Not So Black and White: Uncovering Racial Bias from Systematically Misreported Trooper Reports

(Latest Version Available Here)

Elizabeth Luh

April 30, 2020

Abstract

Highway police officers, or troopers, may misreport the race of people that they engage with in order to evade detection of racial bias. I propose a new test of racial bias in the presence of misreporting that is well-suited to explore the rich heterogeneity in bias behavior. Using a unique event in Texas where troopers were caught deliberately misreporting minority motorists as white, I find bias against all minority motorists, but especially against Hispanic motorists. I estimate bias for each trooper and find that over 30% of troopers were engaging in this behavior. Using my trooper level measure of bias, I identify causal relationships between bias and labor outcomes using a panel data set of trooper employment outcomes. I show troopers were able to use misreporting to evade detection of bias, with bias having no effect on labor market outcomes when the misreporting was possible. I find that after a rule change to trooper stop recording policy in response to the discovery of misreporting led to negative labor outcomes for biased troopers, specifically, lower rates of promotion and lower rates salary growth. I further test how individual trooper bias changes in response to changes in peer, demographic compositions. In particular, black or Hispanic troopers are sensitive to changes in peer composition, while white troopers are unaffected.

1 Introduction

In the United States, individuals experience differential treatment by law enforcement officers depending on their race through all steps of the criminal justice system. Recent, highly publicized and deadly interactions between motorists and police has raised concern by policy makers and the general public alike.[1] According to a series of surveys in 2017, titled *Discrimination in America*, 27% of Latinos and 50% of blacks surveyed felt personally discriminated against when interacting with police compared to only 10% of white respondents.[2] Recent research has also linked these disparities to racial bias and discrimination in other steps of the justice system from airport screening (Persico and Todd, 2005), ticketing (Anbarci and Lee, 2014; Goncalves and Mello, 2017), stop and frisk participation (Coviello and Persico, 2013), bail decisions (Arnold et al., 2018), sentencing (Shayo and Zussman, 2011; Depew et al., 2017), parole (Anwar and Fang, 2015), and capital punishment (Alesina and Ferrara, 2014).

Despite the focus of policy makers and researchers on testing for racial bias and discrimination in the criminal justice system, little attention has been given to the response of law enforcement officers to this heightened scrutiny. These behavioral responses are important as law enforcement officers have considerable discretion when recording interactions. Discriminating officers may find it more costly to be fair in their treatment of individuals by race and instead misreport their interactions with motorists to appear less discriminatory. Thus, if the most discriminatory officers are most likely to misreport, then any policies targeting discrimination may miss the officers who need the most discipline and training. Even more concerning, with widespread misreporting, agencies and institutions could be underestimating the presence of discrimination within their department.

In this paper, my objective is to link this discriminatory misreporting behavior to racial bias and to construct an individual measure of bias using the frequency of the

[1]For example, Sandra Bland in 2015, Philando Castile in 2016, and Terence Crutcher in 2016.

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

misreporting behavior. I use a documented instance where Texas Highway Police Officers (henceforth "troopers") were found to be misreporting minority motorists' race as white (Collister, 2015b). This misreporting was possible because from 2010 to 2015, Texas troopers were allowed to record motorist race based on their own best judgment. Thus, a discrepancy between motorists' actual race and the recorded race was possible. Troopers were able to miscategorize a portion of their minority motorist searches as white, which decreased the observed stop rate of non-white motorists (Collister, 2015b).

Studying racial bias in the presence of misreporting is advantageous for several reasons. One concern when testing for bias, is that unobservable variables, such as driver demeanor, can affect the outcome of trooper-motorist interactions. In the context of this paper, troopers' choice to misreport a portion of their searches creates an observable variation within minority searches that I can link to bias. Another issue is that a trooper's racial bias is always measured relative to their peers, since the true culpability of the motorist is unknown. This only allows for measures of relative bias or aggregate bias. Using the frequency of misreporting creates an absolute measure of bias that is independent of the bias of the trooper's peers. Further, these repeated observations of misreporting within each trooper allow me to create a measure of racial bias that is on the individual, trooper level.

The first step of this paper is to demonstrate how misreporting can be used as a measure of racial bias. In my model, I show that only troopers biased against minority motorists have incentive to misreport minority searches as white. Further, unbiased troopers have no incentive to misreport at all. The intuition being that misreporting reduces the likelihood of punishment for racial bias. Specifically, my model shows that biased troopers will only misreport their failed, minority searches thereby appearing less biased. The major challenge in existing literature is disentangling statistical discrimination, or the difference in criminality by motorist race, from racial bias (Knowles et al., 2001; Anwar and Fang, 2006; Antonovics and Knight, 2009). I am able to overcome this challenge since misreporting is linked to only biased behavior.

The second step is to determine which searches in the data were purposefully misreported. Troopers may accidentally misreport motorists' race for reasons aside from bias (i.e poor visibility) and some motorists who appear non-white in the data, may self-identify as white. While the data can provide estimates of the driver's race, the true race and whether the misreporting was purposeful is unobserved. To deal with this, I estimate the driver's race using the data and measure the rate of mismatch within each trooper, where mismatch is when the estimated race does not match the recorded race. I then compare the different rates of mismatch of failed searches and successful searches. From my model, unbiased troopers' rate of mismatch should not differ across search outcome. Whereas for biased troopers, the rate of mismatch will be higher for failed searches than for successful searches. Therefore troopers' bias is the difference in mismatch rate by search outcome.

To test the validity of my measure of trooper bias, I exploit the discovery of the misreporting and the subsequent rule change at the end of 2015 requiring troopers to verbally ask motorists' for their race. I show that biased troopers were statistically more likely than unbiased troopers to record a minority motorist as white when the search ended in failure prior to the rule change. After the rule change, this difference disappears. These results supports the interpretation of my measure as capturing bias and that troopers were indeed misreporting motorists who did not self-identify as white.

Using a simple regression framework, I find that misreporting was used most frequently against Hispanic motorists, who were 3.6 percent more likely to be misreported when the search ended in failure compared to when the search ended in success. I estimate that over 32,000 searches were mismatched for Hispanic motorists, which is more than half of all estimated-Hispanic searches. Misreporting was used more effectively against black and Asian motorists with a differential misreporting rate of 20 percent and 40 percent respectively. But, because Asian and black motorists make up a small proportion of mismatched motorists in my data set, I focus my results on the individual level to

Hispanic bias.[3]

My main contribution is to provide an estimate of racial bias at the individual, trooper level. I use each trooper's differential likelihood of misreporting a Hispanic motorist as white by search outcome as the trooper's measure of bias. This measure of bias is robust when including controls for seasonality, county characteristics, and stop characteristics. I find that the average trooper is 17% more likely to misreport Hispanic motorists as white when the search ends in failure compared to success. Correlating trooper characteristics to bias, I find that lower paid, recently hired troopers are less racially biased. Surprisingly, I find that racial bias is not significantly correlated with trooper race.

I also identify effects of bias on troopers' labor market outcomes, with misreporting and without misreporting. I find that prior to the discovery of misreporting by various media outlets starting in 2015, bias had no effect on labor market outcomes for troopers who remained in the force. After the discovery of misreporting, every standard deviation in bias, which is measured using stops pre-rule change, decreased the likelihood of promotion by 27 percent and decreased salary growth by 8 percent. Since these negative outcomes occur after the rule change, I interpret this as suggestive evidence that misreporting was effective in shielding biased troopers from punishment for bias before the change.

I show that past tests of bias would underestimate the presence of bias in Texas when using the observed races in the data. By comparing average search outcomes by motorists' race with recorded and estimated races, I find that misreporting distorted search success rates significantly for Asian, Hispanic, and white motorists, with strongest effects on the search success rate for white motorists.[4] While the misreporting increased the search success rates for Asian and Hispanic motorists by 5 percent and 3.3 percent, respectively; the white search success rates decreased by 12 percent. The disproportionate decrease for white motorists is due to significantly fewer failed white searches compared

[3]Asian and black motorists make up only $< 2\%$ of all mismatched searches in the data set.

[4]Hispanic is technically an ethnicity and not a race. But, the context of this study is in Texas, which codifies Hispanic as its own race when recording motorist stops. Therefore for simplicity, I will oftentimes refer to Hispanic as a race.

to non-white searches.

Finally, I use the trooper level measure of Hispanic to understand how the trooper's work environment affects own bias. I find suggestive, but not significant evidence of differing effects of work environment, specifically peer's demographic composition, across troopers' race. My results show that black troopers' level of bias increased when proportion of white troopers they work with rises. Whereas white and Hispanic troopers' bias was unaffected by demographic changes in peers.

This paper contributes to the literature on motorist stops and whether the decision to search is influenced by racial bias. There are many earlier contributions to the literature that 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. The main issue with this is that since these tests rely on the comparison across groups, they can only test for relative bias and not absolute bias. Further, they cannot measure the magnitude of bias on the individual, trooper level. Relative to this literature, my approach does not rely on the relative behavior to measure racial bias and can also measure racial bias on the individual level. My paper is also the first to address the possibility of the data being purposefully misreported to hide bias.

Another popular method of identifying racial bias in highway and traffic stops uses plausibly exogenous characteristics of the stop. One example is the 'veil of darkness,' which uses the diminished ability of trooper's to observe the motorist race after sunset. Kalinowski et al. (2017) use the differing speed distributions in daylight and darkness to test for bias against African Americans. Another method in the context of police ticketing decision, by West (2018), uses the plausibly exogenous assignment of police officers to traffic accidents to identify a causal relationship between the actions of police officers by driver race. The major drawback to both of these identification strategy is that the results are context specific and may not apply to a wider range of motorist and officer interactions.

My research is also related to empirical research in cheating behavior since misreporting in the context of trooper reports is a form of cheating. 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 (2017) also identifies racial bias at the individual level using police officers' choice to be lenient 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 having a comparison group of unbiased officers to measure bias. Further, their framework does not explicitly link the disparity in leniency across motorists' race to racial bias.

Given that mine is possibly the first to document such severity of misreporting behavior in troopers, this may provide motivation to reexamine past work in racial bias in motorist stops. With misreporting, past literature may still be under-detecting the existence of bias. 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 2800 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. Figure 1 shows the division map across Texas.

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 being permanent assignments. After the one year probationary period, troopers take their final exam and are promoted to trooper.

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 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.

