A More Diverse Police Academy Improves Job Performance

The Effect of Minority Peers on Future Arrest Quantity and Quality

Roman Rivera*

Columbia University

Abstract

I exploit the fact that entrance into the Chicago Police Department police academy is determined by randomly assigned lottery numbers to identify the effect of peer racial diversity on officers' future job performance. Using officer-level data on shifts, arrests, and court outcomes of arrests, I construct metrics for individual officer arrest quantity and quality. I find being randomly assigned to an academy cohort with a 10 percent higher share of minorities decreases an officer's future arrests of Blacks by 3-11%. This is driven by a decline in arrests for low-level crimes and is associated with an increase in arrest quality.

*I am grateful to Sandra Black and Bentley MacLeod for guidance and advice. For feedback and comments, I thank Amani Abou Harb, Douglas Almond, Bocar Ba, Michael Best, Felipe Goncalves, Jenny Jiao, Dean Knox, Jonathan Mummolo, Brendan O'Flaherty, Nayoung Rim, Rajiv Sethi, Miguel Urquiola, and Emily Weisburst. I would like to thank Emma Herman, the Invisible Institute, and Sam Stecklow for their contributions to this data set and Rachel Ryley for sharing assignment data. For detailed explanation of the Cook County Court system, I thank Ali Ammoura. Email: r.g.rivera@columbia.edu

1 Introduction

There is substantial evidence that peer racial and ethnic diversity affects individual outcomes in a range of environments, such as classrooms, juries, and dorm rooms.[1] Given the salience of race in policing, the effect of diverse peers on officer behavior is particularly relevant but understudied in economics. As policymakers push for police department to increase diversity in order to better reflect the communities they serve, new officers will be trained in increasingly diverse police academies. Increased academy diversity may lead to changes in officer behavior by causing more interracial friendships to form, reducing officer prejudice, or by causing instructors to train recruits to be more sympathetic to minority civilians. In this paper, I document the long-run effect of academy diversity on officers' enforcement behavior.

I estimate the effect of racially diverse peers in an officer's police academy cohort on their arrest quantity and quality after they exit the academy using data from the second largest police department in the US. The identification of such peer effects is possible because the Chicago Police Department (CPD) assigns new recruits to academy cohorts based on lottery numbers, meaning I observe officers in cohorts not of their own choosing and with varying racial compositions. Officers in Chicago spend 6 months in the academy together, making this treatment significantly more intensive than most diversity-focused interventions. I use a detailed data set on thousands of officers' daily assignments, their arrests, and, using court records linked to arrests, the outcomes of their arrests in court. With these data, I recover officers' individual propensities to make arrests of varying types and qualities, and can identify the effect of academy diversity on these propensities.

A key feature of this paper is my ability to identify a potential mechanism underlying the relationship between officer diversity and policing outcomes, in contrast with previous studies.[2] For example, minority officers may additively change the aggregate behavior of the department, alter how civilians view the department and thus respond to police officers, or cause their peers behave differently with civilians and suspects. In this paper, I isolate the effect of peer diversity on officer behavior, as civilians cannot observe the diversity of a specific officer's cohort after the cohort members have been assigned across the department. Furthermore, in contrast with the literature focused on how officer race is related to differences in enforcement activity,[3] this study identifies peer diversity as a policy-relevant and

[1] See Boisjoly et al. (2006), Sommers (2006), Merlino, Steinhardt, and Wren-Lewis (2019), and Carrell, Hoekstra, and West (2019).

[2] See Donohue III and Levitt (2001), Miller and Segal (2018), McCrary (2007), Garner, Harvey, and Johnson (2019), and Harvey and Mattia (2019). Also, Linos (2018) provides experimental evidence on how to improve diverse recruitment, and Rim et al. (2019) studies racial discrimination in awards which can impact promotion.

[3] See Close and Mason (2007), West (2018), Goncalves and Mello (2018), Weisburst (2020), Hoekstra and

manipulable determinant of differences in officer behavior.

