profiling;
5) Racial Profiling Policy Development;
6) Identifying variables to consider in data analysis;
7) Town Hall meeting agenda; and
8) Developing benchmark process

3 RACIAL PROFILING POLICY DEVELOPMENT

A major part of the success of the Oakland task force was the unanimous consensus on a racial profiling policy. The key to our success was the process used to develop the policy. The responsibility to facilitate this portion of the project was delegated to Maya Harris of PolicyLink, Inc. For a police department to enlist a civilian volunteer to develop a policy, especially one such as racial profiling, is extremely rare. However, this unorthodox approach was in fact the key to our success. Ms. Harris and PolicyLink possessed an expertise in public policy that the Department did not. PolicyLink also involved its staff members and researchers to conduct a "best practices" search that the Department lack the capacity to conduct.

Members of the task force were able to review policies from other agencies across the country and discuss their strengths and weaknesses. Ms. Harris also met with task force members and conducted focus groups with rank and file officers in the Department and community members to ensure all perspectives were considered. As a result, all stakeholders embraced the policy. With this being said, however, there were intense debates surrounding the definition of racial profiling and consent searchers. Nevertheless, the task force was able to come to consensus. In fact, the debate served to educate all stakeholders and increase their understanding and respect for their colleagues' perspectives.

The task force believed that the critical pieces of a racial profiling policy should include a clear definition and prohibition of racial profiling, guidelines for post stop activity, and responsibility for department members to report racial profiling. It is our belief that the Oakland Police Department Racial Profiling policy is one of the most comprehensive policies in the country. At the time we

were developing the policy California Penal Code Section 13519.4(e) already prohibited law enforcement officers to utilize racial profiling. Besides reinforcing the content of the California law and Fourth Amendment protections, the policy includes additional guidelines for Oakland police officers. These include

- Officers must complete a stop-data collection form for every stop they make including stops of vehicles, bicycles, and pedestrians. Forms are due at the end of each shift.

- Consent searches cannot be arbitrary and the officer must complete a field contact form articulating the reason for the suspicion. Officers must also advise individuals that they have a right to refuse the search.

- Supervisors shall regularly monitor officers under their command to ensure compliance with the racial profiling policy. Supervisors will review all stop data forms that their officers submit for accuracy and completeness. Regular audits will ensure that officers document all stops.

- Biannual reports to the chief regarding data collection and an analysis of the data.

Appendix 1 of this report contains the complete text of Oakland's racial profiling policy.

4 ASSESSING COMMUNITY PERCEPTIONS

As part of our overall strategy, the Department was determined to assess whether the process of data collection contributed to community perception, and identify what impact the process had on officer perceptions as well. To accomplish this, the Department, in collaboration with RAND, developed and implemented a community survey instrument, which was administered to both the community and Department (See Appendix 2 and 3).

The survey attempted to examine how the process of data collection may contribute to community perception as it relates to racial profiling. One of the major ways to understand citizens' perception of police services is through the use of a community opinion survey. Using data collected from these surveys, the results will measure perception of the community's police interactions about such issues as racial profiling, response time, capability, ethics, police performance, and overall police/community relationships. This vital information and feedback will reflect community attitudes toward police and their services to gain a better understanding of the police-community relations in the Oakland community

This survey further reflects the impact of the community's perception of policing by analyzing citizens' perception of how they are treated in different Oakland neighborhoods. The survey revealed both positive and negative findings about the community's perception.

Based on the results of the survey, the Department has more to concern themselves with as far as public perception is concerned. This survey is an important tool to gage the perception of the Oakland community. One question asked constituents how responsive they felt the Police Department is for requests of assistance and complaints. Of those that indicated they had an

opportunity to contact/report the incident, most surveyed said the police had been very responsive to requests for assistance.

4.1 RESEARCH VARIABLES

A number of research variables were used in this study to examine citizens' perception of racial profiling and interactions with the police. The variables included (but were not limited to) age, income, education levels, community demographic variables and employment status as perception variables. Most of the performance variables are measured in ordinary and nominal levels. The percentage distributions of respondents' characteristics are presented in another document that contains the raw data (See Appendix 3).

4.2 ANALYSIS AND RESULTS

The citizens' perception of Oakland Police Department is reflected through percentage comparisons by various respondent characteristics. Data results show that over half the citizens surveyed feel that crime has stayed the same in their neighborhood, with a very low percentage reporting a decrease. Furthermore, most reported that they feel very unsafe to somewhat unsafe while walking alone in their neighborhoods and that crime is one of their major concerns. Results also showed that the majority of non-White citizens polled felt that they were more likely to be stopped and or harassed by the police.

Additionally, of the entire group polled, most (over 50%) felt that by definition racial profiling is a problem in the City of Oakland.

The survey also reflects positive perceptions as revealed in the following areas: a considerable number positive perceptions of policing were evidenced in the analysis of citizens' perception of keeping order in their neighborhood, courteous treatment, ethical and friendly officers, and never having been mistreated by an officer. Again, the positive perceptions are shared among and across the demographics.

The community perception of police performance was reflected as follows: the survey reflected that over 35% of citizens had little confidence in the police to treat people of different races equally. While there was nearly an even division by respondents that they believed the police were not capable of performing their job.

Also significant were community responses on how they felt officers respond to White people and their neighborhoods. Over 50% of respondents feel that police are more courteous to, respond faster, and are more friendly and respectful in all interactions with White people than in their interactions with non-White people.

4.3 DISCUSSION

This survey examined the influence of police interaction with citizens' perception of interaction, crime and police work. The survey revealed some positive and negative findings about citizens' perception of crime and police work.

The survey yields a strong perception of the community feeling selectively racialized, which will

determine how the police will treat them. The community is less likely to cooperate with people they mistrust and may develop doubts regarding improvements in community-police interactions.

Regardless of whether profiling can be proven to occur in the context of the data's results, there is definitely widespread Oakland community perception that it is occurring and is cause for concern. This is a substantial perception disparity since…While more than 72% of officers do not believe racial profiling is a problem in the City of Oakland, and 53% believe fellow officers never engage, and another 36 % feel they rarely engage. As a result, a considerable number of officers responded that most citizens are somewhat satisfied to very satisfied.

The disparities of some of the perceptions are found to be statistically significant, which could result in the following: safety concerns for officers and community members may be increased in less safer neighborhoods, and left unchecked a stronger mistrust could develop towards the police department. This could lead to even more civil strife.

We are not offering the results of this survey as scientific, due to the limited number of respondents on certain questions. While the results do bring to light some significant survey results, it was primarily a starting point for the Task Force. As part of the next phase of reporting, the Task Force will work collaboratively with RAND to conduct a second survey.

5 ASSESSING THE POLICE DEPARTMENT PERCEPTION OF RACIAL PROFILING IN THE OAKLAND COMMUNITY

A survey was also conducted to study the process of how data collection contributed to the process of Oakland Police Officer's perception as it relates to the Oakland community. Data was generated from surveys from which polled officers' responses could be used to measure their awareness of how the community perceives their relationships with them, and the presence of racial profiling.

The department was researched and the responses compiled from a questionnaire that housed collected data for researching attitudes of community awareness. The estimation is that the data collected somewhat reflects highly developed information about police opinions relative to their organization and how it impacts community perception. Based on the surveys results, several determinations could be drawn, which are outlined below.

The structure of the respondents was as follows: officers, sergeants and commanders. What the survey reflected was that most Oakland police officers felt racial profiling never occurred, while most sergeants and 100% of the commanders felt it rarely occurred.

5.1 ANALYSIS AND RESULTS OF SURVEY

What the survey reflected is that a considerable percentage of officers, sergeants and commanders felt that fellow officers engaged in racial profiling, but that it was a rare occurrence. Over 50% of officers responded that officers never engaged in profiling, while 100% of the commanders and 67% of sergeants felt that police officers rarely engaged.

Additionally over 70% of officers polled responded that, "…by the definition given of racial

profiling…" it is not a problem at all in the City of Oakland.

Officers, Sergeants and Commanders all reported within the low 40 percentile that the Oakland community was somewhat dissatisfied with the Department. While 15% of the officers felt the community was very satisfied, no sergeants or commanders felt the community was very satisfied.

Officers reported in the low 30 percentile that the community was somewhat dissatisfied with the way the Oakland Police usually treat people. While 67% and 71% of Sergeants and commanders respectively reported that they believed the community was somewhat dissatisfied with their treatment by the police.

Notable, but in the low percentile (14%-16%) are the "very satisfied ratings" officers only assigned to: protection provided, citizens' satisfaction with how they are treated, and overall satisfaction of the department.

Additionally an overwhelming 53% of officers reported they believe officers never engage in racial profiling. While only 13% of sergeants believe the officers did engage, 67% of them felt the officers rarely engaged in racial profiling. Commanders were polled at 100% believing that officers rarely engaged in racial profiling. It is important to report that only 2% of officers felt that officers engaged "all the time."

5.2 DISCUSSION AND CONCLUSIONS

Based on the data's results there is a significant perception disparity between how the Oakland police believe the community feels about them, and how the community reports they are treated by the police, as it relates to their community-police relationships and the presence of racial profiling in the community.

The perception of racial and ethnic groups' feeling that they are being profiled must also be addressed due to the psychological impact of this belief. In other words, the impact of racial profiling has a social cost whether profiling can be proven to be occurring or whether it is based on people's beliefs. It is therefore imperative that steps be taken to address the concerns raised.

Note: We are not offering the results of this survey as scientific, due to the total number of respondents on certain questions. While the results do bring to light some significant survey results, it was primarily a starting point for the Task Force. As part of the next phase of reporting, the Task Force will work collaboratively with RAND to conduct a second survey.

6 TOWN HALL MEETING & MARKETING STRATEGY

Although the task force was comprised of representative groups, we felt it necessary to seek additional community feedback. In order to assess community perceptions about racial profiling, and the efforts of the task force, it was necessary to hold a community forum. To that end, the task force work collaboratively to host a Town Hall meeting.

The process of developing the agenda, selecting the location, and outlining the format of the meeting was extremely positive, and the group, as with most things, came to unanimous consensus

on all matters.

Using grant funds, we placed public service announcements on the local radio stations. Each group also distributed flyers that were developed by the committee. In this sense, the forum was extremely fruitful. However, we did run into a major glitch. We unwillingly scheduled the meeting for the night of the Major League Baseball playoffs in which the Oakland A's were playing. At the time we scheduled the meeting the team had not made the playoffs so this was not a consideration.

Consequently, attendance was low, only about 50 people attended. Nevertheless, we received invaluable feedback from the community members and leaders who did attend. The meeting was held at a High School auditorium, and it was videotaped.

We used a panel format for the meeting. Representatives from PolicyLink, the ACLU, NAACP, Oakland Police Officers' Association, and the Oakland Police Chief sat on the panel. The Department's racial profiling grant manager opened the session with an overview of the grant and our efforts to that point. Each member of the task force gave a 10-minute presentation, and then we allowed a significant period for community members to share their experiences and views, and ask panel members questions.

After the release of this report, the group will hold another Town Hall meeting in which we will discuss the results of the project and the results of the data collection. At this meeting, we will also outline the next steps the Department will take to address racial profiling.

7 IDENTIFY DATA FIELDS

The task force approached data collection somewhat different than many law enforcement agencies. Instead of identifying the data to be collected first, and then attempt to analyze that data later, the task force, in collaboration with RAND, identified what we thought was the ultimate goal of the data collection and identified what questions the task force wanted the data to answer. By doing this, the data collection supported the data analysis instead of the data analysis being shaped or limited by the data collection.

The task force wanted basic questions answered.

1. Are minorities disparately stopped by the police? If so, what are the causes of these disparities? And, are the causes societal-based or police-driven?

2. Does race influence an officers' decision to stop?

3. Are persons of color more likely to be detained longer or be subject to a search?

4. Are there operational and/or enforcement programs in the Department that result in disparate stops of minorities?

Next, the task force identified what data should be captured to answer these questions. For example, to determine whether certain enforcement programs result in disparities, the task force felt it necessary to collect specific data on special operations. To determine if minorities are detained

longer than non-minorities, the task force decided to collect the length of the stop. To determine if race influences an officer's decision to stop, the task force decided to have officers self-report if they could tell whether the driver was of color at the time they made the decision to conduct the stop. This provided the data analysis some ability to study stops when race can and cannot be observed. Section 9 provides a more in-depth explanation of data analysis and the benchmarking process.

As part of the process, task force members also reviewed local and national best practices and identified basic data that should be collected, such as: race, sex, age, reason for stop, search, and search results.

RAND worked with the group to identify additional confounding factors necessary for analysis. It should be noted, however, that this process was collaborative. Our research partner did not just identify the data to be collected in isolation. RAND worked with the group and we engaged in lengthy debate as to the perceived benefit of each data field.

In all, the task force identified over 24 data fields it thought necessary to answer the aforementioned question. To that end, there were numerous occasions in which task force members disagreed. But as with the other phases of this project, the task force came to unanimous consensus. As a result of their efforts, the Oakland Police Department now has the most comprehensive data collection program in the United States. Appendix 3 contains a copy of the data collection form.

This process also demonstrated the value of involving the Police Officers' Association (POA). In fact, it was the POA that recommended we capture whether an officer can tell the race at the time the decision to make the stop is made. This insight is based on their daily experiences of conducting stops in the evening and from observing cars from a distance. Without the POA involvement, something so simple would have most likely escaped other task force members.

There was initial debate on officer identification. It was the Department's initial position that officers' names would not be captured on the data-collection form, at least not in the first iterations of the program. The majority of the task force, except for the police representatives, opposed this decision and offered strong recommendations to collect this information. As mentioned earlier in this report, the Department was forthright at the outset and advised the task force that this was not something the Department was willing to compromise on at this time. However, the task force's recommendations would be noted in our technical guide. The task force was very understanding and accommodating to the Department.

Nevertheless, in January 2003, the Department entered in a Negotiated Settlement Agreement as a result of four Oakland officers, known as the Riders, violating the Constitutional rights of minorities in Oakland. As part of the agreement, the Department agreed to collect officer identification immediately following the release of the first data-analysis report.

The next section describes the technical component of the data collection and processing system.

8 DATA COLLECTION METHODOLOGY

SCANTRON Corporation provided this section

As a part of the project, the Department contracted with SCANTRON Corporation to use scanning technology to capture data. Because of this process the amount of time used to scan forms is significantly reduced, about 300 forms per hour, and we possess the ability to query documents in order to conduct comprehensive analysis and reports.

8.1 PROJECT SCOPE

SCANTRON's goal was to provide Oakland Police Department Officers with an efficient and accurate process for collecting data during traffic stops. The process also needed to address the manual methods of moving data from a paper source into a database. The goal was to automate the process of moving data from the paper forms into an ODBC compliant database for further analysis.

Technology is key to collecting accurate data and because of this, a study was completed to review the steps involved in moving the traffic stop data from a paper process to a total electronic process using a PDA handheld computer.

Below are the high-level steps implemented:

- Created a data collection form using TELEform (paper version)
- Created a data collection form using eListen (PDA version)
- Information captured using the paper solution (TELEform) will be transferred and merged with the .pda data in eListen using the .csv import feature.
- Reports are then generated, as needed, using the eListen Analyzer.

8.2 PAPER DATA COLLECTION (IMAGE DATA COLLECTION) OVERVIEW

1) Oakland Police Department and the task force worked closely with SCANTRON to develop the data collection question set and form layout.
2) A 2-day training session was scheduled at Oakland Police Department. Training covered the Image Data Collection process (all components of the solution).
3) SCANTRON printed the approved forms, which were then distributed to the field officers.
4) Officers complete the forms during traffic stops
5) Completed forms are collected and scanned at end of shift/day.
6) Images/data sets are brought into the TELEform Software for verification-data clean up using the verifier module.
7) Data is then transferred to an ODBC compliant database and/or eListen Survey

Software.

