An Empirical Analysis of Racial Differences in Police Use of Force*

Roland G. Fryer, Jr.†

Draft: July 2016

Abstract

This paper explores racial differences in police use of force. On non-lethal uses of force, blacks and Hispanics are more than fifty percent more likely to experience some form of force in interactions with police. Adding controls that account for important context and civilian behavior reduces, but cannot fully explain, these disparities. On the most extreme use of force – officer-involved shootings – we find no racial differences in either the raw data or when contextual factors are taken into account. We argue that the patterns in the data are consistent with a model in which police officers are utility maximizers, a fraction of which have a preference for discrimination, who incur relatively high expected costs of officer-involved shootings.

Keywords: discrimination, decision making, bias, police use of force

*This work has benefitted greatly from discussions and debate with Chief William Evans, Chief Charles McClelland, Chief Martha Montalvo, Sergeant Stephen Morrison, Jon Murad, Lynn Overmann, Chief Bud Riley, and Chief Scott Thomson. I am grateful to David Card, Kerwin Charles, Christian Dustmann, Michael Greenstone, James Heckman, Richard Holden, Lawrence Katz, Steven Levitt, Jens Ludwig, Glenn Loury, Kevin Murphy, Derek Neal, John Overdeck, Jesse Shapiro, Andrei Shleifer, Jorg Spenkuch, Max Stone, John Van Reenan, Christopher Winship, and seminar participants at Brown University, University of Chicago, London School of Economics, University College London, and the NBER Summer Institute for helpful comments and suggestions. Brad Allan, Elijah De La Campa, Tanaya Devi, and William Murdock III provided truly phenomenal project management and research assistance. Lukas Althoff, Dhruva Bhat, Samarth Gupta, Julia Lu, Mehak Malik, Beatrice Masters, Ezinne Nwankwo, Charles Adam Pfander, Sofya Shchukina and Eric Yang provided excellent research assistance. Financial support from EdLabs Advisory Group and an anonymous donor is gratefully acknowledged. Correspondence can be addressed to the author by email at rolandfryer@edlabs.harvard.edu. The usual caveat applies.

†Department of Economics, Harvard University, and the NBER, (rfryer@fas.harvard.edu);

"We can never be satisfied as long as the Negro is the victim of the unspeakable horrors of police brutality." Martin Luther King, Jr., August 28, 1963.

I. Introduction

From "Bloody Sunday" on the Edmund Pettus Bridge to the public beatings of Rodney King, Bryant Allen, and Freddie Helms, the relationship between African-Americans and police has an unlovely history. The images of law enforcement clad in Ku Klux Klan regalia or those peaceful protesters being attacked by canines, high pressure water hoses, and tear gas are an indelible part of American history. For much of the 20th century, law enforcement chose to brazenly enforce the status quo of overt discrimination, rather than protect and serve all citizens.

The raw memories of these injustices have been resurrected by several high profile incidents of questionable uses of force. Michael Brown, unarmed, was shot twelve times by a police officer in Ferguson, Missouri, after Brown fit the description of a robbery suspect of a nearby store. Eric Garner, unarmed, was approached because officers believed he was selling single cigarettes from packs without tax stamps and in the process of arresting him an officer choked him and he died. Walter Scott, unarmed, was stopped because of a non-functioning third brake light and was shot eight times in the back while attempting to flee. Samuel Du Bose, unarmed, was stopped for failure to display a front license plate and while trying to drive away was fatally shot once in the head. Rekia Boyd, unarmed, was killed by a Chicago police officer who fired into a group of people five times from inside his police car. Zachary Hammond, unarmed, was driving away from a drug deal sting operation when he was shot to death by a Seneca, South Carolina, police officer. He was white. And so are 44% of police shooting subjects.[1]

These incidents, some captured on video and viewed widely, have generated protests in Ferguson, New York City, Washington, Chicago, Oakland, and several other cities and a national movement (Black Lives Matter) and a much needed national discourse about race, law enforcement, and policy. Police precincts from Houston, TX, to Camden, NJ, to Tacoma, WA, are beginning to issue body worn cameras, engaging in community policing, and enrolling officers in training in an effort to purge racial bias from their instinctual decision making. However, for all the eerie similarities

[1] Author's calculations based on ProPublica research that analyzes FBI data between 1980 and 2012.

between the current spate of police interactions with African Americans and the historical injustices which remain unhealed, the current debate is virtually data free. Understanding the extent to which there are racial differences in police use of force and (if any) whether those differences might be due to discrimination by police or explained by other factors at the time of the incident is a question of tremendous social importance, and the subject of this paper.

A primary obstacle to the study of police use of force has been the lack of readily available data. Data on lower level uses of force, which happen more frequently than officer-involved shootings, are virtually non-existent. This is due, in part, to the fact that most police precincts don't explicitly collect data on use of force, and in part, to the fact that even when the data is hidden in plain view within police narrative accounts of interactions with civilians, it is exceedingly difficult to extract. Moreover, the task of compiling rich data on officer-involved shootings is burdensome. Until recently, data on officer-involved shootings were extremely rare and contained little information on the details surrounding an incident. A simple count of the number of police shootings that occur does little to explore whether racial differences in the frequency of officer-involved shootings are due to police malfeasance or differences in suspect behavior.[2]

In this paper, we estimate the extent of racial differences in police use of force using four separate datasets – two constructed for the purposes of this study.[3] The first comes from NYC's Stop, Question, and Frisk program (hereafter Stop and Frisk). Stop and Frisk is a practice of the New York City police department in which police stop and question a pedestrian, then can frisk them for weapons or contraband. The dataset contains roughly five million observations. And, important for the purposes of this paper, has detailed information on a wide range of uses of force – from putting hands on civilians to striking them with a baton. The second dataset is the Police-Public Contact Survey, a triennial survey of a nationally representative sample of civilians, which contains – from the civilian point of view – a description of interactions with police, which includes uses of force. Both these datasets are public-use and readily available.[4]

[2] Newspapers such as the Washington Post estimate that there were 965 officer-involved shootings in 2015. Websites such as fatal encounters estimate that the number of annual shootings is approximately 704 between 2000 and 2015.

[3] Throughout the text, I depart from custom by using the terms "we," "our," and so on. Although this is sole-authored work, it took a large team of talented individuals to collect the data necessary for this project. Using "I" seems disingenuous.

[4] The NYC Stop and Frisk data has been used in Gelman et al. (2012) and Coviello and Persico (2015) to understand whether there is evidence of racial discrimination in proactive policing and Ridgeway (2009) to develop a statistical method to identify problem officers. The Police-Public Contact Survey has been used, mainly in criminol-

The other two datasets were assembled for the purposes of this research. We use event summaries from all incidents in which an officer discharges his weapon at civilians – including both hits and misses – from three large cities in Texas (Austin, Dallas, Houston), six large Florida counties, and Los Angeles County, to construct a dataset in which one can investigate racial differences in officer-involved shootings. Because all individuals in these data have been involved in a police shooting, analysis of these data alone can only estimate racial differences on the intensive margin (e.g., did the officer discharge their weapon before or after the suspect attacked).

To supplement, our fourth dataset contains a random sample of police-civilian interactions from the Houston police department from arrests codes in which lethal force is more likely to be justified: attempted capital murder of a public safety officer, aggravated assault on a public safety officer, resisting arrest, evading arrest, and interfering in arrest. Similar to the event studies above, these data come from arrests narratives that range in length from two to one hundred pages. A team of researchers was responsible for reading arrest reports and collecting almost 300 variables on each incident. Combining this with the officer-involved shooting data from Houston allows us to estimate both the extensive (e.g., whether or not a police officer decides to shoot) and intensive margins. Further, the Houston arrests data contain almost 3,500 observations in which officers discharged charged electronic devices (e.g., tasers). This is the second most extreme use of force, and in some cases, is a substitute for lethal use of force.

The results obtained using these data are informative and, in some cases, startling. Using data on NYC's Stop and Frisk program, we demonstrate that on non-lethal uses of force – putting hands on civilians (which includes slapping or grabbing) or pushing individuals into a wall or onto the ground, there are large racial differences. In the raw data, blacks and Hispanics are more than fifty percent more likely to have an interaction with police which involves any use of force. Accounting for baseline demographics such as age and gender, encounter characteristics such as whether individuals supplied identification or whether the interaction occurred in a high- or low-crime area, or civilian behaviors does little to alter the race coefficient. Adding precinct and year fixed effects, which estimates racial differences in police use of force by restricting to variation within a given police precinct in a given year reduces the black coefficient by 19.4 percent and the

ogy, to study questions such as whether police treatment of citizens impacts the broader public opinion of the police (Miller et al., 2004).

Hispanic coefficient by 26 percent, though both are still statistically larger than zero. Including more than 125 controls available in the data, the odds-ratio on black (resp. Hispanic) is 1.173 (resp. 1.120).

Interestingly, as the intensity of force increases (e.g. handcuffing civilians without arrest, drawing or pointing a weapon, or using pepper spray or a baton), the probability that any civilian is subjected to such treatment is small, but the racial difference remains surprisingly constant. For instance, 0.26 percent of interactions between police and civilians involve an officer drawing a weapon; 0.02 percent involve using a baton. These are rare events. Yet, the results indicate that they are significantly more rare for whites than blacks. In the raw data, blacks are 21.3 percent more likely to be involved in an interaction with police in which at least a weapon is drawn than whites and the difference is statistically significant. Adding our full set of controls reduces the racial difference to 19.4 percent. Across all non-lethal uses of force, the odds-ratio of the black coefficient ranges from 1.163 (0.036) to 1.249 (0.129).

Data from the Police-Public Contact Survey are qualitatively similar to the results from Stop and Frisk data, both in terms of whether or not any force is used and the intensity of force, though the estimated racial differences is larger. In the raw data, blacks and Hispanics are approximately two percentage points more likely than whites to report any use of force in a police interaction. The white mean is 0.8 percent. Thus, the odds ratio is 3.335 for blacks and 2.584 for Hispanics. As the use of force increases, the racial difference remains roughly constant. Adding controls for civilian demographics, civilian behavior, contact and officer characteristics, or year does little to alter the results. The coefficients are virtually unchanged and are all highly significant with the exception of the highest uses of force for which data is sparse.

There are several potential explanations for the quantitative differences between our estimates using Stop and Frisk data and those using PPCS data. First, we estimate odds-ratios and the baseline probability of force in each of the datasets is substantially different. Second, the PPCS is a nationally representative sample of a broad set of police-civilian interactions. Stop and Frisk data is from a particularly aggressive form of policing in a dense urban area. Third, the PPCS is gleaned from the civilian perspective. Finally, granular controls for location are particularly important in the Stop and Frisk data and unavailable in PPCS. In the end, the “Truth” is likely somewhere in the middle and, importantly, both bounds are statistically and economically important.

In stark contrast to non-lethal uses of force, we find no racial differences in officer-involved shootings on either the extensive or intensive margins. Using data from Houston, Texas – where we have both officer-involved shootings and a randomly chosen set of potential interactions with police where lethal force may have been justified – we find, in the raw data, that blacks are 23.8 percent *less* likely to be shot at by police relative to whites. Hispanics are 8.5 percent less likely. Both coefficients are statistically insignificant. Adding controls for civilian demographics, officer demographics, encounter characteristics, type of weapon civilian was carrying, and year fixed effects, the black (resp. Hispanic) coefficient is 0.924 (0.417) (resp. 1.256 (0.595)). These coefficients are remarkably robust across alternative empirical specifications and subsets of the data. Partitioning the data in myriad ways, we find no evidence of racial discrimination in officer-involved shootings. Investigating the intensive margin – the timing of shootings or how many bullets were discharged in the endeavor – there are no detectable racial differences.[5]

Our results have several important caveats. First, all but one dataset was provided by a select group of police departments. It is possible that these departments only supplied the data because they are either enlightened or were not concerned about what the analysis would reveal. In essence, this is equivalent to analyzing labor market discrimination on a set of firms willing to supply a researcher with their Human Resources data! There may be important selection in who was willing to share their data. The Police-Public contact survey partially sidesteps this issue by including a nationally representative sample of civilians, but it does not contain data on officer-involved shootings.

Relatedly, even police departments willing to supply data may contain police officers who present contextual factors at that time of an incident in a biased manner – making it difficult to interpret regression coefficients in the standard way.[6] It is exceedingly difficult to know how prevalent this type of misreporting bias is (Schneider 1977). Accounting for contextual variables recorded by police officers who may have an incentive to distort the truth is problematic. Yet, whether or not we include controls does not alter the basic qualitative conclusions. And, to the extent that there

[5]It is important to recognize that there may be racial bias in the likelihood of appearing in the Houston Arrest Data.

[6]In the Samuel DuBose case at the University of Cincinnati, the officer reported "Mr. DuBose pulled away and his arm was caught in the car and he got dragged" yet body camera footage showed no such series of events. In the Laquan McDonald case in Chicago, the police reported that McDonald lunged at the officer with a knife while dash-cam footage showed the teenager walking away from the police with a small knife when he was fatally shot 16 times by the officer.

are racial differences in underreporting of non-lethal use of force (and police are more likely to not report force used on blacks), our estimates may be a lower bound. Not reporting officer-involved shootings seems unlikely.

Third, given the inability to randomly assign race, one can never be confident in the direct regression approach when interpreting racial disparities. We partially address this in two ways. First, we build a model of police-civilian interactions that allows for both statistical and taste-based discrimination and use the predictions of the model to help interpret the data. For instance, if police officers are pure statistical discriminators then as a civilian's signal to police regarding their likelihood of compliance becomes increasingly deterministic, racial differences will disappear. To test this, we investigate racial differences in use of force on a set of police-civilian interactions in which the police report the civilian was compliant on every measured dimension, was not arrested, and neither weapons nor contraband was found. In contrast to the model's predictions, racial differences on this set of interactions is large and statistically significant. Additionally, we demonstrate that the *marginal* returns to compliant behavior are the same for blacks and whites, but the *average* return to compliance is lower for blacks – suggestive of a taste-based, rather than statistical, discrimination.

For officer-involved shootings, we employ a simple test for discrimination inspired by Knowles, Persico, and Todd (2001) and Anwar and Fang (2006). We investigate the fraction of white and black suspects, separately, who are armed conditional upon being involved in an officer-involved shooting. If the ordinal threshold of shooting at a black suspect versus a white suspect is different across officer races, then one could reject the null hypothesis of no discrimination. Our results, if anything, are the opposite. We cannot reject the null of no discrimination in officer-involved shootings.

Taken together, we argue that the results are most consistent with, but in no way proof of, taste-based discrimination among police officers who face convex costs of excessive use of force. Yet, the data does more to provide a more compelling case that there is no discrimination in officer-involved shootings than it does to illuminate the reasons behind racial differences in non-lethal uses of force.

The rest of the paper is organized as follows. The next section describes and summarizes the four data sets uses in the analysis. Section 3 presents estimates of racial differences on non-lethal uses of force. Section 4 describes a similar analysis for the use of lethal force. Section 5 attempts to reconcile the new facts with a simple model of police-civilian interaction that incorporates both statistical and taste-based channels of discrimination. The final section concludes. There are 3 online appendices. Appendix A describes the data used in our analysis and how we coded variables. Appendix B describes the process of creating datasets from event summaries. Appendix C provides additional theoretical results.

II. The Data

We use four sources of data – none ideal – which together paint an empirical portrait of racial differences in police use of force. The first two data sources – NYC's Stop and Frisk program and the Police-Public Contact Survey (PPCS) – provide information on non-lethal force from both the police and civilian perspectives, respectively. The other two datasets – event summaries of officer-involved shootings in ten locations across the US, and data on interactions between civilians and police in Houston, Texas, in which the use of lethal of force may have been justified by law – allow us to investigate racial differences in officer-involved shootings on both the extensive and intensive margins.

Below, I briefly discuss each dataset in turn. Appendix A provides further detail.

A. New York City's Stop-Question-and-Frisk Program

NYC's Stop-Question-and-Frisk data consists of five million individual police stops in New York City between 2003 and 2013. The database contains detailed information on the characteristics of each stop (precinct, cross streets, time of day, inside/outside, high/low crime area), civilian demographics (race, age, gender, height, weight, build, type of identification provided), whether or not the officers were in uniform, encounter characteristics (reason for stop, reason for frisk (if any), reason for search (if any), suspected crime(s)), and post-encounter characteristics (whether or not a weapon was eventually found or whether an individual was summonsed, arrested, or a crime committed).

