The Black-White Recognition Gap

in Award Nominations[*]

Nayoung Rim[†]   Roman Rivera[‡]   Andrea Kiss[§]   Bocar Ba[¶]

September 18, 2020

There is substantial evidence showing racial bias in firms' hiring decisions, but less is known about bias in career recognition or promotion, which may arguably be more important for the lack of diversity in upper-management positions and, ultimately, the racial wage gap. We construct a novel dataset of police nominations for awards to measure bias against minority employees in career recognition. Exploiting quasi-random variation in supervisor assignment, we find that white supervisors are less likely to nominate black officers for awards than white or Hispanic officers. Increased supervisor-officer interaction reduces, but does not eliminate, the black-white recognition gap. We also conduct an online experiment and find respondents are less likely to acquire information about black officers relative to non-black officers. Our findings suggest bias in career recognition may be important for the black-white earnings gap. In regards to policing, our findings suggest that racial issues in policing are not just at issue between police and the public, but also within departments, and thus that simply hiring minority officers may be limited in its efficacy.

JEL Codes: J71, M51, J45, J48
Keywords: racial bias, police officers, award nominations, supervisors

---

[†]Corresponding author. U.S. Naval Academy, Department of Economics. Email: `rim@usna.edu`
[‡]Columbia University, Department of Economics. Email: `r.g.rivera@columbia.edu`
[§]Duke University, Department of Economics. Email: `a.kiss@duke.edu`
[¶]University of California-Irvine, Department of Economics. Email: `bba@uci.edu`

1 Introduction

For decades, a goal of public policy has been to reduce racial disparities in the labor market. The economics literature has largely focused on firms' hiring decisions because of the ability to experimentally examine hiring (Bertrand and Mullainathan, 2004; Kessler et al., 2019; Neumark et al., 2019). Less is known about racial bias in career recognition and progression, which may arguably be more important for the lack of diversity in upper-management positions and, ultimately, the racial wage gap. This paper uses detailed personnel information, including supervisor assignment, to measure the black-white recognition gap among public sector employees.

Recent studies have shown that, in the face of statistical discrimination, full information or repeated contact may eliminate or even reverse discrimination (Bohren et al., 2019; Rao, 2019) that may arise due to stereotypes (Bordalo et al., 2016). However, in most work settings, evaluators have agency over whether to change their information set or to update their beliefs. For example, Bartoš et al. (2016) find that employers are less likely to open and read resumes from minority candidates. Therefore, an important question for eliminating discrimination and racial gaps in career outcomes is whether supervisors choose to interact with and acquire information about minority colleagues.

We examine this question in the context of the second largest police department in the US, where supervisors do not necessarily observe the officer's day-to-day activities but are required to evaluate the officer's performance annually. Because supervisors do not directly monitor officers, they must exert effort to gather information on officers when it comes to the annual evaluation. This may manifest in a racial gap in career recognition and progression. Although our application focuses on law enforcement, this organizational structure (autonomous workers operating within a hierarchical organization) is common across all industries.

We construct a novel panel dataset of all Chicago Police Department (CPD) officers between 2009 and 2015 containing detailed personnel information. Most importantly, this

dataset contains information on officer performance, such as use of force, arrests, and misconduct—crucial information in an empirical study of bias in the workplace. Using supervisor nominations for departmental awards, we examine whether white supervisors are less likely to acquire information about and nominate their minority officers.

Our identification strategy exploits two institutional features of the CPD that allows us to obtain plausibly causal estimates of the black-white recognition gap. First, officers are assigned a new supervisor every January. We confirm as-good-as-random assignment by analyzing supervisor-officer assignments and find that officers do not sort to same-race supervisors based on work performance measures.[1] Second, all officers must be evaluated annually by their supervisor and the quarter of evaluation is randomized across officers.[2] Under the assumption that supervisors are more likely to interact with and gather information about officers closer to the evaluation period, this institutional feature allows us to exploit the randomly assigned evaluation quarter and estimate the causal impact of an interaction, which would normally be endogenous.

Building off of Rim et al. (2020), we first document that an officer's arrest record has a positive and significant impact on the likelihood of being nominated for an award. But, conditional on the number of arrests, black officers are less likely to be nominated for an award relative to white officers, and the negative black-white gap widens with the number of arrests. Specifically, the black-white gap nomination gap among officers with five or more arrests is -2.6 percentage points, an effect size of 82.6 percent as the mean nomination probability for a white officer by a white supervisor is 3.1 percent. The black-white recognition gap is largest among less experienced supervisors and among supervisors with a higher share of black employees, and does not vary much by the number of supervisees. We also find that supervisors are more likely to nominate all officers in the quarter of evaluation, suggesting

---

[1]In particular, we may be concerned that more-productive white officers and/or less-productive black officers sort to white supervisors. In this case, we would see a negative black-white nominations gap from white supervisors even in the absence of racial bias.

[2]The evaluation must be held in the quarter prior to the quarter that the officer joined the CPD, and the quarter in which officers join CPD is determined by lottery number.

that statistical discrimination may be at play (Altonji and Pierret, 2001). However, this increase in nomination likelihood is larger for white officers relative to black officers and persists only for white officers, suggesting that statistical discrimination is not the only explanation for this racial disparity. If lack of information were the only reason for the racial gap in award nominations, we would expect that full information in the evaluation quarter would eliminate the racial gap.

Because our observational CPD data do not capture detailed interactions between supervisor and officer, we conduct an online experiment to measure the review process in the nomination decision. We ask Amazon Mechanical Turk (MTurk) workers to evaluate officer profiles and nominate one for an award. In addition to being able to experimentally examine the nomination decision process, the online experiment allows us to generalize our findings to a broader evaluator group than Chicago police supervisors.

In one task, MTurk workers choose between a black officer and a non-black officer. Although officer performance levels are randomly chosen, MTurk workers are 6-9.3 percentage points less likely to nominate black officers over white officers. In another task, officer profiles display only demographic information and workers must mouse over the profile to reveal full information about the officer. We monitor mouse movements across the screen and find that black officers are less likely to be moused over. This result is more pronounced when workers are choosing among three white officers and one black officer. We do not see similar patterns for Hispanic officers when workers are choosing among three white officers and one Hispanic officer. Taken together, our findings suggest that the negative black-white nomination gap is due to a lack of interest about black officers.

Our paper relates to the literature on social networks in the workplace. Prior research documents the importance of gender homophily on career outcomes (Cullen and Perez-Truglia, 2020; Sarsons, 2019; Zeltzer, 2020). This paper expands the literature by examining the importance of race homophily in career recognition. For example, we find that CPD supervisors and MTurk workers, who are mostly white, are less likely to gather information

on minority officers, leading to a racial disparity in award nominations. Furthermore, CPD supervisors continue nominating their white officers but not their black or Hispanic officers after the annual evaluation. These findings also speak to the literature that documents benefits of same-race matching on racially disparate outcomes. There is mounting evidence that race-matching leads to better outcomes, such as in education (Carrell et al., 2010; Gershenson et al., 2018; Kofoed et al., 2019) and health (Alsan et al., 2019), but less is known about why. Our results suggest that networks formed through race homophily are important for success in the workplace.

Our paper is also similar to other papers that find that discrimination may arise because biased managers interact less with minorities (Glover et al., 2017) or because managers pay less attention to information about minority groups (Bartoš et al., 2016). Our experimental evidence finds that black officers are less likely to be moused over, particularly when they are in a pool with three other white officers. Additionally, MTurk workers spend more time evaluating black officers' profiles but are not more likely to nominate them. These findings are consistent with studies that find that minorities are less likely to be acknowledged for their work (Hengel, 2019; Sarsons, 2020) and a strand of literature that establishes the existence of bias among managers and work colleagues (Bertrand and Mullainathan, 2004; Bohren et al., 2019; Egan et al., 2018; Giuliano et al., 2009; Glover et al., 2017; Sarsons, 2019). By analyzing the black-white recognition gap among police officers, our paper links this literature to studies on racial disparities in law enforcement.

With respect to law enforcement, our study adds to the growing research that is uncovering racial bias in policing.[3] Prior studies largely use data on officer-initiated encounters, which may be biased because they do not include the universe of all possible police interactions (Knox et al., 2020). Notable exceptions exploit randomly assigned officer dispatches to 911 calls or to investigate automobile crashes (Weisburst, 2018; West, 2018). Similarly,

---

[3]See, for example, Ajilore and Shirey (2017); Antonovics and Knight (2009); Anwar and Fang (2006); Bacher-Hicks and de la Campa (2020); Close and Mason (2006); Cunningham and Gillezeau (2018); Goncalves and Mello (2020); Hoekstra and Sloan (2020); Horrace and Rohlin (2016); Knowles et al. (2001); Mason (2007); Nix et al. (2017).

our paper bypasses the truncated data problem by focusing on supervisor nominations of quasi-randomly assigned officers. Additionally, we ask whether racial bias on the part of officers carries over to their colleagues, a question that was previously unanswered due to a lack of detailed personnel data.

We begin the rest of the paper with a short description of CPD's organizational structure (Section 2). Section 3 describes our data collection efforts and presents summary statistics on our CPD analysis sample. Section 4 presents our empirical strategy and tests the identifying assumption that officers do not sort to same-race supervisors based on work performance measures. We present results using observational CPD data in Section 5. Section 6 discusses the experimental evidence. We conclude with a discussion of the policy implications for law enforcement agencies in Section 7.

2   Basic Facts about CPD's Structure

After passing a written exam, all Chicago Police Department candidates are placed on an eligibility list according to a randomly assigned lottery number and called off in lottery order to enroll in police academy. Upon graduation from Police Academy, all Police Officers begin their career in one of the 25 geographic districts spanning the city of Chicago.[4] These initial assignments are permanent and outside the officer's control, with the exception of a small number of officers who received academic and other distinctions in the Academy (Police Accountability Task Force, 2016). An officer may transfer to another district after the initial assignment pursuant to a bidding process, which is generally based on seniority.

Police officers are supervised by Sergeants, who are normally assigned according to the officer's district assignment.[5] Sergeants prepare officers for duty and roll call, monitor officer activity, evaluate officers' performance annually.[6] Promotions to Sergeant are based on a

---

[4]Between 2012-2014, three districts were dissolved leaving 22 geographic districts.

[5]Among supervisors for whom we have unit information, 82% are assigned to officers in the same unit.

[6]Chicago Police Department, Career Development Directive, Employee Resource E05-01, retrieved from http://directives.chicagopolice.org/directives/data/

written exam and assessment exercise. However, up to 30% of promotions may be made through a merit selection process. Police Officers are eligible to take the Sergeant exam after five years of service.

The Chicago Police Department distributes department awards to recognize the accomplishments, performance, and service of its Department members. In addition to highlighting officers' accomplishments, awards are used for officer evaluations and merit promotions.[7] Most awards require a nomination process. Nominations may originate from any higher-ranking officer, including one's supervisor. Our analysis focuses on nominations by officially assigned supervisors.

3 Data

This section describes administrative police records and district-level crime information that are used for our empirical analysis. We first describe the data sources and the linked analysis dataset. Then, we provide descriptive statistics of Police Officers in the Chicago Police Department between 2009 and 2015.

3.1 Police Officer Data

Administrative records and information on sworn Chicago Police Department members were obtained by Freedom of Information Act requests through a collaboration with Invisible Institute. In order to connect different datasets, officers are first identified within a dataset using unique characteristics available, like name, appointed date, birth year, and race, and then matched with identified officers in different datasets.

