

In-group Bias and the Police: Evidence from Award Nominations*

Nayoung Rim[†]

Bocar Ba[‡]

Roman Rivera[§]

October 13, 2019

Abstract

This paper examines the impact of in-group bias on the internal dynamics of a police department. Prior studies have documented racial bias in policing, but little is known about bias against officers due to lack of available data. We construct a novel panel dataset of Chicago Police Department officers with detailed personnel information. Exploiting quasi-random variation in supervisor assignment, we find that white supervisors are less likely to nominate black officers than white or Hispanic officers. There is weaker evidence that male supervisors are less likely to nominate female officers than male officers. We explore several theories of discrimination that can explain our main findings. Requiring interaction between supervisors and officers reduces the minority nomination gap, but white supervisors still exhibit in-group favoritism. Our findings suggest departments should focus on policies that address in-group bias due to its effect on career advancement.

JEL Codes: J71, J15, J48

Keywords: bias, police officers, award nominations, supervisors

*We would like to thank Peter Arcidiacono, Pat Bayer, Kerwin Charles, Kareem Haggag, Damon Jones, and Robert Moffitt for helpful comments. We thank Sam Stecklow, the Invisible Institute, and Craig Futterman for help with the data. Weizhao Sun provided excellent research assistance. We thank the Duke Economics Department for generous financial support. The views expressed herein do not reflect the position of the Chicago Police Department, the U.S. Naval Academy, or the Department of the Navy. All errors are our own.

[†]United States Naval Academy, Department of Economics. Email: rim@usna.edu

[‡]University of Pennsylvania Law School. Email: bocarba@law.upenn.edu

[§]Columbia University, Department of Economics. Email: rgr2122@columbia.edu

1 Introduction

The recent high-profile cases of unarmed black men fatally shot by police officers raise the question of whether racial bias exists in police use of force (Nix et al. 2017; Fryer 2018; Knox, Lowe, and Mummolo 2019).¹ Even in more commonplace interactions, such as receiving speeding tickets and citations for traffic violations, a growing literature is uncovering racial bias against minority civilians (West 2018; Goncalves and Mello 2018; Coviello and Persico 2015). These racial disparities in law enforcement present an important social policy concern in the United States, particularly as there need not be a trade-off between effectiveness and fairness in policing (Persico 2002).

One proposed solution is increased diversity in police ranks, which traditionally are overwhelmingly white and male.² In addition to improving policing quality (Donohue III and Levitt 2001; McCrary 2007; Miller and Segal 2018; Hoekstra and Sloan 2019; Bulman 2019)³, minority representation, particularly at higher ranks of office, may help 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 (Langan 2018; Kofoed and mcGovney 2019). At the same time, these policies may be constrained by the extent to which officer bias carries over to their colleagues. An extensive literature documents racial and gender bias in the workplace, which may hinder minorities' career progression.⁴

¹See <http://time.com/4404987/police-violence/>.

²For 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/>)

³McCrary 2007 finds that municipal police departments that were ordered by courts to implement racial hiring quotas saw a decrease in the fraction black among violent and property crime arrestees though the overall effect on police performance is mixed. Bulman 2019 finds that the ratio of black-to-white arrests is higher under white sheriffs.

⁴For example, Egan, Matvos, and Seru 2018 find that female financial advisors are 20% more likely

The theoretical predictions for law enforcement are unclear. On one hand, studies detecting racial bias in policing imply that officers may engage in similar behavior towards their minority colleagues. On the other hand, the organizational culture of policing is that of a homogenous identity with a uniform goal of reducing crime (Sklansky 2011; Sklansky 2005). This insider identity, that “blue is blue” and the resulting irrelevance of the individual identity of the person in uniform, is shaped by the nature of the job and may limit bias to interactions with “outsider” subjects. Empirically, little is known about in-group bias against officers due to the lack of available data. The primary purpose of this paper is to examine the impact of bias on the internal dynamics of a police department.

We construct a novel panel dataset of all Chicago Police Department (CPD) officers between 2009 and 2015 containing detailed personnel information. This was a huge undertaking involving more than ten separate datasets gathered from different sources. Most importantly, this dataset contains information on officer use of force, arrests, and misconduct—crucial information in an empirical study of bias in the workplace.

Using supervisor nominations for departmental awards, we test whether police officers are biased against their minority colleagues.⁵ Our identification strategy exploits an institutional feature of the Chicago Police Department that assigns a new supervisor every January. Similar to Chetty, Friedman, and Rockoff 2014, which uses quasi-random variation induced by teacher turnover, we use a new supervisor assignment to approximate random assignment of a supervisor’s race or sex to an officer. We confirm random assignment by analyzing supervisor-officer assignments and find that officers do not sort to majority-group supervisors

than male financial advisors to lose their jobs following a misconduct. In medicine, Sarsons 2018 finds that physicians are less likely to refer to new female surgeons after a bad patient outcome but not to new male surgeons. Beaman, Keleher, and Magruder 2018 find that women are harmed in a referral-based hiring process as workplace networks tend to be gender homophilous. Glover, Pallais, and Pariente 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.

⁵We analyze award nominations rather than promotion decisions for two reasons. First, promotions to Sergeant are largely based on results from a written exam, making it difficult to capture the subjective opinions of officers towards their colleagues. In contrast, award nominations must originate from supervisors or superior officers. Second, departmental awards are highly predictive of receiving overtime pay and are used in officer evaluations and merit promotions.

based on work performance measures, suggesting that we should not see any racial or gender differences in nominations in the absence of supervisor bias.

We find that after switching to a white supervisor, white officers are 4.2 percentage-points more likely to be nominated compared to when they were assigned to a non-white supervisor. This equates to a 24% increase and is statistically significant at the 5% level. On the other hand, black officers are nearly 20% less likely than white officer to be nominated after switching to a white supervisor. We find similar patterns when we define in-group by gender. Male officers are 5 percentage-points or 28% more likely to be nominated after they switch to a male supervisor, and this estimate is significant at the 1% level. Female officers, in contrast, are 13% less likely to be nominated though this estimate is not significant. These results control for officer characteristics, including experience, arrest record, and misconduct, and time-varying neighborhood factors, like district-level crime rates.

Our second analysis exploits the richness of our panel data and examines the impact of arrests on officer's likelihood of being nominated for an award. We find that there are increasing returns to having more arrests, with a marked increase in nomination likelihood for those with 5 or more arrests. Black officers under white supervisors also see increasing returns to having more arrests, but the return is less. For example, between a black officer and a white officer, each with one arrest, the same years of experience, and in the same district, the average white supervisor is nearly 80% less likely to nominate the black officer for an award. This gap diminishes as the number of arrests increases, but does not go away. For example, the average black officer with 5 or more monthly arrests—who is at the 94th percentile of his distribution—is still 47% less likely to be nominated by a white supervisor compared to the average white officer with 5 or more monthly arrests, who is at the 81st percentile of his distribution.

We explore several mechanisms that can explain our main findings. Up to about half of the minority nomination gap can be explained by work performance, neighborhood characteristics, and supervisor interaction, with the other 50% remaining unexplained. We consider

whether implicit preferences can explain the rest. Although it is difficult to detect implicit bias without a scientific measure, we do find that white supervisors who are less likely to nominate black officers are also more likely to use force against black subjects relative to white subjects.

An alternative explanation concerns supervisor-officer interactions. White supervisors may be less likely to nominate black officers because they are less likely to interact with blacks, and, therefore, are less likely to be informed of their accomplishments. We find that interaction between the supervisor and the officer is an important factor in nomination decisions. Specifically, officers are more likely to be nominated in the quarters leading up to their annual evaluation. However, this boost falls after their evaluation, likely due to the decentralized nature of supervision in the CPD. One implication is that officers may benefit from more stable and consistent relationships with their supervisors. However, we also find that black officers benefit less from the required interaction with white supervisors relative to their white peers. Therefore, a race-blind policy of requiring consistent supervisor assignments may not be enough to eliminate the black-white nomination gap.

Our findings are consistent with studies that find that minorities are less likely to be acknowledged for their work ([Hengel 2019](#); [Sarsons 2019](#)) and a literature that establishes the existence of bias among managers and work colleagues ([Bertrand and Mullainathan 2004](#); [Giuliano, Levine, and Leonard 2009](#); [Glover, Pallais, and Pariente 2017](#); [Egan, Matvos, and Seru 2018](#); [Sarsons 2018](#); [Bohren, Imas, and Rosenberg 2018](#)). The current evidence is mixed on whether bias is due to tastes (preferences) or statistical discrimination, as results vary depending on the study setting.⁶ Our paper analyzes bias among police officers, linking this literature to studies on racial disparities in law enforcement. Consistent with [Glover, Pallais, and Pariente 2017](#), we find that interactions with minority officers are an important mechanism to reduce disparities. Like [Sarsons 2019](#), we find that statistical discrimination cannot fully explain supervisor behavior since controlling for evaluation quarter does not

⁶Additionally, [Bohren et al. 2019](#) suggest the accuracy of the individual's beliefs is important and argues for a distinction between accurate statistical discrimination and inaccurate statistical discrimination.

eliminate the minority nomination gap.

With respect to law enforcement, our study adds to the growing research that finds racial bias in policing. Prior studies largely use data on officer-initiated encounters (Knowles, Persico, and Todd 2001; Anwar and Fang 2006; Antonovics and Knight 2009; Nix et al. 2017; Goncalves and Mello 2018), which may be biased because they do not include the universe of all possible police interactions (Knox, Lowe, and Mummolo 2019). Two notable exceptions are Weisburst 2018, which uses data on officers randomly dispatched to respond to 911 calls, and West 2018, which uses data on officers randomly dispatched to investigate automobile crashes. Similarly, our paper bypasses the truncated data problem by focusing on nomination rates of quasi-randomly assigned officers.

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 analysis sample. Section 4 presents our empirical strategy and tests the identifying assumption that supervisors are randomly assigned to officers conditional on race and sex. We present our main results in Section 5, showing that across different analyses, white supervisors are less likely to nominate black officers than white or Hispanic officers. Section 6 discusses several mechanisms that can explain our findings. Section 7 concludes.

