Variation in Racial Disparities in Police Use of Force

Carl Lieberman*

May 13, 2022

Abstract

I examine how racial disparities in police use of force vary using new data covering every municipal police department in New Jersey. Along the intensive margin of force, I find disparities that disfavor Black subjects and are larger at higher force levels, even after adjusting for incident-level factors and using new techniques to limit selection bias. I then extend empirical Bayes methods to estimate department-specific racial disparities and observe significant differences across and within these hundreds of departments. Finally, I find that certain municipal factors are useful predictors of whether a department has a large racial disparity against Black civilians, but the most informative variables can change when considering different levels of force. These findings suggest that ignoring heterogeneity in police use of force misrepresents the problem and masks the existence of both departments with very large disparities and those without apparent disparities against Black civilians, but the variation even within departments may make identifying and treating inequitable departments difficult.

Keywords: policing, police use of force, race, racial disparities

JEL: J15, K42.

*Center for Economic Studies, U.S. Census Bureau. carl.lieberman@census.gov. https://carl.ac.

Racial inequities in police use of force are among the most significant social issues in the United States, with numerous calls for change but little widespread reform. Yet policing is not monolithic: it is a public good typically provided at the local level, and thus the extent to which there are racial problems, or the solutions to those problems, may vary across across departments. Moreover, these disparities might change as we consider the different levels of force intensity officers use, from restraining the subject to discharging a firearm. One municipality's department might use excessive force against Black subjects concentrated at the highest levels of force, while a neighbor's department could have disparities concentrated on lower levels of force, and another may not have any apparent racial differences. To what extent is this true, what factors distinguish these departments, and what does this mean for identifying areas for policing reform?

In this paper, I explore three fundamental aspects of how racial disparities in police use of force vary to better understand these problems. First, I start by exploring aggregated data to examine key estimands of race and policing in my setting: how large are overall racial disparities in police use of force, and how do they change along the spectrum of force severity, from restraining the subject to using deadly force? Second, I consider variation in these disparities by department and place: how do these racial differences vary across and within departments or localities? And third, I focus on identifying possible links to racially inequitable policing: what characteristics of the departments or their municipalities are associated with these department-specific racial disparities?

I answer these questions using novel administrative data from New Jersey. These data offer an unprecedented level of breadth and depth in use of force data in a unified setting, containing every use of force report by every officer in every municipal police department, as well as the state police, over the five-year span from 2012 through 2016. Notably, the data include the entire force spectrum, from restraining the subject with a compliance hold to discharging a firearm. Moreover, because the data cover the entire state, they feature substantial variation in locations: New Jersey has more than 450 police municipal departments serving more than 560 municipalities, and these locations range from small, racially homogeneous towns to dense, diverse cities. This presents a significant advantage over data used in prior research that are often based on a handful of large urban departments willing (or required by the legal system) and able to share their use of force data, enabling not only more credible and representative estimates of racial disparities, but also new measures of how they vary.

For the initial question of how overall racial disparities vary over the force spectrum, I focus on the intensive margin of force: conditional on force being used and adjusting for event-level observables such as incident type and subject actions, do people of different races have different levels of force used against them? This parameter is relevant both in the empirical literature and in public policy, reflective of the higher levels of force experienced by Black civilians in my and others' data analyses and in numerous high-profile incidents of police violence recently and historically. I find significant disparities in police use of force across races that are larger at more severe levels of force. Compared to outcome means, Black subjects are 3% more likely than people of other races to have force more severe than compliance holds, the lowest level of force, used against them, and 19% more likely when looking at enhanced types of force: pepper spray, batons, canines, stun guns, and firearms.

In an effort to better understand the causes and correlates of these gaps, I explore several mechanisms that may drive the disparities. To test whether the disparities might be caused by taste-based or statistical discrimination (police officers are not allowed to engage in either in this setting, but policy solutions to each may be different), I focus on incidents involving male officers and Black female subjects, where statistical discrimination based on the belief that Black subjects are more threatening (whether accurate or not) could be mitigated by lower levels of force required to restrain a female subject. However, I observe no such interaction, which is more consistent with taste-based discrimination. I also observe no time trends in these data that would suggest that increased public scrutiny or filming of police (chilling effects; note that police officer Darren Wilson shot and killed Michael Brown in Ferguson, Missouri during the middle of my sample period) impacted my results. Finally, looking at a set of officers who switch departments, I find that officers' levels of force against Black subjects increase or decrease commensurate with whether their department has a disparity above or below the median, suggesting that the departments themselves (or their municipalities) play a significant role.

Estimating disparities in police use of force is empirically difficult, and I take several steps to address the possibility of bias from selection into my sample. First, building on the techniques of Grogger and Ridgeway (2006) and Horrace and Rohlin (2016), I identify a set of incidents that plausibly reduces the effects of race outside of the decision of what level of force to use, in particular the officer's decisions of whom to investigate or engage with and whether they should use force at all. These papers are built around the "veil of darkness" hypothesis, arguing that in poor lighting such as nighttime traffic stops, a subject's race is harder to observe, and hence less likely to factor into an officer's decision to stop a vehicle. I extend this method by restricting the dataset to "always stop" incidents, those in which the subject likely would have been stopped regardless of their race, and "always force" incidents where the outside of option of not using any force is less viable. Together, these restrictions should limit the role of race in the officers' extensive margin force usage while preserving interactions between race and the intensity of force used. Second, I employ a Heckman correction-style estimator adapted from Goncalves and Mello (2017) that can account for differences in the "first stage" of each department's race-specific propensity to use force at all by identifying the expectation of the bias introduced from selection into my sample. The results from both checks suggest that selection bias prior to observation is not problematic in my setting and support my main estimates.

Next, I explore how racial disparities vary across and within departments (or, almost equivalently, municipalities), leveraging New Jersey's large number of independently operated police departments. To better conduct this analysis, I adapt empirical Bayes techniques to estimate departmental group-specific differences. Using this new procedure, I find that focusing on overall disparities masks significant variation in these disparities across departments that often dwarfs the full-sample magnitudes. A minority of departments, ranging from 19% to 36% depending on the force level considered, do not have estimated racial disparities that disfavor Blacks for a given force level, but there are also long tails of departments where Black subjects face disparities significantly larger than the overall estimates. For example, enhanced force is used in 15% of all force incidents, and the 95th percentile department is more than six percentage points more likely to use enhanced force against a Black subject than a non-Black one conditional on force being used, about double the disparity of the median department. I also find that while the presence of a disparity in force usage against Black subjects at one level of force severity makes a disparity at another force level much more likely, many departments do not seem to have disparities at all levels of force, suggesting practical difficulties in identifying and treating departments engaging in racialized policing.

Finally, I treat the presence of racial disparities as the outcomes of interest themselves and explore whether departmental or municipality characteristics such as officer diversity, economic inequality, crime, and political preferences are able to predict whether a department will have no estimated disparity against Blacks or an especially large one. Through random forest and penalized logit machine learning models, I find that certain municipal characteristics such as household income, the Black share of the population, the violent crime rate, and Mitt Romney's proportion of the 2012 presidential election vote are the most valuable predictors, while traits of the departments are not as informative. When looking at the smaller subset of departments present in the 2016 Law Enforcement Management and Administrative Statistics (LEMAS) survey, the cultural and policy measures available there also tend to be outperformed by municipal characteristics, but this analysis is greatly limited by sample size, and it is conceivable that more idiosyncratic characteristics of a department or its officers that would not be reflected in the LEMAS could play a role. But even if harder-to-measure factors such as culture are indeed playing an important role in disparity magnitudes across departments, the variation present within departments still complicates efforts to find a handful of factors that universally predict racially inequitable policing.

Although there is relatively little work specifically about variation in racial differences in police use of force instead of overall gaps, economists and other social scientists have long grappled with the question of how race interacts with the criminal justice system. Recent efforts have made progress on this methodologically thorny issue using new data or empirical strategies to obtain better estimates. Nix et al. (2017), for example, examine fatal police shootings using new data collected by *The Washington Post* and find racial differences in whether a shot subject was attacking the officer and whether they were armed. Weisburst (2019) looks at data from the Dallas Police Department and finds that Black civilians are disproportionately likely to be involved in an incident involving any level of force, stemming from differences in likelihood of arrest. Fryer (2019) estimates overall disparities by combining several different data sources, including New York City's Stop, Question, and Frisk program and officer-involved shootings from 16 police departments. He finds racial disparities in police use of force for Blacks relative to Whites to be consistent over the nonlethal force spectrum with odds ratios of around 1.2, while Hispanic-White disparities decrease as the force level increases; he also controversially finds no disparities for police shootings. Knox, Lowe, and Mummolo (2020) argue that the estimates from Fryer (2019) understate the causal effect of race by not accounting for racial differences in police-civilian interactions prior to the encounter, such as patrolling habits and the decision of whether to stop a subject.[1] Many other papers explore related aspects of race and the criminal justice system.[2] This paper revisits some of these key estimates with new data and methods, but differs from prior works by focusing on *variation* in racial differences, specifically by force severity and departments or municipalities, rather than treating race as having a single, homogeneous effect.

There is also a growing interdisciplinary literature on variation in racial disparities in police use of force that asks what can predict or explain that variation, offering different answers. Ross (2015), in the project most similar to this aspect of my work, conducts an analysis of county-level police shootings with local racial populations as the benchmark, finding significant variation in risk ratios at this level of aggregation associated with "larger metropolitan counties with low median incomes and a sizable portion of Black residents, especially when there is high financial inequality." Nicholson-Crotty, Nicholson-Crotty, and Fernandez (2017), using data on police homicides collected by the Mapping Police Violence organization and *The Washington Post*, suggest that there may be a critical mass for increased racial diversity of police officers to reduce police killings of Blacks, below which there will be little impact. Hoekstra and Sloan (2020) exploit plausibly exogenous variation from 911 dispatch assignment rules from two

[1] See also comments by Durlauf and Heckman (2020) and the response by Fryer (2020).

[2] Examples include police stops and searches (Knowles, Persico, and Todd, 2001; Anwar and Fang, 2006; Persico and Todd, 2006; Antonovics and Knight, 2009; Coviello and Persico, 2015; West, 2018), speeding tickets (Goncalves and Mello, 2017), police reform after racially unjust incidents (Heaton, 2010; Devi and Fryer, 2020; Luh, 2020), whether subjects are formally booked (Raphael and Rozo, 2019), court decisions (Rehavi and Starr, 2014; Arnold, Dobbie, and Yang, 2018; Bielen, Marneffe, and Mocan, 2019; Tuttle, 2021), and student discipline (Barrett et al., 2019).

cities and observe that White officers increase their use of force more when dispatched to more minority neighborhoods than minority officers do, also finding that racial differences are more severe for firearms than other force types. Similarly, Ba et al. (2021), in exploring the possible effects of officer diversification based on data from patrol assignments in Chicago, note that Black and Hispanic officers use force less often than White officers against Black civilians and that these differences are larger in areas with more Black residents. Cunningham and Gillezeau (2021) and Cunningham, Feir, and Gillezeau (2021) use variation stemming from Black protests and uprisings during the second half of the 20th century (the former) and the introduction of duty to bargain requirements with police unions (the latter), with both finding significant increases police killings of non-White civilians relative to Whites. The variation present in my setting and the new empirical Bayes estimator are key to this paper's value-added, as they allow me to explore how racial differences vary across and within departments in a unified, expansive setting with detailed microdata on every level of force and hundreds of independently operated police departments.

In identifying these variations in racial disparities, I move forward our understanding of race and policing. Beyond simply observing heterogeneity in police use of force, the variation present in department-level disparities underscores the difficulties in remedying the differences and how a one-size-fits-all intervention may not be optimal (or that there may be gains from more targeted or adaptable solutions). Some departments do not appear to have racial disparities against Blacks, but many more do, some very severely, and the identities of these departments change across force levels. Large amounts of unexplained variation in these differences, even with a rich set of predictors, coupled with other estimates emphasizing the importance of the department, may also suggest that unobserved factors play an important role in determining how race and police use of force interact. These issues are important to understand on their own, but they also contribute to a much needed holistic understanding of race and public policy. Consider, for example, the combined impacts of disparities in police use of force and bail or sentencing in the courts (see footnote 2), or how higher levels of force against Black civilians could affect the efficiency of a police department given evidence of decreased public cooperation and engagement after high-profile incidents of police violence (Ang et al., 2021).

Identifying and addressing racial inequities in police use of force is a vital step, but it is one on a much longer journey.

1 Institutional Background and Data

1.1 Use of Force and Force Reporting in New Jersey

In 2001, the New Jersey Attorney General's Office began to require that police officers document all incidents in which they use force. Appendix Figure A.1 shows a model form for these force reports. Plans for a centralized system for reporting, oversight, and analysis of these reports by the state did not materialize, and most of these reports ended up in storage, unused and inaccessible by the public (NJ Advance Media, 2019).[3]

Police in New Jersey are authorized to use several types of force in the line of duty, and these standard force types can be ordered along a spectrum of severity. The lowest level of force is a compliance hold, which uses pressure points to restrain the subject and includes actions such as arm bars and wrist locks. Next are unenhanced types of force: takedowns (forcing a subject to the ground, including tackling them or otherwise knocking them down), hands/fists (punches and slaps), and kicks or other leg strikes. Then there are enhanced force outcomes like pepper spray and other chemical agents, baton strikes, canines, and stun guns (Tasers). Finally, as a subset of enhanced force, officers may use deadly force and discharge firearms. Constructive authority, the threat of force without its actual use, such as brandishing a firearm, is permitted, but warning shots are prohibited. Stun guns are mostly absent from the data due to state regulations against their use, though departments began rolling them out after a change in the law in 2016. Canines are also uncommon, as not all departments have canine units. Police may very rarely employ nonlethal firearm ammunition like beanbag rounds in settings such as riots to incapacitate subjects; I drop these incidents, as the situations in which they are used are not representative of typical police-civilian interactions.

[3] In the wake of the murder of George Floyd by police in Minneapolis, the New Jersey Attorney General announced that a statewide use of force portal would be ready later in 2020; it launched in April 2021.

1.2 Data

This project uses all known force reports from every municipal police department in New Jersey and the New Jersey State Police from 2012 through 2016. ProPublica, a nonprofit newsroom, and NJ Advance Media (NJAM), a news provider that operates NJ.com and *The Star Ledger* newspaper, obtained the reports through more than 500 public records requests and several legal threats. Following substantial data entry and cleaning, ProPublica and NJAM made the resulting dataset available in January 2019.

These reports contain rich incident-level information. Among other variables, the reports record the time, date, and location of the incident, the nature of the incident (such as a crime in progress or traffic stop), the officer(s) involved and their demographics, the subject(s) involved, their demographics, and their actions that led to force being used, the types of force an officer used against the subject(s), and whether an officer or subject was injured or killed.

The final dataset from ProPublica, NJAM, and their partner data entry firm requires additional processing before it can be used in my analyses. For example, some types of force used by police are not neatly categorized, and instead consist of irregular descriptions such as "grabbed rock out of her hand" or "in foot pursuit grabbed suspect left hand." Appendix A documents in detail how I clean and process the data into a consistent format. I structure the data so that each observation represents one subject who had force used against them by one officer in an incident. For cases where multiple officers use force against a single subject, I keep the officer who uses the greatest level of force, choosing randomly in the event of ties.[4] I remove 44 subjects whose races recorded on the force reports do not fall within the categories of White, Black, Hispanic, or Asian/Pacific Islander, such as people marked as "mixed." After cleaning, there are 39,322 incidents that I use in my analysis sample, with the exact number varying slightly depending on the model used due to 197 observations missing exact date or officer identifier information.

I supplement the force reports with data on New Jersey's police departments and municipal-

[4]There are several justifiable ways to structure the data. I choose this format, as it is the most natural for my empirical strategy in which the outcome of each incident is based on the greatest level of force used. When multiple officers use force, it is then sensible to keep the officer who used the greatest level of force, particularly because other officers may only be able to use lower levels of force due to the officer using a more severe type (for example, after Officer A pepper sprays a subject, Officer B is able to place him in a compliance hold).

ities. Characteristics of the police departments themselves, such as racial diversity, come from ProPublica and NJAM. Local variables come from the 2010 Decennial Census, 2012-2016 American Community Survey (ACS) five-year estimates (Ruggles et al., 2020), the FBI's Uniform Crime Reporting (UCR) program, the New Jersey Division of Elections, and Harvard's Project Implicit. Additional data on a smaller sample of departments comes from the 2016 LEMAS survey conducted by the Department of Justice (Bureau of Justice Statistics, 2016).

1.3 Summary Statistics and Stylized Facts

Table 1 presents summary statistics for the cleaned force reports. It includes the most extreme type of force used by police in each incident, the most extreme action the subject took that may have prompted the force, whether an officer was injured, the type of incident, and subject demographics.

The frequency with which a force level is used as the most severe type in an incident decreases with its severity. In 50% of force incidents, officers use only compliance holds, the lowest level of force. 36% of incidents have unenhanced force types such as punches and kicks as the highest level. 14% of incidents have non-deadly enhanced force, i.e. mechanically enhanced force excluding firearms, as their most severe force type. Officers discharge firearms in under half a percent of observations, representing over 160 shootings between 2012 and 2016.

Subject actions are concentrated on the lower end of the spectrum. Resisting an officer, which I also use as a catchall for non-missing actions that do not fit elsewhere, is the most extreme subject behavior in 63% of force incidents. Physically threatening or attacking an officer, such as punching, kicking, or spitting, follows with 34% of observations. The remaining 3% of the data are divided approximately evenly between blunt weapon, knife, and vehicular threats/attacks, as well as threats with firearms. 0.1% of incidents involve a subject actually discharging a firearm. 10% of force incidents result in the officer being injured.

Incident types are varied. Following the model force report from the New Jersey Attorney General, I include indicators for crimes in progress, domestic disputes, other disputes, suspicious persons, traffic stops, and other incidents. An incident may have multiple types, except for the "other" category, which I reserve for incidents that are not classified as belonging to any other category. Consistent with the sample being only incidents in which officers used force, the most common classification is crime in progress, making up more than a quarter of the data. Domestic disputes are the next most frequent at 13% of the sample. Other disputes and suspicious person cases are each 11% of the sample, traffic stops are 9%, and 33% of incidents fall into the "other" category.

The extensiveness of these force data make for an excellent setting to examine fundamental facts about how often police use force. Figure 1 looks at each department's propensity to use force by normalizing the total number of force incidents by arrests. Across all departments, the total number of incidents per arrest resembles a log-normal distribution, slightly left-skewed, with enormous variation. The 90-10 ratio of force incidents per arrest is 5.3 (i.e., the 90th percentile department has 5.3 times as many force incidents per arrest as the 10th percentile department), and the ratio of the maximum to the minimum is 88.0. This variation is especially notable for demonstrating how much departments vary in their force frequencies even after adjusting for arrests. Breaking down the total force incidents per arrest into (non-)Black force incidents per (non-)Black arrest shows that the two distributions are fairly similar. The Black distribution is flatter and more spread out, and the tails are further to the right (meaning that the most extreme departments in either direction use force at a higher rate against Black civilians than non-Black ones). On the whole, Black and non-Black extensive margin force usage are similar in the aggregate, but with greater variation in the Black distribution.

To better understand how and where Black civilians are represented in the force data, Figure 2 compares the proportions of the population ages 18-65 and the force subjects who are Black for each municipality. The results are striking: there is a clear positive relationship, but it is nonlinear. Areas with the lowest percentages of Black residents not only have disproportionately high shares of Black force subjects, but also the steepest marginal increase in those outcome shares as the Black population increases. This relationship flattens somewhat when the Black population percentage is around 30%. Although almost all departments lie above the 45-degree line, and areas with the most Black residents have the most Black force subjects, it is clear that Blacks are most overrepresented in the force statistics in the areas with the fewest Black residents. The slopes of the fit line being greater than one at this point on the curve imply that marginally increasing the Black population in an area with few Black residents is associated with an increase in their share of police violence of more than that increase, an alarming (but non-causal) relationship.