Due to Texas' proximity to the Mexican border, Texas Highway Patrol heavily participates in increased law enforcement along the shared border. Since 2014, DPS has sent troopers from across Texas to the border to serve for approximately one week through various operations. Often, these operations are multi-department efforts such as Operation Strong Safety, which was conducted jointly between DPS, Texas State Guard, and the National Guard (Benen, 2014). The main goal of the operation was to reduce drug trafficking and undocumented immigration across the Texas-Mexico border (HSIN, HSIN).

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. If the officer believes that the driver or passenger has a weapon, he or she may pat down the person and search the vehicle and the surrounding immediate area. Motorists cannot physically resist a search but can notify the officer that he or she does not consent.

Drivers can report troopers if they feel that troopers behaved inappropriately during a stop and troopers can face repercussions if the claim is substantiated. Troopers badge numbers and names are normally provided and drivers can submit complaints to the department. Upon receipt of a complaint, the department assigns the complaint either he Personnel Complaint Investigations or Division Referrals to investigate the complaint. 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. A formal complaint "alleges one or more of either an infraction of Department rules, regulations, or polices, or an illegal act (TxDPS, 2018)." Racial profiling is considered an illegal act under Article 2.132 in the Code of Criminal

Procedures and can be a legitimate reason to file a complaint against the trooper.

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). 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 percent of white motorists being stopped 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, troopers or police officers who participate in misreporting are not checked for accuracy and are less likely to be caught. Usually, only the driver focuses on the content of the ticket. Third, unless the trooper searches the driver and arrests the driver, no other law enforcement officer will see 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.[5] 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 motorists are either white (W) or minority (M) in this section. In my empirical analysis, I allow for troopers and motorists to be W, B (black), 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 $\theta \to 1$ as some motorists may be obviously guilty.

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

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

[5] 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.

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 be able to compare levels of bias across troopers, I need to 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. Troopers incur a cost of μ for misreporting, as it may open the door to greater punishment. I make the following assumptions on μ:

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

As θ, or the culpability of the motorist, increases, the cost of misreporting also rises. Therefore, motorists who appear less guilty are more likely to be misreported. 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

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

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. 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.*[6]

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

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

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.[7]

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,*

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

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

For unbiased troopers, $v = 0$. For biased troopers, $v > 0$.[8] 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 collected over 130 million records from 31 state police agencies (Pierson et al., 2017). The goal of the project is to analyze detailing interactions between police and the public. This information is freely available on the website.

I use the Texas portion of the SOPP data because Texas was the only state where

[7] 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.

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

troopers were caught misreporting. While SOPP provides the data from 2006, Texas troopers were not required to record the driver's last name until 2010, so I cannot test for misreporting behavior prior to 2010. The data contains detailed information on the stop such as latitude and longitude of the stop, make and model of the car, the owner of the car, time and date of the stop, the reason for the stop, whether a search was conducted and why, if contraband was found, whether an arrest was made, first initial and last name of the trooper recording the stop, and the badge number of that trooper. The data set also has limited information on the type of contraband found: currency, weapon, and other. Pierson et al. (2017) courteously provided the raw version of the data, which had the driver's full name and home address. This becomes important when I estimate drivers' true races. I drop Native American and Middle Eastern motorists, which is about 30,000 observations. For reasons I explain in the race estimation section, I also only keep male drivers. Overall, the subset of the data I use contains about 9 million total stops with 3,509 unique troopers.

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. One of the assumptions in my model is that motorists are only guilty through finding contraband. Including searches where the motorist is arrested prior to searching the vehicle violate the assumptions of my model. Because of this, I restrict my definition of search success to only include searches due to probable cause or driver consent.

I also augment the SOPP data with 2016 - 2017 highway stop data from the Texas Department of Public Safety. This data has identical information to the SOPP data, but does not have the driver's full name or addresses in order to protect the privacy of the driver's in the data set. The new data set contains additional information such as whether the driver was a fugitive, the sergeant in charge of the area of the stop, the alleged speed, the judge assigned to the case, and the court date and location. I also drop the female driver's from this data set to maintain consistency with the SOPP data.

Since the stops occurred after the misreporting was revealed in November 2015, I take the driver's races as given. My primary purpose for including the publicly available data is to measure each trooper's change in misreporting behavior after the publication of the article.

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, 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 2,578 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 so I only have 11,819,236 observations. My final number of observations is 7,685,007 after dropping observations after June 2015 for reasons listed in the prior section.

I further the time period of my trooper employment data by adding 2019 trooper employment data, publicly available on the Texas Tribune Salary website. I link the 2019 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 via a FOIA as an objective measure of trooper quality. 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. 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)." Out of the original 1,873 complaints, only 334 had the trooper's badge number in the complaint.

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 et al., 2018). The first method is to use surname analysis, which works well for Hispanic and Asian surnames. I match the driver surnames in my data to the U.S. Census Surnames data set. If the probability of the last name is Hispanic is greater than a certain threshold (75%), I impute the 'estimated' race as Hispanic.[9] For example, Figure 3 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. Since, the probability this driver is Hispanic, conditional on his last name, Mendez, is 92%, I then correct his race to Hispanic. The advantage of this method is that the correction is fairly quick and simple. But, the main drawback is that this method results in errors in the race estimation. For example, even though 92% of persons with the last name Mendez identify as Hispanic, 8% do not. My race estimation would still correct this 8% as Hispanic despite their self-identification of another race. However, even if the race estimate is incorrect, as long as the errors are independent of the search outcome, my results will be unaffected. Another minor drawback is that surname analysis is only suitable for Asian and Hispanic names and is less effective with females since married women tend to change their last names to that of their husband's. Thus, I only keep male drivers in my sample.

The second method I employ is geocoding analysis, which I only use on to uncover black drivers "because at least half of black Americans continue to live in predominantly black neighborhoods (Fiscella and Fremont, 2006; Glaeser and Vigdor, 2001)." I use the recorded address of the driver to geocode to a specific latitude and longitude using

[9]As a robustness check, I raise the threshold to higher levels in later sections.

geocoder.us. I then use that latitude and longitude to map the address to a Block FIPS code using the FCC block finder. I merge this data with the 2010 American Community Survey. If the percentage of Black population in the area is greater than a certain threshold (67%), I correct the race as "black." I use 67% since Fiscella and Fremont (2006) found that with "block groups where more than two-thirds of the residents were black... 89 percent were classified correctly."[10] This method also has a few disadvantages. First, if the trooper did not record the address of the driver (< 7% of the data), I can't geocode it. Second, the address is inputed by the trooper, which is prone to spelling and typing errors. For example, I found 116 different spellings of the city "Houston," which is the largest city in Texas. Third, this method is also computationally expensive so I restrict this analysis to only drivers who live in Texas, which is approximately 90% of the stop data.

For every stop in the data set, I only use one race estimation method. This is to prevent the estimated race from depending on the order of the race estimation. For example, if a driver with a Hispanic last name who lived in a predominantly black block FIPS area was misreported as white, then he would be corrected as black or Hispanic depending on if I used the geocoding or surname analysis first. There are only 427 drivers who were recorded as white with a Hispanic surname living in a predominantly black neighborhood so there's no significant difference if I were to use both methods or changed the order of applying the analysis. I also only correct the races of motorists originally recorded as white or unknown.[11] This allows me to only correct the races of the motorist once. Otherwise, a motorist with the surname Gomez living in a predominantly census block could be corrected as black or Hispanic depending on if I ran the surname analysis first or the geocoding analysis. I will go into further detail my methodology for estimating the race of the motorist in the subsequent section.

[10] I also raise this threshold later as a robustness check

[11] As a robustness check, I also estimate the race for non-white motorists with no change in the results. I also observe only small changes in the number of non-white motorists by race. The number of estimated Hispanic motorists increased by 0.8%, the number of estimated black motorists increased by 0.9%, and the number of estimated Asian motorists increased by 1.8%.

Lastly, I show the how the misreporting affected the observed stop rates of motorists by race by comparing stop rates before and after the DPS rule change. Figure 2 shows the raw, time trend of the stop rate by driver race using the recorded races for Asian, black, Hispanic, and white motorists from 2010 to 2017. The dashed, vertical line indicates the year-month of Sandra Bland's stop. During the time period after her stop and the publication of the article, Texas troopers appeared to significantly increase their misreporting behavior. The second vertical line denotes year-month of the publication of the article and DPS' rule change. After November 2015, the Hispanic stop rate rises to over 40% while the white stop rate falls to nearly equal levels. I observe no noticeable changes for the Asian or black motorist stop rate before and after November 2015. Since only Hispanic motorists have enough observations that are misreported, I focus the rest of my analysis on bias on Hispanic motorists. Since I also cannot discern if the spike in observed the white stop rate and concurrent drop in the observed Hispanic stop rate prior to November 2015 is due to her death or whether there was a shift in behavior, I only use stops from January 2010 to June 2015 for the rest of my analysis.

4.4 Descriptive Statistics

I present summary statistics of motorist characteristics in Table 1 using the estimated races. On average, most motorists stopped are white, but this pattern doesn't carry over to searches. Specifically, I find that conditional on being stopped, Hispanics motorists are searched the most at nearly 40% followed by white motorists at nearly 39%. Black motorists also show a higher search rate compared to stop rate with a difference of 9.5%. The only non-white race of motorist that is stopped at a higher rate than the search rate are Asian motorists. I also find that certain stop characteristics, such as midnight stop, an older car and a luxury brand cars have a higher search rate compared to the stop rate.

Table 2 shows summary statistics of troopers. Of the 2,701 troopers I was able to match to the data, approximately 60% are white, 30% are Hispanic, and almost 9% are black. The last one percent is composed of Asian, American Indian, and other race troopers. The force is predominantly male at 94%. By trooper race, I find that white troopers make up most of the searches at 64%, followed by Hispanic troopers at 21%. 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 less experienced troopers searched more than more experienced troopers since the average hire year for searches was 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, with rank listed in decreasing order.[12] 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 Captain as an example, the interpretation of the probabilities is "If the trooper is a captain, then captains conduct 0% of total searches." I find that troopers make up most of the stops and searches at 70% and 72% respectively.

Trooper characteristics may depend on where they are stationed. For example, troopers may be more likely to be stationed in counties with a higher population of their own race. I measure the correlation of county's trooper composition with county characteristics in Table 14 in the appendix. The county characteristic data is from the 2011-2015 American Community Survey. I also include the average, annual violent crime rate at the county level from the University of Wisconsin Population Health Institute. I find that black troopers tend to be placed in more populated counties with higher percentage of white residents, higher median income, lower rates of health insurance, and lower crime rate. Hispanic troopers tend to stationed in areas with high percentage of Hispanic residents but low percentage of Latin American foreign born. White troopers are more likley to be stationed in areas with high black population, lower educated, and high percentage of Latin American foreign born.