Perhaps the most novel component of the data is that officers are required to document which one of the following seven uses of force was used, if any: (1) hands, (2) force to a wall, (3) handcuffs, (4) draw weapon, (5) push to the ground, (6) point a weapon, (7) pepper spray or (8) strike with a baton.[7] Officers are instructed to include as many uses of force as apply. For instance, if a stop results in an officer putting his hands on a civilian and, later within the same interaction, pointing his weapon, that observation would have both "hands" and "point a weapon" as uses of force. Unfortunately, officers are not required to document the sequence in which they used force.

These data have important advantages. First, the Stop and Frisk program encompasses a diverse sample of police-civilian interactions.[8] Between the years 2003 and 2013, the same period as the Stop and Frisk data, there were approximately 3,457,161 arrests in NYC – 26.3% fewer observations than Stop and Frisk excluding stops that resulted in arrests.[9] Unfortunately, even this robust dataset is incomplete – nowhere is the universe of all police interactions with civilians – or even all police stops – recorded.

Second, lower level uses of force – such as the use of hands – are both recorded in these data and more frequently used by law enforcement than more intense uses of force. For instance, if one were to use arrest data to glean use of force, many lower level uses of force would simply be considered standard operating procedure. Putting hands on a suspect, pushing them up against a wall, and putting handcuffs on them are so un-noteworthy in the larger context of an arrest that they are not recorded in typical arrest descriptions. Yet, because proactive policing is a larger and less confrontational portion of police work, these actions warrant data entry.

The key limitation of the data is they only capture the police side of the story. There have been several high-profile cases of police storytelling that is not congruent with video evidence of the interaction. Another important limitation for inference is that the data do not provide a way to identify officers or individuals. Ideally, one would simply cluster standard errors at the officer level to account for the fact that many data points – if driven by a few aggressive officers – are correlated and classic inference treats them as independent. Our typical regressions cluster standard errors at the precinct level. Appendix table 9 explores the robustness of our results for more disaggregated clusters – precinct x time of day, block-level, and even block x time of day. Our conclusions are

[7] Police officers can also include "other" force as a type of force used against civilians. We exclude "other" forces from our analysis. Appendix Table 3 calculates racial differences in the use of "other" force and shows that including these forces does not alter our results.

[8] Technically, NYC police are only required to record a stop if some force was used, a civilian was frisked or searched, was arrested, or refused to provide identification. Nonetheless, roughly 41 percent of all stops in the database appear to be reported despite not resulting in any of the outcomes that legally trigger the requirement to record the stop.

[9] This number was calculated from the Division of Criminal Justice Services' record of adult arrests by counties in New York City between 2003 and 2013.

unaffected by any of these alternative ways to cluster standard errors.

Summary statistics for the Stop and Frisk data are displayed in table 1A. There are six panels. Panel A contains baseline characteristics. Fifty eight percent of all stops recorded were of black civilians. If police were stopping individuals at random, this number would be closer to 25.5 percent (the fraction of black civilians in New York City according to US Census 2010 records). Hispanics make up twenty-five percent of the stops. The data are comprised predominantly of young males; the median age is 24 years old. The median age in NYC is roughly 11 years older.

Panel B describes encounter characteristics for the full sample and then separately by race. Most stops occur outside after the sun has set in high-crime areas. A surprisingly small number of stops – about three percent – the police report finding any weapon or contraband. Panel C displays variables that describe civilian behavior. Approximately 50 percent of stops were initiated because a civilian fit the relevant description of a person of interest, were assumed to be a lookout for a crime, or the officers were casing a victim or location.

Panel D contains a series of alternative outcomes such as whether a civilian was frisked, summonsed, or arrested. Panel E provides descriptive statistics for the seven forms of force available in the data. Panel F provides the frequency of missing variables.

B. The Police-Public Contact Survey

The Police-Public Contact Survey (PPCS) – a nationally representative sample – has been collected by the Bureau of Justice Statistics every three years since 1996. The most recent wave publicly available is 2011. Across all years, there are approximately 500,000 observations.

The main advantage of the PPCS data is that, unlike any of our other datasets, it provides the civilian's interpretation of interactions with police. The distinction between PPCS data and almost any other data collected by the police is similar to the well-known differences between certain data in the Uniform Crime Reports (UCR) and the National Crime Victimization Survey (NCVS).[10] One explanation for these differences given in the literature is that individuals are embarrassed or

[10] According to the US Department of Justice, UCR and NCVS measure an overlapping but nonidentical set of crimes. The UCR Program's primary objective is to provide a reliable set of criminal justice statistics by compiling data from monthly law enforcement reports or individual crime incident reports transmitted directly to FBI or to centralized state agencies that then report to FBI. The BJS, on the other hand, established the NCVS to provide previously unavailable information about crime (including crime not reported to police), victims and offenders. Therefore, there are discrepancies in victimization rates from the two reports like the UCR which reports 89,000 forcible rapes in 2010 while the NCVS reports 203,830 rapes and sexual assaults in 2010.

afraid to report certain crimes to police or believe that reporting such crimes have unclear benefits and potential costs. Police use of force – in particular for young minority males – may be similar.

Another key advantage is that it approximates the universe of potential interactions with police – rather than limited to arrests or police stops.[11] If a police officer is investigating a crime in a neighborhood and they discuss it with a civilian – this type of interaction would be recorded in the PPCS. Or, if a police officer used force on a civilian and did not report the interaction – this would not be recorded in police data but would be included in the PPCS.

The PPCS also has important limitations. First, data on individual's locations is not available to researchers. Second, the data on contextual factors surrounding the interaction with police or the officer's characteristics are limited. Third, the survey omits individuals who are currently in jail. Fourth, the PPCS only includes the civilian account of the interaction which could be biased in its own way. In this vein, according to individuals in the PPCS data, only 4.18% of them have resisted arrests and only 11% of civilians argued when they were searched despite not being guilty of carrying alcohol, drugs or weapons.

Table 1B presents summary statistics for PPCS data. There are six panels. Panel A contains civilian demographics. Blacks comprise roughly eleven percent of the sample, women are 53 percent. The average age is approximately 17 years older than the Stop and Frisk data. Over 60 percent of the sample reports being employed in the previous week – average income category in the sample is 1.95. Income is recorded as a categorical variable that is 1 for income levels below $20,000, 2 for income levels between $20,000$ and $49,999$, and 3 for income levels greater than $50,000.

Panel B describes self-reported civilian behavior. According to the all PPCS survey respondents, almost no civilians disobey police orders, try to get away, resist, argue or threaten officers. However, since these questions were asked in response to why force was used against respondents, if we restrict the data to civilians who report non-missing use of force from police officers, this percentage rises to 15.3 percent.

Panel C of table 1B includes summary data on the types of contact and officer characteristics. Almost half of the interactions between the public and police are traffic stops, eighteen percent are from street interactions – including the types of street interaction that may not appear in our

[11] Contacts exclude encounters with private security guards, police officers seen on a social basis, police officers related to the survey respondents, or any contacts that occurred outside the United States.

Stop and Frisk data – and thirty percent are "other" which include being involved in a traffic accident, reporting a crime, being provided a service by the police, participating in block watch or other anti-crime programs, or being suspected by the police of something or as a part of a police investigation. Panel D contains alternative outcomes and Panel E describes the five uses of force available in the data. Panel F provides the frequency of missing variables.

C. Officer-Involved Shootings

There are no systematic datasets which include officer-involved shootings (OIS) along with demographics, encounter characteristics, and suspect and police behavior.[12] For the purposes of this project, we compile a dataset on officer-involved shootings from ten locations across America.

To begin, fifteen police departments across the country were contacted by the author: Boston, Camden, NYC, Philadelphia, Austin, Dallas, Houston, Los Angeles, six Florida counties, and Tacoma, Washington.[13] Importantly for thinking about the representativeness of the data – many of these cities were a part of the Obama Administration's Police Data Initiative.[14] We received data from all but three of these police departments – NYC, Philadelphia, and Tacoma, Washington – all of which have indicated a willingness to participate in our data collection efforts but have not yet provided data.[15] This is likely not a representative set of cities. Appendix Table 14 investigates differences between the cities that provided us data and other Metropolitan Statistical Areas on a variety of dimensions such as population demographics and crime rates.

In most cases, OIS data begins as event summaries from all incidents in which a police officer discharged their firearm at civilians (including both hits and misses). These summaries, in many cases, are more than fifty page descriptions of the factors surrounding an officer-involved shooting. Below is an extract from a "typical" summary:

[12] Data constructed by the Washington Post has civilian demographic identifiers, weapons carried by civilian, signs of mental illness and an indicator for threat level but no other contextual information.

[13] Another approach is to request the data from every police department vis-a-vis a freedom of information request. We attempted this method, but police departments are not obliged to include detailed event summaries. In our experience, the only way to obtain detailed data is to have contacts within the police department.

[14] The White House launched the Police Data Initiative as a response to the recommendations made by the Task Force on 21st Century Policing. The Initiative was created to work with police departments to leverage data on police-citizen interactions (e.g., officer-involved shootings, use of force, body camera videos and police stops) to increase transparency and accountability.

[15] Camden and Boston each had one OIS during the relevant time frame, so we did not use their data for this analysis. Camden provided remarkable data on police-civilian interactions which will be used in future work.

> "As I pointed my rifle at the vehicle my primary focus was on the male passenger based on the information provided by the dispatcher as the person who had been armed inside the store. As the vehicle was driving past me I observed the male passenger in the truck turn around in the seat, and begin pointing a handgun at me through an open rear sliding glass window. When I observed this I was still yelling at the female to stop the truck! The male suspect appeared to be yelling at me, but I could not hear him. At that point the truck was traveling southbound toward the traffic light on Atlantic Boulevard, and was approximately 30-40 feet away from me. The car had already passed me so the driver was no longer in my line of fire. I could also see my back drop consisted of a wooded area of tall pine trees. It appeared to me at that time that his handgun was moving in a similar fashion of being fired and going through a recoil process, but I could not hear gunshots. Fearing for my life, the lives of the citizens in the area and my fellow officers I began to fire my rifle at the suspect."

To create a dataset out of these narratives, a team research assistants read each summary and extracted data on 65 pre-determined variables in six categories: (A) suspect characteristics, (B) suspect weapon(s), (C) officer characteristics, (D) officer response reason, (E) other encounter characteristics, and (F) location characteristics.[16] Suspect characteristics include data on suspect race, age and gender. Suspect weapon variables consist of dummy variables for whether the suspect used a firearm, sharp object, vehicle, or other objects as a weapon or did not have a weapon at all. Officer characteristics include variables that determine the majority race of the officer unit, whether there were any female officers in the unit, average tenure of the shooting officer and dummy variables for whether the officer was on duty and was accompanied by two or more officers on the scene. Officer response reason variables determine the reason behind the officer being present at the scene. They include dummy variables on whether the officer was present as a response to a robbery, a violent disturbance, traffic related stop, or was responding to a warrant, any suspicious activity, a narcotics transaction, a suicide, responding because he was personally attacked or other reasons. Other encounter characteristics gather information on whether the shooting happened during the day or night and a variable that is coded 1 if the suspect attacked the officer or drew

[16] Appendix B provides a detailed, step-by-step, account of how the OIS dataset was created and was explicitly designed to allow researchers to replicate our analysis from the original source materials.

a weapon or attempted to draw a weapon on the officer. The variable is coded 0 if the suspect only appeared to have a weapon or did not attack the officer at all. Finally, location characteristics include dummies to represent the jurisdiction that we collected data from. Appendix B contains more details of how the variables were coded.

As a crucial check on data quality, once we coded all OIS data from the event summaries, we wrote Appendix B. We then hired eight new research assistants who did not have any involvement in creating the first dataset. We provided them the event summaries, Appendix B, and extremely minimal instructions – the type of simple clarification that would be provided to colleagues attempting to replicate our work from the source material – and they created a second, independent, dataset. All results remain qualitatively unchanged with the alternatively coded dataset.[17]

The most obvious advantage of the OIS data is the breadth and specificity of information contained in the event summaries. Descriptions of OIS are typically long and quite detailed relative to other police data. A second advantage is that officer-involved shootings are non-subjective. Unlike lower level uses of force, whether or not an officers discharges a weapon is not open to interpretation. Officers are also required to document anytime they discharge their weapon. Finally, OIS are subject to internal and often times external review.

The OIS data have several notable limitations. Taken alone, officer-involved shootings are the most extreme and least used form of police force and thus, in isolation, may be misleading. Second, the penalties for wrongfully discharging a lethal weapon in any given situation can be life altering, thus, the incentive to misrepresent contextual factors on police reports may be large.[18] Third, we don't typically have the suspect's side of the story and often there are no witnesses. Fourth, it is impossible to capture all variables of importance at the time of a shooting. Thus, what appears to be discrimination to some may look like mis-measured contextual factors to others.

A final disadvantage, potentially most important for inference, is that all observations in the OIS data are shootings. In statistical parlance, they don't contain the "zeros" (e.g., set of police interactions in which lethal force was justified but not used). To the extent that racial bias is prevalent on the extensive margin – whether or not someone is ever in an officer-involved shooting – these data would not capture it.

[17]Thanks to Derek Neal for suggesting this exercise.

[18]From interviews with dozens of current police officers, we gleaned that in most all police shootings – even when fully justified and observed by many – the officer is taken off active-duty, pending an investigation.

We address this concern both directly and indirectly in two ways. First, given the data we have, we investigate the intensive margin by defining our outcome variable as whether or not the officer shoots the suspect before being attacked. Second, we collected unprecedented data from the Houston Police Department on all arrest categories in which officers may have used justifiable force as a way to obtain the "zeros." These data are described in the next subsection.

Table 1C displays summary statistics for OIS data, divided into four locations and six categories of data. Column (1) contains observations from the full sample – 1,332 shootings between 2000 and 2015. Forty-six percent of officer-involved shootings in our data are blacks, thirty percent are Hispanic, and twenty-four percent are other with the majority in that category being whites. Given the spate of video evidence on police shootings – all of which are of blacks – it is a bit surprising that they are less than half of the observations in the data.

Columns (2) and (3) displays data from 507 officer-involved shootings with firearms and over 4,000 instances of an officer-involved shooting with a taser, in Houston, Texas. Most police officers in the Houston Police Department carry Glock 22, Glock 23 or the Smith & Wesson M&P40 .40 (S&W) caliber semi-automatic handguns on their dominant side, but many carry an X26 taser on their non-dominant side. We exploit this choice problem to understand how real-time police decisions may be correlated with suspect race.

Columns (4) through (6) contain OIS data from Austin and Dallas, Texas, six Florida counties (Brevard, Jacksonville, Lee, Orange, Palm Beach and Pinellas), and Los Angeles County. Panel F demonstrates that Houston accounts for 38% of all officer-involved shootings. Austin and Dallas, combined, provide 20% of the data while Florida provides 27% of the data. Panel G provides the frequency of missing variables.

D. Houston Police Department Arrests Data

The most comprehensive set of OIS data is from the Houston Police Department (HPD). For this reason, we contacted HPD to help construct a set of police-civilian interactions in which lethal force may have been justified. According to Chapter 9 of the Texas Penal Code, police officers' use of deadly force is justified "when and to the degree the actor reasonably believes the force is immediately necessary." Below, we describe the task of implementing this obtuse definition in data in an effort to develop a set of police-civilian interactions in which the use of lethal force may have been justified by law.

There are approximately 100,000 arrests per year in Houston; 1.6 million total over the years we have OIS data. If the data were more systematically collected, the tasks of creating potential risk sets would be straightforward. Data in HPD is the opposite – most of it is narrative reports in the form of unstructured blocks of text that one can link to alternative HPD data with unique case IDs.[19]

We sample case IDs from five arrest categories which are more likely to contain incidence in which lethal force was justified: attempted capital murder of a public safety officer, aggravated assault on a public safety officer, resisting arrest, evading arrest, and interfering in arrest.[20] This process narrowed the set of relevant arrests to 16,000 total, between 2000 and 2015. We randomly sampled five percent of these arrest records and manually coded 290 variables per arrest record. This process took between 30 and 45 minutes per record to manually keypunch and includes variables related to specific locations for calls, incidents, and arrests, suspect behavior, suspect mental health, suspect injuries, officer use of force, and officer injuries resulting from the encounter.

These data are merged with data on officer demographics and suspect's previous arrest history to produce a comprehensive incident-level dataset on interactions between police and civilians in which lethal force may have been justified.

We also collected 4,250 incident reports for all cases in which an officer discharged their taser. These data form another potential risk set. It it important to note: technology allows for HPD to centrally monitor the frequency and location of taser discharges.