Demographics   Data on officer race, sex, birth year, and appointment date are obtained

a7a56e3d-12887ea9-ce512-887e-c3dce7cd73e28d57.html?hl=true

[7]Chicago Police Department, Career Development Directive, Employee Resource E05-01, retrieved from http://directives.chicagopolice.org/directives/data/a7a56e3d-12887ea9-ce512-887e-c3dce7cd73e28d57.html?ownapi=1

from aggregated data, using the most common observation across datasets.[8] Officer rank is taken from salary data provided by the Chicago Department of Human Resources (DHR), covering 2002 to 2017.

Supervisors   This dataset provides information about the supervisor who conducted each officer's annual evaluation between 2009 and 2017. Our analysis focuses on those at the rank of Police Officer, meaning their supervisors are at the rank of Sergeant. In this paper, the term "supervisor" refers to a Sergeant who is officially assigned to conduct a Police Officer's annual evaluation in a given calendar year.

Awards   The awards dataset provides information on all department award nominations between 2004 and 2017. The dataset includes the award name, the individual being nominated, the requester, request date, and the final status of the nomination (approved, deleted, or denied).[9] We consider all performance awards that are open to all sworn Department members and require a supervisor's nomination.[10] After these restrictions, our analysis considers 18 awards. Appendix Table A1 provides a description of these awards.

Unit Assignment   Historical unit assignment data lists all units to which an officer was assigned since the beginning of his or her career, as well as start- and end-dates in each unit. We focus our analysis on Police Officers assigned to the 25 geographic districts.

Arrests   The arrests dataset contains information on all arrests made by Department members. The dataset includes detailed information about the subject, crime, and arrest location and time. These data cover 2001-2017 but arrest day and month are only provided from 2010 onwards. For arrests made in 2009, we use the date the subject was released from the local police station as a proxy for the arrest date.[11] For our analysis, we use total officer

---

[8]Not all demographic information is complete in each file, so an aggregation of demographic variables across multiple files is necessary. Over 99% of officers are matched to a unique gender, race, and appointment date.

[9]An award may be deleted for various reasons, including: the form was not filled out correctly; supporting evidence was not included; or the nomination does not meet the eligibility requirements of the award. This differs from an award denial, which means the officer did not win the award.

[10]Most awards are open to all Department members. One example of an exception is the Thomas Wortham IV Military and Community Service Award, which is awarded to current or former members of the U.S. Armed Services.

[11]In 96.9% of cases, the release date is on the same day or the day after the arrest date, and 100% of

arrests as well as arrests in three aggregated crime categories: violent crime, property crime, and "non-index" crime. The Federal Bureau of Investigation classifies violent and property crimes as "index crimes" because they are more serious offenses.[12] Non-Index crimes capture crimes that are not related to violence or property, such as municipal code violations, traffic violations, warrants, drugs, prostitution, gambling, etc.[13]

Complaints The complaints data contain all recorded allegations of misconduct filed against an officer from 2000 to 2016. Allegations may originate from the public or from other officers in the department.

Tactical Response Reports Data on officer use of force come from 2004-2016 Tactical Response Reports (TRR). Officers are required to file a TRR if they used any force while performing their duties. A TRR filing requirement can be triggered by three things: the subject's actions; the officer's actions; or a subject who is injured or alleges injury resulting from the officer's use of force option. CPD publishes a Use of Force Model, which provides guidelines on the appropriate level of force to be used in response to a subject's actions and levels of resistance. Using the Use of Force Model as a guide, we classify officer force options into two broad categories of "weak use of force" and "strong use of force." Weak use of force includes force mitigation efforts, such as verbal direction and tactical positioning (which involve no physical touch), and control tactics, such as escort holds and wristlocks. Strong uses of force involve elevated levels of force that are generally intended to enact harm on or injure the subject.[14] The data only report use of force against adult persons. Appendix Table A2 outlines force options and our classification.

Sample restrictions. To construct a complete dataset on all officers in the Chicago Police Department, we require that officers receive a salary from DHR and appear in the

---

release dates are within four days of the arrest.

[12]Violent crimes are crimes related to violence, such as murder and assault. Property crimes are crimes related to property, such as burglary and motor vehicle theft.

[13]A comprehensive list of crime categories can be found at http://gis.chicagopolice.org/clearmap_crime_sums/crime_types.html.

[14]Strong use of force may or may not use weapons. Examples of strong use of force without weapons are take-downs, kicks, and punches. Examples of non-lethal weapons are chemical weapons and long-range acoustic devices. Examples of lethal weapons are tasers, batons, and firearms.

unit assignment dataset. We focus on years 2009 to 2015 to maximize overlap across the different datasets. We further restrict our sample to officers at the rank of Police Officer who are always assigned to a geographic district[15] and officer-supervisor relationships that lasted for 12 months. Our final analysis dataset has 6,518 Police Officers and 1,284 supervisors.

3.2 Crime Data

We use crime data from the Chicago Data Portal (https://data.cityofchicago.org), which contains reported incidents of crime that occurred in the City of Chicago since 2001. The dataset contains the primary type of crime, the date, location, and whether the crime led to an arrest. We construct monthly crime rates[16] for each district, separately for total crimes, property crimes, and violent crimes. To capture a district's productivity, we also construct rates for crimes that led to an arrest.

3.3 Summary Statistics

This section provides descriptive statistics of Police Officers in our analysis sample. From Table 1, we see that most officers are male (73.7%) and white (46.4%), but blacks and Hispanics are also well-represented (23-27%). In fact, these three racial groups make up nearly 97% of our sample. The average CPD officer in our sample joined the force in 2000 at age 30. This indicates that at the start of our analysis dataset (2009), the average officer had been on the force for 9 years.

Relative to Police Officers, the racial makeup of supervisors in our analysis sample is more homogeneous. About 81% of supervisors are male, and 70% are white. Blacks and Hispanics each make up around 14% of supervisors. At the start of our analysis dataset, the average supervisor had worked for 17 years or 8 years longer than the average Police Officer.

---

[15]We remove the three districts that closed between 2012-2014 (13, 21, and 23) from our analysis sample because we do not have crime statistics for these districts.

[16]Crime rate is defined as the total number of reported incidents of crime divided by the population and multiplied by 1000.

Table 1: Summary Statistics

Sample:	Police Officers	Supervisors
Male	73.7%	80.8%
Race		
White	46.4%	69.7%
Black	26.8%	14.7%
Hispanic	23.2%	14.0%
Asian	3.1%	1.6%
Native American	0.4%	0.1%
Birthyear	1970.3	1965.3
Start Year	2000.0	1992.2
N	6,518	1,284

Source: CPD analysis sample.

Table 2 presents racial differences in various work measures. The first row is the probability of being nominated for an award in a particular month. For example, the average officer has a 2.5% chance of being nominated in a given month, which equates to about a 30% chance of being nominated in a given year. Whites and Hispanics have slightly higher than average likelihoods at 3% and 3.2%, respectively, while the likelihood for black officers is half the sample average (1.3%). The black-white difference is statistically significant at the 1% level.

The second row in Table 2 lists the number of monthly complaints. The average officer receives about 0.04 complaints in a given month, equating to about 1 complaint every two years. This statistic is similar across race.

The remaining rows in Table 2 depict the number of monthly arrests by arrest type. For example, the average officer makes 1.8 arrests every month. White and Hispanic officers are slightly over this average at 2 and 2.2 arrests, respectively, while black officers are below this average at 1.2 arrests. The black-white difference equates to 10 fewer arrests a year. This is statistically significant at the 1% level.

When comparing summary statistics for the different types of arrests, we see that the black-white difference in total arrests is driven by arrests for non-index crimes, which make

Table 2: Racial and Gender Differences in Work Measures

Sample:	All Officers	White Officers	Black Officers	Hispanic Officers	B-W Difference (p-value)	H-W Difference (p-value)
Nominated	2.5%	3.0%	1.3%	3.2%	-1.7 (0.000)	0.2 (0.016)
Complaints	0.04	0.04	0.04	0.04	0.00 (0.937)	0.00 (0.075)
Total Arrests	1.82	2.04	1.19	2.16	-0.85 (0.000)	0.12 (0.000)
Violent	0.37	0.37	0.31	0.42	-0.06 (0.000)	0.05 (0.000)
Property	0.27	0.29	0.20	0.30	-0.09 (0.000)	0.01 (0.017)
Non-Index	1.19	1.38	0.68	1.44	-0.69 (0.000)	0.07 (0.000)
Drug	0.31	0.37	0.14	0.41	-0.23 (0.000)	0.03 (0.000)
Traffic	0.12	0.15	0.06	0.16	-0.09 (0.000)	0.01 (0.002)
TRR filings	0.05	0.05	0.03	0.06	-0.02 (0.000)	0.00 (0.039)
Observations	250,872	111,876	70,572	59,148		

Source: CPD analysis sample.
Notes: This table lists monthly summary statistics for 6,518 police officers. Sample is at the officer-month level. Non-index arrests include arrests for non-property and non-violent crimes. B-W Difference reports the percentage-point difference between black officers and white officers. H-W Difference reports the percentage-point difference between Hispanic officers and white officers. p-values are the p-value from a t-test of a difference in means. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

up around 65% of all arrests. Here, the difference is about -0.70 arrests per month or 8.4 fewer arrests per year and is statistically significant at the 1% level.

Although the data reveal a disparity in number of arrests, we caution the reader from jumping to the conclusion that black officers are less productive than white or Hispanic officers. Arrests are not a comprehensive measure of policing quality and may be a biased measure. For example, a comparison of white and black officers' arrest records would suggest that black officers are less productive than white officers. However, a study by Harvey and Mattia (2020) finds that police departments that increased their share of black officers subsequently reduced black crime victimization. Similarly, female officers have fewer arrests than male officers but Miller and Segal (2018) finds that increasing the number of female police officers decreased the number of intimate partner homicides and increased the number of reports of domestic violence in the U.S. These outcome measures, which are important measures of social welfare, are not captured by arrests nor would they appear on an officer's record.

Another example is to consider drug and traffic arrests, which are presumably proactive in that they are more likely to have originated from an officer-initiated incident. This classification of "proactive arrests", which allows for greater officer discretion, can also be seen as a delineation between appropriate and inappropriate uses of police authority.[17] In Table 2, we see that white officers are about 2.4-2.6 times more likely than black officers to arrest someone for drugs or traffic violations. In contrast, the black-white difference for more serious crimes, like violent crimes, is economically small at -0.06 arrests a month. Similarly, Ba et al. (2020) examine daily patrol assignments of CPD officers and find that black officers make fewer stops, arrests, and use force less often than their white colleagues. This disparity is driven by a decreased focus on discretionary contact, such as stops for "suspicious behavior". These facts suggest that although it is important to control for work measures in our analysis, we should not automatically interpret differences in overall arrests as differences in policing quality.