[12] I excluded the rank of major as only two troopers were majors and they conducted no searches and only four stops during 2010 to 2015.

5 Empirical Results

My test for racial prejudice

I first test if troopers in my data are indeed misreporting motorists as white differentially by search outcome. Table 3 shows the results of this test, which compares the probability of misreporting conditional on search outcome. The key identification of my test is that biased troopers should differentially misreport based on search outcome; specifically, the trooper should misreport only when the search ends in failure. Since my race estimation method cannot perfectly estimate driver's race or identify when a trooper is misreporting versus making a mistake, I rely on the differential misreporting behavior across search outcome to measure bias. The key assumption is that any other possible reason for trooper error or driver race mis-identification will occur at equal rates across search failure or search success. Then, if mismatch between the estimated race and the observed race is driven by bias, then the probability of mismatch should be greater when the search ends in failure compared to success.

Columns (1) and (2) of Table 3 show the misreporting rate conditional on search outcome. Indeed, I find that troopers are 40 percent more likely to misreport Asian motorists when the search ends in failure than when the search ends in success. For black motorists, I find that black motorists are 22 percent more likely to be misreported as white when the search ends in failure compared to when the search ends in success, which is a significant difference. For Hispanic motorists, I find that failed searches are 3.6 percent more likely to be misreported compared to successful searches. Since I observe bias with both black and Asian motorists, this also shows that my method of uncovering bias is robust to not only the race estimation technology, but is not unique to just Hispanic motorists, since I detect bias against Asian motorists using the same race estimation technique.

In order to control for confounding variables, such as seasonality and county characteristics, I regress the following equation:

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

where $I(Mismatch_{i,c,t})$ is an indicator variable equal to one if the recorded race did not equal the estimated race for stop i at time t in county c. $I(Failure_{i,c,t})$ is an indicator variable equal to one if the search ends in in failure and equal to zero if the search ends in success. $X_{i,c,t}$ is vector of controls of the stop, such as vehicle characteristics, whether the driver is in-state, and if the owner is the driver. I also include county fixed effects and control for seasonality by including month fixed effects. The coefficient of interest is β_1. If a trooper is biased, then he is more likely to misreport the minority motorist as white when the search ends in failure thus this coefficient will be greater than 0 if troopers are misreporting in a biased manner. I run this regression separately for Asian, black, and Hispanic motorists using the estimated races.

From my results in Table 4, I find that the rate of misreporting was 50 percent more likely to occur when the search ended in failure for Asian motorists, 21 percent more likely for black motorists, and 3.6 percent more likely for Hispanic motorists. These results are robust to county and month fixed effects. The results for Asian and black motorists are robust when including year fixed effects, but the results for Hispanic motorists are not. This is not surprising given the time trend observed in Figure 2, which shows that misreporting was falling steadily during 2010-2015. These results are slightly smaller, but similar to the rates shown in Table 3. This estimation strategy also masks the heterogeneity across troopers because the estimates only reflect the average level of bias. In the next part of my analysis, where I estimate bias on the trooper level, I focus on Hispanic motorists because there are fewer Asian/black stops by trooper. And I do not have a sufficient number of misreported Asian or black stops to control for possible confounders.

To measure the magnitude of Hispanic bias for each officer, I use Equation (6), but allow for each trooper to have his own disproportionate misreporting rate depending on failure. The more biased the trooper is, the more he will misreport motorist of race-r. I rely on the difference because my method of race estimation also corrects successful searches so the differential rate of misreporting based on search outcome will identify bias. For every stop i by trooper$j$ at time t:

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

Each β_1^j will measure each officer's differential misreporting behavior based on search outcome. A positive estimate indicates bias against Hispanics. Since troopers with more searches will have a more precise estimate of bias than troopers with few searches, I exclude troopers with less than 5 searches of Hispanic or white motorists using the observed races. I include controls for county characteristics in $X_{i,c}$, rather than using county fixed effects. These characteristics include median income, percentage Hispanic, percentage black, employment rate, percentage with high school diploma, and population size.[13] I also include month specific fixed effects. I show the distribution of β_1^j in Figure 4.

One notable characteristic of the distribution of bias is the heterogeneity in the measures of bias for troopers. While most troopers are concentrated at no bias, I find a large mass with positive levels of bias, with a mean and a median of 0.175 and 0.071, respectively. This means the average trooper is 17.5% more likely to have the estimated race not match the recorded race when the search ends in failure compared to success. Officers at the 90th percentile of the distribution are 95 percent more likely to misreport with failed searches compared to successful searches while at the bottom 10th percentile are 31% more likely to misreport successful searches than failed searches.

Another characteristic of the distribution is the negative side. Troopers here are more likely to misreport when the search ends in success compared to failure. This can occur for the following reasons. First, some troopers may only misreport when the search ends

[13] I use county fixed effects and have similar results, but less troopers.

in success for reasons that may or may not be related to bias. If that is the case, then few troopers engage in this behavior, as evidenced by the small mass on the left hand side of the distribution (with 934 troopers with bias less than 0 and 1,236 with bias greater than or equal to zero). Second, the race estimation method is an estimate of troopers bias, thus there will be troopers who are negative. Given the smoothness of the negative side compared to the 'hump' on the positive side, troopers with 'negative bias' may just show the natural distribution of bias across Texas troopers.

Next, I use the publication of the news article by KXAN revealing the misreporting as a natural experiment. The article was published in November 8th, 2015, a hearing was conducted by November 18th, and by November 23rd, DPS changed its policies to now require troopers to ask drivers for their race. I can test the effect on stop behavior of troopers after the changes are implemented by augmenting the SOPP data with the publicly available stop data from DPS. Since I observe an inexplicable change in misreporting behavior from July 2015 to November 2015, I will only use stops preceding July 2015 to January 2010 as my pre-data.[14] One possible reason for this change is that prior to the article, some investigation of trooper's misreporting behavior had already occurred. Without any further information, I consider stops after July 2015 as post-rule change data. For my post data, I am using the publicly available data, which has the recorded driver's races but no driver's names or addresses. If my measure of bias is correct, troopers who were using misreporting to hide their bias should have the greatest changes in stop behavior with Hispanic and white motorists. Biased troopers are more likely to misreport the search if the search was unsuccessful. Thus, I should observe a negative relationship between search success rates and motorists being recorded as white relative to the trooper's bias prior to November 2015 and no effect after November 2015.

[14] I so far have been unable to uncover any reasons for this change in troopers' misreporting behavior. The only significant event I can find is Sandra Bland's death in jail after she was pulled over for failing to signal a lane change, which occurs in July 2015.

For ever stop i by trooper j in county c at time t:

$$
\begin{aligned}
I(RecRace = White_{i,c,j,t}) = & \alpha + \sum_{t=2010}^{2017.5} \Big[\beta_1^t Hisp\ Bias_j \times I(Year\ Half = t) + \\
& \beta_2^t I(Failure_{i,c,j,t}) \times I(Year\ Half = t) \\
& + \beta_3^t Hisp\ Bias_j \times I(Year\ Half = t) \times I(Failure_{i,c,j,t}) \Big] + \\
& \boldsymbol{X}_{i,c,t}\gamma + \theta_{c,t} + \phi_c + \alpha_m + \epsilon_{i,c,j,t}
\end{aligned}
\tag{8}
$$

I use the *recorded* race rather than the estimated race because the recorded races will show the greatest change after June 2015 if my measure, $Hisp\ Bias_j$, actually quantifies officer j's bias, where $Hisp\ Bias_j$ is derived from Equation (7) using stops from January 2010 to June 2015. $I(Failure)$ is an indicator variable if stop i ends in failure. I also control for county, year, and month fixed effects along with county specific time trends. The primary coefficient of interest is β_3^t for $t > 2015$, which is the interaction between officer j's estimate of bias measured using stops before 2015, the search outcome, and the years after the changes were implemented. β_3^t will reflect the differential probability in being recorded as white when the search ends in failure for biased officers. If my measure is capturing bias, then officers with higher levels of bias should change their search behavior more than officers with low level of bias.

I use a different specification than with Equation (6) since I cannot test for misreporting using the publicly available data because the publicly available data doesn't contain driver's names or driver's home addresses. But, even if I cannot measure misreporting directly, DPS changes its policies to require troopers to verbally ask for driver's race, thus the recorded race should be the estimated race after November 2015. Thus, I can assume that stops after November 2015 are recorded correctly.

Figure 5 shows the results for both Hispanic and white motorists. If my measure bias captures each trooper's true measure of bias, then the coefficient for β_5^t for $t < 2015$ should be positive and should go to zero for $t > 2015$ for motorists recorded as white. The intuition being that biased officers will disproportionately misreport Hispanic motorists as white depending on the search outcome. And therefore are more likely to record the motorist as white when the search ends in failure compared to unbiased officers prior to the publication of the article. Specifically, a standard deviation away from the mean level of bias leads to a 8 percentage point increase in the probability of being recorded white when the search ends in failure from 2010 to 2014. But, from 2016 to 2017, the coefficients are not significantly different from zero. For Hispanic motorists, I find the opposite pattern. Motorists are 9 percentage points less likely to be recorded as Hispanic if the search ended in failure. But, by 2015, there is no significant difference in the probability of being recorded as Hispanic depending on the search outcome.

This reveals two important characteristics of trooper behavior. The first is that the biased troopers complied to the rule change and began correctly reporting motorists race after 2015. While DPS was not clear as to how they would enforce their new policy, it appears effective in changing biased troopers' behavior. The second is that motorists who self-identified as Hispanic were the ones predominantly being misreported. This is especially important since Hispanic is technically an ethnicity and many driver's may self-identify as white rather than as Hispanic despite having a Hispanic last name (Lopez et al., 2017). If, on the other hand, troopers appeared to misreport Hispanic motorists because the motorists self-identified as white while the data estimated the race to be Hispanic, I would observe no change in behavior as a result of the range. Thus, troopers would not have been misreporting in biased manner.

Past tests for racial prejudice and misreporting

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). Under their test, if troopers were biased against Hispanic motorists, then the search success rate for Hispanic motorists would be lower than the search success rate for white motorists. In Table 5, 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 30.7% compared to the white, Asian, and black search success rate, at approximately 42.0%. Therefore, Knowles et al. (2001) test would conclude that troopers are most biased against Hispanic motorists. When I use the *estimated* races, I find different search success rates, which are shown in Column (2). Using the estimated races and applying KPT's test again, I instead find that Asian, black, and Hispanic motorists are biased against and that the magnitude of bias against Hispanics is actually much larger. This is driven by the white search success rate risking by 6%. The race estimation also slightly changes the Hispanic search success rate, which falls significantly by 1%. I find that the race estimation does not significantly change the black search success rate. The difference in effectiveness of each race estimation is driven by the different methods used to uncover motorist race, but also because Texas troopers were predominantly misreporting Hispanic motorists (Collister, 2015a).