I first document that increased cohort diversity in the police academy reduces a new officer's average arrests of Blacks, but this effect is in part due to the diversity of one's cohort influencing where they are assigned to work. By extending my sample of cohorts and employing a panel structure to control for unit assignment, I provide evidence for the peer effect of diversity driving the results and the assignment effect being relatively small. Then, in order to study the effect of diversity on officer heterogeneity and overcome any assignment effects, I recover individual officer propensities for making arrests of various types and qualities using highly granular daily shift level data. Consistent with the previous results, cohort diversity, specifically cohort shares of Blacks and Hispanics, reduces officers' future propensities to arrest Blacks. In my preferred specification, a 10 percentage point increase in cohort share of minorities decreases officers' propensities to arrest Blacks by 0.22 standard deviations. This decrease is driven entirely by a reduction in officers' propensities to make arrests of Blacks for low-level crimes. Arrest quality is also impacted: more minority peers cause larger declines in low-quality (not found guilty) arrests than high-quality (found guilty) arrests for low-level crimes, resulting in an increase in average arrest quality. These results are persistent across a variety of tests, including those for alternate specifications, assignment sorting, and cohort timing.

Taken together, these results show that diversity in the police academy changes how officers police later in their careers, causing them to make fewer low-level arrests of Blacks. This effect is not due to assignment type or entrance timing being influenced by cohort diversity, and the effect is present across almost 10 years of cohorts (over 2,000 officers) with a wide range of cohort diversity. I also find that white and minority officers are similarly affected; this is not consistent with more diversity increasing interracial contact and reducing white officers' prejudice. Rather, it suggests that cohort diversity influences how all officers learn to police. Furthermore, the reduction in arrest quantity and increase in average arrest quality is likely due to improved policing and discretion, not simply reduced effort, as I find that cohort diversity does not negatively impact arrests for serious crimes.

This paper contributes to two literatures. Most centrally, this paper adds to and combines strands of the peer effects literature in economics. First, the literature on the effect of racially diverse peers has found that increased interracial socialization alters perceptions of minorities (Boisjoly et al. (2006)) and whites' openness to future contact with minorities (Carrell, Hoekstra, and West (2019)).[4] In more task-oriented environments, such as juries,

Sloan (2020), and Ba et al. (2021).

[4]See also Pettigrew (1998), Laar et al. (2005), Pettigrew and Tropp (2006), Baker, Mayer, and Puller (2011), Burns, Corno, and Ferrara (2015), and Merlino, Steinhardt, and Wren-Lewis (2019)

the presence of Blacks jurors changes how whites discuss Black defendants and improves the quality and breadth of information communicated (Sommers (2006)). Second, long-run peer effects are a growing focus of educational studies, often finding that classmate characteristics, such as gender (Black, Devereux, and Salvanes (2013)) or immigration status (Gould, Lavy, and Paserman (2009)), influence future educational or economic outcomes. Peer behavior, in particular, is crucial: Carrell, Hoekstra, and Kuka (2018) find that students from homes linked to domestic violence reduce the long-run earnings of their peers. Similarly, Lavy and Schlosser (2011) find that girls improve contemporaneous educational attainment in classrooms by reducing male misbehavior and thus improving teaching.

In the present study, I contribute to the long-run peer effects literature by documenting persistent effects of cohort diversity on officers' policing outcomes after they exit the academy. As I focus on the effects of cohort diversity, I advance the peer effects of diversity literature by employing a common identification strategy (random assignment of students to classes or grades)[5] in a new setting, policing. Identifying long-run peer effects of diversity in policing also advances the peer effects on workplace performance literature (Guryan, Kroft, and Notowidigdo (2009), Mas and Moretti (2009)). My results indicate that whites and non-whites are similarly affected by minority peers, consistent with the literature on how peers influence learning environments, instructor behavior, and communication, thereby altering the outcomes of all group members.

This research also contributes to the literature on the effect of departmental diversity on arrests and crime (Donohue III and Levitt (2001), McCrary (2007), and Garner, Harvey, and Johnson (2019)). Only recently have researchers been able to study how civilians respond to departmental diversity (Miller and Segal (2018), Harvey and Mattia (2019)) or how officers of different racial groups arrest, stop, use force, and ticket differently (West (2018), Weisburst (2020), Hoekstra and Sloan (2020), Goncalves and Mello (2018), and Ba et al. (2021)). I advance this literature by isolating peer diversity in the academy as a causal determinant of officer behavior on the job, distinguishable from civilian perceptions and officer characteristics. As police academy diversity can be influenced by policymakers, this paper also contributes to the new literature on policing interventions. For example, Owens et al. (2018) finds that officers assigned to procedural justice meetings were 12% less likely to make an arrest; by comparison, I find a similar reduction can be achieved through a 10 percentage point increase in minority peers in the police academy.