Table 1D provides descriptive statistics for the Houston Arrest Data. Compared to the officer-involved shootings dataset, civilians sampled in the arrest dataset carry far fewer weapons – 95% do not carry weapons compared to 21% in the OIS dataset. The other variable that is significantly different between the two datasets is the fraction of suspects who attacked or drew weapon – 56% in the HPD arrest dataset compared to 80% in the OIS dataset.

[19] In conversations with engineers and data scientists at Google, Microsoft Research, and several others in Artificial Intelligence and Machine Learning, we were instructed that current natural language processing algorithms are not developed for the level of complexity in our police data. Moreover, one would need a "test sample" (manually coded data to assess the algorithm's performance) of several hundred thousand to design an algorithm. This is outside the scope of the current project.

[20] Our original request to HPD was for a dataset similar to OIS for all arrests between 2000 and 2015. The response: "we estimate that it will take 375 years to fulfill that request."

III. Estimating Racial Differences in Non-Lethal Use of Force

NYC's Stop, Question, and Frisk Data

Table 2A presents a series of estimates of racial differences in police use of force using the Stop and Frisk data. We estimate logistic regressions of the following form:

$$\text{Force}_{i,p,t} = \text{Race}_i'\alpha + X_{i,t}'\beta + Z_{p,t}'\mu + \nu_t + \psi_p + \epsilon_{i,p,t} \tag{1}$$

where $\text{Force}_{i,p,t}$ is a measure of police use of force on individual i, in precinct p, at time t. A full set of race dummies for civilians are included in the regressions, with white as the omitted category. Consequently, the coefficients on race capture the gap between the named racial category and whites – which is reported as an Odds Ratio.[21] The vectors of covariates included in the specification, denoted $X_{i,t}'$ and $Z_{p,t}$, vary between rows in table 2A. As one moves down the table, the set of coefficients steadily grows. We caution against a causal interpretation of the coefficients on the covariates, which are better viewed as proxies for a broad set of environmental and behavioral factors at the time of an incident. Standard errors, which appear below each estimate, are clustered at the precinct level unless otherwise specified.

The first row in table 2A presents the differences in means for any use of force. These results reflect the raw gaps in whether or not a police stop results in any use of force, by race. Blacks are 53% more likely to experience any use of force relative to a white mean of 15.3 percent. The raw gap for Hispanics is almost identical. Asians are no more likely than whites to experience use of force. Other race – which includes American Indians, Alaskan natives or other races besides white, black, Hispanic and Asian – is smaller but still considerable.

The raw difference between races is large – perhaps too large – and it seems clear that one needs to account for at least some contextual factors at the time of a stop in order to better understand, for example, whether racial differences are driven by police response to a given civilian's behavior or racial differences in civilian behavior. Yet, it is unclear how to account for context that might predict how much force is used by police and not include variables which themselves might be

[21] Appendix Tables 2A through 2G runs similar specification using ordinary least squares and obtains similar results. Estimating Probit models provides almost identical results.

influenced by biased police.[22]

Row (2) adds baseline civilian characteristics – such as age and gender – all of which are exogenously determined and not strategically chosen as a function of the police interaction. Adding these variables does almost nothing to alter the odds ratios. Encounter characteristics – whether the interaction happened inside, the time of day, whether it occurred in a high or low crime area, and whether the civilian provided identification – are added as controls in row 3. If anything, adding these variables increase the odds ratios on each race, relative to whites. Surprisingly, accounting for civilian behavior – row 4 in the table – does little to alter the results.

The final row in table 2A includes both precinct and year fixed effects. This significantly changes the magnitude of the coefficients. Blacks are seventeen percent more likely to incur any use of force, accounting for all variables we can in the data. Hispanics are roughly twelve percent more likely.[23] Both are statistically significant. Asians are slightly less likely, though not distinguishable from whites.

These results have two potential takeaways: precincts matter and, accounting for a large and diverse set of control variables, black civilians are still more likely to experience police use of force. Of the 112 variables available in the data, there is no linear combination that fully explains the race coefficients.[24] From this point forward, we consider the final specification, including precinct and year fixed effects as our main specification.

Inferring racial differences in the types of force used in a given interaction is a bit more nuanced. Police report that in twenty percent of all stops, some use of force is deployed. Officers routinely record more than one use of force. For instance, a stop might result in an officer putting their hands on a civilian, who then pushes the officer and the officer responds by pushing him to the ground. This would be recorded as "hands" and "force to ground". In 85.1% of cases, exactly one use of force is recorded. Two use of force categories were used in 12.6% of cases, 1.8% report three

[22] The traditional literature in labor economics – beginning with Mincer (1958) – dealt with similar issues. O'Neill (1990) and Neal and Johnson (1996) sidestep this by demonstrating that much of the racial wage gap can be accounted for by including only pre-market factors such as test scores.

[23] Even accounting for eventual outcomes of each stop – which include being let go, being frisked, being searched, being arrested, being summonsed, and whether or not a weapon or some form of contraband was found – blacks are twenty-two percent more likely to experience force and Hispanics are twenty-seven percent more likely. We did not include these control variables in our main specification due to the fear of over-controlling if there is discrimination in the probability of arrests, conditional on behavior.

[24] Using data on geo-spatial coordinates, we also included block-level fixed effects and the results were qualitatively unchanged.

use of force categories, and 0.6% of all stop and frisk incidents in which force is used record more than three uses of force.

There are several ways to handle this. The simplest is to code the max force used as "1" and all the lower level uses of force in that interaction as "0". In the example above in which an officer recorded both "hands" and "forced to the ground" as uses of force, one would ignore the use of hands and code forced to the ground as "1." The limitation of this approach is that it discards potentially valuable information on lower level uses of force. When analyzing racial differences in the use of hands by police, one would miss this observation. A similar issue arises if one uses the parallel "min."[25]

Perhaps a more intuitive way to code the data is to treat each use of force as "at least as much". In the example above, both hands and forced to the ground would be coded as "1" in the raw data. When analyzing racial differences in the use of hands by police, this observation would be included. The interpretation would not be racial differences in the use of hands, per se, but racial differences in the use of "at least" hands. To be clear, an observation that records only hands would be in the hands regression but <u>not</u> the regression which restricts the sample to observations in which individuals were at least forced to the ground. This is the method we use throughout.

Results using this method to describe racial differences for each use of force are displayed in Figure 1. The x-axis contains use of force variables that range from at least hands to at least the use of pepper spray or baton. The y-axis measures the odds ratio for blacks (panel A) or Hispanics (panel B). The solid line is gleaned from regressions with no controls, and the dashed line adds precinct and year fixed effects (equivalent to row 5 in table 2A).

For blacks, the consistency of the odds ratios are striking. As the use of force increases, the frequency with which that level of force is used decreases substantially. There are approximately five million observations in the data – 19 percent of them involve the use of hands while 0.04 percent involve using pepper spray or a baton. The use of high levels of force in these data are rare. Yet, it is consistently rarer for whites relative to blacks. The range in the odds ratios across all levels of force is between 1.163 (0.036) and 1.236 (0.058).

Interestingly, for Hispanics, once we account for our set of controls, there are small differences

[25]Appendix table 8 demonstrates that altering the definition to be "at most" or using the max/min force used in any given police interaction does not alter the results.

in use of force for the lower level uses of non-lethal force, but the differences converge toward whites as the use of force increases both in the raw data and with the inclusion of controls.

One may be concerned that restricting all the coefficient estimates to be identical across the entire sample may yield misleading results. Regressions on a common support (for example, only on males or only on police stops during the day) provide one means of addressing this concern. Table 3 explores the sensitivity of the estimated racial gaps in police use of force across a variety of subsamples of the data. I report only the odds-ratios on black and Hispanic and associated standard errors. The top row of the table presents baseline results using the full (any force) sample and our parsimonious set of controls (corresponding to row 5 in table 2a). The subsequent rows investigate racial differences in use of force for high/low crime areas, time of day, whether or not the officer was in uniform, indoors/outdoors, gender of civilian, and eventual outcomes.

Most of the coefficients on race do not differ significantly across these various subsamples with the exception of time of day and eventual outcomes. Black civilians are 7 percent more likely to have any force used against them conditional on being arrested. They are 15 percent more likely to have any force used against them conditional on being summonsed and 11.1 percent more likely conditional on having weapons or contraband found on them. Results are similar for Hispanics. Additionally, for both blacks and Hispanics, racial differences in use of force are more pronounced during the day relative to night.

To dig deeper, Panel A in figure 2 plots the odds ratios of any use of force for black civilians versus white civilians for every hour of day. Panel B displays the average use of force for black civilians and white civilians for every hour of day. These figures show that force against black civilians follows approximately the same pattern as white civilians, though the difference between average force between the two races decreases at night.

Police-Public Contact Survey

One of the key limitations of the Stop and Frisk data is that one only gets the police side of the story, or more accurately, the police entry of the data. It is plausible that there are large racial differences that exist that are masked by police misreporting. The Police-Public Contact Survey is one way to partially address this weakness.

Table 2B presents a series of estimates of racial differences in police use of force using the PPCS data. The specifications estimated are of the form:

$$\text{Force}_{i,t} = \text{Race}_i'\alpha + X_{i,t}'\beta + \nu_t + \epsilon_{i,t},$$

where $\text{Force}_{i,t}$ is a measure of police use of force reported by individual i in year t. A full set of race dummies for individuals and officers are included in the regressions, with white as the omitted category. The vectors of covariates included in the specification vary across rows in table 2B. As one moves down the table, the set of coefficients steadily grows. Standard errors, which appear below each estimate, account for heteroskedasticity.

Generally, the data are qualitatively similar to the the results using Stop and Frisk – namely, despite a large and complex set of controls, blacks and Hispanic are more likely to experience some use of force from police. A key difference, however, is that the share of individuals experiencing any use of force is significantly lower. In the Stop and Frisk data, 15.3 percent of whites incur some force from police. In the PPCS, this number is 1%. There are a variety of potential reasons for these stark differences. For instance, the PPCS is a nationally representative sample of interactions with police from across the U.S., whereas the Stop and Frisk data is gleaned from a rather aggressive proactive policing strategy in a large urban city. This is important because in what follows we present odds-ratios. Odds-ratios are informative, but it is important for the reader to know that the baseline rate of force is substantially smaller in the PPCS.

Blacks are three times more likely to report use of force by police in the raw data. Hispanics are 2.6 times more likely. Adding controls for demographic and encounter characteristics, civilian behavior, and year fixed effects reduces the odds-ratio to roughly 2.7 for blacks and 1.7 for Hispanics. Differences in quantitative magnitudes aside, the PPCS paints a similar portrait – large racial differences in police use of force that cannot be explained using a large and varied set of controls.

One important difference between the PPCS and the Stop and Frisk data is in regards to racial differences on the more extreme uses of non-lethal force: using pepper spray or striking with a baton. Recall, in the Stop and Frisk data the odds ratios were relatively consistent as the intensity of force increased. In the PPCS data, if anything, racial differences on these higher uses of force disappear. For kicking or using a stun gun or pepper spray, the highest use of force available, the black coefficient is 1.867 (0.589) and the Hispanic coefficient is 1.228 (0.468), though because of the rarity of these cases the coefficients are barely statistically significant at the 5% level.

Table 4 explores the heterogeneity in the data by estimating racial differences in police use of force in the PPCS on various subsamples of the data: civilian income, gender, civilian, time of contact, and officer race. Civilian income is divided into three categories: less than \$20,000, between \$20,000 and \$50,000, and above \$50,000. Strikingly, both the black and Hispanic coefficients are statistically similar across these income levels suggesting that higher income minorities do not price themselves out of police use of force – echoing some of the ideas in Cose (1993). Racial differences in police of force does not seem to vary with civilian gender or officer race. Consistent with the results in the Stop and Frisk data, the black coefficient is 3.17 (0.85) for interactions that occur during the day and 1.68 (0.48) for interactions that occur at night. The p-value on the difference is marginally significant.

Putting the results from the Stop and Frisk and PPCS datasets together, a pattern emerges. Relative to whites, blacks and Hispanics seem to have very different interactions with law enforcement – interactions that are consistent with, though definitely not proof of, some form of discrimination. Including myriad controls designed to account for civilian demographics, encounter characteristics, civilian behavior, eventual outcomes of the interaction and year reduces, but cannot eliminate, racial differences in non-lethal use of force in either of the datasets analyzed.

IV. Estimating Racial Differences in Officer-Involved Shootings

We now focus on racial differences in officer-involved shootings. We begin with specifications most comparable to those used to estimate racial differences in non-lethal force, using both data from officer-involved shootings in Houston and data we coded from Houston arrest records that contains interactions with police that might have resulted in the use of lethal force.[26] Specifically, we estimate the following empirical model:

$$\text{shooting}_{i,t} = \text{Race}_i'\alpha + X_{i,t}'\beta + \nu_t + \epsilon_{i,t},$$

[26]Because of this select set of "0s" the non-black, non-Hispanic mean, displayed in column 1, is drastically larger than a representative sample of the population – which would be approximately .0001%. 45.5 percent of whites in our data were involved in an officer-involved shooting.

where $\text{shooting}_{i,t}$ is a dichotomous variable equal to one if a police officer discharged their weapon at individual i in year t. There are no accidental discharges in our data and shootings at canines have been omitted. A full set of race dummies for individuals and officers are included in the regressions, with non-black non-Hispanics as the omitted category for individuals. The vectors of covariates included in the specification vary across rows in table 5. As one moves down the table, the set of coefficients steadily grows. As one moves across the columns of the table, the comparison risk set changes.[27] Presenting the results in this way is meant to underscore the robustness of the results to the inclusion of richer sets of controls and to alternative interpretations of the risk sets. Standard errors, which appear below each estimate, account for heteroskedasticity.

Given the stream of video "evidence", which many take to be indicative of structural racism in police departments across America, the ensuing and understandable outrage in black communities across America, and the results from our previous analysis of non-lethal uses of force, the results displayed in Table 5 are startling.

Blacks are 23.8 percent *less* likely to be shot by police, relative to whites. Hispanics are 8.5 percent less likely to be shot but the coefficient is statistically insignificant.

Rows (2) through (6) add various controls, identical to those in table 1D. Accounting for basic suspect or officer demographics, does not significantly alter the raw racial differences. Including encounter characteristics – which one can only accomplish by hand coding the narratives embedded in arrests reports – creates more parity between blacks and non-black non-Hispanic suspects, rendering the coefficient closer to 1. Finally, when we include whether or not a suspect was found with a weapon or year fixed effects, the coefficients still suggest that, if anything, officers are less likely to shoot black suspects, ceteris paribus, though the racial differences are not significant.

Columns (4) and (5) of table 5 include 4504 incident reports from 2005-2015 for all arrests during which an officer reported using his taser as a risk set, in addition to all OIS in Houston from that time period. The empirical question here is whether or not there are racial differences in the split-second decision as to whether to use lethal or non-lethal force through the decision to shoot a pistol or taser.

Consistent with the previous results, the raw racial difference in the decision to employ lethal

[27]Appendix Table 6 investigates the sensitivity of the main results to more alternative compositions of the risk sets.

force using this taser sample is negative and statistically significant. Adding suspect and officer demographics, encounter characteristics and year controls does little to change the odds ratios for black versus non-black suspects. Including all controls available from the taser sample, table 5 shows that black civilians are 30.9 percent *less* likely to be shot with a pistol (rather than a taser) relative to non-black suspects. Columns (6) and (7) pool the sample from hand coded arrest data and taser data. Results remain qualitatively the same. Controlling for all characteristics from incident reports, black suspects are 21.6 percent less likely to be shot than non-black suspects.

To be clear, the empirical thought experiment here is that a police officer arrives at a scene and decides whether or not to use lethal force. Our estimates suggest that this decision is not correlated with the race of the suspect. This does not, however, rule out the possibility that there are important racial differences in whether or not thse police-civilian interactions occur at all.

Appendix Table 5 explores the sensitivity of the results for various subsamples of the data: number of officers who respond to the scene, whether the suspect attacked an officer first, whether the officer was on-duty, whether the unit that responded was majority black or Hispanic or majority white or Asian, and the type of call the officer was responding to (a partial test of the selection issue described above). Equations identical to (3) are estimated, but due to the smaller sample sizes inherent in splitting the sample, we estimate Ordinary Least Squares regressions.

None of the subsamples explored demonstrate much difference of note. We find no evidence that racial differences in the use of lethal force varies in a statistically meaningful way between the number of officers at an incident. We find no differences in the use of lethal force across different call slips – the p-value for equality of race coefficient across different calls slips is 0.557 for black suspects – suggesting that officers seeking confrontation in random street interactions in a way that causes important selection bias into our sample is not statistically relevant. Subsampling on the number and racial composition of the officer unit also shows no evidence of racial differences.