4 Empirical Methodology

This section outlines the empirical methodology to examine whether a black-white recognition gap exists in award nominations. We use the assignment to a new supervisor at the start of a calendar year to approximate random assignment of a supervisor's race to an officer.[18] Thus, to test for the existence of racial bias, we can simply compare nomination rates of a white officer vs. a black officer assigned to a given supervisor. This identification strategy is similar to Chetty et al. (2014), which develops a quasi-experimental method by exploiting student exposure to teachers of varying value-added levels that is induced through

---

[17]We borrow this term and classification from Worden et al. (2013). We do not know whether an arrest stemmed from an incident that the officer initiated on his or her own authority, but we assume that drug and traffic arrests are more likely to have stemmed from officer-initiated traffic stops as compared to arrests for violent crimes. Importantly, proactive arrests should be considered as a very noisy measure of quality policing. For example, Worden et al. (2013) analyzed the impact of a police agency's early intervention system, which aims at monitoring and managing police misconduct among officers who exhibit patterns of problematic behavior, and found that it lowered the number of proactive arrests with little impact on productivity.

[18]About 96% of officers are assigned to a supervisor in January of each calendar year.

teacher turnover. In this paper, we use officer exposure to supervisors of different races that is induced through annual re-assignment of supervisors to officers.

Several facts suggest that officers do not game the supervisor assignment system. First, about 89% of all officer-supervisor relationships are new relationships. Second, 78% of all supervisor relationships between 2009 and 2015 lasted exactly one year. Although the vast majority of supervisor relationships last one year, because it is not a totality, we may be concerned that some officer-supervisor relationships may have been arranged outside of the random assignment system. Therefore, we restrict our analysis sample to all supervisor-officer relationships that last one year in order to minimize the number of endogenously formed supervisor relationships. In the next section, we test whether officers are randomly assigned to supervisors in the data.

4.1   Exogeneity of Officer Performance and Supervisor Assignment

Throughout the paper, we want to interpret any change in nomination likelihood when white supervisors are assigned to white officers relative to when they are assigned to black officers as a causal effect of officer race. The key assumption is that minority officers were not systematically assigned to white supervisors in years when officer performance would have been particularly low for other reasons. For example, if high-performing white officers sort to white supervisors while high-performing black officers sort to black supervisors, then we would see a negative black-white nomination gap among white supervisors and a positive black-white gap among black supervisors. This may appear to be in-group favoritism by both white and black supervisors but in reality it would be the result of sorting of high-performing police officers to same-race supervisors.

To test this, we examine whether officer performance is correlated with supervisor race.[19] Because officers are assigned to supervisors at the Department level, we use all patrol officers

---

[19]In the data, officers are more likely to be assigned to a supervisor of the same race. The sorting concern, however, is not simply about white officers matching with white supervisors or black officers matching with black supervisors. It is about *high-performing* officers matching with same-race supervisors. It is selection on officer's characteristics *and* race together that we care about.

assigned to a supervisor rather than the analysis sample that is restricted to officers whose supervisor assignment lasted one year.

We estimate the following regression model separately for white supervisors and black supervisors:

$$Y_{it} = \beta_0 + X'\beta + Z'\alpha_1 + (B_i \times Z')\,\alpha_2 + (H_i \times Z')\,\alpha_3$$
$$+ (A_i \times Z')\,\alpha_4 + (N_i \times Z')\,\alpha_5 + e_{it}$$
(1)

where Y_{it} is equal to 1 if officer is assigned to a white (black) supervisor in month t and 0 otherwise. X is a vector of baseline controls, such as officer sex, birth year, tenure, an indicator for whether the officer and supervisor were in the same district last month, district assignment, year fixed effects, and district-year fixed effects. Z is a vector of lagged work performance measures, such as the number of arrests (e.g., violent crimes, property crimes, and non-index crimes), complaints, TRR filings, and levels of force (e.g., strong use of force and weak use of force). We include second-order and third-order terms of officer tenure and all work performance measures in Z to capture non-linear relationships between these variables and supervisor assignment. B_i is a binary indicator variable if the officer is black, H_i if Hispanic, A_i if Asian, and N_i if Native American. White officers are the reference group.

To test our identification strategy, we are interested in whether there is differential sorting by officer race and work performance to a white supervisor. Because we are conducting multiple hypothesis tests to see which measures are statistically significant, we adjust the p-values using the Holm-Bonferroni procedure.[20]

Table 3 presents estimates for α_2, the black-white difference in officer work measures, and for α_3, the Hispanic-white difference in officer work measures, when the dependent variable is

---

[20]The issue with multiple testing is that there is typically a large probability that some of the true null hypotheses will be rejected. Therefore, some of the statistically significant estimates we see may not truly be statistically significant. The Holm-Bonferroni method controls the probability that one or more Type I errors will occur by adjusting the rejection criteria for the individual hypotheses.

a white supervisor (columns 1 and 2) and when the dependent variable is a black supervisor (columns 3 and 4). Asterisks in Table 3 denote adjusted p-values.

The evidence suggests that white, black, and Hispanic police officers do not differentially sort to white supervisors based on their work performance. Most of the work measures that are statistically significant have trivial estimates that are essentially zero. The black-white difference in property arrests is -0.02, but this is weakly significant at the 10% level and is a lot smaller than the unconditional black-white gap in property arrests, which is -0.09. This suggests that white supervisors are assigned to officers who are, on average, similar in terms of their work performance. Similarly for black supervisors, most of the statistically significant work measures have trivial magnitudes. Black and Hispanic officers assigned to black supervisors have fewer violent-crime arrests than white officers assigned to black supervisors, but again these magnitudes are relatively small: -0.016 for the Hispanic-white gap and -0.027 for the black-white gap off a white mean of 0.37 represent four percent and seven percent, respectively.

In our analysis, we focus on officers assigned to white supervisors, as the majority of supervisors are white, and our findings provide an indirect test of random assignment of officers to white supervisors as we would not expect to see any correlation between officer work measures and supervisor assignment if officers are randomly assigned to supervisors. It also suggests that, in the absence of supervisor bias, we should not see any racial differences in nominations once we control for officers' work performance.

Table 3: Officer Work Measures and Supervisor Race

Dependent Variable:	Supervisor is White		Supervisor is Black	
Coefficients for:	Black-White Diff. (1)	Hispanic-White Diff. (2)	Black-White Diff. (3)	Hispanic-White Diff. (4)
Arrests				
Violent	0.0150	0.0109	-0.0270***	-0.0158***
	(0.00690)	(0.00669)	(0.00558)	(0.00360)
Violent^2	-0.00787	-0.000879	0.00856**	0.00641**
	(0.00351)	(0.00363)	(0.00278)	(0.00192)
Violent^3	0.00110*	-6.23e-05	-0.000869**	-0.000637*
	(0.000367)	(0.000436)	(0.000276)	(0.000216)
Property	-0.0224*	-0.00651	-0.00669	0.00200
	(0.00778)	(0.00437)	(0.00650)	(0.00255)
Property^2	0.00872	0.00218	1.03e-06	-0.000794
	(0.00409)	(0.00127)	(0.00343)	(0.000821)
Property^3	-0.000503	-9.47e-05	-7.91e-05	3.50e-05
	(0.000457)	(5.53e-05)	(0.000384)	(3.57e-05)
Non-Index	0.00449	-0.00962***	-0.00787***	0.00396***
	(0.00206)	(0.00182)	(0.00167)	(0.00105)
Non-Index^2	0.000363	0.000613*	0.000675***	-0.000524***
	(0.000228)	(0.000216)	(0.000181)	(0.000129)
Non-Index^3	-1.76e-05**	-1.35e-05	-9.33e-06	9.88e-06*
	(5.49e-06)	(5.72e-06)	(4.03e-06)	(3.48e-06)
Complaints	-0.107	-0.0406	0.0325	-0.0184
	(0.0815)	(0.0507)	(0.0706)	(0.0273)
Complaints^2	0.115	0.0664	-0.0331	0.0136
	(0.111)	(0.0629)	(0.0965)	(0.0328)
Complaints^3	-0.0248	-0.0204	0.00910	-0.00235
	(0.0324)	(0.0162)	(0.0284)	(0.00784)
TRR Filings	-0.0121	-0.120	0.00327	0.00349
	(0.118)	(0.0988)	(0.0988)	(0.0477)
TRR Filings^2	0.0446	0.0421	0.00853	0.0172
	(0.0866)	(0.0575)	(0.0781)	(0.0291)
TRR Filings^3	-0.00906	-0.00364	-0.00423	-0.00344
	(0.0192)	(0.00987)	(0.0184)	(0.00522)
Strong Force Ratio	0.0187	0.103	0.0656	0.0185
	(0.104)	(0.0675)	(0.0894)	(0.0350)
Strong Force Ratio^2	0.0427	-0.0660	-0.130	-0.0467
	(0.134)	(0.0766)	(0.118)	(0.0384)
Strong Force Ratio^3	-0.0183	0.0160	0.0436	0.0152
	(0.0408)	(0.0208)	(0.0367)	(0.00982)
Weak Force Ratio	0.306	-0.256	-0.197	0.344
	(0.388)	(0.301)	(0.325)	(0.160)
Weak Force Ratio^2	-0.530	0.512	0.286	-0.516
	(0.552)	(0.420)	(0.464)	(0.226)
Weak Force Ratio^3	0.183	-0.184	-0.0945	0.169
	(0.182)	(0.138)	(0.153)	(0.0746)
Observations	396,420		396,420	

Source: CPD data.
Notes: This table reports the black-white difference and the Hispanic-white difference from a regression of officer race on having a white supervisor (columns 1 and 2) and having a black supervisor (columns 3 and 4). Additional controls include officer sex, birth year, tenure, unit, whether the officer and supervisor were in the same unit last month, year FE, and unit x year FE. Robust standard errors are in parentheses. Asterisks denote Holm-adjusted p-values. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5 Results

5.1 Main Results

In this section, we examine whether an officer's arrest record affects the supervisor's likelihood of nomination and whether there are any differential effects for minority officers. That is, conditional on the officer's arrest record, are there racial differences in the probability of nomination? The regression sample for this analysis is at the officer-month level. We estimate the following model, separately for white supervisors and black supervisors:

$$
Nom_{it} = \beta_0 + \left(\sum_{c=1}^{5} \mathbb{1}\{Arrests_{i,t-1} = c\} \times \beta_1^c \right) + \left(B_i \times \sum_{c=1}^{5} \mathbb{1}\{Arrests_{i,t-1} = c\} \right) \beta_2^c
$$

$$
+ \left(H_i \times \sum_{c=1}^{5} \mathbb{1}\{Arrests_{i,t-1} = c\} \right) \beta_3^c + \left(A_i \times \sum_{c=1}^{5} \mathbb{1}\{Arrests_{i,t-1} = c\} \right) \beta_4^c \quad (2)
$$

$$
+ \left(N_i \times \sum_{c=1}^{5} \mathbb{1}\{Arrests_{i,t-1} = c\} \right) \beta_5^c + X'\alpha + \tau_t + \varepsilon_{it}
$$

where i denotes officer and t denotes month. Nom_{it} is equal to 1 if officer i was nominated for an award in month t and 0 if not. $Arrests_{i,t-1}$ is the number of arrests officer i made last month. Rather than including this as a continuous measure, we break up the number of arrests into six categories: 0 arrests, 1, 2, 3, 4, and 5 or more arrests. We do this because the average number of monthly arrests is 2, and so we want to estimate impacts for smaller arrest categories. The reference category is 0 arrests.

B_i is a binary indicator variable if the officer is black, H_i if Hispanic, A_i if Asian, and N_i if Native American. White officers are the reference group.