8) Reports can be created using the eListen Survey Software (mentioned below).

9) Scantron will continue to support and maintain software and hardware through Tech. and service support.

8.3 ELECTRONIC PDA COMPONENT (ELISTEN SURVEY SOFTWARE) OVERVIEW

1) A 2-day training session was scheduled at SCANTRON Headquarters. Training covered the use of eListen Survey Software, deployment to PDA handheld computers, combining data from paper data collection process (Imaging Solution) and then providing one data file for analysis.

2) Creation of the survey form template in the TELEform system into the eListen Survey Software begins the process

3) Forms are deployed to the PDA handheld devices.

4) Officers complete the forms during traffic stops using the PDA handheld device.

5) At the end of each day/shift, the submitted data is pumped to the eListen database from the PDA using a sync function.

6) Data from the paper-based form and the .pda are merged in eListen using the .csv import feature.

7) Reports can be created using the eListen Analyzer and/or data can be sent to ODBC compliant database for storage.

8) Scantron will continue to support and maintain software and hardware through Tech. and service support.

8.4 SOLUTION COMPONENTS:

Paper Solution Components (Software, Hardware, Technical Support and Professional Services)

- SCANTRON's Image Data Capture Software
  - TELEform Enterprise Software
  - TELEform eForm Module
- Scantron's Image Scanner Hardware
  - Panasonic KV2065 Image Scanner
- Scantron's Technical Support

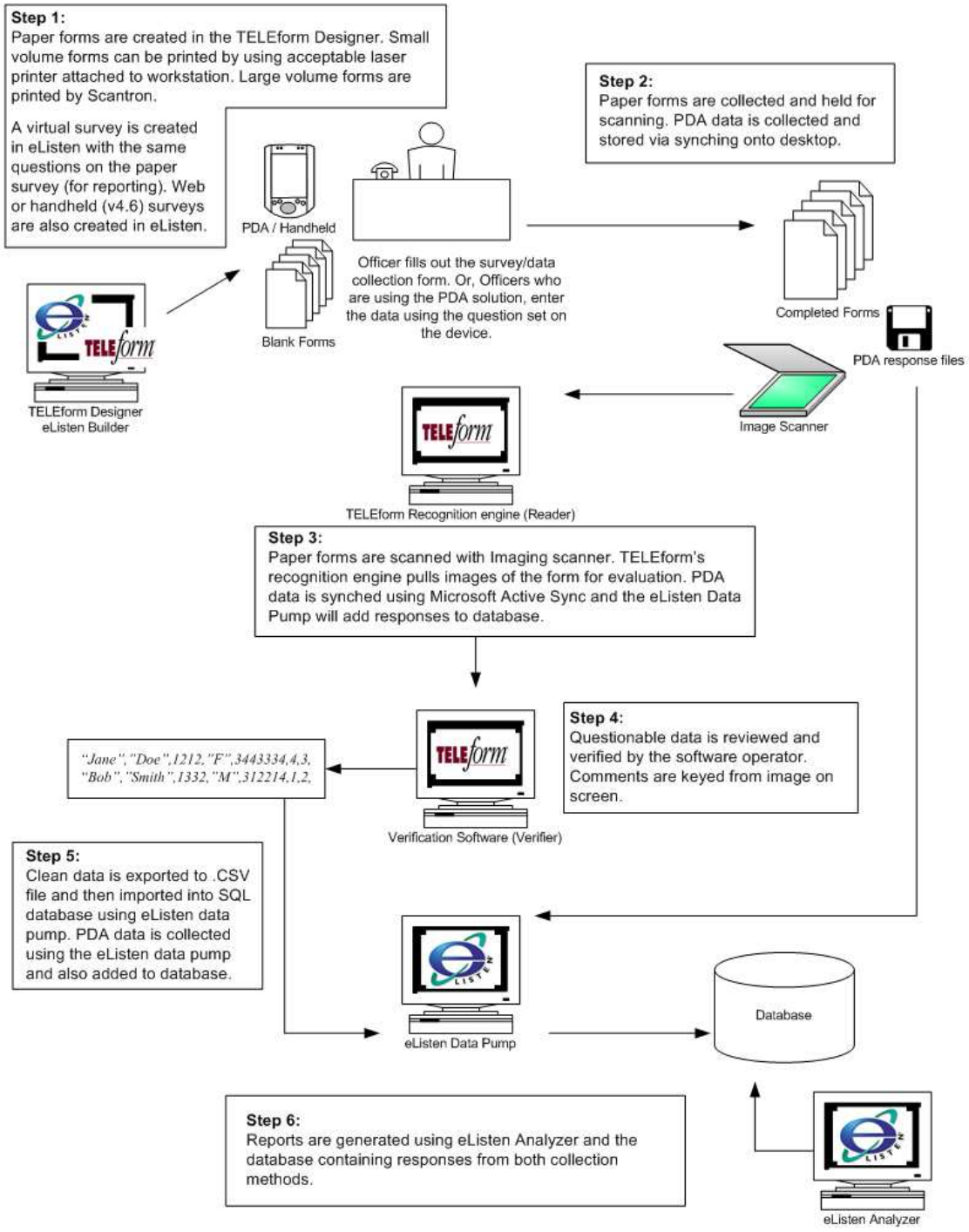
- Software Support
- Scanner "On-site" Service Support

- Scantron's Professional Services
  - Form Design and Implementation (programming)
  - Printing of Forms
  - Training on total solution components

Electronic Solution Components (Software, Technical Support and Professional Services)

- Scantron's Electronic Survey Software eListen
  - eListen Enterprise Software
- Scantron's Technical Support for Software
  - Software support
- Scantron's Professional Services
  - Training on eListen Software
  - Training on implementation with TELEform Components

8.5 PROJECT FLOW CHART (VISUAL)


Step 1:
Paper forms are created in the TELEform Designer. Small volume forms can be printed by using acceptable laser printer attached to workstation. Large volume forms are printed by Scantron.
A virtual survey is created in eListen with the same questions on the paper survey (for reporting). Web or handheld (v4.6) surveys are also created in eListen.
Step 2:
Paper forms are collected and held for scanning. PDA data is collected and stored via synching onto desktop.
PDA / Handheld
Officer fills out the survey/data collection form. Or, Officers who are using the PDA solution, enter the data using the question set on the device.
Completed Forms
PDA response files
Blank Forms
TELEform Designer
eListen Builder
Image Scanner
TELEform Recognition engine (Reader)
Step 3:
Paper forms are scanned with Imaging scanner. TELEform's recognition engine pulls images of the form for evaluation. PDA data is synched using Microsoft Active Sync and the eListen Data Pump will add responses to database.
Step 4:
Questionable data is reviewed and verified by the software operator. Comments are keyed from image on screen.
"Jane","Doe",1212,"F",3443334,4,3,
"Bob","Smith",1332,"M",312214,1,2,
Verification Software (Verifier)
Step 5:
Clean data is exported to .CSV file and then imported into SQL database using eListen data pump. PDA data is collected using the eListen data pump and also added to database.
eListen Data Pump
Database
Step 6:
Reports are generated using eListen Analyzer and the database containing responses from both collection methods.
eListen Analyzer

8.6 PAPER COMPONENT SOLUTION- TELEFORM PROCESS DESCRIPTION

8.6.1 FORM DESIGN

Although the system is designed to handle almost any kind of document, the best return comes from those documents that have been made friendlier to the customer and the technology in the system. These include things like plenty of room to write, some character segmentation, non-interfering registration targets, and clean colors.

Scantron encourages every customer to realize the true potential of automated data entry. Understanding the importance, Scantron provides their customers with additional forms optimization services. By optimizing your form design, your image processing will increase in speed. Increasing speed and streamlining your verification process provides you will see a quicker collection and approval process and therefore shortens the timeframe from when you will see your return on investment. The system is designed to determine the optimal balance between recognition performance and customer acceptance; you may find that an image friendly form is also very customer friendly. Scantron can conduct a hands-on class with the people in charge of form design to lay the foundation for this special type of form design.

8.6.2 FORMS RECOGNITION

The form recognition process consists of two parts: Forms Definition and Forms Processing:

Forms Definition is performed by the system administrator but is typically performed only once per form. When a form is created or modified, the new blank form is scanned into the system and "trained" or defined in preparation for use by the Forms Recognition process. New or changed forms may be added as required. During the form definition process, the operator defines the form to the system and identifies the location of the textual data fields that are contained in the form using simple click-and-drag operations without programming. The system records the topology of the form and uses this topology map to recognize forms as completed forms are scanned and passed to the recognition process.

Forms processing is performed by scanning forms containing data and passing it to the forms recognition process. The form is then compared against templates defined to the system. When a match of the template to the current form is found, the template information is passed to the OCR/ICR process to extract text data from the image bitmap. The OCR/ICR process is responsible for performing image pre-processing and cleanup, print recognition, data validation, and data formatting.

The engine uses information entered during the form definition process to extract specified fields from the form. Machine print (OCR), handprint (ICR), bar code, and mark sense information is automatically captured. Data successfully extracted will be stored in a file along with its associated image. The engine will attempt to extract data from all fields defined on the form. Any characters that have not been correctly recognized by the engine, as well as any validation errors detected, will be sent to an edit station for correction.

8.6.3 DEFINITION

Definition is the process where forms are trained in advance on how to read the information from a specific document. Complete this simple step using easy to follow click-and-drag drawing techniques, as described earlier. This only needs to be done once per document type, after which the software knows how to properly handle those documents.

8.6.4 DOCUMENT PREPARATION

The process starts with the forms received each day. Forms can be separated, batched, and ultimately prepared for the scanning function. These batches can then be scanned for processing. Forms must be unfolded, unstapled, etc. before scanning.

8.6.5 FORM SCANNING

Each stack of documents will be individually placed in the scanner's feeder. At the command of the scanner operator, the scanner will then automatically feed each stack, which then becomes a batch.

In the event that a jam occurs, the scanner control software will warn the operator and provide the opportunity to correct the problem and restart the process. Batches can be canceled and re-scanned at any point in time. Consistent document widths, consistent paper texture and thickness, and proper scanner cleaning and maintenance will ensure reliable performance and quality image capture.

8.6.6 FORM IDENTIFICATION

When the software detects that a batch has been scanned successfully, the first operation that is performed is document identification. Every image in every batch is identified to be one of the types of documents that the system has been trained to process. Various forms are detected by the software, allowing the system to know how to read the required information from each document.

8.6.7 AUTOMATIC DATA CAPTURE

When the form identification is complete for each batch, the system applies the previously defined form template to each document and extracts the data from each image. This process is done automatically.

Any low-confidence characters, fields, groups or validation checks that fail are flagged for an edit operator to review and correct. Typically, the system's initial accuracy varies from about 80% to 99.5%. Those documents that are difficult for the machine to read will require more verification by edit operators.

There are several factors that can influence the initial accuracy, such as the following:

- type of symbologies are being used
- quality of the actual data being filled in (e.g., dot-matrix vs. laser-printed OCR)

- number of pre-assigned validations
- quality of the form design and printing

8.6.8 DATA EDITING

Rather than keying all information contained on a document, edit operators simply have to handle those items that were poorly written or faulty in some way. The system flags each field that has been questioned for some reason, each of which is automatically brought to an operator's attention. Those fields and documents that have no questions pass through the system untouched. It is typical in most environments to view at least some portion of most documents.

Editing is a three-step process. A high-speed character mode is used first. One character after another is displayed for the operator. Characters can either be corrected at this stage or held for review for a later stage.

Any characters that are held at the character level are sent to a field-level context review. The image of an entire field appears on the screen. All an operator has to do is accept what was produced or retype those items that need correction. The net effect is that the overall labor required to process an order is typically about 50% of that required to key the whole thing manually.

Any fields that are held at the field level correction are sent to the form level correction where the operator has a full view of the form. The form level correction is also where the TrueAddress window will appear if this feature is being used.

8.6.9 DATA VALIDATION AND FORMATTING

One of the most powerful features of TELEform is the ability to validate data as it is being captured, as well as reformat data to a required style. Data validation checks can include table lookups, spell checks, math checks, validity checks, etc. Reformatting can include case changing, justification, trimming and padding, and the like. More complex checks or formats can easily be specified as well using the system's Basic Script functions.

All validation and formatting checks are performed during processing and data editing. No field that has an error can be passed through the system without intentionally doing so.

8.6.10 DATA TRANSFER

When the data passing through the system has been processed, cleaned up, and validated, it is ready to be transferred to the target database. The field order and file format to be used can be specified at the time of design and can be easily changed at any time. The transfer process can be set up to be automatic or manual. The transfer of data can either be via ASCII files, ODBC to a database such as Oracle or MS SQL Server or to a database such as Access or FoxPro.

8.6.11 RELEASE

The final stage in the production capture process is to release each document in the batch for storage and retrieval. After the batch of images has been enhanced, indexed and QA'd, the software will automatically release the images and defined index fields to be managed by the designated

imaging database.

8.6.12 IMAGE STORAGE

After the scanning and indexing processes are complete, the data is uploaded to a host database and the images will be automatically routed to the specified destination. Almost any repository can be used for image storage, including a database, RAID, CD, DVD or tape.

8.6.13 IMAGE RETRIEVAL

Authorized users can search, display, and print image files using a "thick" Windows client program or a "thin" client browser. End users will be able to access specific documents by information such as patient number, contact date, trip ID, name, phone number, payee code, and/or any other desired index criteria.

Thumbnails of multi-page, read-only documents are displayed for convenient reference, and full-size images can be rotated and displayed at various levels of zoom enlargement. If the document images have been annotated (electronic sticky notes, highlighting, etc.), these annotations will be displayed along with the original image. Once displayed, the user can keep the images current by selecting any of the annotation tools from the second toolbar at the top of the screen, including highlighting, graphical lines, boxes, electronic "sticky notes," redaction, etc.

9 ANALYSIS OF OAKLAND'S STOP AND SEARCH DATA

The RAND Corporation completed the analysis described in this section. The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. The principal contributors to this section were Greg Ridgeway, Ph.D., Statistician, K. Jack Riley, Ph.D., Director of RAND's Public Safety and Justice Unit, and Jeffrey Grogger, Ph.D., Professor of Policy Studies and Economics at the University of California, Los Angeles. This section has been peer reviewed and has received extensive comments from task force members.

9.1 SUMMARY

We analyzed 7,607 recorded vehicle stops in the City of Oakland, California between June 15, 2003 and December 30, 2003. We found that:

1) There appears to be evidence of substantial underreporting of stops. There are days in which OPD collects no stop forms and the volume on other days seems much too small. Judging only by the number of stop forms, compliance appears to have greatly increased in November, but an audit should attempt to verify complete reporting. Since we know nothing about the characteristics of the unreported stops, *all* results presented in this section are sensitive to this underreporting problem. OPD needs regular audits of reporting compliance and a program for improving compliance.

2) We consider methods for comparing the share of black drivers in the reported stops between times when the officer knew and did not know the driver's race in advance. The findings are mixed. We consider two measures of officers' ability to identify race in

advance and study their effect on the race of the drivers stopped. One measure indicates that race visibility increases the chance that an officer stops a black driver while another measure indicates that race visibility has no effect on the risk of an officer stopping a black driver. In addition, the large fraction of unreported stops prevents this finding from being conclusive.

3) Black drivers receive citations 68% of the time. Non-black drivers (all drivers that are not black including white drivers) stopped in similar circumstances received citations 72% of the time and similarly situated white drivers received citations 65% of the time. There is little evidence that officers cite black drivers at substantially different rates than other similarly situated drivers.

4) Black drivers were more likely to have stops lasting more than 10 minutes when compared with similarly situated non-black drivers and similarly situated white drivers. Black drivers had stops lasting less than 10 minutes 47% of the time as opposed to 53% for similarly situated non-black drivers and 55% for similarly situated white drivers.