2 Basic Facts about CPD’s Structure

The Chicago Police Department is composed of five bureaus: Bureau of Patrol, Bureau of Detectives, Bureau of Organized Crime, Bureau of Organizational Development, and Bureau of Technical Services. These bureaus are further sub-divided into units or, in the case of the Bureau of Patrol, districts. The Bureau of Patrol is responsible for “general field operations, including the protection of life and property; apprehension of criminals; and enforcement of traffic laws and ordinances (Chicago Police Department, Department Organization Directive,

General Order G01-02-03).”⁷ This bureau is composed of 25 geographic districts spanning the city of Chicago and a Special Functions Division, whose mission is to support Bureau of Patrol operations.⁸

Upon graduation from Police Academy, all Police Officers begin their career in one of the 25 geographic districts. These 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.⁹ Sergeants prepare officers for duty and roll call, monitor officer activity, evaluate officers’ performance annually.¹⁰ Promotions to Sergeant are based on a 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.¹¹ 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.

⁷Retrieved from <http://directives.chicagopolice.org/directives/data/a7a57be2-1291da66-88512-91e6-521a90347177e975.html>

⁸Examples of Special Functions Division within the Bureau of Patrol are the special weapon and tactics unit, mounted unit, and canine unit. Between 2012-2014, districts 13, 21, and 23 were dissolved and merged with neighboring districts. See Appendix Figure A1 for a map.

⁹Among supervisors for whom we have unit information, 82% are assigned to officers in the same unit.

¹⁰Chicago Police Department, Career Development Directive, Employee Resource E05-01, retrieved from <http://directives.chicagopolice.org/directives/data/a7a56e3d-12887ea9-ce512-887e-c3dce7cd73e28d57.html?h1=true>

¹¹Chicago 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>

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 Officers 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 from aggregated data, using the most common observation across datasets.¹² Officer rank is taken from salary data provided by the Chicago Department of Human Resources (DHR), covering 2002 to 2017.

Supervisors. The supervisor 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).¹³ We consider all performance awards that are open to all sworn De-

¹²Not all demographic information is complete in each file, so an aggregation of demographic variables across multiple files is necessary.

¹³An award may be deleted for various reasons, including: the form was not filled out correctly; supporting

partment members and require a supervisor’s nomination.¹⁴ 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 to Cook County jail as a proxy for the arrest date.¹⁵ For our analysis, we use total officer 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.¹⁶ 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.¹⁷

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. For civilian allegations, the complainant must sign a sworn affidavit to certify that the allegation is true and correct. Only then is the complaint filed and investigated. We use all allegations about an officer, regardless of whether or not they were investigated.

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.

¹⁴Most 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.

¹⁵In 96.9% of cases, the release date is on the same day or the day after the arrest date, and 100% of release dates are within four days of the arrest.

¹⁶Violent 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.

¹⁷A comprehensive list of crime categories can be found at http://gis.chicagopolice.org/clearmap_crime_sums/crime_types.html.

Tactical Response Reports (TRR). 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.¹⁸ 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 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 district¹⁹ 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

¹⁸Strong 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.

¹⁹We 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.

led to an arrest. We construct monthly crime rates²⁰ for each district, separately for total crimes, property crimes, and violent crimes.²¹ 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. There are no large differences between men and women in terms of age and start-year.

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.

Table 2 presents racial and gender 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 2.3% chance of being nominated in a given month, which equates to about a 28% chance of being nominated in a given year. Whites and Hispanics have slightly higher than average likelihoods at 2.7% and 2.9%, respectively, while the likelihood for black officers is half the sample average (1.1%). The black-white difference is statistically significant at the 1% level. Similarly, the average female officer has a 1.2% chance—about the same odds as an average black officer—while the average male officer has a 2.7% chance of being nominated in a given month. The female-male difference is also statistically significant

²⁰Crime rate is defined as the total number of reported incidents of crime divided by the population and multiplied by 1000.

²¹FBI does not record statistics for non-index crimes.

Table 1: Summary Statistics

	Police Officers			Supervisors		
	All	Male	Female	All	Male	Female
Race						
White	46.4%	48.6%	40.3%	69.7%	71.7%	60.7%
Black	26.8%	23.1%	37.3%	14.7%	12.4%	24.7%
Hispanic	23.2%	24.3%	20.2%	14.0%	14.0%	13.8%
Asian	3.1%	3.6%	1.8%	1.6%	1.8%	0.8%
Native American	0.4%	0.4%	0.4%	0.1%	0.1%	
Female	26.3%			19.2%		
Birthyear	1970.3	1970.7	1969.1	1965.3	1965.4	1965.1
Start Year	2000.0	2000.1	1999.6	1992.2	1992.2	1992.4
N	6,518	4,769	1,698	1,284	1,037	247

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, but women receive slightly fewer complaints than men (0.03 vs 0.04). The female-male difference is statistically significant at the 1% level.

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. Similar to the patterns for nominations and complaints, arrest rates by female officers are similar to those by black officers (1.1 vs 1.2, respectively), while male officers are similar to white officers (2.1 vs 2.0, respectively). The female-male difference 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 up around 65% of all arrests. Here, the difference is about -0.70 arrests per month or 8.4 fewer

Table 2: Racial and Gender Differences in Work Measures

Sample:	Everyone	White	Black	Hispanic	Male	Female
Nominated	0.0231	0.0272	0.0115***	0.0291	0.0271	0.0123***
Complaints	0.0425	0.0421	0.0422	0.0441	0.0467	0.0311***
Total Arrests	1.8246	2.0414	1.1939***	2.1641	2.0767	1.1308***
Violent	0.3672	0.3718	0.3117*	0.4201**	0.4010	0.2740***
Property	0.2713	0.2941	0.2015***	0.3034	0.2944	0.2078***
Non-Index	1.1861	1.3754	0.6807***	1.4406	1.3813	0.6489***
Drug	0.3125	0.3723	0.1411***	0.4050	0.3782	0.1315***
Traffic	0.1239	0.1471	0.0610***	0.1573	0.1411	0.0765***
Observations	250,872	111,876	70,572	59,148	184,020	66,852

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. Asterisks denote p-values from hypothesis tests of equal means between the minority group and the majority group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

arrests per year and is statistically significant at the 1% level. Similarly, the female-male difference for non-index crimes is also large at -0.73 arrests per month.

Although the data reveal a disparity in number of arrests, we caution the reader from jumping to the conclusion that black and/or female officers are less productive than majority-group officers. Arrests are not a comprehensive measure of policing quality and may be a biased measure. For example, a comparison of female and male officers’ arrest records would suggest that female officers are less productive than male officers. However, a study by [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 allow for greater officer discretion, can also be seen as a delineation between appropriate and inappropriate uses of police authority.²² In

²²We borrow this term and classification from [Worden et al. 2013](#). We do not know whether an arrest

the last two rows of Table 2, we see that white officers are about 2.4-2.6 times more likely to arrest someone for drugs or traffic violations relative to black officers. Male officers are also about 1.8-2.8 times more likely than female officers to make these arrests. In contrast, the black-white difference for more serious crimes, like violent crimes, is economically small at -0.06 arrests a month and is weakly significant at 10%. These facts suggest that it is important to control for work measures in our analysis, though 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 majority-group supervisors exhibit in-group favoritism 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 or sex to an officer.²³ Thus, we can simply compare nomination rates when an officer is assigned to a white supervisor vs. a non-white supervisor, to test for the existence of in-group favoritism. This identification strategy is similar to [Chetty, Friedman, and Rockoff 2014](#), which develops a quasi-experimental method by exploiting student exposure to teachers of varying value-added levels that is induced through teacher turnover. In this paper, we use officer exposure to supervisors of different races and sex 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

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.

²³About 96% of officers are assigned to a supervisor in January of each calendar year. [Rim, Ba, and Rivera 2019](#) establishes that officers are as-good-as-randomly assigned to districts, implying that they are also quasi-randomly assigned to supervisors.

supervisor relationships between 2009 and 2015 lasted 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. Next, we examine whether the data support the argument that officers are randomly assigned to supervisors.

4.1 Exogeneity of Officer Performance and Supervisor Assignment

Throughout the paper, we want to interpret any change in nomination likelihood when minority officers are assigned to majority-group supervisors relative to when they are assigned to minority supervisors as a causal effect of supervisor race or sex.²⁴ The key assumption is that minority officers were not systematically assigned to majority-group 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 in-group supervisors.

To test this, we examine whether officer performance is correlated with supervisor race or sex.²⁵ Because officers are assigned to supervisors at the Department level, we use the full officer sample rather than the analysis sample that is restricted to officers whose supervisor assignment lasted one year.²⁶

²⁴Conversely, we would like to attribute any change in nomination likelihood when majority-group supervisors are assigned to minority officers vs. majority-group officers as a causal effect of officer race or sex.

²⁵In 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 in-group supervisors. It is selection on officer's characteristics *and* race together that we care about.

²⁶We also only consider officers who are assigned to a geographic district.

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

$$Maj_{it} = \beta_0 + W'_{it}\beta + X'_{it}\alpha + e_{it} \quad (1)$$

where Maj_{it} is equal to 1 if officer is assigned to a majority-group supervisor in month t and 0 otherwise. For example, when in-group is defined by race, $Maj_{it} = 1$ if the officer is assigned to a white supervisor and 0 otherwise. When in-group is defined by sex, $Maj_{it} = 1$ if the officer is assigned to a male supervisor and 0 otherwise. W is a vector of baseline controls, such as birth year, race, sex, tenure²⁷, district fixed effects, year fixed effects, district-year fixed effects, and an indicator for whether the officer and supervisor were in the same district last month. X is a vector of 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). All of the variables in X are lagged by one month and include second-order and third-order terms to capture non-linear relationships between these variables and supervisor assignment. Standard errors are clustered at the supervisor level.

To test our identification strategy, we are interested in whether officer work performance, X , is predictive of an assignment to a majority-group 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.²⁸ Table 3 presents adjusted p-values from officer measures in X .²⁹ Columns 1-3 present estimates for race, while columns 4-5 present

²⁷We also include squared and cubed measures of tenure.