As I focus on individual arrest quantity and quality, this paper also builds on Weisburst (2020), which documents large differences in Dallas police officers' individual propensities

[5]See Hoxby (2000) and Sacerdote (2001). Angrist (2014) discusses various studies in the educational peer effects literature.

to make arrests following 911 calls and also employs court outcomes to explore how officer propensities relate to arrest quality. Due to the rarity of arrests resulting from 911 calls, the relationship between arrest propensity and arrest quality cannot directly estimated. By contrast, my data allows me to directly estimate officers' arrest propensities across arrestee race, crime type, and court outcome—e.g., an officer's propensity to arrest Blacks for index crimes which are not found guilty in court. The inclusion of arrest quality is particularly important as it provides a rough metric for arrests which wastefully divert public resources and unjustly damage private ones.

This paper proceeds as follows. In Section 2, I describe the background and data for this paper. In Section 3, I present results for the effect of diversity on arrests. In Section 4, I present results for the effect of diversity on officer heterogeneity, and Section 5 contains related robustness checks. I discuss potential mechanisms in Section 6, and Section 7 concludes.

2 Background and Data

2.1 Chicago Police Department and Recruitment

Comprised of over 10,000 officers, the Chicago Police Department (CPD) is the second largest police force in the US. It polices the nation's third largest city, which is racially diverse and economically segregated. To recruit new officers, the CPD issues a call for officers, and applicants take a written exam, which they must pass in order to enter the academy. As this form ("CPD 2017 FAQ" 2017) explains:

> All applicants who pass the exam are placed on an eligibility list based on a randomly assigned lottery number. You will be referred to the Chicago Police Department in lottery order as vacancies become available.

After an applicant's number is called, and if they pass required physical and psychological tests, they are permitted to start at the police academy (see Appendix A.1 for more discussion). Academy start dates, "appointed dates", correspond to officers beginning their time at the police academy. In Appendix A.2, I provide empirical support for the random assignment of officers to cohorts. I define a cohort as the group of officers with the same appointed date—in the main sample, cohorts are separated by about 1 month. During the academy, officers must complete 900 hours (about 6 months) of training[6] in multiple

[6]This encompasses and surpasses the training required to pass the Illinois State Peace Officer's Certification Exam.

areas, such as “firearms, control tactics, physical training, [and] classroom training” (“Education and Training Division (ETD) Chicago Police Department” 2020). Notably, recruits may be asked to speak about their previous experiences with police and participate in role-playing/scenario-based exercises.

After the academy, the recruits in a cohort enter an on-the-job-training period for one year as “probationary police officers” where they work in multiple areas of the city and are evaluated under the supervision of a Field Training Officer. After meeting the various requirements, completing their time as a probationary officer, and becoming “field qualified” (“Field Training and Evaluation Program” 2018), a recruit exits their probationary period and becomes a full (sworn) Chicago police officer. New sworn officers are then assigned to more permanent units.

Transferring between assignments and filling vacancies is determined by a seniority-based bidding process and is only eligible for non-probationary sworn officers, meaning new officers have little to no choice in where and when they work (“Personnel Transfer and Assignment Procedures – (FOP)” 2011). New officers are generally assigned to units 1-25 which correspond to geographical districts in Chicago.[7] These units occupy most CPD officers and correspond to what is commonly considered police work. There are many other units for specialized work which contain far fewer and more experienced officers, such as training units, detective units, etc., which are not studied in this paper.

2.2 Data

The data for this study come from the Chicago Police Department, Chicago’s Department of Human Resources, and the Circuit Court of Cook County. By combining data sets on CPD officers obtained over four years, I construct a detailed panel data set on officer assignments, arrests, and arrest outcomes in court between 2010 and 2016. This contains officers’ demographic information (race,[8] gender, birth year), start dates, when officers exited the training unit (after the academy and probationary period), and other administrative information. Daily assignment and attendance data, from 2010-2016, includes daily records for whether an officer was present for duty, their assigned unit, beat, shift, and car number (if applicable), as well as start and end times. Collectively, this data permits highly granular analysis of an officer’s working environment. I restrict my analysis to observations of police officers (the lowest and most common rank, i.e. not detectives or sergeants, etc.) working on shift numbers 1 through 4, and assigned to either a car or foot or bike patrol (excluding

[7] During 2012, three units/districts were collapsed into other districts, reducing the total number to 22.

[8] The CPD’s demographic data often combines race and ethnicity into a single variable. For expositional purposes and due to the data used, I will refer to ‘Hispanic’ as a distinct racial group.

desk duty).

