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

Elizabeth Luh†

August 8, 2023

Abstract

Individuals or organizations may attempt to hide biased actions by intentionally misreporting. I develop a model of highway searches, highlighting the incentive for biased troopers to misreport failed minority searches as White in an effort to appear less biased under commonly used tests of racial bias. Applying my model to highway searches in Texas from 2010–2015, I document widespread misreporting. Using the public backlash to the discovery of misreporting to study the effect of policy reform on policing, I find that the policy is effective in reducing misreporting, making the most biased troopers appear more biased relative to unbiased troopers.

JEL Classification: J15, K42

Keywords: Racial Bias, Systemic Misreporting, Traffic Enforcement

*This research uses confidential information from the the Texas Department of Public Safety, graciously provided by the Stanford Open Policing Project. I am grateful to Vikram Maheshri, Elaine Liu, Gergely Ujhelyi, Aimee Chin, and Chinhui Juhn for their continual support and invaluable guidance throughout this project. I am also grateful for helpful comments and suggestions from the University of Houston Economics Faculty, Jennifer Doleac, Keith Finlay, Ben Bushong, James Reeves, Ben Pyle, Steve Mello, Mike Mueller-Smith, Emily Owens, Brittany Street, and Cody Tuttle.

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

1 Introduction

In a 2017 survey, 27% of Latinos and 50% of Blacks felt personally discriminated against by police compared to only 10% of White respondents.[1] This perception is supported by a large and growing body of research identifying racial bias in nearly all aspects of the US justice system from airport screening (Persico and Todd 2005), ticketing (Anbarci and Lee 2014; Goncalves and Mello 2021), police stops (Coviello and Persico 2013), bail decisions (Arnold, Dobbie, and Yang 2018), sentencing (Shayo and Zussman 2011; Depew, Eren, and Mocan 2017), parole (Anwar and Fang 2015), use of force (Hoekstra and Sloan 2022), prosecutorial decisions (Tuttle 2021; Sloan 2022), to capital punishment (Alesina and Ferrara 2014). Recent events following the deaths of Trayvon Martin in 2012, Michael Brown in 2014, Philando Castile in 2016, George Floyd in 2020 along with many others at the hands of law enforcement have led to widespread calls for action making criminal justice reform a top priority for policymakers at all levels of government.

Despite this large academic literature on racial bias and discrimination in the criminal justice system, relatively little attention has been given to the response and tactics of law enforcement officers, especially biased officers, to this heightened scrutiny. These behavioral responses may be important as law enforcement officers control how civilian interactions are recorded and face little oversight on the accuracy of their record. Indeed, if officers deliberately misreport their interactions with civilians in order to avoid appearing biased then such systematic measurement could lead researchers to underestimate the extent of racial bias in the criminal justice system and could also hamper the efforts of those in charge of holding law enforcement accountable. Both of these adverse outcomes would be exacerbated if, as is plausible, the most biased officers have the most incentive to misreport. Thus, any policy intended to ameliorate bias may miss the target officers if this misreporting is unaccounted for. This issue extends to other settings where incentives to make target goals leads to manipulation of performance measures (e.g., test scores in education).

Moreover, since policing reform is oftentimes an endogenous response to public outcry, disentangling the two is an empirical challenge. For example, the death of George Floyd

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

in Minneapolis, Minnesota on May 25, 2020 ignited a nationwide protest which called for major police reform in the United States. As one example, the Minnesota state legislature passed major police legislation that banned chokehold restraint along with many other changes in police training (Bakst 2020). One could argue that these significant reforms might not have been passed in the absence of the public scrutiny and major calls to action by the public. Given the simultaneous timing, changes in officer behavior could be due to the public outcry, rather than the policy itself. This is an important distinction for determining the external validity of the impacts of policing reform as policy adoption across other agencies may not replicate the same change in behavior if the driver of the change was the public outcry.

In this paper, I develop a new model of racial bias in highway searches, building on seminal work of Knowles, Persico, and Todd (2001) and Anwar and Fang (2006), that links incentives to misreport to racial bias; specifically, biased troopers will use misreporting of motorist's race to appear less biased and thereby evade detection and punishment for bias. The model makes sharp predictions of the (mis)behavior of troopers: (1) only racially biased troopers will find misreporting to be profitable, (2) biased troopers will misreport if they fear punishment for racial bias, and lastly, (3) biased troopers must balance the risk of punishment for bias with the risk of punishment for misreporting.[2] While troopers may misreport for a variety of reasons (e.g. unsure about motorist race, poor visibility), I show that only biased troopers will misreport in the following way: by misreporting a portion of their failed minority searches as failed White searches. Through this specific pattern of misreporting, biased troopers can improve their reported minority search success rate and appear less biased. In addition, by misreporting minorities as White, troopers can also appear less likely to search minorities overall.

Another finding of the model is that misreporting increases with bias; thus, the level of misreporting itself can be used as a measure of racial bias at the trooper level as well. One major advantage of this novel measure and test of racial bias is that it is independent of the underlying propensity for criminality by race. Thus, I circumvent the inframarginality problem, a major empirical challenge for other past tests of racial

[2]Here punishment can mean a negative response (e.g. being skipped for promotion or not receiving the raise in salary) rather than an explicit punishment (e.g. administrative leave).

bias, such as the hit rate test (Knowles, Persico, and Todd 2001; Anwar and Fang 2006; Antonovics and Knight 2009; Feigenberg and Miller 2021; Arnold, Dobbie, and Yang 2018; Simoiu, Corbett-Davies, and Goel 2017).

I apply my model to a unique event in Texas where highway troopers were accused of widespread misreporting largely Hispanic, but also other non-White motorists, as White from 2010–2015 by local media in November 2015, leading to significant backlash from the public resulting in department wide policy change on trooper's stop behavior (Collister 2015b). Using a restricted data set of the universe of Texas highway searches from the Stanford Open Policing Project (SOPP) from 2010 to 2015 augmented with publicly available highway stop data from 2016–2018.[3] To recover the "true" race in the context of misreported administrative records, I use the driver's full name and residence ZIP code to estimate the true race of each driver.

Applying my model to the data, I use the difference in misreporting likelihood across search outcome (i.e. whether or not contraband was found) to determine if the misalignment between recorded and estimated race was due to racial bias. I show that this misreporting was concentrated on Hispanic motorists and that, on average, likely Hispanic motorists were 8% (relative to a mean of .274) more likely to be reported as White in failed searches than they were in successful searches. I interpret this as evidence that a part of the discordance between recorded race and estimated race was due to intentional misreporting to mask racial bias against minority motorists, especially Hispanic motorists.

Because troopers stop hundreds of motorists over the course of a year, a major contribution of my analysis is that I am able to estimate racial bias at the trooper level. I leverage the public outcry and subsequent policy reform on race recording rules, which required troopers to always verbally ask drivers for their race, as exogenous events to test the robustness of the trooper-level estimates of bias. Since the policy reform made misreporting significantly harder, it provides an intuitive test of the estimated measures of racial bias since only troopers estimated to be misreporting should change their race

[3]Both data sets originated from the Texas Department of Public Safety. The publicly available data (2016–2018) does not contain driver's full names or home addresses but is otherwise identical to the restricted SOPP data. The publicly available data contains stops from 2010–2015 and fully overlaps with the restricted data. I have verified that the restricted and publicly available data are the same.

reporting patterns across search outcome. This is indeed what I find when I compare the gaps in search success rates before and after the policy for estimated biased and unbiased troopers. Specifically, for the most biased troopers, the Hispanic-White gap in search success rates increased by 9.6 percentage points relative to the pre-reform gap, with no response from unbiased troopers. Thus, without being able to misreport failed Hispanic searches as White, the gap in search success rates widened significantly for biased troopers after the policy went into effect. In the absence of misreporting, biased troopers appear more biased under traditional tests of bias (i.e., the hit rate test).

Additionally, public scrutiny resulting from the article publication also affected trooper's search behavior. I find that search success rates rose from 45% in 2014 to almost 60% in 2016, the first year of the policy implementation, regardless of whether troopers were misreporting or not. Thus, the heightened scrutiny seems to have improved all troopers' search behavior, regardless of whether the policy impacted their race recording behavior or not. But, this improvement is short-lived with search success rates falling back to their 2014 levels in 2017 and 2018. Thus, once public scrutiny faded, troopers reverted to their prior search success rates. On the other hand, the policy reform was effective in eliminating the misreporting behavior with the proportion of drivers races within searches remaining stable from 2016 onwards. These results demonstrate the importance of both public outcry and policy changes when considering the external validity of policing reform in the United States.

Lastly, I examine the impact of misreporting and racial bias on trooper's labor market outcomes. I find that in the presence of misreporting, labor market outcomes for troopers are uncorrelated to their level of bias. Exploiting the sudden and plausibly exogenous revelation of misreporting to the public in 2015, I find that Hispanic bias reduced monthly salary growth by 7% for troopers in the top quartile of the racial bias distribution. Surprisingly, I also find that troopers who were engaging in the reverse behavior (misreporting successful Hispanic searches as White) have decreased salary growth of 8%. Since these negative ramifications to misreporting behavior were only found after the rule change, I interpret this as suggestive evidence of (1) misreporting being a punishable behavior, and (2) public scrutiny being a potential driver of policy change.

This paper contributes to the literature on detecting racial bias in the criminal justice system, specifically in contexts where law enforcement officers are responsible for recording the interaction. Many earlier contributions to the literature, notably in motorist stops and racial bias, examine the role of motorist race and trooper race in stop interactions, such as Knowles, Persico, and Todd (2001) and Anwar and Fang (2006) along with Antonovics and Knight (2009) and many others. A majority of these papers build off of Becker (1957) outcome test, which identifies racial bias by comparing the success rates across different groups. Another body of work, Grogger and Ridgeway (2006), West (2018), and Goncalves and Mello (2021), use the context of the search (time of day, vehicular accident, officer leniency) as exogenous variation for measuring racial bias and also find racial bias against non-White motorists.

These findings also build on past work of Jacob and Levitt (2003), Dee et al. (2019), and Mayzlin, Dover, and Chevalier (2014) along with a wealth of other papers on the impact of targeted measures on cheating behavior. Notably, this body of work finds that when individuals know they're being assessed on certain measures, they have incentive to manipulate the results and their behavior to achieve these thresholds rendering these assessments into uninformative measures. Similarly in this context, troopers' misreporting behavior obscure efforts to assess the agency's racial bias. While past work has studied this behavior in the context of schooling (student and teacher behavior) and in firm reviews, this paper is the first to study this cheating behavior in the context of policing.

The miscategorization of Hispanic as White in policing has been found in other research and noticed by the public as well. Recent work, such as Goncalves and Mello (2021) and Feigenberg and Miller (2021), accounts for this potential misreporting for Hispanic motorists by similarly imputing ethnicity using data driven methods or comparing to other administrative data sets (e.g., drivers' licenses or criminal records). But, these works so far have considered the misreporting as unbiased measurement error and have not linked this misreporting itself to racial bias. On the other hand, the public have begun to link misreporting to racial bias. For example, journalists in Louisiana have uncovered systematic misreporting in parishes across Louisiana and have directly linked the the misidentification as an attempt to obscure efforts to measure racial disparities in police contact and racial bias.

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

2 Background

2.1 Misreporting and Highway Troopers in Texas

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

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

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

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

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

Due to Senate Bill 1074 passed in 2001, Texas DPS is required to publish an annual traffic stop data report to provide "background pertaining to the issues of racial profiling (Public Safety 2012)." This report, aptly named 'Racial Profiling Report,' breaks down search, stop, and citation statistics across race and ethnicity. While the report doesn't explicitly call this the 'hit rate test', the report also includes the number of criminal arrests resulting from a traffic search across race.[4] Thus, troopers were aware of how their search and stop patterns, especially their hit rate, may be used to determine racial bias, further motivating potential misreporting behavior.

2.2 2015 Misreporting Incident

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

Texas troopers were already under scrutiny due to the death of Sandra Bland in jail after being pulled over for failing to signal a lane change (Sanchez 2015). One week

[4] See https://www.tcole.texas.gov/content/racial-profiling-reports to see reports from 2016–2023.

after the misreporting was uncovered, the House Committee on County Affairs held a hearing where DPS blamed the error on a computer glitch. As a result of the hearing, DPS changed its policies to require troopers to ask drivers to provide their race, rather than recording it based on the trooper's best judgment. This policy went in effect by November 23rd; as a result of the policy, the proportion of stopped motorists recorded as White fell from 18% to 4% by 2016 (Collister 2015a; Oyeniyi 2015). The timeline of events was quick with only 15 days between article publication and the policy change.

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.

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

3 Model

Motorists of race m travel on highways; a fraction π^m of them are carrying contraband. Trooper t may stop motorists without observing their race. Conditional on stopping a motorist, a trooper receives a signal θ that contains all available information on whether the motorist is carrying contraband.[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,

[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.