Another way to investigate the robustness of our coefficients is to analyze the odds ratios across time. These data are displayed in figure 4. Racial differences in OIS between 2000 and 2015 are remarkably constant. This interval is interesting and potentially informative as it is 9 years after the public beatings of Rodney King and includes the invention of Facebook, the iPhone, YouTube, and related technology that allows bystanders to capture police-civilian interactions and make it publicly available at low costs. Crudely, the period between 2000 and 2005 one might think to be years in which police misconduct could more easily go unnoticed and for which the public attention was relatively low. Thus, the disincentive to misreport was likely lower. After this period, misreporting costs likely increased. Yet, as we see from figure 3, this does not seem to influence racial differences in the use of lethal force.

Are there Racial Differences in the Timing of Lethal Force?

The above results, along with the results on use of force, are about racial differences on the extensive margin: whether or not an officer uses a particular type of force or decides to use lethal force on a suspect. Because of the richness of our officer-involved shooting database, we can also investigate the intensive margin – whether there are racial differences in how quickly a police officer shoots a suspect. In particular, given the narrative accounts, I create a dichotomous variable that is equal to one if a police officer reports that she (he) shoots a suspect before they are attacked and zero if they report shooting the suspect after being attacked. These data are available for Houston as well as the other nine locations where we collected OIS data. An important caveat to these data is that the sequence of events in a police-civilian interaction is subject to misreporting by police. Thus, the dependent variable is subjective.

Table 6 presents a series of estimates of racial differences in the timing of police shootings using the OIS data. The specifications estimated are of the form:

$$\text{Shoot First}_{i,c,t} = \text{Race}_i'\alpha + X_{i,t}'\beta + Z_{c,t}'\mathcal{T} + \nu_t + \psi_c + \epsilon_{i,c,t},$$

where Shoot $\text{First}_{i,c,t}$ is a measure of whether a police officer reports shooting individual i, in city c, in year t, before being attacked. Standard errors, which appear below each estimate, are clustered at the location level unless otherwise specified.

The results from these specifications are consistent with our previous results on the extensive margin. Row (1) displays the results from the raw data. Blacks are 1.3% less likely to be shot first by police. Hispanics are slightly more likely. Neither coefficient is statistically significant. Adding suspect or officer demographics does not alter the results.[28]

Row (4) accounts for important context at the time of the shooting. For instance, whether

[28] We also estimate the "intensity" of force used in officer-involved shootings by estimating racial differences in the total number of bullets used in a given police shooting. The average number of bullets in officer-involved shootings involving blacks is 0.256 (0.508) more relative to shootings that involve whites [not shown in tabular form].

the shooting happened during day time or night time and whether the suspect drew weapon or attacked the officer. Including these variables decreases the black coefficient to 0.693 (0.096) which is statistically significant. The Hispanic coefficient is similar in size but less precisely estimated. Adding whether the suspect was eventually found to have a weapon and its type or including location and year fixed effects only strengthens the results in the unexpected direction. Including all controls available, officers report that they are 47.4% less likely to discharge their firearms before being attacked if the suspect is black. The Hispanic coefficient is strikingly similar (43.6% less likely).

Appendix Table 7 explores the heterogeneity in the data across various subsamples: number of officers who arrive at a scene, whether or not officers report that the suspect clearly drew their weapon or whether they "appeared" to draw their weapon, whether the officer was on-duty, the call type, and the racial composition of the responding unit. The final panel provides results disaggregated by location.

Estimated race coefficients across call types – whether officers were dispatched because of a violent crime, robbery, auto crime, or other type of call – are statistically identical. This is particularly interesting in light of the potential selection into the sample of OIS cases discussed earlier. Indeed, the majority of police shootings in our data occur during violent crimes or robberies and there are no racial differences on these call types.

One of the more interesting subsamples is whether or not a suspect "appeared" to have a weapon versus an officer indicating that it was clear he had a weapon. This dovetails with many of the anecdotal reports of police violence and is thought to be a key margin on which implicit bias, and the resulting discriminatory treatment, occur. Eberhardt et al. (2004) finds that police officers detect degraded images of crime related objects faster when they are shown black faces first.

Yet our data from the field seem to reject this lab-based hypothesis, at least as regards officer-involved shootings. The coefficient on black for the subsample who police report clearly drew their weapon first is -0.105 (0.020). The same coefficient estimated on the set of interactions were police assumed an individual had a weapon is -0.038 (0.033). The Hispanic coefficients are nearly identical.

More generally, the coefficients are uncommonly consistent across all subsamples of the data. Of the 5 tests of equality performed in the table, not one is significant. We cannot detect racial differences in officer-involved shootings on any dimension.

V. Interpretation

A number of stylized facts emerge from the analysis of the preceding sections. On non-lethal uses of force, there are racial differences – sometimes quite large – in police use of force, even after controlling for a large set of controls designed to account for important contextual and behavioral factors at the time of the police-civilian interaction. As the intensity of use of force increases from putting hands on a civilian to striking them with a baton, the overall probability of such an incident occurring decreases but the racial difference remains roughly constant. On the most extreme uses of force, however – officer-involved shootings with a Taser or lethal weapon – there are no racial differences in either the raw data or when accounting for controls.

In this section, we explore the extent to which a model of police-civilian interaction that encompasses both information- and taste-based discrimination – can successfully account for this set of facts. The model is an adaptation of Coate and Loury (1993a, 1993b).

A. A Model of Police-Civilian Interactions

Basic Building Blocks

Imagine a large number of police officers and a weakly larger population of civilians. Each police officer is randomly matched with civilians from this population. Civilians belong to one of two identifiable groups, B or W. Denote by λ the fraction of W's in the population. Police officers are assumed to be one of two types: "biased" or "unbiased." Let $\delta \in (0, 1)$ denote the fraction of biased police officers.

Nature moves first and assigns a cost of compliance to each civilian and a type to each police officer. Let $c \in [\underline{c}, \overline{c}]$, represent the cost to a civilian of investing in compliance. An alternative way to think about this assumption is that individuals contain inherent dangerousness and those who are dangerous have higher costs of compliance.

After observing his cost, the civilian makes a dichotomous compliance decision, choosing to become either a compliant type or a non-compliant type with no in-between. Then, based on this decision, nature distributes a signal $\theta \in [\underline{\theta}, \overline{\theta}]$ to police officers regarding whether or not a civilian is likely to comply.[29] Next, the police officer observes θ and decides whether or not to use force,

[29]This model is a simplified version of a more general model in which individuals invest in a "compliance identity"

which we denote $h \in \{0, 1\}$.[30]

The distribution of θ depends, in the same way for each race, on whether or not a civilian has invested in compliance. This signal is meant to capture the important elements of initial interactions between police and civilians; clothing, demeanor, attitude, posture, and so on. Let $F_1(\theta)$ [resp. $F_0(\theta)$] be the probability that the signal does not exceed θ, given that a civilian has invested in compliance (resp. non-compliance) and let $f_1(\theta)$ and $f_0(\theta)$ be the related density functions. Define $\mu(\theta) \equiv \frac{f_0(\theta)}{f_1(\theta)}$ to be the likelihood ratio at θ. We assume that $\mu(\theta)$ is non-increasing on $[0, 1]$, which implies that $F_1(\theta) \leqslant F_0(\theta)$ for all θ. Thus, higher values of observed θ are more likely if the civilian is compliant, and for a given prior, the posterior likelihood that a civilian will be compliant is larger if his signal takes a higher value.

Payoffs

For the civilian, payoffs depend on whether or not force is used on him and whether he chose to invest in compliance. Specifically, if force is used on the civilian, he receives a payoff of $-\gamma - c$ if he invested in compliance and $-\gamma$ if not. If force is not used on the civilian, he receives a payoff of $-c$ if he invests and the payoff is normalized to zero if he did not invest.

It is assumed that police officers want to use force on civilians who are non-compliant and prefer not to use force on those that are compliant. In addition, we allow for "biased" police officers to gain utility from using force on Bs.

Thus, for police officers, payoffs depend on their type, whether or not they use force, and whether or not the civilian is compliant. We begin with unbiased officers. If force is used, the officers payoff is $-K - \phi_F$ if the civilian is compliant and $\chi_F - \phi_F$ if the civilian is non-compliant. If no force is used, the officer receives a payoff of 0 if the civilian is compliant and $-\chi_{NF}$ if the civilian is non-compliant. These payoffs are identical for biased officers when they interact with W civilians.

When biased police officers interact with B civilians they derive psychic pleasure from using force, independent of whether they are compliant or not. We represent this by, τ a positive term

ala Akerlof and Kranton (2000) and then, in any given interaction with police, decide whether to comply or escalate. For those who have a compliance identity, there is an identity costs of escalation. This model is more intuitive, but delivers the same basic results.

[30]We model the police officer's decision as deciding to use force rather than what type of force to use for two reasons: analytical convenience and for most of our analysis the dependent variable is whether or not to use force. Extending our analysis to allow for N potential uses of force does not alter the key predictions of the model.

in the biased officer's payoff when he uses force on B civilians. Note: This is similar to the taste parameter pioneered in Becker (1957).

Strategies

A civilian's strategy is a mapping $I : [\underline{c}, \overline{c}] \rightarrow \{0, 1\}$. Without loss of generality, the civilian's strategy can be represented by a cut-off point, c^*, such that the civilian will invest in compliance if and only if their cost is below c^*. A strategy for the police officer is a decision of whether or not to use force, conditional upon what he can observe, $h : \{0, \tau\} \times [B, W] \times [\underline{\theta}, \overline{\theta}] \rightarrow \{0, 1\}$.

Expected Payoffs

Let $\pi \in [0, 1]$ denote the officer's prior belief that a civilian will be compliant. Expected payoffs for the police officer are functions of her beliefs, her type, and the signal she receives. Given π and observed signal θ, she formulates a posterior probability (using Bayes' rule) that the civilian will be compliant: $\Psi(\pi, \theta) \equiv \frac{\pi f_1(\theta)}{\pi f_1(\theta) + (1-\pi) f_0(\theta)}$.

The expected payoff of using force for an unbiased police officer (and, equivalently, a biased police officer when interacting with Ws) is:

$$\Psi(\pi, \theta)(-K - \phi_F) + (1 - \Psi(\pi, \theta))(\chi_F - \phi_F). \tag{2}$$

The expected payoff of using force for an biased officer interacting with Bs is:

$$\Psi(\pi, \theta)(-K - \phi_F) + (1 - \Psi(\pi, \theta))(\chi_F - \phi_F) + \tau. \tag{3}$$

Relatedly, the expected payoffs of *not* using force, for both types of officers, can be written as:

$$-(1 - \Psi(\pi, \theta))(\chi_{NF}). \tag{4}$$

Combining equation (2) and equation (4), and using a bit of algebra, an unbiased officer uses force only if

$$\theta \leqslant \theta^*_{ub} \equiv min\{\theta | \Psi(\pi, \theta)(-K - \phi_F) + (1 - \Psi(\pi, \theta))(\chi_F + \chi_{NF} - \phi_F)\} > 0) \tag{5}$$

In words, equation (5) provides a threshold, θ^*_{ub}, such that for any θ below this threshold unbiased officers always use force. Similarly, using the corresponding expected payoffs for a biased officer, one can derive θ^*_b.

Now, consider the civilian's expected payoff. W civilians receive $F_1(\theta^*{}_{ub})(-\gamma) - c$ if they invest and $F_0(\theta^*_{ub})(-\gamma)$ if they choose not to invest. When optimizing, a civilian will invest in compliance if and only if the cost of compliance is less than the net benefit of compliance. In symbols, $c \leqslant c^*_W \equiv \{F_{nc}(\theta^*_{ub}) - F_c(\theta^*_{ub})\}\,\gamma$. Similarly, Bs invest if $c \leqslant c^*_B \equiv \gamma\,\{\delta(F_{nc}(\theta^*_{ub}) - F_c(\theta^*_{ub})) + (1-\delta)(F_{nc}(\theta^*_b) - F_c(\theta^*_b))\}$. Note – given we assume $\delta > 0$ – it follows that $c^*_B < c^*_W$.

Definition 1 *An equilibrium consists of a pair (θ^*, π^*) such that each is a best response to the other.*

B. Understanding the Data Through the Lens of the Model

Assuming the distribution of costs (c) and the signal (θ) are independent of race, racial disparities can be produced in this model in two (non-mutually exclusive) ways: different beliefs or different preferences.[31] To see this formally, suppose all racial differences were driven by information-based discrimination and there was no taste-based component. In this case, equation (3) simplifies to (2) and both B and W individuals' net benefit of investment becomes $\{F_{nc}(\theta^*_{ub}) - F_c(\theta^*_{ub})\}\,\gamma - c$. Thus, one needs differences in π to generate discriminatory equilibrium.

In contrast, one can also derive an equilibrium for cases in which we turn off the information-based channel and only allow differences through preferences. In this case, police officers observe investment decisions perfectly. When police officer bias is sufficiently large, any equilibrium will contain discrimination against Bs.

Distinguishing between these two cases, empirically, is difficult with the available data. In what follows, we attempt to understand whether the patterns in the data are best explained by an information-based or taste-based approach to discrimination – recognizing that both channels may be important.

Statistical Discrimination

[31] It is also plausible that racial differences arise due to differences in costs of compliance (for instance, through peer effects) or in the signal distributions. Incorporating these assumptions into the model is a trivial extension.

To better understand whether statistical discrimination might explain some of the patterns in the data, we investigate two possibilities.[32] First, we explore whether racial differences in mean characteristics across police precincts predicts racial differences in use of force. The key – untestable – assumption is police officer beliefs about the compliance of a civilian – π in our model – is partly driven by local variation in variables such as education or income levels.[33]

Table 7 explores racial differences in any use of force – using the Stop and Frisk data – for various proxies for "dangerousness" including education, income, and unemployment. Education is represented by the fraction, by race, in each precinct of individuals with a high school diploma. Income is measured as median income. Unemployment is measured as the fraction of civilians in the labor force who are unemployed. For each of these variables, we take the difference between the white population and black population and rank the precincts by this difference, individually. We then divide the data into terciles. The first tercile is always the one in which racial differences between our proxies are the lowest. The third tercile represents precincts in which there are relatively large racial differences on a given proxy.

Statistically larger racial differences in use of force for the third tercile (first tercile for unemployment), relative to tercile one or two (tercile two or three for unemployment), would be evidence consistent with statistical discrimination. This would imply that racial differences in use of force are correlated with racial differences in proxies for dangerousness. Table 7 demonstrates no such pattern. The odds-ratio of having any force used on a black civilian versus a white civilian remains statistically the same across terciles. [34]

A second prediction of the statistical discrimination model that is testable in our data is how racial differences in use of force change as signals about civilian compliance become more clear.[35]

[32] Appendix C considers the extent to which discrimination based on categories can explain the results (Fryer and Jackson 2008). We argue categorical discrimination is inconsistent with the fact that black officers and white officers interact similarly with black civilians. See Appendix Table 11.

[33] Ideally, one might use variables more directly correlated with dangerousness such as racial differences in crime rates, by precincts. Despite repeated formal Freedom of Information Law requests, the New York Police Department refused to supply these data.

[34] We performed a similar exercise exploiting the *variance* across space in proxies for dangerousness (see Appendix Tables 10A-10C for results). We also investigated whether more weight in the bottom quintiles of the distribution of our proxies predicted police use of force. These empirical exercises were meant as a partial test of Aigner and Cain (1977). We find no evidence of this sort of statistical discrimination on any of the dimensions tested.

[35] Another potential test of statistical discrimination was pioneered by Altonji and Pierret (2001). They investigate racial differences in wage trajectories, conditional upon being hired. To the extent that statistical discrimination drives wage differences between racial groups, one would expect the wage trajectory for blacks to be higher than whites – as employers learn. We performed a similar, though imperfect, test by estimating the probability that a civilian is arrested, conditional upon force being used. Consistent with a discrimination story, on the lowest level use of

If statistical discrimination is the key driver of racial differences in use of force, the model predicts that as θ becomes perfectly predictive of compliance behavior, there will be no racial differences. We test this using officer recorded data on the compliance behavior of civilians.

The NYC Stop and Frisk data contains officer recorded information on the compliance of civilians during a stop. These variables include: whether the civilians refused to comply with officers' directions, whether the civilian verbally threatened an officer, whether they were evasive in their response to questioning or whether they changed direction at the sight of an officer. If statistical discrimination is a key driver of racial differences, on the set of interactions in which officers report perfect compliance (and, to capture potentially important unobservables – the civilian was not arrested or was not guilty of carrying weapons or contraband) racial differences should be close to zero. And, on the set of interactions in which civilians engage in questionable behavior, racial differences should be statistically larger.

