Not so black and white: uncovering racial bias through systematically misreported trooper reports *

Elizabeth Luh†

July 29, 2022

Abstract

Biased highway troopers may intentionally misreport the race of the stopped motorists in order to evade detection. I develop a new model of traffic stops that highlights the incentive for biased troopers to misreport their failed minority searches as White. Applying my model to the universe of highway searches in Texas from 2010–2015, I find evidence of widespread bias that varies substantially across troopers. When misreporting became more difficult due to public scrutiny, biased troopers faced worse labor outcomes. This suggest an important role for increased accountability in data collection by law enforcement agents.

JEL Classification: J15, K42

Keywords: Racial Bias, Systemic Misreporting, Traffic Enforcement

*This research uses confidential information from the the Texas Department of Public Safety, graciously provided by the Stanford Open Policing Project. I pledge that the views expressed in this paper are my own and not of the State of Texas or of the Stanford Open Policing Project. I am grateful to Vikram Maheshri, Elaine Liu, Gergely Ujhelyi, Aimee Chin, and Chinhui Juhn for their continual support and invaluable guidance throughout this project. I am also grateful for helpful comments and suggestions from Nathan Canen, Steven Craig, Willa Friedman, Andrea Szabo, Fan Wang, Andy Zuppan, Germann Cubas, Bent Sorensen, Jennifer Doleac, Keith Finlay, Steve Mello, Mike Mueller-Smith, Emily Owens, and Brittany Street.

†CJARS, University of Michigan, 426 Thompson, Ann Arbor Michigan 48103. Email: eluh@umich.edu.

1 Introduction

In a 2017 survey, 27% of Latinos and 50% of Blacks felt personally discriminated against by police compared to only 10% of White respondents.[1] This perception is supported by a large and growing body of research identifying racial bias in nearly all aspects of the US justice system from airport screening (Persico and Todd, 2005), ticketing (Anbarci and Lee, 2014; Goncalves and Mello, 2021), police stops (Coviello and Persico, 2013), bail decisions (Arnold, Dobbie, and Yang, 2018), sentencing (Shayo and Zussman, 2011; Depew, Eren, and Mocan, 2017), parole (Anwar and Fang, 2015), use of force (Hoekstra and Sloan, 2022), prosecutorial decisions (Tuttle, 2021; Sloan, 2022), and capital punishment (Alesina and Ferrara, 2014). Recent events following the death of George Floyd at the hands of law enforcement have led to widespread calls for action making criminal justice reform a top priority for policymakers at all levels of government.

Despite this large academic literature on racial bias and discrimination in the criminal justice system, relatively little attention has been given to the response of law enforcement officers to heightened scrutiny. These behavioral responses may be important as law enforcement officers control how civilian interactions are recorded and face little oversight on the accuracy of their record. Indeed, if officers deliberately misreport their interactions with civilians in order to avoid appearing biased then such systematic measurement error could lead researchers to underestimate the extent of racial bias in the criminal justice system, and it could also hamper the efforts of those in charge of holding law enforcement accountable. Both of these adverse outcomes would be exacerbated if, as is plausible, the most biased officers have the most incentive to misreport.

In this paper, I develop a new model of racial bias in highway searches that builds on the seminal work of Knowles, Persico, and Todd (2001) and Anwar and Fang (2006). Motorists of different races differ in their *ex ante* likelihood of carrying contraband. Highway officers (henceforth "troopers") may be racially biased, and troopers have an incentive to misreport their motorist interactions to appear less biased. The model makes sharp predictions of the (mis)behavior of troopers: (1) only racially biased troopers will find misreporting to be profitable, (2) biased troopers must balance the risk of punishment

[1] See https://cdn1.sph.harvard.edu/wp-content/uploads/sites/94/2018/01/NPR-RWJF-HSPH-Discrimination-Final-Summary.pdf for details.

for bias with the risk of punishment for misreporting. By misreporting a portion of their failed minority searches as failed White searches, troopers can improve their reported minority search success rate and appear less biased.

I use my model to study the behavior of Texas Highway troopers during 2010-2015, a unique period in which highway troopers were free to record the race of motorists following their "best judgement" until local media brought scrutiny to this practice, which led to immediate changes in department policy (Collister, 2015b). Indeed, the recorded search rate for White motorists was significantly higher than the recorded search rate for non-White motorists during this time period. This fact was immediately reversed following the change in policy, as one would expect in the presence of racial bias.

The major empirical challenge in identifying biased behavior in policing is disentangling it from statistical discrimination because the underlying propensity for criminality may vary by race (Knowles et al., 2001; Anwar and Fang, 2006; Antonovics and Knight, 2009; Feigenberg and Miller, 2021). I circumvent this challenge by showing that a trooper's choices to misreport some of their searches can only be interpreted as evidence of racial bias – not statistical discrimination – under fairly weak assumptions. As a result, the task of identifying racial bias is reduced to searching for systematic misreporting by driver race. I show in my model that misreporting is only profitable when misreporting failed, minority searches as failed White searches.

I analyze a restricted data set of the universe of Texas highway searches from the Stanford Open Policing Project (SOPP) from 2010 to 2015, which I combine with trooper employment data from the Texas Department of Public Safety (DPS). My data set contains the recorded race, full name, and home address of every driver in every recorded highway search in Texas during the sample period in addition to annual employment records of every trooper employed by the DPS during a portion of the sample period.

One major challenge for my analysis is that intentional misreporting is unobservable. By leveraging the driver's full name and residence ZIP code, I estimate the true race of each driver, which allows me to compute misreporting rates at the trooper level. Of course, troopers may unintentionally misreport the race of a driver for reasons besides racial bias (ex. poor visibility or the driver's race may be visually ambiguous). But only biased troopers will consistently misreport their failed, minority searches. Thus, I use the difference in misreporting likelihood across search outcome (i.e. whether or not contraband was found) as my measure of misreporting. I show that on average, likely Hispanic motorists were 9% (relative to a mean of .257) more likely to be reported as White in failed searches than they were in successful searches. I interpret this is evidence that misreporting was used to hide racial bias against Hispanic motorists.

Because troopers stop hundreds of motorists over the course of a year, a major contribution of my analysis is that I am able to estimate racial bias at the trooper level. I find that racial bias is negatively correlated to the misreporting rates of troopers. I also find that troopers who are one standard deviation more biased against Hispanic motorists are 30% more likely to be biased against Black motorists.

Are biased troopers held accountable? I find that prior to 2015, labor market outcomes for troopers are uncorrelated to their level of bias. By exploiting the sudden and plausibly exogenous revelation of misreporting to the public in 2015, I find that one standard deviation in estimated Hispanic bias reduced monthly salary growth by 10% (relative to an average monthly salary of $8,000) and reduced the likelihood of increasing in rank by 80% (relative to a mean of 10 percentage points). Since these negative ramifications to biased behavior were only found after the rule change, I interpret this as suggestive evidence of (1) misreporting being an effective shield for biased troopers, and (2) public scrutiny being an antidote for this problem.

1.1 Literature Review

This paper contributes to the literature on detecting racial bias in the criminal justice system, specifically in contexts where law enforcement officers record the interaction. Many earlier contributions to the literature, notably in motorist stops and racial bias, examine the role of motorist race and trooper race in stop interactions, notably Knowles et al. (2001) and Anwar and Fang (2006) along with Antonovics and Knight (2009). These papers use Becker's (1957) outcome test, which identifies racial bias by comparing the success rates across different groups. These tests cannot measure the magnitude of bias on the individual, trooper level. Relative to this literature, my paper is able to measure bias on the individual level and also is the first to address the possibility of the data being purposefully misreported to hide bias.

This research contributes to past empirical research on cheating behavior since misreporting in the context of trooper reports is a form of cheating (Jacob and Levitt, 2003; Dee, Dobbie, Jacob, and Rockoff, 2019). The most prominent of these papers, Jacob and Levitt (2003) used patterns in students' test scores to uncover cheating. Similar to my paper, the true measure of cheating was unobserved, thus Jacob and Levitt (2003) had to test for cheaters by measuring cheating rates across different thresholds. An advantage of my method for uncovering misreporting is that the measure is within trooper and does not rely on the existence of a comparison group to measure cheating. This paper also is able to examine trooper consequences before and after the cheating was possible, which Jacob and Levitt (2003) cannot.

A recent paper by Goncalves and Mello (2021) accounts for this potential misreporting for Hispanic motorists and also identifies racial bias at the individual level using police officers' choice of leniency when giving speeding tickets. They find that officers are more likely to be lenient when ticketing speed violations with White drivers compared to non-White drivers, which they argue is proof of biased behavior. Using the individual officer's difference in leniency across motorist race, they identify discrimination by comparing these lenient officers to non-lenient officers. Compared to their paper, a strength of my approach is that my test does not rely on identifying a comparison group of unbiased officers to measure bias.

Given the severity of misreporting behavior in troopers observed in this paper, this motivates the need to carefully consider the accuracy of policing data when studying racial bias, especially when the agents themselves are responsible for recording the data. Knox, Lowe, and Mummolo (2020) show that even the choice of choosing to interact, thereby observing the interaction in the data, can itself be biased. New work by Campbell and Redpath (2022); Ba, Moreno-Medina, Ouss, and Bayer (2022) show that word choice by media and by police can also show racial slant.

In the presence of misreporting, even the most robust tests of racial bias will be under-detecting the existence of bias if they fail to account for this misreporting. The results also motivate other law enforcement departments to require their officers to ask for driver's race in all interactions to prevent misreporting and biased behavior.

The rest of the paper is organized as follows. In Section 2, I outline the background of my research. Section 3 outlines my theoretical model of racial bias. In Section 4, I explain my data construction. Section 5 shows my empirical results and other testable implications of my model. I finally conclude in Section 6.

2 Background

2.1 Misreporting and Highway Troopers in Texas

Texas Highway Patrol is a division of the Texas Department of Public Safety, which is responsible for enforcing state traffic laws and commercial vehicle regulation on highways of Texas. They currently employ over 2,800 troopers in Texas divided across 6 regions in Texas, with a separate region for their headquarters in Austin. The department is responsible for licensing of drivers, vehicle inspections, and handgun licensing.

To become a trooper, a person must complete recruit school or transfer from prior law enforcement service. New hires spend some at least one year as probationary troopers before receiving their permanent assignments. After the one year probationary period, troopers take their final exam and are promoted to trooper, conditional on passing.

With every four years, troopers can be promoted to different level of trooper classes and to different ranks, which include salary increases. Salary amounts are determined by years in the force and rank. Ranks or classes of troopers are similar to military ranks and go from trooper, corporal, sergeant, lieutenant, captain, and major. In general, only troopers in good standing (no sustained complaints, no disciplinary actions, no demotions) are promoted. Unlike other state police agencies, Texas legislature sets the salary of troopers, rather than the individual agencies. With each salary promotion, troopers can be moved to different stations across the state to fill availability. Troopers are allowed to have some say in the choice of where they are stationed after significant changes in DPS in 2012. Prior to 2012, station assignment was based on availability and need.

In a motorist stop, troopers are allowed to investigate the passenger and the driver. While drivers are not required to answer questions, they are required to provide their driver's license and if arrested, they must also provide their name, residence address, and date of birth. Law enforcement officers may ask for consent to search the vehicle or person, which the driver can grant or deny. "... however, if an officer has probable cause to believe that your vehicle contains evidence of crime, it can be searched without your consent (DPS, DPS)." To search a vehicle without the driver's consent, the trooper must either have: probable cause, arrested the driver *prior* to searching the vehicle, reasonably believes the motorist has weapons, or has a warrant.