Table 2 breaks down correlates of race and policing further. Column 1 regresses the share of Black force subjects against some basic municipality and departmental characteristics. Incorporating population size and density, violent crime rates, and the size of the police force does not change the previously observed curved relationship, now parameterized by a quadratic on share of the 18-65 population that is Black, and indeed none of the new variables is statistically significant. Looking instead at the (logged) number of incidents rather than their racial makeup in column 2, it is interesting to note that while the aforementioned covariates now all have positive correlations with this outcome, the same curved relationship with the Black population proportion persists. Again, the number of incidents in which a department uses force rises disproportionately with the share of Black residents when starting at lower shares. These results are not causal, and causality likely runs not only between these factors, but also between other closely linked ones, but they do serve to highlight how closely race and police use of force are linked. Finally, the third column of Table 2 regresses the number of force incidents per arrest for each department on the same set of variables. Under the assumption that arrests constitute an appropriate "denominator" to represent the universe of incidents where force might be used, this exercise can be interpreted as looking at factors related to departmental differences in the extensive margin of whether force is used at all. Unlike before, there is no relationship between the racial shares of the local population and the department's propensity to use force, while higher population densities and violent crime rates are both correlated with greater incident-arrest ratios.

Turning attention to statistics on the subjects who have force used against them, demographics are not at all representative of New Jersey's population. 48% of the subjects in the sample are White, 41% are Black, 10% are Hispanic, and 1% are Asian/Pacific Islander. From the ACS five-year estimates for 2012-2016, 57% of the state's population at the time was non-Hispanic White and 9% was non-Hispanic Asian/Pacific Islander, making these two groups underrepresented in the force reports. Blacks are severely overrepresented, with non-Hispanic

Blacks comprising only 13% of the population. Due to data limitations, I treat Hispanic status as a distinct racial category; one cannot be Hispanic and another race in the data.[5] 19% of the state population was Hispanic of any race, but it is impossible to know whether they are overrepresented or underrepresented in the force data, as the criteria used to classify individuals as Hispanic in the force reports do not match the race and ethnicity definitions in the ACS. Only 20% of subjects are female, and the average subject is 31 years old.

Table A.2 contains summary statistics for the police departments in the data. Over the five-year span 2012-2016, the median department has 23 full-time employees, the average has almost 43, and the maximum has over 1,000. Racial diversity is low: the median values for a department's racial share of each group are 90% for Whites and 0% for each of Blacks, Hispanics, and Asians/Pacific Islanders (not all departments' racial breakdowns sum to 100% due to inconsistencies in reporting). Arrest and force incident distributions are both highly right-skewed, and arrests are orders of magnitude more common. The median and mean number of arrests over these years are 1,327 and 2,959, but the analogous numbers for force incidents are only 34 and 86.

Table A.1 presents summary statistics for the municipalities with their own police departments that are present in the cleaned data. For New Jersey's more than 560 municipalities, 461 police departments are present in the uncleaned data. A handful of small municipalities with a combined 58 force reports do not appear in my subsequent analyses because their reports are missing data in key variables. These hundreds of municipalities offer significant heterogeneity over margins such as race, income, size, and political preferences over which policing might vary. Populations for municipalities in the sample range from the hundreds to the hundreds of thousands. As of the 2010 Census, New Jersey has seven of the 10 densest incorporated places in the United States and is overall the densest state in the country, though this too varies greatly within the state. New Jersey is among the top states by median income, but there are significant areas of poverty and it has high economic inequality as measured by the Gini coefficient relative to other states. Although New Jersey is mostly White, many areas have barely any

[5] Although some departments' force reports do distinguish between a subject's race and Hispanic status, many do not, and the cleaned version of the race variable in the dataset made available to researchers treats Hispanic status as a race. Note that officers use their own judgment when recording a subject's race.

Whites while others are almost exclusively White. Violent crime rates range massively, from 0.03 to 55 crimes per 1000 people. Political preferences, as proxied by Mitt Romney's share of the vote in the 2012 presidential election, are consistent with the state leaning Democratic overall despite the presence of more conservative areas. County-level Black-White implicit association test (IAT) *D*-scores (normalized summary statistics) from Project Implicit, where more positive values indicate stronger implicit biases against Blacks, show large variation: the standard deviation of 0.05 is approximately equal to the difference between the median and 70th percentile counties at the national level.

2 Empirical Strategy

2.1 Model Specifications

2.1.1 Overall Disparities

For the question of overall racial disparities, my primary econometric specification estimates the following equation via ordinary least squares (OLS):

$$Force_{iopt} = \beta \cdot \text{Black}_i + X'_{iopt}\gamma + \psi_p + \nu_t + \varepsilon_{iopt} \tag{1}$$

where the subscripts i, o, p, and t denote the incident-subject pair, officer, department, and year, respectively. *Force* is a binary measure of whether the level of the force used in an incident was at least of a certain severity. I consider four levels of the force intensity outcome, ordered from least severe to most severe: compliance holds, unenhanced force, enhanced force, and firing a weapon (which is a subset of enhanced force types). For example, if a subject is struck with a baton, the outcomes for at least compliance holds, at least unenhanced force, and enhanced force would be one, but would be zero for fired weapons. Because compliance holds are the lowest level of force, the at least compliance hold measure is always one, and I omit it as an outcome. As a result, the three outcomes in my analyses are at least unenhanced force (was force greater than compliance holds used?), enhanced force (was any type of enhanced force used, including firearms?), and fired weapons (did an officer discharge a firearm?). I use this binary measure of force because it is the most intuitive way to interpret outcomes. Alternative parameterizations of force intensity such as an indicator for the minimum force level used incompletely describe an event, and analyzing every force level separately as when only looking at only the maximum can result in estimates without clear interpretations (Fryer, 2019). The coefficient of interest is β, the difference in the observed probabilities of Black subjects having more severe types of force used against them conditional on force being used at all and relative to people of other races and after adjusting for incident-level factors. X is a vector of incident characteristics including time, type of incident, officer rank, subject behaviors, subject sex, and a quadratic of the subject's age. Incidents may have multiple types, and I include indicators for each unique combination of types rather than each individual type. Department fixed effects ψ_p capture time-invariant aspects of each department's propensity to use higher levels of force, such as the severity of crimes within their jurisdiction. Time fixed effects ν_t adjust for year-month specific changes in overall force usage and capture seasonality in crime. I cluster standard errors at the department level.

Although I am able to identify many department and officer fixed effects simultaneously due to officers switching departments, I cannot identify all such effects simultaneously, and smaller departments are especially impacted. Instead, I estimate an additional series of models using officer fixed effects instead of department fixed effects with clustering at the officer level. These estimates allow me to investigate how disparities vary when adjusting for the officer's identity, which may be useful if there is significant heterogeneity in officers' severity of force within departments. Further, comparing these results to the ones using department fixed effects allows me to understand better the role of the individual officer as opposed to the department.[6]

The linear probability models above are appropriate for estimating the marginal "effect" of race, but one may be concerned about their limitations and assumptions, especially for rarer events such as shootings. As a robustness check, in addition to the OLS estimates from Equa-

[6]The officer-based models could hypothetically be sensitive to the data cleaning process in which, for each incident where multiple officers use the most severe level of force in that incident, I keep one randomly chosen force report. However, this is an edge case, and results with a different draw are similar.

tion 1, I estimate conditional logit models of the following form:

$$\ln\left(\frac{Pr(Force_{iopt}=1)}{1-Pr(Force_{iopt}=1)}\right)=\beta\cdot\text{Subject Black}_i+X'_{iopt}\gamma+\psi_p+\nu_t+\varepsilon_{iopt} \tag{2}$$

stratified by department. To prevent separation, I do not include officer fixed effects, use year fixed effects instead of year-month, use event type indicators for each individual event type rather than each combination of event types, and change the officer rank indicators to a dummy for the officer having a superior rank or not. By exponentiating β, I obtain odds ratios for Black subjects compared to others. Logit-based estimators offer several advantages over OLS, in particular probabilities bounded between zero and one, and conditional logit further addresses issues surrounding the inconsistency of logits with numerous fixed effects. However, as I move up the force spectrum, separation becomes a larger concern. More departments will have only zeros in the outcome, and these observations must be completely dropped from the analysis.

2.1.2 Department-Specific Disparities

Next, I examine variation in racial disparities across and within departments. The simplest approach would be to make a slight modification to Equation 1 to estimate equations of the form

$$\text{Force}_{iopt}=\beta_p\cdot\text{Black}_i\times\text{Department}_p+X'_{iopt}\gamma+\psi_p+\nu_t+\varepsilon_{iopt} \tag{3}$$

with the interest being in the distribution of β_p. Although these estimates may be unbiased and consistent, in general, OLS will tend to generate the most extreme point estimates for departments with the fewest observations. I propose an alternative approach that can better handle estimates from departments with many observations and those from much smaller samples.

To improve upon OLS for identifying these hundreds of disparities, I modify empirical Bayes estimators, used most frequently by economists for estimating teacher value-added (see Kane and Staiger, 2008; Chetty, Friedman, and Rockoff, 2014, among others), to estimate these group-specific differences for each department. Empirical Bayes estimators use the overall distribution of estimates to inform each individual point estimate, with a prior distribution calibrated on the observed data instead of being fixed in advance as with more traditional Bayesian methods. Although these techniques are not new (Morris, 1983), there is almost no work on extending this estimator to a setting where we are interested in estimates of a "treatment" on only a subset of the population, like a department's propensity to use more intense force against Black subjects relative to others, as opposed to overall effects (see related work by Kline and Walters, 2019 in the context of audit studies of discrimination).

As is standard practice, I start by fitting a normal distribution for the prior and apply Bayesian updating to obtain a posterior distribution for each department's point estimate. Less reliable estimates, such as those from departments with few observations, are shifted more towards the population mean, resulting in a "shrinkage" estimator. After applying the updating procedure, I record the centers of the posterior distributions and use them as my estimates of β_p.

To conduct these estimates, for each level of force, I begin by running the following "pooled" regression. Note that there is only one overall coefficient for the subject being Black, and so all departments are effectively pooled together.

$$\text{Force}_{iopt} = \beta_0 \cdot \text{Black}_i + X'_{iopt}\gamma + \psi_p + \nu_t + u_{iopt} \tag{4}$$

I subsequently use the following distribution as the prior

$$\beta_p \sim N(\hat{\beta}_0, \sigma_p^2)$$

where $\sigma_p^2 \equiv \text{Var}(u_{iopt}^B - \varepsilon_{iopt}^B)$, the variance of the difference of residuals between Equations 3 and 4 using only Black observations rather than the full sample. I then take the estimated racial disparities β_p from Equation 3 and compute the empirical Bayes estimates, which are effectively weighted averages between the prior mean and the department estimate.

$$\hat{\beta}_{p,EB} = w_p \cdot \hat{\beta}_p + (1 - w_p) \cdot \hat{\beta}_0 \tag{5}$$

where

$$w_p \equiv \frac{\sigma_p^2}{\sigma_p^2 + \text{Var}(\varepsilon_p^B)/n_p^B}$$

are the departmental empirical Bayes weights, ε_p^B is the residual for an observation with a Black subject in department p (residuals for other subjects are omitted), and n_p^B is the number of observations with Black subjects for department p.

2.1.3 Predicting Department-Specific Disparities

Finally, I treat the estimated department-specific disparities as the outcomes of interest in an effort to identify the factors that best predict racial inequities in police use of force. I use two machine learning techniques here, random forests and penalized logit, to handle large numbers of covariates relative to the number of observations. To make this problem more tractable and better reflect the decisions or classifications stakeholders make, such as whether a police department should be investigated for problematic practices,[7] instead of attempting to predict the exact value of a department's disparity, I consider a binary classification problem. The positive outcome of interest is a department having a high disparity against Black subjects, which I define as being in the top decile for a given force level. The null outcome I use is a department having a weakly negative disparity, i.e., not having a disparity against Blacks, and I omit from consideration departments in the middle, those with positive but less extreme disparities. When splitting the data into training and evaluation samples, I oversample observations with positive outcomes (using $2/3$ of them in the training sample, compared to $1/2$ of the null observations) to reduce class imbalance and the problems that may come with it, such as lazy algorithms that effectively never predict positive cases.

As predictors, I include a bevy of department and municipality-level characteristics: log median household income, Gini coefficient, log population, log population density, log number of police, the percentage of officers who are Black, the percentage of the population ages 18-65 that are Black, the violent crime rate, Mitt Romney's vote share in the 2012 presidential election, and county-level D-scores for the Black-White IAT from Project Implicit. I then repeat this exercise on the non-random subset of departments in the 2016 LEMAS survey using additional measures of department policies and culture plausibly linked to racial disparities or policing quality: whether the department engages in asset forfeiture, the use of a written

[7]For example, Newark's police department entered a consent decree in 2016 following an investigation by the Department of Justice that found unconstitutional policing.

aptitude test in selecting recruits, having a written community policing plan, formally surveying local residents to improve policing, having written policies on stop and frisk, foot pursuits, and unbiased policing, and whether the department's chief executive is female or non-Hispanic White.[8] These models collectively provide suggestive results for possible contributors to racial inequities in police violence and serve as starting points for future projects from across the social sciences as to why different police departments treat race differently.

2.2 Identification and Limitations

I estimate disparities in the intensity of police use of force experienced by subjects of different races conditional on force being used and after adjusting for incident characteristics and subject behaviors. Here I discuss this parameter and its interpretation in detail and outline the efforts I take to prevent the effects of confounding factors in my analyses.

Every incident undergoes two "treatments" prior to the officer's decision of what level of force to use. First is the decision to engage with a subject, such as whether to stop someone on the street or pull over a vehicle. Second is the extensive margin of whether to use force at all. Each of these treatments is plausibly affected by race, and some analyses in other settings have found evidence that they are (Gelman, Fagan, and Kiss, 2007; Fryer, 2019). Because my data are observed after these decisions, selection bias in who is in the sample, that is, in who has force used against them at all, could make interpretation of these disparities difficult.

If officers engage with civilians of different races at different rates, this could affect force rates, and depending on the nature of the interactions, also the severity of force used.[9] For example, if officers are overly suspicious of Black civilians, there could be many unwarranted stops that end without force or with only compliance holds, of which only the latter would appear in my data. If Whites are then only stopped for committing violent offenses that require high levels of force from officers, my empirical strategy could be biased towards a disparity that disfavors *Whites*. Or it could be that officers patrol in a manner that makes them more likely

[8]Some variables we may be interested in are inappropriate for this analysis due to insufficient variation. For example, all departments present in both my sample and the 2016 LEMAS allow at least some of their sworn officers with general arrest powers collective bargaining abilities.

[9]This concept is related to the more general idea of differences in the marginal subject across races discussed, for example, by Becker (1957).

to engage with Blacks committing violent crimes and Whites committing nonviolent offenses, which could threaten my results in the opposite direction. These problems are complicated by the presence of the second pre-observation treatment in my setting: the extensive margin of whether to use force at all.

I take two approaches to addressing possible bias from selection into the sample. First, although it cannot be determined ex post whether a particular stop or decision to use force was motivated by a subject's race, under milder assumptions, I can limit the role of racial disparities prior to the intensive margin of force severity. Using information on incident characteristics and subject actions, I repeat my analysis of overall racial disparities on a subset of observations where the subject's race was less likely to factor into the decisions whether to investigate and whether to use force. Specifically, I use the intersection of two subsets of the data, building on the ideas of Grogger and Ridgeway (2006) and Horrace and Rohlin (2016). First, I define a set of "always stop" incidents where the subject's race is less likely to factor into the officer's decision to engage with the subject. These are incidents where race is difficult to observe or the incident's severity is sufficiently high that race should not affect the engagement decision: crimes in progress, disputes, and traffic stops at night (8:00 PM through 5:59 AM), discarding suspicious person incidents, daytime traffic stops, and "other" incidents. Second, I restrict the sample to "always force" incidents where the outside option of not using force is less viable: those where the subject at least physically threatened or attacked an officer or another individual, dropping incidents in which subjects only "resist." The former set should contain fewer racial stops and the latter should be missing fewer incidents where no force was used, limiting problematic correlations between race and the error term and improving my estimates.

My second approach is adapted from the Appendix of Goncalves and Mello (2017) and uses the logic of the Heckman correction. I formally derive the model in Appendix B. In short, if police departments differ in their propensities to use force at all, subsequent estimates of racial disparities would be inconsistent, with the intuition being that the subjects being compared are now dissimilar across races. This is a classic estimation problem analogous to identifying the effects of a labor training program when the people who opt into such a treatment are non-random. Using data for each police department on the number of arrests by subject race as a benchmark for the number of incidents where force might be used, I compute each department's race-specific probability of using force against a subject. This allows me to identify the expectation of any selection bias by including the Mills ratio of these propensities as a regressor. An insignificant coefficient on the Mills ratio would then be evidence supporting a lack of selection bias in intensive margin analyses.

A separate concern is that the data I use come from the police themselves, and officers could withhold or misreport information. NJAM has attempted to solve any such problems through a crowdsourcing effort (McCarthy, 2019). This has uncovered minor discrepancies in some reports, often around officer names, and the only missing reports uncovered are 70 from Jersey City (prior to cleaning, there are more than 1,100 force reports from Jersey City in the data). Reporting by NJAM's NJ.com and a discussion with an active New Jersey police officer also indicate that there is some disagreement about when officers are required to fill out force reports. Because my coefficient of interest is the racial disparity after adjusting for incident characteristics, a necessary condition for these to bias my estimates is that any errors in the data such as missing reports are correlated with the subject's race. If officers systematically misreport Black subjects as posing greater threats than they actually do, this would cause me to underestimate racial disparities. However, incentives to misreport are greatly diminished by the lack of central oversight during the study period. Force reports existed mostly as physical copies, many only at the departments themselves despite guidelines that all force should be reported to county prosecutors, making external monitoring difficult (Nelson, 2019; McCarthy and Stirling, 2019).

3 Results

3.1 Overall Racial Disparities

Before estimating the disparities adjusted for the factors in the full model, it is helpful to understand the raw racial disparities that exist prior to adding any covariates. Figure 3 plots the results from regressing the binary force severity outcomes on an indicator for the subject being Black on top of the outcome mean, with tabular results in Appendix Table A.3.[10] Black subjects are unconditionally much more likely to have higher levels of force used against them when force is used, and these gaps increase with force severity. In percentage terms relative to outcome means, the disparities for Blacks are about 16.2% for at least unenhanced force, 31.5% for enhanced force, and 50.0% for firing weapons. This positive disparity for lethal force contrasts with data from Fryer (2019), where even in the raw data there are not disparities against Blacks for shootings.

Figure 4 now plots the racial disparities from the full model in Equation 1 for Blacks relative to others on top of outcome means, with corresponding regression tables in Appendix Table A.4. The racial differences are diminished compared to the unadjusted ones, but they remain large and meaningful. For the models using department fixed effects and clustering, the disparities estimated for Black individuals in the at least unenhanced, enhanced, and fired weapons outcomes are 1.9, 3.3, and 0.1 percentage points, respectively, with all except the firearms level statistically significant at the 1% level. While the positive disparities for firing weapons have economically significant point estimates given the rarity of these events, that rarity also greatly limits statistical power. Relative to the frequencies with which each force type is used, racial disparities are increasing with the severity of the force: in percentage terms, the department fixed effects-based disparities for Black subjects are 3.8%, 22.6%, and 25.0%. Note that the results are similar for models using department and officer fixed effects and clustering. Point estimates from the two are close, with the officer models having larger standard errors. Being able to adjust for individual officers' propensities to use force and/or the incidents they encounter on their beats does not change estimates of racial disparities, providing evidence about the relative importance of the department and the individual officer, an idea I will explore further later in this section. Estimates with a full set of racial indicators in Appendix Figure A.2 and Table A.5 show that relative to Whites, Hispanics and Asians/Pacific Islanders do not have significantly different outcomes, except that Asians/Pacific Islanders have a (relatively) large *negative* disparity for the fired weapons outcome that is highly significant with the more precise standard errors in the department-based model.