This also shows that misreporting reduces the appearance of racial bias for troopers by reducing the search success rate of white motorists. While I find no large increases in the search success rate after race estimation for black and Hispanic motorist search success rate, the white search success rate has risen significantly by 6% to nearly 50%. Thus the difference in search success rates across motorist race decreases as a result of the misreporting.

Robustness Checks

To ensure that the relationship between my measure of bias is not dependent on my census surname cutoff. I vary the threshold I use in the surname analysis at 50%, 75% (the measure I use throughout my analysis), and 85%, and re-estimate my trooper level measure of bias.[15] Figure 6 shows the distribution for each cutoff. These distributions show a similar shape; it also shows that even with tighter thresholds, the distribution of

[15] I also use 95% as a threshold, but at 95%, there are only 1,242 Hispanic surnames compared to the 4,647 surnames at 85%. Most of the bias estimates were concentrated at no bias, making the other densities hard to see on the graph.

officer level bias is robust.

Since my main estimation of Hispanic motorists relies on the distribution of Hispanic surnames in the United States from the 2000 census, I test if my results are driven by the unequal distribution of last names across search outcome. I randomly assign the probability a last name is Hispanic based off the normal distribution. Using different Z-score cutoffs, I re-estimate bias using Equation (6). I repeat this procedure a thousand times to get the distribution of average bias in Figure 7. The average level of bias is -0.04, which is far less than my estimate in Table 4. In fact, that natural distribution implies troopers on average are *less* likely to misreport unsuccessful searches by 4 percentage points. Thus, I can reject the null hypothesis that my results are merely driven by the distribution of last names.

I also test if biased troopers were able to effectively evade detection of bias using Knowles et al. (2001) and Becker (1957)'s test of bias. I regress:

$$\begin{aligned} Y_{icjt} =& \alpha + \beta_1 Hisp\ Bias_j + DriverRace_i \boldsymbol{\beta_2} + DriverRace_i \times Hisp\ Bias_j \boldsymbol{\beta_3} \\ &+ m + c + t + c * t + \epsilon_{icjt} \end{aligned} \tag{9}$$

separately using the recorded races and the estimated races. The outcome of interest is whether a search was conducted and if the search was successful conditional on search. I control for officer level Hispanic bias along with the recorded (estimated) race of the driver. The variable of interest is $\boldsymbol{\beta_3}$, which shows the differential probability of search or success for each driver race compared to white motorists.

Table 6 shows the results using the recorded races in Panel A and the estimated races in Panel B. From Panel A Column (1), biased troopers appeared less likely to search black and Hispanic motorists by .2 to .3 percentage points compared to white motorists. Biased troopers also appeared to have significantly higher search success rates, specifically Hispanic motorists, with one standard deviation of bias being associated with a 12 percentage point higher search success rate than white motorists. I find similar positive effects for black motorists with 3 percentage points higher search success rate associated with one standard deviation in bias.

In contrast, Panel B, shows different results using the estimated races. While the estimated races are not the true race of the motorists, these results give some insight in the true search behavior of troopers by driver's race. My results show no significant difference in the probability search by driver race compared to the results using the recorded races. In contrast, I find different results using the estimated races on the probability of search success. I find a negative, but not significant, coefficient for the effect of Hispanic bias on the Hispanic motorist search success rate where one standard deviation is associated with a one percentage point decrease in the probability of search success for Hispanic motorists compared to white motorists. This coefficient is significantly different than the results in Panel A Column (2). The results from the estimated races and recorded races show that biased officers appeared to have a much higher search success rate when searching minority motorists compared to unbiased officers, but in actuality, they were equally good if not worse than unbiased officers.

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 test whether bias varies by trooper demographics. Additionally I will address how discrimination varies with other employment characteristics such as promotions, salary, and officer transfers.

Table 7 shows Column (3) of Table 3 broken down by trooper race. Using a model similar to Equation (6), I run the regression:

$$\begin{aligned} I(Mismatch_{i,j,t}) = \alpha + \beta_1 I(Failure_{i,j,t}) + I(TroopRace_j)\beta_{2j} + \\ I(Failure_{i,j,t}) \times I(TroopRace_j)\beta_{3j} + X_{i,t} + c + m + \epsilon_{i,j,t} \end{aligned} \tag{10}$$

where I now control for the race of trooper j. The coefficient of interest is β_{3j}, which is the differential misreporting behavior by search outcome for black and Hispanic troopers. To control for the county characteristics and seasonality, I also include county and month fixed effects. If white troopers are more biased against Hispanic motorists, then I $\beta_{3,black}$ and $\beta_{3,Hisp}$ should be negative since I omit white troopers.

I report my results in Table 7. I find that black troopers are 2.5 percentage points less likely to misreport Hispanic motorists compared to white troopers, but this coefficient is not significant. This may be a power problem since the only 236 troopers are black ($< 8.5\%$ of troopers). Surprisingly, Hispanic troopers are equally biased against Hispanic motorists compared to white troopers, which is contrary to past tests of bias (Goncalves and Mello, 2017; Antonovics and Knight, 2009).

Next, I test if employment outcomes, such as salary and experience, are related to bias where experience is measured using the hire year of the trooper. I omit troopers with negative bias since I am testing the relationship between having Hispanic bias compared to having no bias. I can test how bias is correlated with labor outcomes. For example, if biased troopers have a higher rank than unbiased troopers.

I show my results in Table 8. I find positive correlations of trooper salary and experience to Hispanic bias, but these results are not significant. For trooper rank, I find that troopers with rank Sergeant or higher have Hispanic bias 0.2 standard deviations higher than troopers. No other rank has significant different levels of bias compared to troopers.

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, I do not have trooper employment data prior to 2013 so 2013 is the earliest year I can use. Second, the measure of bias is noisy since it's measured on the differential misreporting behavior across search outcomes. Therefore, officers with few searches have a noisier estimate of bias. By dividing the trooper's career into two sections rather than by year, my estimate of bias is more efficient and more consistent. Lastly, 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. Moving cities is a proxy for salary. Rather than a salary increase, a trooper can be compensated for good behavior by being stationed at a preferred city. I omit troopers with negative bias again for the same reasons as before.

I show my results in Table 9. I regress the likelihood of leaving the force, moving cities, and increasing rank on the standardized measure of bias from the first half of the trooper's career including controls for the trooper's rank before 2014 and for their work city. I find no relationship between an officer's measure of Hispanic bias prior to 2014 on any of the labor outcomes from 2014 to 2015. Not only are the point estimates insignificant with large standard errors, but the estimates are also close to zero. I interpret this as evidence that misreporting allows biased officers to evade punishment for bias since 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.

I also decompose Column (3) from Table 9 into each rank. I regress the probability of increasing rank for each rank of trooper interacted with the level of bias measured from stops conducted 2010 - 2013. From my results in Table 10, I find no evidence that more biased troopers are more or less likely to be promoted regardless of rank.

I next test if biased troopers also perform worst in other aspects of their job by using complaint data obtained from DPS. The results in Table 11 show a positive relationship between trooper level bias and the probability of receiving a complaint. One standard deviation of bias is associated with a one percentage point higher likelihood in having a complaint filed against the trooper. This is possibly an underestimate of the actual association of bias and complaints since complaints where the trooper was exonerated or the complaint was considered unfounded after investigation repressed the trooper's badge number. From the 1,873 complaints, only 334 included the trooper's badge number.

I also test to see how the employment outcomes of troopers were affected by the publication of the article. I use publicly available 2019 salary data published by the Texas Tribune. My results in Table 12 bias has no effect on the probability of leaving the force by 2019. But, I do find significant effects in salary growth and on the probability of ranking up. Specifically, each standard deviation of bias decreases the likelihood of increasing in rank by 27 percent and decreases salary growth by 8 percent. Bias only affects labor outcomes once the misreporting is outlawed, which I interpret as evidence of misreporting being an effective way of evading punishment for bias.

Bias and the Work Environment

Another important question is whether a trooper's work environment plays a role in their level of bias. Many characteristics of the environment can affect own bias, such as social norms, demographic and labor characteristics of the environment, or characteristics of peers. Additionally, in the context of Texas, troopers may have different incentives (for example: reducing drug trafficking, reducing undocumented immigration) to search motorists depending on the needs of the area they patrol. Understanding how characteristics of the environment can affect bias can inform on policy prescriptions to reduce bias within troopers.

I focus my analysis on two observable measures of the trooper's patrolling environment, peer characteristics and environment characteristics using county as the relevant geographic area. This makes sense as troopers are assigned to counties to patrol in within their region. From Figure 1, regions occupy a huge amount of area, thus within regions, troopers only patrol within certain counties, rather than across the entire region.

To construct measures of peer characteristics within a county, I restrict my sample of troopers to those who have conducted 65% of their searches within the same county. This excludes troopers who patrol multiple counties who may not be present enough within the county to be considered a peer. Using these troopers, I construct estimates of the demographic and labor composition of troopers working within the county. Since counties with only a few troopers may not have any peers, I also restrict the counties in my sample to counties with at least 4 troopers who have conducted 65% of their searches within that county. In total, I have 1,440 troopers across 91 counties. These 1,440 troopers have also conducted at least 5 searches.

To determine if the composition of peers within the county has any effects on troopers' own bias, I study troopers who have conducted at least 5 searches in two different counties. These troopers were excluded from the measurement of peer composition in Table 14. I create a measure of Hispanic bias for each trooper-county pair. I have 236 troopers with enough observations across multiple counties to measure bias for each trooper on the county level. Of these 236 troopers, 14 are black, 70 are Hispanic, and 157 are white. I regress trooper-county bias on peer composition variables with trooper fixed effects.

My results in Figure 8 show differential effects of peer composition on own bias by trooper race. Overall, I find that county characteristics have no effect on a trooper's own bias, even when run separately for race. Interestingly, the estimate of the effect of peer demographics on white troopers' bias is close to zero, implying that white troopers' bias is impervious to changes in peer demographics.

I also test if county characteristics have a effect on own bias. For example, whether patrolling in a county with higher crime increase bias. Figure 9 shows the estimates of effects of county characteristics on own bias. Overall, I find that county characteristics have no effect on a trooper's own bias, even when run separately for race. This provides suggestive evidence that the county characteristics of the patrol environment does not affect troopers' bias.