Drivers can report troopers who can face repercussions if the claim is substantiated. Troopers badge numbers and names are normally provided during the stop and drivers can submit complaints to the department. The investigation can have one of four outcomes: unfounded, exonerated, not sustained, or sustained. A sustained complaint can result one or more of the following: formal written reprimand, disciplinary probation, time off without pay, reduction of salary rate, demotion, and or discharge.

Due to Senate Bill 1074 passed in 2001, Texas DPS is required to publish an annual traffic stop data report to provide "background pertaining to the issues of racial profiling (DPS, 2012)." This report breaks down search, stop, and citation statistics across race. Notably, the report also includes the number of criminal arrests resulting from a traffic search across race. Thus, troopers were aware of how their search and stop patterns may be used to determine racial bias, further motivating potential misreporting behavior.

2.2 2015 Misreporting Incident

On November 8th, 2015, KXAN published the results of their investigation of DPS, which found that troopers were " inaccurately recording the race of large numbers of minority drivers, mostly Hispanic, as White" (Collister, 2015b). For example, Figure 2 shows an actual ticket from a stop. The driver, with last name Mendez, is pulled over for speeding by Officer Salinas and is recorded as a White, male driver.

Texas troopers were already under scrutiny due to the death of Sandra Bland in jail after being pulled over for failing to signal a lane change (Sanchez, 2015). One week after the misreporting was uncovered, the House Committee on County Affairs held a hearing where DPS blamed the error on a computer glitch. As a result of the hearing, DPS changed its policies to require troopers to ask drivers to provide their race, rather than recording it based on the trooper's best judgment. This policy went in effect by November 23rd; as a result of the policy, the proportion of stopped motorists recorded as White fell from 18% to 4% by 2016 (Collister, 2015a).

An important result of the KXAN investigation was that misreporting was also found in other law enforcement departments in Texas, namely the Houston and Austin police departments. Thus, it is not out of the question to test for possible misreporting behavior in police or trooper forces in other state and law enforcement agencies. This raises the question if whether past reports and research of racial bias are possible under-measuring and under-detecting the existence of bias. Less than a month after the publication of the article, DPS changed its policies to require Texan troopers to now ask drivers for their race rather than using their own best judgment (Oyeniyi, 2015).

Misreporting is easy in motorist stops compared to other points of the criminal justice system. First, the trooper is not required to ask the driver for his or her race. Instead, the trooper is supposed to infer the race based on observable characteristics of the driver. Second, due to the high frequency of stops, stop reports of troopers or police officers who misreport are not checked for accuracy. Usually, only the driver focuses on the content of the ticket. Third, unless the trooper searches the driver and arrests the driver, it is unlikely another law enforcement officer (i.e judge or attorney) will look at the recorded race.

3 Model

Motorists of race m travel on highways; a fraction π^m of them are carrying contraband. Trooper t may stop motorists without observing their race. Conditional on stopping a motorist, a trooper receives a signal θ that contains all available information on whether the motorist is carrying contraband.[2] θ is collapsed to a single index $\theta \in (0, 1)$ and is drawn from distributions $f_g^m(.)$ if the driver does carry contraband and from $f_n^m(.)$ if the driver does not carry contraband. For ease of exposition, I assume that troopers and

[2]Some examples of these characteristics are age, height, address, gender, the interior of the vehicle, the smell of the driver, whether the driver is under the influence, whether the license plate is in-state, the time and place of the stop, whether the vehicle is rented, and the attitude of the driver.

motorists are either White (W) or minority (M) in this section. In my empirical analysis, I allow for motorists to be W or H (Hispanic).

Similar to past papers on racial bias (notably, Alesina and Ferrara (2014); Anwar and Fang (2006)), I make the following assumption:

Assumption 1. $f_n^m(.)$ *and* $f_g^m(.)$ *are continuous and satisfy the strict monotone likelihood ratio property (MLRP). Specifically,* $\frac{f_g^m}{f_n^m}$ *is strictly increasing in* θ

This implies the following properties of the distribution. First, a higher index of θ implies a higher probability of driver guilt. Second, the cumulative distribution, $F_g^m(.)$ stochastically dominates $F_n^m(.)$. In other words, motorists who carry contraband are more likely to appear more suspicious, or signal higher θ's. Lastly, $\frac{f_g^m}{f_n^m} \to +\infty$ as $\theta \to 1$.

3.1 Bias and Misreporting

Having observed (m, θ), a trooper decides whether to search the motorist in order to find contraband. Searching a driver incurs a cost of $c_{m,t} \in (0,1)$; troopers obtain a normalized benefit of 1 if drivers are guilty. The *ex ante* probability that a motorist is guilty is

$$\Pr\left(G = 1 | m, \theta\right) = \frac{\pi_m f_g^m\left(\theta\right)}{\pi_m f_g^m\left(\theta\right) + \left(1 - \pi_m\right) f_n^m\left(\theta\right)} \tag{1}$$

Trooper t will search a race-m motorist if and only if

$$\Pr\left(G = 1 | m, \theta\right) \geq c_{m,t} \tag{2}$$

This yields the search threshold, $\theta^*_{m,t}$.
Search thresholds that vary by m may reflect either statistical discrimination or bias on the part of troopers. A trooper may choose different thresholds purely because motorists θ's are drawn from different distributions or because π_m varies by race.

Definition 1. Trooper, with $c_{M,t} = c_{W,t}$, exhibits *statistical discrimination* against race M motorist if $\theta^*_{M,t} < \theta^*_{W,t}$.

Alternatively, a trooper may choose different thresholds because they incur different costs of failed searches. Following Knowles et al. (2001) and Anwar and Fang (2006), I define racial bias as

Definition 2. A trooper of race-t exhibits *racial bias* against motorist of race-M if $c_{M,t} < c_{W,t}$.

Given Definition 2, let $b = c_{W,t} - c_{M,t}$ be the magnitude of bias against race-M motorists for trooper-t. b is in terms of the trooper t's search cost across motorists' race and is unobservable. Thus, to compare levels of bias across troopers, I transform b into measurable units.

Definition 3. v is a measure of bias if $b > b' \iff v(b) > v(b')$

v is a monotonic transformation of b. Since $f^m_{g,n}$ and π_m are unobservable, proving that the measure of v is driven by b (racial bias) and not $\theta^*_{M,t} - \theta^*_{W,t}$ (statistical discrimination) is key to identifying v as a measure of b.
Troopers may face punishment for biased policing with probability P, which is monotonically increasing in $|b|$. In order to evade detection, a trooper may intentionally misreport the race of a motorist following a search, which will reduce the appearance of bias and thereby the likelihood of detection. But, troopers incur a cost of μ for misreporting, as it may open the door to greater punishment. I make the following assumptions on μ, the cost of misreporting:

Assumption 2. *$\mu(\theta, G) > 0$ is increasing in θ.*

As θ, increases, the cost of misreporting also rises. Therefore, motorists who appear less guilty are more likely to be misreported. One intuitive reason for this is that motorists with higher θ are in general more likely to be searched. Thus, by misreporting motorists who appear less guilty, the trooper is less likely to be caught misreporting.[3]

Since troopers misreport to reduce the appearance of bias and because of Assumption 2, troopers will misreport the race of a motorist if and only if

$$c_{M,t} + \mu_{M,t}(\theta, G) \leq c_{W,t} \tag{3}$$

[3] θ is likely positively correlated to other criminal behavior, further exposing the trooper to risk of punishment for misreporting.

Therefore, only troopers who are biased against race M motorists will misreport motorists of race M as W. If a trooper is unbiased, there exists no θ such that Equation (3) will hold.

Assumption 3. $0 < \mu(\theta, G = 0) < 1$, $\mu(\theta, G = 1) > 1$ *for all* $\theta \in (0, 1)$.

Guilty searches are more likely to end up in court where another person (i.e. a judge) will view the search report with the incorrect driver's race exposing the trooper to risk of punishment for misreporting. Thus, misreporting searches is only profitable when the search ends in failure.

Assumption 2 and 3 implies that troopers will misreport the race of a motorist if and only if

$$c_{M,t} + \mu_{M,t}(\theta, G = 0) \leq c_{W,t} \tag{4}$$

This yields the misreporting ceiling, $\theta^{\mu}_{M,t}$.

Given this set up, I obtain the following result:

Proposition 1. *Under Assumption 1, 2, and 3, troopers will misreport motorists with characteristics* (M, θ) *if and only if* $\theta \in (\theta^{*}_{M,t}, \theta^{\mu}_{M,t})$ *and the search ends in failure.*[4]

Troopers will only misreport their failed searches. Because the misreporting decision is conditional on search, any misreported motorists must have $\theta > \theta^{*}$. Troopers also will not misreport motorists over a certain threshold, specifically $\theta > \theta^{\mu}$. That is, motorists who appear more guilty than the search threshold will not be misreported.[5]

The fact that only biased troopers will misreport their searches provides an attractive criterion to identify bias. In particular, biased troopers will only misreport their unsuccessful searches and correctly report the motorists' race in successful searches, creating an observable difference in search behavior across motorists race between biased troopers and unbiased troopers:

Proposition 2. *Under Assumption 1,2, and 3, the difference in the average misreporting rate of race* M *motorists for trooper* t *across search outcome* G*,*

$$v_{M,t} = (1 - \pi_M)[F^M_n(\theta^{\mu}_{M,t}) - F^M_n(\theta^{*}_{M,t})] \tag{5}$$

[4] The proof of Proposition 1 is in the appendix.

[5] One intuitive reason for this is that the searching motorists who appear more guilty (have higher θ) are more justifiable if the trooper is accused of discrimination.

is a measure of bias against race M motorists for trooper t.

For unbiased troopers, $v = 0$. For biased troopers, $v > 0$.[6] The magnitude of $v_{M,t}$ itself will also be trooper t's measure of bias against race M motorists. This forms the basis of my measure of racial bias for trooper t against race M motorist that I use throughout the rest of the paper.

4 Data

4.1 Stop Data

The Stanford Open Policing Project (SOPP) has a restricted version of highways stops conducted from 2005 to 2015 from the Texas Department of Public Safety. The restricted version contains personally identifiable information of the driver such as full name, home address, owner's full name, and license plate of the stopped vehicle.Pierson, Simoiu, Overgoor, Corbett-Davies, Ramachandran, Phillips, and Goel (2020) courteously provided the raw version of the data.[7] As DPS did not record the driver's last name prior to 2010, only stops from 2010 onward are included in the study.

The data also has rich stop information such as the latitude, longitude of the stop, the badge number of the officer who recorded the stop, the race of the driver, the state in which the driver's license was issued, and the make and model of the vehicle. The data also has information on the violation such as reason for the stop, the outcome of the stop (citation, warning), whether a search was conducted, the search reason, and the outcome of the search. The highway stop data is publicly available on the TX DPS website from 2013 - 2019.[8]

In Texas, troopers can legally search a vehicle for many reasons aside from probable cause or driver consent. Some of these situations, such as search incident to arrest, after the car is impounded, or with a warrant, do not fit the framework of the model. Because

[6]The proof of this and Proposition 2 is in the appendix.

[7]SOPP collected over 130 million records from 31 state police agencies (Pierson et al., 2020). The goal of the project is to analyze detailing interactions between police and the public. This data is freely available on the website.

[8]The SOPP data is originally from the TX DPS. I have verified that the data is the same for overlapping years.

of this, I restrict my definition of search success to only include searches due to probable cause or driver consent. For my main analysis and estimates of trooper bias, I also drop any stops that occur after the publication of the article, November 15, 2015.