I allow for motorists to be W or H (Hispanic).

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

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

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

3.1 Bias and Misreporting

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

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

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

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

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

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

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

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

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

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

v is a monotonic transformation of b. Since $f_{g,n}^{m}$ and π_m are unobservable, proving that the measure of v is driven by b (racial bias) and not $\theta^*_{M,t} - \theta^*_{W,t}$ (statistical discrimination) is key to identifying v as a measure of b.

Troopers may face punishment for biased policing with probability P, which is monotonically increasing in $|b|$. In order to evade detection, a trooper may intentionally misreport the race of a motorist following a search, which will reduce the appearance of bias and thereby the likelihood of detection. But, troopers incur a cost of μ for misreporting, as it may open the door to greater punishment. I make the following assumptions on μ, the cost of misreporting:

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

As θ, increases, the cost of misreporting also rises. Therefore, motorists who appear less guilty are more likely to be misreported. One intuitive reason for this is that motorists with higher θ are in general more likely to be searched. Thus, by misreporting motorists who appear less guilty, the trooper is less likely to be caught misreporting.[6]

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

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

hold. The presence of misreporting is a clear test of racial bias that is not associated with statistical discrimination and is independent on π_m, the likelihood of carrying contraband for race group m, thus avoiding the inframarginality problem.

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

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

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

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

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

Given this set up, I obtain the following result:

Proposition 1. *Under Assumption 1, 2, and 3, troopers will misreport motorists with characteristics (M, θ) if and only if $\theta \in (\theta^*_{M,t}, \theta^{\mu}_{M,t})$ and the search ends in failure.*[7]

Troopers will only misreport their failed searches. Because the misreporting decision is conditional on search, any misreported motorists must have $\theta > \theta^*$. Troopers also will not misreport motorists over a certain threshold, specifically $\theta > \theta^{\mu}$. That is, motorists who appear more guilty than the search threshold will not be misreported.[8]

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

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

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

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

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

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

For unbiased troopers, $v = 0$. For biased troopers, $v > 0$.[9] The magnitude of $v_{M,t}$ itself will also be trooper t's measure of bias against race M motorists. This forms the basis of my measure of racial bias for trooper t against race M motorist that I use throughout the rest of the paper.

4 Data

4.1 Stop Data

The Stanford Open Policing Project (SOPP) has a restricted version of highways stops conducted from 2005 to 2015 from the Texas Department of Public Safety. The restricted version contains personally identifiable information of the driver such as full name, home address, owner's full name, and license plate of the stopped vehicle. Pierson et al. (2020) courteously provided the raw version of the data.[10] As DPS did not record the driver's last name prior to 2010, only stops from 2010 onward are included in the study.

The data also has rich stop information such as the latitude, longitude of the stop, the badge number of the officer who recorded the stop, the race of the driver, the state in which the driver's license was issued, and the make and model of the vehicle. The data also has information on the violation such as reason for the stop, the outcome of the stop (citation, warning), whether a search was conducted, the search reason, and the outcome of the search. The highway stop data is publicly available on the TX DPS website from 2013–2019.[11]

I also augment the SOPP data with 2016–2019 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 drivers in the data set. Since the stops occurred after the misreporting was revealed in

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

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

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

November 2015, I take the driver's races as given. Given that the proportion of driver's race is stable from 2017, as shown in Figure 1, it appears that the recorded driver's race actually reflects the driver's true race.

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

4.2 Trooper Employment Data

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

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

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

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

I also include trooper complaint data from 2010 to 2015, which I obtained using a FOIA, as a secondary measure of trooper work behavior. The complaint data contains information on the date the incident occurred, the date the complaint was received, the allegation of the complaint, the trooper's badge number (if applicable), and the investigator of the complaint. The badge number is not always included due to Texas' privacy laws.[12] Out of the original 1,873 complaints, only 334 had the trooper's badge number in the complaint.

4.3 Race Estimation

The ideal way to uncover driver's true races would be to link the stop data to an administrative data set of race and ethnicity. Unfortunately, the cost to link these data sets is prohibitively high. Other Texas governmental agencies were also unwilling to share their administrative data for this research project as well.

To circumvent the lack of alternative data on driver's true race, I use two main methods supported by past literature on using observable characteristics to determine race. These methods are predominantly used in social science and health research to infer patient race (Fiscella and Fremont 2006; Freedman, Owens, and Bohn 2018). Recent work by Goncalves and Mello (2021) and Feigenberg and Miller (2021) also use similar methods to impute Hispanic ethnicity when measuring racial bias in highway stops. Furthermore, Feigenberg and Miller (2021) using the exact data set used in this paper verify their ethnicity imputation using arrest records from the Texas DPS and find a correlation of 0.74.[13]

[12] 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)."

[13] DPS is the same agency overseeing the Texas highway troopers. Thus the misreporting present in highway stop data may also be present in the administrative data on criminal records. Furthermore, the

I use surname analysis combined with the same home address analysis for predicting drivers' race and ethnicity. I match each driver's last name to a surname using the 2010 Census Surnames data set. Using Hispanic motorists as an example, if the probability of the last name is Hispanic is greater than a certain threshold (90%) and the proportion of Hispanic residents within the ZCTA5 area is greater than 75% or the home address is from a Spanish speaking country (e.g Mexico), I impute the 'estimated' race as Hispanic. Assuming this driver resides in a ZCTA5 area with proportion Hispanic greater than 75%, given the probability this driver is Hispanic, conditional on his last name, Mendez, is 92%, I estimate his actual race to be Hispanic.[14] While this may raise concerns of bias in coding races (e.g. officers local knowledge of neighborhoods), DPS troopers mainly police interstate highways of unincorporated areas, which will have more non-residential travel compared to local neighborhoods. Thus, it is unlikely that troopers are stopping populations within their local community.

The surname imputation performs poorly with Black individuals as they tend to have less distinctive last names. Notably, only Washington and Jefferson exceed 75% likelihood of identifying as Black given that surname. Thus, I impute Black racial identity if an individual resides in a ZCTA5 neighborhood exceeding 90% proportion Black. Similarly, the ZCTA5 imputation performs poorly with Asian individuals as Texas does not have enough neighborhoods with high concentrations of Asian motorists. Thus, I only apply the surname threshold for imputing Asian motorists.[15]

The goal of this race imputation exercise is to uncover intentional race misreporting, not driver's true race, an important distinction as even with better race imputation strategies, estimated drivers' race may still be mis-measured as driver's can self-identify as any race/ethnicity regardless of their home address or last name.[16] Furthermore, even

authors can only verify their race imputation for motorists with arrest records, a selected subsample of the universe of highway drivers in Texas.

[14] In Appendix Figure A.9, I vary the surname and ZIP code cutoffs from 75% to 90% in 5 percentage point intervals to ensure that the thresholds are not driving my estimates of racial bias. My results are robust to these various thresholds.

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

[16] Goncalves and Mello (2021) also estimate driver's race in their study and are able to corroborate their estimates of driver's race with driver's license data in Florida. Unfortunately, TX driver's license data does not report driver's race nor are they willing to share the data. Feigenberg and Miller (2021) corroborate their estimates of driver's race using criminal history reports from TX DPS, which is also

with perfect race estimation, troopers could misreport drivers' race for reasons aside from bias (e.g., poor visibility, human error). The key assumption for my identification strategy is that errors in race imputation and troopers' unintentional misreporting are independent of search outcome. Thus, I can uncover unbiased estimates of intentional misreporting by comparing the difference in misreporting behavior across search outcome. Based on my model of racial bias, only troopers who are misreporting to hide racial bias will be more likely to misreport when the search ends in failure, compared to success.

4.4 Descriptive Statistics

I present summary statistics of motorist characteristics in Table 1 using the recorded races. On average, I find that White motorists are over represented in both searches and all stops. From the 2010 Decennial Census, only 45% of Texas residents were non-Hispanic White, but make up nearly 70% of the stops and 60% of searches. Black and Hispanic motorists are searched at nearly equal rates of 10% and 13% respectively and are under-represented given the 2010 Decennial Census which reports 11.9% and 40% respectively. I also find that certain stop characteristics, such as stops occurring from 8 PM - 5 AM and whether the car is older than 5 years are more likely to occur in searches compared to stops.

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

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

In the bottom part of the table, I break down the stop and search statistics by trooper position. Troopers with rank of lieutenant or greater are aggregated to the same rank

the source of traffic ticket data used to estimate the race misreporting in this paper. Thus, the accuracy in the race reporting in the DPS criminal history data is likely erroneous as well.

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

5 Empirical Results

5.1 Test for racial prejudice

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

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

The change in proportions of reported driver's race within searches may be driven by reasons aside from bias. Troopers may have not known about the importance of accurate race data and thus did not expend effort in ensuring its accuracy. Troopers could also have unintentionally misreported driver's race for reasons aside from bias, such as poor visibility conditions or poor race identification ability. The model again proves useful for helping separate the previous examples from the misreporting behavior of interest. If an officer's misreporting is intentional and linked to racial bias, then misreporting is only beneficial if unsuccessful minority searches are misrecorded as unsuccessful White searches. In the examples of misreporting that are unrelated to bias, those examples should be equally likely to occur regardless of the search outcome. Thus, testing the likelihood of race mismatch across search outcome provides an intuitive test for whether the mismatch is driven by racial bias or not.

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

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

The coefficient of interest is β_1, which indicates the increased likelihood of race mismatch for failed searches. In other words, how much more likely does mismatch occur when searches end in failure compared to success? $X_{i,c,t}$ is a vector of controls for the stop, including hour of the stop, month of the stop, year of the stop, county fixed effects, and vehicle type and vehicle age. I also include the full interaction for hour of the stop with month of the stop and year of the stop to control for seasonal and darkness variation that may impact driver's race visibility and the full interaction of vehicle characteristics (vehicle age and vehicle type (e.g., truck, sedan, van)) since these may be inputs in an officer's search decision.

Table 3 shows the estimates of β_1 for each Hispanic, Black, and Asian motorists in Columns (1), (2), and (3) respectively. The estimated coefficients show that most of the mismatch between recorded and estimated race is concentrated on Hispanic motorists with an average 27.4% of estimated Hispanic searches resulting in mismatch. Furthermore, for estimated Hispanic motorists, this mismatch appears to be linked to misreporting with estimated Hispanic motorists 2.3 percentage points significantly more likely to be misrecorded as White when searches end in failure compared to success. Thus, failed Hispanic searches are 8% more likely to be misreported compared to success Hispanic searches.

In contrast, only .2% and of 13% estimated Black and Asian searches are mismatched on average, respectively. This could be due to the high threshold for race estimation, with very few ZCTA5 neighborhoods exceeding the 90% proportion of Black residents. At the same time, as noted earlier in Figure 1 much of the misreporting was concentrated on Hispanic motorists, which could also contribute to the the lower estimates of mismatch rates for Black motorists. Troopers do not appear to be systematically misreporting Black and Asian motorists at similar rates to Hispanic motorists as I do not find any significant differences in misreporting rates across search outcome. Furthermore, these coefficient estimates are near zero at 0.01 percentage points and 2 percentage points for Black and Asian motorists, respectively.

There are a few main reasons why misreporting was concentrated on Hispanic motorists rather than other the race groups. The first is that Hispanic is technically an ethnicity and not a race. Thus, Hispanic drivers could technically be recorded as White, despite Hispanic being the more accurate race code. Another reason is that Texas' proximity to the border and the contentious immigration flows may lead to greater animus towards Hispanic motorists compared to other non-White groups. Lastly, unlike Black or Asian, Hispanic was not always included as a recorded race group. Prior to 2010, DPS Hispanic motorists were recorded as White. Thus, troopers may have recorded Hispanic motorists as White out of habit.

5.2 Estimating Officer-level Hispanic Bias

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

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

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

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

From prior work using these value-added models,[17] the distribution of $\hat{\beta}_1^j$ will have a higher variance relative to the true distribution due to estimation error. Compounding on this, the few number of searches the trooper-level estimate of bias introduces potential measurement error, further attenuating the estimates. To correct for this, I follow the Bayes shrinkage procedure from Morris (1983) to estimate the distribution of bias accounting for the estimation error in each $\hat{\beta}_1^j$.

Formally, I calculate the shrinkage estimate by assuming that $\beta_1^j \sim \mathcal{N}(\beta_1, \sigma)$, which I estimate directly.[18] Using the standard errors associated for trooper j's estimated $\hat{\beta}_1^j$ and $\hat{\sigma_j}$ and following Morris (1983), I estimate $B = \frac{\hat{\sigma}^2}{\hat{\sigma_j}^2 + \hat{\sigma}^2}$, with B being the shrinkage factor. Lastly, I calculate the shrinkage estimator as $\tilde{\beta}_1^j = B\hat{\beta}_1^j + (1 - B)\beta_1$.