Lastly, I test whether peer bias affects own bias. I use the average level of bias of the same troopers from the peer composition measures within the county as the measure of county-level bias. My results in Figure 10 shows the correlation of trooper-county bias to county level bias without trooper fixed in effects in the first graph and with trooper fixed effects in the second graph. The results with no trooper fixed effects shows a positive correlation between county bias and Hispanic troopers' own bias. But, these results are not significant after including trooper fixed effects. The inclusion of trooper fixed effects allows for a more direct interpretation of the effect of county bias on own bias.[16] My

[16] I cannot identify the causal effect of peer bias on own bias because of Mansky's reflection problem; Despite this, the direction of the estimates are informative.

results show county bias has no effect on own bias since the estimates are small and insignificant even when separating the results by race.

The combination of all of these results suggests that peer composition and county characteristics do not significantly affect trooper's own bias. Notably, I find that white troopers' bias is impervious to changes in characteristics in the environment, including peer composition, peer bias, and county characteristics. While, Hispanic and black troopers show varying levels of sensitivity to the composition of peers, these results are not significant.

6 Conclusion

Recent news on the disparate treatment of motorists by race in the criminal justice system have increased claims of systemic racial bias as the cause. I show in this paper that this scrutiny may induce law enforcement officers to purposefully misreport their interactions with citizens to appear less biased. Further, I use this misreporting as an absolute measure of an individual's racial bias to understand some causes and effects of bias on labor outcomes. I study highway search and the decision of highway troopers to misreport a portion of their failed minority searches.

By calculating the increased probability of reporting a motorist as white when the search ends in failure, I create a distribution of trooper level measure of bias. I find that 60% of troopers were engaging in this behavior, with the average trooper being 17% more likely to misreport his failed Hispanic searches white compared to his successful Hispanic searches.

I also test whether bias can predict labor outcomes for troopers before and after a rule change in motorists' race recording policy in response to the misreporting. I find that prior to the rule change, troopers' racial bias did not affect their labor outcomes. I also find that biased troopers were objectively worse troopers with one standard deviation of bias correlated to an 18% higher likelihood of receiving a complaint. But after the rule change, one standard deviation in bias led to a 27% decrease in the likelihood of increasing in rank and salary growth also fell by 8%. Thus, I conclude that misreporting was an effective way of evading detection and punishment for racial bias. I also find that this rule change was highly effective in eliminating misreporting.

Taking advantage of the large number of counties in Texas and the heterogeneity in trooper placement across counties, I also measure how well trooper composition and county characteristics can predict trooper's own bias. I find that the level of prediction depends on the race of the trooper. Specifically, Hispanic and black troopers' own bias was sensitive to the proportion of white peers, while white troopers' bias did not significantly change in response to changes in peer composition by race.

7 Appendix

References

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

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

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*.

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

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

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

Becker, G. (1957). *The economics of discrimination.* University of Chicago Press.

Benen, S. (2014, July). Rick Perry's 'Operation Strong Safety. *MSNBC*.

Collister, B. (2015a). DPS troopers getting race right after KXAN investigation. *KXAN*.

Collister, B. (2015b). Texas troopers ticketing Hispanic drivers as white. *KXAN*.

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.

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

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

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

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

Glaeser, E. L. and J. L. Vigdor (2001). Racial segregation in the Census 2000: Promising news. *Brookings Institute Center on Urban & Metropolitan Policy*.

Goncalves, F. and S. Mello (2017). A few bad apples? Racial bias in policing. Working Paper 608, IRS Working Papers, http://arks.princeton.edu/ark:/88435/dsp01z890rw746.

Harris, D. A. (1999). Driving while black: racial profiling on our nation's highways. *American Civil Liberties Union Special Report*.

Horrace, W. C. and S. M. Rohlin (2016). How dark is dark? Bright lights, big city, racial profiling. *The Review of Economics and Statistics*.

HSIN. HSIN supports border security operations during Operation Strong Safety.

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

Kalinowski, J., S. L. Ross, and M. B. Ross (2017). Endogenous driving behavior in veil of darkness tests for racial profiling. Working Papers 2017-017, Human Capital and Economic Opportunity Working Group, https://ideas.repec.org/p/hka/wpaper/2017-017.html.

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

Lopez, M. H., A. Gonzalez-Barrera, and G. López (2017). Hispanic identity fades across generations as immigrant connections fall away. *Pew Research Center*.

Makarechi, K. (2016). What the data really says about police and racial bias. *Vanity Fair*.

Oyeniyi, D. (2015). State troopers will now just ask drivers their race. *TexasMonthly*.

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

Pierson, E., C. Simoiu, J. Overgoor, S. Corbett-Davies, V. Ramachandran, C. Phillips, and S. Goel (2017). A large scale analysis of racial disparities in police stops across the United States.

Roland G. Fryer, J. (forthcoming). An empirical analysis of racial differences in police use of force. Working Paper 22399, NBER, https://www.nber.org/papers/w22399.

Sanchez, R. (2015). Who was Sandra Bland? *CNN*.

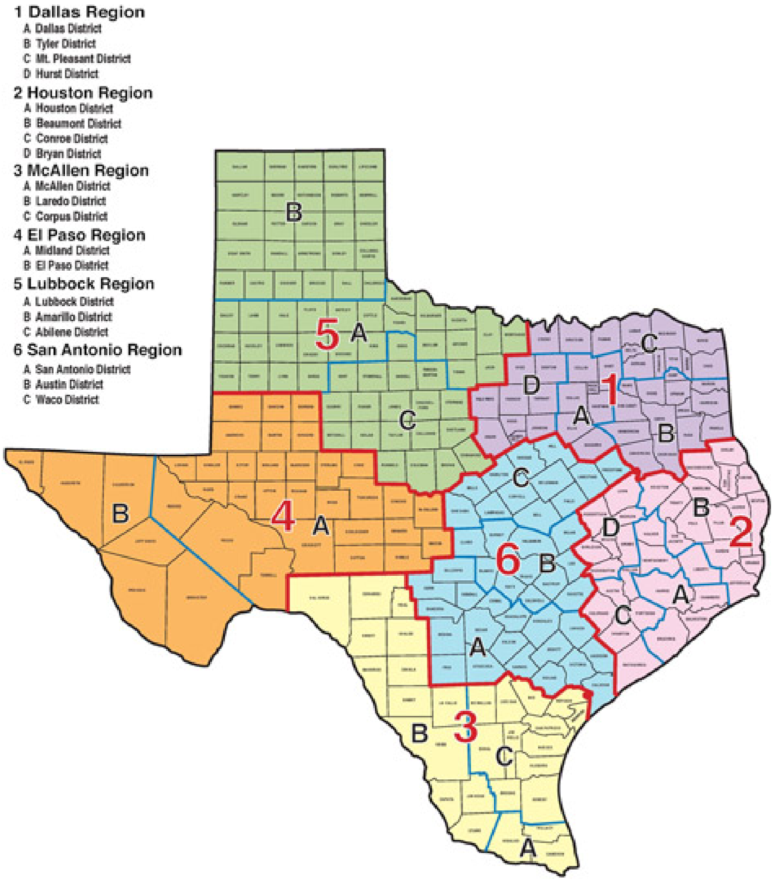
Shayo, M. and A. Zussman (2011). Judicial ingroup bias in the shadow of terrorism. *Quarterly Journal of Economics*.

TxDPS (2018). Complaint investigation and resolution.

West, J. (2018). Racial bias in police investigations.

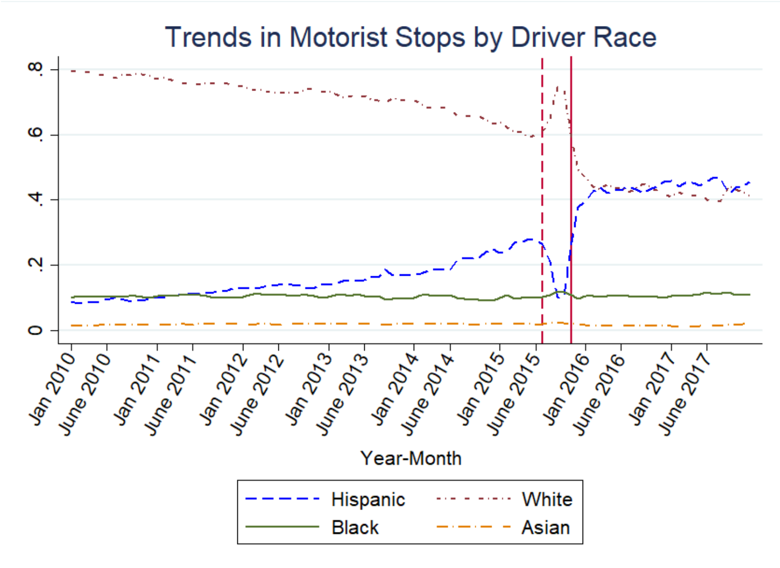
Figures

Figure 1: Trooper Division Map



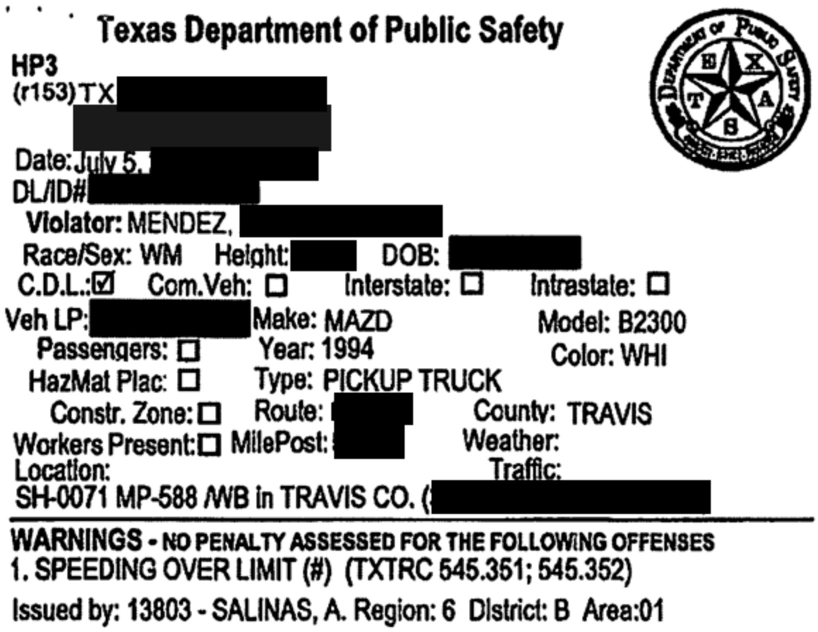
Notes: Source: Texas Department of Public Safety. The 7th region is not shown on the map, but its jurisidiction is limited to only Austin, TX.