5) Black drivers are as likely to be pat searched for weapons than non-black drivers (2.7% of stops) but similarly situated white drivers were much less likely to be pat searched (0.4% of the stops).

6) There were no statistically significant differences in the rates of consent searches. Officers did consent searches of 2.2% of black drivers, 1.6% of similarly situated non-black drivers, and 1.7% of similarly situated white drivers. However, the comparison with similarly situated white drivers was underpowered due to few white drivers being stopped in similar times and places as the black drivers.

7) Probable cause searches occurred in 3.2% of stops involving black drivers. Similarly situated non-black drivers and similarly situated white drivers were involved in probable cause searches at less than half the black driver rate (1.4%). Only 18% of the probable cause searches resulted in an arrest. Searches based on probable cause should almost always result in arrests. There were no discernable differences across the race groups in the rate of arrest following a probable cause search. However, black drivers account for 75% of the probable cause searches and, therefore, this inconsistency has the greatest effect on them.

8) Hit rates for high discretion searches were consistent across races so that there is no evidence of a disparity in the officers' search decisions. However, the sample sizes were fairly small so that the analysis was underpowered to discerning differences in hit rates.

9.2 INTRODUCTION

This section of the report describes the analysis of the 7,607 recorded vehicle stops that the Oakland Police Department reported between June 15, 2003 and December 30, 2003. We investigated whether there was evidence of racially-biased policing at any point in the stop process, in the decision to stop, to cite, and to search a driver. Assessing a race bias at each of these stages

requires a different set of analytical tools since we have different kinds of data available at each stage.

For the analysis of the decision to stop, we apply a method of analysis developed in Grogger and Ridgeway (2004). This method directly links the officers' ability to identify the race of the driver in advance with their decision to stop the vehicle. Next we examine post-stop activity for evidence of racial bias including the duration of the stop and the decision to cite. Lastly we study the decision to search and search outcomes to determine whether officers apply an equal level of suspicion when deciding whom to search. As with all studies based on observational data, we stress that all of the findings for or against racial bias in the vehicle stops are subject to various assumptions. We articulate those assumptions in presenting each method, give our assessment for how sensitive the findings are to the assumptions, and provide information on why the assumptions are likely to be practical for the analysis of Oakland's data.

Other sections of this document define racial profiling and race bias in vehicle stops. The analysis provided focuses on assessing whether officers treat drivers in similar situations equally regardless of race. However, in Oakland residents of different races find themselves in different situations. They drive at different times of day and in different parts of the city. They might even commit different kind of offenses at different rates. Police allocate their law enforcement efforts in ways they deem to be most effective. Since drivers of different races drive in different parts of town, allocating additional law enforcement effort in certain parts of the city will naturally expose members of certain races to more law enforcement. In Oakland, officers are more heavily deployed in non-white neighborhoods. In addition to differences in deployment, police practices might differ in various neighborhoods, perhaps due to special crime prevention programs in high crime neighborhoods (e.g. directed patrols) or additional concerns for officer safety (e.g. more frequent use of pat searches for weapons). Even if police practices are equal for similarly situated drivers regardless of race, when drivers of the various races do not traverse similar streets at similar times, there will be differences at the aggregate level in the likelihood of being stopped or pat searched. Some may legitimately argue that it is these differences in exposure that are the root of the racial profiling problem. The department's officer deployment policy is a topic worthy of discussion and negotiation with the various neighborhoods. It does not, however, address whether patrol officers differentially treat drivers of different races. This is the issue on which this section focuses. Here we exclusively focus on whether similarly situated drivers are treated equally regardless of race.

9.3 DESCRIPTION OF THE DATA

In this section we give an overview of the available data. The data itself presents the greatest challenge in offering a reasonable assessment of racial profiling in Oakland. During the study period the stops were substantially underreported. As a result *every* figure and number in this report relates to the reported stops, which may differ greatly from the numbers we would have obtained had we also observed the currently undocumented stops. The Oakland Police Department needs to implement a plan to improve reporting compliance and regular auditing to measure compliance.

The number of stop forms produced on any given day varies greatly as shown in Figure 1. The daily number of stop forms varies from 0 to 216 with an average daily rate of 38 stop forms. This number should be compared with the number of citations or communication logs to gauge how close this is to the true number of stops that the department is making. Surely the irregularity of the number of stops indicates that compliance has been a problem. The volume of stops appears to be increasing over time, presumably as a result of increasing compliance with the department's

reporting policy. Even in November, where the number of stops peaks we still do not know whether there are another hundred undocumented stops. Unfortunately, we cannot assume that the unreported stops are unrelated to factors of great interest to us, such as race and the outcomes of the stop. As a result, if the unreported stops look much different than the stops on which we have data then the analyses provided in this report can overstate or understate the magnitude of racial profiling.

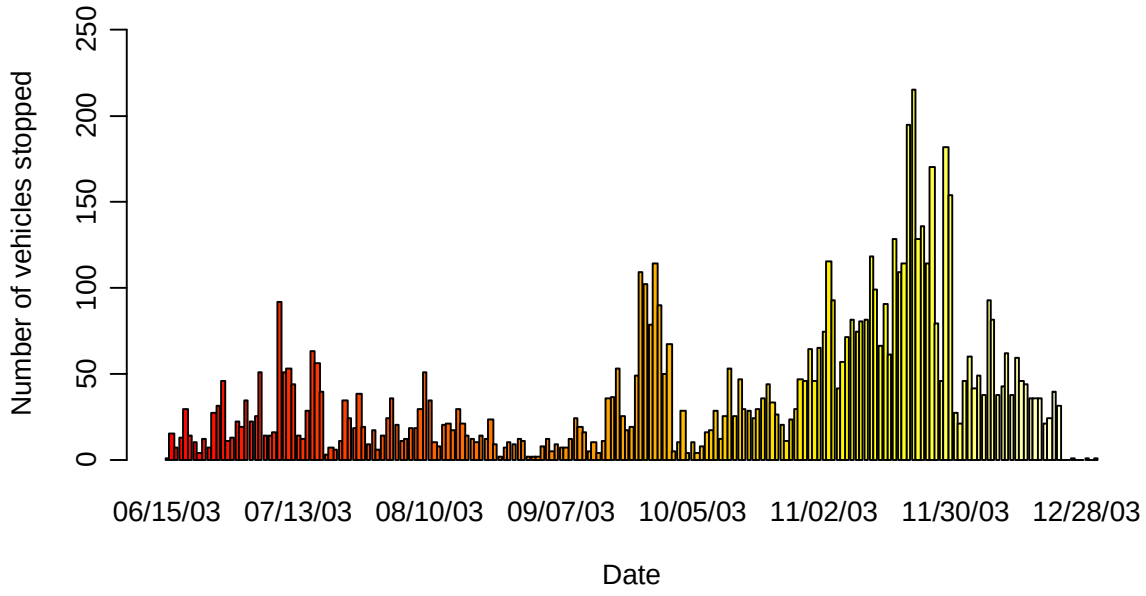


Figure 1: Number of vehicles stopped by date

Missing data on individual items do not seem to be too large of an issue. For example, 4.8% of the forms do not indicate the race of the driver, 2.0% do not give a reason for the stop, and almost 3.7% do not indicate the outcome of the stop. Needless to say this reduces the resolution at which we can analyze the data and introduces questions of why certain items are left unanswered. However, the data indicate that the rate of incomplete items is decreasing over time.

Table 1: Race distribution of the city, stopped drivers, and searched drivers

	Asian	Black	Hispanic	White	Other
Oakland census *N*=399,484	15%	35%	22%	21%	7%
Stopped drivers *N*=7,607	9%	56%	15%	14%	5%
Searched drivers *N*=1,600	3%	75%	16%	2%	5%
Frequency of search following a stop	10%	40%	31%	10%	9%

The greatest concern in Oakland is that black drivers, while representing 35% of the residential population, are involved in a large number of stops, searches, and are frequently searched when stopped. Table 1 displays these percentages. The other category includes Middle Eastern, Native American, and Pacific Islander race groups. At first glance we would assume that in the absence of a race bias, that the race distributions of stopped drivers and of searched drivers would be similar to the residential census. The large differences shown in Table 1 is certainly cause for closer inspection of Oakland's vehicle stops. In addition, in the absence of a race bias we would assume that the frequency of searches following a stop should be roughly the same across the race groups, yet again Table 1 shows large differences across the races. In reality the differences shown here say little if anything about racial profiling. The remainder of this report aims to present methods that more accurately assess the effect of race bias in stops, searches, and other outcomes.

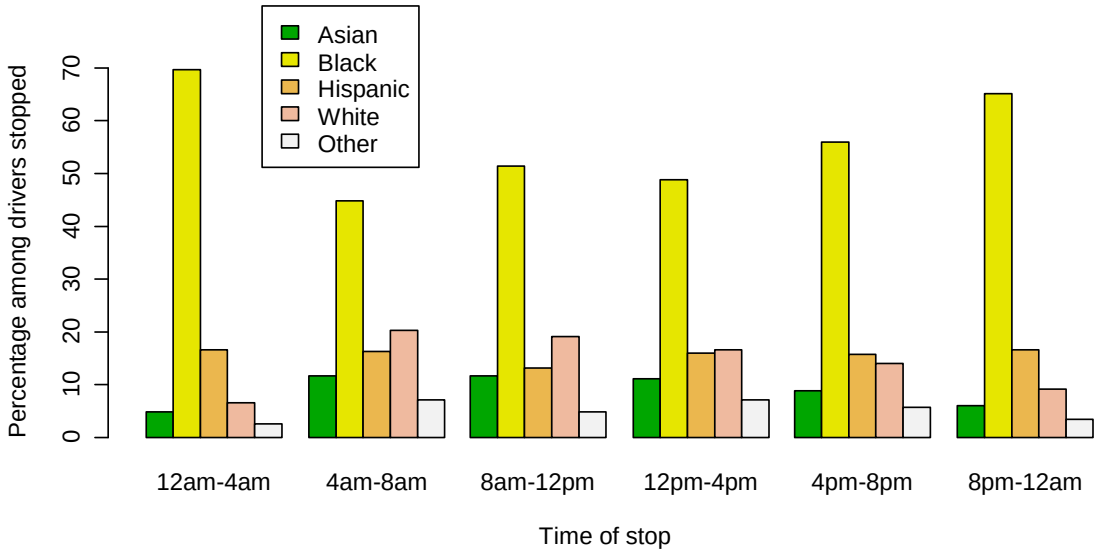


Figure 2: Distribution of the race of stopped drivers by time of day

Black drivers remain the majority race among stopped drivers at all times during the day as shown in Figure 2 below. Late at night and in the early hours of the morning the difference becomes even more pronounced. The variation in these percentages over the course of the day is likely a product of changes in exposure at different times of the day. Changes in exposure are due to changes in the deployment of police officers and the rates at which members of the different races travel on Oakland streets at different hours of the day. Note that the mix of Hispanic and white drivers among stopped drivers changes throughout the day. From 4am to 4pm, there are more white drivers than Hispanic drivers stopped while from 4pm to 4am there are more Hispanic drivers than white drivers stopped. Officers seem to rarely stop Asian, white, or drivers in the Other race category from 8pm to 4am. Possibly these drivers change their driving behavior during these hours, perhaps driving more carefully after dark, maintaining headlights in an operable condition, taking streets that are less busy (and perhaps less policed), or have work hours that allow them to be off the road early. We do not know the reasons for these observed stop patterns in Figure 2, but assessments of racial profiling have to allow for the fact that driving behavior, driving patterns, and exposure to police may figure in to the differences across race groups in the stop rates.

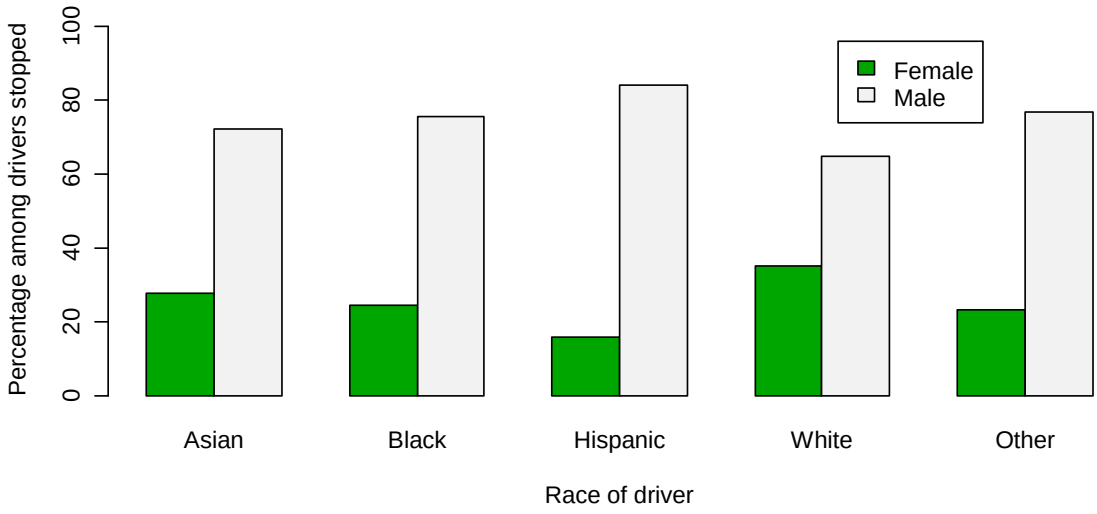


Figure 3: Distribution of the race and sex of stopped drivers

Stopped Hispanic and black drivers are more likely to be men as shown in Figure 3, especially when compared with white drivers. Men overall are more likely to be stopped, composing 77% of the vehicles stopped. If men's driving frequency and behavior put them at greater risk of being stopped then races with a greater prevalence of male drivers can increase that race's representation in the stop dataset. For example, note that over 80% of stopped Hispanic drivers are male compared with 65% male for white drivers. If indeed male drivers are more prone to violations, a predominantly male Hispanic driving population could cause the stop rates for Hispanics to be higher than would otherwise be expected.

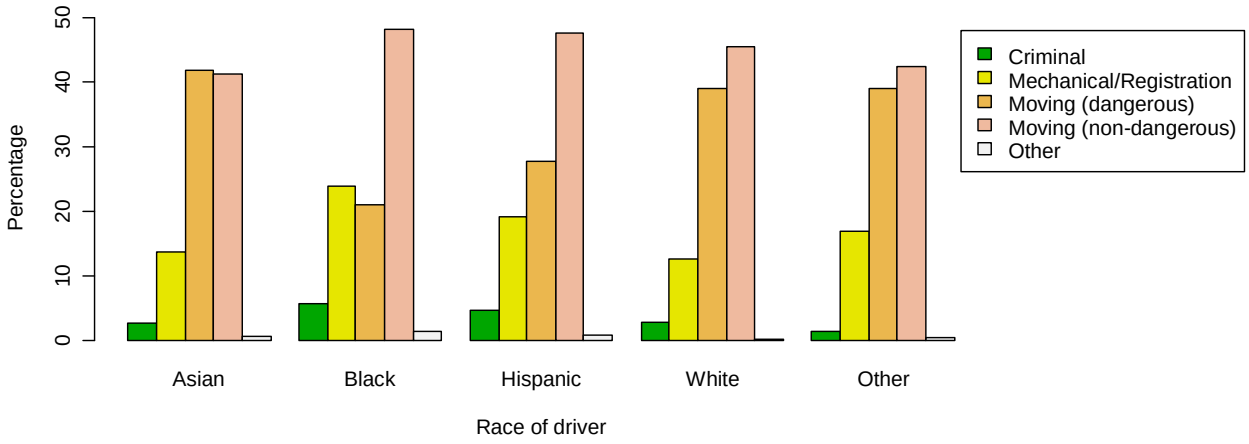


Figure 4: Reason for the stop by race

Figure 4 shows the distribution of the reason for the stop broken down by race. In general there does not seem to be much of a difference across the races with moving violations by far the most common. White drivers are the least likely to be stopped for mechanical or registration violations. Black drivers are more likely than the other races to be stopped for criminal violations (5.6%), which include violations of local ordinances (2.4%). Black drivers were also the only group of drivers to

have a substantial number of stops for other reasons (1.4%), which includes 23 probation/parole stops and 32 be-on-the-lookout (BOLO) stops.