4.2 Trooper Employment Data

The employment data is from the Texas Department of Public Safety, which I obtained using a Freedom of Information Act (FOIA). Unfortunately, DPS only has this information for employees after 2013. If a trooper left DPS prior to 2013, I do not have his or her employment information. For troopers in the data, I have the year the trooper was hired, if he or she left the position and why, the salary for each year, which work city he or she was stationed at, the work position for each year, race/ethnicity of the trooper, the full name of the trooper, and the badge number. I have approximately 2,789 unique troopers of which I can match 2,466 to the stop data.

I merge the stop data to the trooper data together using the badge number of the trooper. I can match all but 10% of the stop data to the trooper ending with approximately 12 million total stops and nearly 220,000 total searches.

I further the time period of my trooper employment data by adding 2019 trooper employment data, which is publicly available on the Texas Tribune Salary website. I link both of the Tribune's employment data to my trooper data using the full name of the trooper. I include this data as a measure of a trooper's long-term employment outcomes.

I also include trooper complaint data from 2010 to 2015, which I obtained using a FOIA, as a secondary measure of trooper work behavior. The complaint data contains information on the date the incident occurred, the date the complaint was received, the allegation of the complaint, the trooper's badge number (if applicable), and the investigator of the complaint. The badge number is not always included due to Texas' privacy laws.[9] Out of the original 1,873 complaints, only 334 had the trooper's badge number in the complaint.

[9] Specifically, "Employee names and ID numbers are not releasable unless the complaint resulted in disciplinary action such as discharge, suspension, or demotion (Government Code 411.00755)."

4.3 Race Estimation

I use two main methods supported by past literature on using observable characteristics to determine race. These methods are predominantly used in social science and health research to infer patient race (Fiscella and Fremont, 2006; Freedman, Owens, and Bohn, 2018). To estimate Black drivers, I use the concentration of Black residents in the driver's home ZIP code to predict race. Specifically, I match the ZIP code of the driver's home address to a ZIP code tabulation area (ZCTA5) using the 2011–2015 5-Year American Community Survey. If the proportion of Black residents within that ZCTRA5 is greater than 90%, I impute the driver's estimated race as Black.

For Hispanic motorists, I use surname analysis combined with the same home address analysis for predicting Black drivers. I match each driver's last name to a surname using the 2000 Census Surnames data set. If the probability of the last name is Hispanic is greater than a certain threshold (90%) and the proportion of Hispanic residents within the ZCTA5 area is greater than 75%, I impute the 'estimated' race as Hispanic. Assuming this driver resides in a ZCTA5 area with proportion Hispanic greater than 75%, given the probability this driver is Hispanic, conditional on his last name, Mendez, is 92%, I estimate his actual race to be Hispanic. I follow the same thresholds for imputing Asian motorists.

In order to simplify the later analysis, I combine all other race groups into the other category. This is a small of observations, making up only 7% of all stops with the largest category being race unknown.

4.4 Descriptive Statistics

I present summary statistics of motorist characteristics in Table 1 using the recorded races. On average, I find that White motorists are over represented in both searches and all stops. From the 2010 Decennial Census, only 45% of Texas residents were non-Hispanic White, but make up nearly 70% of the stops and 60% of searches. Black and Hispanic motorists are searched at nearly equal rates of 10% and 13% respectively and are under-represented given the 2010 Decennial Census which reports 11.9% and 40% respectively. I also find that certain stop characteristics, such as stops occurring from 8

PM - 5 AM and whether the car is older than 5 years are more likely to occur in searches compared to stops.

Table 2 shows summary statistics of troopers. Of the 2,466 troopers I was able to match to the stop data, approximately 63% are White, 26% are Hispanic, and almost 8% are Black. Native American and Asian troopers along with other race troopers make up the remaining force. The force is predominantly male at 96%.

When compared to searches, I find that White troopers make up most of the searches at 70%, followed by Hispanic troopers at 23%. I find that only White troopers search at a higher rate compared to the stop rate while Black and Hispanic troopers search at a lower rate. I also find that troopers with less experience search at higher rates with the average hire year for searches being greater than the average hire year for stops.

In the bottom part of the table, I break down the stop and search statistics by trooper position. Troopers with rank of lieutenant or greater are aggregated to the same rank as ranked officers make up only 20% of the highway patrol. I find as rank increases, troopers are less likely to search. Using the rank of Lieutenant+ as an example, the interpretation of the probabilities is "troopers of lieutenant rank or higher conduct 5.5% of total searches." I find that troopers make up approximately 80% of searches and stops. Probationary troopers make a small portion of searches and stops at only 1%. But, since most probationary troopers in the employment data do not have badge numbers and therefore can't be linked to the stop data, this may reflect poor data linkages within that rank, rather than overall probationary trooper behavior.

For a small percentage of troopers, the employment data is missing employment or demographic information. This seems to occur at higher rates in stops compared to searches, but on average approximately 1% stops and searches are conducted by troopers with no employment information. For a higher percentage, approximately 4%, the stop or search was conducted by a trooper without race or sex information. For this subset of troopers, I imputed race and sex using the full name of the trooper merged with the 2000 Census Surnames data set and the 2000 Census Names by Sex data set. If the likelihood of that name being associated with a certain race or sex is greater than 75%, I impute the race or sex with that sex or race. I keep all troopers, even if the trooper is missing employment information, thus some of the results on trooper outcomes may vary in sample size.

5 Empirical Results

5.1 Test for racial prejudice

From Eq. (5), troopers' decision to misreport the motorists' race as White will vary by search outcome. Furthermore, race misreporting is only profitable for biased troopers' when the search fails to find any contraband. Thus, identifying biased troopers and measuring their misreporting rate is of policy interest because it provides an intuitive measure of racial bias on the individual level.

Figure 1 shows the raw time trend of the search rates by recorded driver's race from 2010–2017. Prior to the rule change, marked with the red-dashed line in the figure, White motorists were the most likely to be searched with a quarterly average search success rate ranging from 60% to just over 40%. The search rates for motorists recorded as Hispanic ranges from 10% to 30%. After the rule change, the Hispanic search rate surpasses the White search rate with the White search rate decreasing simultaneously. Given the changes in the average search rates, the figure shows that most of the misreporting occurred between Hispanic and White motorists.

Measuring purposeful misreporting at the stop level is impossible with the given data for a few reasons. The first is that driver's true race is unobservable, even with the best race estimation techniques. The second is that troopers can unintentionally misreport driver's race for reasons aside from bias, such as poor visibility conditions or poor race identification ability.

The model again proves useful for helping separate the previous examples from the misreporting behavior of interest. If an officer's misreporting is intentional and linked to racial bias, then misreporting is only beneficial if unsuccessful minority searches are misrecorded as White unsuccessful searches. In the examples of misreporting that are unrelated to bias, those examples should be equally likely to occur regardless of the search outcome. Thus, testing the likelihood of race mismatch across search outcome provides an intuitive test for whether the mismatch is driven by racial bias or not.

To formally test whether the misreporting was intentional prior to the rule change, I use a linear probability model where the outcome is whether the recorded race does not match the estimated race, or mismatch regressed on whether the search ended in failure by each estimated race group:

$$I(Mismatch_{i,c,t}) = \beta_0 + \beta_1 I(Failure_{i,c,t}) + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,t} \tag{6}$$

The coefficient of interest is β_1, which indicates the increased likelihood of race mismatch for failed searches. In other words, how much more likely does mismatch occur when searches end in failure compared to success? $X_{i,c}$ is a vector of controls for the stop, including hour of the stop, month of the stop, year of the stop, county fixed effects, and vehicle type and vehicle age. I also include the full interaction for hour of the stop with month of the stop and year of the stop to control for seasonal and darkness variation that may impact driver's race visibility. I also include the full interaction of vehicle characteristics since these may be inputs in an officer's search decision.

Table 3 shows the estimates of β_1 for each Hispanic, Black, and Asian motorists in Columns (1), (2), and (3) respectively. The estimated coefficients show that most of the mismatch between recorded and estimated race is concentrated on Hispanic motorists with 27% of estimated Hispanic searches resulting in mismatch. Furthermore, for estimated Hispanic motorists, this mismatch appears to be linked to misreporting with estimated Hispanic motorists 2.3 percentage points significantly more likely to be misrecorded as White when searches end in failure compared to success.

In contrast, only 12% and 2% of estimated Asian and Black searches are mismatched with the recorded race. This could also be due to the high threshold for Black race estimation combined with very few ZIP codes and few last names having greater than 90% proportion of Black residents or surnames identifying as Black. I find that estimated Asian motorists are also more likely to be misreported during failed searches, but the estimate is not significant.

There are a two main reasons why misreporting was concentrated on Hispanic motorists rather than other race groups. The first, and perhaps most importantly, Hispanic is technically an ethnicity and not a race. Thus, Hispanic drivers could technically be recorded as White, despite Hispanic being the more correct race code. Another reason is that Texas' proximity to the border and the contentious immigration flows may lead to greater animus towards Hispanic motorists compared to other non-White groups.

5.2 Estimating Officer-level Hispanic Bias

To estimate trooper level bias, I focus the rest of the analysis on Hispanic motorists as there is not a sufficient number of searches for Asian or Black drivers to identify misreporting at the individual trooper level. Specifically, the data has nearly 49,000 searches with Hispanic motorists and only 38,000 and 1,500 searches with Black or Asian motorists respectively.

To measure the magnitude of Hispanic bias for each officer, I allow for each trooper to have his own misreporting rate depending on the search outcome. For every estimated Hispanic driver stop i by trooper$j$ at time t:

$$I(Mismatch_{i,c,j,t}) = \beta_0 + \beta_1^j I(Failure_{i,c,j,t}) + \delta_j + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,j,t} \tag{7}$$

β_1^j measures officer j's differential misreporting behavior based on search outcome. A positive estimate indicates that trooper j is more likely to have mismatch between the observed and estimated race when the search ends in failure, which implies bias against Hispanics. δ_j is the officer fixed effect, which can also be interpreted as the average rate of mismatch for each trooper. $X_{i,c}$ is the same vector of controls included in Equation 6.

From prior work using these value-added models (see Aaronson, Barrow, and Sanders (2007); Goncalves and Mello (2021); Koedel, Mihaly, and Rockoff (2015); Weisburst (2022), the distribution of $\hat{\beta}_1^j$ will have a higher variance relative to the true distribution due to estimation error. Compounding on this, the few number of searches the trooper-level estimate of bias introduces potential measurement error, further attenuating the estimates. To correct for this, I follow the Bayes shrinkage procedure from Weisburst (2022) to estimate the distribution of bias accounting for the estimation error in each $\hat{\beta}_1^j$.

Figure 3 shows the raw bias estimates (solid line, black) plotted with the shrunken estimates of bias from Weisburst (2022) (dashed line, blue). The further right the trooper is in the distribution of bias, the higher his level of bias. The measurement of bias is the difference in likelihood of misreporting between his failed searches and his successful searches. For example, a trooper with estimated bias of 0.5 is 50 percentage points more likely to misreport his failed searches of estimated Hispanic motorists compared to his successful searches. The average Hispanic bias using the original estimates is 0.042 and shrinks to 0.038 after applying the Bayes shrinkage procedure. Thus, the average officer is approximately 4 percentage points more likely to misreport his failed Hispanic searches as White compared to his successful Hispanic searches. The average standard error for $\hat{\beta}_i^j$ is 0.498. This estimate of average Hispanic bias is higher than the average Hispanic bias estimated in Table 3. For my main results in Section 5.3, I use the unshrunk estimates, but results on trooper's labor outcomes are robust to using the shrunk estimates.[10]

Given the size of the average standard error of officer's $\hat{\beta}_1^j$ in relation to the standard deviation of $\hat{\beta}_1^j$, some of the variation is due to estimation error. This could be driven by a few possible reasons. First, officer's do not often conduct searches in their highway stops. From Table 1, troopers search only 2% of stops from 2010 to 2015. The sample to estimate bias of each trooper is further restricted to motorists of estimated Hispanic ethnicity. Thus, mechanically, the standard errors for each officer's estimated bias will be large.