Figure 2: Monthly Stop Rate by Driver's Race using Recorded Races



Notes: Dot-dash line shows the recorded Asian stop rate, solid line shows the recorded black motorist stop rate; dashed line shows the Hispanic motorist stop rate using the recorded races, and the dotted line shows the stop rate for white motorists using the recorded races. Average, unweighted stop rates for a given month-year from January 2010 to December 2017 are shown. The vertical red line indicates the year-month the article was released. The dashed red line indicates the year-month Sandra Bland died after a trooper stop.

Figure 3: Example of misreported Highway Ticket



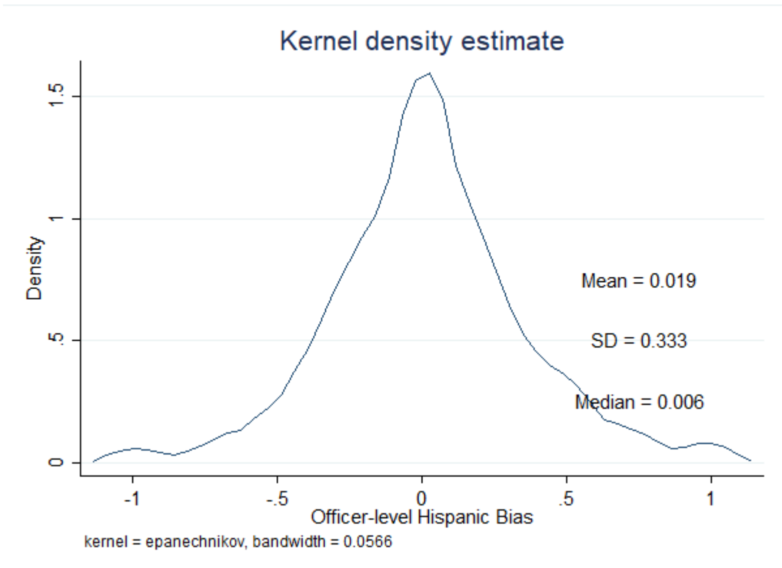
Texas Department of Public Safety

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]

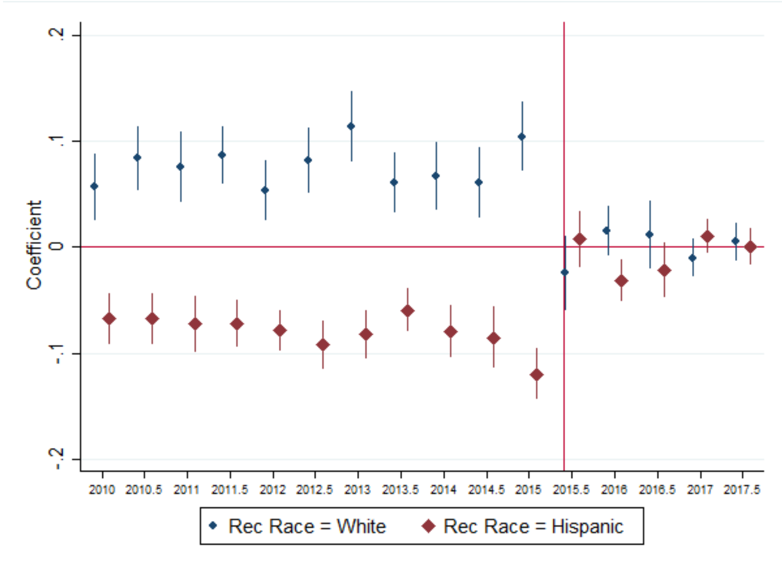
WARNINGS - NO PENALTY ASSESSED FOR THE FOLLOWING OFFENSES
1. SPEEDING OVER LIMIT (#) (TXTRC 545.351; 545.352)
Issued by: 13803 - SALINAS, A. Region: 6 District: B Area:01

Figure 4: Distribution of officer level measure of Hispanic Bias



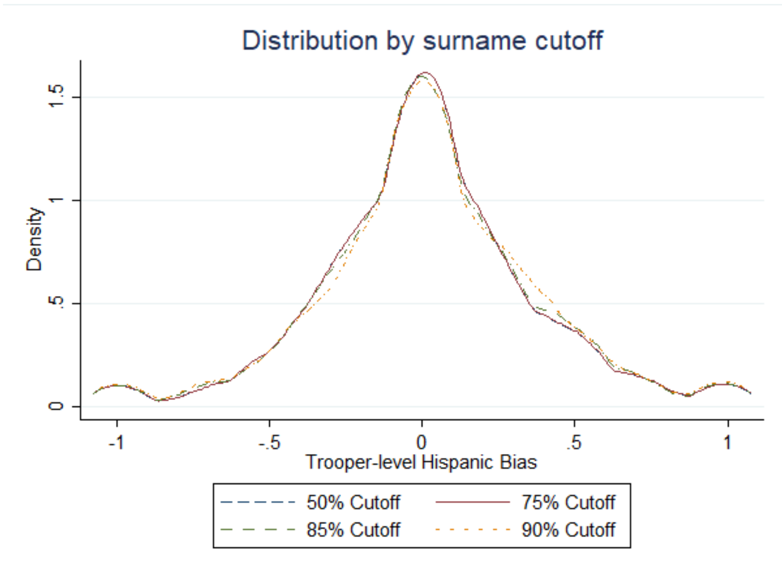
Notes: Kernel density distribution of officer-level Hispanic bias. The figure plots each officer's β^j from the regression $I(Mismatch_{i,t}) = \alpha + \beta^j I(Failure)_{i,t} + \delta_j + X_{i,c,t}\gamma + \epsilon_{j,t}$. Mean reports the average β_1^j and Avg S.E. reports the average standard error for each β^j.

Figure 5: Natural Experiment - White and Hispanic Motorists



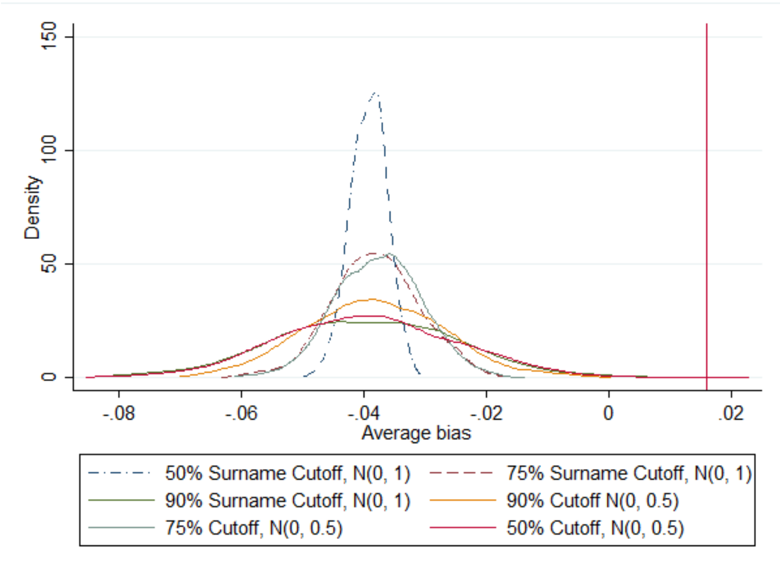
Notes: Figure plots the coefficient of interaction $I(Failure_i) \times Hisp\ Bias_j \times I(Year\ Half = t)$, β_5^t, and with 5% confidence intervals. $Hisp\ Bias_j$ is standardized. Points to the right of the vertical line are after the article publication. Diamond points are when the dependent variable is $I(RecRace = White)$ and circle points are when the dependent variable is $I(RecRace = Hispanic)$.

Figure 6: KPT and misreporting Measure of Bias with different thresholds



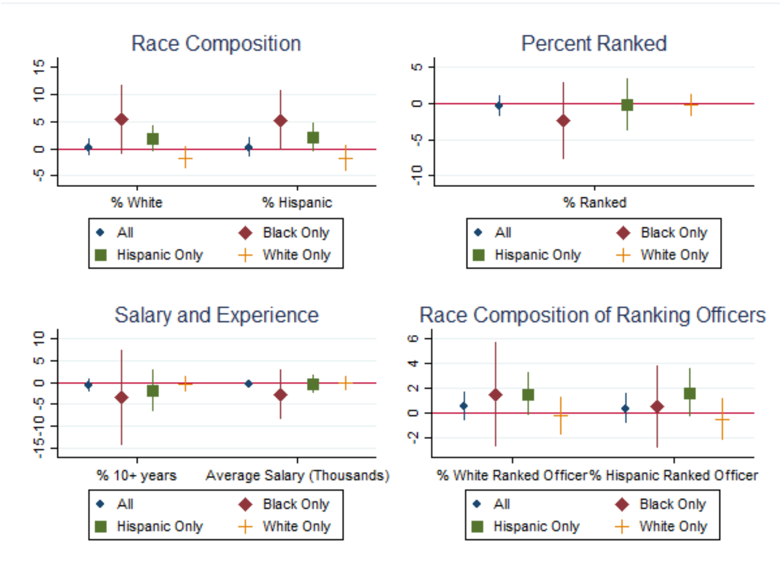
Notes: Each density shows the officer level bias using different levels of surname cutoff and weighted by the total searches.The estimate of bias is from each officer's β^j from the regression $I(Mismatch_{i,t}) = \alpha + \beta^j I(Failure)_{i,t} + \delta_j + X_{i,c,t}\gamma + \epsilon_{j,t}$.