[10]In this and other analyses, I use outcome means as the reference as they are more interpretable and natural quantities once I add covariates to the models than alternatives such as the intercept.

As a robustness check to functional form, Figure 5 and Appendix Table A.6 present the odds ratios from the conditional logit models in Equation 2. The results are very similar to the OLS ones. Point estimates are positive and small at the lower levels of force, and they increase with the severity of the force, with the odds ratios for the at least unenhanced, enhanced, and fired weapons outcomes being 1.05, 1.29, and 1.06, respectively. All results except for the fired weapons disparity are statistically significant.

Next, to address concerns about possible racial differences in the officers' unobserved decisions whether to stop a subject or to use force against them at all biasing results, I take two distinct approaches as described in Section 2.2. First, I estimate Equation 1 on the intersection of two subsets of the data designed to limit pre-observation racial differences in the extensive margin of force. This new sample consists of incidents where a subject's race was less likely to affect the officer's decision whether to engage with the subject and incidents where the outside option of not using force was less viable and contains about 7,800 incidents. Figure 6 and Appendix Table A.7 contain the results from regressions with this new sample. Note that although the point estimates are larger in magnitude than their full sample analogs, outcome means are naturally higher here than in the full sample, as these incidents tend to be more severe than the omitted ones, prompting higher levels of force. Relative to the new outcome means, the estimated racial disparities using department fixed effects represent increases of about 6.6% and 26.1% for at least unenhanced and enhanced force, respectively. These figures are similar to those using the full sample and are not statistically significantly different despite being slightly larger. The disparity for firing weapons is a statistically insignificant 11.1%, with the smaller sample size affecting power and precision for this outcome the most. Note that in all models, standard errors increase severalfold due to the decreased sample size. Again, the estimates using officer fixed effects are similar to the ones based on department fixed effects, though their standard errors are even larger due to estimating officer fixed effects on such a reduced sample. Like in the full sample results, the enhanced force disparity is larger than the at least unenhanced one when benchmarked to outcome means. The fired weapons outcome is then no longer the largest relatively, though the outcome mean is so small and the standard errors sufficiently large that a wide range of disparities are plausible in the reduced sample. Still, the similarity of these estimates using the subsetted data to their full data counterparts support the use of the entire dataset for subsequent exercises and is consistent with any racial differences in events leading up to the officer's decision about the intensive margin of force not having large effects in this setting.

My second approach to possible selection bias uses the Heckman correction-style estimator based on Goncalves and Mello (2017) described earlier and detailed in Appendix B. Combining force and arrests data to compute the department-race probabilities of an incident resulting in force and adding the expectation of any subsequent bias allows for a tractable solution to weakening the assumptions needed to identify the intensive margin racial disparities. Table 3 contains the results of these regressions. In all three models, the coefficient on the Mills ratio is statistically insignificant at the 5% level, with two of the three also insignificant at the 10% level (the p-value for the fired weapons model is 0.06), suggesting minimal impact of any extensive margin racial differences. Taken with the prior limited sample estimates, these two approaches support the validity of intensive margin racial comparisons and the idea that these analyses compare like subjects across races. They emphatically do not, however, suggest that there are not extensive margin racial differences in police use of force, only that any such differences do not correlate with race and force intensity in such a manner or magnitude as to undermine intensive margin comparisons.

Having established the existence and validity of these racial disparities, to better to understand them and make them more transparent, Figure 7 plots smoothed actual outcomes against predicted outcomes by race. The predicted probabilities on the x-axis are for force outcomes based on Equation 1 with the indicator for the subject's race removed to allow for race-neutral context, and on the y-axis are the actual race-specific proportions of incidents smoothed by cubic splines. The top row of figures shows overall disparities for the at least unenhanced and enhanced force levels. I omit the fired weapons outcome, as the relative rarity of police shootings results in the predicted probabilities effectively having a mass point at 0. Axis endpoints are 0 and 1 to prevent overfitting with small numbers of observations with predicted probabilities outside of that window (1.7% for at least unenhanced and 8.3% for enhanced, the latter naturally concentrated below 0).

For at least unenhanced force, the Black and non-Black curves track very closely. The Black curve is slightly higher almost everywhere, and the largest disparities are at the bottom and middle of the predicted probability distribution. Because these predicted force probabilities are centered near 0.5, the overall disparity I find against Black subjects at this level is then coming from the incidents in the middle where officers may have had the most discretion in their level of force to use. Although it contributes less to the overall gap, the disparity at the lowest predicted levels of force shows that Black subjects have higher levels of force used against them in situations when non-Black subjects rarely do. For enhanced force, race gaps again grow towards the middle of the predicted probability support but are largest in the right tail, where the probability that this level of force is used is the greatest. The overall estimated disparity stems from the first divergence of the curves around probability 0.25, as there is much more mass there than in the right tail, but given that the slope of the Black curve is always near 1 while the non-Black curve flattens at the high end, Black subjects may not be getting "breaks" or benefit of the doubt that other subjects do in those situations. I return to the idea of police discretion momentarily when considering mechanisms that may cause these disparities.

The two graphs on the bottom of Figure 7 show two individual municipalities' enhanced outcomes as examples of how overall results do not reflect the heterogeneity present across departments. Municipality 1 on the left has one of the largest estimated racial disparities at this force level. Both Black and non-Black subjects have similar outcomes there until the predicted outcome is about 0.1, at which point Black subjects rapidly approach probability one of having enhanced force used against them, while non-Black actual outcomes track closely with their race-neutral predicted outcomes. In contrast, Municipality 2 on the right has an estimated disparity near 0. Although there are some minor racial differences in the middle of the predicted probability distribution that even out over the full support, there are not consistent disparities the way there are in the overall graphs.

While it is difficult to identify causal mechanisms for racial differences with observational police data, we may ask whether these disparities are the result of taste-based or statistical discrimination, as the policy implications for treating each may be different. Note that racially discriminatory policing for any reason is illegal in New Jersey and should not be considered justified, regardless of motivation. Statistical discrimination might manifest due to officers' belief or stereotype that, all else equal, Black subjects pose greater threats than subjects of other races do, prompting higher levels of force to address that threat (these beliefs need not be accurate; see Bohren et al., 2020). Under the assumption that male officers should not believe that they need this extra force to respond to female Black subjects, interacting race with the female indicator for incidents with male officers can differentiate between these channels of discrimination. A positive coefficient for the subject being Black is consistent with both statistical and taste-based discrimination. A negative coefficient on the interaction between being Black and female would then suggest statistical discrimination, while a null coefficient would be more consistent with taste-based discrimination. Table 4 shows that coefficients from this augmented model on the subject being Black, the subject being female, and their interaction are positive, negative, and approximately zero, respectively, for both the at least unenhanced and enhanced models, supporting the taste-based discrimination hypothesis (I do not include the fired weapons outcome, as only five women are shot in my sample).

Continuning to consider mechanisms, there may be reasons to expect disparities to shrink over time, such as increased scrutiny towards police (so-called "Ferguson effects," or chilling responses on policing after the shooting of Michael Brown in Ferguson, Missouri in 2014) incentivizing misreporting, or perhaps from departments improving through learning and reducing excessive force against Black subjects. Figure 8 presents point estimates from the full model for the subject being Black interacted with the year using 2012 as a reference to test whether disparities change over time. However, there is no compelling evidence of any such changes, casting doubt on these hypotheses: racial disparities appear fairly static over the sample period despite the possible external pressures.

Given the similarity of results from models based on department and officer fixed effects, I now explore the relative importance of the department/municipality and officer more closely. On a set of 1,041 force incidents by 174 manually identified officers who switch police departments during my sample period, I fit Equation 1 with officer fixed effects and an additional indicator for whether the department in which an officer is working has an above-median racial disparity for the relevant level of force as estimated in Section 3.2.[11] Table 5 presents the results. For at least unenhanced force, the interaction between a subject being Black and the department having an above-median disparity is positive, large, and statistically significant, supporting the idea that the department or municipality in which an officer works is a larger determinant of racial disparities than the identity of the officer. For enhanced force, the interaction is 0. One explanation would be that more severe force is less discretionary, leaving less room for any marginal effects of the new department's norms or differences in the municipality's setting on force intensity. In any case, these and other results suggest that departments or municipalities play an important role in racial disparities, which leads to the next set of results exploring to what extent these departmental disparities may vary.

3.2 Variation in Racial Disparities

In this section, I focus on how racial disparities vary across and within departments. Overall racial disparities treat these differences as monolithic, while they may vary both across and within departments. Furthermore, overall estimates put the most weight on the places with the most incidents, typically large, urban areas. As explained in Section 2, I estimate department-level racial disparities β_p using the empirical Bayes estimator defined in Equation 5. Due to small sample issues for certain departments, I winsorize the estimated disparities at the 1st and 99th percentiles.[12]

Figure 9 presents kernel density estimates of the winsorized empirical Bayes departmental disparities, with corresponding summary statistics in Table 6. These distributions have long tails that often dwarf the overall racial disparities estimated before. For the at least unenhanced force, enhanced force, and fired weapons outcomes, the standard deviations of the winsorized disparities compared to the (mean disparity) are very large: 0.07 (0.02), 0.05 (0.01), and 0.01 (0.00), respectively. Moreover, the tails are extreme. The 99th percentile department is esti-

[11] I do not analyze firearms here, as no switching officers use that level of force in my data. Additionally, I include year fixed effects instead of month-year due to the greatly reduced sample size.

[12] Although empirical Bayes estimates are designed to shrink less reliable point estimates to the center of the distribution, in extremely small samples such as with the smallest departments, this can break down. Consider a department that uses force against only two Black subjects in the data with similar demographics and actions taken as part of the same incident with the same level of force used against them. The model would fit these data almost perfectly, i.e., Var(ε_p^B) from Equation 5 is near zero, and so the Bayesian updating process is very confident that the disparity is accurately estimated and minimal shrinkage occurs.

mated to use at least unenhanced force 37 percentage points more often against a Black subject than it would against a comparable White subject, while unenhanced or greater force is used in just over half of all incidents. And the 99th percentile department for the enhanced force outcome is 9 percentage points more likely to use this level of force against a Black subject than a counterfactual White one, while enhanced force is used in only 14 percent of all force incidents. However, some departments do not seem to have racial disparities against Blacks, or at least not for certain force outcomes. 19% of departments have weakly negative racial disparities for the at least unenhanced force outcome, 24% for enhanced, and 36% for fired weapons. Although these numbers are small, they serve as an existence proof for departments without racial disparities against Blacks in their use of force, and may be a starting point for researchers and stakeholders seeking to improve other departments.

The process of estimating these empirical Bayes disparities also allows for determining how much variance in force outcomes these department-race interactions can explain. First, I look at the change in model fit from the pooled regression without the interactions in Equation 4 to the unpooled regression in Equation 3. I then compare this difference to the corresponding change between the unpooled regression and new ones with additional variables removed as a benchmark for explanatory power. Across the three force outcomes, adding department-specific racial differences increases the models' R^2 values by an average of 7%. In comparison, the unpooled regression is an improvement of an average of only 4% over models without the subject's sex and 2% over models missing incident type indicators. Hence departmental racial differences are a major contributor to overall variation in force intensity.

Having seen that racial disparities in police use of force vary across departments, how do they vary within departments? Figure 10 presents a series of probabilities that a department does not have a disparity disfavoring Blacks (i.e., that its estimated disparity is weakly negative). For example, as seen earlier, the unconditional probability that a department does not have such a disparity in the at least unenhanced force outcome is 19%. This rises to 31% if we condition on the department not having a disparity for enhanced force (a 72% increase) and rises further still if we additionally condition on the department not having a disparity for firearms. Similar results at other levels of force corroborate this within-department disparity linkage: at every level of force, the probability that a department does not have a disparity against Blacks increases greatly when conditioning on not having a disparity at another force level and increases even more when conditioning on not having a disparity at both of the other outcomes. Although departments with racially equitable use of force practices at one level are much more likely to be similarly fair at another level, many departments do have disparities at force level despite not having them at other levels.[13] This makes identification, understanding, and treatment of racial disparities more complicated. Certain departments do not have racial disparities in their use of force, but others might have large disparities only at the highest levels of force, while still more might have smaller disparities at every level.

Both across and within departments, we see variation in racial disparities. This variation highlights the difficulties in identifying problematic departments and designing interventions targeting racialized policing. It may be difficult to treat departments with such heterogeneous disparities, and the variation in disparities within departments further confounds efforts. At the same time, the existence of departments without disparities, and the knowledge that they are also less likely to have racial disparities at other levels, is a promising sign that better policing is feasible.

3.3 Predicting Departmental Racial Disparities

Having established the heterogeneity present in department-specific racial disparities, I turn to the question of what departmental or municipal factors can predict these disparities. For example, plotting a map of the disparities in Appendix Figure A.3 rules out geographic factors such as a municipality's proximity to New York City or Philadelphia (the nearest major metropolitan areas) or the presence of clusters of departments with larger racial differences. I use machine learning techniques to identify the best predictors for these departmental disparities as described in Section 2.1.3. These models answer the question "what factors differentiate the departments with the highest disparities from those that do not have disparities

[13] These proportions of departments without disparities would all be greater if using an appropriate cutoff $0+\varepsilon$ instead of simply 0 such that departments below the cutoff, even if the estimated disparities are positive, are unlikely to be engaging in racially inequitable use of force, essentially accounting for a margin of error. Identifying such a value is beyond the scope of this paper, and I take the stricter approach of using only weakly negative disparities.

against Black subjects?" Because the mapping from police departments to municipalities is almost one-to-one, I include predictors based on both the police departments themselves and the municipalities they cover. Doing so requires dropping the New Jersey State Police, as that department covers state highways and some municipalities that lack their own department. To ease reading coefficients in the penalized logit models, I follow Gelman (2008) and divide non-binary variables by two times their standard deviations, so that all coefficients, including binary ones, are on approximately the same scale.

Figure 11 presents the results of the three random forest models, one for each outcome. The models for at least unenhanced force, enhanced force, and fired weapons have error rates of 36%, 31%, and 25%, respectively.[14] In the graphs, the x-axes indicate the mean decrease in accuracy of the classification decision (high/no disparity) associated with removing that predictor from the model. For the at least unenhanced force model, median household income is the most valuable predictor by a large margin, followed by the percentage of the population ages 18-65 that is Black and Mitt Romney's share of the 2012 presidential election vote. The enhanced force model shows several of the same predictors near the top: the Romney vote share is now the most important predictor, with the violent crime rate a moderate gap behind, the Black population share a similar distance behind that, and household income just behind that. The model for firing weapons should be taken with caution given the rarity of these events, but the percentage of officers that are Black and the Romney share are the two top predictors, followed by the municipality's Gini coefficient and the violent crime rate.

Penalized logit summary results are displayed in Figure 12. These plots trace the coefficient paths as the penalty parameter changes, from larger penalties (resulting in fewer, smaller nonzero coefficients) on the left to smaller penalties (more, greater nonzero coefficients) on the right. At the cross-validation-selected optimal values of this parameter, the weighted out-of-sample error rates in classification are 37%, 31%, and 22% for the at least unenhanced, enhanced, and fired weapons outcomes, respectively. In the at least unenhanced model, the first predictors to turn on and the (sign) of their coefficients are log median household income (+)

[14] Because the training and evaluation data are not split identically between departments in the top decile of disparities and those without disparities, I compute the overall error rates as a weighted average of the type I and type II error rates to match the population proportions.

and the percentage of officers that are Black (+), followed by the size of the police force (-) and the Black population share (-). For the enhanced model, the first predictors are the violent crime rate (+), Romney vote share (-), and total population (+). Finally, the fired weapons model has the size of the police force (- then +), Gini coefficient (-), population (-), and household income (+).

The results from these two sets of predictions are suggestive of the characteristics of police departments and their communities that are associated with racial inequities in police use of force. For at least unenhanced force, household income and Black population share appear as top predictors in both the random forest and penalized logit models. For enhanced force, the Romney vote share and violent crime rates are both present. While the fired weapons models are likely the least informative due to the smaller number of events underlying them, the violent crime rate is a key predictor in both. The similarity of variables across the random forest and penalized logit methods is encouraging for their validity. While the different force levels have several different predictors emerge, this is not inherently problematic, and it is plausible that different force levels have different factors associated with their disparities, or even that the same factors have different relationships at different levels of force. For example, at lower levels of force, being in richer areas with fewer Black civilians could result in officers tending to use excessive force against the Black subjects that they do encounter (consistent with the signs of the coefficients from the penalized logistic model), while at higher levels of force with and the more severe threats that tend to prompt them, the effects of the setting might be diminished compared to those from factors like the officer's training.

It is notable that the top predictors that emerge from both prediction algorithms are characteristics of the municipalities rather than the departments, but more can be done to explore the extent to which departmental characteristics might impact these disparities by attempting to capture more intangible characteristics such as policies or culture. While these aspects of the departments are inherently difficult to measure, I can supplement the prior models with additional variables for the subset of departments present in the 2016 LEMAS. These new variables, such as the race of the police chief and whether the department has a written policy on racial profiling, may help illustrate the predictive value of less surface-level departmental characteristics. Being limited to the departments that responded to that vintage of the LEMAS does greatly diminish the number of available data points, to below 100 prior to filtering out those with positive disparities outside of the top decile, meaning these models are based on far fewer observations and are likely to be less robust. Appendix Figures A.4 and A.5 plot the random forest and penalized logit results incorporating both the original variables and the new LEMAS variables. Indeed, the error rates for the three models, in the typical order, are 51%, 29%, and 17% for the random forests and 58%, 46%, and 19% for the penalized logits. None of the three random forest select any of the LEMAS variables as one of their top predictors, with the highest-ranked predictor being sixth in the at least unenhanced model. The penalized logit models also bear this out. For the at least unenhanced model, having a White, non-Hispanic chief is the third variable to turn on, and having a written policy on cultural awareness training is third for enhanced force. Having a female chief is actually the second variable to turn on in the fired weapons model, but it actually turns back off as the penalty parameter decreases. Although these results certainly do not suggest that these policy and cultural indicators are especially valuable predictors of racial disparities in police use of force, the limited sample size and possible importance of other variables not included mean that I cannot rule it out.

In summary, this prediction exercise yields several important takeaways about the factors that differentiate high- and no-disparity departments. First, predicting departmental disparities purely from characteristics of the departments and the municipalities they cover is difficult, and this is an area where social scientists have much more to contribute; in addition to quantitative analyses such as this, more qualitative work may help us understand what separates departments. This is also an obvious area for policy improvements in the form of centralized data collection like New Jersey has implemented, and better data availability across states and time can only improve our understanding of the most important departmental (or even officer) characteristics. Second, certain factors are useful predictors of a department having a large disparity as opposed to none, in particular median household income and the Black share of the population ages 18-65 when considering disparities at the lower levels of force (at least unenhanced), and political preferences and the violent crime rate for the disparities at the higher levels (enhanced). Because these predictions are not causal, causality may run in many different directions, including between predictors, from outcome to predictor, or from unseen confounders. Third, those top predictors vary across force levels, which may suggest that not all racial disparities have the same underlying causes. And fourth and finally, the most informative predictors tend to be characteristics of the municipality rather than the department, with the caveat that it is conceivable that there are some difficult-to-measure departmental characteristics that I am unable to account for. Overall, these findings underscore the variation and heterogeneity in disparities seen throughout this project and will hopefully serve as a starting point for future research or policy to identify and ameliorate racial inequities in police use of force.