²⁸The 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.

²⁹In addition, we estimate a version of equation (1) with and without the officer characteristics. The R-squared does not vary much between these two versions. It increases from 0.0923 for the baseline model to 0.1147 with the full set of controls when examining supervisor sorting by race, and from 0.0418 to 0.0585 when examining sorting by sex. That is, a flexible specification with officer performance measures interacted with race/sex increases the predictive power by an additional 1-2 percentage-points. This suggests that

estimates for gender. Coefficients and unadjusted standard errors are reported in Appendix Table A3.

There is no evidence that white, black, or Hispanic police officers sort to white supervisors based on their work performance. All of the adjusted p-values in columns 1-3 are above the 5% and even 10% threshold. Similarly, there is no evidence that female officers sort to male supervisors (Column 5). For male officers, non-index arrests may predict assignment to male supervisors, but the effect size is small; an additional non-index arrest increases the likelihood that a male officer is assigned to a male supervisor by 1.35 percentage-points the subsequent month.³⁰

In summary, we do not find strong evidence that minority officers sort to majority-group supervisors based on their work performance. This suggests that, in the absence of supervisor bias, we should not see any racial or gender differences in nominations. It also provides an indirect test of random assignment of officers to supervisors as we would not expect to see any correlation between officer work measures and supervisor assignment if officers are randomly assigned to supervisors. In Appendix Table A4, we check whether officer work performance is correlated with a supervisor switch (specifically a nonwhite-to-white supervisor switch and a female-to-male supervisor switch). Again, we do not see strong evidence that high-performing workers are more likely to switch supervisors.

officer characteristics do not explain much of the variation in supervisor race, which is reassuring.

³⁰We also estimate a logit regression and find very similar results. In fact, with a non-linear specification, we do not see any Holm-adjusted p-value that is less than 10%.

Table 3: Impact of Officer Characteristics on Supervisor Race and Sex

Dependent Variable: Officer Sample:	Supervisor is White			Supervisor is Male	
	White (1)	Black (2)	Hispanic (3)	Male (4)	Female (5)
Arrests					
Violent	1.0000	1.0000	1.0000	0.7726	1.0000
Violent ²	1.0000	1.0000	1.0000	1.0000	1.0000
Violent ³	1.0000	1.0000	1.0000	1.0000	1.0000
Property	0.3660	1.0000	0.9970	1.0000	1.0000
Property ²	1.0000	1.0000	1.0000	1.0000	0.9826
Property ³	1.0000	1.0000	1.0000	1.0000	1.0000
Non-Index	1.0000	1.0000	1.0000	0.0007	1.0000
Non-Index ²	1.0000	0.9039	1.0000	0.0816	1.0000
Non-Index ³	1.0000	1.0000	1.0000	0.2859	1.0000
Complaints	1.0000	1.0000	1.0000	1.0000	1.0000
Complaints ²	0.9713	1.0000	1.0000	1.0000	1.0000
Complaints ³	1.0000	1.0000	1.0000	0.9765	1.0000
TRR Filings	1.0000	1.0000	1.0000	0.5133	1.0000
TRR Filings ²	1.0000	1.0000	1.0000	0.5375	1.0000
TRR Filings ³	1.0000	1.0000	1.0000	0.6943	1.0000
Strong Force Ratio	1.0000	1.0000	0.8261	1.0000	1.0000
Strong Force Ratio ²	1.0000	1.0000	0.6586	1.0000	1.0000
Strong Force Ratio ³	1.0000	1.0000	0.3167	1.0000	1.0000
Weak Force Ratio	1.0000	1.0000	1.0000	0.4568	1.0000
Weak Force Ratio ²	1.0000	1.0000	1.0000	0.6848	1.0000
Weak Force Ratio ³	1.0000	1.0000	1.0000	0.7402	1.0000
Observations	9,175	2,923	4,636	15,080	2,337
Controls for:					
Race/Ethnicity				Yes	Yes
Sex	Yes	Yes	Yes		
Birth year	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes
District x Year	Yes	Yes	Yes	Yes	Yes

Notes: This table reports Holm-adjusted p-values for coefficients on officer performance measures. Standard errors are clustered at the supervisor level. Performance measures are lagged by one month. Non-index crimes are composed of municipal code violations, traffic violations, warrants, drug-related crimes, and other.

5 Results

In this section, we conduct two analyses to examine whether majority-group supervisors exhibit in-group bias. First, we test whether an officer’s probability of nomination changes when the officer switches supervisors, and whether there are any differential changes if the officer is in the supervisor’s in-group vs. out-group. Second, 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.

5.1 Supervisor Switch

This analysis uses a difference-in-differences approach to test whether majority-group officers’ nomination likelihood changes when they switch from a minority to a majority-group supervisor, and whether minority officers are differentially affected by the same supervisor switch. Assuming equal performance between minority and majority-group officers, there should be no difference in nomination rates in the absence of bias. Further, the DID estimates may be interpreted as causal if supervisors are randomly assigned to police officers. In other words, we use a supervisor switch to approximate random assignment of a supervisor’s race or sex to an officer.

The regression sample is unique at the officer-switch level. A switch is defined as a change from one supervisor to a different supervisor. For example, say an officer has three different supervisors during the sample period. Then, this officer has two switches: Supervisor A to Supervisor B, and Supervisor B to Supervisor C.³¹ To examine whether majority-group supervisors exhibit in-group bias, we focus on two types of switches: (1) from a non-white supervisor to a white supervisor, and (2) from a female supervisor to a male supervisor.

³¹It is possible that an officer is assigned to Supervisor A in year 1, Supervisor B in year 2, and back to Supervisor A in year 3. These are defined as two unique switches: A-B and B-A.

The regression model for examining non-white to white supervisor switches is³²:

$$\begin{aligned}
Nom_{isj} = & \beta_0 + \beta_1 PostSwitch_{isj} + \beta_2(PostSwitch_{isj} \times B_i) + \beta_3(PostSwitch_{isj} \times H_i) \\
& + \beta_4(PostSwitch_{isj} \times A_i) + \beta_5(PostSwitch_{isj} \times N_i) + X'_{isj}\alpha + \varepsilon_{isj}
\end{aligned} \tag{2}$$

where i denotes officer, s denotes switch, and j denotes supervisor. Nom_{isj} is equal to 1 if officer i was nominated for an award by supervisor j and 0 otherwise. $PostSwitch_{isj}$ is a binary variable equal to 1 if supervisor j is the new supervisor and equal to 0 if supervisor j is the old supervisor. For example, when examining non-white to white supervisor switches, $PostSwitch_{isj} = 1$ if supervisor j is white and 0 if non-white.

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. When in-group is defined by sex (e.g., when the sample is restricted to female to male supervisor switches), we include a single binary indicator variable, F_i , for whether the officer is female, and male officers are the reference group.

X_{isj} is a vector of officer and district characteristics measured during the officer's assignment to supervisor j . Officer characteristics include officer tenure (measured in years) at the start of switch s , total complaints, and total arrests. District characteristics include the average crime rate, average violent crime rate, and average arrest rate. If the officer changed districts during his or her assignment to supervisor j , then we use measures from the modal district. Standard errors clustered at the supervisor level.

The parameters of interest are β_1 and δ . β_1 tells us how nominations for majority-group (white or male) officers changed after they switched from a minority supervisor to an in-group supervisor, while δ tells us whether nominations for minority officers are differentially affected, relative to majority-group officers, from the same supervisor switch. Table 4 column 1 presents estimates for a non-white to white supervisor switch, while column 2 present

³²The regression model for examining female to male supervisor switches is: $Nom_{isj} = \beta_0 + \beta_1 PostSwitch_{isj} + \beta_2(PostSwitch_{isj} \times F_i) + X'_{isj}\alpha + \varepsilon_{isj}$.

estimates for a female to male supervisor switch. Columns 3 and 4 present results from placebo regressions of a white to white supervisor switch and a male to male supervisor switch, respectively.

White officers are 4.2 percentage-points more likely to be nominated after switching to a white supervisor, and this estimate is statistically significant at the 5% level. This is an effect size of 24% as the average white officer has a 17.5% chance of being nominated by his or her supervisor during the year. Black officers, however, do not benefit from the same switch. Black officers are 0.07 percentage-points more likely to be nominated following a switch to a white supervisor, which is 3.5 percentage-points smaller than the increase for white officers. Put another way, black officers are nearly 20% less likely ($= -3.5/17.5$) than white officers to be nominated after switching to a white supervisor. This estimate is significant at the 10% level. In contrast, Hispanic officers are almost 7% ($= 1.2/17.5$) more likely than white officers to be nominated after switching supervisors, but this estimate is not statistically significant.

We see similar patterns when we define in-group by gender and examine how nominations change when officers switch from a female to a male supervisor (Table 4 column 2). The average male officer has a 17.9% chance of being nominated. Switching from a female supervisor to a male supervisor increases the probability of nomination by 5 percentage-points for male officers, or an increase of 28% ($= 5/17.9$). Compared to male officers, female officers are 2.4 percentage-points less likely to be nominated, an effect size of -13% ($= -2.4/17.9$). This estimate is not statistically significant.

Columns 3 and 4 present results from a placebo test of same-race and same-sex switches. If majority-group supervisors are exhibiting in-group bias, then we would not expect to see any changes in nomination likelihood from a supervisor switch of the same race or gender. Column 3 says that the change in nomination likelihood for both white and black officers is minimal at 0.17 percentage-points and 0.64 percentage-points, respectively. Neither estimate is statistically significant. Similarly for gender, switching to a new male supervisor has

Table 4: Effect of Supervisor Switch on Nomination Likelihood

Switch Type:	Dependent Variable: Officer Nominated for Award			
	Non-White → White (1)	Female → Male (2)	White → White (3)	Male → Male (4)
Post-Switch	0.0420** (0.0183)	0.0502*** (0.0190)	-0.00166 (0.00844)	-0.000319 (0.00654)
Black Officer x Post-Switch	-0.0345* (0.0208)		0.00642 (0.0118)	
Hispanic Officer x Post-Switch	0.0121 (0.0322)		0.0121 (0.0135)	
Female Officer x Post-Switch		-0.0235 (0.0237)		0.00432 (0.00864)
Observations	4,706	3,350	11,876	15,600
Reference Group Mean	0.175	0.179	0.175	0.179
R-squared	0.517	0.537	0.436	0.409
Controls for:				
Officer Characteristics	Yes	Yes	Yes	Yes
District Characteristics	Yes	Yes	Yes	Yes

Notes: Sample is at the officer-switch level. The reference group in column 1 is white officers and male officers in column 2. Officer characteristics include officer tenure at the start of switch, total complaints, and total arrests. District characteristics include the average crime rate, average violent crime rate, and average arrest rate. Standard errors are clustered at the supervisor level and are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

minimal impacts on nomination rates for male and female officers (0.03 and 0.4 percentage-points, respectively). Again, neither estimate is statistically significant.