In order to recover individual officer arrest quantity and quality metrics, I use arrest data from 2010-2016 and court data. The arrest data contains all arrest of adults by Chicago police officers including arrest date and time, crime description, primary arresting officer(s),[9] and arrestee race. By connecting the arrest data to court records, I construct a metric for arrest quality by determining if the arrest was associated with any guilty finding indicating high quality and no guilty finding indicating low quality. Guilty findings include plea deals, which account for over 90% of convictions. Combined, these data allow me to construct a measure for individual officers' arrest quantity and quality after extensively controlling for their working environments.

2.2.1 Sample Selection

A total of 3,146 recruits joined the CPD between January of 2006 and February of 2015. As defined above, an academy cohort is all the recruits who started at the CPD academy on the same date, resulting in 96 cohorts during this period. I focus my primary analyses on the cohorts starting between July of 2012 and May of 2014 (the "Main Sample") because I can observe their assignments and arrests from their probationary periods onward, these cohorts originated from the same entrance exam issued in December of 2010 (see Appendix A.1), and new officers have almost no choice in assignments.[10] In Sections 3.3 and 5.4, I include additional cohorts (the extended sample) as a robustness check and find similar results, though the precise entrance exam that each of these cohorts took is not certain.

Both the main and the extended sample of officers were subject to a series of filters.[11] Notably, I drop recruits in cohorts who were not matched in the assignment data, recruits with invalid durations in the academy or probationary period, and recruits not matched in the salary and unit assignment data. I also drop a few recruits for whom no fixed effects were able to be recovered during their time as full-officers and, for the main sample, those that had fewer than 15 observations in the assignment panel. So, attrition from the initial cohort to the final sample can occur for multiple reasons. If attrition is related to cohort diversity, it may contaminate the results, but, as I show in Appendix A.3, cohort diversity

[9] Almost all arrests have at most two primary officers listed.

[10] Additionally there is over a half-year gap between the first cohort in 2012 and the last cohort in 2011, and all following cohorts are separated by about one month; and, the period 2012-2014 was without any major scandals and no superintendent changes for the CPD.

[11] As previously discussed, the academy requires around 6 months of training before recruits start working (as probationary officers) then another year before they become full officers. Due to this 18 month lag between beginning the academy and working as a full officer, I ensure that I can observe the officers for at least 6 months by ending my sample at the cohort starting on May 27, 2014. I exclude 9 recruits who started during the sample period but were in cohorts with less than 3 recruits.

has no significant impact on attrition for the main sample. After filters, the main sample of cohorts contain 962 new officers (initially 1,139 recruits) in 21 cohorts with 322,729 total officer-shift observations over 37 months.

2.3 Summary Statistics

2.3.1 Cohort Composition

Table 1 displays the demographic composition of the main sample of recruits (Column 1), the average composition of the main sample cohorts before attrition (Column 2), and the demographics of all officers in the panel data (Column 3). By comparing Columns (1) and (2), it is apparent that the sample of recruits is very similar to that of the average cohort, which is expected due to the random assignment of recruits to cohorts, and that attrition after entering the academy did not significantly alter the demographic composition of the pool of officers. The comparison between all officer demographics (Column 3) and recruit demographics (Column 1) illustrates the changing nature of the police department in Chicago. More recent recruits are less likely to be female (19% vs. 24%). While minorities make up roughly half of both groups, the composition of minorities has changed: Black officers are almost twice as common among all officers (24%) compared with recruits (14%); the sharp decline in Black recruitment has been made up for by a surge in Hispanic recruitment (33% vs. 23%). This pattern is generally representative of police departments across the country in the last 30 years (Keller (2015)).

2.3.2 Policing Outcomes

Arrests are a common metric when studying individual officer and departmental performance,[12] and in the light of concerns about over-policing—excessive and detrimental interactions between law enforcement and civilians—arrests are the main metric I will use to measure officer enforcement activity. To distinguish between the seriousness of arrests, I divide them based on crime: index arrests, which I define as arrests for official index crimes (aggravated assault, robbery, murder, rape, burglary, larceny, motor vehicle theft, arson), and I also include domestic violence (if the description indicates domestic battery or assault) and sexual assault (if the description indicates criminal sexual assault); and (less serious) non-index arrests, which I define as all arrests for crimes not classified as index—e.g., warrant, traffic, or drug crimes. I also classify arrests based on arrestee race/ethnicity (white, Black, Hispanic, or other). Using Cook County court data, I determine if the arrest is

[12] See Donohue III and Levitt (2001), Mas (2006), McCrary (2007), Shi (2009), Coviello and Persico (2015), Weisburst (2020), Owens et al. (2018), and Garner, Harvey, and Johnson (2019).

associated with a guilty finding, and I interpret this as a measure of arrest quality.[13] Opportunities for officers to make arrests are dependent upon the crime rates where they work, which influence the quantity, quality,[14] and kind of arrests.