4 Conclusion

In this paper, I combine new analytical strategies with incident-level data on all recorded uses of force by municipal and state police in New Jersey between 2012 and 2016 to estimate racial disparities in the severity of force used on a subject, conditional on force. This rich dataset allows me to adjust for factors such as the type of incident or the subject's actions and address selection into the dataset to better examine the role of race in police violence. Overall, I find large disparities against Black subjects in police use of force that increase along the spectrum of force severity. Based on an analysis of officers who switch departments and comparisons of models using departmental or officer fixed effects, the department (or municipality) seems to play a larger role in determining these disparities rather than the individual officer. Further, the disparities are more consistent with taste-based, as opposed to statistical, discrimination. I then extend empirical Bayes methods and document substantial heterogeneity in these racial differences across and within departments, finding some departments without disparities against Black subjects, but also a long tail with especially large ones. A department with a disparity at one level of force is more likely to have a disparity at the others, but many do not, highlighting the complexities of investigating racial inequities in police use of force. Finally, I explore the factors that differentiate departments without disparities against Black subjects from those with the largest ones. The strongest evidence is in favor of municipal characteristics, as opposed to departmental observables or a select set of policy or cultural indicators, but additional research is required to differentiate correlation and causality and whether these factors are root causes themselves or are reflective of others.

Much work remains to be done on race and police use of force, with policy and research closely intertwined. The presence of departments without estimated racial disparities is a sign that improvements are possible. But the variation in disparities both across and within departments may make identifying departments for treatment and implementing effective reform more difficult and less efficient. Police departments and public officials must ensure that data availability and transparency continue to improve to aid progress. Initiatives such as New Jersey's new statewide data portal combined with its mandatory force reporting and the FBI's National Use-of-Force Data Collection will facilitate future studies of how race and policing interact that can better consider how these issues are not uniform, varying across departments, areas, and force levels. Better understanding of variation in racial disparities in police use of force is a vital step towards greater racial justice and equitable public services that can only help us move forward.

Acknowledgments

I am grateful to Leah Boustan, Jesse Bruhn, Janet Currie, Milena Djourelova, Felipe Goncalves, Thomas Fujiwara, Bo Honoré, Ilyana Kuziemko, David Lee, Alex Mas, Jonathan Mummolo, Amy Wickett, seminar participants at Princeton University, the U.S. Census Bureau, and the Young Economists Symposium, and members of the Princeton Police Department for helpful comments. This work is based on research started while the author was a graduate student at Princeton University and was supported by The Princeton University Industrial Relations Section. Any opinions and conclusions expressed herein are those of the author and do not reflect the views of the U.S. Census Bureau.

Table 1: Summary Statistics for Force Reports

Statistic	N	Mean
Max force: compliance hold	39,322	0.50
Max force: unenhanced	39,322	0.36
Max force: enhanced (non-firearms)	39,322	0.14
Max force: fired weapon	39,322	0.004
Max subject action: resisted	39,322	0.63
Max subject action: physical threat/attack	39,322	0.34
Max subject action: blunt weapon threat/attack	39,322	0.01
Max subject action: knife threat/attack	39,322	0.01
Max subject action: vehicular threat/attack	39,322	0.01
Max subject action: firearm threat	39,322	0.01
Max subject action: fired weapon	39,322	0.001
Officer injured	39,322	0.10
Incident: crime in progress	39,322	0.27
Incident: domestic dispute	39,322	0.13
Incident: other dispute	39,322	0.11
Incident: suspicious person	39,322	0.11
Incident: traffic stop	39,322	0.09
Incident: other	39,322	0.33
Subject: White	39,322	0.48
Subject: Black	39,322	0.41
Subject: Hispanic	39,322	0.10
Subject: Asian/Pacific Islander	39,322	0.01
Subject: female	39,322	0.20
Subject: age	39,322	31.08

Notes: Data cover all police departments in New Jersey from 2012 through 2016. Data have been restructured so that each observation is a subject who had force used against them by police and cleaned as described in Appendix A. “Max force” refers to the highest level of force used in an incident; lower levels of force are not reported. “Max subject action” indicates the most severe action a subject took that could justify an office using force. Force used and subject actions are ordered from least severe to most severe. Incidents may have multiple types.

Table 2: Departmental Regressions

	Prop. Subjects Black	Log(Incidents)	Force Incidents/Arrest
Prop. Pop. 18-65 Black	1.843^{***}	2.935^{***}	0.022
	(0.153)	(0.782)	(0.018)
Prop. Pop. 18-65 Black^2	-1.150^{***}	-4.731^{***}	-0.040^{*}
	(0.176)	(1.163)	(0.022)
Log(Pop.)	0.001	0.325^{***}	-0.000
	(0.017)	(0.099)	(0.002)
Log(Pop. Density)	0.002	0.067^{*}	0.003^{***}
	(0.007)	(0.034)	(0.001)
Violent Crime Rate	-0.182	152.551^{***}	1.227^{**}
	(3.152)	(28.490)	(0.475)
Log(Avg. Police)	0.016	0.663^{***}	0.002
	(0.021)	(0.119)	(0.003)
R^2	0.568	0.697	0.112
Num. obs.	454	454	454

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from regressions run at the departmental level using cleaned force data for New Jersey from 2012-2016.

Table 3: Racial Disparities by Police in Force of At Least Specified Severity, Conditional on Force, with Heckman Correction

	Unenhanced	Enhanced	Fired Weapons
Subject Black	0.021^{***}	0.034^{***}	0.001
	(0.006)	(0.005)	(0.001)
Heckman correction	0.068	0.022	0.008^{*}
	(0.045)	(0.030)	(0.004)
Fixed effects	Dept.	Dept.	Dept.
Clustering	Dept.	Dept.	Dept.
Outcome mean	0.505	0.146	0.004
R^2	0.181	0.118	0.181
Num. obs.	39182	39182	39182

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from Equation 1, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, and department or officer fixed effects. These regressions additionally incorporate a Heckman correction-style adjustment adapted from Goncalves and Mello (2017) to examine the possibility of extensive margin racial differences in force usage impacting results.

Table 4: Effect of Female Subject on Police Use of Force of At Least Specified Severity, Conditional on Force

	Unenhanced	Enhanced
Subject Black	0.017^{***}	0.033^{***}
	(0.006)	(0.005)
Subject female	-0.172^{***}	-0.063^{***}
	(0.008)	(0.005)
Subject Black $\times$ female	0.010	-0.000
	(0.010)	(0.008)
Fixed effects	Dept.	Dept.
Clustering	Dept.	Dept.
Outcome mean	0.505	0.146
R^2	0.181	0.118
Num. obs.	39182	39182

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from Equation 1, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race interacted with an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, and department fixed effects.

Table 5: Effect of Being in Department with Above-Median Empirical Bayes Racial Disparity, Officers Switching Departments Only

	Unenhanced	Enhanced
Subject Black	-0.063	-0.006
	(0.047)	(0.029)
Department above median	-0.117^{*}	-0.027
	(0.065)	(0.041)
Subject Black $\times$ department above median	0.218^{***}	0.049
	(0.063)	(0.049)
Fixed effects	Off.	Off.
Clustering	Off.	Off.
R^2	0.412	0.289
Num. obs.	1047	1047

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from Equation 1, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, department fixed effects, and an indicator for the department in which the officer served having above-median racial disparities from the empirical Bayes analysis. This regression only includes incidents from officers identified as switching departments.

Table 6: Summary Statistics for Winsorized Empirical Bayes Department x Black Subject Interactions

	Unenhanced	Enhanced	Fired Weapons
SD	0.070	0.051	0.007
Min	-0.260	-0.199	-0.037
P05	-0.047	-0.101	-0.009
P25	0.008	0.002	-0.001
Median	0.019	0.033	0.001
P75	0.032	0.037	0.001
P95	0.108	0.064	0.009
Max	0.366	0.094	0.026
Mean	0.023	0.014	-0.000
$\% \leq 0$	0.185	0.238	0.360

Notes: Table reports empirical Bayes estimates of departmental racial disparities from Equation 5, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race interacted with a department indicator, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, and department fixed effects. Winsorization is done at the 1st and 99th percentiles.

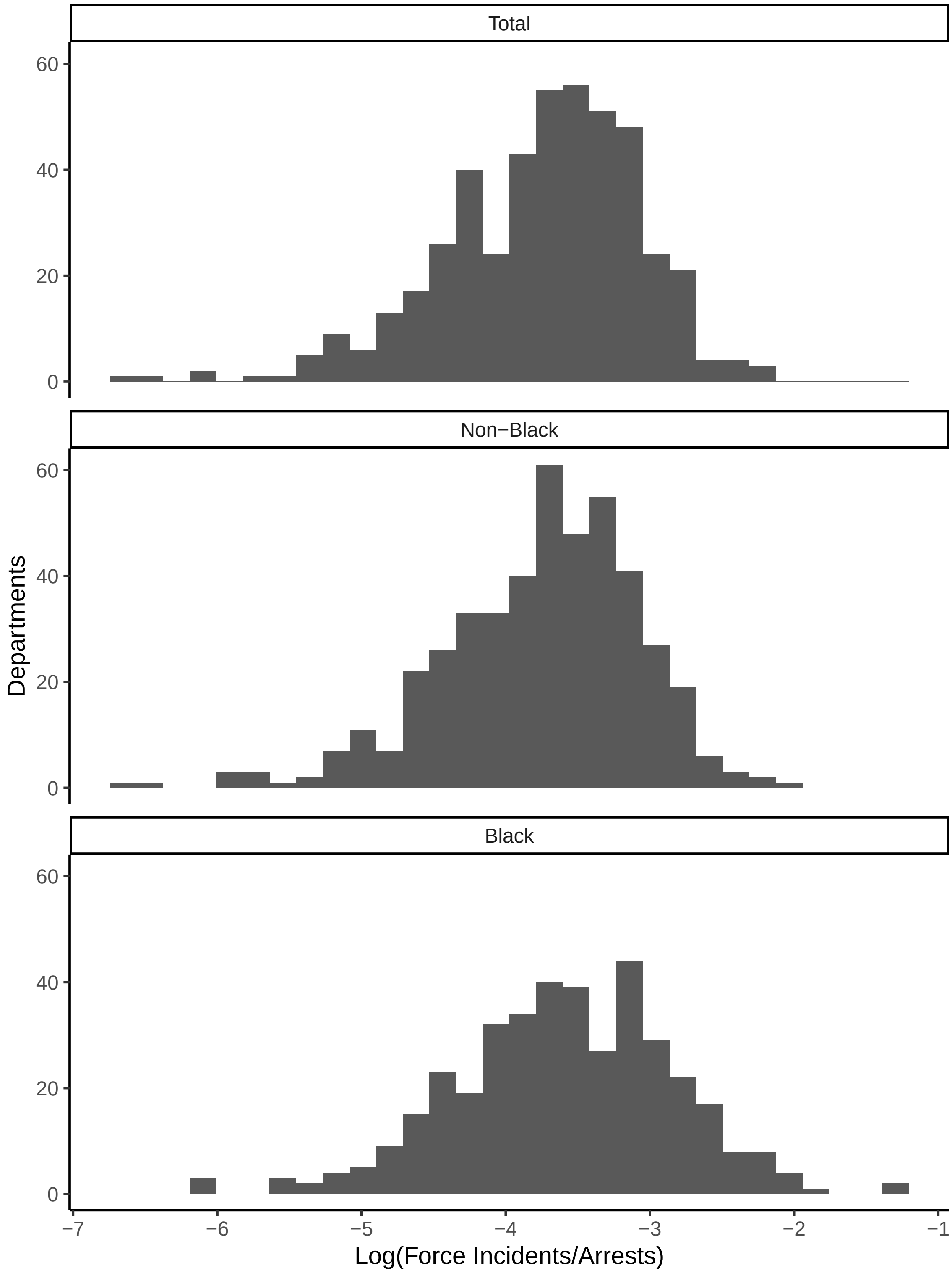

Figure 1: Departmental Force Incidents Normalized by Arrests

Notes: Figure displays summary statistics for the departments present in the cleaned use of force data from New Jersey for 2012-2016.

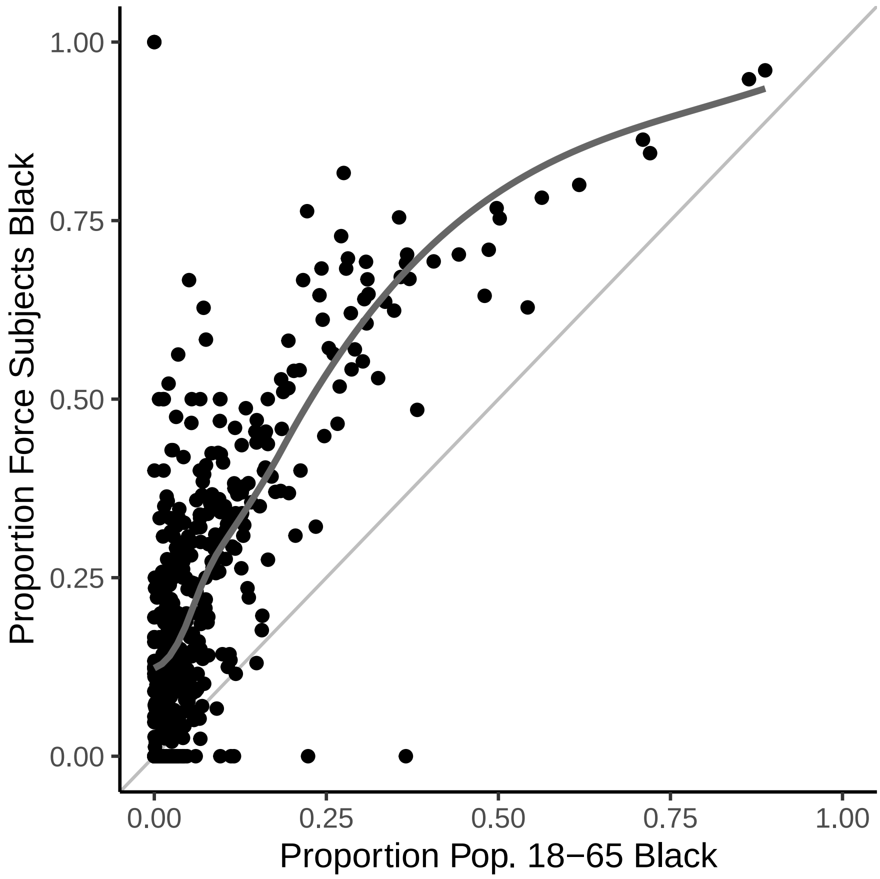

Figure 2: Racial Composition of Local Population and Force Suspects
Notes: Figure plots the Black proportion of the local population ages 18-65 against the Black proportion of firce subjects for the departments present in the cleaned use of force data from New Jersey for 2012-2016, except for the New Jersey State Police. The fit line is created via loess.

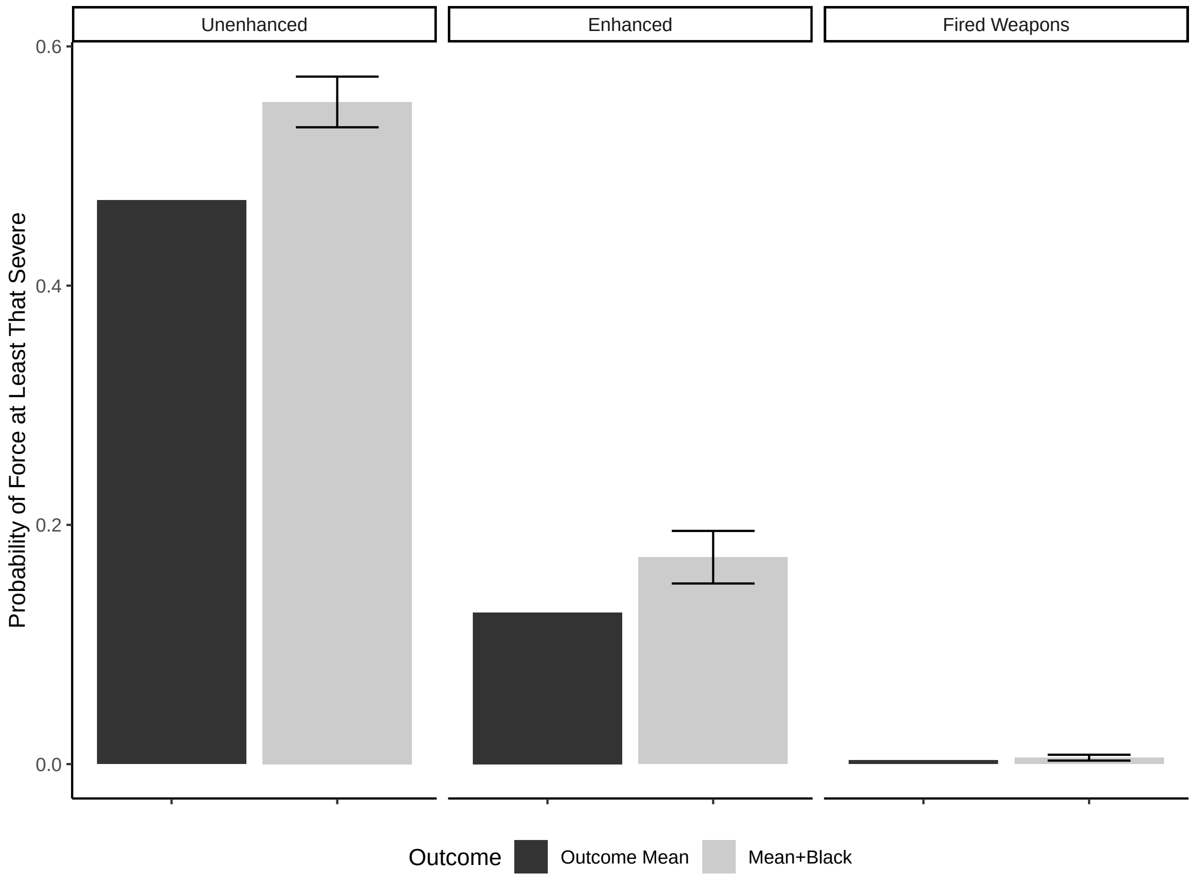

Figure 3: Overall Racial Disparities (No Controls)

Notes: Figure presents results from a series of OLS models regressing outcomes on an indicator for the subject being Black. Each heading represents a different outcome: whether, conditional on any force being used, force of at least the specified severity was used. Black values are obtained by taking the outcome mean and adding the coefficient on being Black, with confidence intervals based on the corresponding standard errors.