This test also ensures that our results are not an artifact of supervisor switch, in general. That is, an alternative explanation is that the switch itself may have resulted in an increase in nomination likelihood, rather than a change in supervisor race or sex. For example, perhaps white officers are better at socializing with new supervisors whereas black officers don't do well with change. If this is true, then we would expect to see similar estimates with *any* change in supervisor. However, our estimates in columns 3 and 4 do not support this argument.

Next, we discuss the validity of our results. The identifying assumption for DID is that of parallel pre-trends. In our context, that means the minority nomination gap under minority

supervisors should be zero, or that we have parallel leads.³³ Because our regression model includes officer fixed effects (equation 2), it cannot estimate the black-white gap or the female-male gap in the pre-switch period. However, we can manually estimate this through the Frisch-Waugh Theorem.³⁴ We estimate that the black-white difference is 0.002 and the female-male difference is essentially 0.³⁵ Neither are statistically significant.

A limitation of the supervisor-switch analysis is that it can only identify relative in-group bias. Since we are comparing nominations by majority-group supervisors to nominations by minority supervisors, an increase in nominations for white officers may be due to in-group favoritism by white supervisors or in-group bias by minority supervisors in the previous period. For example, in the Appendix we examine how nomination likelihoods change when officers switch from a majority-group supervisor to a minority supervisor. We find that white officers are 4.5 percentage-points less likely to be nominated after switching to a minority supervisor, while black officers are 4.3 percentage-points more likely to be nominated relative to white officers under white supervisors (Appendix Table A5).³⁶ Put another way, white officers see their nomination likelihoods decrease when they switch from white supervisors to non-white supervisors, while black supervisors do not necessarily see a change in their nomination likelihoods. For officers switching from male supervisors to female supervisors, male officers are 2.11 percentage-points less likely to be nominated after a switch and there is no differential effect for female officers. These estimates, however, are not statistically significant. Because the switch analysis conducts a relative analysis, it is difficult to determine the direction of the bias. To better answer this question, we examine whether there are differential returns to having more arrests for minority vs. majority-group officers.

³³One cannot actually test for parallel pre-trends without data on the counterfactuals. However, one *can* empirically test for parallel leads.

³⁴We do this separately to estimate the black-white nomination gap under non-white supervisors and the female-male nomination gap under female supervisors. First, we wipe the variation in nominations of all other covariates except officer race/sex (for race, we use a black dummy variable). Then we wipe the variation in the black/female dummy variable of the same covariates. Last, we regress the residuals from the first regression against the residuals from the second regression to obtain the black-white or female-male difference in nominations.

³⁵The analogous coefficients for the placebo regression are effectively zero.

³⁶These estimates are significant at the 5% level and 10% level, respectively.

5.2 Nomination Likelihood by Arrest Record

To supplement our supervisor switch analysis, we exploit the panel nature of our data to examine the impact of officers' work performance on their likelihood of being nominated for an award. That is, conditional on the officer's arrest record, are there racial or gender differences in the probability of nomination? We estimate the following model, separately for white supervisors and male supervisors. The regression sample for this analysis is at the officer-month level.

$$\begin{aligned}
 Nom_{it} = & \beta_0 + \beta_1 Arrests_{i,t-1} + \beta_2 (Arrests_{i,t-1} \times B_i) + \beta_3 (Arrests_{i,t-1} \times H_i) \\
 & + \beta_4 (Arrests_{i,t-1} \times A_i) + \beta_5 (Arrests_{i,t-1} \times N_i) + X'_{it} \alpha + \tau_t + \varepsilon_{it}
 \end{aligned} \tag{3}$$

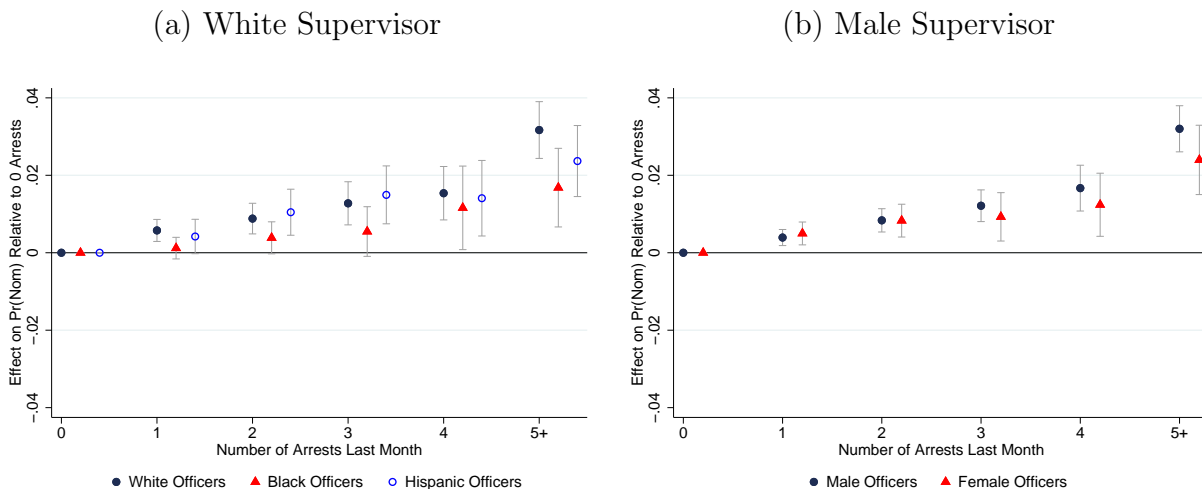
where i denotes officer and t denotes month. Nom_{it} is equal to 1 if officer i was nominated for an award in month t and 0 if not. $Arrests_{i,t-1}$ is the number of arrests officer i made last month. Rather than including this as a continuous measure, we break up the number of arrests into six categories: 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 at lower arrests. The reference group 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. When in-group is defined by sex (e.g., when the sample is restricted to female to male supervisor switches), we include a single binary indicator variable, F_i , for whether the officer is female, and male officers are the reference group.

X_{it} is a vector of officer and district characteristics. These include officer fixed effects and time-varying characteristics, such as district assignment, tenure, and the number of complaints made against the officer. It also includes time-varying district characteristics, such as overall crime rate, violent crime rate, and arrest rate. All time-varying variables except for district assignment are lagged by one month. We also include fixed effects for year and month in τ_t . Standard errors are clustered at the supervisor level.

The parameters of interest are β , which tell us how arrests affect the nomination likelihood for majority-group officers (white, male), and δ , which tells us the differential in likelihood between minority officers and their majority-group colleagues by number of arrests. We expect β 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 majority supervisors do not exhibit in-group bias, then we would expect δ to be zero. A negative δ indicates that majority-group supervisors favor officers in their in-group (or, are biased against minority officers), whereas a positive δ indicates that majority supervisors favor officers not in their in-group.

Figure 1: Effect of Arrests on Nomination Likelihood



Notes: This graph depicts the impact of arrests on the probability of nomination, separately for white supervisors (Panel A) and male supervisors (Panel B). Sample is at the officer-month level. N is 176,552 for Panel A and 204,988 for Panel B. The reference group is officers with no arrests. Estimates control for officer fixed effects, district, tenure, complaints, overall crime rate, violent crime rate, and overall arrest rate, as well as year and month fixed effects. Standard errors are clustered at the supervisor level. Wings depict 95% confidence intervals.

Figure 1a presents estimates from equation (3) when in-group is defined by race.³⁷ It illustrates how white supervisors' nomination behavior is affected by the officer's arrest record. There are increasing returns to having more arrests, with a marked increase for

³⁷Specifically, we plot estimates for β_1 when the reference group is white officers. Then, we re-estimate equation (3) with black officers as the reference group and plot β_1 on the same graph. Then we do the same with Hispanic officers as the reference group.

those with five or more arrests. The average white officer with zero arrests has about a 1% chance of being nominated by his white supervisor for a departmental award the subsequent month. If the officer made one arrest last month, then the nomination likelihood increases by 0.58 percentage-points (57.5%). This estimate is statistically significant at the 5% level. If the officer made two arrests last month, which is the average number of monthly arrests across all officers, the nomination likelihood increases by 0.88 percentage-points (88%). This estimate is also statistically significant at the 5% level. For three arrests, the nomination likelihood increases by 1.3 percentage-points (127%), and by 1.5 percentage-points (153%) for four arrests. White officers with five or more arrests triple their probability of receiving a nomination (an increase of 3.2 percentage-points) relative to those with zero arrests.

Black officers under white supervisors also see increasing returns to having more arrests, but the return is less. Relative to black officers with zero arrests, the nomination likelihood increases by 0.12 percentage-points (24.5%) if black officers have one arrest last month, 0.39 percentage-points (79.6%) with two arrests, 0.55 percentage-points (90%) with three arrests, 1.16 percentage-points (237%) with four arrests, and 1.68 percentage-points (241%) with five or more arrests.

When compared to the increase in nomination probability for white officers with one arrest, black officers with one arrest are 0.46 percentage-points less likely to be nominated. Put another way, a black officer with the same arrest record, same years of experience, and in the same district as his white colleague, is 79.3% percent ($= -0.0046/0.0058$) less likely to be nominated for a departmental award by his white supervisor. This estimate is statistically significant at the 5% level.