Figure 5 shows that even when we take perfectly compliant individuals and control for civilian, officer, encounter and location variables, black civilians are 21.1 (0.041) percent more likely to have any force used against them compared to white civilians with the same reported compliance behavior. As the intensity of force increases, the odds ratio for perfectly compliant individuals decreases.

Ultimately, it is difficult to know if statistical discrimination is an important component of racial differences in use of force. Though our tests have quite limited power, we find no evidence that statistical discrimination plays an important role.

Taste-Based Models of Discrimination

Similar to any large organization, police departments surely have individuals who hold biased views toward minority citizens and those views may manifest themselves in biased treatment of individuals based solely on their race. Yet, as Becker (1957) argued, individual discrimination does not necessarily equate to market (or systemic) discrimination.

Taste-based discrimination is consistent with the data from the direct regression approach on non-lethal uses of force if, among those who discriminate, the preference for discrimination is greater than the expected costs of wrongly using force. In other words, the expected price of

force, blacks and Hispanics are less likely to be arrested conditional upon force being used. As the intensity of force increases, if anything, minorities are more likely to be arrested conditional upon force being used.

discrimination is not large enough – either through low penalties or low probabilities of detection – to alter behavior of those who have biased preferences. This model is also consistent with the lack of racial differences in officer-involved shooting if there is a discrete increase in the costs of being deemed a discriminator, relative to the costs incurred with non-lethal uses of force.[36]

Below, we explore the extent to which two additional implications of the taste-based channel of our model are borne out in the data. The first uses the predictions on average versus marginal returns of compliant behavior. The second is inspired by the seminal work in Knowles, Persico, and Todd (2001) and Anwar and Fang (2006).

In any equilibrium model of discrimination, officer behavior influences the incentive to invest in compliance behavior. This is made explicit in equations (6) and (7). Figure 5 provides some suggestive evidence that the returns to compliance may be different across races. We can test this a bit more directly. One issue in this setting, which does not arise in labor markets, is it's not obvious how to aggregate non-compliance into a monotonic index. It may be considered more dangerous from a police officer perspective that a civilian shouted verbal threats than refusing to comply with an officer's directions and being evasive regarding questioning. A simple aggregation of the number of non-compliant activities is likely misleading.

To sidestep this important potential issue of aggregating non-compliance, we create an index equal to 1 if a civilian changes direction at the sight of an officer, 2 if a civilian is non-compliant on any other, but not all dimensions of measured compliance, and 3 if a civilian is non-compliant on all four dimensions we can measure. The regression estimated, then, is whether or not an officer uses any force – accounting for our full set of controls – and including our measure of non-compliance interacted with race. Racial differences in the *marginal* return to non-compliance behavior would manifest itself in statistically different coefficients on the compliance variable. For a given race, adding both the race coefficient and the interaction term with compliance behavior provides an estimate of the net benefit of investment (equations (6) and (7)).

The results of this exercise [not shown in tabular form] are consistent with racial differences in police use of force being driven by taste-based discrimination. Black civilians have statistically

[36]While purely anecdotal, in police departments across the country, any officer-involved shooting – no matter how "justified" – results in the temporary confiscation of the officer's weapon until an investigation of the incident is complete This is a potentially high cost relative to other non-lethal uses of force. Moreover, in informal interviews with dozens of police officers in Boston, Cambridge, Camden, and Houston – almost all police officers described pulling the trigger of their weapon as a "life altering event."

similar marginal returns to compliance as white civilians. In other words, the probability of force being used as θ increases is statistically identical between blacks and whites. Yet, black civilians always have a higher likelihood of force being used on them compared to white civilians, for all θ. Further, the net benefit of investment in compliance is lower for blacks relative to whites. This is precisely what the model predicts if racial animus is an important factor in explaining racial differences in use of force.

We conclude our statistical analysis by developing a test for discrimination based on Knowles, Persico, and Todd (2001) [hereafter KPT] and Anwar and Fang (2006) to complement the direct regression approach described in the previous sections. KPT tests for racist preferences by looking at officers' success rate of searches across races. Their model assumes that police maximize the number of successful searches net of the cost of searching motorists. If racial prejudice exists then the cost of searching drivers will be different across races. This, in turn, implies that the rate of successful searches will be different across races.

Anwar and Fang (2001) build upon the theory of KPT; arguing that the KPT results might not hold if police officers are non-monolithic in their behavior. They test this by investigating search rates of civilians of a particular race, across officer races. Under the null hypothesis that none of the racial groups of officers has racial prejudice, it must be true that the ranking of search rates for white civilians across officer races is the same as the ranking of search rates for black civilians across officer races.

We adopt this approach by investigating whether or not a suspect was eventually found to have a weapon during the interaction with police. In other words, we calculate the probability, for each race, that a suspect has a weapon conditional upon being involved in an officer-involved shooting. Given the level of detail in our data, one can perform this test for weapons generally – guns, knives or other cutting objects, or assault weapons – or for guns specifically, including pistols, rifles, or semi-automatic machine guns, specifically. Moreover, following the insights in Anwar and Fang (2006), we disaggregate the data by officer race.

The null hypothesis is no racial discrimination in officer-involved shootings. The null could be rejected in several ways. First, according to KPT, the null could be rejected if the fraction of suspects carrying weapons or firearms is different across suspect races. Second, according to Anwar and Fang, the null could be rejected if the ranking of "being armed" rates for black suspects across officer races is different from the ranking of being armed rates for white suspects.

Consistent with our direct regression approach and the findings in Knowles, Persico, and Todd (2001), and Anwar and Fang (2006), we fail to reject the null of no discrimination. The data are displayed in Table 8. For white officers, the probability that a white suspect who is involved in officer-involved shooting has a weapon is 85.1% percent. The equivalent probability for blacks is 81%. A difference of 4%, which is not statistically significant. For black officers, the probability that a white suspect who is involved in an officer-involved shooting has a weapon is surprisingly lower, 62.5%. The equivalent probability for black suspects is 74%. The only statistically significant differences by race demonstrate that black officers are more likely to shoot unarmed whites, relative to white officers.

We perform a similar exercise for non-lethal uses of force, recognizing that as the use of force gets less extreme the application of that force and whether or not a suspect has a weapon is more tenuous. For instance, investigating racial differences in whether or not officers use "hands" on civilians who are unarmed is not a valid test of discrimination as there are myriad legitimate reasons for police officers to place hands on civilians who are unarmed. Yet, racial differences in the use of a baton – after accounting for suspect behavior – seem less justifiable. Unfortunately, where to draw the line on the continuum of potential uses of force is ad hoc. Thus, we present our modified KPT test for all uses of force while acknowledging that for the low level uses, it does not seem appropriate.

Table 9 presents these results. Each row is a different level of force which begins with "at least hands" and increases in severity of force until "use of pepper spray or Baton." Column (1) contains the white mean. Columns (2) and (3) display the coefficient on black and Hispanic, respectively. Column (4) displays the number of observations which range from over one million for the use of hands to 1,745 for the use of pepper spray or baton.

Blacks are 1.3 (0.4) percentage points less likely to have a weapon, conditional upon a police officer using any force. Hispanics are 0.8 (0.3) less likely to have a weapon. Both are statistically significant. Interestingly, on all other non-lethal uses of force, the probability that a weapon is found – conditional upon force being used – is statistically identical across races. Taken at face value, these data are consistent with discrimination against minorities on the lowest level uses of non-lethal force.

VI. Conclusion

The issue of police violence and its racial incidence has become one of the most divisive topics in American discourse. Emotions run the gamut from outrage to indifference. Yet, very little data exists to understand whether racial disparities in police use of force exist or might be explained by situational factors inherent in the complexity of police-civilian interactions. Beyond the lack of data, the analysis of police behavior is fraught with difficulty including, but not limited to, the reliability of the data that does exist and the fact that one cannot randomly assign race.

With these caveats in mind, this paper takes first steps into the treacherous terrain of understanding the nature and extent of racial differences in police use of force. On non-lethal uses of force, there are racial differences – sometimes quite large – in police use of force, even after accounting for a large set of controls designed to account for important contextual and behavioral factors at the time of the police-civilian interaction. Interestingly, as use of force increases from putting hands on a civilian to striking them with a baton, the overall probability of such an incident occurring decreases dramatically but the racial difference remains roughly constant. Even when officers report civilians have been compliant and no arrest was made, blacks are 21.3 (0.04) percent more likely to endure some form of force. Yet, on the most extreme use of force – officer-involved shootings – we are unable to detect any racial differences in either the raw data or when accounting for controls.

We argue that these facts are most consistent with a model of taste-based discrimination in which police officers face discretely higher costs for officer-involved shootings relative to non-lethal uses of force. This model is consistent with racial differences in the average returns to compliant behaviors, the results of our tests of discrimination based on Knowles, Persico, and Todd (2001) and Anwar and Fang (2006), and the fact that the odds-ratio is large and significant across all intensities of force – even after accounting for a rich set of controls. In the end, however, without randomly assigning race, we have no definitive proof of discrimination. Our results are also consistent with mismeasured contextual factors.

As police departments across America consider models of community policing such as the Boston Ten Point Coalition, body worn cameras, or training designed to purge officers of implicit bias, our results point to another simple policy experiment: increase the expected price of excessive force on lower level uses of force. To date, very few police departments across the country either collect data on lower level uses of force or explicitly punish officers for misuse of these tactics.

The appealing feature of this type of policy experiment is that it does not require officers to change their behavior in extremely high-stakes environments. Many arguments about police reform fall victim to the "my life versus theirs, us versus them" mantra. Holding officers accountable for the misuse of hands or pushing individuals to the ground is not likely a life or death situation and, as such, may be more amenable to policy change.

****

The importance of our results for racial inequality in America is unclear. It is plausible that racial differences in lower level uses of force are simply a distraction and movements such as Black Lives Matter should seek solutions within their own communities rather than changing the behaviors of police and other external forces.

Much more troubling, due to their frequency and potential impact on minority belief formation, is the possibility that racial differences in police use of non-lethal force have spillovers on myriad dimensions of racial inequality. If, for instance, blacks use their lived experience with police as evidence that the world is discriminatory, then it is easy to understand why black youth invest less in human capital or black adults are more likely to believe discrimination is an important determinant of economic outcomes. Black Dignity Matters.

References

[1] Allport, G.W. 1954. *The Nature of Prejudice.* Reading, MA: Addison Wesley

[2] Aigner, D.J. and Cain, G.G., 1977. "Statistical theories of discrimination in labor markets." *Industrial and Labor relations review*, pp.175-187.

[3] Altonji, J.G. and Pierret, C.R., 2001. Employer Commitment and Statistical Discrimination. *The Quarterly Journal of Economics*, 116.

[4] Anwar, S. and Fang, H., 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review*, 96(1): 127-151.

[5] Becker Gary, S., 1957. *The economics of discrimination.*

[6] Coate, S. and Loury, G.C., 1993a. Will affirmative-action policies eliminate negative stereotypes?. *The American Economic Review*, pp.1220-1240.

[7] Coate, S. and Loury, G., 1993b. Antidiscrimination enforcement and the problem of patronization. *The American Economic Review*, 83(2), pp.92-98.

[8] Coviello, D. and Persico, N., 2015. An Economic Analysis of Black-White Disparities in NYPDs Stop and Frisk Program. *Journal of Legal Studies*, 44(2), pp. 315-360.

[9] Cose, E., 1993. *The rage of a privileged class: Why are middle-class Blacks angry? Why should America care.* New York: *HarperPerennial.*

[10] Eberhardt, J.L., Goff, P.A., Purdie, V.J. and Davies, P.G., 2004. Seeing black: race, crime, and visual processing. *Journal of Personality and Social Psychology*, 87(6), p.876.

[11] Fiske, S.T. 1998. "Stereotyping, Prejudice, and Discrimination" in D.T. Gilbert, S.T. Fiske and G. Lindzey, eds, *Handbook of Social Psychology, vol 2.* New York: Oxford University Press, 357-414.

[12] Fryer, R. and Jackson, M.O., 2008. "A categorical model of cognition and biased decision making." *The BE Journal of Theoretical Economics*, 8(1).

[13] Gelman, A., Fagan, J. and Kiss, A., 2012. "An analysis of the New York City police department's stop-and-frisk policy in the context of claims of racial bias." *Journal of the American Statistical Association.*

[14] Goff, P.A., Jackson, M.C., Di Leone, B.A.L., Culotta, C.M. and DiTomasso, N.A., 2014. "The essence of innocence: Consequences of dehumanizing Black children." *Journal of Personality and Social Psychology*, 106(4), p.526.

[15] Knowles, J., Persico, N., and Todd, P., 2001. "Racial Bias in Motor-Vehicle Searches: Theory and Evidence." *Journal of Political Economy*, 109(1), pp. 203-29.

[16] Miller, J., Davis, R.C., Henderson, N.J., Markovic, J. and Ortiz, C.W., 2004. Public opinions of the police: The influence of friends, family and news media. *New York: Vera Institute of Justice.*

[17] Mincer, J., 1958. "Investment in human capital and personal income distribution." *The Journal of Political Economy*, pp.281-302.

[18] Neal, D. A., and Johnson, W. R. 1996. "The Role of Premarket Factors in Black White Wage Differences." *Journal of Political Economy*, 104, 869-895.

[19] O'Neill, J. 1990. "The Role of Human Capital in Earnings Differences Between Black and White Men." *Journal of Economic Perspectives*, 4, 25-45.

[20] Ridgeway, G., and MacDonald, J. M. 2009. Doubly robust internal benchmarking and false discovery rates for detecting racial bias in police stops. *Journal of the American Statistical Association*, 104(486), 661-668.

[21] Schneider, A., 1977. Portland (Or) Forward Record Check of Crime-Victims Final Report, December 1977. *Oregon Research Institute and United States of America, 1977.*

[22] Sporer, S. 2001. "Recognizing Faces of Other Ethnic Groups: An Integration of Theories," *Psychology, Public Policy, and Law*, 7, 36-97.

Table 1A
Summary Statistics for New York City Stop, Question and Frisk, 2003 - 2013

	Full Sample	White	Black	Hispanic	*p-val*	*p-val*
	(1)	(2)	(3)	(4)	(2) = (3)	(2) = (4)
Panel A: Baseline Characteristics						
White	0.10	1.00	0.00	0.00	.	.
Black	0.58	0.00	1.00	0.00	.	.
Hispanic	0.25	0.00	0.00	1.00	.	.
Asian	0.03	0.00	0.00	0.00	.	.
Other race	0.04	0.00	0.00	0.00	.	.
Age	28.00	29.25	27.96	27.57	0.000	0.000
Male	0.93	0.90	0.93	0.93	0.000	0.000
Panel B: Encounter Characteristics						
Indoors	0.23	0.16	0.26	0.21	0.000	0.000
Daytime	0.36	0.39	0.35	0.36	0.000	0.000
High-crime Area	0.56	0.52	0.57	0.55	0.000	0.000
High-crime Time	0.37	0.36	0.38	0.36	0.000	0.931
Police in Uniform	0.72	0.64	0.73	0.72	0.000	0.000
Photo ID	0.53	0.63	0.51	0.54	0.000	0.000
Verbal ID	0.43	0.34	0.45	0.43	0.000	0.000
Refused ID	0.02	0.01	0.03	0.02	0.000	0.000
Other ID	0.02	0.01	0.02	0.01	0.000	0.144
With Others Who Were Stopped	0.23	0.30	0.21	0.26	0.000	0.000
Wpn or Contraband Fnd	0.03	0.04	0.03	0.03	0.000	0.000
Panel C: Civilian Behavior						
Carrying Suspicious Obj	0.03	0.04	0.02	0.03	0.000	0.000
Fit Relevant Descr	0.17	0.19	0.17	0.17	0.000	0.000
Preparing for Crm	0.29	0.35	0.27	0.30	0.000	0.000
Lookout for Crm	0.17	0.20	0.16	0.18	0.000	0.000
Dressed in Crm Attire	0.04	0.03	0.05	0.04	0.000	0.000
Appearance of Drug Tran	0.09	0.09	0.10	0.09	0.000	0.000
Suspicious Mvmnts	0.44	0.37	0.46	0.43	0.000	0.000
Engaging in Vlnt Crm	0.08	0.06	0.08	0.09	0.000	0.000
Concealing Suspicious Obj	0.09	0.04	0.10	0.08	0.000	0.000
Other Suspicious Bhvr	0.21	0.21	0.21	0.20	0.162	0.000
Panel D: Alternative Outcomes						
Frisked	0.55	0.43	0.57	0.56	0.000	0.000
Searched	0.08	0.08	0.08	0.09	0.022	0.000
Arrested	0.06	0.06	0.06	0.06	0.000	0.000
Summonsed	0.06	0.06	0.06	0.06	0.010	0.000
Wpn or Contraband Fnd	0.03	0.04	0.03	0.03	0.000	0.000
Panel E: Use of Force						
Hands	0.19	0.13	0.19	0.20	0.000	0.000
Push to Wall	0.03	0.03	0.03	0.03	0.000	0.000
Handcuffs	0.04	0.04	0.04	0.03	0.085	0.019