Figure 3 shows the raw bias estimates (solid line, black) plotted with the shrunken estimates of bias (dashed line, blue). The further right the trooper is in the distribution of bias, the higher his level of bias. The measurement of bias is the difference in likelihood of misreporting between his failed searches and his successful searches. For example, a trooper with estimated bias of 0.5 is 50 percentage points more likely to misreport his failed searches of estimated Hispanic motorists compared to his successful searches. The average Hispanic bias using the original estimates is 0.042 with standard deviation of 0.276. Thus, the average officer is approximately 4 percentage points more likely to misreport his failed Hispanic searches as White compared to his successful Hispanic searches. The average standard error for $\hat{\beta}_i^j$ is 0.498. Since troopers on average only search

[17]See Aaronson, Barrow, and Sanders (2007), Guarino et al. (2015), Goncalves and Mello (2021), Jackson (2018), Koedel, Mihaly, and Rockoff (2015), Morris (1983), and Weisburst (2022).

[18]β_1^j represents the true value of trooper level Hispanic bias. β_1 is the average trooper Hispanic bias.

approximately 2% of all stops, the high standard error is expected. Thus, the distribution shrinks significantly (to standard deviation of 0.058) after applying the Morris' Bayes shrinkage procedure.

Given the high rates of shrinkage, I collapse the distribution of bias into even quartiles. Thus, while each of the bias estimates are measured with error, the categorizations into quartiles are less prone to error. For example, a trooper with an estimated bias of 0.5 in the 4th quartile might not be more biased than a trooper with estimated bias of 0.48 in the 4th quartile but is likely to more biased than a trooper of estimated bias 0.05 in the 3rd quartile. This will reduce the measurement error while retaining the ordinal ranking of bias. Table 4 shows the range of bias values within each of the quartiles for both the unshrunk and shrunken estimates of bias. Largely, troopers in Quartile 2 are centered on 0, or no bias regardless of using the unshrunk or shrunken estimates. Thus, I consider troopers in quartiles 3 and 4 as 'semi-biased' and 'biased' troopers, respectively. Since troopers in the first quartile have negative estimates of bias, I label these troopers as 'negatively' biased troopers.

5.3 Using the 2015 policy change as a robustness check for the validity of the trooper level bias estimates

Taking the publication of the article revealing the misreporting in 2015 as a plausibly exogenous shock to troopers, I assess how the revealment of the misreporting to the public along with the subsequent rule change affected troopers' misreporting behavior conditional on their estimated bias estimates. Notably, the rule change should only affect troopers who were purposefully misreporting minorities, thus the 2015 policy change should only impact the differential recording of race across search outcome for those I estimate to be biased. Specifically, the search success rate for recorded Hispanic motorists should fall significantly after 2015 relative to recorded White motorists if the trooper was misreporting their failed Hispanic searches as White. If on the other hand, troopers were only misreporting due to poor abilities to perceive race, then I should find no significant change in recording behavior across drivers' race conditional on search outcome after the rule change. Given the change in proportions of drivers searched by race in Figure 1, the policy did significantly impact troopers race recording behavior.

An ideal test would be to reevaluate the differential rate of misreporting across search outcome using searches conducted from 2016–2019 using Eq. 6 and 7. Unfortunately, the restricted data from SOPP ends in 2015. From 2016 onwards, the publicly available data only contains recorded driver's race and no identifying characteristics of the driver (i.e., full name, home address) that I can use to impute race. To circumvent this, I leverage how the impact of misreporting, if used to hide bias, distorted the gap in the observed Hispanic-White search success rate. Specifically, with misreporting, the gap in search success rates across race will be smaller for biased troopers. Thus, the policy change will increase the Hispanic-White gap in search success rates using the recorded races for biased troopers only. Specifically, the article publication and rule change should widen the observed gap in search success rates for biased troopers only and have no impact for non-misreporting troopers.

To assess this, I evaluate the impact of the article publication on the difference in search success rates between recorded Hispanic and White motorists using an event study framework for troopers who are estimated to have no bias (quartile 2) and troopers who are estimated with bias (quartile 4), separately. I run the following linear probability model for each recorded White or Hispanic driver i stopped in county c at time t:

$$\begin{aligned} I(Success_{i,c,t}) = &\beta_0 + \sum_{t=2010}^{2019} \beta_1^t I(Year = t) + \beta_2 I(Race_i^{Recorded} = Hispanic) + \\ &\sum_{t=2010}^{2019} \left[\beta_3^t I(Year = t) \times I(Race^{Recorded} = Hispanic)\right] + X_{i,c,t}\gamma + \epsilon_{i,c,t} \end{aligned} \tag{8}$$

where $I(Success_{i,c,t})$ is an indicator variable equal to one if the search ends in success. I include the same set of controls included in Eq. 6.

The coefficient of interest is β_3^t, which describes the change in the difference in search success rates between recorded Hispanic and White motorists relative to the omitted year, 2014. I do not use 2015 as the omitted year since trooper behavior may be endogenous to the policy adoption. As Figure 1 shows, the proportion of Hispanics in searches dropped sharply in the first quarter or 2015 coinciding with an increase in White searches, which could indicate an increase in misreporting in that time period leading to the article publication. Given this behavior, I choose 2014 as the comparison year.

If the policy change had no impact on race recording behavior across search outcome, then $\beta_3^t = 0$ for $t > 2015$, which should be the case for troopers I estimate as having 'no bias.' Since these troopers were not misreporting, the policy should not impact the difference in search success rates across race. If troopers were indeed using misreporting to mask their failed Hispanic searches as White, then I should observe $\beta_3^t < 0$ for $t > 2015$. Specifically, since misreporting shifts failed Hispanic searches into White searches, the recorded White search success rates should increase relative to the Hispanic search success rate once the policy goes in effect and misreporting ends.

However, the heightened public scrutiny from the article publication may also influence troopers' search behavior across race, especially for biased troopers, who are the target of the race reporting policy change. Specifically, biased troopers may become more judicious about their search decisions, especially across race. Thus, if troopers improved their Hispanic search decisions (↑ Hispanic search success rate) while also reducing their misreporting of Hispanic motorists (↓ Hispanic search success rate, ↑ White search success rate) without changing their search behavior with White motorists, this could cancel out any observed changes in the Hispanic-White search success rate.

Figure 4 panels A and B shows each of the β_3^t estimates with the unshrunk and shrunken estimates of bias, respectively. Each of the graphs in the panel plots the β_3^t estimates for misreporting troopers (Quartile 4, red circles) and non-misreporting troopers (Quartile 2, blue diamonds) for each year, t.[19] I also estimate the average effect of the policy using a two-way fixed effects framework and plot the implied difference-in-difference coefficient in the maroon line.[20] Here I find interesting dynamics when comparing the changes in the gap in search success rates across ethnicity between the two groups of troopers. First, in 2016, the first year the policy was in effect, the gap observed in 2015 for misreporting troopers increases to -6.3 percentage points, indicating a fall in the search success rates for Hispanic motorists relative to White motorists. The

[19] For completeness, Appendix Figure A.10 shows the same figure with negatively biased troopers (Quartile 1) and 'semi'-biased troopers (Quartile 3).

[20] In this regression, I estimate the average change in Hispanic search success rates relative to the White search success rate due to the policy. I define the after period as 2016–2018 and omit searches conducted in 2015 from the regression. I include the same set of controls, $X_{i,c,t}$ and cluster my standard errors at the county level.

gap continues to widen to -11.5 percentage points by 2017. In 2018, the gap shrinks to -3.5 percentage points but is still significant, indicating a sustained lower Hispanic search success rate relative to the White search success rate for misreporting troopers. This change in search success rates for biased troopers could be driven by either the policy change or the increased public scrutiny from the article publication. I explore this further detail in Section 5.5.

This reduction in the gap in search success rates observed in 2018 between Hispanic and White motorists may indicate a reversion to misreporting behavior. This is plausible as the department did not give explicit details for how they would enforce the policy. Furthermore, public scrutiny may have faded at that point, further reducing pressure on trooper behavior. However, the stability in the proportions of driver's races within searches after 2015 in Figure 1 assuages concerns that troopers reverted to misreporting after 2017. As shown in the figure, the proportion of Hispanic motorists within searches remains steady at around 40%. Thus, I interpret the results in Figure 4 as evidence that biased troopers used misreporting to artificially boost Hispanic search success rates prior to 2015 thereby reducing the appearance of racial bias. Furthermore, the use of misreporting as estimates of bias, while noisy, does identify bias, especially for troopers in the 4th quartile.

For non-biased troopers, or troopers in the second quartile of the estimated bias distribution, I do not find a significant change in the difference in search success rates between Hispanic and White motorists as a result of the policy change regardless of the comparison year. Notably, nearly all of the estimates are insignificant and close to 0 indicating no change in the difference in search success rates across White and Hispanic motorists. The point estimates for non-misreporting troopers are not only insignificant, but also relatively small, with nearly all estimates less than 2 percentage points. The exception is the estimated Hispanic-White gap in 2018, an insignificant 3 percentage points. Overall, the patterns emerging from the regression support the interpretation that the quartiles of bias estimates are indeed describing trooper's misreporting behavior as only troopers with estimated positive amounts of bias responded to the policy change.

Reassuringly, Figure A.10 also shows that troopers in Quartile 1 ('negatively' biased troopers) do not significantly respond to the 2015 policy change and in fact behave similarly to unbiased troopers. Notably, I do not observe a significant change in the gap in Hispanic-White search success rates after the policy change. While I do find a marginally significant gap in search success rates in 2011 and 2012 of 5 percentage points, this relationship may be mechanical as troopers in Quartile 1 by definition are misreporting their successful Hispanic searches as White, thus their reported White search success rate will be higher than their reported Hispanic search success rate.

For troopers in Quartile 3, I observe a pattern similar to troopers in the fourth quartile. Here, I find that troopers in the year of the policy enactment (2015), the Hispanic-White gap in search success rates increases by 4.2 percentage points relative to 2014 and is marginally significant at 90%. As hypothesized earlier from Figure 1, the drop in the proportion of Hispanic motorists within searches in 2015 prior to the policy change may be from increased misreporting. If this misreporting was to hide bias, then the search success rates for Hispanics would increase relative to the White search success rate. This increase may be suggestive evidence that troopers in Quartile 3 increased their misreporting just prior to the events in November 2015. Supporting this hypothesis, for 2016–2017, I do not find any significant change in the Hispanic-White search success rate gap relative to 2014. In 2018, the gap in search success rates does decline and is marginally significant with the Hispanic search success rate being 7.4 percentage points lower than the White search success rate relative to 2014. Thus, they do appear more biased under the hit rate test relative to their pre-policy performance.

5.4 Other robustness checks

To test the validity of my measure, I regress $\hat{\delta}_j$, each trooper's overall mismatch rate, with $\hat{\beta}_1^j$ from Eq. (9). This yields the correlation between trooper's average mismatch rate and the trooper's Hispanic bias. Figure 5 shows the scatter plot of $\hat{\beta}_1^j$ against $\hat{\delta}_j$. The trooper's average mismatch rate can be interpreted as a measure of the trooper's ability to identify driver's race accurately. Higher rates of mismatch, $\hat{\delta}_j$, indicate that the trooper's own imputation of race does not match the estimated race imputation, regardless of search outcome. Thus, a positive correlation would indicate that trooper's with lower ability to accurately identify race are more likely to have higher rates of bias, undermining the estimates of bias. I find the opposite relationship; biased troopers have lower rates of mismatch overall, indicating higher ability to accurately measure driver's race.

As a second robustness check, I estimate officer-level Black bias using the same specification in Eq. (7), but restricted to estimated Black motorists. If a trooper is using misreporting to hide their Hispanic bias, this trooper may also be using the same misreporting to hide bias for other minority motorists. Using the main thresholds in the prior analysis, I only estimate approximately 80 drivers as Black who were originally recorded with another race. In order to increase the variation needed to estimate officer-level bias, I lower the thresholds for Black race estimation to 75%. Lowering the threshold introduces measurement error but will not bias the estimates as long as the error is orthogonal to the search outcome.

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

Lastly, I ensure that the results are not driven by the race estimation thresholds. Figure A.9 in the Appendix shows the distribution of officer level Hispanic bias across various thresholds. The distribution does not change significantly across the different thresholds and 90% has the highest concentration at 0 bias. Although the distribution widens slightly as the threshold lowers, the increase is minimal and does not affect the overall distribution.

5.5 Impact of misreporting rule change on search success rates

Given the public furor over the misreporting, a natural question to explore is how trooper's search behavior changed after the rule change. In addition to the policy change, troopers were now under more scrutiny from the public, which could impact their search behavior overall. Specifically, troopers might be more judicious about their search decision and thus only conduct searches where likelihood of finding contraband is higher.