X is a vector of officer, supervisor, and district characteristics. Officer controls include officer's birth year, district assignment, tenure, and the number of complaints made against the officer. Supervisor controls include supervisor fixed effects and the share of black supervisees. District characteristics include overall crime rate and violent crime rate. All time-varying variables except for district assignment and the share of black supervisees are lagged by one

month. We also include fixed effects for year and month in τ_t. We estimate robust standard errors to account for heterogeneity introduced by the binary dependent variable.

The parameters of interest are β_2^c, which tells us how the black-white difference changes by the number of arrests, and β_3^c, which tells us how the Hispanic-white difference changes by the number of arrests. We expect β_1 to be positive and increasing in the number of arrests. This is based on our belief that departmental awards are based on officer work performance. If supervisors do not exhibit racial bias, then we expect β_2^c to be zero. A negative β_2^c indicates that supervisors favor white officers (or, are biased against black officers), whereas a positive β_2^c indicates that supervisors favor black officers.

Table 4 reports estimates for β_2^c and β_3^c separately for white supervisors in Panel A and for black supervisors in Panel B. We report estimates for white officers in column 1, the black-white difference in column 2, and the Hispanic-white difference in column 3. There are increasing returns to having more arrests, with a marked increase for those with five or more arrests (column 1). Although we do not assert that arrests are an accurate measure of policing quality, we do the analysis this way because police departments seem to value and reward arrest quantity. It is interesting, therefore, that the return to having more arrests is less for black officers compared to white officers (Panel A, column 2). The black-white difference in nomination probability for officers with one arrest widens by 0.5 percentage-points compared to the black-white difference among officers with no arrests last month. This estimate is significant at the 5% level.

The black-white nomination gap widens even more as the number of arrests increases. Among officers with five or more arrests, the relative black-white difference widens by 2.4 percentage-points and is significant at the 1% level. It is informative to interpret this disparity in the context of racial differences in work performance. For example, black officers with 5 or more monthly arrests are at the 94th percentile of their distribution, while white officers are at the 81st percentile of their distribution. Yet, white supervisors are even *less* likely to nominate black officers over white officers compared to if both had zero arrests.

18

Table 4: Impact of Arrest Record on Nomination Likelihood by Officer Race

	Outcome Variable: Nominated		
Estimates for:	White Officer	Black-White Gap	Hispanic-White Gap
	(1)	(2)	(3)
Panel A: White Supervisors			
1 arrest	0.00913***	-0.00496**	-0.000886
	(0.00135)	(0.00193)	(0.00239)
2 arrests	0.0147***	-0.00758***	0.000313
	(0.00190)	(0.00278)	(0.00329)
3 arrests	0.0216***	-0.0105***	0.000170
	(0.00260)	(0.00408)	(0.00447)
4 arrests	0.0250***	-0.00248	-0.000702
	(0.00326)	(0.00586)	(0.00566)
5 or more arrests	0.0566***	-0.0235***	-0.0156***
	(0.00256)	(0.00481)	(0.00428)
Observations		171,094	
Panel B: Black Supervisors			
1 arrest	-0.000463	0.00354	0.00737
	(0.00309)	(0.00354)	(0.00639)
2 arrests	0.00857*	-0.00640	0.000598
	(0.00509)	(0.00570)	(0.00997)
3 arrests	0.0101	-0.000213	0.00949
	(0.00659)	(0.00800)	(0.0137)
4 arrests	0.00851	-0.0151	0.00817
	(0.00889)	(0.0102)	(0.0171)
5 or more arrests	0.0451***	-0.00302	-0.0222*
	(0.00762)	(0.0113)	(0.0125)
Observations		29,413	
Mean Pr(Nom) for White Officers		0.031	

Source: CPD analysis sample.
Notes: This table reports estimates for the impact of an officer's arrest record on the probability of nomination by white supervisors (Panel A) and by black supervisors (Panel B). Each panel is a single OLS regression with estimates for white officers in column 1, the black-white difference in column 2, and the Hispanic-white difference in column 3. All estimates control for month and year FE, supervisor FE, officer birth year, tenure, lagged complaints, unit FE, share of black supervisees, lagged crime rate, lagged violent crime rate. The mean nomination probability of white officers by white supervisors is 3.1%. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We also examine whether white supervisors are less likely to nominate Hispanic officers, another racial minority in the Chicago Police Department (Panel A, column 3). The Hispanic-white difference is pretty trivial and not statistically significant until the five or more arrests category. Among officers with at least five arrests, the Hispanic-white gap in

nomination probability widens by 1.6 percentage-points and is significant at the 1% level. As Hispanics are more likely to be nominated than white officers in the zero arrest category, this means the Hispanic-white gap becomes negative.

When comparing between the two racial minority groups, the black-white difference is statistically significantly different from the Hispanic-white difference in the 1-3 arrests categories and not for those with 4 or more arrests. This suggests that white supervisors penalize black officers more than white or Hispanic officers among those with average arrest records (recall the average number of arrests is two), but favor white officers when comparing officers with more arrests.

Panel B reports estimates for officers assigned to black supervisors. None of the point estimates are significant, though this may be due to the fact that there are few black supervisors (190 compared to 893 white supervisors). However, the magnitudes of the point estimates are also very small. The one exception is that the Hispanic-white gap among officers with five or more arrests is weakly significant at -0.0222.

In Appendix Table A4, we estimate a version with officer fixed effects instead of supervisor fixed effects. The results are very similar. We also examine how the black-white nominations gap changes by supervisor characteristics (Appendix Figure A1). The gap is larger among supervisors with a higher share of black supervisees and among less experienced supervisors. It does not change much by the number of supervisees.

5.2   Information Acquisition

In this section, we examine a potential mechanism for the black-white nomination gap. The 2016 report by the Police Accountability Task Force found little stability in supervisor-officer relationships. First, officers may work with different Sergeants over the course of their shift, any of whom may or may not be their officially assigned supervisor. Second, personnel information does not necessarily get transferred to supervisors when officers switch assignments. Therefore, one potential explanation for why white supervisors may be less

likely to nominate black officers is because they are less likely to interact with them and, therefore, are less likely to be informed of their accomplishments (Glover et al., 2017). To test this theory, we exploit an institutional feature that randomizes the quarter in which officers are evaluated by their supervisor. Although there appears to be little interaction between officers and supervisors on a daily basis, we assume that the annual evaluation requires supervisors to acquire information about the officer's work record.

All supervisors are required to conduct annual evaluations of their assigned officers, and this evaluation must take place during the quarter prior to the quarter in which the officer joined the Department. Appendix Table A3 lists the evaluation quarters and evaluation due dates by start month. For example, if an officer started his career in July (Q3), then his annual evaluation must take place in the second quarter of every calendar year. Because start dates are randomly determined by a lottery number, this means that the evaluation quarter is essentially randomly assigned across officers.[21]

We exploit this institutional feature and compare nomination likelihoods of black vs. white officers assigned to white supervisors across quarters. Because the evaluation quarter is randomly assigned, this simple comparison allows us to isolate the effect of acquiring information. If a lack of information acquisition is the reason for a black-white nomination gap, then we would expect this to disappear in the quarter when supervisors are required to evaluate their assigned officers. For this analysis, the sample is at the officer-month level, and the regression model is:

$$
\begin{aligned}
Nom_{it} = \beta_0 + \sum_{q=-2}^{3} \mathbb{1}\{EQ = q\}\delta^q + \left(B_i \times \sum_{q=-2}^{3} \mathbb{1}\{EQ = q\} \right) \beta_1^q \\
+ \left(H_i \times \sum_{q=-2}^{3} \mathbb{1}\{EQ = q\} \right) \beta_2^q + \left(A_i \times \sum_{q=-2}^{3} \mathbb{1}\{EQ = q\} \right) \beta_3^q \\
+ \left(N_i \times \sum_{q=-2}^{3} \mathbb{1}\{EQ = q\} \right) \beta_4^q + X'\alpha + \tau_t + e_{it}
\end{aligned}
\tag{3}
$$

---

[21]After passing a written exam, all CPD candidates are placed on a eligibility list according to a randomly assigned lottery number and called off in lottery order to enroll in the police academy.

where i denotes officer and t denotes month. Nom_{it} is equal to 1 if officer i was nominated for an award in month t and 0 if not. B_i is a binary indicator variable if the officer is black, H_i if Hispanic, A_i if Asian, and N_i if Native American. White officers are the reference group.

The second term is a set of binary indicator variables for each quarter relative to the evaluation quarter, which is denoted as $EQ = 0$. The reference quarter is $EQ = -3$, or three quarters prior to the evaluation quarter. The coefficients δ^q tell us how nomination likelihoods for white officers change across quarters. If information acquisition is an important mechanism, then we expect it to be enhanced in the quarter that supervisors evaluate their officers (δ^0).

The third term in parentheses interacts the black indicator variable and the relative-quarter indicator variables. The coefficients β_1^q depict how the black-white nomination gap evolves relative to $EQ = -3$. If white supervisors are equally likely to nominate their black and white officers, then we expect β_1^q to be zero. Likewise, the coefficients β_2^q tell us how the Hispanic-white nomination gap evolves over time.

X is a vector of controls for officer characteristics (e.g., officer fixed effects, tenure, district, complaints, arrests) and district characteristics (e.g., overall crime rate, violent crime rate, overall arrest rate). All time-varying variables except for district assignment and tenure are lagged by one month. τ_t includes month and year fixed effects. We estimate robust standard errors.

Figure 1 plots the estimates for δ^q separately for white officers, black officers, and Hispanic officers.[22] The hump-backed shape suggests that the nomination probability increases as we grow closer to the evaluation quarter then falls afterwards. This pattern exists for both white and black officers assigned to white supervisors, but the hump is less steep for black officers. Estimates and standard errors for the black-white difference and Hispanic-white difference are reported in Table 5.

---

[22]Specifically, we plot estimates for δ^q when the reference group is white officers. Then, we re-estimate equation (3) with black officers as the reference group and plot δ^q on the same graph.

Figure 1: Probability of Nomination by Quarter
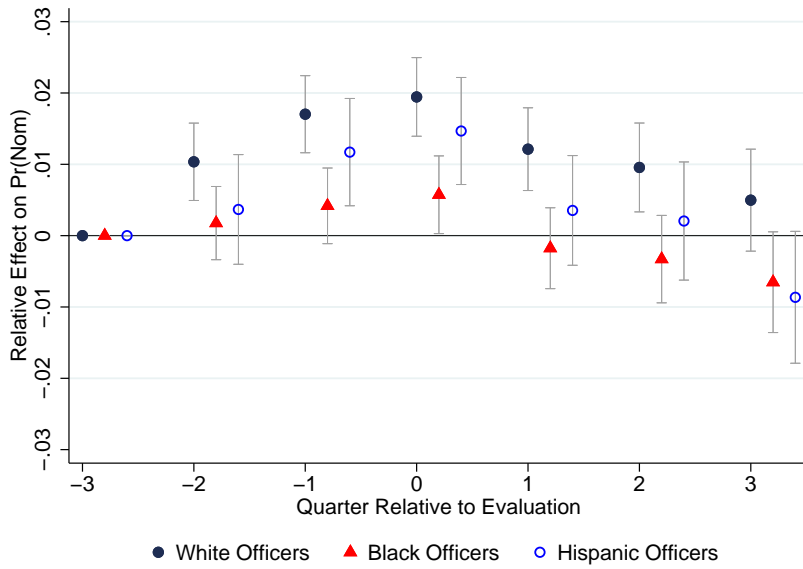
Source: CPD analysis sample.
Notes: This graph shows how the probability of nomination changes by quarter relative to the quarter of the officer's evaluation, separately for white, black, and Hispanic officers. Sample is restricted to all officers assigned to a white supervisor. Estimates control for supervisor fixed effects, officer district, tenure, arrests, complaints, overall crime rate, violent crime rate, and overall arrest rate. Wings depict 95% confidence intervals using robust standard errors. N = 171,094.