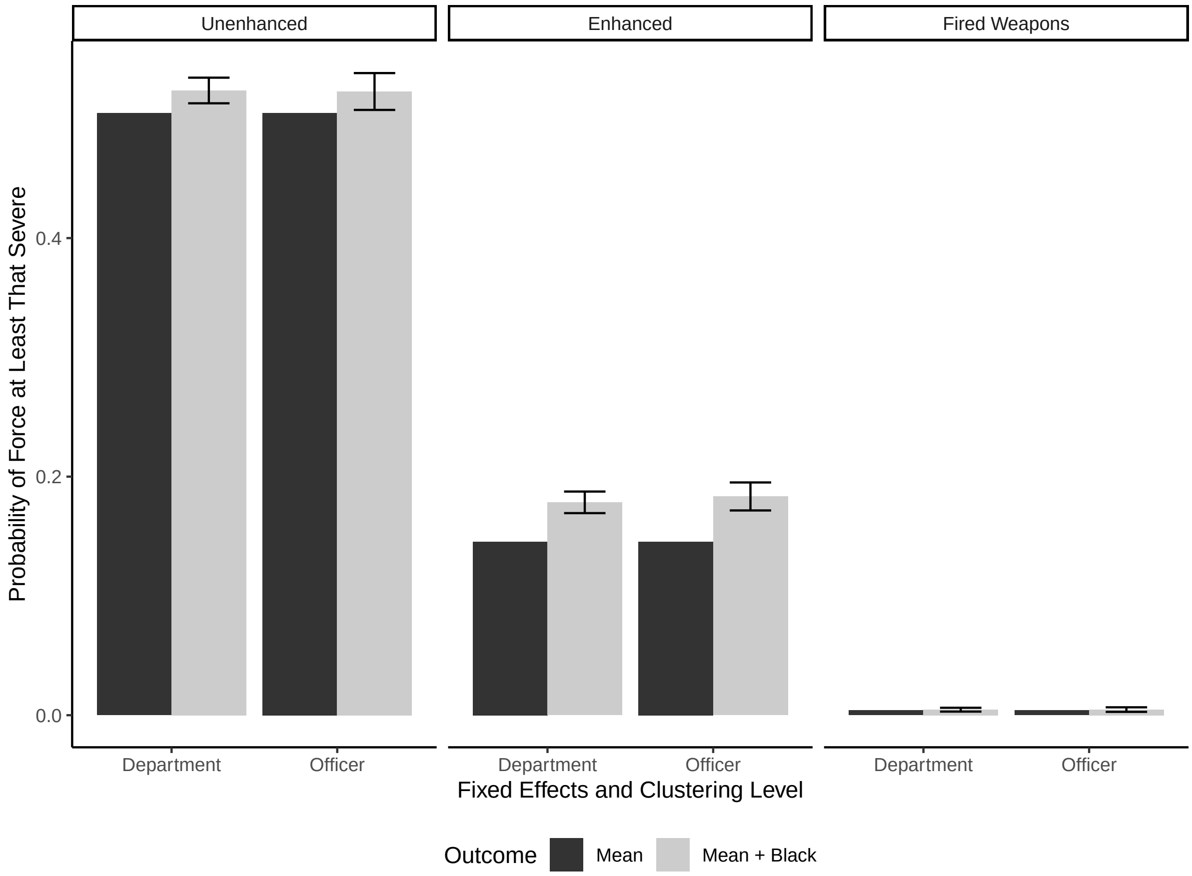

Figure 4: Overall Racial Disparities for Subject Being Black on Probability of Force of at Least Specified Severity

Notes: Figure presents results from a series of OLS models fit via Equation 1. Each heading represents a different outcome: whether, conditional on any force being used, force of at least the specified severity was used. Bars labeled "Department" include department fixed effects, and bars labeled "Officer" instead include officer fixed effects. Confidence intervals are based on the Black coefficient.

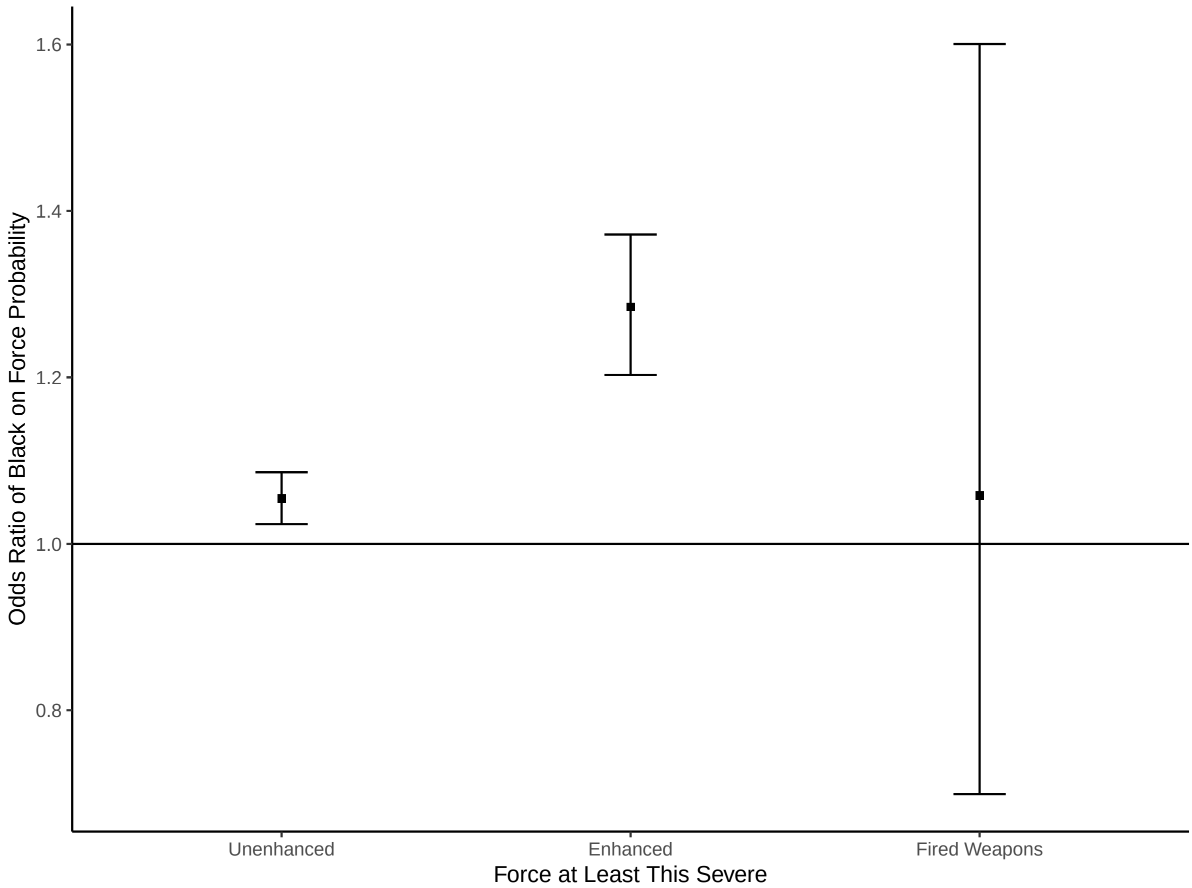

Figure 5: Conditional Logit Odds Ratios of Subject Being Black on Probability of Force of at Least Specified Severity

Notes: Figure presents results from a series of conditional logit models fit via Equation 2. Each heading represents a different outcome: whether, conditional on any force being used, force of at least the specified severity was used. Points indicate the odds ratio of a Black subject. Confidence intervals clustered at the department level are obtained by exponentiating the logit coefficients from the regression and are asymmetric.

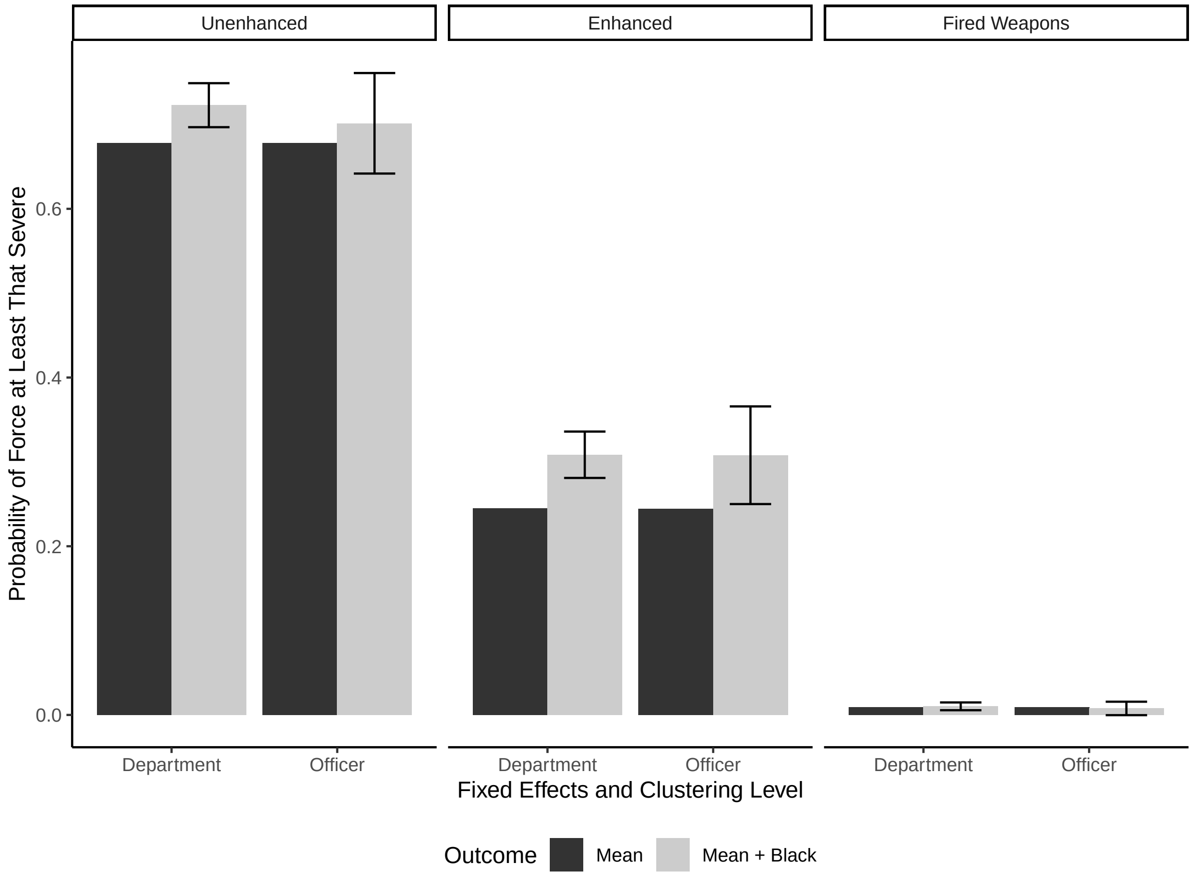

Figure 6: Overall Racial Disparities of Subject Being Black on Probability of Force of at Least Specified Severity (Subset Where Race is Unlikely to Affect Decision to Engage Subject and Subject at Least Physically Threatened/Attacked Officer or Another)

Notes: Figure presents results from a series of OLS models estimated on the subset of the data where a subject's race was less likely to have influenced the officer's decision to engage with the subject: crimes in progress, disputes, and traffic stops at night, and where the subject at least physically threatened or attacked an officer or another. The latter restriction is equivalent to dropping observations where the most severe actions by the subject was resisting. Regressions are fit via Equation 1. Each heading represents a different outcome: whether, conditional on any force being used, force of at least the specified severity was used. Bars labeled "Department" include department fixed effects, and bars labeled "Officer" instead include officer fixed effects. Confidence intervals are based on the Black coefficient.

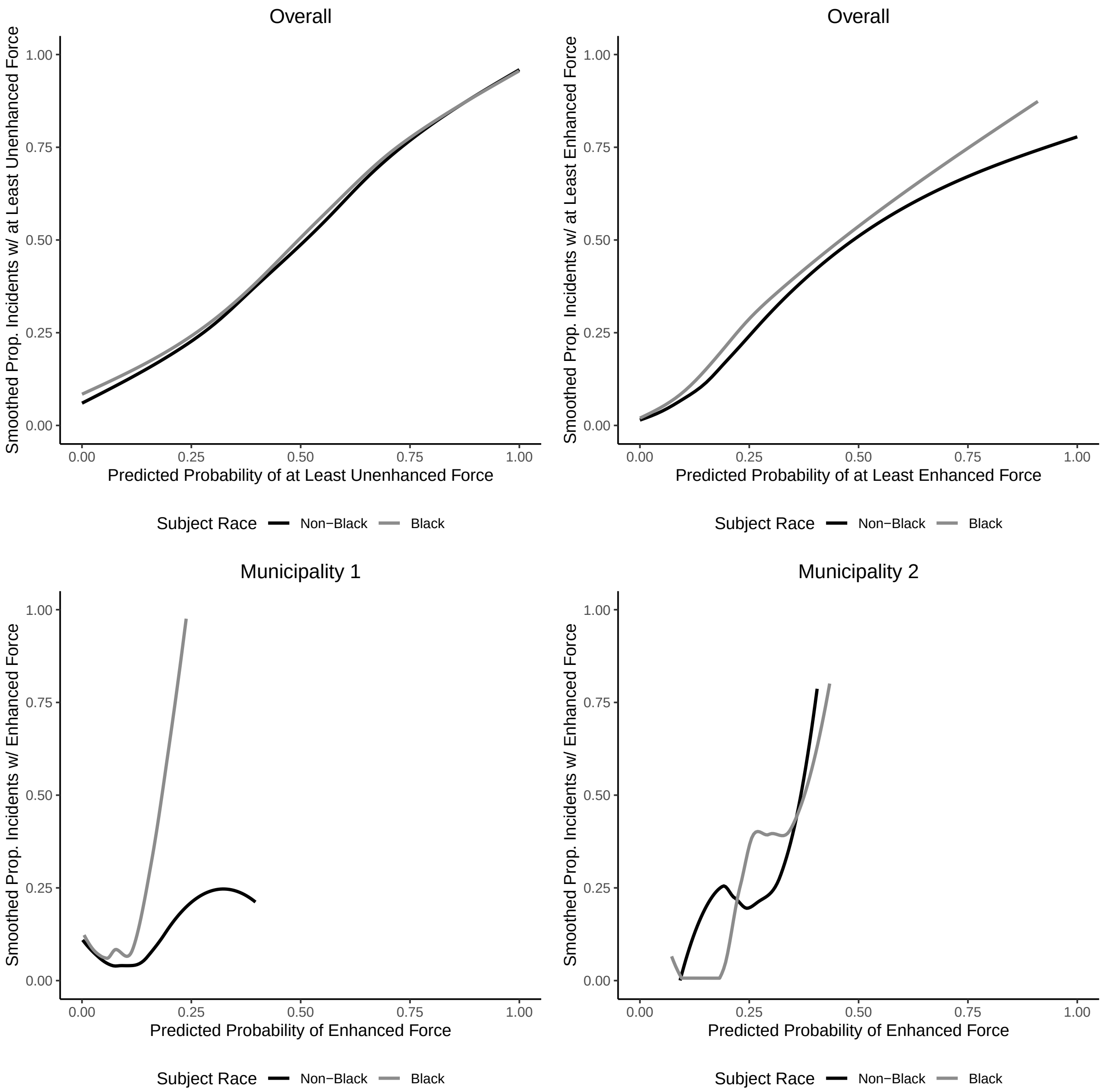

Figure 7: Predicted Probabilities of Force Outcomes vs Actual Outcomes, by Race

Notes: Figure presents the predicted probabilities of each outcome fit via Equation 1 excluding the racial indicator against the actual proportions of incidents with that outcome, by race. Outcomes are smoothed via cubic splines. The top row presents overall plots for at least unenhanced force, enhanced force, and firing weapons. The bottom row presents example plots for two individual municipalities for the enhanced level: Municipality 1 on the left has one of the largest disparities for this outcome, and Municipality 2 on the right has a disparity near zero.

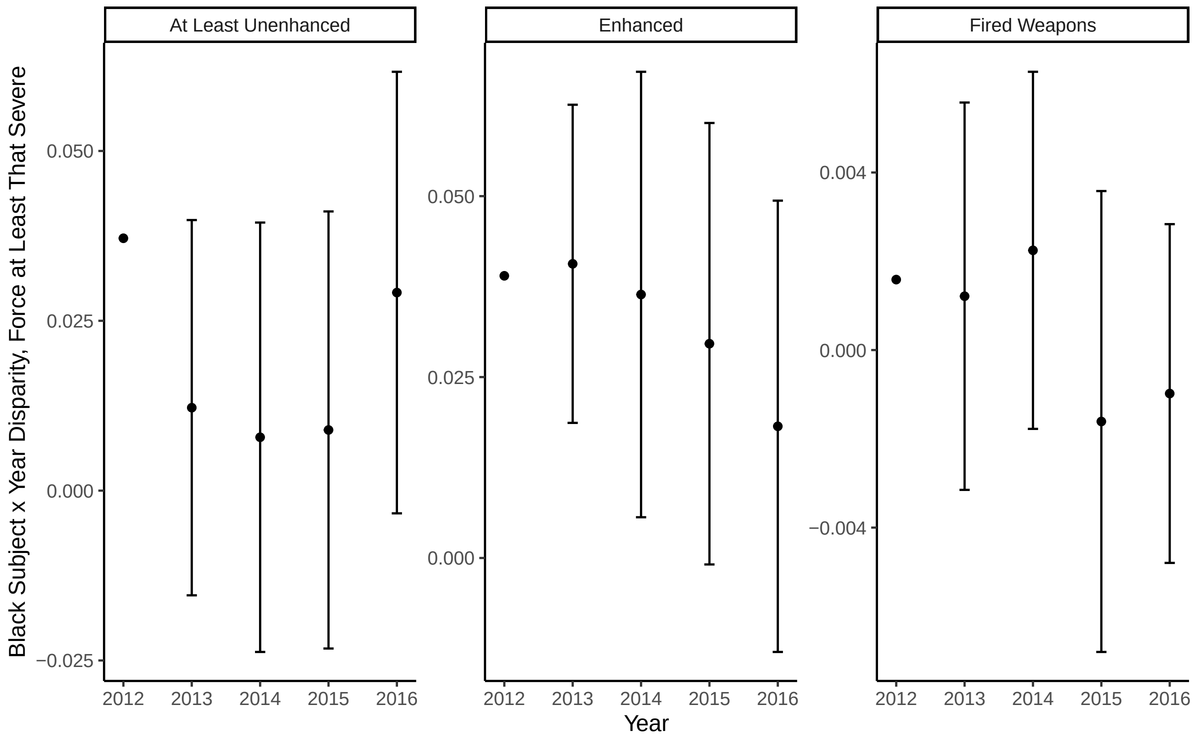

Figure 8: Overall Racial Disparities of Subject Being Black on Probability of Force of at Least Specified Severity by Year

Notes: Figure presents results from a series of OLS models fit via Equation 1 with the Black indicator interacted with year (2012 is omitted). Each heading represents a different outcome: whether, conditional on any force being used, force of at least the specified severity was used. 95% Confidence intervals are based on the corresponding year's coefficient. Standard errors are clustered at the department level.