Figure 7: Monte Carlo Simulation of Bias



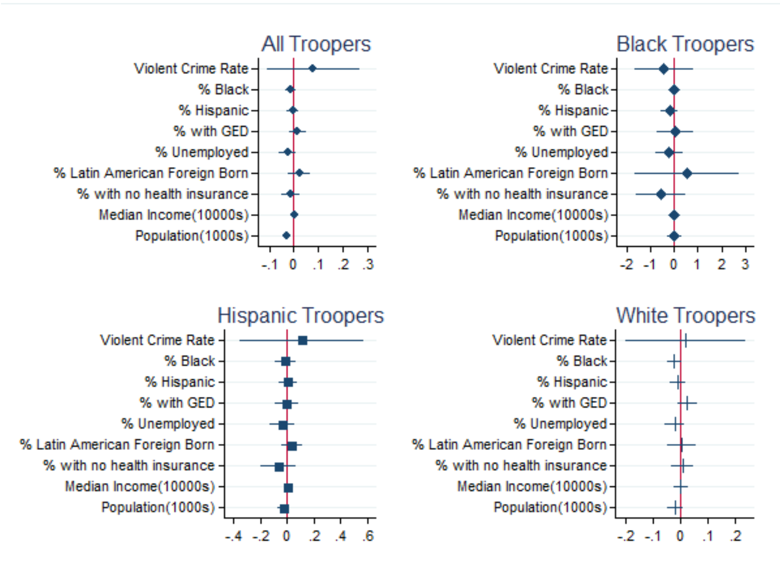
Notes: The graph plots the density of estimates of β_1 from the regression $I(Mismatch_{i,t} = \alpha + \beta_1 I(Failure_{it}) + X_i\gamma + c + t + c*T + m + \epsilon_{i,t}$ where rates of mismatch are simulated using different normal distributions. Horizontal intercept at 0.16 is the bias estimate from the regression using the actual data.

Figure 8: Effects of Peer Composition on Own Bias



Notes: Each graph shows the coefficient estimate with 5% confidence intervals for the variable in the x-axis. Regression is run separately for all troopers and by race with robust standard errors. Troopers who had conducted 75% of searches within the county were used to construct the independent variables. Dependent variable is the trooper-county level of bias for troopers who had conducted at least 5 searches within the county and were excluded from the construction of the independent variables (less than 75% of searches within the county). All regressions include trooper fixed effects.

Figure 9: Effects of County Characteristics on Own Bias



Notes: Each graph is a separate regression and shows the coefficient estimate with 5% confidence intervals for the variable in the y-axis with robust standard errors. Each regression includes trooper fixed effects. Measurements of county characteristics are from the 2011-2015 5-year American Community Survey on the county level. Estimates of violent crime rate are from the University of Wisconsin Population Health Institute. Dependent variable is the trooper-county level of bias for troopers who had conducted at least 5 searches within the county and were excluded from the construction of the independent variables (less than 75% of searches within the county).

Figure 10: Effects of Peer Bias on Own Bias

Notes: Each graph shows the coefficient estimate with 5% confidence intervals for the variable in the x-axis. Regression is run separately for all troopers and by race with robust standard errors. Average county bias is the unweighted, average bias within county of troopers who have conducted at least 75% of searches within that county. Dependent variable is the trooper-county bias of troopers who had conducted at least 5 searches within the county and were excluded from the construction of the independent variables (less than 75% of searches within the county).

Tables

Table 1: Mean of Variables Related to Drivers

Driver Characteristics	(1) All	(2) Searches Only	(1)-(2) Δ
Estimated Asian	.022	.011	.011
	(.147)	(.103)	(0)
Estimated Black	.104	.199	-.095
	(.306)	(.4)	(.001)
Estimated Hispanic	.341	.399	-.058
	(.474)	(.49)	(.001)
Estimated White	.53	.389	.141
	(.499)	(.488)	(.001)
Midnight	.086	.131	-.045
	(.281)	(.337)	(.001)
Owner Driver	.202	.142	.06
	(.401)	(.349)	(.001)
Texas Driver	.9	.848	.052
	(.299)	(.359)	(.001)
Old Car	.304	.426	-.122
	(.46)	(.494)	(.001)
New Car	.338	.179	.159
	(.473)	(.384)	(.001)
Luxury Car	.085	.102	-.017
	(.279)	(.302)	(.001)
Observations	8045487	145730	

Standard deviations are in parentheses. Unweighted means are shown. Stops made from July 2015 to December 2015 were omitted. Midnight is defined as a stop from 12 am to 6 am. Owner information was missing for 38% of the stops. Vehicle was coded as old if made 10 or more years before the stop year and coded as new if made within 3 years of the stop year. Vehicle was considered luxury car if on the Forbes list of top 20 best selling luxury vehicles in 2010. Only 54.5% of the stops were geocoded to a Texas block FIPS.

Table 2: Mean of Variables Related to Troopers

Troopers' Characteristics	(1) All Stops	(2) Searches Only
Black	.087	.04
	(.282)	(.207)
Hispanic	.287	.205
	(.453)	(.404)
White	.606	.633
	(.489)	(.482)
Male	.946	.979
	(.226)	(.142)
Hire Year	2004	2006
	(7.244)	(4.676)
Trooper Rank		
Captain	.007	0
	(.084)	(.018)
Lieutenant	.023	.004
	(.15)	(.059)
Sergeant	.125	.063
	(.33)	(.242
Corporal	.1	.104
	(.3)	(.305)
Trooper	.697	.723
	(.46)	(.447)
Probationary Trooper	.018	.004
	(.133)	(.063)
No Rank	.031	.102
	(.173)	(.303)
Total Troopers	2,701	

Notes: Only merged observations are shown. Trooper rank uses the highest rank the trooper obtained during 2010 - 2015. Stops from July 2015 to December 2015 were omitted. Stops are considered mismatched if the recorded race does not equal the corrected race. 10.5% of the troopers in the employment data were not matched to the stop data. 22% of the troopers in the stop data were not matched to the trooper employment data.

Table 3: Difference in Misrecording Rate by Search Success

Estimated Driver Race	(1) $Pr(Mismatch \mid Failure)$	(2) $Pr(Mismatch \mid Success)$	(3) Δ
Asian	.192	.113	.079
	(.394)	(.316)	(.018)
Black	.018	.014	.004
	(.133)	(.117)	(.001)
Hispanic	.578	.557	.021
	(.494)	(.497)	(.004)

Notes: Unweighted means are shown stops from January 2010 to June 2015. Standard deviations are in parantheses. Mismatch is defined as 1 if the recorded race does not equal the estimated race. Search is defined as success if the trooper found contraband (drugs, weapons, high amounts of currency, and drug paraphernalia)

Table 4: Main Test of bias

	(1)	(2)	(3)	(4)	(5)	(6)
	Asian Motorists		Black Motorists		Hispanic Motorists	
I(Failure)	0.06^{**}	0.05^{**}	0.00^{*}	0.003^{*}	0.02^{**}	-0.01
	(0.02)	(0.02)	(0.00)	(0.00)	(0.008)	(0.01)
Constant	0.13^{***}	0.128^{***}	0.01^{***}	0.02^{***}	0.56^{***}	0.58^{***}
	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
County, Month FE	X	X	X	X	X	X
Year FE		X		X		X
Observations	1520	1520	29031	29031	58098	58098
F	6.679	5.965	3.699	3.420	5.850	1.692

Notes: Dependent variable is an indicator variable equal to one if the recorded race of the motorist in stop i is does not equal the estimated race. Column (1) and (3) have only county fixed effects, Column 2 and 4 have year and county FE. The F-test reports the joint hypothesis test that variables $I(Failure)$ through the fixed effects are equal to zero. Standard errors are clustered at the county level. Regression uses data from January 2010 - June 2015 * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 5: Search Success Rates across Driver's Race

	Search Success Rate		
	(1)	(2)	(3)
	Recorded	Estimated	Δ
Driver Race			
Asian	.426	0.403	0.023
	(.495)	(.491)	(0.018)
	1287	1530	-243
Black	.422	.421	.001
	(.494)	(.494)	(.004)
	27999	28401	-318
Hispanic	.307	.297	.01
	(.461)	(.457)	(.004)
	23868	56530	-32662
White	.421	.482	-.061
	(.494)	(.5)	(.003)
	82451	55631	26736

Notes: Unweighted means are shown. Standard deviations are in the parentheses. Columns (1)-(3) use data from January 2010 - June 2015. Row 3, 6, 9 show the total number of searches using the recorded and estimated races respectively.

Table 6: My measure of Bias and the Becker test of bias

	(A) Recorded Races		(B) Estimated Races	
	(1) I(Vehicle Searched)	(2) I(Success)	(1) I(Vehicle Searched)	(2) I(Success)
Recorded Asian	-0.008^{***}	-0.015		
	(0.001)	(0.017)		
Recorded Black	0.021^{***}	-0.008		
	(0.001)	(0.008)		
Recorded Hispanic	0.009^{***}	-0.022^{***}		
	(0.001)	(0.007)		
Recorded Asian x Hispanic Bias	0.001^{***}	0.039		
	(0.000)	(0.025)		
Recorded Black x Hispanic Bias	-0.003^{***}	0.033^{***}		
	(0.001)	(0.008)		
Recorded Hispanic x Hispanic Bias	-0.002^{***}	0.119^{***}		
	(0.000)	(0.012)		
Hispanic Bias	-0.003^{***}	-0.036^{***}	-0.002^{***}	-0.008
	(0.000)	(0.007)	(0.000)	(0.007)
Estimated Asian			-0.005^{***}	-0.078^{***}
			(0.001)	(0.016)
Estimated Black			0.023^{***}	-0.049^{***}
			(0.001)	(0.007)
Estimated Hispanic			0.012^{***}	-0.110^{***}
			(0.001)	(0.008)
Estimated Asian x Hispanic Bias			0.001^{**}	0.007
			(0.000)	(0.021)
Estimated Black x Hispanic Bias			-0.003^{***}	0.006
			(0.001)	(0.008)
Estimated Hispanic x Hispanic Bias			-0.002^{***}	-0.010
			(0.000)	(0.007)
Constant	0.018^{***}	0.401^{***}	0.015^{***}	0.449^{***}
	(0.000)	(0.002)	(0.000)	(0.004)
Observations	6570205	143306	6570205	143306

Notes: Hispanic Bias is the normalized measure of Hispanic bias for each trooper. The measure for Hispanic bias comes from Equation (7). The regression includes county FE, month FE. Standard errors are clustered at the county level. Regression uses data from January 2010 - June 2015. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 7: Hispanic Bias and Trooper Race

	(1)
	$I(Mismatch)$
I(Failure)	0.0141*
	(0.0082)
Failure X Black Troopers	−0.0257
	(0.0607)
Failure X Hispanic Troopers	0.0199
	(0.0156)
Observations	51059

Notes: Dependent variable is an indicator variable equal to 1 if the recorded race does not equal to the estimated race. Standard errors are clustered at the county level. Regression uses data from January 2010 - June 2015. HispanicTroopers and Black Troopers are indicator variables for the trooper's race with white troopers being the omitted category. Each regression is run separately for motorists of each race, where race is identified using the estimated race. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 8: Correlates of Hispanic Bias