Three quarters before their evaluation ($EQ = -3$), white officers have a 1.89% chance of being nominated by their white supervisors. The following quarter, this likelihood increases by about 1 percentage-point (53%), then 1.7 percentage-points (90%) the quarter before their evaluation, then 1.94 percentage-points (103%) in the evaluation quarter. All of these estimates are significant at the 1% level.

After their evaluation, white officers are still more likely to be nominated relative to $EQ = -3$, but not as likely as they were in the quarter of their evaluation. Specifically, white officers are 1.2 percentage-points more likely to be nominated in the quarter following their evaluation—about a 38% decrease from the previous quarter—and 0.96 percentage-points more likely to be nominated in $EQ = 2$ relative to $EQ = -3$. These estimates are significant at the 1% level and 10% level. Although white officers are less likely to be

nominated after their evaluation quarter, their chances of nomination are still higher than the initial likelihood in $EQ = -3$. There also appears to be a sustained effect from the information acquisition that persists for two quarters after the officer's evaluation.

The story for black officers is a different one. Although the point estimates for black officers are positive leading up to their evaluation, they are not statistically significant until the quarter of their evaluation. In the evaluation quarter ($EQ = 0$), black officers are 0.57 percentage-points (61%) more likely to be nominated relative to the reference quarter, and this is significant at the 5% level. Immediately after their evaluation, the point estimates turn negative and are not statistically significant. Three quarters after their evaluation, the point estimate is negative (-0.0065) and significant at the 10% level. As the average black officer has a 0.94 chance of being nominated in $EQ = -3$, this means that their chances of nomination have been reduced by 69% in $EQ = 3$. Contrary to their white colleagues, black officers do not benefit from information acquisition after their evaluation.

Hispanic officers experience a statistically significant relative increase in nomination likelihood in the quarters before and of their evaluation: 1.2 percentage-points (67%) and 1.5 percentage-points (83%), respectively.[23] After their evaluation, the point estimates are positive but not statistically significant. Like their black colleagues, the nomination likelihood three quarters after the evaluation for Hispanic officers is negative (-0.0086) and significant at the 10% level.

To summarize, white and Hispanic officers are more likely to be nominated in the quarters leading up to and including their evaluation. This boost grows steadily as we move closer to the evaluation quarter, and then drops after the evaluation quarter. In contrast, black officers assigned to white supervisors see a jump in their nomination likelihood only in the quarter of evaluation. Further, the negative black-white gap in all of the quarters are statistically significant at either the 5% or 1% levels (Table 5).

---

[23]The mean nomination probability for Hispanic officers assigned to white supervisors in $EQ = -3$ is 1.8 percent.

Table 5: Racial Difference in Nomination Likelihood by Quarter

Estimates for:	Outcome Variable: Nominated		
	White Officers	Black-White Gap	Hispanic-White Gap
	(1)	(2)	(2)
Quarter relative to three quarters before evaluation			
Two quarters pre-eval. (EQ=-2)	0.0104***	-0.00859**	-0.00668
	(0.00277)	(0.00372)	(0.00473)
One quarter pre-eval. (EQ=-1)	0.0170***	-0.0128***	-0.00532
	(0.00276)	(0.00360)	(0.00454)
Evaluation quarter (EQ=0)	0.0194***	-0.0137***	-0.00478
	(0.00281)	(0.00356)	(0.00445)
One quarter post-eval. (EQ=1)	0.0121***	-0.0139***	-0.00859*
	(0.00296)	(0.00354)	(0.00444)
Two quarters post-eval. (EQ=2)	0.00956***	-0.0129***	-0.00751
	(0.00318)	(0.00378)	(0.00474)
Three quarters post-eval. (EQ=3)	0.00498	-0.0115***	-0.0136**
	(0.00365)	(0.00440)	(0.00536)
Observations		171,094	

Source: CPD analysis sample.
Notes: This table reports estimates of how the nomination likelihood changes by quarter. Estimates for white officers are in column 1, the black-white difference in column2, and the Hispanic-white difference in column 3. See notes in Figure 1.

6   Experimental Evidence

Because our observational data do not capture interactions between supervisor and officer, we run an online experiment to measure the review process in the nomination decision.[24] Participants were asked to review CPD officer profiles and nominate one for an award. By using the same officers from the CPD analysis sample, we are able to generalize our findings to a broader evaluator group than Chicago police supervisors.

6.1   Experimental Design

Survey participants were given two different types of tasks. In the first type of task, participants chose between a black officer and a non-black officer, where the black officer

---

[24]The experiment was pre-registered in the AEA RCT Registry, AEARCTR-0005929.

was randomly assigned to be either "high-quality" or "low-quality" and the non-black officer was assigned the converse. In judging officer profiles, we used the number of civilian complaints and arrests. These classifications are admittedly subjective but they were made independently of officer race and sex. "High-quality" profiles were those with zero civilian complaints and an above-average number of arrests. "Low-quality" profiles were those with one or two civilian complaints and a below-average number of arrests.

In the second type of task, participants were shown four officer profiles and asked to nominate one for an award. In this task, officer profiles displayed only demographic information (e.g., race, sex, and age) and participants had to mouse over a profile to reveal full information about the officer. All officers were of "average quality", defined as having zero or one civilian complaints and an average number of arrests. There were two iterations of this task. In the first iteration, the officer pool was racially heterogeneous. Three of the four profiles always featured a white officer, a black officer, and an Hispanic officer. The race of the fourth profile was randomly chosen amongst these three races. In the second iteration, three of the four profiles were always white officers and, again, the race of the fourth profile was randomly chosen amongst white, black, and Hispanic. The display ordering of officer profiles was randomly determined. All tasks were time-constrained to limit the amount of time the participant had to review profiles. See Appendix B for more information about the online experiment.

6.2 Sample Selection and Data

The experiment was conducted on Amazon Mechanical Turk in July 2020.[25] We recruited 411 MTurk workers (hereafter "workers") who were 18 years of age or older, based in the United States with English language proficiency, and who had access to a computer with a

---

[25]It is possible that the George Floyd incident on May 25, 2020 and subsequent protests may have altered people's perceptions of the police and blacks. Specifically, the incident may have increased MTurk workers' interest in and affinity towards black officers because they are black. This would work against our results, which find that black officers are less likely to be moused over and are less likely to be nominated when paired against a non-black officer.

mouse and Javascript. The technical requirements were necessary in order to capture mouse movements on the screen. The survey had three data quality checks to identify bots and to ensure workers paid attention during the survey. For the analysis, we decided to include workers who passed at least two of the three data quality checks. This restriction reduces our final analysis sample to 407 workers.

We study how officer race affects two types of choices: attention to an officer profile and the nomination decision. First, we measured which profiles workers hovered over, the order in which workers hovered over the four profiles, and how long workers hovered over each profile. Second, we measured which officer the worker ultimately nominated for an award.

6.3 Results

Are black officers less likely to be nominated for an award?

Column 1 of Table 6 reports results from the pairwise comparison of a black and non-black (white or Hispanic) officer. Low-quality black officers are 9.3 percentage-points ($p < 0.01$) less likely than low-quality white officers to be nominated. This gap largely persists with high-quality officers. Although high-quality officers are almost 50 percentage-points more likely to be nominated for an award ($p < 0.01$) than low-quality officers, high-quality black officers are still 6 percentage-points less likely to be nominated than high-quality white officers.

In column 2, we conduct a robustness check wherein the two officer profiles are of the same race and sex and differ only in terms of quality. As expected, high-quality profiles are more likely to be nominated–about 52 percentages-points–and this is significant at the 5% level. This also provides an indirect test that MTurk workers were able to discern the quality difference between the two officer profiles. Interestingly, when the officers are both black or both female, workers are even more likely to nominate the high-quality officer (12.8 additional percentage-points for black officers and 14.7 additional percentage-points for female officers) relative to when the officers are white males. Although these estimates are not statistically

27

significant, the positive point estimates, together with the results from column 1, suggest that either white males are given some slack even if they do not meet a certain standard or that minorities are held to a higher standard.

Table 6: Impact of being Black on Nomination Likelihood

| Pairwise Comparison: | Outcome Variable: Nominated | |
| | Black vs Non-Black | High-Quality vs Low-Quality |
	(1)	(2)
High-Quality Profile	0.478***	0.517**
	(0.0919)	(0.169)
Black Officer	-0.0929***	-0.0268
	(0.0309)	(0.119)
High-Quality x Black Officer	0.0333	0.128
	(0.0446)	(0.172)
Female Officer		-0.0724
		(0.0682)
High-Quality x Female Officer		0.147
		(0.107)
Observations	1,592	802

Source: MTurk survey data.
Notes: All estimates control for officer traits and profile location on screen. Officer traits include officer age, tenure, arrests, and complaints. Robust standard errors are in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Do MTurk workers choose different levels of attention based on officer's race?

Table 7 presents summary statistics on information acquisition by MTurk workers for white officers, black officers, and Hispanic officers. Panel A combines results from the racially heterogeneous officer pool and the white-majority officer pool. Workers tend to mouse over most of the officer profiles: about 81.5 to 84% of profiles were moused over. However, black officers are about 2.8 percentage-points less likely to be moused over, and this difference has a p-value of 0.107. Hispanic officers are also less likely to be moused over, but the difference is not statistically significant.

Conditional on being moused over, there does not appear to be a significant difference in terms of race regarding which officer is moused over first. However, there is a difference in the amount of time spent reviewing the profiles. Workers spend around half a second more

reviewing black and Hispanic profiles, and these are significant at the 1% level.

Table 7: Information Acquisition by Officer Race, Comparison of Means

	White Officer	Black Officer	Hispanic Officer	W-B Difference (p-value)	W-H Difference (p-value)
Panel A: All					
Percent Moused Over	84.2%	81.5%	81.8%	0.028 (0.107)	0.024 (0.163)
Percent First Mouseover	30.0%	31.9%	28.4%	-0.019 (0.419)	0.016 (0.490)
Mean Mouseover Duration (seconds)	2.33	2.78	2.89	-0.448 (0.000)	-0.559 (0.000)
Panel B: Heterogenous-Race Officer Pool					
Percent Moused Over	85.2%	82.7%	80.6%	0.025 (0.276)	0.046 (0.054)
Percent First Mouseover	32.8%	31.0%	26.5%	0.018 (0.581)	0.063 (0.048)
Mean Mouseover Duration (seconds)	2.27	2.64	2.85	-0.366 (0.006)	-0.576 (0.000)
Panel C: White-Majority Officer Pool					
Percent Moused Over	83.8%	76.6%	86.7%	0.073 (0.036)	-0.028 (0.420)
Percent First Mouseover	28.9%	35.7%	35.6%	-0.069 (0.155)	-0.069 (0.155)
Mean Mouseover Duration (seconds)	2.36	3.40	3.07	-1.041 (0.000)	-0.714 (0.001)

Source: MTurk survey data.
Notes: B-W Difference reports the percentage-point difference between black officers and white officers. H-W Difference reports the percentage-point difference between Hispanic officers and white officers. p-values are the p-value from a t-test of a difference in means.