The black-white nomination gap diminishes as the number of arrests increases, but does not go away. For example, a black officer with two or three arrests is 56-57% less likely to be nominated than a white officer with two-to-three arrests. This is statistically significant at the 10% level. For four arrests, it is -24.7% ($= -0.004/0.015$) though not statistically significant. For five or more arrests, it is -46.9% ($= -0.149/0.032$), which is statistically

significant at the 5% 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 47% less likely to nominate a black officer at the 94th percentile compared to a white officer at the 81st percentile, all else equal.

We also examine whether white supervisors are less likely to nominate Hispanic officers, another racial minority in the Chicago Police Department. The returns to an additional arrest is similar to that for white officers, particularly for those with 1-4 arrests. For five or more arrests, Hispanic officers are 2.37 percentage-points (188%) more likely to be nominated, relative to Hispanics with zero arrests. Compare this to white officers with five or more arrests, who are 3.2 percentage-points more likely to be nominated.

The magnitude of the Hispanic-white nomination gap, which ranges from 0.12 to 0.8 percentage-points, is smaller than the black-white gap. Further, while there is a negative Hispanic-white gap for 1 arrest and 4+ arrests, the gap turns positive for 2-3 arrests. Last, none of the estimates are statistically significant. These results suggest that white supervisors do not exhibit similar behavior towards Hispanic officers and black officers, which is consistent with our findings from Table 4.

Figure 1b graphs the same relationship for officers under male supervisors. As with Figure 1a, there are increasing returns to more arrests. The average male officer with zero arrests has a 1.1% probability of being nominated by his male supervisor the following month. This is very close to the likelihood for the average white officer assigned to a white supervisor because almost half of all male officers are white. It is not surprisingly, therefore, that the pattern for majority-group officers is similar between Figure 1a and 1b. A male officer with one arrest is 0.39 percentage-points more likely (35.6%) to be nominated relative to when he had no arrests. This increases to 0.84 percentage-points (75.8%) with 2 arrests, 1.2 percentage-points (110%) with 3 arrests, 1.7 percentage-points (151%) with 4 arrests, and

3.2 percentage-points (290%) with 5 or more arrests. All of these estimates are statistically significant at the 5% level.

When we examine female officers assigned to male supervisors, we find that women are similar to men, particularly for those between 0 and 2 arrests. Starting with three arrests, women are 1.2 percentage-points (240%) more likely to be nominated. This increases to 1.67 percentage-points (334%) for four arrests and 3.2 percentage-points (640%) for five or more arrests. Comparing these estimates to men's, women are 24-27% less likely than their male colleagues to be nominated, though the differences are not statistically significant.³⁸ Focusing on the right-tail of the officer distribution, however, sharpens the picture. A female officer with five or more arrests is at the 96th percentile among women, while a male officer with 5 arrests is at the 81st percentile among men. However, the male supervisor is 25% less likely to nominate the woman than the man, all else equal. This estimate is weakly significant with a p-value of 0.107.

6 Evidence on Mechanisms

In this section, we discuss several theories of discrimination that can explain our main findings. To streamline the analysis, we focus on white supervisors because we see stronger evidence of in-group bias by white supervisors.

6.1 Less Interaction

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 assign-

³⁸We check whether these results mask heterogeneity by race. We do not find strong evidence that either male supervisors or white male supervisors nominate white female officers at different rates from black female officers and Hispanic female officers.

ments. 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, Pallais, and Pariente 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 “interact” with the officer by reviewing his or her personnel 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 A6 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.³⁹

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 interaction channel. If less interaction 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:

$$Nom_{it} = \beta_0 + \sum_{q=-2}^3 \mathbb{1}\{EQ = q\} \delta^q + \left(black_i \times \sum_{q=-2}^3 \mathbb{1}\{EQ = q\} \right) \beta_1^q + X'_{it} \alpha + \tau_t + e_{it} \quad (4)$$

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. $black_i$ is a dummy variable equal to 1 if officer i is black and 0 if white. X contains controls for officer characteristics (e.g., officer fixed

³⁹After 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.

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. Standard errors are clustered at the supervisor level.

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 interaction between officers and supervisors are 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.

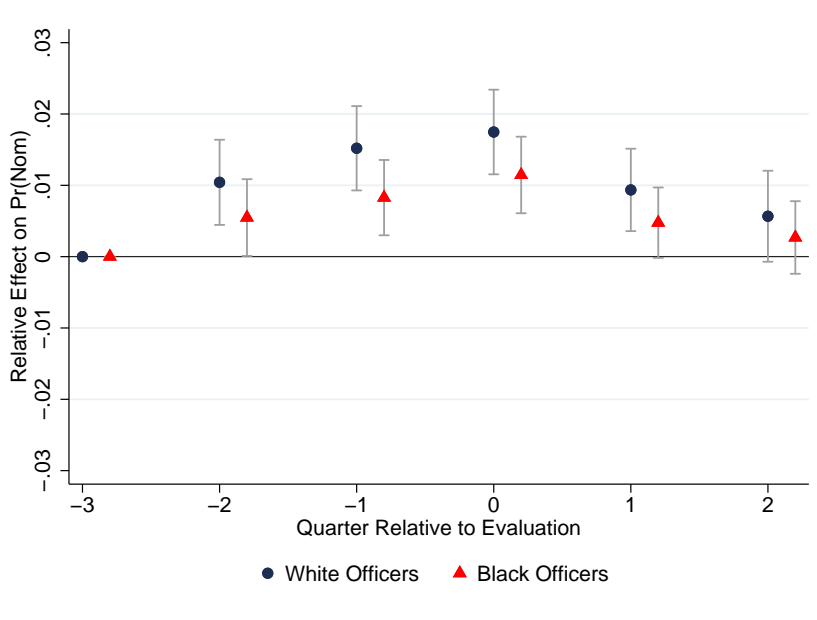
Figure 2 plots our estimates for δ^q .⁴⁰ The hump-backed shape suggests that the nomination probability increases as we grow closer to the evaluation quarter then falls afterwards. This pattern exists for both white and black officers assigned to white supervisors, but the hump is less steep for black officers.

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 percentage increases to 2.9%, then 3.45%, and in their evaluation quarter, white officers have a 3.72% chance of being nominated by their white supervisors. Relative to $EQ = -3$, these numbers equate to an initial increase of 53% (1 percentage-point), an increase of 80% (1.52 percentage-points) in $EQ = -1$, and an increase of nearly 93% (1.75 percentage-points) in $EQ = 0$. All of these estimates are statistically significant at the 1% level.

After their evaluation, white officers are still more likely to be nominated relative to

⁴⁰Specifically, we plot estimates for δ^q when the reference group is white officers. Then, we re-estimate equation (4) with black officers as the reference group and plot δ^q on the same graph.

Figure 2: Probability of Nomination by Quarter



Notes: This graph depicts the impact of arrests on the probability of nomination by white supervisors. Sample is at the officer-month level and restricted to black and white officers assigned to white supervisors. $N = 128,233$. The reference group is officers in $EQ = -3$. Estimates control for officer fixed effects, district, tenure, arrests, complaints, overall crime rate, violent crime rate, and overall arrest rate. Standard errors are clustered at the supervisor level. Wings depict 95% confidence intervals.

$EQ = -3$, but the increase is less. For example, white officers are now only 0.94 percentage-points more likely to be nominated in $EQ = 1$ relative to $EQ = -3$ and 0.57 percentage-points more likely to be nominated in $EQ = 2$ relative to $EQ = -3$. Although these are significant at the 1% level and 10% level, respectively, the nomination likelihood fell by nearly 46% ($= 0.94 / 1.75 - 1$) between the evaluation quarter and the subsequent quarter.

The story for black officers is a slightly different one. Although black officers also see an increase in nomination likelihood leading up to their evaluation quarter, they do not receive the same boost as their white peers. For example, two quarters before their evaluation, black officers are 0.55 percentage-points (68.8%) more likely to be nominated relative to the reference quarter. This is significant at the 5% level. However, this increase is 54.5% (0.57 percentage-points) less than their white peers though it is not significant. In the subsequent quarter leading up to their evaluation ($EQ = -1$), black officers are 0.83 percentage-points

(104%) more likely to be nominated, and this is significant at the 1% level. But again, the increase is 55% (0.84 percentage-points) less than the increase for their white peers, and this black-white difference becomes significant at the 10% level.

In the evaluation quarter ($EQ = 0$), black officers are 1.15 percentage-points (144%) more likely to be nominated relative to the reference quarter, and this is significant at the 1% level. But again, the increase for black officers is 47% (0.82 percentage-points) less than the increase for white officers. This black-white difference is significant at the 10% level.

After the evaluation quarter, relative nomination likelihoods fall for black officers, just like they did for white officers. In $EQ = 1$, black officers are 0.48 percentage-points more likely to be nominated relative to $EQ = -3$, and this is significant at the 10% level. In $EQ = 2$, however, the 0.27 percentage-point estimate is no longer significant. This is not true for white officers, who saw sustained, albeit smaller, benefits after the evaluation quarter.

To summarize, white 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 markedly after the evaluation quarter. For black officers assigned to white supervisors, we see a more stable nomination likelihood with smaller increases leading up to the evaluation quarter. These patterns suggest that early interaction between supervisors and officers is an important mechanism and that it is not sustained after the required evaluation. Importantly, it also suggests that one possible avenue for correcting statistical discrimination, continued evaluations of a given officer, may be closed to law enforcement due to its decentralized nature of supervision (Bohren, Imas, and Rosenberg 2018).⁴¹

⁴¹Bohren, Imas, and Rosenberg 2018 suggest that continuous and sustained positive evaluations may eliminate discrimination that is driven by biased beliefs. This mechanism is most effective in an environment where there is continuity in supervisor-officer relationships. 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.