Table 2 displays arrests per shift, violent crime rates, and observations in the daily panel data for all main sample recruits as full officers in Column (1), and Columns (2) and (3) divide these officers by whether their cohort had high ($\geq 50\%$) or low ($< 50\%$) minority share. The vast majority of new officer arrests in Chicago are of Black civilians (80%), with Hispanic arrests being far less common at 14.7%—for this reason I will focus my analysis on Black arrests. Recruits in high-minority cohorts make fewer arrests per shift than those in low-minority cohorts, driven by a difference in arrests of Blacks (0.1361 vs. 0.1582). About two-thirds of arrests are for non-index crimes, and recruits in low-minority cohorts make about a 4% smaller share of their arrests for non-index crimes. Recruits in low- and high-minority cohorts have similar guilty arrest rates at 23.23% and 23.03% guilty, respectively. Recruits in low-minority cohorts work, on average, in slightly lower crime districts relative to recruits in high-minority cohort, yet both groups work in Chicago's most dangerous areas.[15] While this table documents differences between new officers in terms of arrest quantity, quality, and type, as well as working environment based on cohort diversity, whether or not cohort diversity is actually changing officer enforcement behavior requires more detailed analysis.

3 Effect of Peer Diversity on Arrests

3.1 Empirical Strategy

The aim of this paper is to estimate the long-run effect of peer diversity on officer behavior. The identification strategy for this paper borrows heavily from the education literature on long-run peer effects, leveraging the random assignment of students (officers) to classrooms (academy cohorts). As a first step, I adapt the regression specification from the long-run peer

[13] I define an arrest to be 'guilty' if the central booking number (CBN) is associated with any guilty finding; I consider an arrest not guilty if the CBN is associated with no guilty findings and at least one not guilty finding. If a CBN is associated with no guilty findings and no not guilty findings, and it has any dismissed cases, then I consider it dismissed. If a CBN does not appear in the court data, I classify the case as dropped. I group not guilty, dismissed, and dropped cases together and label them as 'non-guilty'. If a CBN is not classified as guilty, not guilty, or dismissed, but it is in the court data, then it only has incomplete/open cases, so it is classified as neither guilty nor non-guilty. A single CBN may have multiple charges or cases associated with it, and I use the method discussed above to provide a single outcome of an arrest which is conservative as only one guilty verdict on any charge is sufficient for an arrest to be 'guilty'.

[14] For example, lower crime may mean the marginal arrest is less likely to be high quality if officers value making arrests.

[15] A monthly violent crime rate of 14 per 10,000 population is the 6th most violent district in Chicago.

effects in education literature (Chetty et al. (2011), Carrell, Hoekstra, and Kuka (2018)) by regressing outcomes on the characteristics of randomly assigned peers. Specifically, I estimate:

$$\overline{Arrest}^k_{icp} = \alpha^k_{cp} + \pi^k_1 \overline{X}_{c(i)} + \pi^k_2 X_i + v^k_i \tag{1}$$

where $\overline{Arrest}^k_{icp}$ is the average arrests per shift of type k (e.g., Black non-index guilty arrests)[16] made by officer i randomly assigned to cohort c in period p. Variable α^k_{cp} is a fixed effect for the time period p during which cohort c started, such as 2008 to 2011 or 2012-07 to 2014-05 (the main sample), and is a proxy for the entrance exam officers took (see Appendix A.1). X_i contains the demographic characteristics (e.g., race, start age) for officer i. $\overline{X}_{c(i)} = \frac{\sum_{j \neq i}^{n_c} X_j}{n_c - 1}$,[17] contains the leave-out mean of the demographic characteristics of members of officer i's cohort c.