Officer's may also have reasons to misreport for reasons aside from racial bias, which may also explain the the sizeable portion of troopers with $\hat{\beta}_1^j < 0$. Here, troopers are more likely to misreport their successful minority searches as White, thus making themselves appear more biased against Hispanic motorists in the recorded data. One possible explanation for this behavior is that this is randomly generated by the race estimation process. The second is that some troopers may not want to appear better at searching White motorists compared to Hispanic motorists for reasons aside from bias.

To explore this second possibility, Table 4 shows the correlation of trooper characteristics and stop behavior with the likelihood of being estimated with negative bias. Column (1) shows that Hispanic troopers and those of corporal rank are significantly more likely to have estimated negative bias. Column (2) is restricted to just Hispanic troopers; here, corporal troopers are still significantly more likely to misreport their successful Hispanic

[10]See Appendix (A.3 for results using the shrunk estimates.

searches as White. When examining the correlation for negative bias across trooper rank for non-Hispanic troopers in Column (3), corporals do not behave significantly differently compared to their peers. Thus, it seems that higher ranked Hispanic troopers use misreporting to appear worse at searching Hispanics. Determining the motives for why these troopers choose to misreport in this fashion is beyond the scope of this paper.

Regressing $\hat{\delta}_j$ with $\hat{\beta}_1^j$ yields the correlation between trooper's average mismatch rate and the trooper's Hispanic bias. Figure 4 shows the scatter plot of $\hat{\beta}_1^j$ against $\hat{\delta}_j$. The trooper's average mismatch rate can be interpreted as a measure of the trooper's ability to identify driver's race accurately. Higher rates of mismatch indicate that the trooper's own imputation of race does not match the estimated race imputation, regardless of search outcome. Thus, a positive correlation would indicate that trooper's with lower ability to accurately identify race are more likely to have higher rates of bias, undermining the estimates of bias. I find the opposite relationship; biased troopers have lower rates of mismatch overall, indicating higher ability to accurately measure driver's race.

As a robustness check, I estimate officer-level Black bias using the same specification in Eq. (7), but restricted to estimated Black motorists. If a trooper is using misreporting to hide their Hispanic bias, this trooper may also be using the same misreporting to hide bias for other minority motorists. Using the main thresholds in the prior analysis, I only estimate approximately 80 drivers as Black who were originally recorded with another race. In order to increase the variation needed to estimate officer-level bias, I lower the thresholds for Black race estimation to 75%.

Table 5 shows the correlations between Hispanic bias and Black bias along with the average Black mismatch rate of each trooper. I find a significant, positive correlation between Hispanic bias and Black bias. Specifically, an increase in Hispanic bias is associated with a 30% increase in the likelihood of having positive Black bias (relative to a mean of 51%), shown in Column 1. Thus, officers with higher levels of Hispanic bias are likely to be biased against other minorities. Similarly to Figure 4, Column 2 shows a strong, negative correlation between the officer's average Black motorist mismatch rate and the estimated levels of Black bias. Thus, the estimates of Black bias are less likely to be driven by troopers who generally are poorly skilled at Black race identification.

Lastly, I ensure that the results are not driven by the race estimation thresholds.

Figure A.5 in the Appendix shows the distribution of officer level Hispanic bias across various thresholds. The distribution does not change significantly across the different thresholds and 90% has the highest concentration at 0 bias. Although the distribution widens slightly as the threshold lowers, the increase is minimal and does not affect the overall distribution.

Since the policy question of interest is the employment characteristics of biased troopers, I restrict most of the rest of my analysis to troopers with 0 or greater estimates of Hispanic bias. Furthermore, troopers with estimated negative bias may have different motives for misreporting motorist race, which may attenuate any estimates on the relationship between bias and labor outcomes. Results using all troopers are included in the appendix.

5.3 Bias and Trooper Characteristics

One contribution of this paper is to be able to generate trooper level estimates of discrimination and to identify effects of bias on labor outcomes. In this section, I will address how discrimination varies with other employment characteristics such as promotions, salary, and officer transfers. I will also test how troopers' employment outcomes were affected by the change in driver race identification method in 2015.

First, I test if employment outcomes, such as salary and experience, and trooper demographic characteristics, such as race and sex, are correlated to bias where experience is measured using the number of years employed by 2015. Table 6 shows the correlation between officer-level estimates of bias, normalized, with employment characteristics using employment information from 2010–2015. Column (1) shows the correlation for all troopers and Column (2) is restricted to troopers with 0 or greater level estimates of Hispanic bias. I find that both experience and salary are not correlated to Hispanic bias with near zero and insignificant estimates. Hispanic troopers are also significantly less biased compared to their White peers, but the estimate attenuates and is insignificant when restricted to only positive troopers. Black troopers' estimates of bias, when restricted to positive bias, is 0.24 standard deviations greater than White troopers.

For trooper rank, increasing in rank has no significant difference in bias compared to trooper rank except for probationary troopers. I find that probationary troopers' bias is 0.3 to 0.5 standard deviations greater than troopers' bias. This result is robust across both samples. One possible reason is that probationary troopers are inexperienced and may use misreporting to cover poor search decisions.

To examine the effect of trooper bias on the trooper's career across time, I divide the trooper's career into two sections: pre-2013, and 2014–2015. This has a few advantages; first, DPS does not have trooper employment data available prior to 2013 so 2013 is the earliest possible year. Second, combining the years increases the number of searches used to measure bias, which increases the precision of the estimates of bias. Third, with the panel-like structure, I can test if changes in employment outcomes are related to bias, specifically outcomes such as increasing in rank, moving cities, and leaving the force. If bias in 2013 has no effect on employment outcomes from 2014 to 2015, this could imply that misreporting is effective in making biased troopers appear unbiased thereby avoiding punishment of bias.

This is indeed what I find, as shown in Table 7. Specifically, I find that bias measured using stops from 2010–2013 has no significant effect on employment outcomes 2014–2015. Not only are the point estimates insignificant with large standard errors, but the estimates are also close to zero indicating a precise, null effect. I observe no difference in the likelihood of leaving the force, increasing in rank, or changes in salary with respect to the trooper's level of bias. Thus, in the presence of misreporting, biased and unbiased troopers have similar labor outcomes.

I next test if biased troopers also perform worst in other aspects of their job by using complaint data obtained from DPS. While misreporting may help troopers evade negative employment outcomes, drivers may find cause to report the trooper. The results in Table 8 show a positive relationship between trooper level bias and the probability of receiving a complaint. One standard deviation of bias is associated with a 50% higher likelihood in having a complaint filed against the trooper (relative to mean of 6.1%). This estimate is likely an underestimate of the actual association of bias and complaints since not sustained or unfounded complaints repressed the trooper's badge number. From the 1,873 complaints, only 334 included the trooper's badge number.

Lastly, I test to see how the employment outcomes of troopers were affected by the publication of the article relative to their level of bias. I use publicly available 2019 salary data published by the Texas Tribune. My results in Table 9 show that troopers with one standard deviation of positive Hispanic bias are 3.7 percentage points more likely to leave the force, but this estimate is not significant. For those who remain in the force, I find negative and significant estimates of the correlation between Hispanic bias on salary growth and likelihood of ranking up. Specifically, one standard deviation is associated with an 8.9 pp decrease in likelihood of ranking up, which is equivalent to a 90% decrease. Given the significantly reduced likelihood of ranking up, I also find that one standard deviation decreases monthly salary growth by $80, or 10% relative to an average growth of $800.

Overall, I interpret this as evidence that misreporting reduces agency's ability to identify biased troopers. Once misreporting becomes significantly harder, biased troopers are less likely to be promoted, which reduces their salary growth relative to their unbiased peers. I do find minimal, but insignificant, evidence that biased troopers are also less likely to remain in the force.

5.4 Relationship to past tests for racial bias in highway stops

Misreporting can distort the results when using past statistical tests of bias, notably, Becker's outcome test which is used by Knowles et al. (2001); Anwar and Fang (2006); Antonovics and Knight (2009). Under this test, a lower search success rate for Hispanic motorists would imply bias. In Table 10, Column (1) shows the search success rate by motorist race using the recorded races. From Column (1), I find Hispanic motorists have the lowest success rate of 27.7% compared to the White (38.1%), Asian (37.2%), and Black search success rate (41.1%). Under Knowles et al. (2001), the test would detect bias. When restricted to troopers with estimates of positive bias only, the differences between the recorded race and estimated race become even larger as shown in Column (3) of Table 11. Furthermore, differences across race also become even more pronounced.

How does the hit rate test change in the presence of misreporting? Table 12 shows the likelihood of search success across driver's race fully interacted with each trooper's estimated levels of Hispanic bias. Similar to Table 10, I find that motorists recorded as

Hispanic or Asian or Other have significantly lower search success rates when compared to motorists recorded as White. When restricted to just troopers with estimated bias greater than or equal to 0, I find that only motorists recorded as Other or Asian have a significantly different search success rate compared to motorists recorded as White.

The interaction of Hispanic bias with motorist race shows how the use of misreporting changed the appearance of racial bias tested by the hit rate test. Here, I find that one standard deviation of bias increased the Hispanic search success rate by 5 percentage points, making them appear unbiased when compared to White search success rate. Thus, this misreporting trooper would effectively appear unbiased against Hispanic motorists.

When restricted to troopers with 0 or greater levels of estimated Hispanic bias, one standard deviation of bias increased the Hispanic search success rate by 3 percentage points. Thus, compared to their unbiased peers, these troopers would actually appear to have a higher Hispanic search success rate. Overall, it appears that biased troopers did use misreporting to evade detection of racial bias using the hit rate test.

6 Conclusion

Recent events have highlighted disparate treatment by race in the criminal justice system by law enforcement officers. In this paper, I show how racially biased officers take systematic measures in order to appear less biased. Crucially, the findings of the paper bring into question outcome based tests, notably by Knowles et al. (2001) and Anwar and Fang (2006) that are at risk of manipulation by law enforcement officers. My statistical model of highway searches that explicitly allows for misreporting reveals that because biased troopers have an incentive to misreport their searches, evidence of misreporting can be interpreted as evidence of racial bias.

One positive outcome of the misreporting is the public response and ability to change DPS' policies. Notably, from the time of the misreporting publication to the race recording rule change was only 15 days. Furthermore, once misreporting became significantly harder, DPS also positively responded to racial bias by promoting unbiased troopers. Overall, the public played a significant role in alleviating the racial bias present in Texas' DPS troopers.

While my paper is the first to find a relationship between race misreporting and racial bias, the geographic scope of this paper is limited and further study of misreporting in other levels of policing and varying geographic contexts will require further study. Inputs such as trooper peers and supervisors, can explain the distribution of trooper behavior and raise important policy implications, which are beyond the scope of this paper. Lastly, evidence for what other factors, aside from punishment, may induce misreporting are important for future policies and research.

References

Aaronson, D., L. Barrow, and W. Sanders (2007). Teachers and student achievement in the chicago public high schools. *Journal of Labor Economics 25*, 95–135.