Racial profiling can take the form of more harshly enforcing laws concerning minor traffic infractions based on race. As Figure 4 indicates, white drivers are more likely to be involved in dangerous moving violations, the kind of stop requiring little officer discretion. Mechanical and registration stops are disproportionately associated with black drivers. It is possible that black drivers are more likely to have expired registration or vehicles in disrepair. It is also possible that officers are using these conditions as a pretext to stop these drivers. We have no information on the actual condition, year, or make/model of these vehicles. Unfortunately, our analysis of bias in the decision to stop cannot use stops for mechanical or registration violations as described in the next section. However, we will be able to assess bias in post-stop activity (e.g. search, citation) utilizing these stops.

9.4 RACE BIAS IN THE DECISION TO STOP

As previously shown in Table 1, we found that 56% of the vehicle stops during the study period involved a black driver. The number of stops involving a black driver is more than three times the number involving a Hispanic driver and almost four times the number involving a white driver. The question of concern to Oakland residents, and the question the police department needs to address, is whether or not it is "appropriate" for black drivers to make up 56% of the drivers that the Department stops. Since only 35% of Oakland's residential population is black, residents may be suspicious about the department's practices.

The difference between 56% and 35% may not necessarily be the result of racial profiling.[1] Some other possible causes for the disparity are:

1. *Driving behavior may vary by race*. Black drivers may be stopped more often because they are more likely to commit some kind of traffic infraction. This may include speeding, running stop signs, mechanical, and registration violations.

2. *Exposure to law enforcement may vary by race*. Black drivers may be stopped more often because they are more likely to drive longer distances or through areas of the city that have a greater law enforcement presence. Naturally, the greater the law enforcement presence the more likely that an infraction will be noticed.

If the percentage of black drivers stopped equals the percentage of black drivers among the "at-risk population[2]," then racial profiling is not occurring. The "at-risk population" is defined as drivers in Oakland who:

---

[1] Analogously, in the same dataset we found that 75% of the drivers stopped are male. Even though this figure differs greatly from the residential rate, one could surmise that much of this disparity is due to men driving in the city more often and more likely to break traffic laws when they do drive. It is possible that police also stop men more frequently due to suspicion, as officers are more likely to be asked to be on the lookout (BOLO) for men rather than women. The delicate boundary between good police practice and profiling lies somewhere in the vague spectrum ranging from evidence in a specific suspect description and acting on general opinion held about the danger of male drivers.

[2] Health epidemiologists use the term at-risk population to describe individuals who are both exposed to a disease-causing agent (e.g. a virus, radiation) and have characteristics (e.g., age, sex, diabetes) that make them more prone to contract the disease when exposed.

1. exhibit characteristics that would make an officer stop the driver (e.g., speeding, mechanical violation); and

2. are exposed to law enforcement

How do we determine the percentage of black drivers among the at-risk driving population? Measures of the race distribution of the at-risk population are commonly called "benchmarks" and the process of creating these estimates "benchmarking." The following section provides various methods for formulating benchmarks for analysis of racial profiling in vehicle stops.

9.4.1 TRADITIONAL BENCHMARKING METHODS

Census data, traffic surveys (Lamberth, 1994), and traffic accident data (Alpert and Smith, 2003) have been used to estimate the race distribution of the driving population for use as a benchmark. Each of these is unlikely to be sufficient to gauge the effect of racial bias in Oakland's vehicle stops.

9.4.1.1 Census data

As previously discussed, census data can be very different from the driving population, especially during heavy commute hours when residents of outlying areas pass through predominantly non-white neighborhoods to get to jobs in downtown Oakland. In addition, difference in car ownership, time on the road, and distances traveled all affect whether a resident is likely to be exposed to police. While racial profiling reports continue to use census data as a benchmark, researchers almost uniformly admit that its use is inappropriate if not irresponsible. The census benchmark can potentially exaggerate or understate the effect of race bias on vehicle stops (Fridell, 2004).

9.4.1.2 Traffic Surveys

Traffic surveys, while potentially effective, can be very expensive and have limited scope in an urban environment. Traffic surveys usually select a small set of intersections or road segments for analysis. If a neighborhood with a large racially-biased policing problem is not selected for the survey, then the problem can go unreported. A large, well-designed traffic survey can bound the probability that such a problem neighborhood exists.

Traffic surveys have been used to assess some aspects of the differences in driving behavior by race. For example, a study of racial profiling on the New Jersey turnpike (Lange, Blackman, and Johnson, 2001) found that black drivers were twice as likely as white drivers to exceed the posted speed limit by more than 15 mph. In studies of vehicle stops on highways, assessing which vehicles should be stopped is more or less straightforward since speeding is the primary offense and is easy to quantify. In the urban environment where drivers commit a wide variety of infractions some of which are subject to greater officer discretion, assessing the race distribution of drivers committing stoppable offenses is extremely difficult. Deciding which vehicles are committing a stoppable offense is likely to be subjective.

Since assessing differences in driving behavior is difficult, as well as a sensitive subject, analyses often assume that violation rates simply do not differ by race. Defenders of this assumption argue that nearly all vehicles are violating some part of the vehicle code and violating frequently. Lamberth (2003) reports on a study in which officers spotted violations for 94% of vehicles with an average

"time-to-identification" of 28 seconds. While this may seem compelling, many of these violations are high-discretion violations for which officers simply do not stop vehicles, unless some other factors enter the picture. This 94% figure, therefore, does not represent the fraction of vehicles that are committing a violation that would actually cause an officer to initiate a stop. Under the assumption of no behavior differences, the analysis needs only to determine the distribution of race of drivers on the street exposed to law enforcement.

9.4.1.3 Not-at-fault car crashes

Data on the not-at-fault driver involved in traffic accidents provides an easy and inexpensive method of forming an estimate of the race distribution of the driving population (Alpert and Smith, 2003). Presumably all cars traveling in a particular area at a particular time are equally at risk for being struck by another vehicle (e.g. rear-ended at a stop sign, sideswiped). Although it is possible that certain defensive driving behaviors reduce the risk of being in a car accident, we must assume that these skills are equally distributed across the races in order for a car crash analysis to give meaningful estimates of disparities in stop rates. Furthermore, the car crash benchmark does not account for differences in exposure to law enforcement, which can vary greatly in the city by time and location. Oakland does record the not-at-fault driver's race and sex on its car crash accident reports and future analyses may include an investigation of these data.

9.4.1.4 Disadvantage of the current benchmarking methods

Accounting for exposure to police is one of the primary difficulties in current benchmarking methods. While some of the benchmarking methods previously discussed can be adjusted to account for exposure, proper adjustment remains a difficult task to complete correctly.

As opposed to studies of racial profiling on highways where all drivers are exposed, in urban environments differences in exposure may account for a large part of the apparent disparity. The high crime area of Oakland known as the Flatlands is nearly 80% non-white. Officers are regularly called away from their patrols in the predominantly white Hills neighborhoods to handle the volume of calls in the Flatlands. As a result, a driver might be able to run stoplights across the Hills without exposure to law enforcement, but such behavior would likely be noticed immediately in the Flatlands. As a result, part of the disparity could be a product of the increased law enforcement effort in non-white areas of the city caused by greater calls-for-service. While recognizing the exposure issue is the first step, actually measuring exposure is difficult. We considered using man-hours by policing beat extracted from patrol logs, but officers' assignments are frequently reshuffled during the shift as situations arise. In the end, it has not been feasible to know where officers spend their time, compounding the problem of not knowing where drivers of different races spend their time. As a result measuring the at-risk driving population is extremely expensive and logistically complex. Nevertheless, police and the citizens that they serve need accurate and inexpensive methods for assessing whether officers practice racially profiling.

For Oakland we have developed an approach that sidesteps many of the main difficult benchmarking issues. We essentially construct a benchmark using the race distribution of stops when officers are unable to identify the race of the driver in advance. The next section describes this method.

9.4.2 AN ALTERNATIVE TO BENCHMARKING

As previously discussed, developing a solid benchmark is challenging. All of the benchmarks available have shortcomings in one respect or another. The complexity we face in determining the race distribution of the at-risk driving population made us reconsider whether the fixation with this benchmarking method was warranted.

Grogger and Ridgeway (2004) proposed an alternative approach based on an ideal experiment. They recommended comparing the race distribution of vehicle stops when the officers know the race of the driver against the race distribution of vehicle stops when the officers do not know the race of the driver. If there is no racial bias then:

% black drivers among those where the officer *knew* the race in advance =
% black drivers among those where the officer *did not know* the race in advance.

That is, advance knowledge of the driver's race should not influence the race distribution of stopped drivers. The advantage of such an approach is that it bypasses the questions of unequal law enforcement exposure and differential driving behavior. It attempts to directly answer the question of whether an officer's knowledge of a driver's race influences the rate at which drivers are stopped.

Practically speaking, though, how can we determine an officer's knowledge prior to a stop? We propose two possible methods for addressing this issue. First, Oakland's task force asked officers to self-report their knowledge of the driver's race by including the following question on the data collection form: "Could you determine prior to the stop whether the person was of color?" Within the data sample, officers completed this field on the form 94% of the time. The next section offers an initial assessment of that method. The section also argues that relying on officer self-reports is problematic. Section 9.4.3 provides a solution that does not require self-reports from officers.

9.4.2.1 Analysis with officer self-reported advanced knowledge of drivers' race

We first discuss the analysis relying on officers reporting whether they knew the race of the driver in advance. This analysis is *insufficient* to measure racial profiling since it relies on officer self-reports that are impossible to verify, but this method offers an introduction to the general strategy. We discuss its shortfalls afterwards and present a solution that avoids relying on officer reported advance knowledge of the driver's race making the analysis more robust to officer reporting.

Table 2: The rate at which black driver's are stopped depending on whether the officer knew the driver's race in advance. At all hours of the day, the two rates are statistically significantly different (*p-value* < *0.0001*).

Officer knew the driver's race in advance	Number of black drivers stopped / Number of drivers stopped	Percent black	95% confidence interval
Morning hours 8am-12 noon			
No	422/934	45%	(42%, 48.4%)
Yes	230/351	66%	(60.6%, 70.5%)
Afternoon hours 12 noon-4pm			
No	482/1166	41%	(38.5%, 44.2%)
Yes	232/347	67%	(61.9%, 71.8%)
Evening hours 4pm-8pm			
No	636/1272	50%	(47.3%, 52.7%)
Yes	265/379	70%	(65.3%, 74.5%)
Night hours 8pm-12 midnight			
No	596/969	62%	(58.4%, 64.6%)
Yes	196/258	76%	(70.8%, 81.2%)

For each period of the day, we compute the percentage of drivers that are black when the officer knew the race of the driver in advance and the percentage of drivers that are black when the officer did not know the driver's race in advance. Table 2 shows the results for four 4-hour periods of the day. In the morning, when officers indicated that they could not tell the driver's race in advance, 45% of the drivers stopped were black. This should represent the benchmark rate at which a race-blind police force should stop black drivers. When officers reported knowing the race of the driver in advance, 66% of the drivers stopped were black. Thus, when officers report knowing the race of the driver in advance, black drivers are 1.5 times more likely to be stopped than when officers cannot tell the race of the driver in advance. Similar results hold for the afternoon drivers. The disparity seems to remain for evening and night stops, however, the race effect may be confounded during nighttime hours since darkness introduces other kinds of violations (e.g. headlight violations) and inhibits the officers from identifying the race of the driver, a feature we will take advantage of in the next section.

This analysis indicates that there may be a race effect in the decision to stop black drivers. According to the officers' reporting, instances in which the officers know the driver's race in advance are much more likely to involve black drivers. This may indicate that race visibility increases the chances of officers stopping black drivers. However, there are some caveats to this analysis discussed next.

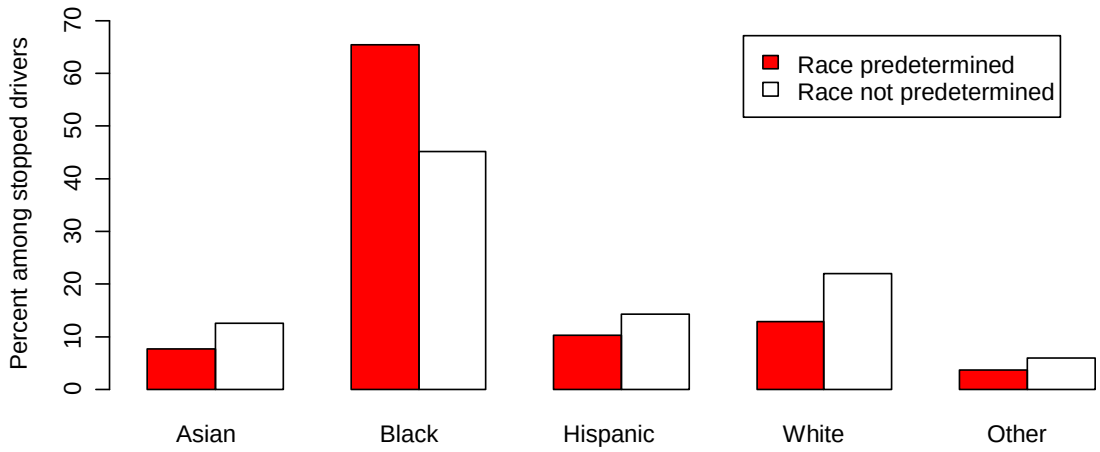


Figure 5: Percentage of morning hour (8am to 12 noon) drivers stopped, separated by race and whether the officer knew the race of the driver in advance

The analysis so far focused only on black drivers. Figure 5 shows the results for all of the race groups for the morning hours (the afternoon results were similar). Only black drivers appear to be strongly affected when the officer identifies the race of the driver.

9.4.2.2 Disadvantages of using self-reported officer responses

A more subtle assumption in this analysis is that black drivers are at equal risk of being stopped when officers can and cannot identify their race in advance. Certainly identifying a driver's race in advance is more difficult at night, but for the first two time periods in Table 2, natural lighting is about the same yet a disparity still exists.

Relying on self-reported data from the officer concerning advance knowledge of the driver's race is problematic. For example, the officers may be identifying black drivers more easily. The actual question on the form asks whether the officer could tell in advance whether the driver was "of color." At distances, white, Asian, Hispanic, Middle Eastern, and other races might easily be confused but black drivers might be more easily identified. If white drivers are exposed to officers when traveling at higher speeds than black drivers then the observed disparity could be a product of differences in speeds rather than differences in race. As a result the observed differences could be a result of officers being more likely to indicate advance knowledge of the race when they see a black driver. In addition, this analysis *critically* depends on the officers answering the form questions honestly and accurately. If officers in predominantly black neighborhoods always answer accurately but officers in other neighborhoods always answer no then this will skew the analysis.

The benefit of this style of analysis is that it did not require us to estimate the race distribution of the driving population or any kind of benchmark. It directly addresses the question of whether the driver's race influences an officer's decision to stop a vehicle. The differences documented in Table 2 could result from racial bias and may provide some evidence of racial profiling. However, this analysis is *not* completely satisfactory because of the problems with self-reports. The next section aims to relax that assumption using natural lighting to moderate an officer's ability to identify the race of a driver.