When we differentiate between a more heterogeneous racial pool (Panel B) and a more homogenous racial pool (Panel C), we see that the patterns are more salient in the more homogenous racial pool–that is, when black or Hispanic officers are the only minority officer. When workers are choosing amongst three white officers and one black officer, the black officer is 7.3 percentage-points less likely to be moused over (significant at the 5% level). Conditional on being moused over, black officers are about 7 percentage-points more likely than white officers to be the first mouseover, though this is not statistically significant, and workers spend one second longer on black profiles than white profiles (significant at the 1% level). When the sole minority officer is Hispanic, workers are more likely to mouse over the Hispanic officer relative to a white officer and mouse over the Hispanic officer first, though neither of these estimates are statistically significant. However, like with black profiles, workers spend about 7/10s of a second longer reviewing Hispanic profiles. This is significant at the 1% level. Table 8 reports the findings in a regression framework, where we control for

the profile location on the computer screen and the worker's starting mouse position.

Table 8: Impact of Officer Race on Information Acquisition

Officer Pool:	All (1)	Het. Race (2)	White Maj. (3)
Panel A: Outcome Variable: Ever Moused Over			
Black Officer	-0.0289	-0.0271	-0.0707*
	(0.0177)	(0.0230)	(0.0386)
Hispanic Officer	-0.0252	-0.0475**	0.0332
	(0.0179)	(0.0237)	(0.0329)
Observations	2,992	1,500	1,492
Mean Nomination Prob. for White Officer	0.842	0.852	0.838
Panel B: Outcome Variable: First Mouseover			
Black Officer	0.00355	-0.0285	0.0399
	(0.0210)	(0.0293)	(0.0409)
Hispanic Officer	-0.0136	-0.0578**	0.0681
	(0.0212)	(0.0293)	(0.0424)
Observations	2,488	1,243	1,245
Mean Nomination Prob. for White Officer	0.300	0.328	0.289
Panel C: Outcome Variable: Mouseover Duration (seconds)			
Black Officer	0.431***	0.347**	0.984***
	(0.107)	(0.135)	(0.230)
Hispanic Officer	0.570***	0.575***	0.697***
	(0.107)	(0.137)	(0.224)
Observations	2,488	1,243	1,245
Mean Nomination Prob. for White Officer	2.335	2.270	2.361

Source: MTurk survey data.
Notes: All estimates control for profile location on screen and worker's starting mouse position. Standard errors are in parentheses. Panels A and B report robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Next, we examine how information acquisition affects the probability of nomination. Table 9 reports the probability of nomination conditional on officer race and the type of information acquisition: first mouseover (Panel A) or mouseover duration (Panel B).

Being moused over first increases the probability of nomination by 2.5 percentage-points though this is not statistically significant. Relative to white officers, however, black officers who are moused over first are 9.2 percentage-points even more likely to be nominated. This estimate is significant at the 10% level. Column 3 indicates that this finding is driven by a

more racially homogenous officer pool. Since a minority officer will stand out in a mostly-white officer pool, if a worker chooses to mouse over the minority officer first, then the worker is also decidedly more likely to nominate the minority officer (19 percentage-points for black officers and 17 percentage-points for Hispanic officers). As the mean nomination rate for white officers is 21.6%, these estimates mean that being moused over first increases the probability of nomination by 89% for black officers and 79.6% for Hispanic officers in comparison to white officers.

Panel B looks at the impact of mouseover duration on nomination likelihood. The longer a worker spends viewing an officer's profile, the higher the chance of a nomination: an increase of 4.5 percentage-points for each additional second. This is significant at the 1% level. This estimate does not differ for black officers; each additional second on a white officer's profile increases the probability of nomination the same as an additional second spent on a black officer's profile. However, Hispanic officers do benefit; each additional second increases the nomination likelihood by an additional 2.3 percentage-points relative to white officers. This estimate is significant at the 5% level.

To summarize the results, we find that black officers are less likely to be nominated compared to white officers regardless of profile quality. In terms of the information acquisition process, we find that black officers are less likely to be moused over, and this is more salient when workers are choosing among three white officers and one black officer. We do not see similar patterns when workers are choosing among three white officers and one Hispanic officer. Taken together, these findings suggest that the negative black-white nomination gap is due to a lack of interest about black officers on the part of workers when choosing between black and white officers.

We also find that, conditional on being moused over, black officers are as equally likely as white officers to be the first mouseover, but workers that choose to mouse over a black profile first are also more likely to nominate black officers over white officers. This suggests that black officers will benefit from having supervisors who are interested in interacting with

31

them. At the same time, conditional on being moused over, black officers are scrutinized for longer but there is no racial difference in the impact of mouseover duration on nomination likelihood. Since workers spend more time on black officers' profiles, on average, this suggests that black officers are under more scrutiny but do not benefit from the additional information acquisition.

Table 9: Impact of Information Acquisition on Nomination Likelihood

| | Outcome Variable: Nominated Officer | | |
| Officer Pool: | All | Heterogeneous Race | White Majority |
	(1)	(2)	(3)
Panel A: First Mouseover			
First Mouseover	0.0251	0.109**	-0.00466
	(0.0259)	(0.0483)	(0.0313)
Black Officer	-0.00694	-0.205	-0.138*
	(0.0268)	(0.302)	(0.0834)
Hispanic Officer	-0.0330	0.0977	0.110*
	(0.0326)	(0.163)	(0.0600)
First Mouseover x Black Officer	0.0922*	-0.0102	0.194*
	(0.0493)	(0.0669)	(0.106)
First Mouseover x Hispanic Officer	0.0242	-0.113	0.172*
	(0.0528)	(0.0701)	(0.102)
Panel B: Mouseover Duration			
Mouseover Duration (seconds)	0.0451***	0.0569***	0.0403***
	(0.00734)	(0.0133)	(0.00887)
Black Officer	-0.000822	-0.247	-0.149
	(0.0333)	(0.281)	(0.0916)
Hispanic Officer	-0.0961**	0.0423	0.129*
	(0.0398)	(0.154)	(0.0753)
Mouseover Duration x Black Officer	0.00608	-0.0141	0.0224
	(0.0109)	(0.0172)	(0.0175)
Mouseover Duration x Hispanic Officer	0.0226**	0.0156	0.00782
	(0.0109)	(0.0163)	(0.0178)
Observations	2,488	1,243	1,245
White Officer Nomination Mean	0.225	0.245	0.216

Source: MTurk survey data.
Notes: All estimates control for officer traits, profile location on screen, and worker's starting mouse position. Officer traits include officer age, tenure, arrests, and complaints. Standard errors are in parentheses. Panel A reports robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

7 Discussion and Conclusion

Racial bias has been extensively documented in a variety of settings, including hiring decisions (Agan and Starr, 2017; Bertrand and Mullainathan, 2004; Craigie, 2020; Doleac and Hansen, 2018), sports umpires (Parsons et al., 2011), judicial and sentencing decisions (Flanagan, 2018; Mueller-Smith and Schnepel, 2017; Park, 2017; Rehavi and Starr, 2014)[26], and bail decisions (Arnold et al., 2018). The increasing availability of police administrative data has allowed researchers to carefully examine and detect bias in law enforcement as well.

A potential solution that has been put forth is to increase racial and gender diversity among officers, which are traditionally homogenous.[27] A diverse police force may improve policing quality in various ways (Sklansky, 2005). Outwardly, it may improve the police's relationship with the community through unique skills that minority officers may possess (Anwar et al., 2012; Harvey and Mattia, 2020; Miller and Segal, 2018).[28] Inwardly, it may alter the internal dynamics of the department.

This paper asks how racial bias affects career progression, which is of particular relevance to law enforcement, where minorities are less represented at higher ranks. For example, white males comprised 40 percent of all entry-level police officers in the Chicago Police Department in 2015, and 56 percent of those at the rank of Sergeant or higher.[29] In addition to improving policing quality[30], minority representation—particularly at higher ranks of office—may help

---

[26]Mueller-Smith and Schnepel (2017) finds that the practice of diversion, or a halt or termination of one's progression through the justice system, reduces re-offending rates and improves labor market outcomes among young black men charged with misdemeanors.

[27]For example, in their investigative report of the Ferguson Police Department, the U.S. Department of Justice called for a more diverse police force as part of a broader reform effort (United States Department of Justice, 2015, p. 58). Several cities, including Chicago, Indianapolis, and Knoxville, have followed this lead and pursued diversity initiatives (Chicago: https://www.chicagotribune.com/news/breaking/ct-met-chicago-police-hiring-20180503-story.html; Indianapolis: https://www.indystar.com/story/opinion/columnists/suzette-hackney/2018/09/27/impd-leads-charge-toward-diversity-columnist-suzette-hackney-writes/1433649002/; Knoxville: https://www.knoxnews.com/story/news/local/2017/01/09/knoxville-police-department-recruits-remain-diverse-group/96345092/)

[28]McCrary (2007) and Garner et al. (2020) do not find that court-ordered affirmative action litigation affects offense and arrest rates, but Garner et al. (2020) acknowledges that there may be racially heterogeneous effects that offset each other.

[29]These numbers do not include civilian Department members.

[30]See, for example, Bulman (2019); Donohue III and Levitt (2001); Miller and Segal (2018).

to recruit more minorities and close promotion gaps, which may further attract minority applicants. Prior research has found that minorities in management positions can address wage gaps and occupational gaps (Kofoed and mcGovney, 2019; Langan, 2018). At the same time, an extensive literature documents racial and gender bias in the workplace, which may hinder minorities' career progression.[31] In the context of policing, diversity initiatives may be constrained by the extent to which officer bias carries over to their colleagues. Indeed, in the Chicago Police Department, 98 percent of CPD officers believe that promotions are due to connections not merit (Police Accountability Task Force, 2016).

To examine the extent of racial bias in law enforcement, we construct a panel dataset of all CPD officers containing their personnel information. We exploit quasi-random variation in supervisor assignment and find that white supervisors are less likely to nominate black officers than white or Hispanic officers. These results control for officer characteristics, including experience, arrest record, and misconduct. We find that supervisors are more likely to nominate officers in the evaluation quarter, suggesting that interactions and information acquisition is important for career recognition, but also that black officers benefit less than their white peers.

To supplement our CPD analysis, we run an online experiment using Amazon Mechanical Turk workers and, again, find that black officers are less likely to be nominated than their non-black peers. In terms of the information acquisition process, we find that black officers are less likely to be moused over and, conditional on being moused over, are reviewed for longer.

Our findings have two important policy implications for law enforcement. First, we find that interactions between supervisors and officers are an important mechanism for career recognition, suggesting that the observed racial gap in award nominations may be due to

---

[31]For example, Egan et al. (2018) find that female financial advisors are 20% more likely than male financial advisors to lose their jobs following a misconduct. In medicine, Sarsons (2019) finds that physicians are less likely to refer to new female surgeons after a bad patient outcome but not to new male surgeons. Beaman et al. (2018) find that women are harmed in a referral-based hiring process as workplace networks tend to be gender homophilous. Glover et al. (2017) find that manager bias can cause a self-fulfilling prophecy in that biased managers interact less with minority cashiers, leading them to exert less effort.

statistical discrimination. One way to overcome discrimination that is driven by biased beliefs is continuous and sustained positive evaluations (Bohren et al., 2019). But, as we find a persistent negative black-white gap even in the evaluation quarter, our results suggest that the decentralized nature of supervision and oft-changing supervisor assignment in the CPD present a challenge for discrimination to be reversed.