Alesina, A. and E. L. Ferrara (2014). A test of racial bias in capital sentencing. *American Economic Review 104*(11), 3397–3433.

Anbarci, N. and J. Lee (2014). Detecting racial bias in speed discounting: evidence from speeding tickets in Boston. *International Review of Law and Economics 38*, 11–24.

Antonovics, K. and B. G. Knight (2009). A new look at racial profiling: evidence from the Boston police department. *The Review of Economics and Statistics 91*(1), 163–177.

Anwar, S. and H. Fang (2006). An alternative test of racial prejudice in motor vehicle searches: theory and evidence. *American Economic Review 96*(1), 127–151.

Anwar, S. and H. Fang (2015). Testing for racial prejudice in the parole board release process: theory and evidence. *Journal of Legal Studies 44*(1).

Arnold, D., W. Dobbie, and C. S. Yang (2018). Racial bias in bail decisions. *Quarterly Journal of Economics 133*(4), 1017–1055.

Ba, B., J. Moreno-Medina, A. Ouss, and P. Bayer (2022, July). Officer-involved: The media language of police killings.

Campbell, R. A. and C. Redpath (2022, June). Officer language and subject race: A text analysis of police reports.

Collister, B. (2015a). DPS troopers getting race right after KXAN investigation. *KXAN*. March 1. https://www.kxan.com/news/dps-troopers-getting-race-right-after-kxan-investigation/.

Collister, B. (2015b). Texas troopers ticketing Hispanic drivers as white. *KXAN*. July 21. https://www.kxan.com/investigations/texas-troopers-ticketing-hispanic-drivers-as-white/.

Coviello, D. and N. Persico (2013). An economic analysis of black-white disparities in NYPD's stop and frisk program. Working Paper 18803, NBER, https://www.nber.org/papers/w18803.

Dee, T. S., W. Dobbie, B. A. Jacob, and J. Rockoff (2019). The causes and consequences of test score manipulation: evidence from New York Regents Examination. *American Economic Journal: Applied Economics 11*(3), 382–423.

Depew, B., O. Eren, and N. Mocan (2017). Judges, juveniles, and in-group bias. *Journal of Law and Economics 60*(2).

DPS. *When stopped by law enforcement.* Texas DPS.

DPS, T. (2012). 2012 traffic stop data report. *Texas Department of Public Safety*.

Feigenberg, B. and C. Miller (2021, 05). Would Eliminating Racial Disparities in Motor Vehicle Searches have Efficiency Costs?*. *The Quarterly Journal of Economics 137*(1), 49–113.

Fiscella, K. and A. M. Fremont (2006). Use of geocoding and surname analysis to estimate race and ethnicity. *HSR: Health Services Research 41*(4), 1482–1500.

Freedman, M., E. Owens, and S. Bohn (2018). Immigration, employment opportunities, and criminal behavior. *American Economic Journal: Economic Policy 10*(2), 117–151.

Goncalves, F. and S. Mello (2021). A few bad apples? racial bias in policing. *American Economic Review 111*(5), 1406–41.

Hoekstra, M. and C. Sloan (2022, March). Does race matter for police use of force? evidence from 911 calls. *American Economic Review 112*(3), 827–60.

Jacob, B. A. and S. D. Levitt (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics 118*(3), 843–877.

Knowles, J., N. Persico, and P. Todd (2001). Racial bias in motor vehicle searches: theory and evidence. *Journal of Political Economy 109*(1), 203–229.

Knox, D., W. Lowe, and J. Mummolo (2020). Administrative records mask racially biased policing. *American Political Science Review*.

Koedel, C., K. Mihaly, and J. E. Rockoff (2015). Value-added modeling: A review. *Economics of Education Review 47*, 180–195.

Oyeniyi, D. (2015). State troopers will now just ask drivers their race. *Texas Monthly*. November 23. https://www.texasmonthly.com/the-daily-post/state-troopers-will-now-just-ask-drivers-their-race/.

Persico, N. and P. E. Todd (2005). Passenger profiling, imperfect screening, and airport security. *American Economic Association Papers and Proceedings 95*(2), 127–131.

Pierson, E., C. Simoiu, J. Overgoor, S. Corbett-Davies, V. Ramachandran, C. Phillips, and S. Goel (2020). A large scale analysis of racial disparities in police stops across the United States. *Nature Human Behavior 4*.

Sanchez, R. (2015). Who was Sandra Bland? *CNN*. July 23. https://www.cnn.com/2015/07/22/us/sandra-bland/index.html.

Shayo, M. and A. Zussman (2011). Judicial ingroup bias in the shadow of terrorism. *Quarterly Journal of Economics 126*(3), 1447–1484.

Sloan, C. (2022, January). Racial bias by prosecutors: evidence from random assignment.

Tuttle, C. (2021, August). Racial disparities in federal sentencing: Evidence from drug mandatory minimums.

Weisburst, E. (2022). Whose help is on the way? The importance of individual police officers in law enforcement outcomes. *Journal of Human Resources*.

Tables

TABLE 1. Driver Summary Statistics

Driver Characteristics	All Stops	Searches	Δ
Recorded Asian	.017	.008	.009
	(.129)	(.09)	(0)
Recorded Black	.101	.174	-.073
	(.302)	(.379)	(.001)
Recorded Hispanic	.132	.163	-.031
	(.338)	(.369)	(.001)
Recorded Other Race	.064	.073	-.009
	(.245)	(.26)	(.001)
Recorded White	.686	.582	.104
	(.464)	(.493)	(.001)
Male	.677	.812	-.135
	(.468)	(.39)	(.001)
Luxury Car	.075	.071	.004
	(.263)	(.256)	(.001)
Vehicle Age > 5 years	.593	.748	-.155
	(.491)	(.434)	(.001)
Stop between 8 PM – 5 AM	.277	.399	-.122
	(.448)	(.49)	(.001)
N	11,897,213	218,813	

Notes: Unweighted means are shown. Standard deviations are in the parentheses for columns (1)–(3). Sample is restricted to stops and searches by troopers with employment information (90% of the sample). All statistics are generated using information reported on the stop by the trooper. Estimates are generated using stops from 2010 to November, 2015.

TABLE 2. Trooper Summary Statistics

Trooper Characteristics	All Stops	Searches	Δ
Total Stops	7242.643	7082.567	160.076
	(3453.623)	(3174.635)	(6.86)
Total Searches	130.036	328.904	-198.868
	(181.628)	(325.032)	(.697)
Year Hired	2005.3	2005.818	-.518
	(5.726)	(4.798)	(.01)
Monthly Salary	5455.014	5413.763	41.251
	(642.409)	(611.153)	(1.327)
Native American	.01	.007	.003
	(.097)	(.084)	(0)
Asian	.011	.012	-.001
	(.104)	(.109)	(0)
Black	.081	.05	.031
	(.273)	(.218)	(0)
Hispanic	.26	.226	.034
	(.439)	(.418)	(.001)
White Trooper	.634	.703	-.069
	(.482)	(.457)	(.001)
Trooper	.801	.795	.006
	(.4)	(.404)	(.001)
Probationary Trooper	.011	.01	.001
	(.104)	(.101)	(0)
Corporal	.12	.116	.004
	(.325)	(.321)	(.001)
Lieutenant+	.055	.068	-.013
	(.229)	(.252)	(.001)
Male	.962	.975	-.013
	(.19)	(.155)	(0)
Missing Rank	.014	.011	.003
	(.117)	(.106)	(0)
Missing Salary	.013	.01	.003
	(.112)	(.102)	(0)
Race/Sex Imputed	.047	.037	.01
	(.211)	(.188)	(0)
N	11897213	218813	

Notes: Unweighted means are shown. Standard deviations are in the parentheses for columns (1)–(3). Sample is restricted to stops and searches conducted from 2010–2015 by troopers with employment information (90% of searches). All statistics are generated using information reported on the stop by the trooper from stops conducted from 2010 to November 2015.

TABLE 3. Misreporting and Search Outcome by Driver's Estimated Race

	(1) Hispanic	(2) Black	(3) Asian
I(Failure)	0.0226	-0.000149	0.0156
	(0.00870)	(0.000732)	(0.0137)
Constant	0.257	0.00214	0.116
	(0.00515)	(0.000365)	(0.00626)
Observations	48780	37989	1454
F	6.729	0.0413	1.304

Notes: Dependent variable is an indicator variable equal to 1 if the recorded race does not equal to the estimated race, which I often refer to as mismatch. Standard errors are clustered at the county level. Regression uses all stops conducted from 2010 to November 2015. Each regression is run separately for motorists of each race, where race is identified using the estimated race. The regression includes fixed effects for hour of the stop, month of the stop, year of the stop, county fixed effects, and vehicle type and vehicle age. I also include the full interaction for hour of the stop with month of the stop and year of the stop and the full interaction of vehicle characteristics (vehicle type and vehicle year). Standard errors are clustered at the county FIPS and year.

TABLE 4. Correlates of estimated 'negative' bias

	(1)	(2)	(3)
Probationary Trooper	-0.037	-0.116	0.002
	(0.073)	(0.124)	(0.091)
Corporal	0.094	0.137	0.071
	(0.045)	(0.078)	(0.055)
Lieutenant+	-0.050	-0.111	-0.019
	(0.049)	(0.082)	(0.061)
Native American	0.079		
	(0.148)		
Asian	0.037		
	(0.136)		
Black	0.031		
	(0.058)		
Hispanic	0.066		
	(0.029)		
I(Male)	-0.088	-0.113	-0.059
	(0.062)	(0.090)	(0.087)
Experience	0.002	-0.002	0.003
	(0.003)	(0.005)	(0.003)
Constant	0.448	0.570	0.409
	(0.065)	(0.091)	(0.089)
Observations	1402	424	978
F	1.769	1.686	0.904

Notes: Dependent variable is likelihood the trooper's estimated Hispanic bias is negative. Trooper level estimates of Hispanic bias are generated from β_1^j from the regression $I(Mismatch_{i,c,j,t}) = \beta_0 + \beta_1^j I(Failure_{i,c,j,t}) + \delta_j + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,j,t}$, detailed in Section 5.2. All regressions are estimated with robust standard errors. Only troopers with non-missing estimates of Black and Hispanic bias are included in the regression. Column (1) is for all troopers that the data has employment, race, and rank information; column (2) is restricted to Hispanic troopers from Column (1); column (3) is restricted to all non-Hispanic troopers from Column (1).

TABLE 5. Estimated Hispanic Bias and Estimated Black Bias

	(1) Any Black Bias	(2) Black Bias
Hispanic Bias	0.153	
	(0.0650)	
Average Black Mismatch rate		-0.882
		(0.0696)
Constant	0.513	0.00587
	(0.0164)	(0.00125)
Observations	972	972
F	5.555	160.6

Notes: Dependent variable is in the column. Regressions in this table are restricted to troopers with estimates of Black bias and estimates of Hispanic bias. Estimates of Hispanic and Black bias are generated from β_1^j (Columns 1 and 2) and δ_j (average mismatch rate by race, Column 2) from the regression $I(Mismatch_{i,c,j,t}) = \beta_0 + \beta_1^j I(Failure_{i,c,j,t}) + \delta_j + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,j,t}$, detailed in Section 5.2. All regressions are estimated with robust standard errors. Only troopers with non-missing estimates of Black and Hispanic bias are included in the regression.