To explore this, I use an event study framework to estimate the change in search success rates before and after the policy across each quartile of bias. Specifically, for recorded Hispanic and White motorists,

$$\begin{aligned} I(Success_{i,c,t}) =& \beta_0 + \sum_{t=2010}^{2019} \Big[\beta_1^t I(Year_i = t) \times I(Bias_i^{Q1}) + \beta_3^t I(Year_i = t) \times I(Bias_i^{Q3}) \\ &+ \beta_4^t I(Year_i = t) \times I(Bias_i^{Q4})\Big] + \sum_{\substack{t=2010, \\ \neq 2015}}^{2019} \Big[\beta_2^t I(Year_i = t) \times I(Bias_i^{Q2})\Big] \\ &+ X_{i,c,t}\gamma + \epsilon_{i,c,t} \end{aligned} \tag{9}$$

where $I(Success_{i,c,t})$ is an indicator variable equal to 1 if the search ends in success regressed on indicator variables for if the search occurred in year t, $I(Year = t)$ fully interacted with the quartiles of bias. The omitted group is searches conducted by troopers in the second quartile in 2014. I include the same set of fixed effects as in Eq. (6).

I plot the coefficient estimates with 95% confidence intervals in Figure 8 adjusted with the 2014 average search success rate for the omitted group, troopers in Quartile 2, which I also plot in each graph (point without confidence interval). Panels A and B show the change in search success rates for motorists recorded as Hispanic or White using the unshrunk and shrunken estimates of bias, respectively. The results show an interesting pattern. Over time, troopers regardless of bias, significantly improved their search behavior overall. On average, the search success rates rose by approximately 10 percentage points from 2010 to 2014. Troopers in the first quartile of bias had the greatest improvement, with the search success rates increasing from 30% in 2010 to 45% in 2014.

Interestingly, the events in 2015 seemed to have shocked all troopers, regardless of bias. Compared to 2014, the search success rate increased by 12 to 17 percentage points in 2016, the first year following the article publication and policy adoption. Since the race recording policy change only impacted misreporting troopers, the improvement in search success rates for the non-misreporting troopers (Quartile 2) may be driven by the public scrutiny from the backlash to troopers misreporting and the death of Sandra Bland. Supporting this hypothesis, the improvement was not sustained, with the search success rates dropping back to the pre-policy search success rates for all troopers in 2017. If the race recording policy change was the main driver of the improved search success rates, then these improvements should be sustained.

Panels B and D in Figure 8 plots the estimates for the same regression, estimated separately for recorded White (hollow shapes) and recorded Hispanic (filled in shapes) motorists. Thus, the vertical distance between the hollow and filled in shape visually depicts the gap in search success rates. Within the same time period, the changing gap size across quartiles of bias represents the impact of misreporting on tests of racial bias, such as the hit rate test.

Separately estimating the change in search success rates across ethnicity reveals more interesting impacts of the events in 2015 and misreporting on observed trooper search behavior. Notably, prior to 2015, the Hispanic-White gap in search success rates largely decreases as the quartiles increase. This relationship is partly mechanical by how troopers were binned into each quartile but also visually demonstrates how misreporting made troopers appear less biased under the hit rate test. Once the policy goes into effect, or after 2015, I find that the trend in the Hispanic-White search success rate gap across quartiles changes and in fact reverses in 2017. Troopers in Quartile 4 and 3 now have the largest gaps in the Hispanic-White search success rate with a gaps of 22 percentage points (Quartile 4) and 17 points (Quartile 3). This pattern is not mechanical as the post-2015 searches are not used to determine the bias distribution. Thus, this provides more supporting evidence that the binning of troopers into the quartiles does actually rank troopers by their misreporting behavior and racial bias.

For the most biased troopers, the gap in 2017 is double the size of the gap in 2014, the year preceding the policy change. Furthermore, prior to 2015, the estimated gap size across ethnicity for troopers in Quartile 4 were never significantly different from each other, thereby appearing unbiased under the hit rate test. After 2015, the estimates across ethnicity for the same troopers are now significantly different from each other at the 95% level and would be considered biased under the hit rate test. Repeating this logic over the other quartiles of bias, troopers in the third quartile follow a similar pattern as troopers in the fourth quartile.

The reverse is true for troopers in the first and second quartile. Although these troopers were not misreporting to hide racial bias, it appears that the policy affected their search behavior as well with improved search success rates of Hispanic motorists for troopers in the first and second quartile. Prior to the policy change, these troopers appeared biased under the rate test with significant differences in search success rates across recorded Hispanic and White motorists. Notably, the gap is the widest in 2015 for troopers in Quartile 2, with the White search success rate of 52%, which is 16 percentage points greater than the Hispanic search success rate at 36%. After the policy change, the gap shrinks to 11 percentage points, and is not not significant for troopers in either Quartile 1 or Quartile 2. This pattern also supports the interpretation and finding of my theoretical model of racial bias in highway searches that non-misreporting troopers were less biased.

5.6 Bias and Trooper Characteristics

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

First, I test if employment outcomes, such as salary and experience, and trooper demographic characteristics, such as race and sex, are correlated to bias where experience is measured using the number of years employed by 2015. Table 7 shows the correlation between officer-level estimates of bias, normalized, with employment characteristics using employment information from 2010–2015 with unshrunk and shrunk estimates in columns (1) and (2), respectively.

Of the measurable trooper characteristics, I find that ethnicity is significantly correlated to bias. Specifically, Hispanic troopers have 0.15 (0.17) lower standard deviations of estimated Hispanic (shrunken) bias. This result aligns with past findings by Antonovics and Knight (2009), which also finds reduced bias with non-White police officers. I do not find any significant differences in bias between Black or Asian troopers relative to their White peers. But, as shown in Table 2, Black and Asian troopers make up approximately 10% of the force, thus I may be underpowered to detect any significant differences. I also find that male troopers are significant more biased than female troopers with 0.25 standard deviations higher bias.

For trooper rank, increasing in rank has no significant difference in bias compared to trooper rank except for higher ranked troopers. Troopers of lieutenant rank and higher are have 0.176 significantly higher bias than troopers, but the estimate attenuates to 0.17 and is insignificant when using the shrunken estimates of bias, which is unsurprising given how few searches ranked officers conduct. For all other ranks, probationary troopers and corporal, I do not find any significant differences in estimated bias when 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 for the following reasons. First, DPS does not have trooper employment data available prior to 2013 so 2013 is the earliest possible year. Second, combining the years increases the number of searches used to measure bias, which increases the precision of the estimates of bias. Third, with the panel-like structure, I can test if changes in employment outcomes are related to bias, specifically outcomes such as increasing in rank, moving cities, and leaving the force. If bias in 2013 has no effect on employment outcomes from 2014 to 2015, this could imply that misreporting is effective in making biased troopers appear unbiased thereby avoiding punishment of bias. I again apply Morris (1983) Bayes' shrinkage to the estimates of bias. I measure labor outcomes on the other quartiles of bias relative to the second quartile of bias since those troopers have near-zero levels of estimated bias.

As shown in Figure 6, with estimates using unshrunk estimates in panel A and shrunken estimates in panel B, I find that Hispanic bias from 2010–2013 has no impact on labor outcomes in 2014.[21] Not only are the point estimates insignificant with

[21] See Tables A3 and A4 for estimates in tabular form.

large standard errors, but the estimates are also close to zero. I also do not observe any significant differences in estimates across the different quartiles of bias relative to Quartile 2 of bias. Overall, I find that bias has no impact on the likelihood of leaving the force, increasing in rank, or changes in salary. Thus, in the presence of misreporting, biased and unbiased troopers have similar labor outcomes. These null results also provides suggestive evidence that misreporting was not driven by laziness or lack of effort rather than racial bias. If troopers were misreporting out of laziness then these troopers would also likely be lazy in other aspects of their job and would face worse labor outcomes than their unbiased peers.

I next test if biased troopers also perform worst in other aspects of their job by using complaint data obtained from DPS. While misreporting may help troopers evade negative employment outcomes, drivers may find cause to report the trooper. The results in Table 6 show a positive relationship between trooper level bias and the probability of receiving a complaint. One standard deviation of bias is associated with a 25% higher likelihood in having a complaint filed against the trooper (relative to mean of 5.99%). This estimate is likely an underestimate of the actual association of bias and complaints since not sustained or unfounded complaints repressed the trooper's badge number. From the 1,873 complaints, only 334 included the trooper's badge number. These results are robust to using the shrunken estimates of bias shown in Columns (3) and (4).

Lastly, I test to see how the employment outcomes of troopers were affected by the publication of the article relative to their level of bias. With the article publication, misreporting became significantly more costly as DPS changed its race recording policy to explicitly reduce officer discretion in recording driver's race. Thus, troopers could no longer use misreporting to mask biased behavior, laying bare differences in search outcomes across race.

To test this, I use publicly available 2019 salary data published by the Texas Tribune and regress the impact of estimated Hispanic bias using stops from 2010–2015 on labor outcomes in 2019. My results in Figure 7 show the effect of Hispanic bias on 2019 labor outcomes using the unshrunk estimates and shrunk estimates in panels A and B, respectively.[22] Here I observe that troopers in the first and fourth quartile of bias

[22]See Tables A5 and A6 for estimates in tabular form.

have lower salary growth of approximately $100 compared to the troopers in the second quartile. One potential interpretation for these results is that misreporting behavior overall led to negative work outcomes once misreporting became significantly harder. I also find that both quartile one and four are approximately 0.05 percentage points less likely to rank up, but these estimates are not significant. This pattern attenuates when using the shrunken estimates, especially for the troopers in the first quartile of bias.

I also find that only troopers in the 4th quartile are also 3 percentage points more likely to leave the force when using either the shrunken or unshrunk estimates. Although not significant, this does provide support to the possibility that troopers misreporting due to bias faced potentially worse labor outcomes than their peers.

Overall, I interpret this as suggestive evidence that misreporting reduced agency's ability to identify biased troopers. Once misreporting became significantly harder, I find modest evidence that misreporting troopers had lower salary growth relative to their honest peers. It appears that troopers who misreported overall, regardless of how they misreported, faced worse outcomes as I observe negative effects for both troopers in Quartile 1 and Quartile 4 of Hispanic bias.

6 Conclusion

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

One positive outcome of the misreporting is the public response and ability to change DPS' policies. Notably, from the time of the misreporting publication to the race recording rule change was only 15 days. Furthermore, once misreporting became significantly harder, I find suggestive evidence that misreporting troopers were less likely to be promoted. Additionally, in the first year following the policy, all troopers improved their search behavior, regardless of their estimated levels of bias.

While this improvement was short-lived, with search success rates returning to their pre-policy levels, the policy was effective in curbing misreporting. While the agency was unclear about how they would enforce such a policy, it appears that troopers adhered to more honest reporting, with the proportion of Hispanic motorists within searches remaining steady at 0.4. Thus, the adopted policy may be an effective way of curbing misreporting in other policing agencies.

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

References

Aaronson, Daniel, Lisa Barrow, and William Sanders. 2007. Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics* 25:95–135.

Alesina, Alberto, and Eliana La Ferrara. 2014. A test of racial bias in capital sentencing. *American Economic Review* 104 (11): 3397–433.

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

Antonovics, Kate, and Brian G. Knight. 2009. A new look at racial profiling: evidence from the Boston police department. *The Review of Economics and Statistics* 91 (1): 163–77.

Anwar, Shamena, and Hanming Fang. 2006. An alternative test of racial prejudice in motor vehicle searches: theory and evidence. *American Economic Review* 96 (1): 127–51.

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

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. Racial Bias in Bail Decisions*. *The Quarterly Journal of Economics* 133 (4): 1885–932. https://doi.org/10.1093/qje/qjy012.

Bakst, Brian. 2020. Legislature passes policing bill, ends special session [in en]. *MPR News,* accessed May 11, 2023. https://www.mprnews.org/story/2020/07/21/legislature-passes-policing-bill-ends-special-session.

Becker, Gary. 1957. *The economics of discrimination.* Edited by Chicago. University of Chicago Press.

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

———. 2015b. Texas troopers ticketing Hispanic drivers as white. July 21. https://www.kxan.com/inves troopers-ticketing-hispanic-drivers-as-white/, *KXAN.*

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

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

Depew, Briggs, Ozkan Eren, and Naci Mocan. 2017. Judges, juveniles, and in-group bias. *Journal of Law and Economics* 60 (2).

DPS. n.d. When stopped by law enforcement. Texas DPS.

Feigenberg, Benjamin, and Conrad Miller. 2021. Would Eliminating Racial Disparities in Motor Vehicle Searches have Efficiency Costs?*. *The Quarterly Journal of Economics* 137 (1): 49–113. https://doi.org/10.1093/qje/qjab018.

Fiscella, Kevin, and Allen M. Fremont. 2006. Use of Geocoding and Surname Analysis to Estimate Race and Ethnicity. *HSR: Health Services Research* 41 (4): 1482–500.

Freedman, Matthew, Emily Owens, and Sarah Bohn. 2018. Immigration, employment opportunities, and criminal behavior. *American Economic Journal: Economic Policy* 10 (2): 117–51.