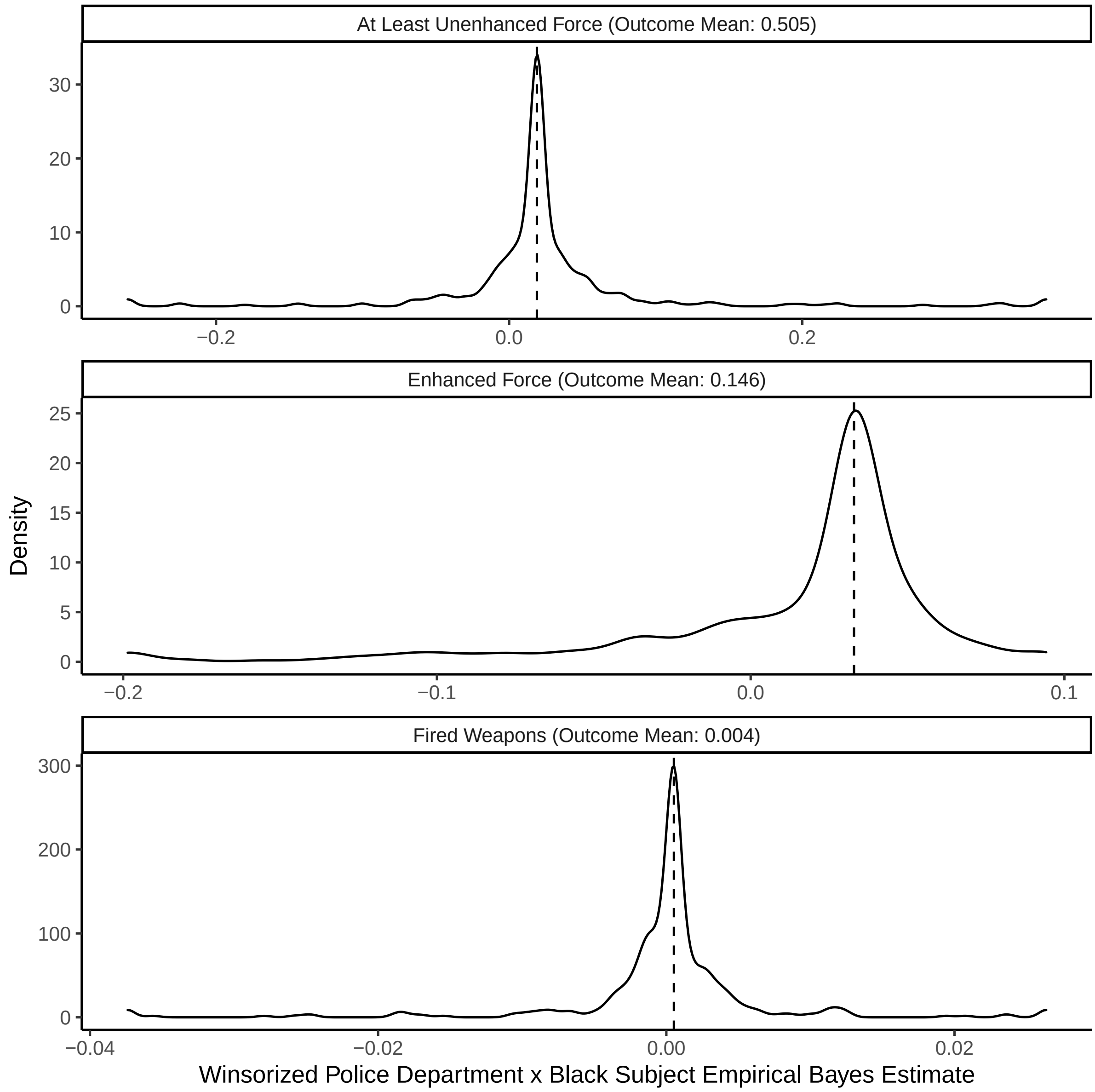

Figure 9: Distribution of Winsorized Empirical Bayes Estimates of Department-Black Interactions

Notes: Figure presents kernel density estimates of department-specific racial disparities β between White/Asian/Pacific Islander subjects and Black subjects as estimated from the empirical Bayes estimator in Equation 5 and winsorized at the 1% and 99% levels with Gaussian kernels and the Silverman (1986) rule-of-thumb bandwidth. Each subgraph shows results from regressions with the specified outcome outcome: whether, conditional on any force being used, force of at least the specified severity was used. The dashed line indicates the mean of the prior.

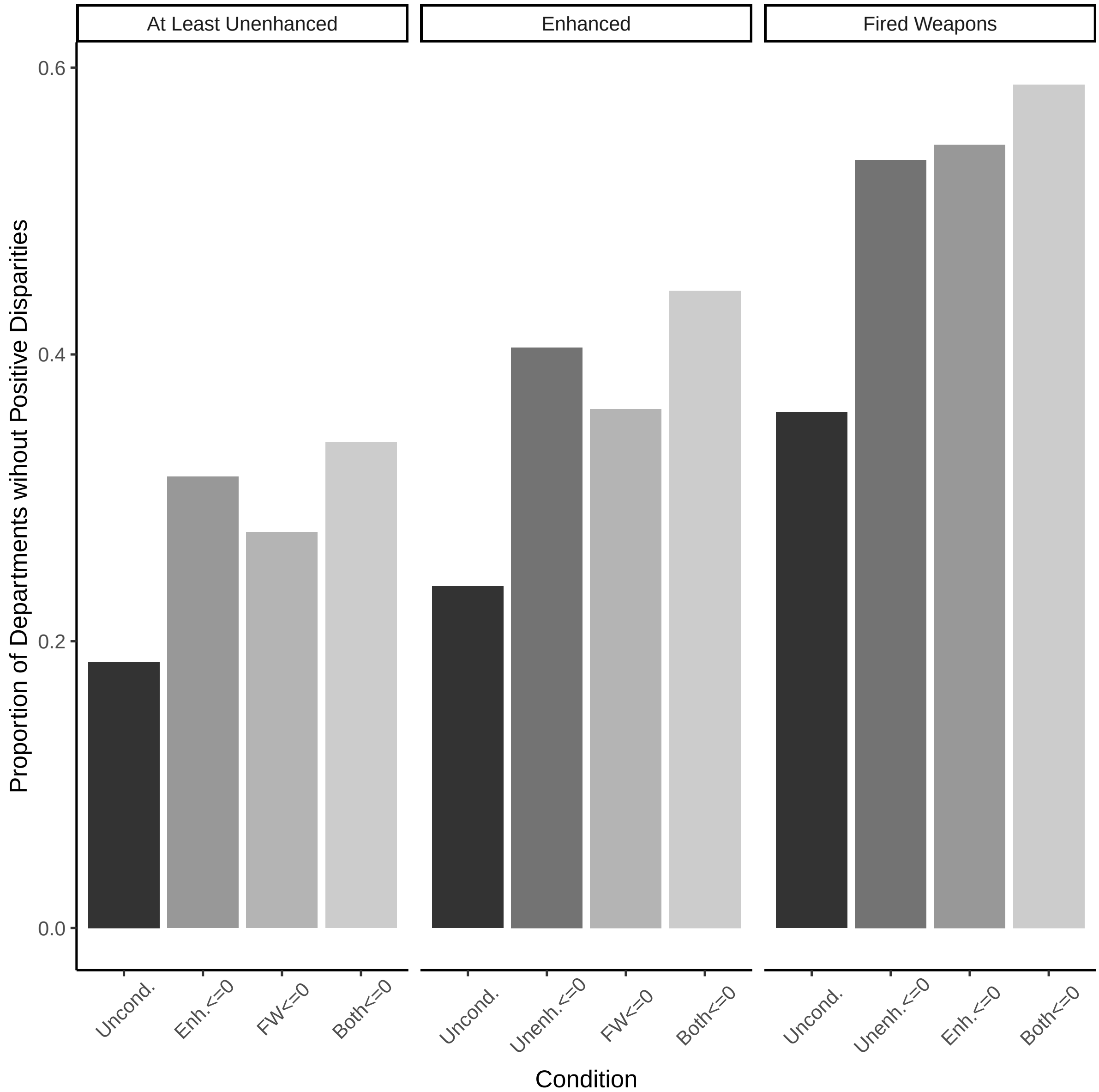

Figure 10: Within-Department Correlations of Winsorized Empirical Bayes Estimates of Department-Black Interactions Across Force Types

Notes: Figure presents conditional probabilities that a given department does not have a disparity in use of force that disfavors Black subjects. For each of the three outcomes (at least unenhanced, enhanced, and fired weapons), there are four bars: one for the unconditional probability that a department does not have a disparity against Blacks, two for the probabilities conditional on each of the other outcomes not having a disparity, and one for the probability conditional on both other outcomes not having disparities.

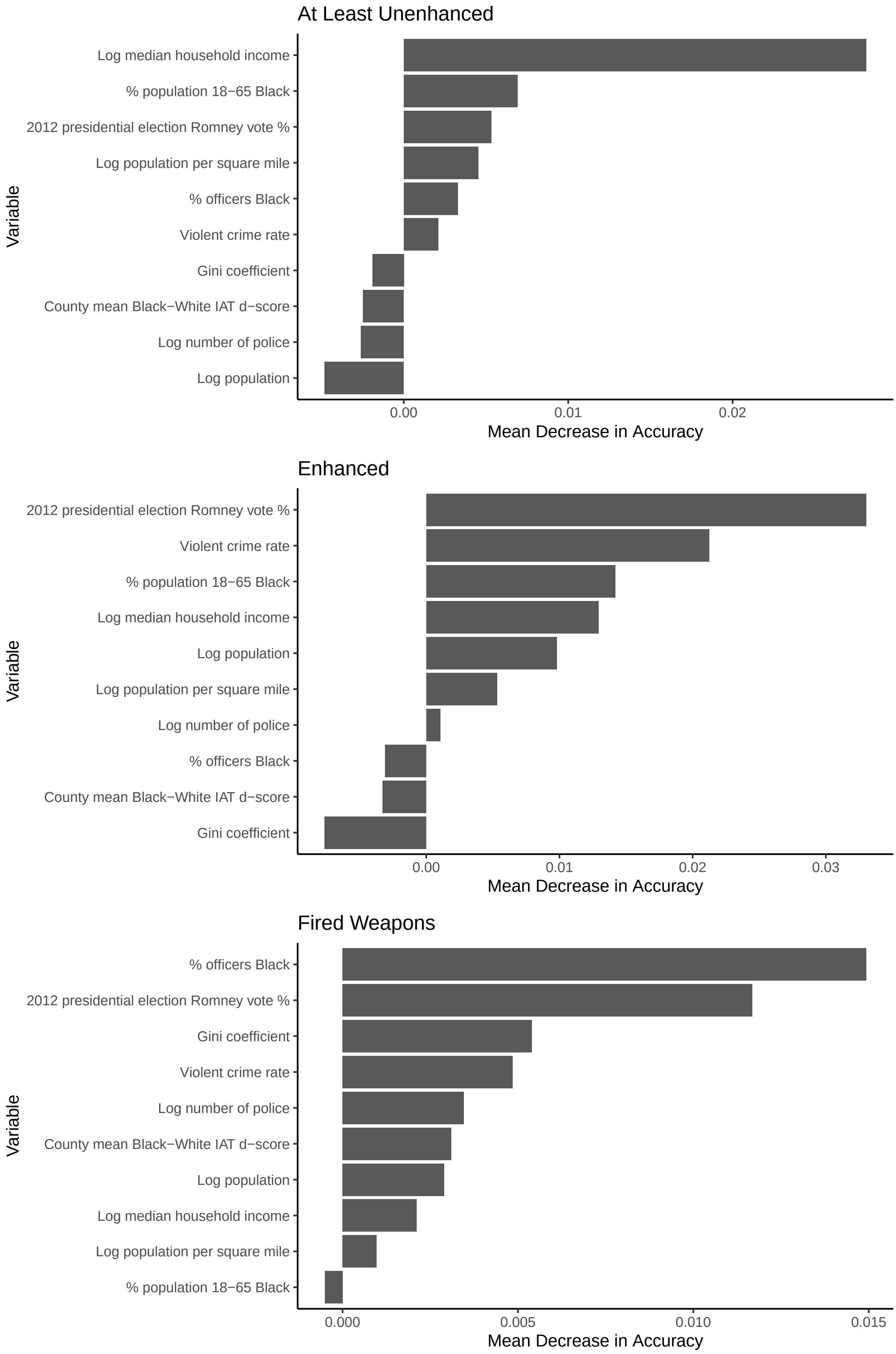

Figure 11: Random Forest Variable Importance Plots

Notes: Figure presents variables' average effects on the accuracy of random forest models on classification models for whether a department has a disparity in use of force against Black subjects for the given force level as described in Section 3.3.

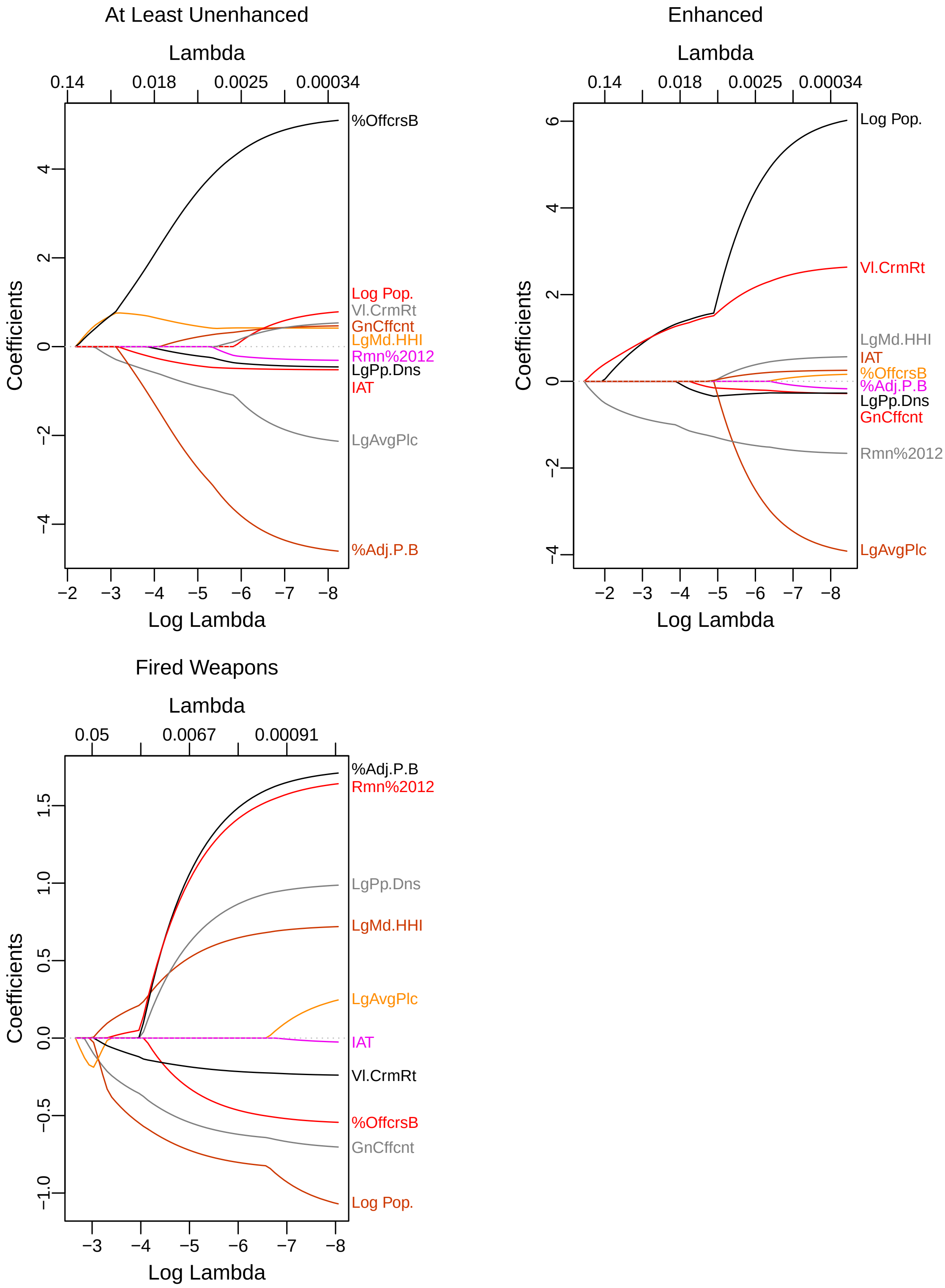

Figure 12: Penalized Logit Coefficient Paths

Notes: Figure presents penalized logit coefficient paths as the value of the penalty parameter lambda is altered for regressions of whether a department has a disparity in use of force against Black subjects for the given force level as described in Section 3.3. Non-binary regressors have been normalized by twice their standard deviation.

References

Ang, Desmond, Panka Bencsik, Jesse Bruhn, and Ellora Derenoncourt (2021). *Police Violence Reduces Civilian Cooperation and Engagement with Law Enforcement*. Working Paper 2021-005. Orlando Bravo Center for Economic Research. URL: https://economics.brown.edu/sites/g/files/dprerj726/files/papers/Bravo%20Working%20Paper%202021-005.pdf.

Antonovics, Kate and Brian G Knight (2009). "A New Look at Racial Profiling: Evidence from the Boston Police Department". *The Review of Economics and Statistics* 91.1, pp. 163–177.

Anwar, Shamena and Hanming Fang (2006). "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence". *American Economic Review* 96.1, pp. 127–151.

Arnold, David, Will Dobbie, and Crystal S. Yang (2018). "Racial Bias in Bail Decisions". *The Quarterly Journal of Economics* 133.4, pp. 1885–1932.

Ba, Bocar, Dean Knox, Jonathan Mummolo, and Roman Rivera (2021). "The Role of Officer Race and Gender in Police-Civilian Interactions in Chicago". *Science* 371.6530, pp. 696–702.

Barrett, Nathan, Andrew McEachin, Jonathan N Mills, and Jon Valant (2019). "Disparities and Discrimination in Student Discipline by Race and Family Income". *Journal of Human Resources*.

Becker, Gary S (1957). *The Economics of Discrimination*. University of Chicago Press.

Bielen, Samantha, Wim Marneffe, and Naci H. Mocan (2019). *Racial Bias and In-group Bias In Judicial Decisions: Evidence from Virtual Reality Courtrooms*. Working Paper 25355. National Bureau of Economic Research. URL: http://www.nber.org/papers/w25355.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope (2020). *Inaccurate Statistical Discrimination: An Identification Problem*. Tech. rep. National Bureau of Economic Research.

Bureau of Justice Statistics (2016). *Law Enforcement Management and Administrative Statistics (LEMAS)*. DOI: 10.3886/ICPSR37323.v1.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff (2014). "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates". *American Economic Review* 104.9, pp. 2593–2632.

Coviello, Decio and Nicola Persico (2015). "An Economic Analysis of Black-White Disparities in the New York Police Department's Stop-and-Frisk Program". *The Journal of Legal Studies* 44.2, pp. 315–360.

Cunningham, Jamein, Donna Feir, and Rob Gillezeau (2021). *Collective Bargaining Rights, Policing, and Civilian Deaths*. Tech. rep. URL: `http://ftp.iza.org/dp14208.pdf`.

Cunningham, Jamein P and Rob Gillezeau (2021). "Don't Shoot! The Impact of Historical African American Protest on Police Killings of Civilians". *Journal of Quantitative Criminology* 37.1, pp. 1–34.

Devi, Tanaya and Roland G. Fryer (2020). *Policing the Police: The Impact of "Pattern-or-Practice" Investigations on Crime*. Working Paper 27324. National Bureau of Economic Research. URL: `http://www.nber.org/papers/w27324`.

Durlauf, Steven N. and James J. Heckman (2020). "An Empirical Analysis of Racial Differences in Police Use of Force: A Comment". *Journal of Political Economy* 128.10, pp. 3998–4002.

Fryer, Roland G. (2019). "An Empirical Analysis of Racial Differences in Police Use of Force". *Journal of Political Economy* 127.3, pp. 1210–1261.

— (2020). "An Empirical Analysis of Racial Differences in Police Use of Force: A Response". *Journal of Political Economy* 128.10, pp. 4003–4008.

Gelman, Andrew (2008). "Scaling Regression Inputs by Dividing by Two Standard Deviations". *Statistics in Medicine* 27.15, pp. 2865–2873.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss (2007). "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* 102.479, pp. 813–823.

Goncalves, Felipe and Steven Mello (Mar. 2017). *A Few Bad Apples? Racial Bias in Policing*. Working Papers 608. Princeton University, Department of Economics, Industrial Relations Section. URL: `https://ideas.repec.org/p/pri/indrel/608.html`.

Grogger, Jeffrey and Greg Ridgeway (2006). "Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness". *Journal of the American Statistical Association* 101.475, pp. 878–887.