9.4.3 USING CHANGES IN NATURAL LIGHTING

This section describes the method for measuring racial profiling in stops proposed by Grogger and Ridgeway (2004). Patrol officers claim that at night they can rarely determine the race of the driver in advance. As a result, as the day transitions from daylight to darkness the officers' ability to identify a driver's race in advance decreases. Rather than rely on officers to self-report whether they had advance knowledge of the driver's race, we can rely on variation in natural lighting to control their ability to observe a driver's race. For the method to correctly estimate the effect of racial profiling, officers do *not* need to be completely "race blind" at night and have complete race identification in advance during the day. The method only needs a *degraded* ability to see the driver's race after dark.

To determine whether an officer's ability to identify the driver's race in advance influences their decision to make a stop, we compare the race distribution of drivers stopped in daylight with the race distribution of drivers stopped after dark. Since the race distribution of drivers on the road may change over the course of the day, directly comparing daytime to nighttime stops confounds the effect of racial profiling with changes in the driving population. To mitigate this problem we compare only those stops occurring near the boundary of daylight and darkness, a time interval during which the driving population cannot change too quickly. To be precise we define darkness to begin at the end of civil twilight.[1] During the study period the end of civil twilight occurs between 5:19pm and 9:06pm depending on the season so we consider stops only in this interval. We do not consider stops occurring between sunset and the end of civil twilight, usually lasting about 20 minutes, since we cannot determine darkness or visibility during this period. To summarize, we label stops occurring between 5:19pm and sunset as "daylight stops" and stops occurring between the end of civil twilight and 9:06pm as "darkness stops." We only included moving violations in this analysis since the likelihood of an officer recognizing a mechanical or registration violation changes from daylight to darkness (e.g. headlight violations only occur at night and may be more likely associated with drivers of a particular race).

Table 3: Comparison between daylight and dark of the percentage of stopped drivers that were black. All drivers stopped for moving violations between 5:19pm and 9:06pm exclusive of the period between sunset and the end of civil twilight, (*p-value=0.29*)

	Light	Dark
N	401	575
Percent of stopped drivers that were black	50%	54%
95% confidence interval	(45.7%, 55.5%)	(50.2%, 58.3%)

Table 3 gives the basic idea of this comparison. During daylight hours black drivers composed 50% of the stops while at night they composed 54% of the stops. This is not a statistically significant

[1] Technically this occurs when the center of the sun is 6 degrees below the horizon, but practically it is when one can see the brightest stars and artificial light is needed to perform most outside activities.

difference and, counter to the racial profiling concerns, black drivers do seem slightly less at risk for being stopped during daylight.

This comparison depends on two key assumptions. First, the substantial underreporting can affect this analysis if the officers that are not reporting their stops are the ones with the greatest difference in their stop rates of black drivers between daylight and darkness. Second, in order for the difference to represent a racial profiling effect, the mix of black and non-black drivers at risk for being stopped must remain the same. The volume of traffic may increase or decrease, but the relative representation in the at-risk population must remain constant. Otherwise, if the proportion of drivers on the road who are white is much greater after dark, then the observed equality in the percentage of black stopped could be due solely to shifts in the driving population and would mask any effect of racial bias. To prevent changes in exposure from causing such problems, we focused the analysis on the hour before and the hour after the end of civil twilight, under the assumption that the mix of black and non-black drivers does not change drastically over that time interval.

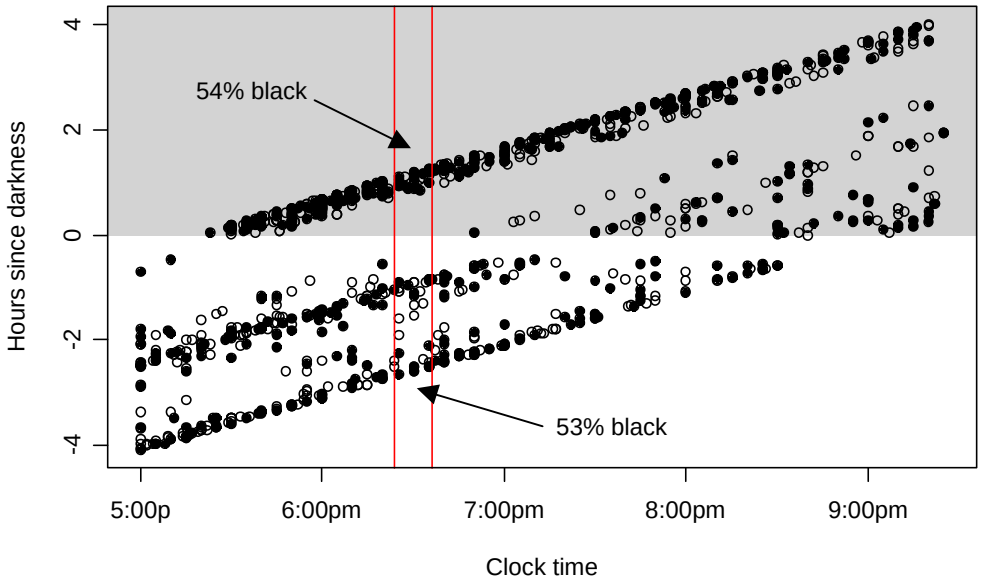


Figure 6: Plot of stops by clock time and darkness. The solid points indicate black drivers and the open circles represent non-black drivers. The shaded region indicates those stops occurring after the end of civil twilight. Note that at each value of clock time some stops occur in daylight and some occur in darkness. The vertical lines mark a period around 6:30pm discussed in the text. When it is daylight in that time interval, 53% of the stops involved black drivers and when it was dark in that time interval 54% of the stops involved black drivers. The gap in the data is due to the daylight savings time ending in October.

To improve upon the analysis we also adjusted for "clock time." For example, we consider drivers stopped at 6:30pm and compare the fraction of black drivers among them on days when 6:30pm is during daylight and on days when 6:30pm is after dark. This refinement relaxes the assumption that the mix of black and non-black drivers remains constant, requiring it to hold only within small intervals of the clock time. Figure 6 depicts this notion. The vertical lines mark a 12-minute interval around 6:30pm. Some stops occurring in this interval occur during daylight (unshaded) and others in darkness (shaded). If we assume that the mix of black and non-black

drivers at risk for being stopped does not change between the dark and daylight stops, then the fraction of black drivers (solid dots) stopped during daylight should be close to the fraction stopped after dark. Essentially we compute the difference in the fraction of black drivers among the daylight and darkness stops at each value of clock time, looking for time intervals in which chance alone cannot explain the observed differences. We found that, on average over clock time intervals, black drivers were no more likely to be stopped during daylight than after dark. In fact, the relative risk of a black driver being stopped (relative to a non-black driver) was slightly less during daylight hours, similar to what Table 3 indicates.

9.4.3.1 Sensitivity of the results to assumptions

Our analysis concluded that the black driver's share of stops remained the same during daylight and after dark. However, equality in their share of daylight and darkness stops could still imply a race bias if there were many fewer black drivers at risk for being stopped during the day. For example, if black drivers comprised 20% of the at-risk drivers during daylight and 70% after dark, equality in their share of daylight and darkness stops would indicate a problem. As described above we took several measures to ensure that black drivers' share of the at-risk population did not change between the stops we used for daylight and darkness stops (adjusting for clock time and considering only those stops occurring near the boundary of daylight and darkness).

We also completed a sensitivity analysis to make sure our results were insensitive to potential differences in exposure. We asked by how much the percentage of black drivers exposed would need to change between daylight and darkness for our conclusion of no racial profiling to be reversed. We found that the percentage of black drivers during daylight would have to be at least 10 percentage points less than the percentage of black drivers after dark (e.g., 50% after dark but only 40% during the daylight) in order for the conclusions to change. We believe a change of this magnitude is unlikely especially since we have controlled for clock time. Work, school, and business hours set the schedules for many, thus it is unlikely that the black driver population would increase by 10 percentage points on days when 6:30pm occurs after dark relative to days when 6:30pm occurs in daylight.

The analysis utilized stops from June through December. As a result many of the daylight stops occur during summer months and darkness stops occur in fall months. While seasonal differences in traffic volume will not affect the analysis, seasonal changes in the race distribution of drivers can. We repeated the analysis using only October and November data since activities in those months are relatively constant (with the exception of Halloween and Thanksgiving). The analysis still concludes that, if anything, black drivers are less likely to be stopped during daylight. This conclusion is still subject to problems in the underreporting of the stops.

9.4.4 OVERALL CONCLUSIONS ABOUT TRAFFIC STOPS

The evidence concerning racial profiling in traffic stops is mixed and depends on the method for measuring the officers' ability to identify race in advance. While analysis utilizing self-reports from officers indicates evidence in favor of a race bias, it appears from our analysis based on variation in natural lighting that an officer's ability to identify the race of the driver in advance does not influence whom they are stopping. However, both of these conclusions may be sensitive to the substantial underreporting of stops in Oakland.

As with other analyses assessing race bias in the decision to stop, the results do rely on certain assumptions. The findings assume that the fraction of black drivers in the at-risk population (driving, exhibiting stoppable behavior, and exposed to law enforcement) does not change from daylight to darkness. We have relaxed this assumption to some degree by controlling for clock time in the analysis, the idea being that clock time is more likely to affect the mix of black and non-black drivers on the road rather than darkness. Fortunately, assessing the sensitivity of the results to this assumption is fairly straightforward and we find that for moderately sized deviations from this assumption that our conclusion does not change. However, large changes in the fraction of black drivers between daylight and darkness at a fixed clock time can bias the conclusions and mask evidence of racial profiling. Also, we noted earlier that stops are substantially underreported in this dataset. The methodology is not sensitive to differences in reporting rates amongst the races. Specifically, if reporting rates for black and non-black drivers do not vary between daylight and darkness, then this is sufficient for underreporting not to affect the estimate of the racial profiling effect. However, if officers not reporting their stops are also the officers with the greatest race bias then this method will understate the extent of racial profiling. Grogger and Ridgeway (2004) offer details of this property.

The accuracy of the racial profiling estimate presented here is contingent on some important assumptions. However, the various benchmarking approaches utilized elsewhere also rely on assumptions. For example, they might require difficult to satisfy assumptions like the age-adjusted race distribution reported in the census must match the at-risk driving population. Even expensive traffic surveys still must deal with issues of differential driving behavior and differential exposure to law enforcement. We believe the analysis based on variation in natural lighting aims directly at the root question, whether knowing the race of the driver influences whether an officer makes a vehicle stop.

9.5 ANALYSIS OF RACE BIAS IN POST-STOP ACTIVITY

In this section we examine whether officers engage in racial profiling in the following post-stop activities:

- issuing warnings;
- issuing citations;
- making arrests;
- affecting the duration of the stops; and
- conducting searches.

As with the analysis of the decision to stop, these results are sensitive to the nature of the substantial number of unreported stops.

To determine whether there is a race bias in post-stop outcomes, we attempt to isolate the effect of race from all other factors that may contribute to differences in the way officers treat drivers. For example, the following factors, referred to as "confounding factors," may independently or jointly contribute to how officers interact with the drivers they stop:

- the location of the stop;
- the time of the stop;
- whether the driver is an Oakland resident;

- the age of the driver;
- the reason for the stop; and
- the driver's sex.

In our analysis, we attempted to match drivers from different races on all of the above factors and then compared the outcomes of the matched drivers' stops. The purpose of the "matching" was to ensure that the differences in stop outcomes were the result of race and not the result of one of the foregoing factors. We matched drivers by using a statistical matching method known as "propensity score adjustment" which is discussed in section 9.5.1.

One limitation of the propensity score adjustment method is that it only allows us to correct for differences in *observed* features of vehicle stops. It does not allow us to correct for unobserved features such as whether the driver was aggressive or whether the officer summoned a K9 unit to the scene. Therefore, other important variables may exist which are not captured in Oakland's stop form and, consequently, the existence of there unobserved variables may bias the results of our analysis.

In order to isolate the effect of race, we cannot naïvely compare the post-stop activities across the race groups. Just as law enforcement effort plays a role in which race group is at most at risk for police to stop, post-stop police practice may vary by neighborhood. In high crime neighborhoods, police may approach vehicles more cautiously regardless of the driver's race. They may be more likely to pat search for weapons regardless of race. They may be more thorough in checking out the vehicle registration and driver's license, might have a longer list of recent suspect descriptions that the stopped driver may match, and may be more likely to develop probable cause, all regardless of race. As a result, the stop location may influence all of the measured post-stop activities even in the absence of a race bias. When black and non-black drivers drive in different neighborhoods we must adjust for differences in post-stop activity that is attributable to the location differences.

Variation in post-stop practices across neighborhoods, "neighborhood profiling," could itself be a reasonable community concern even if within each neighborhood police officers apply their practices equitably to black and non-black drivers alike. To combat race disparities, we must first determine whether disparities are due to the biases of individual officers or policies at the department level. Police executives can manage some of these neighborhood profiling effects by, for example, reallocating police, instituting policies on the length of stop, and training on the use of pat searches. While a race bias in individual officers is difficult to correct, police and neighborhoods can negotiate the level of vehicle enforcement that they wish to have. In this analysis, however, we assume that if officers handle black and non-black drivers equitably within each neighborhood no race bias exists.

Location is one of several non-race factors for which post-stop activity might legitimately vary. We also adjust for the driver's sex, age, time of the stop, whether the driver is an Oakland resident, and the reason for the stop. One may reasonably question whether post-stop activity should vary by any of these characteristics. The goal of this section is to isolate the effect of race and, therefore, the analysis takes all other things into account. A separate debate can consider, for example, whether consent search rates should vary by sex or neighborhood or driver's age, or whether they should be permitted at all. Propensity score analysis (Rosenbaum and Rubin, 1983) is a transparent analytical tool for adjusting for the confounding factors to isolate the effect of race on post-stop activity.

Section 9.5.1 is technical and not essential for understanding the process. Section 9.5.2 shows that the method equalizes the groups on the confounding factors and proceeds from there.

9.5.1 PROPENSITY SCORE ADJUSTMENT (TECHNICAL)

This section is technical and not essential for understanding the process. Section 9.5.2 discusses the purpose of the method, that it equalizes the groups on the confounding factors and proceeds from there.

On October 30, 2003 at 3:25pm police stopped a black, male driver in his 20s in East Oakland for a moving violation. The driver received a citation but the stop lasted more than 20 minutes. In such a situation, ideally we wish to know what would have happened had this driver been white. We wonder if the stop would have been shorter or if the officer would still issue a citation. While we cannot make the direct comparison between what really occurred and the counterfactual world in which this driver is white, we can try to locate white drivers situated similarly to this black driver and examine what happened to them. These white drivers would be stopped in the same area, at the same time, for the same reason, and have the same age, sex, and Oakland residence status. A comparison in post-stop activities between the observed black driver and the matched white drivers would indicate the effect of race if the stop form contains all of the essential information for determining the post-stop activity.

The analysis presented here follows this model locating similarly situated comparison drivers for the collection of black drivers. We then average the differences between what actually occurred and our best estimate of what would have occurred had the driver not been black. Even though the presentation so far has focused on the black drivers, we can also address whether white drivers (driving in the locations that they usually drive in and stopped for the reasons that officers stop white drivers) would be treated differently had they been Hispanic or black. For the subsequent description of the methodology we will focus on finding matches for a "target" group. The reader can initially think of the target group as stopped black drivers but these can be drivers of any particular race.

To find the comparison group we rely on a technique know as *propensity score weighting* (Rosenbaum, 1987). Propensity score weighting "upweights" comparison drivers with stops that are similar to drivers in the target group. At the same time it "downweights" comparison drivers with stops that are dissimilar to those stops in the target group. The amount of weighting depends on the comparison stop's similarity to the stops involving the target group. For the target group we can compute, for example, the average stop duration. To estimate the stop duration for comparison drivers similarly situated we compute the *weighted* average stop length of the comparison group, calculated as the sum over stops in the comparison group of the propensity weight times the stop length divided by the number of observations in the comparison groups. The effect due to race is the difference between the two figures.