Goncalves, Felipe, and Steven Mello. 2021. A Few Bad Apples? Racial Bias in Policing. *American Economic Review* 111 (5): 1406–41.

Grogger, Jeffrey, and Greg Ridgeway. 2006. Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness. 101 (475): 878–87.

Guarino, Cassandra M., Michelle Maxfield, Mark D. Reckase, Paul N. Thompson, and Jeffrey M. Wooldridge. 2015. An Evaluation of Empirical Bayes's Estimation of Value-Added Teacher Performance Measures [in en]. *Journal of Educational and Behavioral Statistics* 40 (2): 190–222. Accessed May 23, 2023. http://journals.sagepub.com/doi/10.3102/1076998615574771.

Hoekstra, Mark, and CarlyWill Sloan. 2022. Does Race Matter for Police Use of Force? Evidence from 911 Calls. *American Economic Review* 112 (3): 827–60. https://www.aeaweb.org/articles?id=10.1257/aer.20201292.

Jackson, C. Kirabo. 2018. What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. Publisher: The University of Chicago Press, *Journal of Political Economy* 126 (5): 2072–107. Accessed May 23, 2023. https://www.journals.uchicago.edu/doi/10.1086/699018.

Jacob, Brian A., and Steven D. Levitt. 2003. Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118 (3): 843–77.

Knowles, John, Nicola Persico, and Petra Todd. 2001. Racial bias in motor vehicle searches: theory and evidence. *Journal of Political Economy* 109 (1): 203–29.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. Value-added modeling: A review. *Economics of Education Review* 47:180–95. https://www.sciencedirect.com/science/article/pii/S0272775715000072.

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier. 2014. Promotional Reviews: An Empirical Investigation of Online Review Manipulation [in en]. *American Economic Review* 104 (8): 2421–55. Accessed May 19, 2023. https://www.aeaweb.org/articles?id=10.1257/aer.104.8.2421.

Morris, Carl N. 1983. Parametric Empirical Bayes Inference: Theory and Applications. Publisher: [American Statistical Association, Taylor & Francis, Ltd.] *Journal of the American Statistical Association* 78 (381): 47–55. Accessed November 10, 2022. https://www.jstor.org/stable/2287098.

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

Persico, Nicola, and Petra E. Todd. 2005. Passenger profiling, imperfect screening, and airport security. *American Economic Association Papers and Proceedings* 95 (2): 127–31.

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

Public Safety, Texas Department of. 2012. 2012 Traffic Stop Data Report. *Texas Department of Public Safety,* https://www.dps.texas.gov/sites/default/files/documents/director_staff/public_information/2012_traffic_stop_data_report.pdf.

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

Shayo, Moses, and Asaf Zussman. 2011. Judicial ingroup bias in the shadow of terrorism. *Quarterly Journal of Economics* 126 (3): 1447–84.

Simoiu, Camelia, Sam Corbett-Davies, and Sharad Goel. 2017. *The Problem of Infra-marginality in Outcome Tests for Discrimination.* ArXiv:1607.05376 [stat], June. Accessed November 17, 2022. `http://arxiv.org/abs/1607.05376`.

Sloan, CarlyWill. 2022. Racial bias by prosecutors: evidence from random assignment. January.

Tuttle, Cody. 2021. Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums. August.

Weisburst, Emily. 2022. Whose Help is on the Way? The Importance of Individual Police Officers in Law Enforcement Outcomes. *Journal of Human Resources.*

West, Jeremy. 2018. Racial bias in police investigations.

Tables

TABLE 1. Driver Summary Statistics

	(1) All Stops	(2) Searches	(3) Δ
Driver characteristics			
Recorded Asian	.017	.008	.009
	(.129)	(.09)	(0)
Recorded Black	.101	.174	-.073
	(.302)	(.379)	(.001)
Recorded Hispanic	.132	.163	-.031
	(.338)	(.369)	(.001)
Recorded Other Race	.064	.073	-.009
	(.245)	(.26)	(.001)
Recorded White	.686	.582	.104
	(.464)	(.493)	(.001)
Male	.677	.812	-.135
	(.468)	(.39)	(.001)
Vehicle and stop characteristics			
Luxury Car	.075	.071	.004
	(.263)	(.256)	(.001)
Vehicle Age > 5 years	.593	.748	-.155
	(.491)	(.434)	(.001)
Stop between 8 PM – 5 AM	.277	.399	-.122
	(.448)	(.49)	(.001)
N	11,897,213	218,813	

Notes: Unweighted means are shown. Standard deviations are in the parentheses for columns (1)–(2). Column (3) shows the difference between Columns (1) and (2) with the two-sample t-statistic in the parentheses. Sample is restricted to stops and searches by troopers with employment information (90% of the sample). All statistics are generated using information reported on the stop by the trooper. Estimates are generated using stops from 2010 to November, 2015.

TABLE 2. Trooper Summary Statistics

	(1) All Stops	(2) Searches	(3) Δ
Productivity, salary, and experience			
Total Stops	7242.643	7082.567	160.076
	(3453.623)	(3174.635)	(6.86)
Total Searches	130.036	328.904	-198.868
	(181.628)	(325.032)	(.697)
Year Hired	2005.3	2005.818	-.518
	(5.726)	(4.798)	(.01)
Monthly Salary	5455.014	5413.763	41.251
	(642.409)	(611.153)	(1.327)
Missing Salary	.013	.01	.003
	(.112)	(.102)	(0)
Demographic characteristics			
Native American	.01	.007	.003
	(.097)	(.084)	(0)
Asian	.011	.012	-.001
	(.104)	(.109)	(0)
Black	.081	.05	.031
	(.273)	(.218)	(0)
Hispanic	.26	.226	.034
	(.439)	(.418)	(.001)
White	.634	.703	-.069
	(.482)	(.457)	(.001)
Male	.962	.975	-.013
	(.19)	(.155)	(0)
Race/Sex Imputed	.047	.037	.01
	(.211)	(.188)	(0)
Rank information			
Trooper	.801	.795	.006
	(.4)	(.404)	(.001)
Probationary Trooper	.011	.01	.001
	(.104)	(.101)	(0)
Corporal	.12	.116	.004
	(.325)	(.321)	(.001)
Lieutenant+	.055	.068	-.013
	(.229)	(.252)	(.001)
Missing Rank	.014	.011	.003
	(.117)	(.106)	(0)
N	11,897,213	218,813	

Notes: Unweighted means are shown. Standard deviations are in the parentheses for columns (1)–(2). Column (3) shows the difference between Columns (1) and (2) with the two-sample t-statistic in the parentheses. Sample is restricted to stops and searches conducted from 2010–2015 by troopers with employment information (90% of searches). All statistics are generated using information reported on the stop by the trooper from stops conducted from 2010 to November 2015.

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

	(1) Hispanic	(2) Black	(3) Asian
I(Failure)	0.0226**	-0.000149	0.0156
	(0.00870)	(0.000732)	(0.0137)
Constant	0.257***	0.00214***	0.116***
	(0.00515)	(0.000365)	(0.00626)
Avg. Mismatch Rate	0.274	0.002	0.126
Observations	48780	37989	1454
r2	0.225	0.0622	0.646
F	6.729	0.0413	1.304

Notes: Dependent variable is an indicator variable equal to 1 if the recorded race does not equal to the estimated race, which I often refer to as mismatch. Standard errors are clustered at the county level. Regression uses all stops conducted from 2010 to November 2015. Each regression is run separately for motorists of each race, where race is identified using the estimated race. The regression includes fixed effects for hour of the stop, month of the stop, year of the stop, county fixed effects, and vehicle type and vehicle age. I also include the full interaction for hour of the stop with month of the stop and year of the stop and the full interaction of vehicle characteristics (vehicle type and vehicle year). Standard errors are clustered at the county FIPS and year level. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

TABLE 4. Quartiles of Estimated Hispanic Bias

	Unshrunk		Shrunken	
	(1)	(2)	(3)	(4)
	Min	Max	Min	Max
	Mean		Mean	
Quartile 1	-1.25	-0.080	-.218	.012
	-.286		-.031	
Quartile 2	-.079	.042	.012	.042
	-0.014		.028	
Quartile 3	.042	.180	.042	.076
	0.106		.057	
Quartile 4	.180	1.60	.076	.271
	.363		.112	

Notes: Table shows unweighted averages of estimated Hispanic bias when troopers are split into even quartiles. Columns (1)–(2) shows the statistics for unshrunken estimates of Hispanic bias while Columns (3)–(4) shows the statistics using the shrunken estimates detailed in Section 5.2.

TABLE 5. Estimated Hispanic Bias and Estimated Black Bias

	(1) Any Black Bias	(2) Black Bias
Hispanic Bias	0.153**	
	(0.0650)	
Average Black Mismatch rate		-0.882***
		(0.0696)
Constant	0.513***	0.00587***
	(0.0164)	(0.00125)
Observations	972	972
r2	0.00550	0.539
F	5.555	160.6

Notes: Dependent variable is in the column. Regressions in this table are restricted to troopers with unshrunk estimates of Black bias and unshrunk estimates of Hispanic bias. Estimates of Hispanic and Black bias are generated from β_1^j (Columns 1 and 2) and δ_j (average mismatch rate by race, Column 2) from Eq. 7, detailed in Section 5.2. All regressions are estimated with robust standard errors. Only troopers with non-missing estimates of Black and Hispanic bias are included in the regression. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

TABLE 6. Positive Hispanic Bias on Complaints

	(1)	(2)	(3)	(4)
	Unshrunk Bias		Shrunken Bias	
	Any Complaint	Sustained Complaint	Any Complaint	Sustained Complaint
Hispanic Bias	0.0149**	0.0150**		
	(0.00630)	(0.00619)		
Hispanic Bias			0.0150**	0.0151**
			(0.00617)	(0.00600)
Constant	0.0599***	0.0565***	0.0600***	0.0565***
	(0.00626)	(0.00609)	(0.00627)	(0.00609)
Observations	1433	1433	1433	1433
r2	0.00394	0.00424	0.00399	0.00427

Notes: Dependent variable is an indicator variable equal to one if the trooper had any complaints (column (1)) or a sustained complaint (column (2)) from 2010 to November 2015. Hispanic bias is normalized and is estimated from Equation (7), β_1^j. Regression has robust standard errors. Column (3) and (4) use the shrunken estimates of Hispanic bias outlined in Section 5.2. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

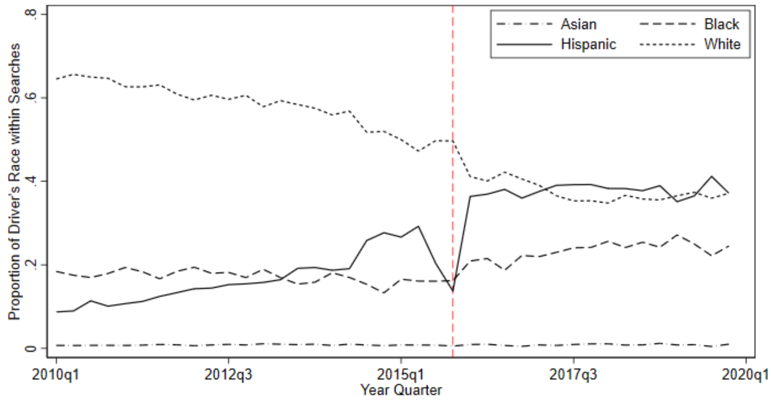
TABLE 7. Correlates of Hispanic Bias

	(1) Bias	(2) Shrunken Bias
Experience	-0.00592	-0.00630
	(0.00600)	(0.00581)
Native American	0.0631	-0.0186
	(0.280)	(0.247)
Asian	-0.00258	0.00741
	(0.195)	(0.248)
Black	0.0166	0.0201
	(0.123)	(0.113)
Hispanic	-0.147**	-0.169***
	(0.0600)	(0.0621)
Probationary Trooper	0.0697	0.0692
	(0.159)	(0.142)
Corporal	-0.111	-0.0882
	(0.0868)	(0.0889)
Lieutenant+	0.176*	0.171
	(0.104)	(0.106)
I(Male)	0.250*	0.230*
	(0.130)	(0.124)
Constant	-0.150	-0.121
	(0.137)	(0.131)
Observations	1394	1394
r2	0.0135	0.0137
F	2.009	2.042

Notes: Regression includes controls for the work city and is clustered at the work city level. Dependent variable is the officer level measure of bias from Equation (7). Troopers with rank equal to or higher than Lieutenant (sergeant, major, captain) were grouped into "Lieutenant +". Salary is monthly salary measured in thousands of dollars. Only troopers with rank and salary information are included in the regression. Column (1) shows the statistics for unshrunken estimates of Hispanic bias while Column (2) shows the estimates using the shrunken estimates of Hispanic bias detailed in Section 5.2. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