TABLE 6. Correlates of Hispanic Bias

	(1) All Troopers	(2) Positive Bias Only
Experience	-0.00335	-0.00327
	(0.00503)	(0.00682)
Native American	-0.0201	0.0463
	(0.140)	(0.262)
Asian	-0.180	0.277
	(0.260)	(0.182)
Black	-0.0226	0.238
	(0.0884)	(0.121)
Hispanic	-0.110	0.0134
	(0.0526)	(0.0813)
Probationary Trooper	0.294	0.530
	(0.141)	(0.162)
Corporal	0.0476	0.177
	(0.0698)	(0.135)
Lieutenant+	0.0610	0.198
	(0.110)	(0.121)
I(Male)	0.221	-0.139
	(0.114)	(0.124)
Constant	-0.214	-0.246
	(0.119)	(0.131)
Observations	1394	827
F	1.685	2.531

Notes: Regression includes controls for the work city and is clustered at the work city level. Troopers with negative levels of bias are omitted from the regression. Dependent variable is the officer level measure of bias from Equation (7). Troopers with rank equal to or higher than Lieutenant (sergeant, major, captain) were grouped into "Lieutenant +". Salary is monthly salary measured in thousands of dollars. Only troopers with rank and salary information are included in the regression. Troopers are weighted by the their total number of searches of estimated Hispanic motorists conducted from 2010 to November 2015.

TABLE 7. Positive Hispanic Bias on Labor Outcomes - Panel Results

	(1)	(2)	(3)	(4)	(5)
	Impute Left	Recorded Left	Fired	Salary Difference	Ranked Up
Hispanic Bias	0.00419	-0.0191	-0.397	-0.0160	-0.0445
	(0.00532)	(0.0160)	(0.275)	(0.0239)	(0.0454)
I(Male)	0.00321	-0.0103	-0.602	0.0848	-0.136
	(0.00282)	(0.0416)	(0.170)	(0.0564)	(0.119)
Native American	0.000676	-0.0311		-0.0366	-0.202
	(0.00340)	(0.0129)		(0.0994)	(0.0659)
Asian	-0.00419	-0.0341		-0.0996	0.0515
	(0.00454)	(0.0159)		(0.0840)	(0.122)
Black	0.00154	-0.0357		0.0748	-0.0806
	(0.00230)	(0.0145)		(0.0307)	(0.0809)
Hispanic	0.00212	0.0183	0.596	0.00854	0.0500
	(0.00380)	(0.0274)	(0.145)	(0.0364)	(0.0658)
Corporal	0.0122	0.0144	-0.184	0.0391	0.439
	(0.0127)	(0.0402)	(0.0980)	(0.0391)	(0.0903)
Lieutenant+	-0.000447	-0.0151	-0.899	0.370	0.564
	(0.000794)	(0.0308)	(0.233)	(0.0707)	(0.0907)
Experience	-0.000808	0.00411	-0.0328	-0.00393	-0.0291
	(0.000598)	(0.00291)	(0.0119)	(0.00377)	(0.00561)
Constant	0.00209	0.0120	1.459	0.803	0.540
	(0.00425)	(0.0468)	(0.141)	(0.0663)	(0.129)
Observations	517	493	22	486	515
F	0.216	1.458	.	5.457	11.74

Notes: Regression has robust standard errors. Dependent variable is the officer level measure of bias from Equation (7) using only stops from 2010 to 2013. Employment outcomes are from 2014–November 2015. Each regression includes controls for the trooper's gender, trooper's rank in 2013, and trooper race. Hispanic bias is normalized with mean 0 and standard deviation of 1. Impute left is defined as not observing any searches in the 2014–2015 traffic stop data. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. Troopers are weighted by the their total number of searches of estimated Hispanic motorists conducted from 2010–2013. Employment data is missing rank and salary for some observations leading to differences in number of observations.

TABLE 8. Positive Hispanic Bias on Complaints

	(1) Any Complaint	(2) Sustained Complaint
Hispanic Bias	0.0389	0.0385
	(0.0166)	(0.0165)
Constant	0.0752	0.0723
	(0.0162)	(0.0161)
Observations	851	851
F	5.519	5.405

Notes: Dependent variable is an indicator variable equal to one if the trooper had a sustained complaint from 2010 to November 2015. Hispanic bias is normalized and is estimated from Equation (7), β_1^j. Regression has robust standard errors and restricted to troopers with positive bias only.

TABLE 9. Positive Hispanic Bias on 2019 Labor Outcomes

	(1) Left Force	(2) Salary Change	(3) Ranked Up
Bias	0.0369	-0.0856	-0.0886
	(0.0283)	(0.0428)	(0.0437)
Experience	0.00533	-0.0407	-0.00239
	(0.00477)	(0.00523)	(0.00743)
Constant	0.193	0.811	0.103
	(0.106)	(0.101)	(0.184)
Observations	827	648	648
F	1.451	19.30	61.92

Notes: Regression has robust standard errors shown in parentheses and uses 2019 employment data posted publicly by the Texas Tribune. Each regression controls for the trooper's gender, trooper's maximum rank from 2013 to 2015, and trooper race. Hispanic bias is normalized and is estimated from Equation (7), β_1^j. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. Troopers are weighted by the their total number of searches of estimated Hispanic motorists conducted from 2010 to November 2015. Regressions in columns (2) and (3) are restricted to being observed in the 2019 employment data.

TABLE 10. Search Success Rates across Driver's Race - All troopers

	Search Success Rate		
	(1)	(2)	(3)
	Recorded	Estimated	Δ
Driver Race			
Asian	.372	.356	.016
	(.484)	(.479)	(.015)
Black	.411	.411	0
	(.492)	(.492)	(.004)
Hispanic	.277	.247	.03
	(.447)	(.431)	(.003)
White	.381	.402	-.021
	(.486)	(.49)	(.002)

Notes: Unweighted means are shown. Standard deviations are in the parentheses. Columns (1) and (2) show the search success rates using the recorded race and the data driven race estimation using driver's last name and residential ZIP code. Race estimation steps are detailed in Section 4.3. Column (3) shows the difference between Columns (1) and (2) with the two-sample t-statistic in the parentheses. Statistics are generated using searches conducted from 2010 to November 2015.

TABLE 11. Search Success Rates across Driver's Race - Positive Bias only

	Search Success Rate		
	(1)	(2)	(3)
	Recorded	Estimated	Δ
Driver Race			
Asian	.38	.362	.018
	(.486)	(.481)	(.021)
Black	.429	.429	0
	(.495)	(.495)	(.005)
Hispanic	.318	.265	.053
	(.466)	(.441)	(.004)
White	.395	.422	-.027
	(.489)	(.494)	(.003)

Notes: The sample is restricted to troopers with any Hispanic bias ($\hat{\beta_i^j} > 0$ from Eq. (7)). Unweighted means are shown. Standard deviations are in the parentheses. All other notes from Table 10 apply.

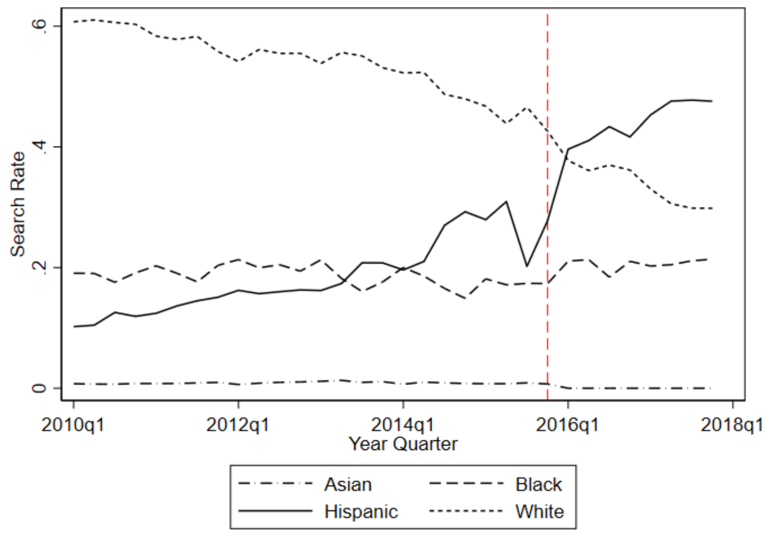
TABLE 12. Hit Rate Test in the presence of misreporting

	(1) All troopers	(2) Positive troopers only
Recorded Asian	-0.0526	-0.0522
	(0.0184)	(0.0144)
Recorded Black	0.00806	0.0119
	(0.00501)	(0.00617)
Recorded Hispanic	-0.0469	-0.0185
	(0.00718)	(0.00992)
Recorded Other	-0.0335	-0.0450
	(0.00349)	(0.00896)
Hispanic Bias	-0.00796	-0.000426
	(0.00601)	(0.00545)
Recorded Asian × Hispanic Bias	0.00324	0.0170
	(0.0224)	(0.0195)
Recorded Black × Hispanic Bias	0.0107	-0.000164
	(0.00619)	(0.00580)
Recorded Hispanic × Hispanic Bias	0.0532	0.0341
	(0.00703)	(0.00734)
Recorded Other × Hispanic Bias	-0.0227	-0.00687
	(0.00959)	(0.00523)
Constant	0.393	0.389
	(0.000764)	(0.00163)
Observations	194717	117497

Notes: Unweighted means are shown. Standard deviations are in the parentheses. Dependent variable is likelihood of search success or finding contraband. Column (2) is restricted to troopers with estimated Hispanic bias ($\hat{\beta_i^j} > 0$ from Equation 7. Hispanic bias is normalized. The regression includes fixed effects for hour of the stop, month of the stop, year of the stop, county fixed effects, and vehicle type and vehicle age. I also include the full interaction for hour of the stop with month of the stop and year of the stop and the full interaction of vehicle characteristics (vehicle type and vehicle year). The omitted race group is Recorded White motorists. Standard errors are clustered at the county FIPS and year. Regression is restricted to searches conducted from 2010 to November 2015

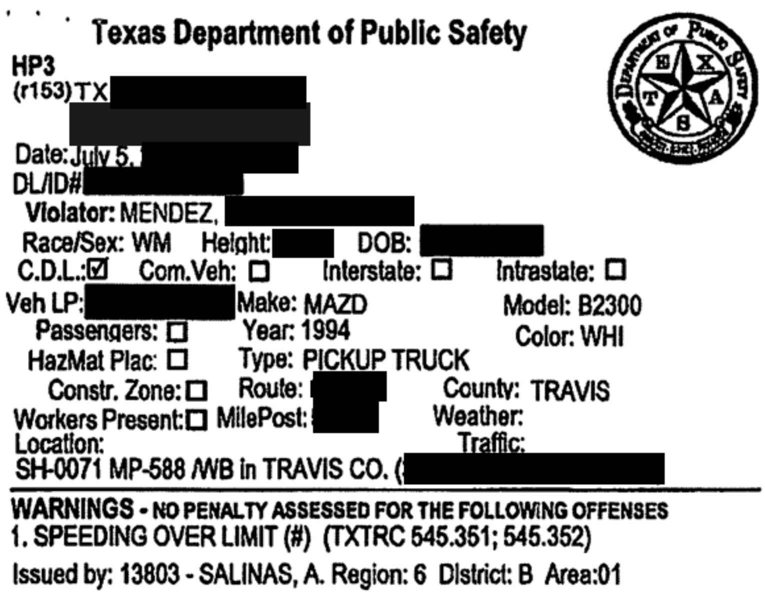
Figures

Figure 1. Quarterly Search Rate by Driver's Race using Recorded Races



Notes: Average, unweighted search rates, by reported motorist race, for a given quarter-year from January 2010 to December 2015 are shown. The dashed red, vertical line indicates the quarter when KXAN published the article revealing the trooper race misreporting..