Second, our finding of a persistent black-white recognition gap suggests that simply increasing the diversity of incoming recruits may not be enough to eliminate racial bias in policing. The argument for a diverse police force stems from the "contact hypothesis", or that outsider bias can be reduced if the integrated group has a common goal. Although there is empirical evidence in support of this theory (Lowe, 2019), another study finds that the improved behavior towards out-group members does not extend beyond the intervention setting (Mousa, 2020). As such, police departments should also focus on policies that address in-group bias due to its effect on career advancement.

References

Agan, A. and S. Starr (2017). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Ajilore, O. and S. Shirey (2017). Do #AllLivesMatter? An evaluation of race and excessive use of force by police. *Atlantic Economic Journal 45*(2), 201–212.

Alsan, M., O. Garrick, and G. Graziani (2019). Does diversity matter for health? experimental evidence from oakland. *American Economic Review 109*(12), 4071–4111.

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The quarterly journal of economics 116*(1), 313–350.

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

Anwar, S., P. Bayer, and R. Hjalmarsson (2012). The impact of jury race in criminal trials. *The Quarterly Journal of Economics 127*(2), 1017–1055.

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

Arnold, D., W. Dobbie, and C. S. Yang (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics 133*(4), 1885–1932.

Ba, B., D. Knox, J. Mummolo, and R. Rivera (2020). Diversity in Policing: The Role of Officer Race and Gender in Police-Civilian Interactions in Chicago. Working paper.

Bacher-Hicks, A. and E. de la Campa (2020). Social Costs of Proactive Policing: The Impact of NYC's Stop and Frisk Program on Educational Attainment. Working paper.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review 106*(6), 1437–75.

Beaman, L., N. Keleher, and J. Magruder (2018). Do job networks disadvantage women? Evidence from a recruitment experiment in Malawi. *Journal of Labor Economics 36*(1), 121–157.

Bertrand, M. and S. Mullainathan (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American Economic Review 94*(4), 991–1013.

Bohren, J. A., A. Imas, and M. Rosenberg (2019). The dynamics of discrimination: Theory and evidence. *American economic review 109*(10), 3395–3436.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *The Quarterly Journal of Economics 131*(4), 1753–1794.

Bulman, G. (2019). Law Enforcement Leaders and the Racial Composition of Arrests. *Economic Inquiry*.

Carrell, S. E., M. E. Page, and J. E. West (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics 125*(3), 1101–1144.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Close, B. R. and P. L. Mason (2006). After the traffic stops: Officer characteristics and enforcement actions. *The BE Journal of Economic Analysis & Policy 6*(1).

Craigie, T.-A. (2020). Ban the box, convictions, and public employment. *Economic Inquiry 58*(1), 425–445.

Cullen, Z. B. and R. Perez-Truglia (2020). The old boys' club: Schmoozing and the gender gap. Working paper, National Bureau of Economic Research.

Cunningham, J. P. and R. Gillezeau (2018, May). Racial differences in police use of force: Evidence from the 1960s civil disturbances. *AEA Papers and Proceedings 108*, 217–21.

Doleac, J. L. and B. Hansen (2018). Does "ban the box help or hurt low-skilled workers? statistical discrimination and employment outcomes when criminal histories are hidden. *Journal of Labor Economics Forthcoming*.

Donohue III, J. J. and S. D. Levitt (2001). The impact of race on policing and arrests. *The Journal of Law and Economics 44*(2), 367–394.

Egan, M. L., G. Matvos, and A. Seru (2018). When Harry fired Sally: The Double Standard in Punishing Misconduct. Working paper, National Bureau of Economic Research.

Flanagan, F. X. (2018). Race, gender, and juries: Evidence from north carolina. *The Journal of Law and Economics 61*(2), 189–214.

Garner, M., A. Harvey, and H. Johnson (2020). Estimating effects of affirmative action in policing: A replication and extension. *International Review of Law and Economics 62*, 105881.

Gershenson, S., C. Hart, J. Hyman, C. Lindsay, and N. W. Papageorge (2018). The long-run impacts of same-race teachers. Working paper, National Bureau of Economic Research.

Giuliano, L., D. I. Levine, and J. Leonard (2009). Manager race and the race of new hires. *Journal of Labor Economics 27*(4), 589–631.

Glover, D., A. Pallais, and W. Pariente (2017). Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores. *The Quarterly Journal of Economics 132*(3), 1219–1260.

Goncalves, F. and S. Mello (2020). A Few Bad Apples?: Racial Bias in Policing. Working paper.

Harvey, A. and T. Mattia (2020). Reducing racial disparities in crime victimization. Technical report.

Hengel, E. (2019). Publishing while female. are women held to higher standards? evidence from peer review.

Hoekstra, M. and C. Sloan (2020). Does Race Matter for Police Use of Force? Evidence from 911 Calls. Working paper.

Horrace, W. C. and S. M. Rohlin (2016). How dark is dark? bright lights, big city, racial profiling. *Review of Economics and Statistics 98*(2), 226–232.

Kessler, J. B., C. Low, and C. D. Sullivan (2019). Incentivized resume rating: Eliciting employer preferences without deception. *American Economic Review 109*(11), 3713–44.

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

Knox, D., W. Lowe, and J. Mummolo (2020). Administrative Records Mask Racially Biased Policing. *American Political Science Review 114*(3), 619–637.

Kofoed, M. and E. mcGovney (2019). The Effect of Same-Gender or Same-Race Role Models on Occupation Choice. *Journal of Human Resources Forthcoming*.

Kofoed, M. S. et al. (2019). The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at west point. *Journal of Human Resources 54*(2), 430–467.

Langan, A. (2018). Female Managers and Gender Disparities: The Case of Academic Department Chairs. Working paper.

Lowe, M. (2019). Types of contact: A field experiment on collaborative and adversarial caste integration. Working paper.

Mason, P. L. (2007). Driving while black: do police pass the test? *Swedish Economic Policy Review 14*, 79–113.

McCrary, J. (2007). The Effect of Court-ordered Hiring Quotas on the Composition and Quality of Police. *American Economic Review 97*(1), 318–353.

Miller, A. and C. Segal (2018). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence. *Review of Economic Studies Accepted*.

Mousa, S. (2020). Building social cohesion between christians and muslims through soccer in post-isis iraq. *Science 369*(6505), 866–870.

Mueller-Smith, M. and K. Schnepel (2017). Diversion in the criminal justice system: Regression discontinuity evidence on court deferrals. Technical report, Working paper.

Neumark, D., I. Burn, and P. Button (2019). Is it harder for older workers to find jobs? new and improved evidence from a field experiment. *Journal of Political Economy 127*(2), 922–970.

Nix, J., B. A. Campbell, E. H. Byers, and G. P. Alpert (2017). A Bird's Eye View of Civilians Killed by Police in 2015. *Criminology and Public Policy 16*(1), 309–340.

Park, K. H. (2017). The impact of judicial elections in the sentencing of black crime. *Journal of Human Resources 52*(4), 998–1031.

Parsons, C. A., J. Sulaeman, M. C. Yates, and D. S. Hamermesh (2011). Strike three: Discrimination, incentives, and evaluation. *American Economic Review 101*(4), 1410–35.

Police Accountability Task Force (April 2016). Recommendations for reform: Restoring trust between the chicago police and the communities they serve. Technical report.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review 109*(3), 774–809.

Rehavi, M. M. and S. B. Starr (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy 122*(6), 1320–1354.

Rim, N., B. Ba, and R. Rivera (2020). Disparities in Police Award Nominations: Evidence from Chicago. *AEA Paper and Proceedings 110*, 447–451.

Sarsons, H. (2019). Interpreting Signals in the Labor Market: Evidence from Medical Referrals. Working paper.

Sarsons, H. (2020). Gender differences in recognition for group work. *Journal of Political Economy Accepted*.

Sklansky, D. A. (2005). Not Your Father's Police Department: Making sense of the new demographics of law enforcement. *J. Crim. L. & Criminology 96*(3), 1209–1243.

United States Department of Justice (March 4, 2015). Investigation of the Ferguson Police Department. Technical report, Civil Rights Office.

Weisburst, E. (2018). Whose Help Is on the Way? The Importance of Individual Police Officers in Law Enforcement. Working paper.

West, J. (2018). Racial Bias in Police Investigations. Working paper.

Worden, R. E., M. Kim, C. J. Harris, M. A. Pratte, S. E. Dorn, and S. S. Hyland (2013). Intervention with problem officers: An outcome evaluation of an eis intervention. *Criminal justice and behavior 40*(4), 409–437.

Zeltzer, D. (2020). Gender homophily in referral networks: Consequences for the medicare physician earnings gap. *American Economic Journal: Applied Economics 12*(2), 169–97.

Online Appendix for

"The Black-White Recognition Gap in Award

Nominations"

A   Additional Tables and Figures

Table A1: Department Awards

1	*Superintendent's Award of Valor* for an act of outstanding bravery or heroism by which the member has demonstrated in great degree the characteristics of selflessness, personal courage, and devotion to duty.
2	*Superintendent's Award of Merit* for an outstanding accomplishment that has resulted in improved administration, improved operations, or substantial savings in manpower or operational costs, wherein the member has gone far beyond the requirements of their normal assignment.
3	*Police Blue Star Award* is granted to any sworn member who has been seriously, critically, or fatally injured while in the performance of police duty.
4	*Police Blue Shield Award* is granted to any sworn member who, as a result of accidental causes, has been seriously, critically, or fatally injured while in the performance of police duty.
5	*Superintendent's Award of Tactical Excellence* for exceptional tactical skills or verbal approaches and techniques to mitigate any deadly force situation resulting in the saving or sustaining of a human life.
6	*Arnold Mireles Special Partnership Award* for making a significant impact upon the quality of life within their community by identifying and resolving problems.
7	*Special Commendation* for making a significant impact on public safety or crime prevention.
8	*Lifesaving Award* for a successful effort in saving a human life that involved exceptional courage or performance.
9	*Police Officer of the Month* for performance of duty during a specific month was characterized by such exceptional professional skill that it merits recognition by the entire Department.
10	*Chicago Police Leadership Award* for exemplary service, dedication, and leadership.
11	*Department Commendation* for an outstanding act or achievement that brings great credit to the Department and involves performance above and beyond that required by the member's basic assignment.
12	*Problem Solving Award* for an exemplary effort to identify, analyze, and successfully respond to causes, conditions, and problems that may lead to crime and neighborhood disorder.
13	*Joint Operations Award* for efforts and participation in a broad multi-agency joint operation/event, spanning several days or more, significantly contributing to the overall successes of the operation.
14	*Unit Meritorious Performance Award* for exhibiting exceptional professional skill and conduct during a coordinated action.
15	*Traffic Stop of the Month Award* for excellence in conducting professional traffic stops that result in quality arrests.
16	*Top Gun Arrest Award* for exceptional commitment to the recovery of illegal firearms.
17	*Special Service Award* for contributing to any event that has a significant impact upon the historical direction and operations of the Department.
18	*Honorable Mention Certificate* for demonstrating outstanding performance above and beyond that required by the member's assignment.