6.2 Implicit Preferences

Another explanation for why white supervisors are less likely to nominate black officers is that they may simply dislike black officers. We find support for this theory in Figure 1, where we find that white supervisors are less likely to nominate black officers *even conditional on the number of arrests*.⁴² Further, it persists when we focus on officers with 5 or more arrests—those in the right-tail of the distribution. These results suggest that a mechanism other than statistical discrimination is at play.⁴³

At the same time, implicit preferences are difficult to ascertain without a scientific measure. We attempt to detect implicit bias using two supervisor-specific measures: the black-white nomination gap and the black-white TRR (use of force) gap. These measures on their own are not enough to detect individual preferences. For example, a negative black-white nomination gap may be due to animus or statistical discrimination. Likewise, a positive black-white gap in TRR filings—indicating that an officer is more likely to use force against black subjects relative to white subjects—may reflect animus, but it may also reflect unobservable racial differences in subject behavior. For example, if black subjects tend to be more aggressive than white subjects on average, then this may explain why officers use more force with black subjects than white subjects.

In contrast, considering both of these measures, which are constructed from interactions with two different populations (colleagues vs. subjects), may help to identify implicit preferences.⁴⁴ If there is implicit racial bias, then we may expect there to be a negative relationship between a supervisor’s nomination record towards black officers and his use of force on black subjects. That is, a supervisor with implicit bias against blacks would be less likely to nom-

⁴²A negative black-white nomination gap does not necessarily indicate differing racial preferences on the part of the supervisor; it may also be the result of statistical discrimination if supervisors believed that the average black officer is less productive than the average white officer. However, the black-white gap is still negative when we compare officers with the same number of arrests.

⁴³Alternatively, it may be that black officers are less likely to publicize their accomplishments to white supervisors, relative to white officers.

⁴⁴We assume implicit preferences hold across aggregate racial groups. For example, an implicit preference for blacks means a preference for all blacks, not a preference for black subjects but not for black colleagues.

inate a black officer relative to a white officer *and* be more likely to use force against a black subject relative to white subject.

Using TRR data from 2004 to 2015, we calculate the black-white difference in TRR filings for each officer conditional on the officer having filed at least one TRR during this time period.⁴⁵ If an officer filed no TRRs between 2004 and 2015, then the officer is not in our sample.⁴⁶ The black-white nomination gap is calculated as the difference between the number of times the supervisor nominated a black officer and the number of times supervisor nominated a white officer, divided by the total number of assigned officers. Only supervisors who have both black and white assigned officers are included in this exercise.

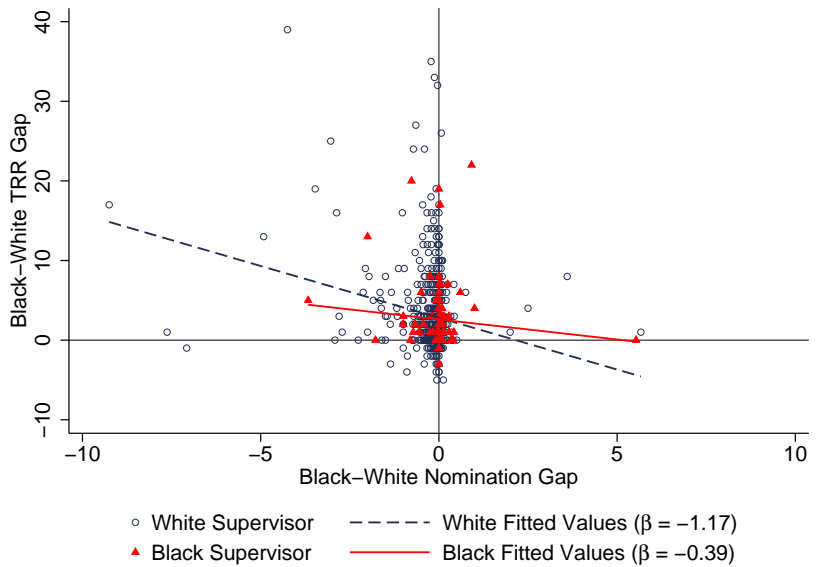
Figure 3 plots these two measures for the 569 supervisors in our analysis sample for whom we have TRR information. For white supervisors, the plot shows a downward sloping line of best fit, indicating that white officers who are less likely to nominate black officers are also more likely to use force against black subjects. The coefficient on the best fit line is -1.17, meaning that a supervisor who nominates all his white officers and none of his black officers also has 1.17 more TRR filings against black subjects, relative to a supervisor who nominates black and white officers at equal rates. This represents a 31% increase in the use of force against black subjects as the average white supervisor in this analysis sample⁴⁷ has a black-white TRR gap of 3.75. This estimate is statistically significant at the 1% level. For black supervisors, the slope coefficient is -0.39 and not statistically significant, suggesting that use of force is independent of their nomination decisions.

⁴⁵The black-white TRR gap is defined as the number of TRR filings where the subject is black less the number of TRR filings where the subject is white. Missings are treated as zeroes. So, for example, if Officer X had filed 4 TRRs where the subject was black but no TRRs where the subject was white, then Officer X's black-white TRR gap is equal to 4.

⁴⁶This is why we were able to match 569 out of the 1,284 supervisors in our analysis sample. Including these officers does not affect our results. We gain an additional 105 white supervisors and an additional 39 black supervisors, and the correlation coefficient changes from -1.17 to -1.30 for white supervisors and remains statistically significant at the 1% level. The correlation coefficient for black supervisors changes from -0.39 to -0.50 and is not significant.

⁴⁷Recall that the sample for this exercise is restricted to supervisors who are assigned to both black and white officers and have filed at least one TRR between 2004 and 2015.

Figure 3: Nominations vs TRR Filings by Supervisor Race



Notes: This graph plots the black-white nomination gap on the x-axis against the black-white TRR (use of force) gap on the y-axis for each supervisor. $N = 569$. The black-white nomination gap is calculated as the difference between the number of times the supervisor nominated a black officer and the number of times supervisor nominated a white officer, divided by the total number of assigned officers. The black-white TRR gap is defined as the number of TRR filings where the subject is black less the number of TRR filings where the subject is white.

6.3 Interpretation

In this section, we examine the extent to which different factors can explain the minority nomination gap. How much of the gap can be explained by officer performance vs. supervisor interaction vs. implicit preferences?

The sample for this exercise is at the officer-month level, restricted to all officers who are assigned to a majority-group supervisor. When focusing on race, the majority supervisor is white. When focusing on gender, the majority supervisor is male. The regression model is a

variant of equation (4) without the interaction between relative quarter and officer race.⁴⁸

$$Nom_{it} = \beta_0 + N_i\gamma + \sum_{q=-2}^3 \mathbb{1}\{EQ = q\}\delta^q + X'_{it}\alpha + \tau_t + e_{it} \quad (5)$$

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. X_{it} is a vector of officer characteristics (e.g., birth year, tenure, complaints, district) and district characteristics (e.g., overall crime rate, violent crime rate, property crime rate, and non-index crime rate).⁴⁹ All time-varying controls except for tenure and district assignment are lagged by one month. τ_t includes fixed effects for year and month. Standard errors are clustered at the supervisor level.

EQ is a set of binary indicator variables for each quarter relative to the evaluation quarter. We use these to proxy for the information channel, and set the reference quarter as the evaluation quarter ($EQ = 0$) since that is the quarter in which supervisors are required to interact with their assigned officers.

N_i is a vector of binary indicator variables indicating the officer's race or sex. When in-group is defined by race (e.g., when the sample is restricted to all officers assigned to white supervisors), N_i contains four binary indicator variables for whether the officer is Asian, black, Hispanic, or Native American. White officers are the reference group. When in-group is defined by sex (e.g., when the sample is restricted to all officers assigned to male supervisors), N_i is a female dummy variable, and male officers are the reference group. The parameters of interest are the coefficient on the black indicator variable, γ^B , and the coefficient on the female indicator variable, γ^F . These will tell us, respectively, how the black-white nomination gap and the female-male nomination gap evolve as we include additional controls.

Table 5 Panel A presents estimates for the black-white nomination gap. We begin with

⁴⁸As a robustness check, we estimate a model where we include interactions and find that the estimates do not change much.

⁴⁹We also include second-order and third-order terms for tenure and overall crime rate.

Table 5: Minority Nomination Gap

Dependent Variable: Officer Nominated for Award				
	(1)	(2)	(3)	(4)
Panel A: Black-White Gap				
Black	-0.0177*** (0.00187)	-0.00789*** (0.00150)	-0.00866*** (0.00208)	-0.00870*** (0.00208)
Observations	176,552	176,552	176,552	176,552
White Officer Mean	0.031	0.031	0.031	0.031
R-squared	0.002	0.035	0.037	0.038
Panel B: Female-Male Gap				
Female	-0.0166*** (0.00142)	-0.00787*** (0.00107)	-0.00742*** (0.00109)	-0.00726*** (0.00109)
Observations	204,988	204,988	204,988	204,988
Male Officer Mean	0.032	0.032	0.032	0.032
R-squared	0.002	0.037	0.038	0.039
Month and Year FE		Yes	Yes	Yes
Officer Characteristics		Yes	Yes	Yes
District Characteristics			Yes	Yes
Evaluation Quarter				Yes

Note: Panel A sample is restricted to all officers assigned to white supervisors, while Panel B sample is restricted to all officers assigned to male supervisors. Officer characteristics include birth year, tenure, complaints, and district assignment. District characteristics include overall crime rate, violent crime rate, and overall arrest rate. Standard errors in parentheses are clustered at the supervisor level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

the unconditional gap in column 1, which states that white supervisors are 1.8 percentage-points (57%) less likely to nominate black officers relative to white officers. In column 2, this estimate reduces by about 45%, from 1.77 to 0.79 percentage-points, when we control for time and officer characteristics. In column 3, we include districts characteristics, such as crime rate, and find that the estimate increases slightly to 0.87 percentage-points. This is because black officers are more likely to be assigned to districts with higher crime rates, and nomination rates are positively correlated with higher crime rates. In column 4, we include the relative quarter in order to control for interactions between supervisors and officers. This does not change the estimate much; the black-white gap remains at around

-0.87 percentage-points. All of the estimates are statistically significant at the 1% level.