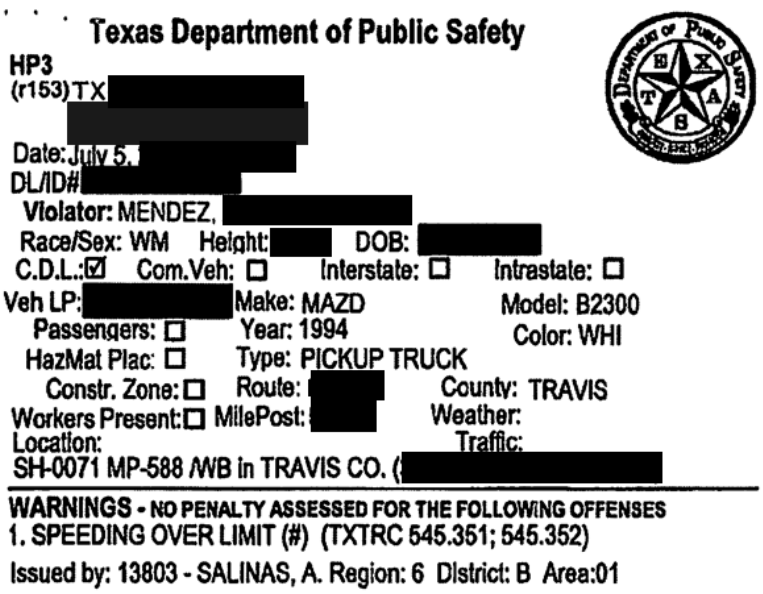
Figures

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



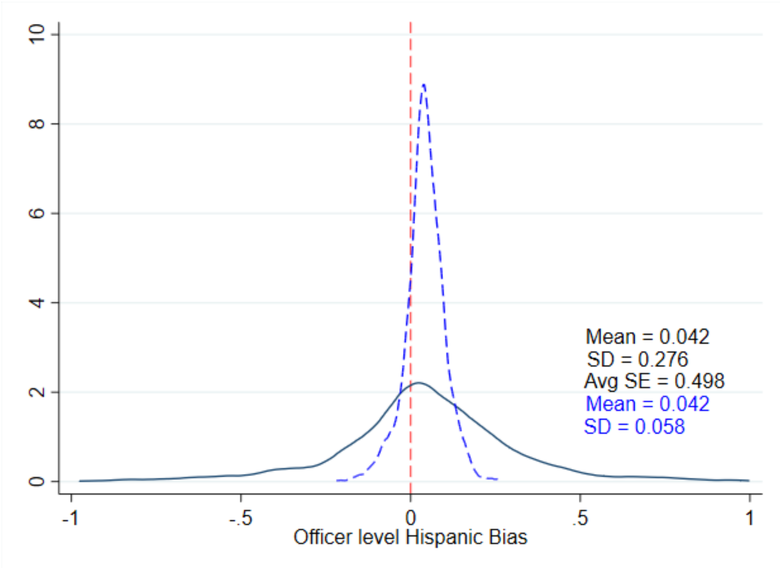
Notes: Average, unweighted search rates, by reported motorist race, for a given quarter-year from January 2010 to December 2015 are shown. The dashed red, vertical line indicates the quarter when KXAN published the article revealing the trooper race misreporting.

Figure 2. Example of a Misreported Highway Ticket



Texas Department of Public Safety
HP3
(r153) TX[redacted]
Date: July 5, [redacted]
DL/ID# [redacted]
Violator: MENDEZ, [redacted]
Race/Sex: WM Height: [redacted] DOB: [redacted]
C.D.L.: ☑ Com.Veh: ☐ Interstate: ☐ Intrastate: ☐
Veh LP: [redacted] Make: MAZD Model: B2300
Passengers: ☐ Year: 1994 Color: WHI
HazMat Plac: ☐ Type: PICKUP TRUCK
Constr. Zone: ☐ Route: [redacted] County: TRAVIS
Workers Present: ☐ MilePost: [redacted] Weather:
Location: Traffic:
SH-0071 MP-588 WB in TRAVIS CO. ([redacted]
WARNINGS - NO PENALTY ASSESSED FOR THE FOLLOWING OFFENSES
1. SPEEDING OVER LIMIT (#) (TXTRC 545.351; 545.352)
Issued by: 13803 - SALINAS, A. Region: 6 District: B Area:01

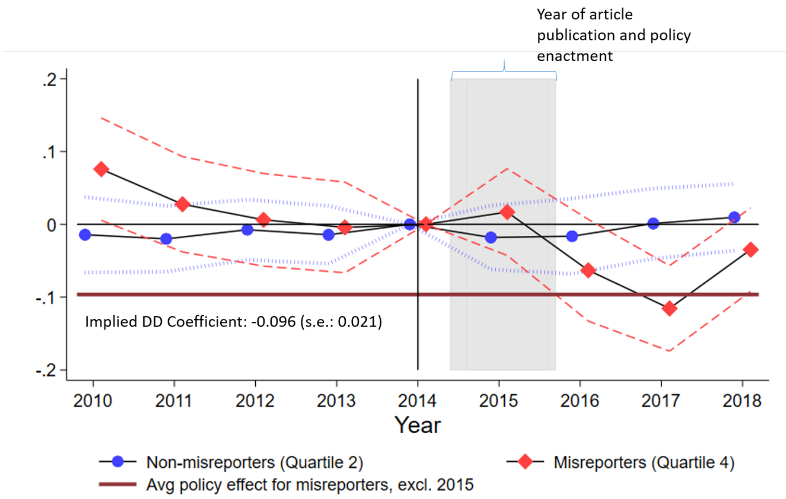
Figure 3. Trooper level estimates of bias



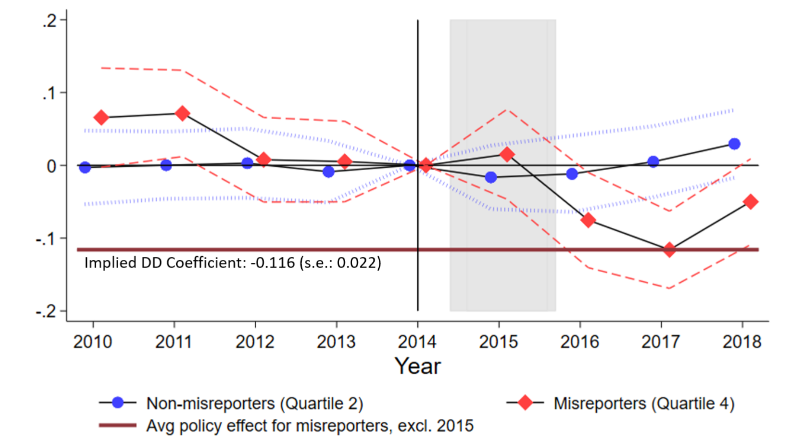
Notes: Kernel density distribution of officer-level estimates of Hispanic bias. The figure plots each officer's β_1^j from the regression $I(Mismatch_{i,c,j,t}) = \beta_0 + \beta_1^j I(Failure_{i,c,j,t}) + \delta_j + X_{i,c}\gamma + \alpha_t + \epsilon_{i,c,j,t}$. The solid, black line is using the original estimates; the dashed, blue line is the Bayes shrinkage procedure (Morris 1983).

Figure 4. Change in the Hispanic-White search success rate gap due to the 2015 rule change for biased and unbiased troopers

Panel A: Unshrunk estimates

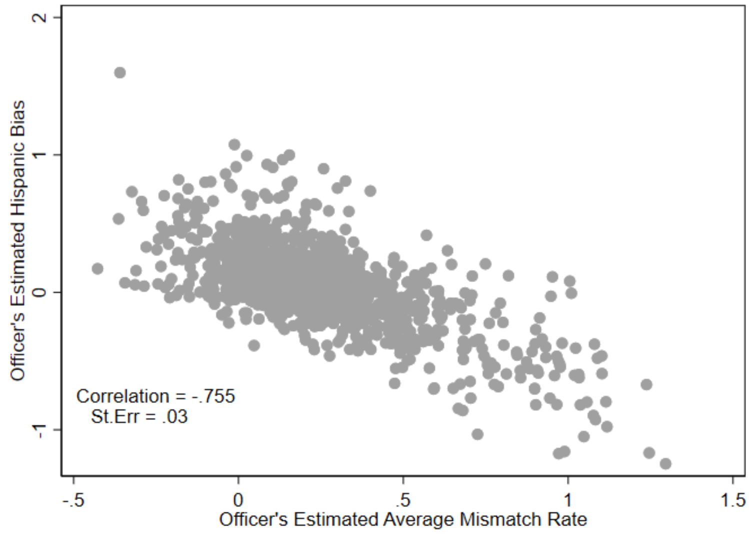


Panel B: Shrunken estimates



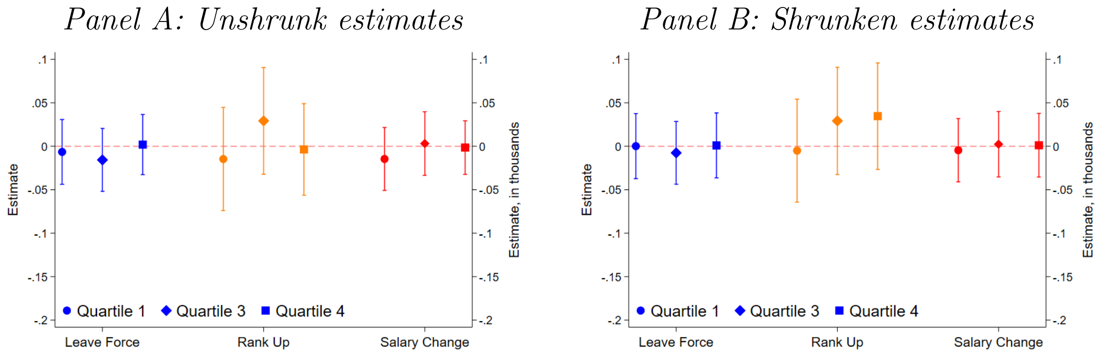
Notes: The figure plots each year's β_3^t from the Eq. 8 for biased troopers (red) and unbiased troopers (blue). Biased troopers are in the 4th quartile of the bias distribution while unbiased troopers are in the 2nd quartile of the bias distribution, using Table 4. Standard errors are clustered county level. 2014, the year prior to the policy change and article publication, is the comparison year. Shaded region denotes the article publication and policy change year, 2015. The thick, maroon, y-intercept indicates the average post-policy effect for misreporters compared to non-misreporters, excluding 2015. Appendix Figure A.10 shows the results for 'negatively' biased troopers (Q1) and 'semi'-biased troopers (Q3). Table A1 shows the results in tabular format.

Figure 5. Correlation of trooper's level of bias with trooper's average race mismatch rate



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

Figure 6. Impact of Hispanic bias on 2014 labor outcomes, by quartiles of bias

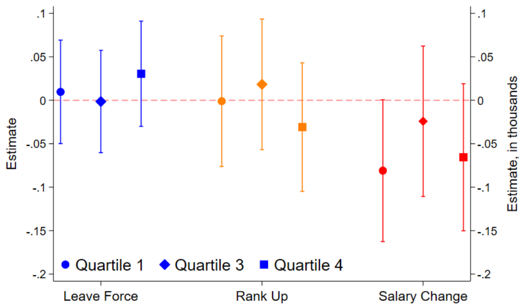


Notes: The figure plots the estimate of each quartile of bias, estimated using stops from 2010–2013, on work outcomes relative to Quartile 2 (no bias) with robust standard errors. 'Leave Force' is an indicator variable equal to one if trooper is no longer employed as a trooper in DPS by 2019. 'Rank up' is defined as increasing in rank in 2014 relative to maximum rank observed from 2010–2013. Salary change is unadjusted and is measured in thousands. For all regressions, controls for years on the force, fixed effects for race and ethnicity, fixed effects for maximum rank in 2010–2013, and sex are included. See Tables A3 and A4 for estimates in tabular form.