Source: Chicago Police Department Special Order S01-01-01 "Description and Eligibility for Department Awards", retrieved from http://directives.chicagopolice.org/directives/

Table A2: CPD Use of Force Options and Member Response

Use of Force Options	Our Classification

Force Mitigation Efforts
  Member Presence
  Zone of Safety
  Verbal Direction/Control Techniques
  Movement to Avoid Attack
  Specialized Units
  Tactical Positioning
  Additional Unit Members
  None
  Other
	Mitigation

Control Tactics
  Escort Holds
  Wristlock
  Armbar
  Control Instrument
  Pressure Sensitive Areas
  Emergency Handcuffing
  Other
	Control tactics

Response without Weapons
  Open Hand Strike
  Take down
  Elbow strike
  Close hand strike/Punch
  Knee strike
  Kicks
  Other
	No Weapon

Response with Weapons
  OC/Chemical Weapon
  OC/Chemical Weapon w/Authorization
  LRAD w/Authorization
	Non-Lethal Weapon

Taser	Taser
Canine	Canine
Baton/Expandable baton Impact munitions	Baton
Revolver Rifle Semi-auto pistol Shotgun	Firearm
Other	Other Use of Force

Source: Chicago Police Department TRR Form

Table A3: Evaluation Quarter and Due Dates by Start Month

Quarter	Anniversary Date Month of the Member	The Quarter the Member Will Be Evaluated	Due Date of the Evaluation
1st	January, February, March	4th	30 January
2nd	April, May, June	1st	30 April
3rd	July, August, September	2nd	30 July
4th	October, November, December	3rd	30 October

Source: Chicago Police Department, Career Development Directive, Employee Resource E05-01, Section IX, B. Retrieved from http://directives.chicagopolice.org/directives/data/a7a56e3d-12887ea9-ce512-887e-c3dce7cd73e28d57.html?ownapi=1

Table A4: Impact of Arrest Record on Nomination Likelihood by Officer Race, With Officer Fixed-Effects

| Estimates for:	Outcome Variable: Nominated		
	White Officer	Black-White Gap	Hispanic-White Gap
	(1)	(2)	(3)
Panel A: White Supervisors			
1 arrest	0.00629***	-0.00495**	-0.00118
	(0.00143)	(0.00206)	(0.00254)
2 arrests	0.00927***	-0.00543*	0.00138
	(0.00202)	(0.00296)	(0.00350)
3 arrests	0.0128***	-0.00736*	0.00243
	(0.00274)	(0.00418)	(0.00474)
4 arrests	0.0154***	-0.00430	-0.00156
	(0.00342)	(0.00588)	(0.00591)
5 or more arrests	0.0326***	-0.0154***	-0.00830
	(0.00308)	(0.00531)	(0.00509)
Observations		171,094	
Panel B: Black Supervisors			
1 arrest	-0.000415	0.00285	0.00215
	(0.00367)	(0.00414)	(0.00760)
2 arrests	0.00562	-0.00393	-0.000198
	(0.00528)	(0.00593)	(0.0108)
3 arrests	0.00573	0.00438	0.00756
	(0.00721)	(0.00848)	(0.0149)
4 arrests	0.00254	-0.00861	0.00880
	(0.0104)	(0.0119)	(0.0189)
5 or more arrests	0.00949	0.0274**	-0.000599
	(0.00898)	(0.0127)	(0.0155)
Observations		29,413	
Mean Pr(Nom) for White Officers		-0.0045	

Source: CPD data.

Notes: This table reports estimates for the impact of an officer's arrest record on the probability of nomination by white supervisors (Panel A) and by black supervisors (Panel B). Each panel is a single OLS regression with estimates for white officers in column 1, the black-white difference in column 2, and the Hispanic-white difference in column 3. All estimates control for month and year FE, officer FE, officer birth year, tenure, lagged complaints, unit FE, share of black supervisees, lagged crime rate, lagged violent crime rate. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure A1: Black-White Nomination Gap by Supervisor Characteristics

(a) Share of Black Supervisees
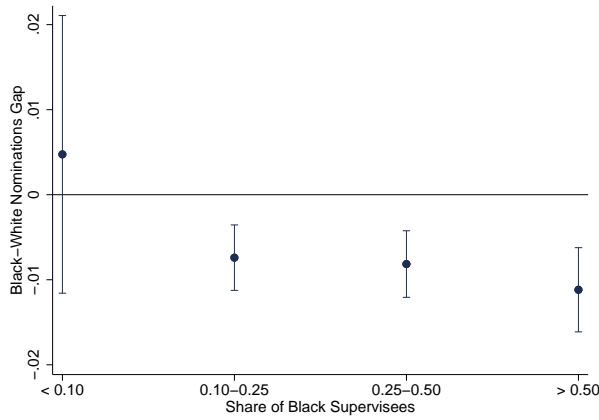
(b) Supervisor Tenure
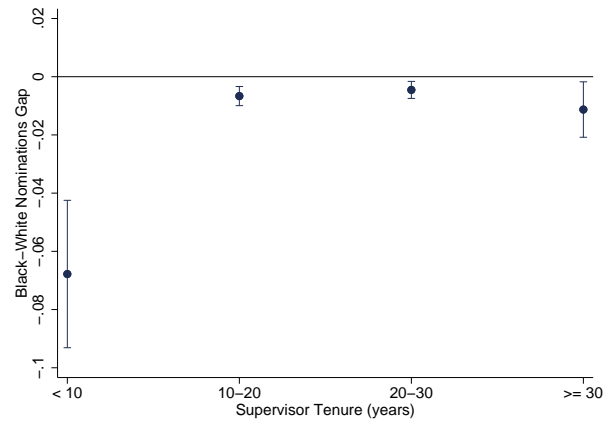
(c) Number of Supervisees
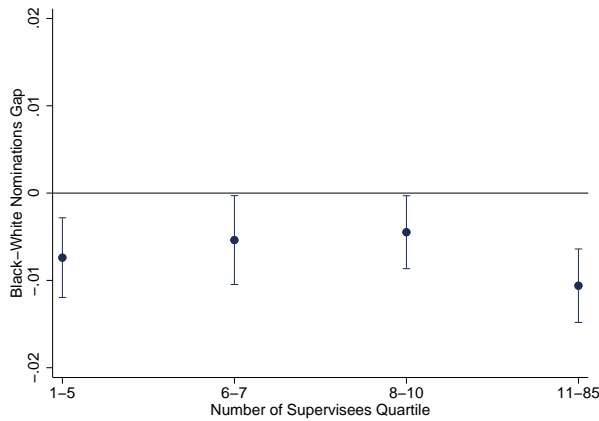
Notes: White Supervisors. Estimates control for officer birth year, officer tenure, lagged arrests, lagged complaints, supervisor fixed effects, district, overall crime rate, violent crime rate, number of supervisees, and share of black supervisees, as well as year and month fixed effects. Wings depict 95% confidence intervals by robust standard errors.

B Online Experiment

The experiment was conducted on Amazon Mechnical Turk (MTurk) in July 2020. It was pre-registered in the AEA RCT Registry, AEARCTR-0005929. We recruited 411 MTurk workers (hereafter "workers") who were compensated $1.20 for completing a survey experiment. Table A5 reports summary statistics on all 411 workers. Figure A2 plots the distribution of workers' states of residence.

We included three data quality checks to identify bots and to ensure workers paid attention during the survey. For the analysis, we decided to include workers who passed at least two of the three data quality checks. This restriction reduces our final analysis sample to 407 workers.

To avoid deception in our survey, we used real officer profiles but used officer initials to preserve officers' identities. Workers were informed that the profiles belonged to real officers but were not told which agency they were from. Further, we informed workers that their nominations would be relayed to the police department. This was to achieve incentive compatibility.
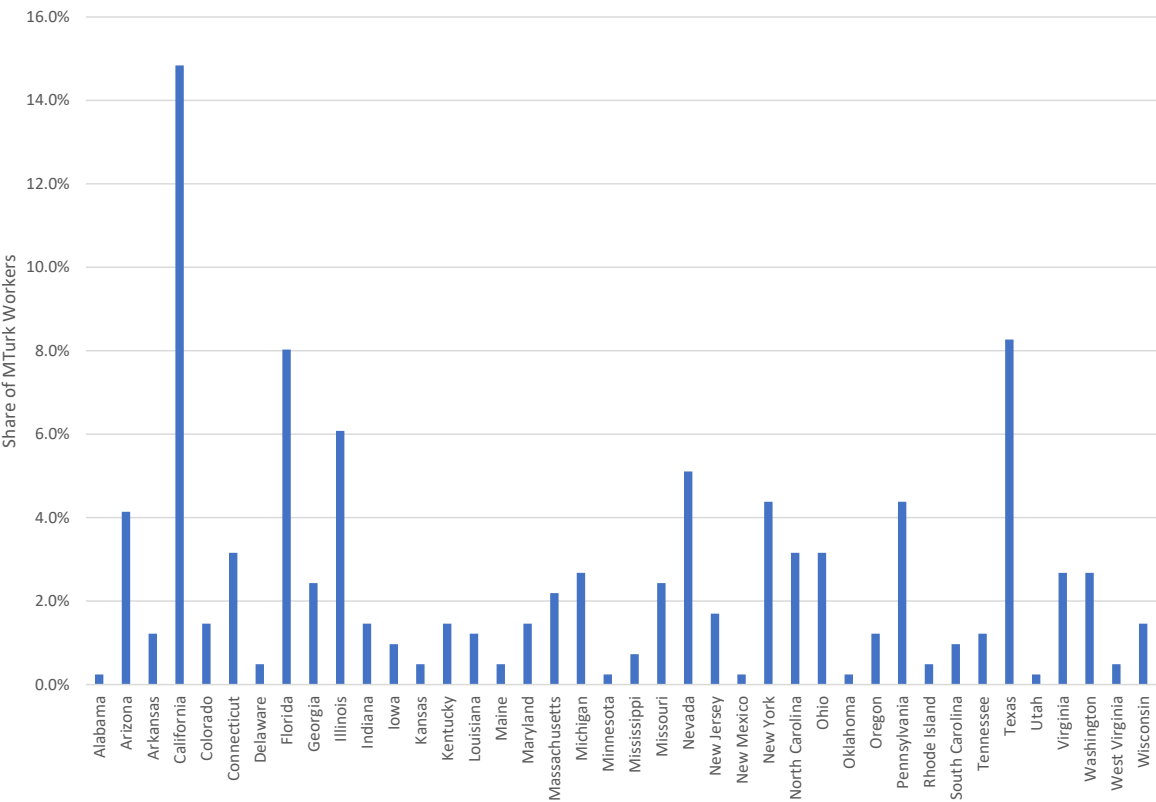
Table A5: Summary Statistics

	N	%
Race		
Black	48	11.7%
Hispanic	66	16.1%
White	263	64.0%
Other	21	5.1%
Prefer not to answer	10	2.4%
Missing	3	0.7%
Female	166	40.4%
Age		
18-25	53	12.9%
26-35	189	46.0%
36-45	78	19.0%
46-55	55	13.4%
56+	35	8.5%
Missing	1	0.2%
Is English your first language?		
Yes	401	97.6%
No	5	1.2%
Missing	5	1.2%
Length of Residency in US		
< 1 yr	6	1.5%
More than 1 yr but less than 3 yrs	21	5.1%
More than 3 yrs but less than 6 yrs	16	3.9%
More than 6 yrs	365	88.8%
Missing	3	0.7%
Number of Surveys (MTurk Workers)	411	

Source: MTurk survey data.

Figure A2: Distribution of MTurk Worker State of Residence

Source: MTurk survey data.