Draw Weapon	0.00	0.00	0.00	0.00	0.000	0.077
Push to Ground	0.01	0.01	0.01	0.01	0.000	0.000
Point Weapon	0.00	0.00	0.00	0.00	0.000	0.027
Pepper Spray/Baton	0.00	0.00	0.00	0.00	0.005	0.385
Panel F: Missing Variables						
Missing Race	0.00	0.00	0.00	0.00	.	.
Missing Age	0.01	0.01	0.01	0.01	0.000	0.000
Missing Gender	0.03	0.01	0.01	0.01	0.000	0.000
Missing Indoors	0.01	0.01	0.00	0.01	0.000	0.000
Missing Daytime	0.00	0.00	0.00	0.00	0.503	0.093
Missing High-crm Area	0.00	0.00	0.00	0.00	.	.
Missing High-crm Time	0.00	0.00	0.00	0.00	.	.
Missing Police Uniform	0.00	0.00	0.00	0.00	0.132	0.314
Missing ID	0.00	0.00	0.00	0.00	0.196	0.093
Missing Others Stopped	0.00	0.00	0.00	0.00	0.000	0.000
Missing Wpn or Contra Fnd	0.00	0.00	0.00	0.00	0.164	0.007
Missing Carry Susp Obj	0.00	0.00	0.00	0.00	.	.
Missing Relevant Descr	0.00	0.00	0.00	0.00	.	.
Missing Preparing Crm	0.00	0.00	0.00	0.00	.	.
Missing Lookout Crm	0.00	0.00	0.00	0.00	.	.
Missing Crm Attire	0.00	0.00	0.00	0.00	.	.
Missing Drug Tran	0.00	0.00	0.00	0.00	.	.
Missing Suspicious Mvmnts	0.00	0.00	0.00	0.00	.	.
Missing Vlnt Crm	0.00	0.00	0.00	0.00	.	.
Missing Conceal Susp Obj	0.00	0.00	0.00	0.00	.	.
Missing Other Susp Bhvr	0.00	0.00	0.00	0.00	.	.
Observations	4,982,426	492,391	2,885,857	1,214,961		

Notes: This table reports summary statistics. The sample consists of all NYC stop and frisks from 2003-2013. The first column includes the entire sample. The second column includes white civilians only. The third column includes black civilians only. The fourth column includes hispanic civilians only. The fifth column reports p-values for a ttest to see whether the mean for white civilians is statistically similar to the mean for black civilians. The sixth column reports p-values for a ttest to see whether the mean for white civilians is statistically similar to the mean for hispanic civilians.

Table 1B
Summary Statistics for Police-Public Contact Survey

	Full Sample	White	Black	Hispanic	*p-val*	*p-val*
	(1)	(2)	(3)	(4)	(2) = (3)	(2) = (4)
Panel A: Civilian Demographics						
White	0.72	1.00	0.00	0.00	.	.
Black	0.11	0.00	1.00	0.00	.	.
Other Race	0.05	0.00	0.00	0.00	.	.
Hispanic	0.12	0.00	0.00	1.00	.	.
Male	0.47	0.48	0.43	0.49	0.000	0.000
Female	0.53	0.52	0.57	0.51	0.000	0.000
Age	45.36	47.19	42.76	38.55	0.000	0.000
Employed last week or not	0.61	0.61	0.59	0.63	0.000	0.000
Income	1.95	2.03	1.64	1.73	0.000	0.000
Population size of Suspect's Address	1.47	1.31	1.89	1.86	0.000	0.000
Panel B: Civilian Behavior						
Disobeyed	0.00	0.00	0.01	0.01	0.013	0.013
Tried to Get Away	0.00	0.00	0.00	0.00	0.047	0.364
Hit Officer	0.00	0.00	0.00	0.00	0.224	0.225
Resisted	0.00	0.00	0.00	0.00	0.057	0.000
Complained	0.06	0.05	0.07	0.06	0.099	0.697
Argued	0.02	0.02	0.02	0.02	0.000	0.366
Threatened Officer	0.00	0.00	0.02	0.01	0.000	0.166
Used Physical Force	0.00	0.00	0.00	0.00	0.437	0.436
Panel C: Contact and Officer Characteristics						
Incident type: Street Stop	0.19	0.18	0.19	0.22	0.000	0.000
Incident type: Traffic Stop	0.49	0.49	0.50	0.50	0.655	0.349
Incident type: Other	0.32	0.32	0.31	0.28	0.000	0.000
Time of Contact was Day	0.68	0.69	0.62	0.66	0.000	0.007
Time of Contact was Night	0.32	0.31	0.38	0.34	0.000	0.007
Hispanic	0.09	0.06	0.11	0.25	0.000	0.000
White	0.86	0.88	0.78	0.83	0.000	0.000
Black	0.08	0.07	0.20	0.07	0.000	0.201
Other Race	0.05	0.04	0.06	0.09	0.000	0.000
Panel D: Alternative Outcomes						
Civilian injured	0.15	0.16	0.12	0.13	0.334	0.483
Civilian perceived excessive force	0.64	0.61	0.69	0.66	0.038	0.279
Civilian searched	0.05	0.04	0.10	0.09	0.000	0.000
Civilian arrested	0.03	0.02	0.05	0.04	0.000	0.000
Civilian guilty of carrying drugs/alcohol/weapon	0.17	0.19	0.11	0.16	0.005	0.327
Panel E: Use of Force						
Any use of force	0.03	0.02	0.06	0.04	0.000	0.000
Grab	0.01	0.01	0.02	0.01	0.000	0.000
Kick	0.00	0.00	0.00	0.00	0.006	0.079
Point Gun	0.00	0.00	0.01	0.00	0.000	0.000

Handcuff	0.03	0.02	0.06	0.04	0.000	0.000
Pepper spray/Stungun	0.00	0.00	0.00	0.00	0.051	0.037
Panel F: Missing Variables						
Missing Gender	0.00	0.00	0.00	0.00	.	.
Missing Age	0.00	0.00	0.00	0.00	.	.
Missing Employed last week or not	0.14	0.13	0.16	0.15	0.000	0.000
Missing Income	0.14	0.14	0.14	0.14	0.509	0.165
Missing Population size of Suspect's Address	0.28	0.28	0.29	0.29	0.000	0.000
Missing Disobeyed	0.95	0.94	0.95	0.96	0.000	0.000
Missing Tried to Get Away	0.95	0.94	0.95	0.96	0.000	0.000
Missing Hit Officer	0.95	0.94	0.95	0.96	0.000	0.000
Missing Resisted	0.95	0.94	0.95	0.96	0.000	0.000
Missing Complained	0.99	0.99	0.99	0.99	0.554	0.721
Missing Argued	0.95	0.94	0.95	0.96	0.000	0.000
Missing Threatened	0.99	0.99	0.99	0.99	0.518	0.803
Missing Used Physical Force	0.95	0.94	0.95	0.96	0.000	0.000
Missing Incident type	0.42	0.42	0.43	0.43	0.000	0.000
Missing Time of Contact	0.96	0.96	0.96	0.96	0.000	0.000
Missing Officer Hispanic	0.99	0.99	0.99	0.99	0.143	0.806
Missing Officer White	0.96	0.95	0.96	0.96	0.000	0.000
Missing Officer Black	0.96	0.95	0.96	0.96	0.000	0.000
Missing Officer Other Race	0.96	0.95	0.96	0.96	0.000	0.000
Observations	566,674	404,380	62,277	67,660		

Notes: This table reports summary statistics. The sample consists of survey respondents of the Police Public Contact Survey from 1996 to 2011. The first column includes the entire sample. The second column includes white civilians only. The third column includes black civilians only. The fourth column includes hispanic civilians only. The fifth column reports p-values for a ttest to see whether the mean for white civilians is statistically similar to the mean for black civilians. The sixth column reports p-values for a ttest to see whether the mean for white civilians is statistically similar to the mean for hispanic civilians.

Table 1C
Summary Statistics for Officer Involved Shootings

	Full Sample	Houston OIS	Houston Taser	Austin + Dallas	Florida	Los Angeles
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Suspect Demographics						
Black	0.46	0.52	0.63	0.46	0.47	0.25
Hispanic	0.30	0.33	0.03	0.30	0.10	0.60
Non-Black, Non-Hisp	0.24	0.14	0.33	0.23	0.43	0.15
Male	0.96	0.96	0.94	0.97	0.97	0.97
Age	30.77	28.86	31.39	32.90	33.24	30.56
Panel B: Suspect Weapon						
Firearm	0.51	0.52	.	0.52	0.45	0.54
Sharp Object	0.08	0.08	.	0.07	0.07	0.09
Vehicle	0.15	0.11	.	0.17	0.24	0.08
None	0.21	0.24	.	0.18	0.17	0.25
Other Weapon	0.06	0.05	.	0.06	0.06	0.03
Panel C: Officer Characteristics						
Officer Unit Majority White	0.51	0.32	0.42	0.53	0.80	0.28
Officer Unit Majority Black	0.09	0.14	0.15	0.14	0.05	0.03
Officer Unit Majority Hisp	0.27	0.40	0.22	0.18	0.07	0.52
Officer Unit Majority Asian/Other	0.03	0.06	0.03	0.04	0.01	0.03
Officer Unit Split Race	0.09	0.08	0.18	0.11	0.07	0.14
Female Officers in Unit	0.06	0.06	0.13	0.06	0.05	0.10
Officer On-duty	0.86	0.75	.	0.90	0.94	0.95
Two+ Officers on Scene	0.29	0.22	0.37	0.28	0.33	0.41
Avg Officer Tenure	10.13	10.21	9.05	8.41	9.93	12.70
Panel D: Officer Response Reason						
Robbery	0.20	0.26	0.07	0.23	0.15	0.08
Violent Disturbance	0.29	0.25	0.15	0.33	0.29	0.34
Traffic	0.18	0.18	0.08	0.09	0.22	0.20
Personal Attack	0.04	0.07	0.00	0.02	0.01	0.04
Warrant	0.05	0.05	0.00	0.05	0.08	0.03
Suspicious Persons	0.07	0.05	0.05	0.06	0.06	0.12
Narcotics	0.06	0.05	0.05	0.07	0.06	0.04
Suicide	0.03	0.02	0.07	0.03	0.04	0.02
Other Response Reason	0.09	0.08	0.53	0.11	0.07	0.11
Panel E: Other Encounter Characteristics						
Daytime	0.37	0.35	0.38	0.38	0.43	0.39
Suspect Attacked or Drew Weapon	0.80	0.79	.	0.79	0.86	0.75
Panel F: Location						
Austin	0.05	0.00	.	0.25	0.00	0.00
Dallas	0.15	0.00	.	0.75	0.00	0.00

Houston	0.38	1.00	1.00	0.00	0.00	0.00
Jacksonville	0.03	0.00	.	0.00	0.11	0.00
Palm Beach County	0.06	0.00	.	0.00	0.23	0.00
Lee County	0.03	0.00	.	0.00	0.10	0.00
Brevard County	0.03	0.00	.	0.00	0.09	0.00
Pinellas County	0.03	0.00	.	0.00	0.11	0.00
Orange County	0.09	0.00	.	0.00	0.35	0.00
LA County	0.15	0.00	.	0.00	0.00	1.00
Panel G: Missing Variables						
Missing Race	0.02	0.04	0.38	0.00	0.02	0.00
Missing Sex	0.02	0.02	0.38	0.00	0.01	0.05
Missing Age	0.23	0.08	0.38	0.75	0.13	0.08
Missing Weapon	0.05	0.03	1.00	0.00	0.07	0.08
Missing Officer Race	0.16	0.38	0.00	0.00	0.06	0.00
Missing Officer Sex	0.06	0.12	0.00	0.00	0.04	0.00
Missing Officer Duty	0.01	0.00	1.00	0.00	0.02	0.06
Missing Num Officers	0.03	0.06	1.00	0.00	0.03	0.00
Missing Officer Tenure	0.20	0.37	0.00	0.01	0.20	0.02
Missing Response Reason	0.01	0.00	0.00	0.00	0.00	0.07
Missing Time of Day	0.36	0.01	0.00	0.75	0.75	0.00
Missing Suspect Behavior	0.00	0.00	1.00	0.00	0.00	0.00
Missing Injury Status	0.04	0.03	1.00	0.01	0.08	0.00
Observations	1,332	507	4,504	269	362	194

Notes: This table reports summary statistics. The sample consists of – (1) All officer involved shootings (OIS) between 2000 and 2015 (though exact time frame varies by location) from Austin, Dallas, six large Florida counties, Houston, and Los Angeles, and (2) Arrests in Houston from 2005 to 2015 during which an officer reported using his or her charged electronic device (taser). The first column includes the entire OIS sample. The second column includes OIS from Houston only. The third column includes arrests from Houston where a taser was discharged. The fourth column includes OIS from Austina and Dallas. The fifth column includes OIS from all Florida counties. The sixth column includes OIS from Los Angeles county.

Table 1D
Summary Statistics for Houston Police Department Arrests Data

	Full Sample	Non-Black/ Non-Hispanic	Black	Hispanic	*p-val*	*p-val*
	(1)	(2)	(3)	(4)	(2) = (3)	(2) = (4)
Panel A: Suspect Demographics						
Black	0.58	0.00	1.00	0.00	.	.
Hispanic	0.30	0.00	0.00	1.00	.	.
Non-black/non-hispanic	0.12	1.00	0.00	0.00	.	.
Male	0.82	0.79	0.83	0.82	0.325	0.478
Age	26.84	32.37	26.85	24.64	0.000	0.000
Panel B: Suspect Weapon						
Firearm	0.03	0.04	0.03	0.01	0.880	0.226
Sharp Object	0.01	0.01	0.02	0.00	0.716	0.488
Vehicle	0.00	0.00	0.00	0.00	0.647	0.534
None	0.95	0.94	0.93	0.97	0.855	0.190
Other Weapon	0.01	0.01	0.01	0.00	0.965	0.488
Panel C: Officer Characteristics						
Officer Unit Majority White	0.42	0.51	0.43	0.38	0.196	0.044
Officer Unit Majority Black	0.16	0.03	0.20	0.15	0.000	0.005
Officer Unit Majority Hisp	0.19	0.13	0.17	0.27	0.460	0.013
Officer Unit Majority Asian/Other	0.04	0.11	0.01	0.06	0.000	0.205
Officer Unit Split Race	0.18	0.22	0.19	0.14	0.504	0.101
Female Officers in Unit	0.10	0.14	0.10	0.08	0.246	0.171
Officer On-duty	0.87	0.86	0.84	0.92	0.781	0.093
Two+ Officers on Scene	0.66	0.63	0.67	0.66	0.483	0.620
Avg Officer Tenure	7.62	9.19	7.20	7.44	0.006	0.035
Panel D: Officer Response Reason						
Robbery	0.06	0.04	0.06	0.07	0.439	0.327
Violent Disturbance	0.21	0.22	0.19	0.26	0.524	0.460
Traffic	0.16	0.14	0.12	0.23	0.617	0.121
Personal Attack	0.01	0.01	0.00	0.01	0.455	0.838
Warrant	0.02	0.02	0.03	0.01	0.867	0.327
Suspicious Persons	0.25	0.24	0.30	0.16	0.255	0.104
Narcotics	0.07	0.04	0.10	0.03	0.057	0.721
Suicide	0.01	0.00	0.00	0.01	0.650	0.279
Other Response Reason	0.22	0.29	0.19	0.23	0.046	0.291
Panel E: Other Encounter Characteristics						
Daytime	0.48	0.47	0.53	0.40	0.427	0.399
Suspect Attacked or Drew Weapon	0.56	0.67	0.50	0.61	0.006	0.339
Panel F: Missing Variables						
Missing Race	0.32	0.00	0.00	0.00	.	.
Missing Sex	0.31	0.00	0.00	0.00	.	.