Heaton, Paul (2010). "Understanding the Effects of Antiprofiling Policies". *The Journal of Law and Economics* 53.1, pp. 29–64.

Hoekstra, Mark and CarlyWill Sloan (2020). *Does Race Matter for Police Use of Force? Evidence from 911 Calls*. Working Paper 26774. National Bureau of Economic Research. URL: `http://www.nber.org/papers/w26774`.

Horrace, William C. and Shawn M. Rohlin (2016). "How Dark Is Dark? Bright Lights, Big City, Racial Profiling". *The Review of Economics and Statistics* 98.2, pp. 226–232.

Kane, Thomas J. and Douglas O. Staiger (2008). "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation". National Bureau of Economic Research working paper 14607.

Kline, Patrick and Christopher Walters (2019). *Audits as Evidence: Experiments, Ensembles, and Enforcement*. Tech. rep. URL: `https://arxiv.org/abs/1907.06622`.

Knowles, John, Nicola Persico, and Petra Todd (2001). "Racial Bias in Motor Vehicle Searches: Theory and Evidence". *Journal of Political Economy* 109.1, pp. 203–229.

Knox, Dean, Will Lowe, and Jonathan Mummolo (2020). "Administrative Records Mask Racially Biased Policing". *American Political Science Review* 114.3, pp. 619–637.

Luh, Elizabeth (2020). *Not So Black and White: Uncovering Racial Bias from Systematically Misreported Trooper Reports*. Tech. rep. URL: `https://elizluh.github.io/files/highwaybias.pdf`.

McCarthy, Craig (2019). *Help us continue to improve N.J's first statewide database of police use of force*. URL: `https://www.nj.com/news/2018/12/help-us-continue-to-improve-njs-first-statewide-database-of-police-use-of-force.html` (visited on 06/01/2019).

McCarthy, Craig and Stephen Stirling (2019). *How we built the most comprehensive statewide database of police force in the United States*. URL: `https://www.nj.com/news/2018/`

11/how_we_built_the_most_comprehensive_statewide_database_of_police_force_in_the_us.html (visited on 06/01/2019).

Morris, Carl N. (1983). “Parametric Empirical Bayes Inference: Theory and Applications”. *Journal of the American Statistical Association* 78.381, pp. 47–55.

Nelson, Blake (2019). *Frequently asked questions about The Force Report*. URL: https://www.nj.com/news/2018/11/frequently_asked_questions_about_the_force_report.html (visited on 06/01/2019).

Nicholson-Crotty, Sean, Jill Nicholson-Crotty, and Sergio Fernandez (2017). “Will More Black Cops Matter? Officer Race and Police-Involved Homicides of Black Citizens”. *Public Administration Review* 77.2, pp. 206–216.

Nix, Justin, Bradley A Campbell, Edward H Byers, and Geoffrey P Alpert (2017). “A Bird’s Eye View of Civilians Killed by Police in 2015: Further Evidence of Implicit Bias”. *Criminology & Public Policy* 16.1, pp. 309–340.

NJ Advance Media (2019). *The Force Report*. URL: http://force.nj.com (visited on 06/01/2019).

Persico, Nicola and Petra Todd (2006). “Generalising the Hit Rates Test for Racial Bias in Law Enforcement, with an Application to Vehicle Searches in Wichita”. *The Economic Journal* 116.515, F351–F367.

Raphael, Steven and Sandra V. Rozo (2019). “Racial Disparities in the Acquisition of Juvenile Arrest Records”. *Journal of Labor Economics* 37.S1, S125–S159.

Rehavi, M. Marit and Sonja B. Starr (2014). “Racial Disparity in Federal Criminal Sentences”. *Journal of Political Economy* 122.6, pp. 1320–1354.

Ross, Cody T. (2015). “A Multi-Level Bayesian Analysis of Racial Bias in Police Shootings at the County-Level in the United States, 2011–2014”. *PLoS One* 10.11, e0141854.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek (2020). *IPUMS USA: Version 10.0 [dataset]*. Minneapolis, MN. URL: https://doi.org/10.18128/D010.V10.0.

Silverman, Bernard W. (1986). *Density Estimation for Statistics and Data Analysis*. Vol. 26. CRC Press.

Tuttle, Cody (2021). *Racial Disparities in Federal Sentencing: Evidence from Drug Mandatory Minimums*. Tech. rep. URL: https://www.google.com/url?q=https%3A%2F%2Fcodytuttle.github.io%2Ftuttle_mandatory_minimums.pdf&sa=D&sntz=1&usg=AOvVaw1DJP9wLtJKQAfO8lktm6Ji.

Weisburst, Emily K. (2019). "Police Use of Force as an Extension of Arrests: Examining Disparities across Civilian and Officer Race". *AEA Papers and Proceedings* 109, pp. 152–56.

West, Jeremy (2018). *Racial Bias in Police Investigations*. Tech. rep. URL: https://people.ucsc.edu/~jwest1/articles/West_RacialBiasPolice.pdf.

Appendix (For Online Publication Only)

A Data Cleaning

Despite the substantial efforts by the teams at ProPublica and NJ Advance Media, the use of force dataset requires additional processing to be used for analysis. Here I outline the changes I make to the data.

The structure of each department's force reports and how many officers or subjects can be put on a single report vary. For consistency, I structure the data so that every observation is one subject who had force used against them by one officer. When there are multiple force reports for a single subject in an incident, I identify duplicates by computing similarity scores based on the Levenshtein distance between subject names, preventing situations such as a missing middle initial in one report or typos from creating repeat incidents. When officers from different departments use force against the same subject in the same incident as identified by the subject identifier and time of the incident, I continue to keep only the highest use of force from any department. For 167 observations where multiple officers are recorded on the same force report, I use the first officer listed. Because my empirical strategies are based on the most severe level of force used against a subject, for incidents where multiple officers use force against a single subject, I keep the officer who uses the greatest level of force, choosing randomly in the event of ties.

Besides restructuring, the most significant modifications I make to the data are for the types of force used. These are text strings in the raw data and contain many irregularities. Some have typos, some do not directly correspond with force categories (for example, "Grabbed Gardening Tool Out Of Her Hand"), and some forms, notably from the New Jersey State Police, have their own names for force levels, such as "physical" and "mechanical." Many of these force classifications require subjective judgments, which I inform from officer narratives whenever possible. For instances where an officer uses a firearm as a blunt weapon ("pistol whipping"), I mark this as a use of a baton, the most similar type of force. Similarly, when blunt objects such as flashlights are used to strike a subject, I record them as batons. When batons are used for leverage in compliance holds, I count these incidents as compliance holds. I do not count restraints such as handcuffs or “the wrap” as force, but maneuvers to facilitate them may be, such as forcing one’s arms behind their back. Unless there is an indication that a subject was punched, slapped, or otherwise struck, I classify hands on the subject as compliance holds rather than the hands/fists level. If an officer pushes a subject to get them to move, such as pushing them into a police vehicle, I classify it as a compliance hold. When the pushing is done to incapacitate a subject, such as pushing them off of a bicycle, I classify it as a use of hands/fists. I classify forcibly moving actively resisting subjects as compliance holds, but do not count moving passively resisting subjects, such as when a subject sits down and does not move.

Two other variables require manual cleaning: the unique officer identifier and the actions of the subject. The officer identifier variable present in the raw dataset does not track officers as they move across departments. Using available information on race, experience, rank, and geographic location, I manually create an officer identifier variable that follows officers across departments. The subject actions variable is structured in the same manner as the force used variable and suffers from the same inconsistency problems. Again I manually map this variable to the maximum action a subject took that justified force. The lowest category is resisting, which includes physically resisting an officer’s control, fleeing officer apprehension, and other actions that do not fit into another category. The next categories, physical, blunt, knife, and vehicle, cover both attacks and threats of attack. I split incidents involving firearms into threats with firearms and actually firing them. I only count subject actions directed at humans, ignoring behaviors such as subjects attempting to kick out the windows of the police vehicle in which they are being held. I consider flailing and spitting or using other bodily fluids as projectiles to be a physical attack. When a subject attempts to disarm an officer, I mark it as a physical attack.

For the incident type variable, I use the categories present on the New Jersey Attorney General’s model form (Figure A.1): crime in progress, domestic, other dispute, suspicious person, and other, reserving “other” for incidents that do not fit into another category. Some incidents have multiple types, which I allow, except for the “other” category.

I remove a small number of irregular or non-representative observations from the data.

I drop 963 observations (more than 600 of them would have been dropped anyway due to other missing variables) where no reason for the force was given, as it is more likely that the officer-stated reason is missing than nonexistent. I remove a handful of incidents involving nonlethal firearms (“riot guns”), as these are not standard police carry. I do not include the constructive authority category of force, incidents where a firearm or stun gun is drawn but not fired to induce compliant behavior, as no physical force is used in these situations. I remove 44 individuals whose indicated races do not fall within the categories of White, Black, Hispanic, or Asian/Pacific Islander, such as people marked as “mixed.”

B Adjustment for Selection into the Data

By design, my data do not contain incidents where force might have been used but was not. Depending on correlations with race, this extensive margin of force can affect my estimates of racial disparities along the intensive margin. Here, I outline an approach to adjust for this possibility adapted from Appendix F of Goncalves and Mello (2017) and similar to a Heckman correction.

Imagine a two-step model of force usage. For narrative simplicity, consider officers as homogeneous within their departments and ignore covariates (my actual implementation of this strategy includes all covariates from the main model). First, an officer from department j must decide whether to use force against a subject of race i at all, the latent variable Z_{ij} (this subsumes all elements that affect the extensive margin of force, such as where an officer patrols).

$$Z_{ij} = \alpha_j^{NonBlack} + \alpha_j^{Black} \cdot Black_i + \eta_{ij}$$

Assume $\eta_{ij} \sim N(0,1)$. If $Z_{ij} > 0$, then the officer decides to use force against the subject, and chooses the level of force D to use against the subject (assume the binary "at least this severe" outcomes from the main text).

$$D^*_{ij} = \theta_j^{NonBlack} + \theta_j^{Black} \cdot Black_i + \varepsilon_{ij}$$

We observe the actual binary outcome D_{ij} of whether or not force of a certain severity was used, if and only if the officer uses force at all ($Z_{ij} > 0$), and so the racial differences θ I estimate are also based only on observed incidents.

$$\begin{aligned}\hat{\theta}_j^{Black} &= E[D^*_{ij}|Black_i = 1, Z_{ij} > 0] - E[D^*_{ij}|Black_i = 0, Z_{ij} > 0] \\ &= \theta_j^{Black} + E[\varepsilon_{ij}|\eta_{ij} > -\alpha_j^{NonBlack} - \alpha_j^{Black}] - E[\varepsilon_{ij}|\eta_{ij} > -\alpha_j^{NonBlack}]\end{aligned}$$

This two-stage model can lead to problems where the observed subjects are not comparable across races. There are two requirements for this to occur. First, there must of course be a racial difference in the extensive margin of force: $\alpha_j^{Black} \neq 0$. Second, there must also be a correlation in the residuals of the two stages: $corr(\varepsilon_{ij}, \eta_{ij}) \neq 0$. Under those conditions, my estimate of θ_j^{Black} is inconsistent. Unlike in Goncalves and Mello (2017), because I am not estimating a causal treatment effect, my estimates would still be interpretable as the race gap in force experienced conditional on force being used, but this moves away from my target parameter of the difference in force experienced solely from this intensive margin of force with Black and non-Black subjects directly comparable (after adjusting for incident factors).

Suppose that arrests represent an appropriate "denominator" for use of force, i.e., that they approximate the number of possible incidents in which in officer might use force, regardless of whether they actually do. Then define the department-race-specific probability that a subject has force used against them.

$$P_{ij} \equiv P(Z_{ij} = 1) = \frac{\#\text{Force}_{ij}}{\#\text{Arrests}_{ij}}$$

As shown in Goncalves and Mello (2017), it follows that P_{ij} can be used to compute the expectation of the error term η_{ij} for subjects who had force used against them.

$$\begin{aligned}
P_{ij} &= P(\alpha_j^{NonBlack} + \alpha_j^{Black} \cdot Black_i + \eta_{ij} \geq 0) \\
&= \Phi(\alpha_j^{NonBlack} + \alpha_j^{Black} \cdot Black_i) \\
\implies E(\eta_{ij} | Z_{ij} = 1) &= \frac{\phi(\alpha_j^{NonBlack} + \alpha_j^{Black} \cdot Black_i)}{\Phi(\alpha_j^{NonBlack} + \alpha_j^{Black} \cdot Black_i)} \\
&= \frac{\phi(\Phi^{-1}(P_{ij}))}{P_{ij}}
\end{aligned}$$

One can then compute the expectation of this selection bias in the intensive margin disparities.

$$\begin{aligned}
E(\varepsilon_{ij} | \eta_{ij} > -\alpha_j^{NonBlack} - \alpha_j^{Black} \cdot Black_i) &= \rho \cdot E[\eta | \eta_{ij} > -\alpha_j^{NonBlack} - \alpha_j^{Black} \cdot Black_i] \\
&= \rho \cdot E[\eta_{ij} | Z_{ij} = 1]
\end{aligned}$$

Thus, adding the expectation of this term (the Mills ratio) to my regressions adjusts for possible selection bias. An insignificant estimate of the ρ coefficient, as I find in Table 3, suggests minimal impact of any such bias on my estimates.

Table A.1: Summary Statistics for Municipalities Represented in Use of Force Data

Statistic	N	Mean	St. Dev.	Median	Min	Max
Population	454	18,424.43	25,402.93	10,373	296	277,140
Population/square mile)	454	4,020.84	5,616.57	2,612.12	39.13	55,880.00
Median household income	454	87,296.63	31,652.80	83,006	26,214	190,625
Gini coefficient	454	0.43	0.05	0.43	0.33	0.60
Land area (sq. miles)	454	11.16	16.50	3.67	0.10	111.13
Pop. 18-65 % White	454	68.52	22.53	74.44	1.82	99.29
Pop. 18-65 % Black	454	8.41	12.17	3.87	0.00	88.77
Pop. 18-65 % Hispanic	454	14.16	14.30	9.19	0.00	82.94
Pop. 18-65 % Asian/PI	454	7.66	9.15	4.51	0.00	58.60
Violent crimes per 1000	454	1.67	2.59	0.85	0.00	25.66
Romney vote share 2012 presidential election	454	45.20	14.79	47.74	1.31	81.82
County mean Black-White IAT *D*-score	21	0.35	0.05	0.35	0.24	0.41

Notes: Data cover all municipalities present in the force reports after cleaning and processing that are served by their own police department in the force reports data, i.e. not those served by New Jersey State Police. Data have been cleaned as described in Appendix A. Data come from the 2010 Census, 2012-2016 American Community Survey five-year estimates, FBI's Uniform Crime Reporting program, the New Jersey Division of Elections, and Project Implicit.

Table A.2: Summary Statistics for Police Departments Represented in Use of Force Data

Statistic	N	Mean	St. Dev.	Median	Min	Max
Avg. num. full-time police	455	42.75	76.41	23	0	1,088
% officers White	455	72.42	35.01	90.58	0.00	100.00
% officers Black	455	2.64	6.80	0.00	0.00	77.81
% officers Hispanic	455	2.04	5.69	0.00	0.00	53.49
% officers Asian/Pacific Islander	455	0.53	2.13	0.00	0.00	30.43
Arrests 2012-2016	455	2,959.03	7,190.05	1,327	67	117,680
Force incidents 2012-2016	455	86.42	181.12	34	1	1,818

Notes: Data cover all police departments in New Jersey from 2012 through 2016 with force reports after cleaning and processing. Not all departments' racial breakdowns sum to 100%. Data from ProPublica and NJ Advance Media.

Table A.3: Racial Disparities by Police in Force of At Least Specified Severity, Conditional on Force (No Controls)

	Unenhanced	Enhanced	Fired Weapons
Subject Black	0.082^{***}	0.046^{***}	0.002^{*}
	(0.011)	(0.011)	(0.001)
Clustering	Dept.	Dept.	Dept.
Outcome mean	0.505	0.146	0.004
R^2	0.007	0.004	0.000
Num. obs.	39322	39322	39322

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from regressions where the outcome is a binary measure of whether police used force of at least the specified severity on indicators for subject race with no other covariates.

Table A.4: Racial Disparities by Police in Force of At Least Specified Severity, Conditional on Force

	Unenhanced		Enhanced		Fired Weapons	
Subject Black	0.019^{***}	0.018^{**}	0.033^{***}	0.038^{***}	0.001	0.001
	(0.005)	(0.008)	(0.005)	(0.006)	(0.001)	(0.001)
Fixed effects	Dept.	Off.	Dept.	Off.	Dept.	Off.
Clustering	Dept.	Off.	Dept.	Off.	Dept.	Off.
Outcome mean	0.505	0.505	0.146	0.145	0.004	0.004
R^2	0.181	0.512	0.118	0.466	0.181	0.539
Num. obs.	39182	39125	39182	39125	39182	39125

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$ *Notes:* Table reports OLS estimates from Equation 1, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, and department or officer fixed effects.

Table A.5: Racial Disparities by Police in Force of At Least Specified Severity, Conditional on Force

	Unenhanced		Enhanced		Fired Weapons	
Subject Black	0.019***	0.020**	0.034***	0.040***	0.001	0.001
	(0.006)	(0.009)	(0.005)	(0.006)	(0.001)	(0.001)
Subject Hispanic	−0.000	0.008	0.004	0.006	0.001	0.002
	(0.010)	(0.013)	(0.007)	(0.010)	(0.002)	(0.002)
Subject Asian/PI	−0.022	0.003	0.004	0.010	−0.005***	−0.003
	(0.026)	(0.040)	(0.017)	(0.026)	(0.002)	(0.002)
Fixed effects	Dept.	Off.	Dept.	Off.	Dept.	Off.
Clustering	Dept.	Off.	Dept.	Off.	Dept.	Off.
Outcome mean	0.505	0.505	0.146	0.145	0.004	0.004
R^2	0.181	0.512	0.118	0.466	0.181	0.539
Num. obs.	39182	39125	39182	39125	39182	39125

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from Equation 1, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, and department or officer fixed effects.

Table A.6: Conditional Logit Odds Ratios on Race in Intensity of Force Used, Conditional on Force

	Unenhanced	Enhanced	Fired Weapons
Subject Black	1.054***	1.285***	1.058
	(1.024, 1.086)	(1.203, 1.372)	(0.699, 1.601)
Fixed effects	Dept.	Dept.	Dept.
Clustering	Dept.	Dept.	Dept.
Num. obs.	39322	39322	39322

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Notes: Table reports conditional logit estimates from Equation 2, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, and year fixed effects, stratified by department. Asymmetric 95% confidence intervals with clustering at the department level based on exponentiating the log odds confidence interval are in parentheses.

Table A.7: Racial Disparities by Police in Force of At Least Specified Severity, Conditional on Force, Incidents where Race Likely Unrelated to Decision to Engage with Subject and Subject at Least Physically Threatened/Attacked Officer or Another

	Unenhanced		Enhanced		Fired Weapons	
Subject Black	0.045***	0.024	0.064***	0.063**	0.001	−0.001
	(0.013)	(0.030)	(0.014)	(0.030)	(0.002)	(0.004)
Fixed effects	Dept.	Off.	Dept.	Off.	Dept.	Off.
Clustering	Dept.	Off.	Dept.	Off.	Dept.	Off.
Outcome mean	0.678	0.678	0.245	0.245	0.009	0.009
R^2	0.190	0.736	0.140	0.714	0.276	0.847
Num. obs.	7789	7777	7789	7777	7789	7777

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: Table reports OLS estimates from Equation 1, which regresses a binary measure of whether police used force of at least the specified severity on indicators for the subject's race, an indicator for the subject being female, a quadratic of the subject's age, time of day fixed effects, indicators for the subject's actions, an indicator for whether an officer was harmed, an indicator for whether the subject was an emotionally disturbed person, indicators for the officer's rank, year fixed effects, and department or officer fixed effects. These regressions are run on a subset of the data where the following two conditions hold. 1) The subject at least physically threatened or attacked an officer or another person. 2) The incident's type is one where the subject's race was less likely to have influenced the officer's decision to engage with the subject: crimes in progress, disputes, and traffic stops at night.