The remaining detail involves the computation of the correct weights for the comparison stops so that there are no systematic differences in observed characteristics between those stops and the stops involving drivers in the target group except for race. The propensity score for a particular vehicle stop is the percentage of stops with the same characteristics as the observed stop that involved a target group driver. Stops with propensity scores near 1 have features (e.g. location, stop reason) that are characteristic of target group drivers. Stops with propensity scores near 0 likely

occur in places where officers rarely stop drivers from the target group. We will denote the propensity score for stop i as p_i. Wooldridge (2001) notes that weighting observations in the comparison group with $p_i/(1-p_i)$ will match all of the stop characteristics of the comparison group with the characteristics of the target group. That is, the percentage of men in the target group will match the *weighted* percentage of men in the comparison group. The percentage of drivers in the target group stopped in downtown will match the *weighted* percentage of drivers in the comparison group stopped downtown. The two groups will even match on multiple factors simultaneously like the percentage of male drivers stopped in downtown. Table 4 summarizes the formulas needed for the propensity score analysis.

Table 4: Formulas for propensity analysis. N_t is the number of stops of the target group. y_i represents the outcome relating to stop i (stop duration or a citation indicator). w_i is the propensity weight. $p(\mathbf{x}_i)$ is the fraction of stops with characteristics $\mathbf{x}_i$ that involved drivers of the target group.

Description	Formula
Propensity weights for the comparison stops	$w_i = \frac{p(\mathbf{x}_i)}{1-p(\mathbf{x}_i)}$
Average outcome for the target group	$\frac{\sum_{i \text{ in target}} y_i}{N_t}$
Propensity weighted outcome for the comparison group	$\frac{\sum_{i \text{ in comparison}} w_i y_i}{\sum_{i \text{ in comparison}} w_i}$

To estimate the propensity score we use a form of logistic regression (Hosmer and Lemeshow, 2000). Logistic regression is a standard statistical tool used to estimate the probability of a particular outcome (e.g. that the driver is a member of the target group) from a set of observation features, what we have denoted as $\mathbf{x}_i$. McCaffrey, Ridgeway, and Morral (2004) describe an evaluation of a drug treatment program using the same methodology used in this study to estimate the propensity scores using boosted logistic regression (Friedman 2001, Ridgeway 2004). The interested reader can refer to that article for the exact details of the logistic regression model fitting. Of primary importance for whichever method we use to estimate the propensity score, is whether the resulting weights create a comparison group that is similar to the target group.

The benefit of the propensity score based method of analysis is that we can examine the observed characteristics of the target group and weighted comparison group and note that they are matched on important stop features before we proceed to assess differences in post-stop activities. It gives all parties involved a chance to determine whether the analysis will present an apples-to-apples comparison of, for example, black and non-black drivers. Before seeing the results we can ask whether there are additional factors, not among those used in developing the propensity scores, on which the target and comparison stops should also match. If the stop form captures these items then they may be included in the adjustment. If not, we can consider the merits of including these

additional factors on the next version of the stop form.

For a propensity score analysis to accurately estimate the race effect, the propensity score model must include all factors associated with both race and the outcome of interest. That is, if our outcome of interest was duration of stop and we only included age in the propensity score model, then we have neglected to adjust for the reason for the stop. Figure 4 showed that there are differences in the reason for the stop by race and we can reasonably expect that stops for criminal violations will last longer than other kinds of stops. Failing to include the reason for the stop will result in an analysis that is unable to differentiate between an effect due to race and an effect due to the reason for the stop.

9.5.2 DEVELOPMENT OF THE COMPARISON GROUP FOR BLACK AND WHITE DRIVERS

With the propensity score methodology in hand we can construct a comparison group for a specific target group of drivers. Here we describe the generation of the comparison group for black drivers. The comparison group will match the black drivers on the location of the stop, the driver's sex, age, whether the driver is an Oakland resident, and the reason for the initial stop. For stop location we collapsed the city's policing beats according to Table 5.

Table 5: Collapsing of beats to larger regions of Oakland

Beats	Region
1 through 8	Downtown
9 through 12	North
13	Hills
14 through 19	West
20, 21, 23, 24	Midtown
22, 25	South hills
26 through 35	East

For this analysis we focused only on stops made for dangerous and non-dangerous moving violations and mechanical/registration violations.[1] Utilizing only moving violations and mechanical/registration violations focuses the analysis on the routine traffic stops and avoids mixing these stops with those involving a much more involved decision process.

We will present several comparisons. The first comparison matches black drivers to similarly situated non-black drivers. This comparison is important since, by many measures, the black drivers bear the greatest burden of vehicle stops. Naturally, the non-black driver comparison group will be a mixture of Hispanic, white, Asian, and the other race groups. The mixture is in proportion to how

[1] These kinds of stops make up the majority of vehicle stops (95%). We did not include probation/parole stops (0.3%) since it is difficult to isolate a race bias from the other explanations for these stops including police having additional rights to stop probationers and parolees and the overrepresentation of black residents in this population. We did not include stops for felonies (1.5%) and misdemeanors (0.9%) since these likely were stops for which searches involve little officer discretion. Lastly, we did not include stops for be-on-the-lookout (BOLO), which represent (0.6%) of the stops. These stops require a separate analysis since BOLO stops represent a unique situation. We might expect some of these stops to involve almost no post-stop activity once the officer assesses that they have not stopped the right suspect, while others can be particularly long if they stop the right suspect.

frequently drivers in those race groups are involved in stops similar to the stopped black drivers. In our description of the comparison group we show that all the races are well-represented with white drivers comprising 29% and Asian drivers comprising 18%.

The second comparison matches black drivers exclusively to similarly situated white drivers. Detecting differences in this comparison is less powerful than the black/non-black comparison since there are relatively few stops with white drivers resembling the stops involving black drivers.

The third comparison matches white drivers to non-white drivers. This last comparison addresses whether racial profiling is really taking place in the places and at the times when white drivers drive. As with the black/non-black comparison, the non-white comparison group contains a mixture of drivers of the other races in proportion to the frequency at which their stops resemble the stops involving white drivers.

Table 6 shows that the propensity weighting developed for creating a set of comparison drivers from the set of non-black drivers balances the samples on many important stop characteristics. The first column indicates the variables. The table lists the variables in decreasing order of importance in the propensity score model. This essentially measures the magnitude of the difference between the black and non-black drivers on each variable before adjusting. The second column shows the percentages for the black driver population. The third column shows the weighted percentages for our constructed comparison sample. Critical to making a valid comparison, the weighted percentages for the comparison group are uniformly close to the percentages for the black drivers. Having these percentages close indicates that the stops in the comparison group are nearly identical to the stops involving black drivers, race being the only factor differing between the groups by design. Note that the groups may still differ on an unobserved factor, but at a minimum we know that the groups are similar on these factors. The fourth column shows the raw percentages for the non-black driver sample. In particular we see that very few non-black drivers are involved in stops in east Oakland. Also non-black drivers are three times more likely to be stopped in the hills and almost twice as likely to be stopped for dangerous moving violations than black drivers. Critically the weighted sample has been constructed to downweight non-black drivers stopped in the hills and upweight non-black drivers stopped in east Oakland. Similarly, non-black drivers ages 18-29 are substantially upweighted so that the age distribution of the comparison sample is closer to that of the black driver sample. Non-black drivers are most likely stopped during the evening rush hour whereas black drivers have many stops after 8pm and still many after midnight. We did not include an indicator of whether the day of the stop was a weekend, yet even the percentage stopped on a weekend now matches across the two groups.

Table 6: Assessment of the comparison driver sample for black drivers derived from the propensity weighting

	% Black drivers *N*=3,703	% Non-black drivers (weighted) *N*=2,089	% Non-black drivers (unweighted) *N*=3,033
Region			
Downtown	31%	29%	27%
East	32%	30%	14%
Hills	1%	1%	3%
Midtown	11%	13%	21%
North	9%	9%	8%
South hills	3%	3%	6%
West	14%	15%	21%
Time of day			
12:00am-4:00am	16%	13%	7%
4:00am-8:00am	4%	4%	4%
8:00am-12:00pm	17%	17%	21%
12:00pm-4:00pm	20%	23%	28%
4:00pm-8:00pm	24%	25%	26%
8:00pm-12:00am	20%	18%	13%
Resident	76%	72%	64%
Age			
Under 18	3%	3%	3%
18-29	47%	45%	38%
30-39	22%	25%	26%
40+	28%	27%	33%
Reason			
Mechanical/Registration	26%	23%	16%
Moving (dangerous)	22%	26%	37%
Moving (non-dangerous)	52%	52%	47%
Male	75%	76%	74%
Weekend*	29%	30%	27%

* Weekend was not used in the propensity score model, but balances nonetheless across the two groups.

According to Table 6 the target and comparison groups of interest balance across many important features of the stops. This balance is the critical step when using propensity score techniques. Race is the one feature on which they differ by design. To create a matched set of non-black drivers the propensity scoring slightly downweighted Asian and white drivers and slightly

upweighted Hispanic drivers as shown in Table 7. This table assures us that this comparison has not developed simply into a black and Hispanic comparison since white and Asian drivers still comprise more than 50% of the comparison sample.

Table 7: Race distribution of the comparison group for black drivers

Race	Target group	Weighted comparison group	Unweighted comparison group
Asian	0%	18%	22%
Black	100%	0%	0%
Hispanic	0%	40%	33%
White	0%	29%	33%
Other	0%	13%	12%

We also created a comparison group for black drivers containing only white drivers. However, we were only able to find about 305 similarly situated white drivers so statistical power may be small. Table 8 shows that after weighting, the black drivers group was still slightly more likely to be stopped in East Oakland, but the difference is not statistically large enough to cause concern.

Table 8: Assessment of the comparison sample of white drivers for a target sample of black drivers derived from the propensity weighting

	% Black drivers *N*=3,703	% White drivers (weighted) *N*=305	% White drivers (unweighted) *N*=988
Region			
Downtown	31%	34%	30%
East	32%	28%	8%
Hills	1%	1%	7%
Midtown	11%	11%	8%
North	9%	9%	16%
South hills	3%	3%	11%
West	14%	15%	20%
Time of day			
12:00am-4:00am	16%	16%	5%
4:00am-8:00am	4%	4%	4%
8:00am-12:00pm	17%	17%	26%
12:00pm-4:00pm	20%	22%	30%
4:00pm-8:00pm	24%	23%	25%
8:00pm-12:00am	20%	18%	11%
Resident	76%	72%	52%
Age			
Under 18	3%	1%	1%

18-29	47%	46%	23%
30-39	22%	23%	26%
40+	28%	30%	50%
Reason			
Mechanical/Registration	26%	27%	13%
Moving (dangerous)	22%	24%	40%
Moving (non-dangerous)	52%	49%	47%
Male	75%	75%	64%
Weekend*	29%	29%	21%

* Weekend was not used in the propensity score model, but balances nonetheless across the two groups.

We have designed the previous two comparison groups so that we match stops involving black drivers to stops involving non-black drivers or white drivers. This focuses the racial profiling analysis on the areas that officers stop black drivers. The possibility remains that the race disparity could be greater in the areas that are more common to white drivers. To address this question we constructed a set of stops involving non-white drivers that match the characteristics of stops involving white drivers. We located effectively 1,727 stops of non-white drivers that were similarly situated to the stops of white drivers. Table 9 shows how the propensity weighted comparison group is well matched to the target group. White drivers are much more likely to be stopped in the hills than non-white drivers are, but the propensity score method compensates for that by locating non-white drivers that were stopped in the hills. The propensity score weighting also drastically downweighted stops between midnight and 4am as well as stops of male drivers.

Table 9: Assessment of the comparison driver sample for white drivers derived from the propensity weighting

	% White drivers *N*=988	% Non-white drivers (weighted) *N*=1,727	% Non-white drivers (unweighted) *N*=5,748
Region			
Downtown	30%	31%	29%
East	8%	8%	27%
Hills	7%	6%	1%
Midtown	8%	8%	17%
North	16%	16%	7%
South hills	11%	11%	3%
West	20%	20%	16%
Time of day			
12:00am-4:00am	5%	5%	13%
4:00am-8:00am	4%	4%	4%
8:00am-12:00pm	26%	25%	18%
12:00pm-4:00pm	30%	30%	22%
4:00pm-8:00pm	25%	23%	25%
8:00pm-12:00am	11%	12%	18%
Resident	52%	51%	74%
Age			
Under 18	1%	1%	3%
18-29	23%	23%	46%
30-39	26%	26%	23%
40+	50%	50%	27%
Reason			
Mechanical/Registration	13%	13%	23%
Moving (dangerous)	40%	41%	27%
Moving (non-dangerous)	47%	46%	50%
Male	64%	64%	76%
Weekend*	21%	24%	29%

* Weekend was not used in the propensity score model, but balances nonetheless across the two groups.

Again race is the one feature on which the target and comparison groups differ by design. To create the matched set of non-white drivers the propensity scoring downweighted stops involving black drivers in order to reduce imbalance in the stop location and time of the stop. Table 10, which shows the resulting race distribution of the comparison group, indicates that stops involving black drivers, at 58% of the sample, still hold the majority of the stops in the comparison group.

Table 10: Race distribution of the comparison group for white drivers

Race	Target group	Weighted comparison group	Unweighted comparison group
Asian	0%	18%	11%
Black	0%	58%	64%
Hispanic	0%	13%	18%
White	100%	0%	0%
Other	0%	11%	6%

The analysis of post-stop outcomes will utilize the comparison groups formed in this section. The factors adjusted for here will not be sufficient for analysis of the duration of the stop. Section 9.5.4 shows that we can further adjust for factors such as the stop outcome in order to equalize the groups on these factors.

Having created two comparison groups for stops involving black drivers and a third comparison group for stops involving white drivers, we can turn to assessing disparities in post-stop activity.

9.5.3 CITATIONS, WARNINGS, AND ARRESTS

Twenty-seven percent (27%) of the vehicle stops resulted in the officer issuing the driver a warning and sixty-seven percent (67%) of the vehicle stops resulted in the officer issuing the driver a citation. Only a small percentage of stops resulted in an arrest.[1] The following analyses exclude stops resulting in arrest.

9.5.3.1 Analysis of citation rates using matched driver samples

When drivers of different races are matched by the factors listed in Table 6, the data reveal that officers cite drivers of different races at different rates. After being stopped by an officer:

- Black drivers were cited 68% of the time,
- Non-Black drivers were cited 72% of the time, and
- White drivers were stopped 65% of the time.

The 4% difference between the citation rate of black drivers and the citation rate for non-black drivers is statistically significant (p-value=0.001). The likelihood of a 4% difference in the citation rates being the result of a chance or a sampling variation is less than 0.1%. Therefore, it is safe to say that black drivers are cited less frequently than non-black drivers.

[1] Arrests are slightly more likely to result from mechanical and registration violations (7.4%) than moving violations (4.0%). That difference of 3.4% in arrest rates is statistically significant and not explainable by chance alone. That is, in the long run it seems possible that mechanical and registration stops will be more likely to produce arrests than moving violations. Such a finding might inform the use of mechanical and registration stops as a tool to produce arrests if such practices are being questioned.

The 3% difference between the citation rate for black drivers and the citation rate for white drivers, however, is not statistically significant. This result could be the product of an idiosyncrasy within this particular dataset. Therefore, we are unable to confidently conclude that white drivers are cited less frequently than black drivers. Table 11 below provides the 95% confidence intervals for each racial group.

Table 11: Propensity score estimates of the effect of being black on citation rate. Analysis excludes those stops resulting in arrests.