Figure 2. Example of misreported Highway Ticket



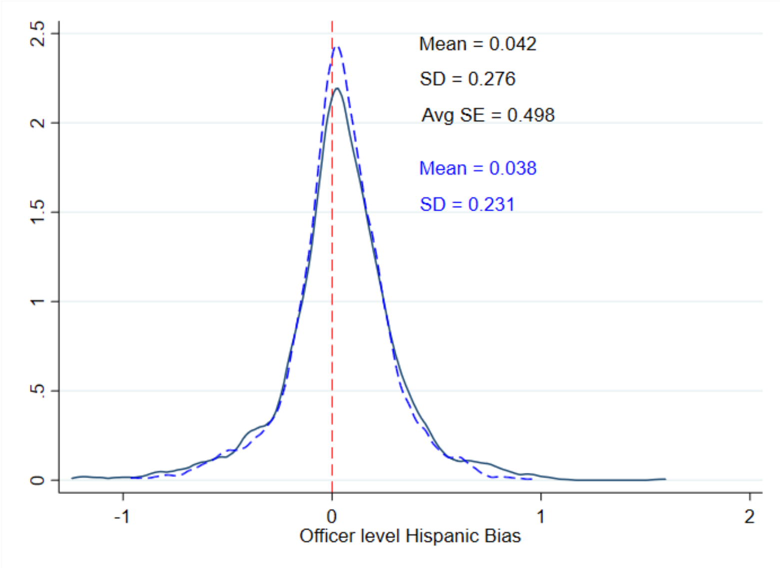
Texas Department of Public Safety

HP3
(r153) TX [redacted]

Date: July 5, [redacted]
DL/ID# [redacted]
Violator: MENDEZ, [redacted]
Race/Sex: WM Height: [redacted] DOB: [redacted]
C.D.L.: ☑ Com.Veh: ☐ Interstate: ☐ Intrastate: ☐
Veh LP: [redacted] Make: MAZD Model: B2300
Passengers: ☐ Year: 1994 Color: WHI
HazMat Plac: ☐ Type: PICKUP TRUCK
Constr. Zone: ☐ Route: [redacted] County: TRAVIS
Workers Present: ☐ MilePost: [redacted] Weather:
Location: Traffic:
SH-0071 MP-588 /WB in TRAVIS CO. ([redacted]

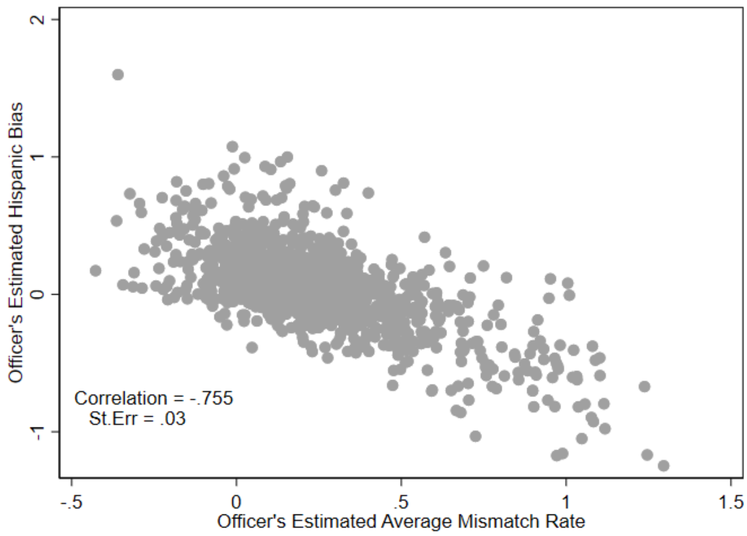
WARNINGS - NO PENALTY ASSESSED FOR THE FOLLOWING OFFENSES
1. SPEEDING OVER LIMIT (#) (TXTRC 545.351; 545.352)
Issued by: 13803 - SALINAS, A. Region: 6 District: B Area:01

Figure 3. Trooper level estimates of bias



Notes: Kernel density distribution of officer-level estimates of Hispanic bias. The figure plots each officer's β_1^j from the regression $I(Mismatch_{i,c,j,t}) = \beta_0 + \beta_1^j I(Failure_{i,c,j,t}) + \delta_j + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,j,t}$. The solid, black line is using the original estimates; the dashed, blue line is the Bayes shrinkage procedure (Weisburst, 2022).

Figure 4. Correlation of trooper's level of bias with trooper's average race mismatch rate



Notes: The figure plots each officer's β_1^j against each officer's estimated δ_j from the regression $I(Mismatch_{i,c,j,t}) = \beta_0 + \beta_1^j I(Failure_{i,c,j,t}) + \delta_j + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,j,t}$. The correlation between the two variables is also shown along with the standard error.

A Appendix

A.1 Discussion of the Model

Proof of Proposition 1

Suppose some motorist, z, with characteristics (M, θ) with $\theta_z > \theta^{\mu}_{M,t}$ is pulled over by trooper t. Then this implies

$$c_{M,t} + \mu_{M,t}(\theta_z) > c_{W,t}$$

Therefore, the trooper will not misreport motorists with $\theta > \theta^{\mu}$, regardless of the search outcome, G. If $G = 1$, under Assumption 3, $\mu_{\theta,G=1} > 1$. This implies

$$c_{M,t} + \mu_{M,t}(\theta, G = 1) > c_{W,t}$$

Therefore the trooper will not misreport motorists if the search ends in success ($G = 1$) regardless of the characteristics, (M, θ), of the motorist.

For motorists with $\theta < \theta^*$, this is not sufficient for search, therefore the trooper will also never misreport motorists with $\theta < \theta^*$.

Proof of Proposition 2

Suppose trooper i and trooper j are biased against race M motorist, but trooper i is more biased such that $c_{M,i} < c_{M,j}, c_{W,i} = c_{W,j}$, and $c_{M,t} < c_{W,t}$ for $t \in \{i, j\}$. Since both troopers face the same population of race-M motorist and race-W motorist, then this implies that $\theta^{\mu}_{M,i} > \theta^{\mu}_{M,j}$, $\theta^*_{M,i} < \theta^*_{M,j}$, and $\theta^*_{W,i} = \theta^*_{W,j}$. From Assumptions 1, 2, and 3, this implies that:

$$\begin{aligned}
&\Rightarrow \theta^{\mu}_{M,i} - \theta^*_{M,i} > \theta^{\mu}_{M,j} - \theta^*_{M,j} \\
&\Rightarrow (1 - \pi_M)[F^M_n(\theta^{\mu}_{M,i}) - F^M_n(\theta^*_{M,i})] > (1 - \pi_M)[F^M_n(\theta^{\mu}_{M,j} - \theta^*_{M,j})] \\
&\Rightarrow v_{M,i} > v_{M,j} > 0
\end{aligned}$$

Thus, since trooper i is more biased than trooper j, trooper i also misreports a higher portion of race M searches than trooper j.

Relaxing Assumption 3

While Assumption 3 is fairly intuitive, specifically that misreporting is only profitable when the search ends in failure, it is not a necessary condition for using misreporting as a measure of bias. From Assumption 1 and 2, the average, misreporting rate for trooper t is:

$$\phi_{M,t} = \frac{\pi_M[F_g^M(\theta^\mu_{M,t}) - F_G^M(\theta^*_{M,t})] + (1-\pi_M)[F_n^M(\theta^\mu_{M,t}) - F_n^M(\theta^*_{M,t})]}{\pi_M[1 - F_g^M(\theta^*_{M,t})] + (1-\pi_M)(1 - F_n^M(\theta^*_{M,t})]} \tag{8}$$

Proposition 3. *From Assumptions 1 and 2, if a trooper exhibits racial bias against race M motorists, then $\phi_{M,t} > 0$.*

Further,

Corollary 1. *The misreporting rate, $\phi_{M,t}$, is the magnitude of bias against race M motorist.*

The proof is in the section below.

Since the distributions, f_g^m and f_n^m, and the true proportion of guilty motorists, π_m are unobservable, the misreporting rate cannot be directly measured. Instead, the misreporting rate can be derived from the observed, average search rate and the true, average search rate. The difference between the observed and true search rate vary depending on whether the trooper is racially biased or not.

The true, average search rate $\gamma_{m,t}$ for race m motorists is as follows:

$$\gamma_{m,t} = \pi_m[1 - F_g^m(\theta^*_{m,t})] + (1-\pi_m)[1 - F_n^m(\theta^*_{m,t})] \tag{9}$$

Let $\gamma^O_{m,t}$ denote trooper t's observed, average search rate of race m motorist.

From Proposition 1, only a portion of race M motorists are misreported, specifically, unsuccessful searches of race M motorists of $\theta \in (\theta^*, \theta^\mu)$ if the trooper is biased. Thus the observed search rate, composed of the correctly recorded race M motorists, is:

$$\gamma^O_{M,t} = \pi_M[1 - F_g^M(\theta^\mu_{M,t})] + (1-\pi_M)[1 - F_n^M(\theta^\mu_{M,t})] \tag{10}$$

Since motorists of characteristics (M, θ) where $\theta \in (\theta^*, \theta^\mu)$ are misreported, the observed search rate for race M motorists is lower than the true search rate for race M motorists.

Misreporting also affects the search rate for race W motorists. The inclusion of race M motorists miscategorized as race W will affect the search rate for race W motorists in the following way:

$$\begin{gathered}\gamma^O_{W,t} = \pi_W[1 - F_g^W(\theta^*_{W,t})] + (1-\pi_W)(1 - F_n^W(\theta^*_{W,t})] \\ +\pi_M[F_g^M(\theta^\mu_{M,t}) - F_G^M(\theta^*_{M,t})] + (1-\pi_M)[F_n^M(\theta^\mu_{M,t}) - F_n^M(\theta^*_{M,t})]\end{gathered} \tag{11}$$

Therefore, the misreporting rate, $\phi_{M,t}$, can be rewritten in terms of the observed search rates:

$$\phi_{M,t} = \frac{\gamma^{O}_{W,t} - \gamma_{W,t}}{\gamma_{M,t}} \tag{12}$$

While this test and measure of racial bias relies on fewer assumptions, it requires knowing the true search rate which is often unobservable. Thus, I will include my results using this measure of racial bias once I better my race estimation methods.

Proof of Proposition 2 and Corollary 3

The magnitude of this misreporting rate also yields a measure of bias. For example, suppose trooper i and trooper j are biased against race M motorist, but trooper i is more biased such that $c_{M,i} < c_{M,j}, c_{W,i} = c_{W,j}$, and $c_{M,t} < c_{W,t}$ for $t \in \{i, j\}$. Since both troopers face the same population of race-M motorist and race-W motorist, then this implies that $\theta^{\mu}_{M,i} > \theta^{\mu}_{M,j}$ and $\theta^{*}_{M,i} < \theta^{*}_{M,j}$. From Proposition 1, this implies that:

$$\Rightarrow \theta^{\mu}_{M,i} - \theta^{*}_{M,i} > \theta^{\mu}_{M,j} - \theta^{*}_{M,j}$$
$$\Rightarrow \phi_{M,i} > \phi_{M,j}$$

Thus, since trooper i is more biased than trooper j, trooper i also misreports a higher portion of race M searches than trooper j.