Figure 7. Impact of Hispanic bias on 2019 labor outcomes, by quartiles of bias

Panel A: Unshrunk estimates

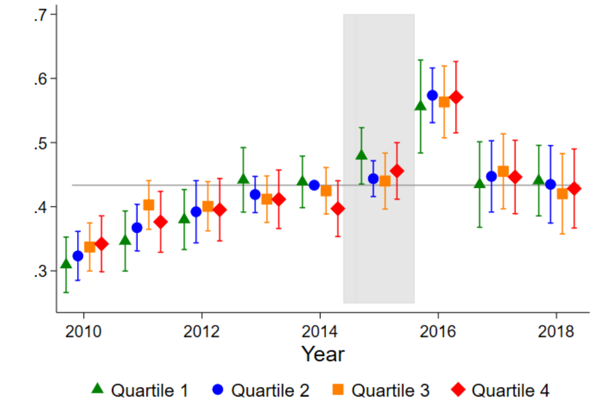
Panel B: Shrunken estimates



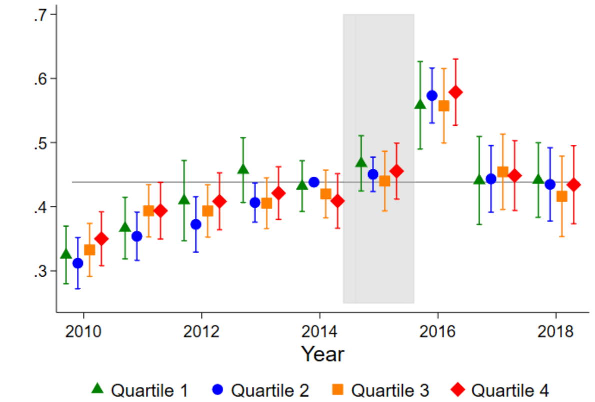
Notes: The figure plots the estimate of each quartile of bias, estimated using stops from 2010–2015, on work outcomes relative to Quartile 2 (no bias) with robust standard errors. 'Leave Force' is an indicator variable equal to one if trooper is no longer employed as a trooper in DPS by 2019. 'Rank up' is defined as increasing in rank in 2019 relative to maximum rank observed from 2013–2015. Salary change is unadjusted and is measured in thousands. For all regressions, controls for years on the force, fixed effects for race and ethnicity, fixed effects for maximum rank in 2010–2015, and sex are included. See Tables A5 and A6 for results in tabular format.

Figure 8. Impact of policy change on overall search success rates, by quartiles of bias

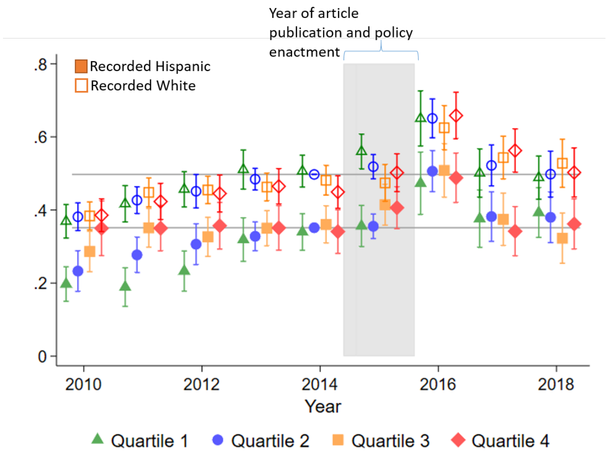
Panel A: Unshrunk estimates for recorded Hispanic and White motorists, combined



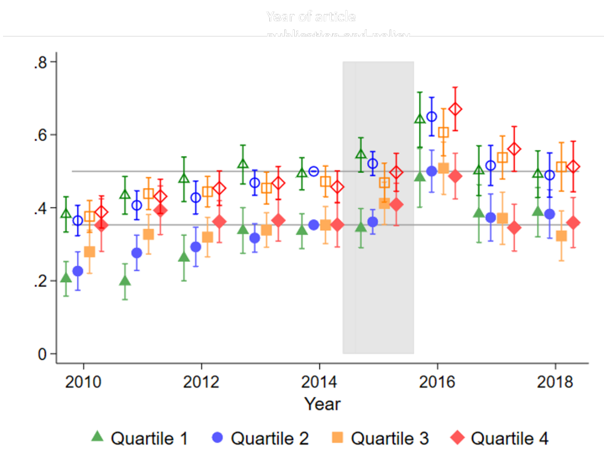
Panel B: Shrunken estimates for recorded Hispanic and White motorists, combined



Panel C: Unshrunk estimates for recorded Hispanic and White motorists, separated



Panel D: Shrunken estimates for recorded Hispanic and White motorists, separated



Notes: The figure plots the coefficient estimates from Eq. 9 adjusted using the average of the excluded group. The excluded group are searches conducted in 2014 by the second quartile of bias. Standard errors are clusted at the county level. The grey bar indicates 2015, the year the policy was enacted.

A Appendix

A.1 Discussion of the Model

Proof of Proposition 1

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

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

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

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

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

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

Proof of Proposition 2

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

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

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

Relaxing Assumption 3

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

$$\phi_{M,t} = \frac{\pi_M[F_g^M(\theta_{M,t}^\mu) - F_G^M(\theta_{M,t}^*)] + (1-\pi_M)[F_n^M(\theta_{M,t}^\mu) - F_n^M(\theta_{M,t}^*)]}{\pi_M[1 - F_g^M(\theta_{M,t}^*)] + (1-\pi_M)(1 - F_n^M(\theta_{M,t}^*)]} \tag{10}$$

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:

$$\gamma_{m,t} = \pi_m[1 - F_g^m(\theta_{m,t}^*)] + (1-\pi_m)[1 - F_n^m(\theta_{m,t}^*)] \tag{11}$$

Let $\gamma_{m,t}^O$ 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:

$$\gamma_{M,t}^O = \pi_M[1 - F_g^M(\theta_{M,t}^\mu)] + (1-\pi_M)[1 - F_n^M(\theta_{M,t}^\mu)] \tag{12}$$

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:

$$\begin{aligned} \gamma_{W,t}^O &= \pi_W[1 - F_g^W(\theta_{W,t}^*)] + (1-\pi_W)(1 - F_n^W(\theta_{W,t}^*)] \\ &+\pi_M[F_g^M(\theta_{M,t}^\mu) - F_G^M(\theta_{M,t}^*)] + (1-\pi_M)[F_n^M(\theta_{M,t}^\mu) - F_n^M(\theta_{M,t}^*)] \end{aligned} \tag{13}$$

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

$$\phi_{M,t} = \frac{\gamma^{O}_{W,t} - \gamma_{W,t}}{\gamma_{M,t}} \tag{14}$$

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

Proof of Proposition 2 and Corollary 3

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

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

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

A.2 Appendix tables

TABLE A1. Change in the Hispanic-White search success rate gap due to the 2015 rule change for biased and unbiased troopers

	Unshrunk bias estimates		Shrunk bias estimates	
	No bias (Quartile 2)	Biased (Quartile 4)	No bias (Quartile 2)	Biased (Quartile 4)
	(1)	(2)	(3)	(4)
$I(Year = 2010)\times$	-0.014	0.076**	-0.003	0.066*
$I(Race_i^{Recorded} = Hispanic)$	(0.026)	(0.036)	(0.026)	(0.035)
$I(Year = 2011)\times$	-0.020	0.028	0.000	0.071**
$I(Race_i^{Recorded} = Hispanic)$	(0.023)	(0.033)	(0.024)	(0.030)
$I(Year = 2012)\times$	-0.007	0.006	0.003	0.008
$I(Race_i^{Recorded} = Hispanic)$	(0.021)	(0.032)	(0.024)	(0.030)
$I(Year = 2013)\times$	-0.014	-0.004	-0.009	0.005
$I(Race_i^{Recorded} = Hispanic)$	(0.020)	(0.032)	(0.022)	(0.028)
$I(Year = 2015)\times$	-0.018	0.017	-0.016	0.015
$I(Race_i^{Recorded} = Hispanic)$	(0.022)	(0.030)	(0.022)	(0.032)
$I(Year = 2016)\times$	-0.016	-0.063*	-0.012	-0.075**
$I(Race_i^{Recorded} = Hispanic)$	(0.026)	(0.036)	(0.027)	(0.033)
$I(Year = 2017)\times$	0.001	-0.115***	0.005	-0.116***
$I(Race_i^{Recorded} = Hispanic)$	(0.025)	(0.030)	(0.025)	(0.027)
$I(Year = 2018)\times$	0.010	-0.035	0.030	-0.050*
$I(Race_i^{Recorded} = Hispanic)$	(0.023)	(0.029)	(0.024)	(0.030)
Observations	81839	53646	76363	58029

Notes: This table displays the estimates of each year's β_3^t from the Eq. 8 for unbiased troopers (columns 1 and 3) and biased troopers (columns 2 and 4) corresponding to Figure 4. Regression estimates using unshrunk estimates of bias are in columns 1 and 2 and using shrunken estimates of bias in columns 3 and 4. Biased troopers are in the 4th quartile of the bias distribution while unbiased troopers are in the 2nd quartile of the bias distribution, using Table 4. Standard errors are clustered county level. 2014, the year prior to the policy change and article publication, is excluded since it's the comparison year. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

TABLE A2. Change in the Hispanic-White search success rate gap due to the 2015 rule change for 'negatively' and 'semi-biased' troopers

	Unshrunk bias estimates		Shrunk bias estimates	
	Neg. bias (Quartile 1)	Semi-biased (Quartile 3)	Neg bias (Quartile 1)	Semi-biased (Quartile 3)
	(1)	(2)	(3)	(4)
$I(Year = 2010)\times$	-0.007	0.002	-0.016	-0.007
$I(Race_i^{Recorded} = Hispanic)$	(0.027)	(0.024)	(0.025)	(0.024)
$I(Year = 2011)\times$	-0.041	0.006	-0.059**	-0.018
$I(Race_i^{Recorded} = Hispanic)$	(0.026)	(0.021)	(0.022)	(0.022)
$I(Year = 2012)\times$	-0.037*	-0.016	-0.047**	-0.017
$I(Race_i^{Recorded} = Hispanic)$	(0.022)	(0.022)	(0.021)	(0.022)
$I(Year = 2013)\times$	-0.004	0.009	-0.010	0.004
$I(Race_i^{Recorded} = Hispanic)$	(0.026)	(0.023)	(0.027)	(0.023)
$I(Year = 2015)\times$	-0.021	0.042*	-0.032	0.043*
$I(Race_i^{Recorded} = Hispanic)$	(0.026)	(0.025)	(0.025)	(0.025)
$I(Year = 2016)\times$	-0.002	0.001	-0.010	0.013
$I(Race_i^{Recorded} = Hispanic)$	(0.035)	(0.027)	(0.035)	(0.027)
$I(Year = 2017)\times$	0.035	-0.037	0.021	-0.034
$I(Race_i^{Recorded} = Hispanic)$	(0.022)	(0.026)	(0.023)	(0.027)
$I(Year = 2018)\times$	0.077**	-0.074**	0.043	-0.058*
$I(Race_i^{Recorded} = Hispanic)$	(0.033)	(0.034)	(0.027)	(0.033)
Observations	54124	78146	59600	73760

Notes: This table displays the estimates of each year's β_3^t from the Eq. 8 for 'negatively' biased troopers (columns 1 and 3) and 'semi-biased' troopers (columns 2 and 4) corresponding to Figure A.10. Regression estimates using unshrunk estimates of bias are in columns 1 and 2 and using shrunken estimates of bias in columns 3 and 4. Biased troopers are in the 4th quartile of the bias distribution while unbiased troopers are in the 2nd quartile of the bias distribution, using Table 4. Standard errors are clustered county level. 2014, the year prior to the policy change and article publication, is excluded since it's the comparison year. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

TABLE A3. Unshrunk Hispanic Bias on Labor Outcomes - Panel Results

	(1) Impute Left	(2) Recorded Left	(3) Fired	(4) Salary Difference	(5) Ranked Up
1st Quartile Bias	0.00376	-0.00654	-0.228	-0.0146	-0.0147
	(0.00364)	(0.0190)	(0.221)	(0.0185)	(0.0303)
3rd Quartile Bias	0.00348	-0.0157	0.123	0.00311	0.0292
	(0.00373)	(0.0185)	(0.240)	(0.0187)	(0.0313)
4th Quartile Bias	0.00455	0.00195	-0.190	-0.00152	-0.00371
	(0.00309)	(0.0176)	(0.177)	(0.0157)	(0.0269)
I(Male)	0.00451*	-0.0390	0.122	-0.0343	-0.107*
	(0.00244)	(0.0396)	(0.214)	(0.0292)	(0.0604)
Native American	-0.00264	0.154	0.296	-0.133*	-0.0625
	(0.00244)	(0.126)	(0.410)	(0.0725)	(0.107)
Asian	-0.00318	0.0165	-0.0412	-0.0136	-0.0186
	(0.00206)	(0.0670)	(0.210)	(0.0529)	(0.0715)
Black	-0.00188	-0.0340	0.218	0.0100	0.0534
	(0.00171)	(0.0218)	(0.436)	(0.0207)	(0.0437)
Hispanic	0.00321	-0.00936	0.200	-0.00705	0.0467**
	(0.00411)	(0.0129)	(0.171)	(0.0138)	(0.0236)
Probationary Trooper	0.00133				0.811***
	(0.00167)				(0.0749)
Corporal	0.00447	-0.0187	-0.0894	0.0182	0.268***
	(0.00714)	(0.0182)	(0.220)	(0.0140)	(0.0390)
Lieutenant+	-0.00280	-0.0482***	-0.0601	0.342***	0.405***
	(0.00185)	(0.0144)	(0.542)	(0.0310)	(0.0430)
Experience	-0.000394	0.00400**	-0.00805	0.00476***	-0.0166***
	(0.000314)	(0.00159)	(0.00776)	(0.00121)	(0.00224)
Constant	-0.00148	0.0607	0.368	0.821***	0.364***
	(0.00243)	(0.0443)	(0.237)	(0.0331)	(0.0675)
Observations	1286	1237	61	1211	1282
r2	0.00408	0.0192	0.128	0.220	0.159