Group	Citation rate	95% confidence interval
Black drivers	68%	(66.6%, 69.7%)
Non-black drivers (weighted)	72%	(70.3%, 74.5%)
White drivers (weighted)	65%	(59.7%, 71.1%)

These findings potentially imply that either police are slightly more hesitant to cite black drivers or that some of the stops involving black drivers were of a level of severity unlikely to result in a citation. We have little detail on the exact reasons for the stops. It is possible that black drivers stopped for non-dangerous moving violations, for example, are the kinds of non-dangerous violations for which non-black drivers are not even detained. Police and residents need to debate and agree on whether the target and comparison groups are adequately matched to one another on all the important factors. Once the parties agree that the target and comparison groups represent an "apples-to-apples" comparison, the subsequent analysis and findings leave little to debate.

Stops involving non-black drivers (unweighted) resulted in citations 79% of the time. Had we not adjusted for factors such as time of the stop and location of the stop we would have concluded that black drivers are much less likely to be cited than non-black drivers. This implies that much of the difference in the raw citation rates, comparing 68% to 79%, is due to the factors in Table 6. In addition to providing a valid assessment of citation rates, this also demonstrates the importance of adjusting for factors other than race that might explain differences in post-stop activity.

We repeated this analysis comparing the citation rate of white drivers to the citation rate of non-white drivers. White drivers are slightly more likely to be cited than non-white drivers (see Table 12). The difference, however, is not statistically significant (p-value=0.52) and, therefore, we are unable to confidently conclude that white drivers are more likely to be cited than non-white drivers.

Table 12: Propensity score estimates of the effect of being white on citation rate. Analysis excludes those stops resulting in arrests.

Group	Citation rate	95% confidence interval
White drivers	79%	(76.6%, 81.8%)
Non-white drivers (weighted)	78%	(76.3%, 80.1%)

As we examine these differences in the citation rates across races, it is unclear what the disparities might imply. If officers cite non-white drivers more frequently than white drivers one might conclude that officers overly cite non-white drivers. If officers cite non-white drivers less

frequently this could be interpreted as officers stopping non-white drivers for reasons for which citations are not usually given. In the latter scenario, officers might be more inclined to use minor traffic violations with black drivers as a pretext to question the driver.

9.5.3.2 Conclusions

We can conclude that officers cite black drivers less frequently than they cite non-black drivers. However, we are unable to conclusively determine whether officers cite white drivers less frequently than black drivers or cite white drivers more frequently than non-white drivers.

As with the other analyses presented, this analysis of citation rates may be sensitive to the underreporting of stops in Oakland.

9.5.4 DURATION OF THE STOP

Presumably, two drivers stopped for the same reason with the same stop outcome should be detained for roughly the same amount of time. Of course, there will be some natural variation in the length of the stop. In the absence of racial bias, on average we expect similarity in stop duration.

We augmented the propensity score adjustment for some of the aspects of post-stop activity. In particular we have included an indicator of whether a search occurred (we will separately analyze the decision to search next) and the outcome of the stop. We include stops that result in arrests but made sure that the comparison group had similar arrest rates. The propensity score model still includes all of the other variables, like stop location and time of the stop, and preserves balance of the target percentages and the weighted comparison percentages for those factors. Table 13, Table 14, and Table 15 show that the target group and weighted comparison group have nearly equal percentages on the two additional factors for black and white driver target groups. Note that the effective sample sizes for the matched groups are smaller than when we did not adjust for whether the officers searched the vehicle or the stop outcome. To improve the comparison group matching the propensity score method had to further downweight those stops that did not also match on these two stop characteristics, reducing the number of suitable comparison stops. After weighting, black drivers are still slightly more likely to be searched than their comparison group but the difference is not substantial.

Table 13: Assessment of the comparison driver sample for black drivers on the additional items included in the analysis of stop duration

	% Black drivers N=3,703	% Non-black drivers (weighted) N=1,544	% Non-black drivers (unweighted) N=3,033
Searched	27%	24%	11%
Stop outcome			
Arrest	6%	6%	3%
Citation	64%	64%	77%
No action/Warning	30%	30%	20%

Table 14: Assessment of the comparison driver sample of white drivers for a black driver target group on the additional items included in the analysis of stop duration

	% Black drivers N=3,703	% White drivers (weighted) N=272	% White drivers (unweighted) N=988
Searched	27%	23%	6%
Stop outcome			
Arrest	6%	7%	1%
Citation	64%	61%	78%
No action/Warning	30%	31%	20%

Table 15: Assessment of the comparison driver sample for white drivers on the additional items included in the analysis of stop duration

	% White drivers N=988	% Non-white drivers (weighted) N=1,707	% Non-white drivers (unweighted) N=5,748
Searched	6%	7%	22%
Stop outcome			
Arrest	1%	2%	5%
Citation	78%	78%	68%
No action/Warning	20%	21%	26%

9.5.4.1 Results of the analysis of stop duration

Using the propensity score adjustment method, Table 16, below, compares the stop durations of black drivers and non-black drivers. Black drivers were detained for longer periods of time than non-black drivers (p-value<0.0001). There is less than a 0.01% chance that this conclusion is the result of sampling variation in the dataset.

Table 16: Propensity score estimates of the effect of being black on stop duration. The 95% CI columns show the 95% confidence intervals for the estimates.

Stop duration	% Black drivers N=3,703	95% CI	% Non-black drivers (weighted) N=1,544	95% CI	% Non-black drivers (unweighted) N=3,033
0-9 minutes	47%	(45.4%, 48.6%)	53%	(51%, 56.1%)	66%
10-19 minutes	38%	(36.2%, 39.3%)	34%	(31.4%, 36.3%)	26%
20-30 minutes	10%	(9.2%, 11.1%)	8%	(6.4%, 9.8%)	5%
Over 30 minutes	5%	(4.4%, 5.8%)	4%	(3.2%, 5.6%)	3%

Table 17 compares the stop times of black drivers and similarly situated white drivers. Black drivers were detained for longer periods of time than white drivers (p-value=0.013). Among stopped black drivers, 47% of the drivers were detained for less than ten minutes. Among the stopped white drivers, 55% of the drivers were detained for less than 10 minutes.

Table 17: Propensity score estimates of the effect of being black on stop duration compared with similarly situated white drivers. The 95% CI columns show the 95% confidence intervals for the estimates.

Stop duration	% Black drivers N=3,703	95% CI	% White drivers (weighted) N=272	95% CI	% White drivers (unweighted) N=988
0-9 minutes	47%	(45.4%, 48.6%)	55%	(49.5%, 62.1%)	74%
10-19 minutes	38%	(36.2%, 39.3%)	31%	(25.4%, 36.6%)	23%
20-30 minutes	10%	(9.2%, 11.1%)	9%	(4.3%, 14.5%)	3%
Over 30 minutes	5%	(4.4%, 5.8%)	4%	(0.3%, 7.2%)	1%

Table 18 compares the stop times for white drivers and non-white drivers. When white drivers are compared to similar non-white drivers, 74% of the white drivers are detained for less than 10 minutes while 67% of non-white drivers are detained for less than 10 minutes. White drivers appear to be overrepresented in the 0-9 minute stop duration class (p-value=0.0002). Stop durations over 10 minutes appear more likely to be from the non-white comparison group.

Table 18: Propensity score estimates of the effect of being white on stop duration. The 95% CI columns show the 95% confidence intervals for the estimates.

Stop duration	% White drivers N=988	95% CI	% Non-white drivers (weighted) N=1,707	95% CI	% Non-white drivers (unweighted) N=5,748
0-9 minutes	74%	(70.7%, 76.3%)	67%	(65.2%, 69.4%)	53%
10-19 minutes	23%	(20.1%, 25.4%)	27%	(25.4%, 29.3%)	34%
20-30 minutes	3%	(1.8%, 3.9%)	4%	(3.2%, 4.6%)	8%
Over 30 minutes	1%	(0.3%, 1.5%)	2%	(1.1%, 1.7%)	4%

9.5.4.2 Conclusions

In conclusion, having adjusted for factors such as location and time of stop and other important stop outcomes (such as whether a search occurred), black drivers were detained for longer periods of time than similarly situated non-black and white drivers. These results only apply to reported stops and may understate or overstate the effect of race. This analysis suggests that 6% to 8% of the stops involving black drivers should not be lasting more than 10 minutes since stops with similar characteristics involving non-black drivers are generally take less time.

9.5.5 ANALYSIS OF RACE BIAS IN THE DECISION TO SEARCH

As with the analysis of citation rates and stop lengths, investigating racially biased policing in the decision to search involves a comparison of search rates across the race groups. Unlike the analysis of citation rates, we can also evaluate the effectiveness of the decision to search by assessing the recovery of some form of contraband (firearms, drugs, etc.).

To begin, we investigate whether the percent of black drivers searched is the same as the percentage of Hispanic drivers searched, is the same as the percentage of white drivers searched and so on. A proper comparison, however, is not quite so simple.

The decision to search can be associated with race even if the officer is not motivated by racial bias. Where exposure and behavior were the essential components of the at-risk population for analysis of stops, for analysis of post-stop actions we need to adjust for the non-race based features of the driver that affect officer decisions and differences in neighborhood driving patterns. Failure to adjust for these important factors can lead to conclusions of racial bias when differences are actually attributable to differences in probation/parole rates and local law enforcement practices that are not based on race.

9.5.5.1 Low discretion searches

The Oakland Police Department has a policy of searching all stopped drivers who are on probation or parole.[1] Most of the people who are on probation or parole in Oakland are black. Consequently, stopped black drivers may be searched more frequently than stopped drivers of other races because of the police department's policy to search all stopped drivers on probation or parole. This result, however, does not indicate racial bias on the part of the officers.

Similarly, stops resulting in the arrest of the driver or an impounding of the vehicle will also result in an automatic search of the driver and/or vehicle. Thus, if non-white drivers are more likely to be arrested, have warrants, or have their vehicles impounded, then stops involving non-white drivers are more likely to result in a search than stops involving white drivers. Apparent disparities in search rates do not necessarily indicate the existence of racial profiling.

Additionally, in Oakland, there are some neighborhoods where search rates are much higher for all drivers, both white and non-white. Those neighborhoods correspond to the predominantly non-white, high-crime areas of the city. Solely looking at search rates aggregated to the city level may produce figures indicative of a disparity where much of that disparity may be due simply to differences in driving patterns.

9.5.5.2 High discretion searches

There are other searches that involve high discretion on the part of the officer. For example**,** officers may ask the driver for consent to search. If officers are racially biased, such high discretion searches are most likely to exhibit disparities. Unfortunately, the analysis of consent searches is somewhat confounded because the data collection form only records those cases in which the officer asks *and* the driver gives consent. Disparities may be due to differences in which officers ask or differences in which drivers accept consent searches.

To ensure officer safety, officers may pat search a vehicle occupant to check for weapons. Such searches involve high officer discretion, may depend on the type of neighborhood the officer patrols, and may be prone to racial bias if officers feel more at risk with a black or Hispanic driver, all else being equal.

Probable cause searches may or may not involve much officer discretion. Officers claim that probable cause searches are not prone to racial bias since these involve clear signs of wrongdoing (driver is drunk, contraband in plain view, etc.) so that any officer in the same position would conclude that an arrest is almost certain to follow.

9.5.5.3 Description of the searches

Officers searched 22% of the drivers they stopped. Among the 1,657 drivers who were searched, the police found:

- narcotics in 125 cases;

[1] As a condition of their release, individuals on probation or parole consent to be searched at anytime and anywhere.

- § firearms in 16 cases; and
- § other evidence in 114 cases.

Note that if unreported stops involve searches with characteristics that differ greatly from the reported searches than these and other figures in this section will not represent the state of searches in Oakland. Many of the searches were low discretion searches in that the officers were required to search the driver because of the situation from which the search arose. The low discretion searches specifically arose because:

- § a vehicle occupant was on probation or parole (28%);
- § the search was incident to an arrest (28%); or
- § the search pursuant to an inventory search (15%).

The high discretion searches arose because:

- § the driver gave his/her consent (8);%,
- § the officer wanted to pat search for weapons (10%);

In addition the officers recorded probable cause 10% of the time and two of the searches were based on search warrants.

Figure 7 shows the number of searches broken down by race and the basis for the search. Clearly the greatest burden of the searches falls on vehicles with black drivers. Black drivers are involved in 75% of the searches. However, most of the searches of vehicles with black drivers were based on reasons with little officer discretion, 34% probation/parole, 27% incident to arrest, and 13% inventory searches. The main concern should lie in the large number of consent (8%) and pat searches (8%) and, perhaps, probable cause searches (10%).

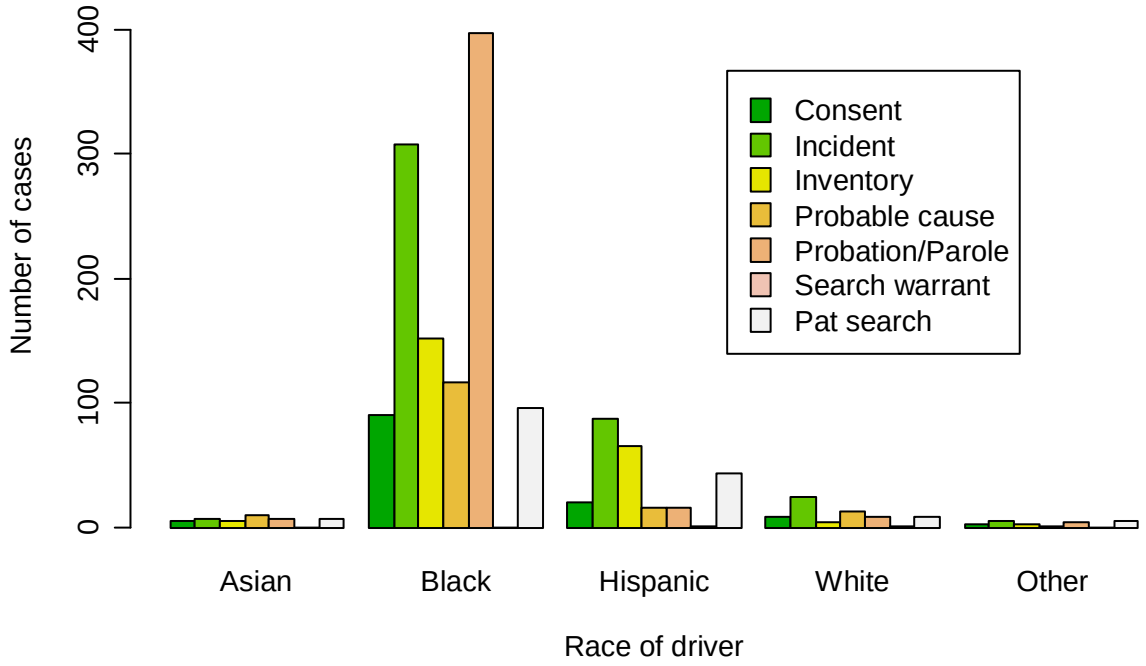


Figure 7: The number of searches by race and the reason for initiating the search

Making comparisons directly from Figure 7 does not account for differences in the reasons for the stop, the conditions in the locale of the stop, and age and sex differences of black and non-black drivers. Since an officer's decision to pat search, ask for consent, or develop probable cause can heavily depend on the conditions of the neighborhood in which the stop occurs, accounting for the neighborhood is an important component.

9.5.5.4 Analysis of search rates using matched driver samples

We compared consent search rates, pat search rates, and probable cause search rates for black drivers with the rates for similarly situated non-black and white drivers. The comparison groups used here are the same as those described in Section 9.5.2 and used in the analysis of citation rates in Section 9.5.3.1. Table 19 shows the results of the analysis.

Pat searches. Officers conducted pat searches of black drivers and non-black drivers at equal search rates. However, officers pat searched white drivers much less frequently than they pat search black drivers. Less than 1% of the stopped white drivers were pat searched.