A.2 Results using all troopers

TABLE A1. Hispanic Bias on Labor Outcomes - Panel Results

	(1) Impute Left	(2) Recorded Left	(3) Fired	(4) Salary Difference	(5) Ranked Up
Bias	-0.000722	-0.0210	0.0262	0.0202	0.00335
	(0.00310)	(0.0179)	(0.0764)	(0.0130)	(0.0224)
I(Male)	0.00263	-0.0681	0.528	0.0216	-0.0615
	(0.00227)	(0.0756)	(0.264)	(0.0564)	(0.0911)
Native American	-0.000172	0.0350	-0.546	-0.0636	-0.0459
	(0.000998)	(0.0709)	(0.156)	(0.0818)	(0.0747)
Asian	0.0000714	-0.0542		-0.0144	0.0484
	(0.00184)	(0.0168)		(0.0457)	(0.0617)
Black	0.000965	0.103	-0.168	-0.0907	0.199
	(0.00144)	(0.117)	(0.144)	(0.0562)	(0.190)
Hispanic	0.00341	0.0427	0.423	-0.00917	0.0839
	(0.00291)	(0.0334)	(0.145)	(0.0221)	(0.0408)
Probationary Trooper	-0.00316				0.790
	(0.00218)				(0.0680)
Corporal	0.00524	-0.000102	-0.126	0.0333	0.339
	(0.00614)	(0.0316)	(0.130)	(0.0277)	(0.0724)
Lieutenant+	-0.00167	-0.0566	-0.735	0.451	0.607
	(0.00142)	(0.0278)	(0.148)	(0.0523)	(0.0677)
Experience	-0.000448	0.00315	-0.0277	0.00426	-0.0139
	(0.000266)	(0.00281)	(0.0103)	(0.00231)	(0.00422)
Constant	0.00158	0.0749	0.227	0.769	0.266
	(0.00173)	(0.0761)	(0.268)	(0.0595)	(0.0958)
Observations	1063	1019	48	1001	1060

Notes: Regression has robust standard errors. Dependent variable is the officer level measure of bias from Equation (7) using only stops from 2010 to 2013. Employment outcomes are from 2013 and 2014. Includes controls for the trooper's gender, trooper's rank in 2013, and trooper race. Hispanic bias is normalized. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. Each trooper is weighted by their total number of searches of estimated Hispanic drivers conducted from 2010 to November 2015. Employment data is missing rank and salary for some observations leading to differences in number of observations.

TABLE A2. Hispanic Bias on Labor Outcomes - after 2015

	(1) Left Force	(2) Salary Change	(3) Ranked Up
Bias	-0.00555	0.00712	0.00501
	(0.0248)	(0.0250)	(0.0260)
Experience	0.00603	-0.0345	-0.00405
	(0.00379)	(0.00440)	(0.00557)
Native American	-0.0735	0.00583	-0.0822
	(0.0818)	(0.155)	(0.226)
Asian	-0.0170	-0.195	-0.291
	(0.120)	(0.0723)	(0.0667)
Black	0.207	-0.00237	-0.00816
	(0.116)	(0.105)	(0.106)
Hispanic	0.0854	-0.0501	0.00950
	(0.0458)	(0.0616)	(0.0695)
Probationary Trooper	0.0271	0.783	0.599
	(0.0820)	(0.0918)	(0.0498)
Corporal	-0.0700	0.0549	0.00899
	(0.0476)	(0.0756)	(0.0870)
Lieutenant+	-0.0480	-0.285	-0.256
	(0.0682)	(0.0885)	(0.0659)
I(Male)	-0.158	-0.116	0.0500
	(0.0975)	(0.0707)	(0.0879)
Constant	0.279	1.188	0.351
	(0.102)	(0.0814)	(0.0997)
Observations	1394	1100	1100
F	1.523	30.47	120.6

Notes: Regression has robust standard errors show in parentheses and uses 2019 employment data posted publicly by the Texas Tribune. Includes controls for the trooper's gender, trooper's maximum rank from 2010 to 2015, and trooper race. Each trooper is weighted by their total number of searches of estimated Hispanic drivers conducted from January 2010 to November 2015. Black and Hispanic are indicator variables equal to one if the trooper is Black or Hispanic, respectively, and equal to zero otherwise. Hispanic bias is normalized. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000.

A.3 Results using shrunken bias estimates

TABLE A3. Correlates of Hispanic Bias

	(1) All Troopers	(2) Positive Bias Only
Experience	-0.00376	-0.00467
	(0.00584)	(0.00799)
Native American	-0.0347	0.00145
	(0.150)	(0.255)
Asian	-0.214	0.331
	(0.303)	(0.217)
Black	-0.0312	0.246
	(0.101)	(0.131)
Hispanic	-0.123	0.0345
	(0.0619)	(0.0965)
Probationary Trooper	0.305	0.491
	(0.151)	(0.177)
Corporal	0.0615	0.222
	(0.0818)	(0.161)
Lieutenant+	0.0663	0.243
	(0.130)	(0.142)
I(Male)	0.252	-0.136
	(0.131)	(0.147)
Constant	-0.237	-0.214
	(0.136)	(0.155)
Observations	1394	827
F	1.611	2.243

Notes: Regression includes controls for the work city and is clustered at the work city level. Troopers with negative levels of bias are omitted from the regression. Dependent variable is the shrunken, officer level measure of bias from Equation (7). Troopers with rank equal to or higher than Lieutenant (sergeant, major, captain) were grouped into "Lieutenant +". Salary is monthly salary measured in thousands of dollars. Only troopers with rank and salary information are included in the regression. Troopers are weighted by the their total number of searches of estimated Hispanic motorists conducted from 2010 to November 2015.

TABLE A4. Hispanic Bias on Labor Outcomes - Panel Results

	(1) Impute Left	(2) Recorded Left	(3) Fired	(4) Salary Difference	(5) Ranked Up
stdbias_p	0.00397	-0.0160	-0.404	-0.0108	-0.0378
	(0.00484)	(0.0143)	(0.242)	(0.0219)	(0.0407)
I(Male)	0.00319	-0.0126	-0.581	0.0827	-0.141
	(0.00286)	(0.0415)	(0.164)	(0.0565)	(0.120)
Native American	0.000699	-0.0279		-0.0310	-0.192
	(0.00305)	(0.0115)		(0.0991)	(0.0658)
Asian	-0.00414	-0.0320		-0.0947	0.0593
	(0.00455)	(0.0153)		(0.0842)	(0.120)
Black	0.00154	-0.0342		0.0760	-0.0740
	(0.00215)	(0.0141)		(0.0310)	(0.0822)
Hispanic	0.00194	0.0191	0.599	0.0149	0.0597
	(0.00343)	(0.0266)	(0.139)	(0.0358)	(0.0642)
Corporal	0.0121	0.0161	-0.195	0.0382	0.448
	(0.0124)	(0.0396)	(0.0918)	(0.0387)	(0.0890)
Lieutenant+	-0.000430	-0.0155	-0.924	0.373	0.585
	(0.000732)	(0.0294)	(0.232)	(0.0676)	(0.0863)
Experience	-0.000804	0.00428	-0.0319	-0.00332	-0.0290
	(0.000597)	(0.00291)	(0.0111)	(0.00376)	(0.00557)
Constant	0.00196	0.00973	1.453	0.794	0.533
	(0.00399)	(0.0466)	(0.130)	(0.0661)	(0.130)
Observations	526	500	22	493	524
F	0.215	1.419	.	6.085	12.99

Notes: Regression has robust standard errors. Dependent variable is denoted in the column. Bias is the the shrunken, officer level measure of bias from Equation (7) using only stops from 2010 to 2013. Employment outcomes are from 2013 and 2014. Each regression includes controls for the trooper's gender, trooper's rank in 2013, and trooper race. Hispanic bias is normalized. Impute left is defined as not observing any searches in the 2014–2015 traffic stop data. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. Each trooper is weighted by their total number of searches of estimated Hispanic drivers conducted from January 2010 to December 2013. Regression is restricted to troopers with positive bias. Employment data is missing rank and salary for some observations leading to differences in number of observations.

TABLE A5. Hispanic Bias on Complaints

	(1) Any Complaint	(2) Sustained Complaint	(3) Any Complaint	(4) Sustained Complaint
Hispanic Bias	0.0268	0.0262		
	(0.0110)	(0.0110)		
Hispanic Bias			0.0366	0.0362
positive only			(0.0158)	(0.0158)
Constant	0.0604	0.0586	0.0730	0.0702
	(0.0107)	(0.0107)	(0.0159)	(0.0158)
Observations	1433	1433	851	851
F	5.908	5.704	5.357	5.256

Notes: Dependent variable is an indicator variable equal to one if the trooper had a sustained complaint from 2010 to 2015. Regression has robust standard errors. Each trooper is weighted by their total number of searches of estimated Hispanic drivers conducted from January 2010 to November 2015. Bias is the the shrunken, officer level measure of bias from Equation (7) using only stops from 2010 to November 2015. Columns (1) and (2) are for all troopers; Columns (3) and (4) is restricted to troopers with positive bias only.

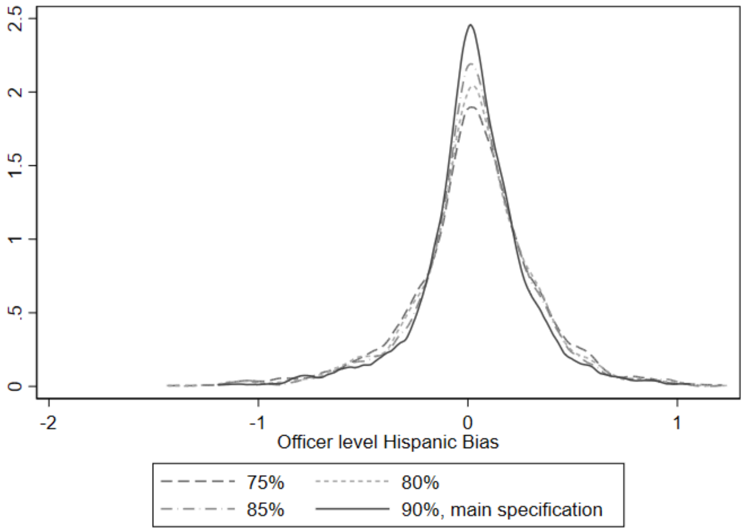
TABLE A6. Hispanic Bias on Labor Outcomes on 2019 labor outcomes

	(1) Left Force	(2) Salary Change	(3) Ranked Up
Bias	0.0327	-0.0795	-0.0781
	(0.0260)	(0.0394)	(0.0403)
Experience	0.00536	-0.0408	-0.00247
	(0.00477)	(0.00522)	(0.00742)
Constant	0.191	0.820	0.109
	(0.106)	(0.101)	(0.183)
Observations	827	648	648
F	1.410	19.21	61.93

Notes: Regression has robust standard errors show in parentheses and uses 2019 employment data posted publicly by the Texas Tribune. Includes controls for the trooper's gender, trooper's maximum rank from 2010 to 2015, and trooper race. Each trooper is weighted by their total number of searches conducted from January 2010 to June 2015. Bias is the the shrunken, officer level measure of bias from Equation (7) using only stops from 2010 to 2015. Hispanic bias is normalized. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. Each trooper is weighted by their total number of searches of estimated Hispanic drivers conducted from January 2010 to November 2015. Regression is restricted to troopers with positive bias.

A.4 Appendix Figures

Figure A.5. Officer-level estimates of Hispanic bias with different thresholds



Notes: Each density shows the unshrunk, officer level bias using different levels of surname cutoff and ZIP code density cutoff (surname threshold - 15%). The estimate of bias is from each officer's β^j from the regression $I(Mismatch_{i,t}) = \alpha + \beta^j I(Failure)_{i,t} + \delta_j + X_{i,c,t}\gamma + \epsilon_{j,t}$. . Bias is estimated using searches conducted from 2010 to November 2015.