Panel B repeats this same exercise for the female-male nomination gap. The unconditional gap is 1.66 percentage-points, or 52%. That is, male supervisors are 52% less likely to nominate a female officer than a male officer. If we control for officer performance, this gap reduces by about 47%. So, even after accounting for the number of arrests and complaints, male supervisors are 25% less likely ($= -0.742/3.2$) to nominate a female officer. These estimates are statistically significant at the 1% level. In column 3, we control for district characteristics, which does not change the estimate very much. In column 4, we control for supervisor interaction, and the estimate, again, changes little (0.73 percentage-points).

To summarize, our findings suggest that officer performance, neighborhood characteristics, and supervisor interaction can explain 44 to 49% of the minority nomination gap.

7 Conclusion

Racial bias has been extensively documented in a variety of settings, including hiring decisions ([Bertrand and Mullainathan 2004](#); [Agan and Starr 2017](#); [Doleac and Hansen 2018](#)), sports umpires ([Parsons et al. 2011](#)), judicial and sentencing decisions ([Park 2017](#); [Flanagan 2018](#); [Mueller-Smith and Schnepel 2017](#); [Rehavi and Starr 2014](#))⁵⁰, and bail decisions ([Arnold, Dobbie, and Yang 2018](#)). The increasing availability of police administrative data has allowed researchers to carefully examine and detect bias in police interactions as well. A potential solution that has been put forth is to increase racial and gender diversity among officers, which are traditionally homogenous.

There are three mechanisms through which a diverse police force may improve policing quality ([Sklansky 2005](#)). First, minority officers may have a unique set of skills that complement their majority-group colleagues (“competency effect”). Second, demographic diversity of a police department may impact its relationship with the community it serves (“commu-

⁵⁰[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 reoffending rates and improves labor market outcomes among young black men charged with misdemeanors.

nity effect”). Past studies have touched upon the first two channels. For example, [Miller and Segal 2018](#) find that the integration of women in the US police force led to increased reports of violent crime and domestic violence against women, but had no impact on reporting for male victims. Relatedly, [Anwar, Bayer, and Hjalmarsson 2012](#) finds that the inclusion of women on jury panels led to an increase in convictions for sex offenses. [McCrary 2007](#) finds that court-ordered racial hiring quotas for police departments during the 1970s were targeted at big cities with growing black populations.

This paper sheds light on the third mechanism, or the extent to which demographic diversity may affect the internal dynamics of the department (“organizational effect”). This question is of particular relevance to law enforcement, where minorities are less represented at higher ranks. For example, white males comprised 40% of all entry-level police officers in the Chicago Police Department in 2015, and 56% of those at the rank of Sergeant or higher.⁵¹ Further, 98% of CPD officers believe that promotions are due to connections not merit ([Police Accountability Task Force 2016](#)). This is likely compounded by the fact that the two most recent Sergeant promotion exams were nearly ten years apart, thus limiting advancement opportunities for many officers.

To examine the extent of in-group 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. We find weaker evidence that male supervisors are less likely to nominate female officers than male officers. These results control for officer characteristics, including experience, arrest record, and misconduct. Further, we find a significant, negative black-white nomination gap even among officers who are the right-tail of the arrests distribution.

We explore several theories of discrimination that can explain our main findings. We find that interaction between supervisors and officers is important, but that black officers benefit

⁵¹These numbers do not include civilian Department members.

less from this interaction with white supervisors relative to their white peers. Further, we find suggestive evidence of implicit preferences towards in-group members; white supervisors who are less likely to nominate black officers are also more likely to use force against black subjects relative to white subjects. Taken together, these findings suggest that simply increasing the diversity of incoming recruits may not be enough. Rather, police departments should also focus on policies that address in-group bias due to its effect on career advancement.

References

- Agan, Amanda, and Sonja Starr. 2017. "Ban the box, criminal records, and racial discrimination: A field experiment." *The Quarterly Journal of Economics* 133 (1): 191–235.
- Antonovics, Kate, and Brian G Knight. 2009. "A new look at racial profiling: Evidence from the Boston Police Department." *The Review of Economics and Statistics* 91 (1): 163–177.
- Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson. 2012. "The impact of jury race in criminal trials." *The Quarterly Journal of Economics* 127 (2): 1017–1055.
- Anwar, Shamena, and Hanming Fang. 2006. "An alternative test of racial prejudice in motor vehicle searches: Theory and evidence." *American Economic Review* 96 (1): 127–151.
- Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial bias in bail decisions." *The Quarterly Journal of Economics* 133 (4): 1885–1932.
- Beaman, Lori, Niall Keleher, and Jeremy Magruder. 2018. "Do job networks disadvantage women? Evidence from a recruitment experiment in Malawi." *Journal of Labor Economics* 36 (1): 121–157.
- Bertrand, Marianne, and Sendhil 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 Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope. 2019. "Inaccurate Statistical Discrimination." Technical Report, National Bureau of Economic Research.
- Bohren, J Aislinn, Alex Imas, and Michael Rosenberg. 2018. "The dynamics of discrimination: Theory and evidence."
- Bulman, George. 2019. "Law Enforcement Leaders and the Racial Composition of Arrests." *Economic Inquiry*.
- Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Review* 104 (9): 2593–2632.
- Coviello, Decio, and Nicola Persico. 2015. "An Economic Analysis of Black-White Disparities in the New York Police Departments Stop-and-Frisk Program." *The Journal of Legal Studies* 44 (2): 315–360.
- Doleac, Jennifer L, and Benjamin 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*, vol. Forthcoming.
- Donohue III, John J, and Steven D Levitt. 2001. "The impact of race on policing and arrests." *The Journal of Law and Economics* 44 (2): 367–394.
- Egan, Mark L, Gregor Matvos, and Amit Seru. 2018. "When Harry fired Sally: The Double Standard in Punishing Misconduct." Technical Report, National Bureau of Economic Research.

- Flanagan, Francis X. 2018. "Race, gender, and juries: Evidence from north carolina." *The Journal of Law and Economics* 61 (2): 189–214.
- Fryer, Roland G. 2018. "An Empirical Analysis of Racial Differences in Police Use of Force." *Journal of Political Economy*, vol. Forthcoming.
- Giuliano, Laura, David I Levine, and Jonathan Leonard. 2009. "Manager race and the race of new hires." *Journal of Labor Economics* 27 (4): 589–631.
- Glover, Dylan, Amanda Pallais, and William Pariente. 2017. "Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores." *The Quarterly Journal of Economics* 132 (3): 1219–1260.
- Goncalves, Felipe, and Steven Mello. 2018. "A Few Bad Apples?: Racial Bias in Policing." Technical Report.
- Hengel, Erin. 2019. "Publishing while Female. Are women held to higher standards? Evidence from peer review."
- Hoekstra, Mark, and CarlyWill Sloan. 2019. "The Effect of Police Officer Race on Use of Force." Technical Report.
- Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial bias in motor vehicle searches: Theory and evidence." *Journal of Political Economy* 109 (1): 203–229.
- Knox, Dean, Will Lowe, and Jonathan Mummolo. 2019. "The Bias is Built In: How Administrative Records Mask Racially Biased Policing." Technical Report.
- Kofoed, Michael, and Elizabeth mcGovney. 2019. "The Effect of Same-Gender or Same-Race Role Models on Occupation Choice." *Journal of Human Resources*, vol. Forthcoming.
- Langan, Andrew. 2018. "Female Managers and Gender Disparities: The Case of Academic Department Chairs." Technical Report.
- McCrary, Justin. 2007. "The Effect of Court-ordered Hiring Quotas on the Composition and Quality of Police." *American Economic Review* 97 (1): 318–353.
- Miller, Amalia, and Carmit Segal. 2018. "Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence." *Review of Economic Studies*, vol. Accepted.
- Mueller-Smith, Michael, and Kevin Schnepel. 2017. "Diversion in the criminal justice system: Regression discontinuity evidence on court deferrals." Technical Report, Working paper.
- Nix, Justin, Bradley A Campbell, Edward H Byers, and Geoffrey P Alpert. 2017. "A Bird's Eye View of Civilians Killed by Police in 2015." *Criminology and Public Policy* 16 (1): 309–340.
- Park, Kyung H. 2017. "The impact of judicial elections in the sentencing of black crime." *Journal of Human Resources* 52 (4): 998–1031.
- Parsons, Christopher A, Johan Sulaeman, Michael C Yates, and Daniel S Hamermesh. 2011. "Strike three: Discrimination, incentives, and evaluation." *American Economic Review* 101 (4): 1410–35.

- Persico, Nicola. 2002. "Racial profiling, fairness, and effectiveness of policing." *American Economic Review* 92 (5): 1472–1497.
- Police Accountability Task Force. April 2016. "Recommendations for Reform: Restoring Trust between the Chicago Police and the Communities they Serve." Technical Report.
- Rehavi, M Marit, and Sonja B Starr. 2014. "Racial disparity in federal criminal sentences." *Journal of Political Economy* 122 (6): 1320–1354.
- Rim, Nayoung, Bocar Ba, and Roman Rivera. 2019. "Community Ties and Police Use of Force." Technical Report.
- Sarsons, Heather. 2018. "Interpreting signals in the labor market: evidence from medical referrals." Technical Report.
- . 2019. "Gender differences in recognition for group work." *Journal of Political Economy*, vol. Accepted.
- Sklansky, David Alan. 2005. "Not Your Father's Police Department: Making sense of the new demographics of law enforcement." *J. Crim. L. & Criminology* 96 (3): 1209–1243.
- . 2011. "The persistent pull of police professionalism." *New Perspectives in Policing, March*.
- United States Department of Justice. March 4, 2015. "Investigation of the Ferguson Police Department." Technical Report, Civil Rights Office.
- Weisburst, Emily. 2018. "Whose Help Is on the Way? The Importance of Individual Police Officers in Law Enforcement." Technical Report.
- West, Jeremy. 2018. "Racial Bias in Police Investigations." Technical Report.
- Worden, Robert E, Moonsun Kim, Christopher J Harris, Mary Anne Pratte, Shelagh E Dorn, and Shelley S Hyland. 2013. "Intervention with problem officers: An outcome evaluation of an EIS intervention." *Criminal justice and behavior* 40 (4): 409–437.