Missing Age	0.33	0.02	0.01	0.03	0.562	0.835
Missing Weapon	0.31	0.02	0.03	0.01	0.855	0.333
Missing Officer Race	0.35	0.10	0.08	0.07	0.752	0.473
Missing Officer Sex	0.34	0.06	0.06	0.06	0.927	0.916
Missing Officer Duty	0.30	0.01	0.00	0.00	0.029	0.111
Missing Num Officers	0.30	0.01	0.00	0.00	0.029	0.111
Missing Officer Tenure	0.36	0.07	0.09	0.08	0.640	0.711
Missing Response Reason	0.30	0.01	0.00	0.00	0.029	0.111
Missing Time of Day	0.63	0.39	0.47	0.46	0.211	0.295
Missing Suspect Behavior	0.31	0.04	0.02	0.02	0.475	0.559
Observations	1,024	84	402	213		

Notes: This table reports summary statistics. The sample consists of a random draw of arrests in Houston for the following offenses, from 2000 - 2015: aggravated assault on a peace officer, attempted capital murder of a peace officer, resisting arrest, evading arrest, and interfering in an arrest. The first column includes the entire sample. The second column includes non-black/non-hispanic civilians only. The third column includes black civilians only. The fourth column includes hispanic civilians only. The fifth column reports p-values for a ttest to see whether the mean for non-black/non-hispanic civilians is statistically similar to the mean for black civilians. The sixth column reports p-values for a ttest to see whether the mean for non-black/non-hispanic civilians is statistically similar to the mean for hispanic civilians.

Table 2A

Racial Differences in Non-Lethal Use of Force, NYC Stop Question and Frisk,Any Use of Force

	White Mean	Black	Hispanic	Asian	Other Race
	(1)	(2)	(3)	(4)	(5)
No Controls	0.153	1.534***	1.582***	1.044	1.392***
		(0.144)	(0.149)	(0.119)	(0.121)
+ Baseline Characteristics		1.480***	1.517***	1.010	1.346***
		(0.146)	(0.146)	(0.122)	(0.114)
+ Encounter Characteristics		1.655***	1.641***	1.059	1.452***
		(0.155)	(0.157)	(0.133)	(0.121)
+ Civilian Behavior		1.456***	1.513***	1.049	1.368***
		(0.128)	(0.136)	(0.124)	(0.107)
+ Fixed Effects		1.173***	1.120***	0.951	1.057**
		(0.034)	(0.026)	(0.033)	(0.028)
Observations			4,927,467		

Notes: This table reports odds ratios by running logistic regressions. The sample consists of all NYC stop and frisks from 2003-2013 with non-missing use of force data. The dependent variable is an indicator for whether the police reported using any force during a stop and frisk interaction. The omitted race is white, and the omitted ID type is other. The first column gives the unconditional average of stop and frisk interactions that reported any force being used for white civilians. Columns (2) through (5) report logistic estimates for black, hispanic, asian and other race civilians respectively. Each row corresponds to a different empirical specification. The first row includes solely racial group dummies. The second row adds controls for gender and a quadratic in age. The third row adds controls for whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or during a high crime time, whether the officer was in uniform, civilian ID type, and whether others were stopped during the interaction. The fourth row adds controls for civilian behavior. The fifth row adds precinct and year fixed effects. Each row includes missings in all variables. Standard errors clustered at the precinct level are reported in parentheses.

Table 2B

Racial Differences in Non-Lethal Use of Force, Police Public Contact Survey, Any Use of Force

	White Mean	Black	Hispanic	Other Race
	(1)	(2)	(3)	(4)
No Controls	0.008	3.335***	2.584***	1.047
		(0.349)	(0.299)	(0.262)
+ Baseline Characteristics		2.609***	1.633***	0.765
		(0.283)	(0.195)	(0.192)
+ Encounter Characteristics		2.582***	1.616***	0.794
		(0.286)	(0.194)	(0.198)
+ Civilian Behavior		2.653***	1.763***	0.758
		(0.314)	(0.221)	(0.195)
+ Year		2.697***	1.747***	0.755
		(0.321)	(0.226)	(0.197)
Observations		48,498		

Notes: This table reports odds ratios by running logistic regressions. The sample consists of all Police Public Contact Survey respondents from 1996 - 2011 with non-missing use of force data. The dependent variable is an indicator for whether the survey respondent reported any force being used in a contact with the police. The omitted race is white. The sixth column gives the unconditional average of contacts that reported any force being used by white civilians. Columns (1) - (4) report logistic estimates for black, hispanic, and other race civilians respectively. Each row corresponds to a different empirical specification. The first row includes solely racial dummies. The second row adds civilian gender, work, income, population size of civilian's address and a quadratic in age. The third row adds controls for contact time, contact type and officer race. The fourth row adds a civilian behavior dummy. The fifth row adds a control for year. Each row includes missing in all variables. Robust standard errors are reported in parentheses.

Table 3
Analysis of Subsamples, Any Use of Force, NYC Stop Question and Frisk

	White Mean	Coefficient on Black	Coefficient on Hispanic	Observations
Full Sample	0.153	1.173***	1.120***	4,927,467
Panel A: Crime Rate in Area				
High Crime	0.143	1.165***	1.115***	2,750,262
		(0.035)	(0.027)	
Low Crime	0.163	1.196***	1.136***	2,177,205
		(0.039)	(0.029)	
p-value		0.260	0.321	
Panel B: Time of Day				
Day	0.126	1.255***	1.162***	1,783,796
		(0.035)	(0.026)	
Night	0.170	1.136***	1.099***	3,143,671
		(0.039)	(0.029)	
p-value		0.001	0.020	
Panel C: Officer in Uniform				
Uniformed Officer	0.132	1.176***	1.124***	3,546,056
		(0.047)	(0.035)	
Non-Uniformed Officer	0.189	1.193***	1.121***	1,381,411
		(0.032)	(0.023)	
p-value		0.759	0.923	
Panel D: Location				
Indoors	0.144	1.125***	1.092***	1,129,443
		(0.044)	(0.033)	
Outdoors	0.154	1.184***	1.125***	3,798,024
		(0.031)	(0.025)	
p-value		0.101	0.252	
Panel E: Civilian Gender				
Male	0.160	1.170***	1.120***	4,446,921
		(0.034)	(0.026)	
Female	0.094	1.224***	1.103***	480,546
		(0.052)	(0.040)	
p-value		0.136	0.614	
Panel F: Eventual Outcomes				
Frisk	0.311	1.031	1.020	2,699,613
		(0.024)	(0.021)	
Search	0.411	1.054	1.039	409,255
		(0.037)	(0.030)	
Arrest	0.327	1.073**	1.038	291,109
		(0.034)	(0.025)	
Summons	0.195	1.150***	1.064*	304,580
		(0.044)	(0.035)	
Weapon/Contra-band found	0.359	1.111***	1.059***	136,894
		(0.025)	(0.024)	
p-value		0.003	0.394	

Notes: This table reports odds ratios by running logistic regressions. The sample consists of all NYC stop and frisks from 2003-2013 in which use of force and reported subgroup variables were non-missing. The dependent variable is whether any force was used during a stop and frisk interaction, with each panel presenting results from the indicated subgroups. We control for gender, a quadratic in age, civilian behavior, whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or during a high crime time, whether the officer was in uniform, civilian ID type, whether others were stopped during the interaction, and missings in all variables. Precinct and year fixed effects were included in all regressions. Standard errors clustered at the precinct level are reported in parentheses.

Table 4
Analysis of Subsamples, Any Use of Force, Police Public Contact Survey

	White Mean	Coefficient on Black	Coefficient on Hispanic	Observations
Full Sample	0.008	2.697***	1.747***	48,498
Panel A: Civilian Income				
$0 < 20,000$	0.012	2.942***	1.639**	11,881
		(0.532)	(0.332)	
$20,000 < 50,000$	0.010	1.983***	1.779**	11,115
		(0.515)	(0.411)	
$50,000+$	0.004	3.989***	1.874*	15,849
		(1.286)	(0.708)	
p-value		0.218	0.937	
Panel B: Civilian Gender				
Male	0.013	2.719***	1.815***	25,194
		(0.373)	(0.259)	
Female	0.004	2.599***	1.488	23,304
		(0.613)	(0.453)	
p-value		0.869	0.555	
Panel C: Officer Race				
Black/Hispanic	0.011	2.646**	3.886***	2,272
		(1.212)	(2.040)	
White	0.008	2.790***	1.743**	20,711
		(0.565)	(0.419)	
p-value		0.916	0.165	
Panel D: Time of Day				
Day	0.004	3.169***	2.178***	16,313
		(0.851)	(0.598)	
Night	0.012	1.678*	2.273***	7,656
		(0.476)	(0.591)	
p-value		0.104	0.910	

Notes: This table reports odds ratios by running logistic regressions. The sample consists of all Police Public Contact Survey respondents between 1996 to 2011 in which use of force and reported subgroup variables were non-missing. The dependent variable is whether any force was used during a contact, with each panel presenting results from the indicated subgroups. We control for civilian gender, a quadratic in age, work, income, population size of civilian's address, civilian behavior, contact time, contact type, officer race, year of survey and missings in all variables. Standard errors are robust and reported in parentheses.

Table 5
Racial Differences in Lethal Use of Force
Extensive Margin, Officer Involved Shootings

	Approx OIS With Narratives			Taser W/O Narratives		Full Sample W/O Narratives	
	Non-Black/ Non-Hispanic Mean	Black	Hispanic	Non-Black Mean	Black	Non-Black Mean	Black
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
No Controls	0.455	0.762	0.915	0.185	0.633^{***}	0.150	0.666^{***}
		(0.137)	(0.176)		(0.062)		(0.065)
+ Suspect Demographics		0.782	0.967		0.648^{***}		0.678^{***}
		(0.150)	(0.176)		(0.066)		(0.067)
+ Officer Demographics		0.779	1.113		0.728^{**}		0.746^{**}
		(0.191)	(0.294)		(0.094)		(0.087)
+ Encounter Characteristics		1.105	0.990		0.686^{***}		0.782^{*}
		(0.366)	(0.344)		(0.098)		(0.105)
+ Suspect Weapon		0.973	1.365		–		–
		(0.428)	(0.600)		(-)		(-)
+ Year		0.924	1.256		0.691^{**}		0.784^{*}
		(0.417)	(0.595)		(0.099)		(0.105)
Observations		1,531			5,011		6,035

Notes: This table reports odds ratios by running logistic regressions. The sample for each regression is displayed in the top row. For columns (1)-(3), the sample consists of all officer involved shootings in Houston from 2000 - 2015, plus a random draw of all arrests for the following offenses, from 2000 - 2015: aggravated assault on a peace officer, attempted capital murder of a peace officer, resisting arrest, evading arrest, and interfering in an arrest. These arrests contain narratives from police reports. For columns (4)-(5), the sample consists of all officer involved shootings in Houston from 2000 - 2015, plus a sample of arrests where tasers were used. These arrests do not contain narratives from police reports. For columns (6)-(7), the sample combines all officer involved shootings in Houston from 2000 - 2015, plus a random draw of all arrests for the following offenses, from 2000 - 2015: aggravated assault on a peace officer, attempted capital murder of a peace officer, resisting arrest, evading arrest, and interfering in an arrest, plus arrests where tasers were used. These arrests do not contain narratives from police reports. Data without narratives have no information on officer duty, civilian's attack on officer and civilian weapon. The dependent variable is whether the officer fired his gun during the encounter. The omitted race is non-blacks (with the exception of the sample with narratives where the omitted race is non-black/non-Hispanic). The first column for each sample gives the unconditional average of omiited race contacts that resulted in an officer firing his gun. The second column for each sample reports logistic estimates for black civilians. Each row corresponds to a different empirical specification. The first row includes solely racial dummies. The second row adds civilian gender and a quadratic in age. The third row adds controls for the split of races of officers present at the scene, whether any female officers were present, whether multiple officers were present and the average tenure of officers at the scene. The fourth row adds controls for the reason the officers were responding at the scene, whether the encounter happened during day time, and whether the civilian attacked or drew a weapon. The fifth row adds controls for the type of weapon the civilian was carrying. The sixth row adds year fixed effects for columns (1)-(3). It adds year as a categorical variable for columns (4)-(7). Each row includes missing in all variables. For arrest data without narratives missing indicators for officer gender, officer tenure, and number of officers on the scene were removed to minimize loss of observations in logistic regressions. For all regressions, missing indicator for response reason was removed for the same reason. Standard errors are robust and are reported in parentheses.

Table 6
Racial Differences in Lethal Use of Force
Intensive Margin, Officer Involved Shootings

	Non-Black/ Non-Hispanic Mean	Black	Hispanic
	(1)	(2)	(3)
No Controls	0.534	0.987	1.114
		(0.135)	(0.257)
+ Civilian Demographics		0.946	1.046
		(0.103)	(0.267)
+ Officer Demographics		0.835	0.896
		(0.094)	(0.221)
+ Encounter Characteristics		0.693^{***}	0.759
		(0.096)	(0.191)
+ Civilian Weapon		0.558^{***}	0.625^{**}
		(0.066)	(0.149)
+ Fixed Effects		0.526^{***}	0.564^{**}
		(0.037)	(0.131)
Observations		1,332	

Notes: This table reports odds ratios by running logistic regressions. The sample consists of officer involved shootings from Dallas, Austin, six Florida counties, Houston and Los Angeles between 2000 to 2015. The dependent variable is based on who attacked first. It is coded as 1 if the officer attacked the civilian first and 0 if the civilian attacked the officer first. The omitted race is non-blacks and non-hispanics. The first column gives the unconditional average of non-black/non-hispanic contacts that resulted in an officer firing his gun. The second column reports logistic estimates for black civilians. The third column reports logistic estimates for hispanic civilians. Each row corresponds to a different empirical specification. The first row includes solely racial dummies. The second row adds civilian gender and a quadratic in age. The third row adds controls for the split of races of officers present at the scene, whether any female officers were present, whether multiple officers were present and the average tenure of officers at the scene. The fourth row adds controls for the reason the officers were responding at the scene, whether the encounter happened during day time, and whether the civilian attacked or drew a weapon. The fifth row adds controls for the type of weapon the civilian was carrying. The sixth row adds city and year fixed effects. Each row includes missing in all variables. Standard errors are clustered at the police department level and are reported in parentheses.

Table 7
Any Use of Force, NYC Stop Question and Frisk

	Education Terciles	Income Terciles	Unemployment Terciles
	(1)	(2)	(3)
Tercile 1	1.174***	1.194***	1.113*
	(0.034)	(0.046)	(0.071)
N	2,275,062	1,941,664	1,472,862
Tercile 2	1.172***	1.100*	1.177***
	(0.070)	(0.060)	(0.044)
N	1,613,725	1,750,137	1,943,285
Tercile 3	1.180***	1.259***	1.246***
	(0.041)	(0.057)	(0.043)
N	1,030,997	1,227,983	1,503,637

Notes: This table reports odds ratios by running logistic regressions, subsampled for precinct demographics. Precinct demographics are calculated by collapsing data across census tracts received from the American Community Survey 2007-2011. For each column, we take the tract's white population demographic minus the black population demographic and collapse the means of the differences over precinct. We then take terciles in differences and calculate odds ratios for each tercile. Column (1) shows odds ratios across education terciles. Education is measured as the fraction of high school graduates in every census tract. Column (2) shows odds ratio across income terciles. Incomes is measured as the median household income. Column (3) shows odds ratios across unemployment terciles. Unemployment is calculated as the total number of unemployed people divided by the total number of people in the labor force. The sample consists of all NYC stop and frisks from 2003-2013 in which use of force and reported subgroup variables were non-missing. The dependent variable is whether any force was used during a stop and frisk interaction, with each panel presenting results from the indicated subgroups. We control for gender, a quadratic in age, civilian behavior, whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or during a high crime time, whether the officer was in uniform, civilian ID type, whether others were stopped during the interaction, and missings in all variables. Precinct and year fixed effects were included in all regressions. Standard errors clustered at the precinct level are reported in parentheses.