______________________ POLICE DEPARTMENT

USE OF FORCE REPORT

A. Incident Information

Date	Time	Day of Week	Location	**INCIDENT NUMBER**
Type of Incident ☐ Crime in progress ☐ Domestic ☐ Other dispute ☐ Suspicious person ☐ Traffic stop ☐ Other (specify)				

B. Officer Information

Name (Last, First, Middle)		Badge #	Sex	Race	Age	Injured Y / N	Killed Y / N
Rank	Duty assignment	Years of service		On-Duty Y / N		Uniform Y / N	

C1. Subject 1 (List only the person who was the subject of the use of force by the officer listed in Section B.)

Name (Last, First, Middle)	Sex	Race	Age	Weapon Y / N	Injured Y / N	Killed Y / N
☐ Under the influence ☐ Other unusual condition (specify)	Arrested Y / N		Charges			

Subject's actions (check all that apply)	Officer's use of force toward this subject (check all that apply)	
☐ Resisted police officer control	☐ Compliance hold	Firearms Discharge
☐ Physical threat/attack on officer or another	☐ Hands/fists	☐ Intentional
☐ Threatened/attacked officer or another with blunt object	☐ Kicks/feet	☐ Accidental
☐ Threatened/attacked officer or another with knife/cutting object	☐ Chemical/natural agent	
☐ Threatened/attacked officer or another with motor vehicle	☐ Strike/use baton or other object	Number of Shots Fired _____
☐ Threatened officer or another with firearm	☐ Canine	Number of Hits _____
☐ Fired at officer or another		[Use 'UNK' if unknown]
☐ Other (specify)	☐ Other (specify)	

C2. Subject 2 (List only the person who was the subject of the use of force by the officer listed in Section B.)

Name (Last, First, Middle)	Sex	Race	Age	Weapon Y / N	Injured Y / N	Killed Y / N
☐ Under the influence ☐ Other unusual condition (specify)	Arrested Y / N		Charges			

Subject's actions (check all that apply)	Officer's use of force toward this subject (check all that apply)	
☐ Resisted police officer control	☐ Compliance hold	Firearms Discharge
☐ Physical threat/attack on officer or another	☐ Hands/fists	☐ Intentional
☐ Threatened/attacked officer or another with blunt object	☐ Kicks/feet	☐ Accidental
☐ Threatened/attacked officer or another with knife/cutting object	☐ Chemical/natural agent	
☐ Threatened/attacked officer or another with motor vehicle	☐ Strike/use baton or other object	Number of Shots Fired _____
☐ Threatened officer or another with firearm	☐ Canine	Number of Hits _____
☐ Fired at officer or another		[Use 'UNK' if unknown]
☐ Other (specify)	☐ Other (specify)	

➤ **If this officer used force against more than two subjects in this incident, attach additional USE OF FORCE REPORTS.**

Signature:	Date:
Print Supervisor Name:	Supervisor Signature:

7/2001

Figure A.1: New Jersey Model Use of Force Report

Notes: Figure obtained from the website of the New Jersey Attorney General.

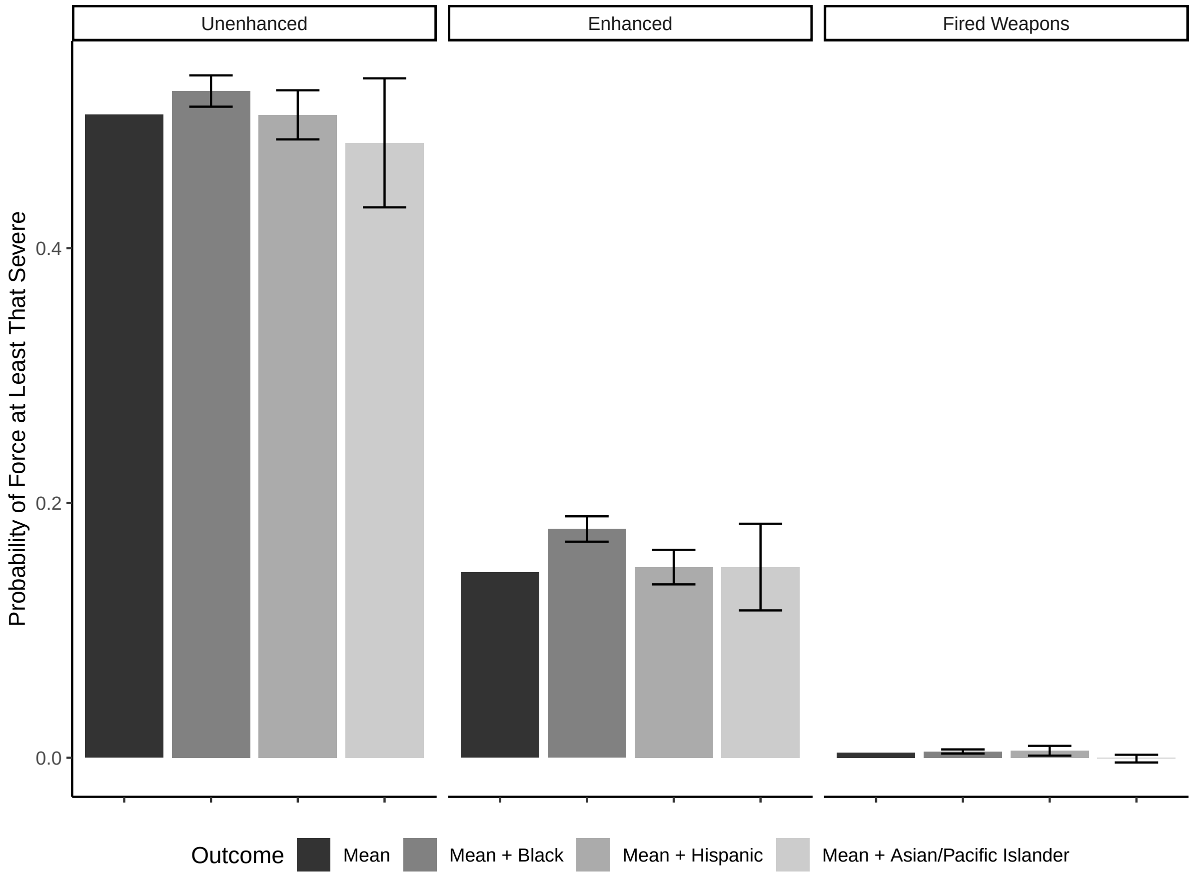

Figure A.2: Overall Racial Disparities of Subject Being Black on Probability of Force of at Least Specified Severity, Full Race Dummies (Department Fixed Effects and Clustering)
Notes: Figure presents results from a series of OLS models fit via Equation 1. Each heading represents a different outcome: whether, conditional on any force being used, force of at least the specified severity was used. Confidence intervals are based on the corresponding race coefficient. Standard errors are clustered at the department level.

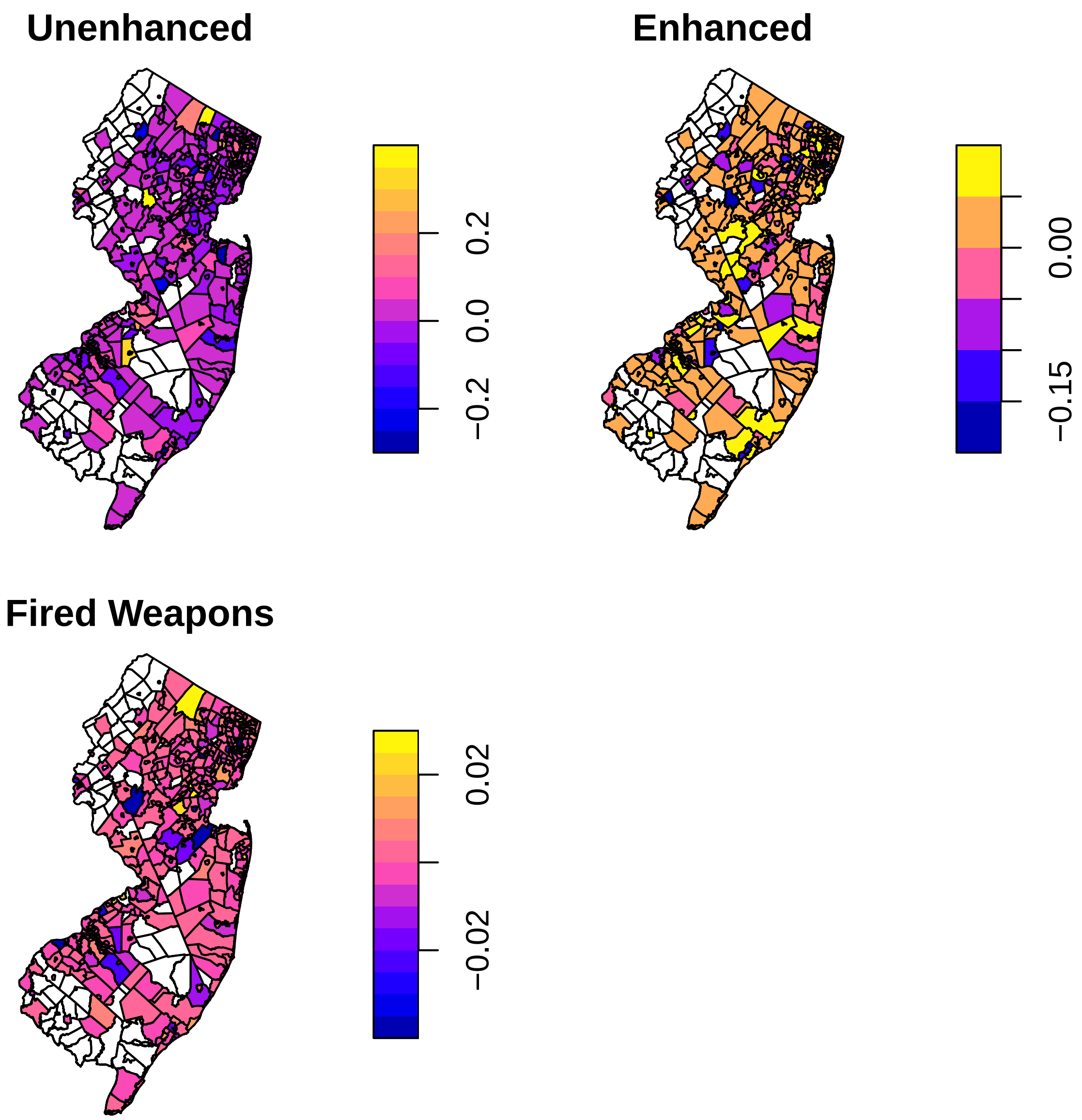

Figure A.3: Heatmap of of Winsorized Empirical Bayes Estimates of Department-Black Interactions

Notes: Figure presents heatmaps of department-specific racial disparities β between White/Asian/Pacific Islander subjects and Black subjects as estimated from the empirical Bayes estimator in Equation 5. Each subgraph shows results from regressions with the specified outcome outcome: whether, conditional on any force being used, force of at least the specified severity was used. Municipalities in white are not present in the data.

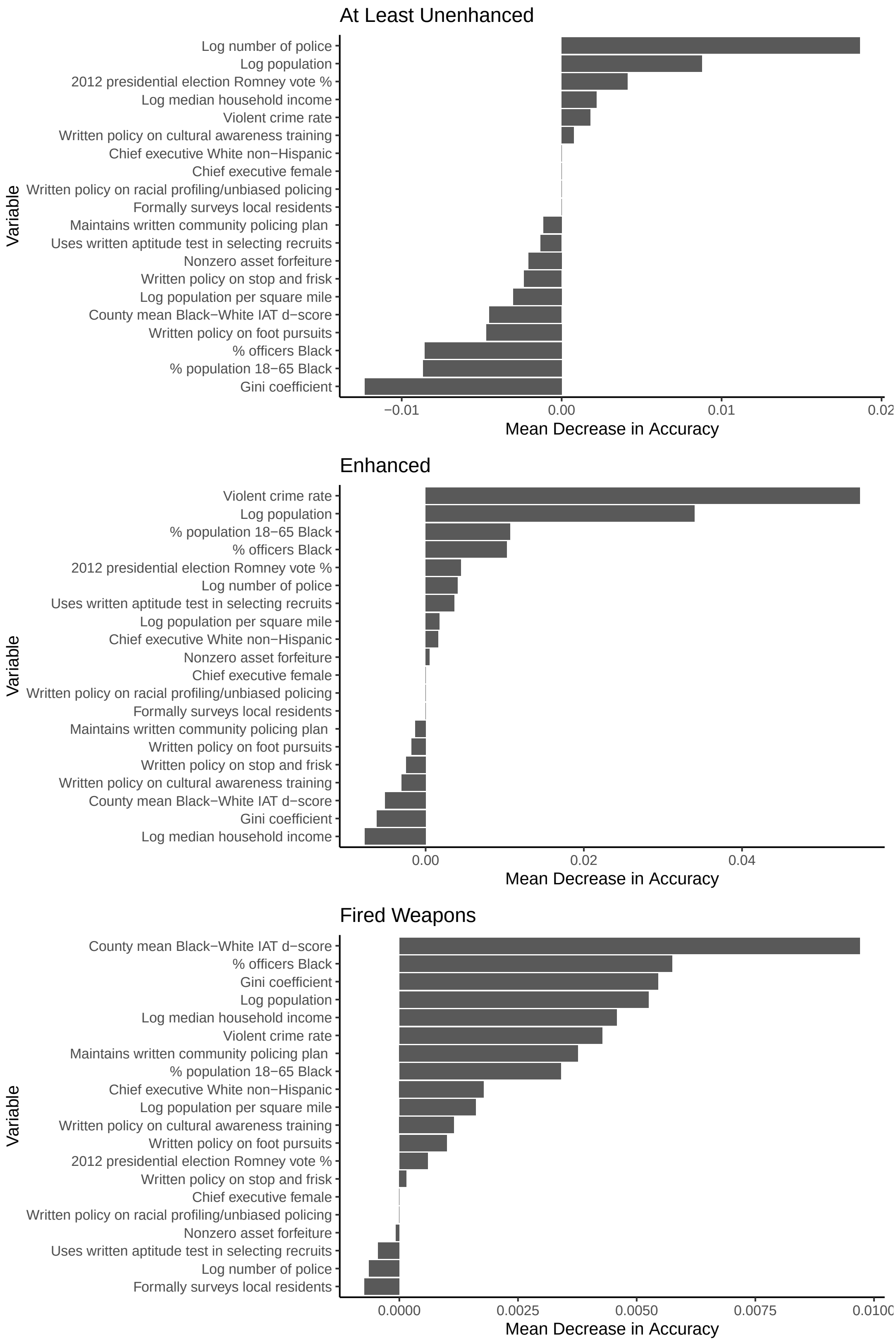

Figure A.4: Random Forest Variable Importance Plots (with LEMAS Variables)

Notes: Figure presents variables' average effects on the accuracy of random forest models on classification models for whether a department has a disparity in use of force against Black subjects for the given force level as described in Section 3.3.

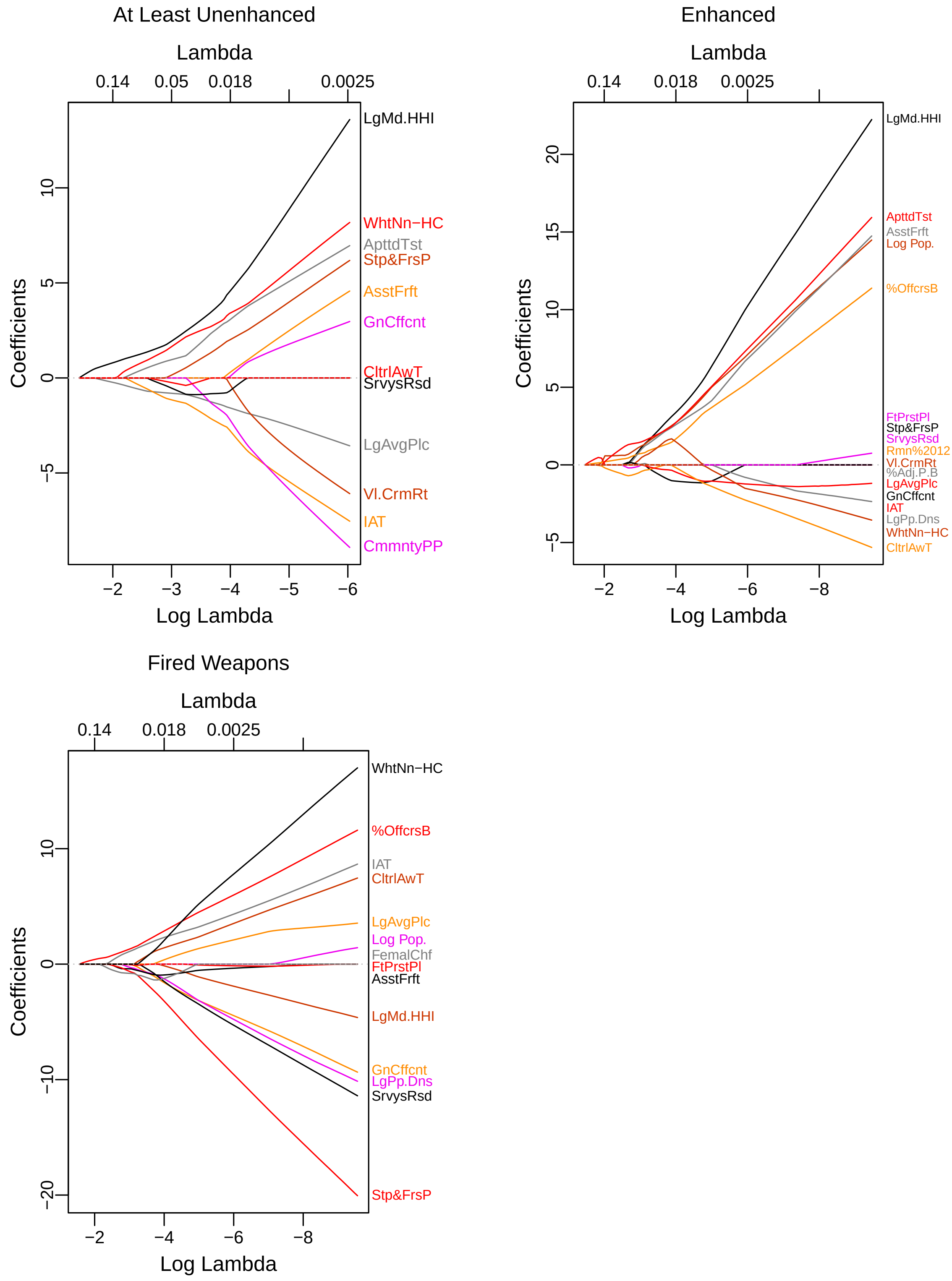

Figure A.5: Penalized Logit Coefficient Paths (with LEMAS Variables)

Notes: Figure presents penalized logit coefficient paths as the value of the penalty parameter lambda is altered for regressions of whether a department has a disparity in use of force against Black subjects for the given force level as described in Section 3.3. Non-binary regressors have been normalized by twice their standard deviation.