Table 19: Propensity score estimates of the effect of being black on the probability of the occurrence of various high discretion searches

	No search	Pat search	Consent search	Probable cause
% Black drivers	91.9%	2.7%	2.2%	3.2%
95% CI	(90.9%, 92.9%)	(2.1%, 3.2%)	(1.7%, 2.7%)	(2.6%, 3.9%)
% Non-black drivers (weighted)	94.4%	2.6%	1.6%	1.4%
95% CI	(93.2%, 95.6%)	(1.8%, 3.4%)	(0.9%, 2.2%)	(0.8%, 2.0%)
% white drivers (weighted)	96.5%	0.4%	1.7%	1.4%
95% CI	(94.2%, 98.7%)	(0.0%, 0.9%)	(0.0%, 3.7%)	(0.4%, 2.5%)

Consent searches. Officers were slightly more likely to search stopped black drivers than similarly situated non-black drivers and similarly situated white drivers although the differences are not statistically significant.

Probable cause searches. Stops involving black drivers were more than twice as likely to result in a probable cause search. Probable cause should imply that the officer had an articulable reason for the search believing that an arrest was imminent. However, further investigation of the dataset found that probable cause searches rarely result in an arrest (18%). The remainder resulted in citations (54%) or no action (28%). Probable cause searches should almost always result in an arrest. Either officers are incorrectly coding the search basis and stop outcome fields, or these searches are not up

to the standard of probable cause. Since black drivers bear most of the burden of probable cause searches the department needs to further investigate the use of probable cause as a search basis.

9.5.5.5 Analysis of hit rates

For assessing race bias in searches, we can compare the rates at which police recover contraband. That is, even though search rates may be high for a certain race group, if the searches regularly turn up contraband, a race bias is less likely to be at issue. On the other hand, policies regarding high discretion searches that place the greatest burden on black and Hispanic drivers and that rarely produce contraband need to be regularly revisited and reassessed for their value to public safety versus their contribution to police/community tension. Searches of vehicles with occupants on probation or parole produce no contraband 85% of the time, yet police and citizens will likely agree that there is no need for changing search policies for such drivers. However, out of the 128 consent searches seven (5%) resulted in a narcotics recovery and four (3%) resulted in other non-weapon, non-narcotics evidence of some kind. Since officers most frequently request consent searches of Black drivers (71% of consent searches) and Hispanic drivers (16% of consent searches), the value of such searches needs to be questioned.

There are too few searches of Asian and drivers of "other" races in the current dataset to compute hit rates. Officers only searched black, Hispanic, and white drivers in sufficient numbers to yield hit rate estimates. However, sample sizes were still too small for us to construct powerful tests of disparities. With additional data collection the hit rates for other races may be estimable.

Table 20: Hit rates (and 95% intervals) for black and Hispanic drivers

Search basis	Black		Hispanic		White	
	Hit rate	95% interval	Hit rate	95% interval	Hit rate	95% interval
Consent	12.5%	(6.8%, 22.1%)	5.9%	(1.4%, 27.3%)	16.7%	(3.7%, 57.9%)
Probable cause	40.2%	(31.2%, 49.9%)	33.3%	(15.2%, 58.7%)	37.5%	(13.7%, 70.1%)
Pat search	4.3%	(1.6%, 11.9%)	9.4%	(3.4%, 24.3%)	16.7%	(3.7%, 57.9%)

Table 20 shows the hit rates for black, Hispanic, and white drivers. In all categories and across all three races, the hit rates appear to be relatively close to one another. Since consent searches and pat searches are relatively rare, estimating the hit rate to any degree of accuracy is difficult. The 95% intervals shown in the table reflect the uncertainty in the hit rate estimate. The intervals for each race group overlap considerably with each other indicating that we cannot tell whether one rate is significantly different from another. Probable cause searches generally appear legitimate with relatively high rates of recoveries of some kind although this should be coupled with other stop outcomes as we already saw that probable cause searches do not often end in an arrest.

9.6 CONCLUSIONS FROM THE VEHICLE STOP DATA

A race bias can reveal itself at every stage of the vehicle stop process, including the decision to stop, the decision to cite or warn, the amount of time to detain the vehicle, and the decision to search. This report has dissected each of these stages of the decision process to assess the degree

that a race bias might present itself.

9.6.1 DECISION TO STOP

We examined two measures of the officers' ability to identify race in advance, officer reported advance knowledge of race and variation in natural lighting. Using the first measure we find that there may be a race bias in the decision to stop while using the second measure we find no evidence of a race bias.

Comparing the stops when officers could and could not identify the race in advance allows us to determine whether race visibility affects the stop decision. In race blind police practices, race visibility should have no effect (except for the handful of cases matching suspect descriptions). Yet the driver is more likely to be black when officers report knowing the race of the driver in advance. There are two explanations for this result, which makes interpretation complex. The first is that this is indicative of racial profiling. Race visibility increases a black driver's risk of being stopped. The second explanation is that black drivers might be more visible to officers. Officers may have difficulty distinguishing white, Hispanic, or Asian drivers at distances but can identify black drivers more easily. As a result, even if acting without a race bias, officers are more likely to indicate advance knowledge of the race when stopping a black driver.

We used variation in natural lighting to study the ability of officers to see the race of the driver in advance. In that analysis we found no evidence of a race bias in the reported stops, finding that black drivers composed the same fraction of stopped drivers regardless of the natural lighting available for officers to detect the driver's race in advance. This is evidence that those officers reporting stops are not involved in bias based policing practices as a group. Given the size of the department, a few officers may be involved in such practices and this analysis would not detect the resulting disparities in their relatively small contribution to the volume of stops.

Especially since the officers substantially underreport the stops, non-reporting officers could be involved in racial profiling. Without information on their stops this analysis cannot draw conclusions for the department as a whole. When the forms begin to consistently record the officer badge numbers we can augment this analysis with an officer level analysis that could become a part of the early warning system.

9.6.2 POST-STOP ACTIVITY

When evaluating citation rates we found that black drivers were slightly less likely to receive a citation when compared with similarly situated non-black drivers. This finding potentially implies that either police are slightly more hesitant to cite black drivers or that some of the stops involving black drivers were of a level of severity unlikely to result in a citation.

Race appears to have the strongest influence on the duration of the stop. Black drivers are much less likely to have stops lasting less than 10 minutes. It is possible that there are unmeasured factors that could legitimately explain the difference. The Oakland racial profiling task force needs to consider what factors could legitimately explain this difference. However, those additional factors themselves could be subject to a race bias as well, such as the decision to call out a K9 unit or call for additional information on a driver. It is also unclear how a race bias would cause stops to last longer especially given that citation rates are, if anything, less frequent for black drivers. Police may

be able to provide some explanation for these differences.

To assess whether there is differential treatment by race in the decision to search we created a comparison group of non-black drivers that were stopped in the same neighborhoods, for the same reason, and have similar age and sex. We found that officers engage in consent searches and pat searches of black drivers and non-black drivers at equal rates. Officers engage in consent searches of black drivers and white drivers at equal rates as well. However, black drivers are pat searched more frequently than white drivers

Officers seem more than twice as likely to conduct a probable cause search of a black driver than a non-black driver. In spite of this difference in search rates, the hit rate for probable cause searches of black drivers is 40%, potentially indicating that the probable cause searches are reasonable. However, only 18% of these searches actually resulted in an arrest casting doubt on either the officers' reporting of probable cause or on the reasons they conduct probable cause searches.

9.6.3 LIMITATIONS

There are some limitations to this study. First, all of the analyses address whether racial profiling is a department-wide practice. The data analyzed for this report did not associate an officer badge number with the vehicle stop. If a small number of officers were racially biased, the disparities that they cause would not likely appear in the analyses we have provided here.

Second, in order to use the natural variation in lighting, the analysis of the decision to stop examines only those stops occurring in the evening. The number of stops near sunrise was too small at this stage to complete the analysis. As a result, we can only estimate the racial profiling effect that might occur during evening hours. This analysis does assume that the race distribution of the at-risk population does not change drastically in a small interval of time near the transition from daylight to darkness.

Third, our analysis of post-stop activity is able to form target and comparison groups of similarly situated drivers, but we can only confirm their similarities on observed features of the stop. It is possible that important unobserved differences between the groups remain that could confound the estimates of the race bias effect.

Lastly, there appears to be substantial underreporting of stops, especially in the data from June through September. The analysis of the stop decisions as well as the post-stop activity may be sensitive to underreporting. If, for example, stops involving pat searches of black drivers are overrepresented in the unreported stops then the disparity we have reported underestimates the problem.

In summary, the substantial underreporting of stops prevents the findings in this report to be conclusive. Even among the reported stops there are some indications of racial disparities in post-stop activity, including differences in stop duration and frequency of pat searches that the Oakland racial profiling task force needs to consider and address.

10 CONCLUSIONS

The Department was extremely ambitious in its grant proposal as outlined in the ten objectives provided in this report. The Oakland Racial Profiling Task Force, which was comprised of a group of highly dedicated, committed and sincere professionals, worked tirelessly to accomplish these goals.

With few exceptions, the Task Force accomplished all of the goals outlined in the grant. Furthermore, the processes developed by the group should serve as a national model. In short, the group demonstrated that the police and the community can work together on any issue, even an issue such as racial profiling, and do so in an effective manner.

Through our efforts, the Oakland Task Force has accomplished the following:

1. Formed a diverse, representative task force
2. Provided extensive training to task force members
3. Conducted two-day retreat for the task force at SCANTRON
4. Conducted monthly task force meetings for over a year
5. Identified 24-fields of data to collect
6. Created data collection form
7. Purchased data collection technology & implement data collection program
8. Conducted survey of community
9. Conducted survey of Departmental personnel
10. Hosted Bay Area Workshop on Racial Profiling
11. Conducted Town Hall Meeting
12. Defined racial profiling
13. Developed racial profiling policy
14. Developed effective benchmarking process
15. Analyzed over 7,000 stops

Based on the experiences of our project, the task force makes the following recommendations for agencies that decide to collect stop-data and do so in a collaborative, comprehensive and credible manner.

10.1 LOCAL TASK FORCE

Form a local advisory group or task force comprised of key stakeholders, including police, community, civil rights, police unions or associations, professional researchers and/or academics.

In identifying "key" stakeholders, consider the following:

1) Racial and ethic make-up the City. The task force should reflect the diversity of the city.

2) Constituency of the interested group. Select groups that are established with large constituencies. This will prove necessary in order to receive input from the community, and to effectively market our program and successes.

3) The stakeholder's prior work in the community. The best predictor of future behavior is past performance. Data collection and analysis programs require extensive work and time commitment. You want to ensure that the representative groups have been successful in the past, and that they were willing to commit the time and resources necessary for the project.

4) Ability to be both fair and objective. Because of the nature of racial profiling, many people of predisposed and lack objectivity. For a project to be successful, representatives must have an open mind, and be accepting to learning new ideas and methods.

5) Provide training to the advisory group so they may obtain an expert level of knowledge and understanding of racial profiling, bias-based policing and the complexities of data collection and analysis. Do not assume task force members understand the issues.

6) Conduct regularly scheduled meetings with the task force, and establish a reasonable agenda for each meeting. Do not become overambitious and try to accomplish too much each meeting.

7) Involve and engage all members of the task force when facilitating the meetings through shared responsibilities and the delegation of meaningful work. Allow members to give presentations and lead discussions.

8) Determine the goal(s) and desired outcomes of data collection before designing the system. Engage the community in this process through marketing strategies, such as Town Hall meetings.

9) Utilize the task force to define racial profiling and bias-based policing and develop an agency policy that complies with applicable local ordinances, state law and CALEA standards.

10.2 DATA COLLECTION & ANALYSIS

1) Partner with a credible, reputable research partner to assist the group identify the locally relevant variables that may skew aggregate data and list all the relevant variables that are necessary in establishing benchmarks. This process must be completed prior to identifying what data should be collected.

2) The task force should work collaboratively with its research partner to identify baseline comparison data and establish benchmarks. This process cannot be relegated to the academic partner or conducted in isolation. Otherwise, the process will create suspicion and cause divisiveness within the group.

3) Identify what data should be collected. Professional research has been conducted in this area so it may not be necessary to reinvent the wheel. It is necessary to identify locally based variables as they can vary between agencies and jurisdictions.

4) Identify "best practices" in data collection, and develop collection methodologies that fit the organization, the community, and the budget.

5) Train officers and the community on racial profiling and bias-based policing; the new policy, the agencies data collection program – its purposes, value and expected outcome (not statistical), and their role in ensuring success.

6) Conduct survey to assess community and officer perceptions about racial profiling and police-community relations.

7) Conduct second survey after release of data-analysis report to assess what impact, if any, the program has had on police and community perceptions.

8) Conduct community forums, such as Town Hall meetings to promote program, announce results, and solicit input and feedback from community.

11 ACKNOWLEDGEMENTS

The Oakland Police Department is grateful to the United States Department of Justice, Community Oriented Policing Services (COPS) for the opportunity and support to participate in this worthwhile project. We especially thank our COPS coordinator, Ed Mixon, for providing continued guidance and assistance in our efforts.

Principal contributors to this report include

Capt. Ronald Davis, Oakland Police Department
Greg Ridgeway, RAND
Christopher Swartz-Edmisten, SCANTRON Corp.
Carolyn Marsh, Oakland Police Department
Ellen Purcell, Oakland Police Department
Jenny Wong, Oakland Police Department

12 REFERENCES

Alpert, G.P., M.R. Smith, and R.G. Dunham (2003). "Toward a better benchmark: Assessing the utility of not-at-fault traffic crash data in racial profiling research," *Confronting Racial Profiling in the 21st Century: Implications for Racial Justice*, Boston.

Attorney General of the State of New York (1999). *The New York City Police Department's 'Stop & Frisk' Practices: A Report to the People of the State of New York From The Office Of The Attorney General*, Civil Rights Bureau.

Fridell, L. (2004). *By the Numbers: A Guide for Analyzing Race Data From Vehicle Stops*. Police Executive Research Forum, Washington, D.C.

Friedman, J. H. (2001). "Greedy function approximation: A gradient boosting machine." *Annals of Statistics* 29(5), 1189-1232.

Grogger, J. and G. Ridgeway (2004). "Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness," Working paper.

Hosmer, D.W. and S. Lemeshow (2000). *Applied Logistic Regression*, Wiley, Boston.

Lamberth, J. (1994). *Revised Statistical Analysis of the Incidence of Police Stops and Arrests of Black Drivers/Travelers on the New Jersey Turnpike Between Exits or Interchanges 1 and 3 from the Years 1988 through 1991*. Lamberth Consulting.

Lamberth, J. (1996). *Expert Report for Wilkins v. Maryland State Police, Civil Action No. CCB-93-483*.

Lange, J.E., K.O. Blackman, and M.B. Johnson (2001). *Speed Violation Survey of the New Jersey Turnpike: Final Report*. Office of the Attorney General of New Jersey.

D.F. McCaffrey, G. Ridgeway, and A.R. Morral (2004, to appear). "Propensity score estimation with boosted regression for evaluating causal effects in observational studies," *Psychological Methods*.

Rickabaugh, C.A. (2003). *A Study to Analyze Traffic Stop Data in Santa Cruz County*. Lamberth Consulting.

Ridgeway, G. (2004). *GBM 1.1-1 package manual*. Available at the R-project web site, http://cran.r-project.org/doc/packages/gbm.pdf.

Rosenbaum, P. R. and D.B. Rubin (1983). "The central role of the propensity score in observational studies for causal effects," *Biometrika, 70*, 41–55.

Rosenbaum, P. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, *82*, 387–394.

US Naval Observatory, Astronomical Applications Department, available at http://aa.usno.navy.mil/.

Wooldridge, J. (2001). *Econometric analysis of cross section and panel data.* Cambridge: MIT Press.

Zingraff, M.T., W.R. Smith, and D. Tomaskovic-Devey (2000). "Driving While Black," *The Criminologist* 25:1-3.