Table 8
Fraction weapon found, conditional on being
in an Officer Involved Shooting

	Civilian White	Civilian Black	*p-value*
	(1)	(2)	(3)
Officer White	0.851	0.810	
	(0.027)	(0.025)	0.268
Officer Black	0.625	0.742	
	(0.125)	(0.054)	0.354
p-value	0.010	0.115	

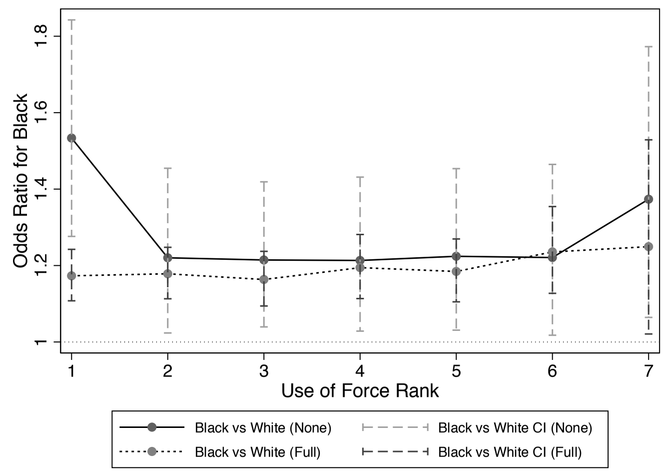
Notes: This table presents results for Anwar and Fang (2006) test. The first column presents the fraction of white civilians carrying weapons in the Officer Involved Shootings (OIS) dataset. The second column presents the fraction of black civilians carrying weapons in the OIS dataset. The third column displays the p-value for equality of means in columns (1) and (2). The first row presents the fractions when the majority of officers present during the encounter were white. The second row presents the fractions when the majority of officers present during the encounter were black.

Table 9
Weapon Found,
Conditional on Force Used

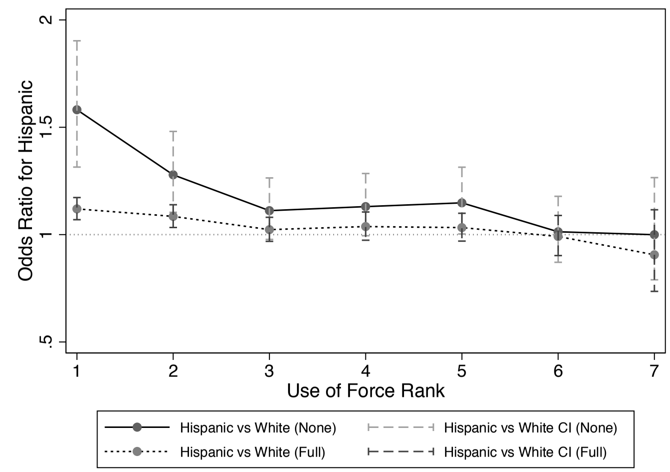
	White Mean	Coefficient on Black	Coefficient on Hispanic	Observations
	(1)	(2)	(3)	(4)
At Least Hands	0.036	-0.013^{***}	-0.008^{**}	1,028,625
		(0.004)	(0.003)	
At Least Pushing to Wall	0.036	-0.002	-0.000	253,643
		(0.002)	(0.002)	
At Least Using Handcuffs	0.040	-0.000	0.000	118,527
		(0.002)	(0.003)	
At Least Drawing a Weapon	0.053	0.003	0.001	58,443
		(0.004)	(0.004)	
At Least Pushing to Ground	0.054	0.005	0.002	51,083
		(0.004)	(0.005)	
At Least Pointing a Weapon	0.083	-0.011	-0.007	19,505
		(0.010)	(0.010)	
At Least Using Spray/Baton	0.092	-0.013	0.007	1,745
		(0.027)	(0.033)	

Notes: This table reports OLS estimates. The sample consists of all NYC stop and frisks from 2003-2013 in which use of force and outcome variable were non-missing. The dependent variable is a binary variable that is coded as 1 whenever a weapon was found on the civilian and 0 if weapon was not found. Each row looks at the fraction of white civilians carrying weapons and racial differences in carrying weapons for black civilians versus white civilians and hispanic civilians versus white civlians, conditional on at least a force level being used. We control for gender, a quadratic in age, civilian behavior, whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or during a high crime time, whether the officer was in uniform, civilian ID type, whether others were stopped during the interaction, and missings in all variables. Precinct and year fixed effects were included in all regressions. Standard errors clustered at the precinct level are reported in parentheses.

Figure 1: Odds Ratios by Use of Force, NYC Stop Question and Frisk



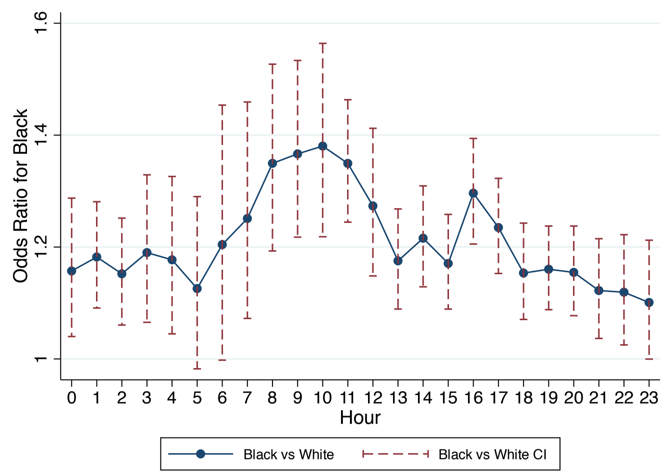
Panel A



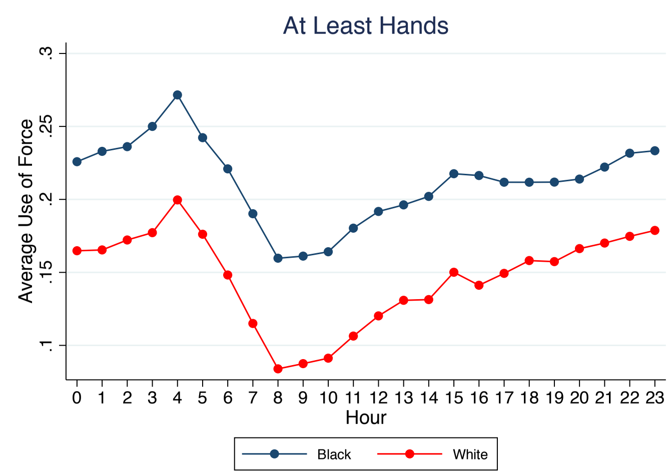
Panel B

Notes: These figures plot odds ratios with 95% confidence intervals from logistic regressions. For the figure on the left, the y-axis denotes the odds ratio of reporting various uses of force for black civilians versus white civilians. For the figure on the right, the y-axis denotes the odds ratio of reporting various uses of force for hispanic civilians versus white civilians. For both figures, the x-axis denotes different use of force types: 1 is an indicator for whether the police reported using at least hands or a more severe force on a civilian in a stop and frisk interaction. 2 is for whether the police reported at least pushing a civilian to a wall or using a more severe force. 3 is for whether the police reported at least using handcuffs or a more severe force. 4 is for whether the police reported at least drawing a weapon on a civilian or using a more severe force. 5 is for whether the police reported at least pushing a civilian to the ground or using a more severe force. 6 is for whether the police reported at least pointing a weapon at a civilian or using a more severe force. Finally, 7 is for whether the police reported at least using a pepper spray or a baton on a civilian. All force indicators are coded as 0 when the police report using no force in a stop and frisk interaction. The line plot with no controls is achieved by regressing the type of force (described above) on civilian race dummies only. The line plot with full controls is achieved by regressing the type of force on civilian race dummies, civilian gender, a quadratic in age, civilian behavior, whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or a high crime time, whether the officer was in uniform, civilian ID type, whether others were stopped during the interaction, and missings in all variables. Precinct and year fixed effects were included in the controlled regression. Standard errors are clustered at the precinct level.

Figure 2: Odds Ratios of Any Use of Force, NYC Stop Question and Frisk



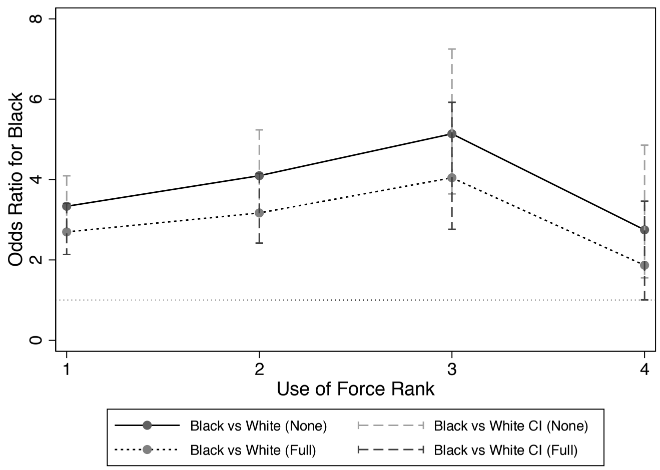
Panel A



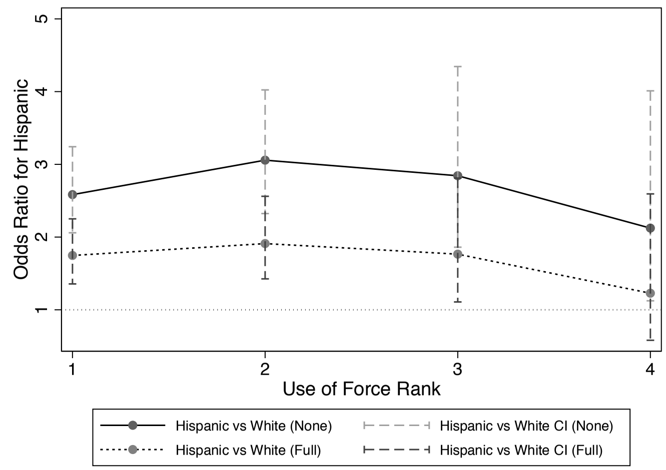
Panel B

Notes: These figures plot odds ratios with 95% confidence intervals from logistic regressions. For the figure in Panel A, the y-axis denotes the odds ratio of reporting any use of force for black civilians versus white civilians. For the figure in Panel B, the y-axis denotes the average fraction of white and black civlians who had any force used against them. For both figures, the x-axis denotes different hours of the day. For Panel A, odds ratios are achieved by regressing any use of force on civilian race dummies, civilian gender, a quadratic in age, civilian behavior, whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or a high crime time, whether the officer was in uniform, civilian ID type, whether others were stopped during the interaction, and missings in all variables, for every hour of day. Precinct and year fixed effects were included in all regressions. Standard errors are clustered at the precinct level.

Figure 3: Odds Ratios by Use of Force, Police Public Contact Survey



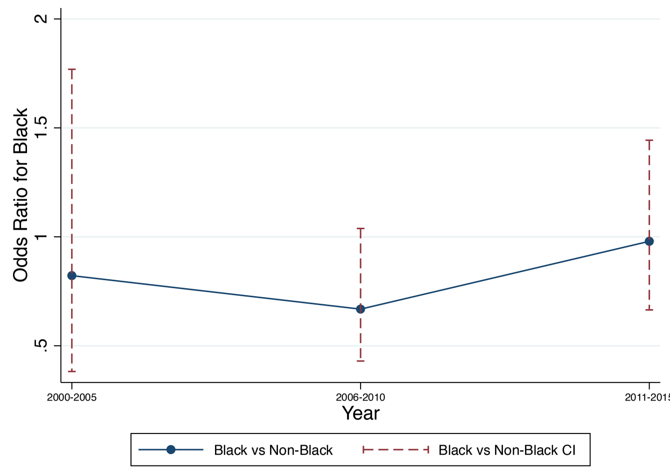
Panel A



Panel B

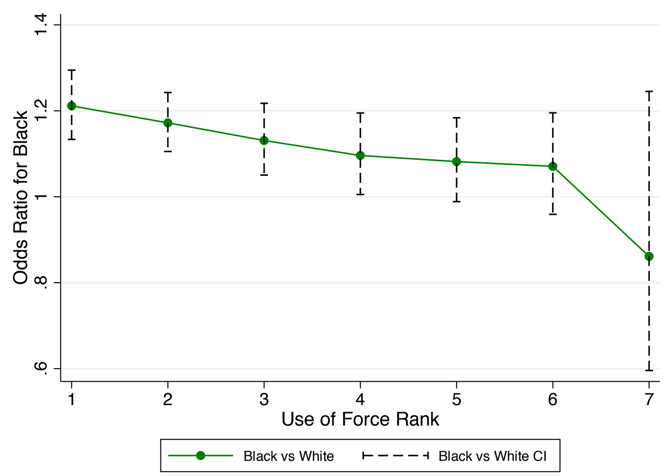
Notes: These figures plot odds ratios with 95% confidence intervals from logistic regressions. For the figure on the left, the y-axis denotes the odds ratio of reporting various uses of force for black civilians versus white civilians. For the figure on the right, the y-axis denotes the odds ratio of reporting various uses of force for hispanic civilians versus white civilians. For both figures, the x-axis denotes different use of force types: 1 is an indicator for whether the survey respondent report the officer at least grabbing him/her in an interaction. 2 is for whether the respondent reported the police handcuffing him/her or using a more sever force in an interaction. 3 is for whether the survey respondent reported the police pointing a gun at him/her or using a more severe force in an interaction. Finally, 4 is for whether the respondent reported the police kicking, using a stun gun or using a pepper spray on him/her or using a more severe force. All force indicators are coded as 0 when the respondent reports the police using no force in an interaction. The line plot with no controls is achieved by regressing the type of force (described above) on civilian race dummies only. We control for civilian gender, a quadratic in age, work, income, population size of civilian's address, civilian behavior, contact time, contact type, officer race, year of survey and missings in all variables. Standard errors are robust.

Figure 4: Odds Ratios for Officer Involved Shootings, Extensive Margin, By Year Categories

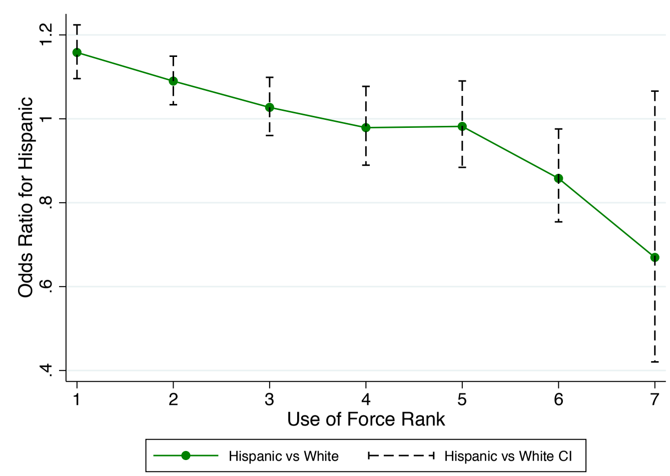


Notes: This figure plot odds ratios with 95% confidence intervals from logistic regressions. The sample consists of all officer involved shootings in Houston from 2000 - 2015, plus a random draw of all arrests for the following offenses, from 2000 - 2015: aggravated assault on a peace officer, attempted capital murder of a peace officer, resisting arrest, evading arrest, and interfering in an arrest ,plus a sample of arrests where tasers were used. The y-axis denotes odds ratios of an officer shooting at a black civlian versus a white civilian. The x-axis denotes the period of years for which the odds ratios were calculated. We control for civilian gender, a quadratic in age, officer demographics, encounter characteristics, and missings in all variables (i.e. all variables included in the final row of Table 5). Year fixed effects are included in all regressions. Robust standard errors are reported in parentheses.

Figure 5: Odds Ratios by Use of Force for Perfectly Compliant Civilians, NYC Stop Question and Frisk



Panel A



Panel B

Notes: These figures plot odds ratios with 95% confidence intervals from logistic regressions. For the figure on the left, the y-axis denotes the odds ratio of reporting various uses of force for perfecly compliant black civilians versus perfect compliant white civilians. For the figure on the right, the y-axis denotes the odds ratio of reporting various uses of force for perfectly compliant hispanic civilians versus perfectly compliant white civilians. For both figures, the x-axis denotes different use of force types: 1 is an indicator for whether the police reported using at least hands or a more severe force on a civilian in a stop and frisk interaction. 2 is for whether the police reported at least pushing a civilian to a wall or using a more severe force. 3 is for whether the police reported at least using handcuffs or a more severe force. 4 is for whether the police reported at least drawing a weapon on a civilian or using a more severe force. 5 is for whether the police reported at least pushing a civilian to the ground or using a more severe force. 6 is for whether the police reported at least pointing a weapon at a civilian or using a more severe force. Finally, 7 is for whether the police reported at least using a pepper spray or a baton on a civilian. All force indicators are coded as 0 when the police report using no force in a stop and frisk interaction. The line plot is achieved by regressing the type of force on civilian race dummies, civilian gender, a quadratic in age, civilian behavior, whether the stop was indoors or outdoors, whether the stop took place during the daytime, whether the stop took place in a high crime area or a high crime time, whether the officer was in uniform, civilian ID type, whether others were stopped during the interaction, and missings in all variables. Precinct and year fixed effects were included in all regressions. Standard errors are clustered at the precinct level.