The random assignment of lottery numbers within a testing pool allows cohort composition, $\overline{X}_{c(i)}$ to be uncorrelated with unobserved characteristics about the officer, v^k_i, permitting consistent estimation of the peer effect of cohort diversity, π^k_1. More formally: $\mathbb{E}[v^k_i | \overline{X}_{c(i)}, \alpha^k_{cp}] = 0 \forall i$.[18] However, the mechanism by which cohort composition influences future arrests is not specified. One part of π^k_1 is the effect of cohort diversity on an individual officer's behavior, their opinions, beliefs, and prejudices. Yet, as cohort diversity influences officer assignments and future peers (discussed more in Appendix A.4 and Appendix A.5), and assignments influence arrest possibilities, the other part of π^k_1 is the assignment effect of diversity. So, though the assignment effect proves to be minor, π^k_1 is a causal estimate of the effect of cohort composition on an officer's future arrests of type k *within* the assignment system of the Chicago Police Department.[19]

[16]More formally, $k \in \begin{pmatrix} All \\ Index \\ Non\ Index \end{pmatrix} \times \begin{pmatrix} All \\ Guilty \\ Non\ Guilty \end{pmatrix} \times \begin{pmatrix} Minority \\ Black \\ NonBlack \\ Hispanic \\ White \end{pmatrix}$.

[17]For computing $\overline{X}_{c(i)}$, I include all recruits beginning in the cohort c excluding i.

[18]Given that cohort composition is randomly determined and $\overline{X}_{c(i)}$ excludes the officer i, cohort composition excluding officer i is independent of officer i's observable characteristics, X_i. So, leaving out X_i should not impact estimates of π^k_1.

[19]In this setting, I cannot distinguish between endogenous and exogenous effects of peers (Manski (1993)), meaning I cannot disentangle the effect of officers being affected by minority peers due to their behavior or their characteristics. I assume there are no correlated effects (e.g., instructor effects); given the large amount of courses recruits are taught during the academy, it is implausible that a cohort with 40% minority composition would receive different institutional environments or instructors than a cohort with 50% minority composition starting a month later.

3.2 Results

Table 3 displays the results of estimating equation (1) on the officers in the main sample cohorts—as a result, the period fixed effect, α^k_{cp}, is simply the intercept. I focus on arrests of Blacks as they make up the vast majority of new officer arrests. Column (1) is the most parsimonious model, only controlling for cohort share minority and cohort mean age, while Column (2) adds in controls for officer race, gender, start age, and cohort size. The coefficient for Cohort Share Minority is similar in Columns (1) and (2), providing further evidence for the random assignment of recruits to cohorts. The coefficient in Column (2) for Cohort Share Minority is -0.2, meaning that officers in cohorts with 10 percentage points (pp) more minorities (moving from the 1st quartile to the 3rd quartile of cohort diversity) make 0.02 fewer arrests of Blacks per shift on average. This is equivalent to a 13.81% decline relative to the mean Black arrests per shift for main sample officers.

The composition of these arrests is important, however, as not all crimes are equal: drug crimes are less serious than robbery or domestic abuse. I divide the arrests of Blacks into non-index and index crime arrests. Non-index crimes are less serious in general and often 'victimless crimes' (e.g., drug and traffic crimes or municipal code violations), where as index crimes always have a victim and are often very serious (e.g., robbery, assault, or burglary). Given policymakers' concerns about over-policing of minority communities for low-level crimes, it may be desirable for officers to make fewer low-level arrests. Peer diversity may cause officers to alter how aggressively they police, which may more strongly impact low-level arrests.

Columns (3) and (4) display the effect of cohort diversity on average Black arrests per shift for non-index and index crimes (low and high severity, respectively). Cohort diversity significantly decreases officers' average arrests of Blacks for non-index crimes per shift but has almost no effect on index arrests. A 10pp increase in an officer's cohort share of minorities reduces their future arrests of Blacks per 100 shifts by 1.8 for non-index crimes ($p < 0.05$) and by 0.2 for index crimes ($p > 0.1$). Relative to their means, this is equivalent to a 20.01% decline for non-index arrests and a 4.3% decline for index arrests of Blacks.

3.3 Extended Sample and Panel Structure

As this study focuses on a single period of cohorts, a concern may be that the period of study was unique and the effects are nonexistent for earlier or later cohorts. A second concern is that while scores on a standardized exam or income at a specific age are straightforward outcomes, as previously discussed, arrests are determined, in part, by an officer's assignment. Assignments are influenced by cohort composition as minorities are more likely to be assigned to high-minority areas—I explore the extent of crowding out and confounding assignment in Appendix A.4 and find it to be present but economically small for new officers.