	(1)	(2)	(3)	(4)	(5)
			Hispanic Bias		
Experience	0.02**	0.02**			
	(0.01)	(0.01)			
Black	−0.01	−0.02	0.00	−0.01	0.02
	(0.18)	(0.17)	(0.18)	(0.18)	(0.18)
Hispanic	0.13	0.15	0.14	0.15	0.17
	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)
Probationary Trooper		0.31		0.40	0.20
		(0.30)		(0.32)	(0.30)
Corporal		−0.29**		−0.28**	−0.23*
		(0.14)		(0.14)	(0.14)
Sergeant+		0.15		0.07	0.23*
		(0.12)		(0.14)	(0.13)
Salary (1000s)			0.12**	0.14**	
			(0.05)	(0.06)	
Observations	1037	1037	1037	1037	1037

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 sergeant (lieutenant, major, captain) were grouped into "Sergeant +". Salary is monthly salary measured in thousands of dollars. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 9: Hispanic Bias on Labor Outcomes - Panel Results

	(1) Pr(Left Force)	(2) Pr(Moved Cities)	(3) Pr(Rank Up)	(4) Δ Salary
Hispanic bias	0.00	0.00	0.01	0.01
	(0.01)	(0.02)	(0.02)	(0.01)
Probationary Troop	0.02	0.36*	−0.78***	0.10
	(0.02)	(0.19)	(0.12)	(0.10)
Corporal	0.00	−0.02	−0.21**	0.04
	(0.04)	(0.09)	(0.09)	(0.06)
Sergeant+	−0.02	0.03	0.02	0.09
	(0.05)	(0.11)	(0.10)	(0.07)
Black Trooper	0.23	0.25*	0.20	−0.19
	(0.41)	(0.13)	(0.27)	(0.17)
Hispanic Trooper	0.01	0.29	−0.06	−0.18
	(0.05)	(0.29)	(0.24)	(0.13)
Constant	−0.05	−0.10	0.90***	0.84***
	(0.05)	(0.13)	(0.08)	(0.06)
Observations	587	533	546	546

Notes: Troopers with negative levels of bias are omitted from the regression. 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. Troopers with rank equal to or higher than sergeant (lieutenant, major, captain) were grouped into "Sergeant +". Omitted categories are white for trooper race and trooper for trooper rank. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 10: Hispanic Bias on Labor Outcomes - Transition Matrix

	(1)
Probationary Trooper	-0.000
	(0.011)
Corporal	0.021
	(0.050)
Sergeant+	-0.014
	(0.051)
Observations	766

Notes: Dependent variable is the probability of increasing in rank conditional on being the rank observed in the row. Each variable in the row is the reported rank of the trooper in 2013 interacted with the trooper's level of bias. Troopers with rank equal to or higher than sergeant (lieutenant, major, captain) were grouped into "Sergeant +". Regression has robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 11: Hispanic Bias on Complaints

	(1) I(Sustained Complaint)
Hispanic Bias	0.012^{**}
	(0.005)
Constant	0.054^{***}
	(0.005)
Observations	2041
R^2	0.003
F	4.756

Notes: Regression has robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 12: Hispanic Bias on Labor Outcomes - after 2015

	(1) Prob(Left Force)	(2) Salary Difference	(3) Prob(Rank Up)
HispBias	0.03	-0.03*	-0.04**
	(0.02)	(0.02)	(0.02)
Black	-0.01	-0.05	0.01
	(0.07)	(0.15)	(0.19)
Hispanic	0.03	0.02	0.12
	(0.04)	(0.13)	(0.18)
Prob. Troop		1.52***	
		(0.20)	
Corporal		-0.12***	
		(0.04)	
Sergeant		0.64***	
		(0.05)	
Lieutenant		1.49***	
		(0.26)	
Constant	0.29***	0.55***	0.27
	(0.09)	(0.13)	(0.21)
Observations	741	586	586

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. Each trooper is weighted by their total number of searches conducted from January 2010 to June 2015. Black and Hispanic are indicator variables equal to one if the trooper is black or Hispanic, respectively, and equal to one otherwise. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

7.1 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:

$$\phi_{M,t} = \frac{\pi_M[F_g^M(\theta^\mu_{M,t}) - F_G^M(\theta^*_{M,t})] + (1-\pi_M)[F_n^M(\theta^\mu_{M,t}) - F_n^M(\theta^*_{M,t})]}{\pi_M[1 - F_g^M(\theta^*_{M,t})] + (1-\pi_M)(1 - F_n^M(\theta^*_{M,t})]} \tag{11}$$

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:

$$\gamma_{m,t} = \pi_m[1 - F_g^m(\theta^*_{m,t})] + (1-\pi_m)[1 - F_n^m(\theta^*_{m,t})] \tag{12}$$

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:

$$\gamma^O_{M,t} = \pi_M[1 - F^M_g(\theta^\mu_{M,t})] + (1 - \pi_M)[1 - F^M_n(\theta^\mu_{M,t})] \tag{13}$$

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:

$$\begin{aligned} \gamma^O_{W,t} = \pi_W[1 - F^W_g(\theta^*_{W,t})] + (1 - \pi_W)(1 - F^W_n(\theta^*_{W,t})] \\ +\pi_M[F^M_g(\theta^\mu_{M,t}) - F^M_G(\theta^*_{M,t})] + (1 - \pi_M)[F^M_n(\theta^\mu_{M,t}) - F^M_n(\theta^*_{M,t})] \end{aligned} \tag{14}$$

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

$$\phi_{M,t} = \frac{\gamma^O_{W,t} - \gamma_{W,t}}{\gamma_{M,t}} \tag{15}$$

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.

7.2 Negatively Biased Troopers

Another important question is how to consider the troopers with negative measure of bias. Are trooper characteristics significantly different for troopers with negative bias compared to troopers with no bias ($Hisp\ Bias = 0$)? From the density plot in Figure 2 the 2,319 troopers, only 456 are biased, and 320 have zero bias. I first test if trooper characteristics vary significantly for troopers with negative bias compared to troopers with no bias. My results in Table 13 show that troopers with negative bias were less likely to be black troopers compared to troopers with zero bias. This is unsurprising since black troopers tended to be less biased, albeit not significantly so, compared to white troopers. When including the trooper rank, I find that probationary troopers were also less likely to be biased, which is surprising given the results in Table 8, which finds higher, but insignificant, levels of Hispanic bias for probationary troopers. I also find that corporals were more likely to also have negative bias, which is significant at the 10% level. In Column (3), I find that troopers with negative bias are not more likely to be paid more after controlling for trooper rank. I also find no correlation with experience.

One major concern of troopers with negative bias is that the negative bias is produced by errors in the race correction. This would be a major concern since this would imply similar issues even with troopers with positive bias. One way for this to occur is if troopers with negative bias have fewer searches and their misreporting measure is governed by over-correction of successes, which occurs by random. Columns (4) and (5) of Table 13 show that troopers with negative bias in fact search more in general and search more Hispanic motorists compared to troopers with no bias.

Further, if there was a systematic issue with the race estimation, then troopers with negative bias would misreport significantly more when searches with Hispanic motorists ended in success. To test this, I regress a version of Equation 6 but include a triple interaction indicating whether the trooper's level of bias is negative. Specifically:

$$\begin{aligned} I(RecRace = White)_{i,j,t} &= \alpha + \beta_0 I(CorrRace = Hispanic)_i + \beta_1 I(Failure)_i + \\ &\beta_3 I(Hisp\ Bias_j < 0) + \beta_4 I(CorrRace = Hispanic)_i \times I(Failure)_{i,j,t} + \\ &\beta_5 I(CorrRace = Hispanic)_i \times I(Hisp\ Bias_j < 0) + \beta_6 I(Failure)_i \times I(Hisp\ Bias_j < 0) \\ &\beta_7 I(CorrRace = Hispanic)_i \times I(Hisp\ Bias_j < 0) \times I(Failure) + \\ &X_{i,j,c,t}\gamma + \epsilon_{i,j,t} \end{aligned} \tag{16}$$

Instead, as seen in Table **??**, troopers with negative bias are not significantly less likely to misreport Hispanics when the search ends in failure compared to troopers with zero bias. Conversely, for troopers with positive bias, the point estimate is larger at 0.09. This means that troopers with positive bias are 9% more likely to record the Hispanic trooper as white when the search ends in failure compared to troopers with no bias.

Tables and Figures

Table 13: Negative Bias and Trooper Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
			I(Hisp Bias < 0)			
Black	0.003	0.002				
	(0.045)	(0.045)				
Hispanic	−0.025	−0.023				
	(0.024)	(0.024)				
Prob. Troop		0.010	−0.037			
		(0.073)	(0.078)			
Corporal		−0.032	−0.021			
		(0.032)	(0.032)			
Sergeant+		−0.006	0.030			
		(0.035)	(0.041)			
Salary			−0.029*			
			(0.017)			
Experience				−0.003*		
				(0.002)		
Total Searches					0.000	
					(0.000)	
Total Searches Hisp						0.000
						(0.000)
N	716	716	716	716	852	852

Notes: Dependent variable is an indicator variable equal to one if the trooper has negative bias and 0 if he has no bias. Regression uses employment data 2013 - 2015. Omitted category for trooper rank is trooper and the omitted category for trooper race is white. Salary is monthly salary in thousands of dollars.* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 14: Correlations of Trooper and County Characteristics

	(1) % Black Troopers	(2) % Hispanic Troopers	(3) % White Troopers
Violent Crime Rate	-1.047***	1.169	-0.148
	(0.372)	(1.134)	(1.232)
% Black	-0.548***	0.139	0.494***
	(0.075)	(0.114)	(0.136)
% Hisp	-0.516***	1.092***	-0.484***
	(0.064)	(0.092)	(0.117)
% HS diploma	0.104	0.528	-0.611**
	(0.099)	(0.328)	(0.309)
% Unemployed	0.147	0.020	-0.152
	(0.109)	(0.215)	(0.247)
% Latin American foreign born	0.008	-0.546***	0.541***
	(0.065)	(0.181)	(0.188)
% no health ins	0.361***	-0.008	-0.268
	(0.108)	(0.291)	(0.299)
Median HH inc (10000s)	0.016***	0.011	-0.027***
	(0.005)	(0.009)	(0.009)
Population (100000s)	0.004***	-0.002**	-0.002*
	(0.001)	(0.001)	(0.001)
Constant	0.334***	-0.521***	1.061***
	(0.078)	(0.191)	(0.208)
Observations	180	180	180
R^2	0.546	0.747	0.674
F	27.325	62.312	49.335

Notes: Dependent variable is in the column and is constructed from troopers who conduct at least 75% of their searches within that county. Counties with less than 4 troopers were excluded.
Regression uses robust standard errors show in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$