Table A1: Department Awards

1	<i>Superintendent's Award of Valor</i> 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	<i>Superintendent's Award of Merit</i> 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	<i>Police Blue Star Award</i> is granted to any sworn member who has been seriously, critically, or fatally injured while in the performance of police duty.
4	<i>Police Blue Shield Award</i> 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	<i>Superintendent's Award of Tactical Excellence</i> 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	<i>Arnold Mireles Special Partnership Award</i> for making a significant impact upon the quality of life within their community by identifying and resolving problems.
7	<i>Special Commendation</i> for making a significant impact on public safety or crime prevention.
8	<i>Lifesaving Award</i> for a successful effort in saving a human life that involved exceptional courage or performance.
9	<i>Police Officer of the Month</i> for performance of duty during a specific month was characterized by such exceptional professional skill that it merits recognition by the entire Department.
10	<i>Chicago Police Leadership Award</i> for exemplary service, dedication, and leadership.
11	<i>Department Commendation</i> 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	<i>Problem Solving Award</i> for an exemplary effort to identify, analyze, and successfully respond to causes, conditions, and problems that may lead to crime and neighborhood disorder.
13	<i>Joint Operations Award</i> 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	<i>Unit Meritorious Performance Award</i> for exhibiting exceptional professional skill and conduct during a coordinated action.
15	<i>Traffic Stop of the Month Award</i> for excellence in conducting professional traffic stops that result in quality arrests.
16	<i>Top Gun Arrest Award</i> for exceptional commitment to the recovery of illegal firearms.
17	<i>Special Service Award</i> for contributing to any event that has a significant impact upon the historical direction and operations of the Department.
18	<i>Honorable Mention Certificate</i> 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
<i>Force Mitigation Efforts</i> Member Presence Zone of Safety Verbal Direction/Control Techniques Movement to Avoid Attack Specialized Units Tactical Positioning Additional Unit Members None Other	Mitigation
<i>Control Tactics</i> Escort Holds Wristlock Armbar Control Instrument Pressure Sensitive Areas Emergency Handcuffing Other	Control tactics
<i>Response without Weapons</i> Open Hand Strike Take down Elbow strike Close hand strike/Punch Knee strike Kicks Other	No Weapon
<i>Response with Weapons</i> 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: Impact of Officer Characteristics on Supervisor Race and Sex

Dependent Variable: Officer Sample:	Supervisor is White			Supervisor is Male	
	White (1)	Black (2)	Hispanic (3)	Male (4)	Female (5)
L_arrest_violent	0.00505 (0.0154)	-0.0352 (0.0280)	0.0200 (0.0197)	-0.0151* (0.00798)	0.0276 (0.0212)
L_arrest_violent2	-0.00197 (0.00769)	0.0102 (0.0143)	-7.55e-05 (0.00909)	0.00390 (0.00355)	-0.0145 (0.00936)
L_arrest_violent3	0.000266 (0.000950)	-0.000809 (0.00181)	-0.000739 (0.00105)	-0.000359 (0.000402)	0.00140 (0.00102)
L_arrest_property	0.0236 (0.00993)	0.0280 (0.0336)	8.20e-05 (0.0219)	-0.000754 (0.00904)	0.00262 (0.0164)
L_arrest_property2	-0.00275 (0.00217)	-0.00740 (0.0186)	0.00342 (0.0104)	-0.00114 (0.00376)	8.87e-05 (0.00407)
L_arrest_property3	9.63e-05 (7.91e-05)	0.000960 (0.00202)	-4.88e-05 (0.00101)	0.000106 (0.000299)	-6.90e-05 (0.000152)
L_arrest_nonfbi_other	-0.000457 (0.00577)	-0.00268 (0.00976)	-0.00226 (0.00749)	0.0135 (0.00326)	0.00576 (0.00873)
L_arrest_nonfbi_other2	-0.000494 (0.000519)	8.70e-05 (0.000720)	-0.000189 (0.000660)	-0.000655 (0.000227)	-0.000256 (0.000958)
L_arrest_nonfbi_other3	1.45e-05 (1.06e-05)	-2.96e-06 (1.29e-05)	4.48e-06 (1.41e-05)	9.84e-06 (4.04e-06)	7.27e-06 (2.41e-05)
L_cmpl	-0.00767 (0.0512)	-0.0785 (0.175)	-0.115 (0.137)	0.00958 (0.0371)	-0.0584 (0.0905)
L_cmpl2	-0.00206 (0.0573)	0.0939 (0.222)	0.130 (0.184)	-0.00757 (0.0434)	0.0709 (0.0869)
L_cmpl3	0.00510 (0.0131)	-0.0180 (0.0609)	-0.0369 (0.0534)	0.000312 (0.0106)	-0.00950 (0.0173)
L_trr_new	-0.0442 (0.111)	-0.125 (0.238)	0.133 (0.183)	-0.149 (0.0685)	-0.0171 (0.178)
L_trr_new2	0.0165 (0.0594)	0.0878 (0.137)	-0.0865 (0.0893)	0.0774 (0.0364)	-0.00381 (0.109)
L_trr_new3	-0.00147 (0.00972)	-0.0139 (0.0249)	0.0148 (0.0123)	-0.0114 (0.00571)	0.00327 (0.0190)
L_strongforceratio	0.0317 (0.0484)	0.0558 (0.0828)	0.108 (0.0535)	0.0423 (0.0351)	0.0349 (0.0961)
L_strongforceratio2	-0.0777 (0.0604)	-0.0651 (0.108)	-0.123 (0.0574)	-0.0109 (0.0469)	-0.0583 (0.128)
L_strongforceratio3	0.0237 (0.0174)	0.0169 (0.0332)	0.0363 (0.0149)	-0.00340 (0.0145)	0.0136 (0.0401)
L_weakforceratio	0.248 (0.183)	0.447 (0.323)	-0.0479 (0.236)	0.260 (0.116)	-0.239 (0.350)
L_weakforceratio2	-0.360 (0.250)	-0.717 (0.454)	0.137 (0.312)	-0.307 (0.156)	0.379 (0.505)
L_weakforceratio3	0.119 (0.0814)	0.245 (0.149)	-0.0521 (0.102)	0.0946 (0.0506)	-0.129 (0.167)
Constant	0.976 (0.167)	0.625 (0.380)	1.017 (0.157)	0.521 (0.224)	1.216 (0.148)
Observations	9,175	2,923	4,636	15,080	2,337
R-squared	0.121	0.178	0.164	0.060	0.138

Notes: Only estimates for officer performance measures are shown. All variables except tenure are lagged by one month. Non-index crimes are composed of municipal code violations, traffic violations, warrants, drug-related crimes, and other. Standard errors are in parentheses and clustered at the supervisor level.

Table A4: Impact of Officer Characteristics on Supervisor Switch Type

Dependent Variable:	Non-White to White Sup Switch			Female to Male Sup Switch	
	White (1)	Black (2)	Hispanic (3)	Male (4)	Female (5)
Arrests					
Violent	0.0311	1.0000	0.1228	1.0000	1.0000
Violent ²	0.0082	1.0000	0.2876	0.9622	1.0000
Violent ³	0.0000	1.0000	0.5240	1.0000	1.0000
Property	1.0000	1.0000	1.0000	1.0000	1.0000
Property ²	1.0000	1.0000	1.0000	1.0000	1.0000
Property ³	1.0000	1.0000	1.0000	1.0000	0.2941
Non-Index	0.0341	1.0000	1.0000	0.0000	1.0000
Non-Index ²	0.5711	0.9554	1.0000	1.0000	1.0000
Non-Index ³	1.0000	1.0000	1.0000	1.0000	1.0000
Complaints	1.0000	1.0000	1.0000	1.0000	1.0000
Complaints ²	1.0000	1.0000	1.0000	1.0000	1.0000
Complaints ³	1.0000	1.0000	1.0000	1.0000	1.0000
TRR Filings	1.0000	1.0000	1.0000	1.0000	1.0000
TRR Filings ²	1.0000	1.0000	0.9437	1.0000	1.0000
TRR Filings ³	1.0000	1.0000	1.0000	1.0000	1.0000
Strong Force Ratio	0.5545	1.0000	0.1645	1.0000	1.0000
Strong Force Ratio ²	0.0893	1.0000	0.3802	1.0000	1.0000
Strong Force Ratio ³	0.0286	1.0000	0.5763	1.0000	1.0000
Weak Force Ratio	1.0000	1.0000	0.7440	1.0000	0.9739
Weak Force Ratio ²	1.0000	1.0000	1.0000	1.0000	1.0000
Weak Force Ratio ³	0.7470	1.0000	1.0000	1.0000	1.0000
Observations	12,008	6,602	5,970	18,833	6,641
Controls for:					
Race/Ethnicity				Yes	Yes
Sex	Yes	Yes	Yes		
Birth year	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes
Same District	Yes	Yes	Yes	Yes	Yes
Number of months	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes

Notes: This table reports Holm-adjusted p-values for coefficients on officer performance measures. Standard errors are clustered at the supervisor level. Performance measures are lagged by one month. Non-index crimes are composed of municipal code violations, traffic violations, warrants, drug-related crimes, and other.

Table A5: Effect of Minority Supervisor Switch on Nomination Likelihood

Dependent Variable: Officer Nominated for Award		
Switch Type:	White → Non-White (1)	Male → Female (2)
Post-Switch	-0.0447** (0.0193)	-0.0211 (0.0189)
Black Officer x Post-Switch	0.0434* (0.0241)	
Hispanic Officer x Post-Switch	0.0391 (0.0331)	
Female Officer x Post-Switch		-0.00222 (0.0222)
Observations	4,640	3,680
Reference Group Mean	0.175	0.179
R-squared	0.498	0.520
Controls for:		
Officer Characteristics	Yes	Yes
Unit Characteristics	Yes	Yes

Notes: Sample is at the officer-switch level. The reference group in column 1 is white officers and male officers in column 2. Officer characteristics include officer tenure at the start of switch, total complaints, and total arrests. District characteristics include the average crime rate, average violent crime rate, and average arrest rate. Standard errors are clustered at the supervisor level and are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6: 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>