As robustness checks, I address the former concern by re-estimating equation (1) on the extended sample of officers in cohorts between 2008 and early 2015 (all of whom I observe near their starts as full officers). Second, in order to reduce the influence of the assignment effect, I alter equation (1) for a panel structure in order to account for officer unit assignments over time by estimating:

$$Arrest^k_{icpt} = \alpha^k_{cp} + \pi^k \overline{X}_{c(i)} + \pi^k_2 X_{it} + \gamma^k_u + \lambda^k_t + v^k_{it} \tag{2}$$

where $Arrest^k_{icpt}$ is the number of arrests made during month t by officer i,[20] randomly assigned to cohort c in period p; X_{it} now contains not only fixed characteristics of officer i but also their number of shifts of each type (1-4) worked and a second-degree polynomial of tenure in month t; γ^k_u is a fixed effect for the unit u officer i was assigned to during month t; and λ^k_t is a fixed effect for month t. This panel regression also uses the extended sample. By controlling for unit and month fixed effects, π^k_1 now identifies the peer effect of diversity on officer behavior conditional on the effect of cohort diversity on district assignment. In order to have sufficient observations to include unit fixed effects, I additionally include the 2006-2007 cohorts in the panel regressions.

Table 4 displays the results of estimating equation (1) on the extended sample in Columns (1) and (2), and Columns (3)-(5) display the results of estimating equation (2) on the monthly panel. Though the point estimates are about 30% smaller, Columns (1) and (2) show the effect of peer diversity on arrests persists after including additional cohort periods, indicating that the previous results were not unique to the main sample or the 2012-2014 training period.

The results of the panel regressions in Column (3) shows that the effect of peer diversity is persistent after controlling for unit assignment and shift counts (and including even older cohorts), as officers in cohorts with 10pp more minorities make about 0.1 ($p < 0.01$) fewer arrests of Blacks conditional on assigned unit per month. This equates to about 7 fewer arrests of Blacks per 100 shifts. This is smaller than the main sample results, possibly a result of assignment controls reducing the assignment effect. Alternatively, this may be due to panel regressions over-weighting the earlier cohorts, which display somewhat smaller peer effects of diversity due to officers with higher tenure being more able to select their working units, shifts, and beat assignments. The effects of peer diversity on Black arrests by crime

[20]Using a linear probability model instead, with the outcome being if officer i made at least 1 arrest of type k during shift i, produces officer fixed effects that are highly correlated with the linear model's fixed effects and very similar results for peer effect coefficients.

severity are also consistent with the main results. While there is no statistically significant effect of minority peers on index arrests of Blacks, there is a large effect on non-index arrests of Blacks ($p < 0.05$). Comparing the point estimates, cohort diversity has a 7 times large effect on low-severity arrests compare to high-severity arrests.

These tests show that the effect of minority peers on their cohort members' outcomes is persistent and not driven by unit assignment or sample selection. They further indicate that having more minority peers in the academy decreases average arrests of Blacks in the future, driven by a decline in arrests for low-level crimes. However, while monthly level assignments are useful in understanding officer working environment, they are only a rough approximation of when and where an officer is actually patrolling– as found in Ba et al. (2021), shift-times and beat assignments are not random within units. Thus they may be limited in their ability to control for the assignment effect. For that, I turn to studying individual officer heterogeneity using the most detailed data available.

4 Effect of Peer Diversity on Officer Heterogeneity

4.1 Empirical Strategy

As documented in the previous section, higher shares of minorities in cohorts cause officers to make fewer arrests of Blacks during their careers. This is in part due to how cohort diversity influences how the Chicago Police Department assigns new officers and how officers choose to bid for assignments. For the effect of peers to be externally valid and relevant for police departments with different priorities and assignment policies, understanding how peers influence an officer's individual *type* is necessary. By an officer's type I mean their individual propensity to make an arrest, a measure for their enforcement activity regardless of their working environment, or their individual contribution to the quantity or quality of arrests they make.