Notes: Regression has robust standard errors. Dependent variable is the officer level measure of bias from Equation (7) using only stops from 2010 to 2013. Employment outcomes are from 2014–November 2015. Each regression includes controls for the trooper's gender, trooper's rank in 2013, and trooper race. Hispanic bias is unshrunk and is split into 4 even quartiles. Impute left is defined as not observing any searches in the 2014–2015 traffic stop data. Salary is measured in thousands, thus Salary Difference = 1 indicates an increase in $1,000. Troopers are weighted by the their total number of searches of estimated Hispanic motorists conducted from 2010–2013. Employment data is missing rank and salary for some observations leading to differences in number of observations. Fired (column (3)) is restricted to troopers whose separation was recorded in the employment data. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

TABLE A4. Shrunken Hispanic Bias on Labor Outcomes - Panel Results

	(1) Impute Left	(2) Recorded Left	(3) Fired	(4) Salary Difference	(5) Ranked Up
1st Quartile Bias	0.00342	0.000155	-0.0878	-0.00456	-0.00507
	(0.00368)	(0.0191)	(0.238)	(0.0186)	(0.0302)
3rd Quartile Bias	0.00325	-0.00761	0.0781	0.00235	0.0292
	(0.00381)	(0.0184)	(0.245)	(0.0192)	(0.0315)
4th Quartile Bias	0.00329	0.00101	-0.105	0.00119	0.0346
	(0.00349)	(0.0191)	(0.226)	(0.0187)	(0.0313)
I(Male)	0.00444	-0.0380	0.124	-0.0140	-0.146**
	(0.00285)	(0.0443)	(0.318)	(0.0357)	(0.0716)
Native American	-0.00163	0.102	-0.356*	-0.185*	-0.0512
	(0.00301)	(0.133)	(0.204)	(0.0965)	(0.155)
Asian	-0.00225	-0.0494***		-0.0395	0.0294
	(0.00237)	(0.00999)		(0.0618)	(0.0823)
Black	0.0000718	-0.0407**	-0.339	0.0148	0.0420
	(0.00111)	(0.0196)	(0.230)	(0.0231)	(0.0495)
Hispanic	0.00470	-0.00247	0.111	-0.00584	0.0506**
	(0.00448)	(0.0145)	(0.188)	(0.0153)	(0.0258)
Probationary Trooper	-0.000432				0.774***
	(0.00144)				(0.0844)
Corporal	0.00685	-0.0140	-0.0912	0.0296*	0.286***
	(0.00843)	(0.0201)	(0.239)	(0.0161)	(0.0426)
Lieutenant+	-0.00205	-0.0381**	-0.00350	0.366***	0.466***
	(0.00169)	(0.0169)	(0.512)	(0.0364)	(0.0486)
Experience	-0.000624*	0.00356**	-0.00745	0.00488***	-0.0171***
	(0.000374)	(0.00168)	(0.0123)	(0.00144)	(0.00253)
Constant	0.000115	0.0585	0.372	0.791***	0.394***
	(0.00227)	(0.0489)	(0.316)	(0.0396)	(0.0771)
Observations	1063	1019	48	1001	1060
r2	0.00758	0.0131	0.0661	0.223	0.183
F	.	3.395	.	11.67	.

Notes: Regression has robust standard errors. Dependent variable is the officer level measure of bias from Equation (7) using only stops from 2010 to 2013. Employment outcomes are from 2014–November 2015. Each regression includes controls for the trooper's gender, trooper's rank in 2013, and trooper race. Hispanic bias is split into 4 even quartiles. Impute left is defined as not observing any searches in the 2014–2015 traffic stop data. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in $1,000. Employment data is missing rank and salary for some observations leading to differences in number of observations. Fired (column (3)) is restricted to troopers whose separation was recorded in the employment data. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

TABLE A5. Unshrunk Hispanic Bias on 2019 Labor Outcomes

	(1) Left Force	(2) Salary Change	(3) Ranked Up
1st Quartile Bias	-0.0112	-0.102**	-0.0583
	(0.0304)	(0.0415)	(0.0383)
3rd Quartile Bias	-0.0225	-0.0273	-0.0123
	(0.0301)	(0.0445)	(0.0392)
4th Quartile Bias	0.0290	-0.0891**	-0.0536
	(0.0316)	(0.0438)	(0.0386)
Experience	0.0140***	-0.0378***	-0.00536**
	(0.00221)	(0.00256)	(0.00263)
Native American	0.0945	0.0375	-0.182
	(0.130)	(0.214)	(0.119)
Asian	0.0214	-0.222**	-0.127
	(0.112)	(0.0982)	(0.121)
Black	0.0418	-0.0290	-0.0461
	(0.0520)	(0.0525)	(0.0582)
Hispanic	0.0474*	0.0237	0.0330
	(0.0248)	(0.0369)	(0.0310)
Probationary Trooper	0.0741	0.659***	0.655***
	(0.0598)	(0.0759)	(0.0260)
Corporal	-0.105***	0.0665	0.0681
	(0.0338)	(0.0455)	(0.0475)
Lieutenant+	-0.0549	-0.136**	-0.147***
	(0.0406)	(0.0571)	(0.0422)
I(Male)	-0.0494	-0.0763	0.0477
	(0.0570)	(0.0633)	(0.0629)
Constant	0.132**	1.215***	0.327***
	(0.0632)	(0.0718)	(0.0718)
Observations	1394	1100	1100
r2	0.0401	0.224	0.0966
F	4.551	38.89	210.9

Notes: Regression has robust standard errors shown in parentheses and uses 2019 employment data posted publicly by the Texas Tribune. Each regression controls for the trooper's gender, trooper's maximum rank from 2013 to 2015, and trooper race. Hispanic bias is normalized, unshrunk, estimated from Equation (7), β_1^j. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. Regressions in columns (2) and (3) are restricted to being observed in the 2019 employment data. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

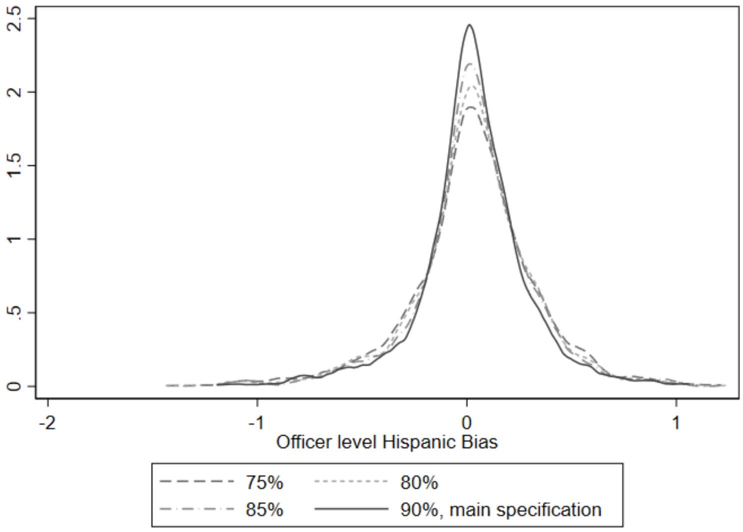
TABLE A6. Shrunken Hispanic Bias on 2019 Labor Outcomes

	(1) Left Force	(2) Salary Change	(3) Ranked Up
1st Quartile Bias	0.0145	-0.0959**	-0.00722
	(0.0304)	(0.0414)	(0.0383)
3rd Quartile Bias	0.000541	-0.0393	0.0114
	(0.0300)	(0.0447)	(0.0384)
4th Quartile Bias	0.0318	-0.0699	-0.0248
	(0.0308)	(0.0433)	(0.0381)
Experience	0.0141***	-0.0380***	-0.00532**
	(0.00220)	(0.00255)	(0.00263)
Native American	0.0957	0.249	-0.0674
	(0.129)	(0.233)	(0.167)
Asian	0.0235	0	0
	(0.112)	(.)	(.)
Black	0.0426	0.173	0.0716
	(0.0515)	(0.110)	(0.131)
Hispanic	0.0456*	0.240**	0.153
	(0.0249)	(0.103)	(0.121)
Probationary Trooper	0.0785	0.649***	0.653***
	(0.0597)	(0.0770)	(0.0258)
Corporal	-0.108***	0.0660	0.0666
	(0.0338)	(0.0456)	(0.0474)
Lieutenant+	-0.0554	-0.144**	-0.152***
	(0.0406)	(0.0571)	(0.0420)
I(Male)	-0.0501		0.0524
	(0.0572)		(0.0629)
Constant	0.119*	0.930***	0.177
	(0.0630)	(0.102)	(0.131)
Observations	1394	1100	1100
r2	0.0389	0.221	0.0945
F	4.335	40.67	216.6

Notes: Regression has robust standard errors show in parentheses and uses 2019 employment data posted publicly by the Texas Tribune. Includes controls for the trooper's gender, trooper's maximum rank from 2010 to 2015, and trooper race. Black and Hispanic are indicator variables equal to one if the trooper is Black or Hispanic, respectively, and equal to zero otherwise. Hispanic bias is normalized. Salary is measured in thousands, thus $\Delta Salary = 1$ indicates an increase in \$1,000. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

A.3 Appendix Figures

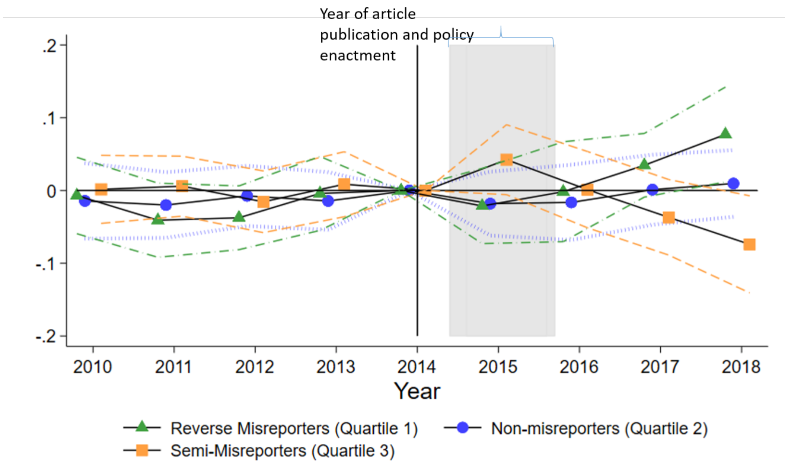
Figure A.9. Officer-level estimates of Hispanic bias with different thresholds



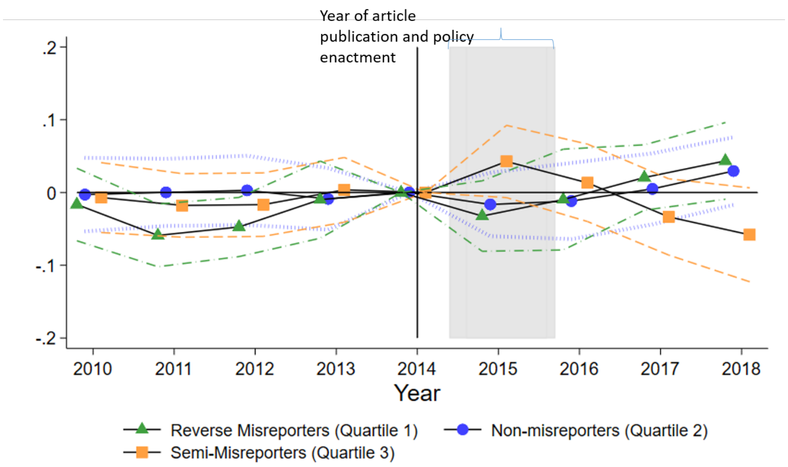
Notes: Each density shows the unshrunk, officer level bias using different levels of surname cutoff and ZIP code density cutoff (surname threshold - 15%). The estimate of bias is from each officer's β^j from Eq. 7. Bias is estimated using searches conducted from 2010 to November 2015. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

Figure A.10. Change in the Hispanic-White search success rate gap due to the 2015 rule change for biased, unbiased, and 'negatively' biased troopers

Panel A: Unshrunk estimates



Panel B: Shrunken estimates



Notes: The figure plots each year's β_3^t from the Eq. 8 for semi-biased troopers (purple), unbiased troopers (blue), and 'negatively' biased troopers (green). Semi-biased troopers are in the 3rd quartile of the bias distribution, unbiased troopers are in the 2nd quartile of the bias distribution, and 'negatively' biased troopers are in the 1st quartile, using Table 4. Standard errors are clustered county level. 2014, the year prior to the policy change and article publication, is the comparison year. Shaded region denotes the article publication and policy change year, 2015. Table A2 shows the results in tabular format.