In this section, I first recover a measure of an officer's 'type', i.e. their propensity to make arrests net of high dimensional daily assignment fixed effects and other factors. Then, I regress these arrest propensities on cohort composition to estimate the long-run effect of peer diversity on individual officer behavior. This allows for data reduction and exploration of heterogeneity, and it permits flexible specifications in the first stage.[21]

I first recover an estimate for all officers' (including those outside the extended sample of

[21] Weisburst (2020) uses an analogous method by first recovering officer fixed effects for making arrests following 911 calls, then regressing these officer fixed effects on officer characteristics, and Card and Krueger (1992) use a similar two-step procedure for studying the determinants of returns to education.

cohorts) propensities to make arrest of type k, θ_i^k, using a first stage regression. I estimate[22] a linear fixed effect regression model:

$$Arrest_{it}^k = \theta_i^k + \gamma_{bsw_t}^k + \beta^k V_{it} + \epsilon_{it}^k \tag{3}$$

where $Arrest_{it}^k$ is the number of arrests of type k officer i made during their on-duty time on date t. I control for assignment and environment characteristics flexibly with high-dimensional fixed effects, $\gamma_{bsw_t}^k$, which interacts officer i's assigned district and truncated beat code, b, their shift number, s, and the year, month, and day of the week, w_t. V_{it} controls for second-degree polynomials of shift duration, officer i's tenure, and the number of crimes of various types (violent, property, domestic violence, sexual assault, and other) reported during i's on-duty time on date t in officer i's assigned district. All random shocks to an officer's arrest participation during their working period are contained in ϵ_{it}^k.

I assume that conditional on polynomials of local crime, tenure, shift duration, and officer and environment fixed effects: 1. current and future shocks to arrest counts are orthogonal to past observables; 2. shocks to arrest counts are not serially correlated across shifts; 3. since the number of daily shifts I observe for each officer grows quickly, the officer fixed effects are consistently estimated. I interpret the recovered $\hat{\theta}_i^k$ as an estimate of the individual officer's propensity for enforcement of type k, which I recover for all officers in the daily assignment panel between 2010 and 2016. The data has a total of 6.5 million officer-shift observations on over 9,000 officers and contains approximately 1 million assignment fixed effects (bsw_t).

With this first stage regression, I control for significant temporal, geographic, demographic, and income variation in where each officer is working, as well as the within-day heterogeneity, officer exposure to different types of civilians,[23] and local crime rates, and the influence of an officer's experience on the force. Another strength of this design is that I am able to leverage data on all officers, not just those in the sample cohorts. This means I have sufficient observations within highly granular assignments to use high-dimensional fixed effects (i.e. $\gamma_{bsw_t}^k$) and allow for interactions between assignment characteristics. Recovering a single metric (per arrest type) for each officer also avoids weighting issues as new cohorts have fewer observations in the panel data than older ones.

Using only the fixed effects of officers in my sample cohorts, I replace $\hat{\theta}_{icp}^k$ as the depen-

[22]Estimation was performed using the R package 'lfe' (Gaure (2013a)), which implements the algorithm introduced in Gaure (2013b) that is designed for estimating linear models with multiple overlapping high-dimensional fixed effects (e.g. officers and who move across shifts or workers who move across firms). Notably, this package also allows for standard errors of the fixed effects to be recovered which is used in Section 5.1.

[23]The CPD's operational schedule reinforces the inability of officers to select shifts on specific days or civilian pools (see Ba et al. (2021) for more detail).

dent variable in equation (1):

$$\hat{\theta}^k_{icp} = \alpha^k_{cp} + \pi^k_1 \overline{X}_{c(i)} + \pi^k_2 X_i + v^k_i \tag{4}$$

Now, π^k_1 can be interpreted as the peer effect on an officer's propensity to make arrests of type k. As before, the racial composition of one's cohort is independent of one's own pre-existing characteristics, but now the outcome is the result of extensively controlling for working environment such that $\hat{\theta}^k_{icp}$ is officer i's individual contribution to make arrests of type k regardless of when or where they work. The minor effect of assignment crowding out due to cohort diversity is removed from this measure, and the exogeneity assumption ($\mathbb{E}[v^k_i|\overline{X}_{c(i)}, \alpha^k_{cp}] = 0 \forall i$) holds, making π^k_1 the causal effect of cohort diversity on officer enforcement propensity.

In my setting, the two-step procedure offers a number of benefits. First, recovering fixed effects for enforcement allows for the exploration of officer heterogeneity with dimension reduction and flexible first-stage specifications. Second, as discussed above, I believe understanding how peer diversity influences an officer's type, net of assignment effects, enhances the external validity of this paper. Third, if assignments are influenced by cohort diversity, a panel regression controlling for both officer and assignment fixed effects will return a consistent estimate of effect of officer type on arrests. Lastly, the two-step procedure allows me to overcome various difficulties with the institutional structure of the CPD as well as exploit the detail of my data to its fullest extent.[24] Running a single regression replacing officer fixed effects with officer characteristics and cohort diversity will suffer from weighting issues due to the unbalanced nature of the panel data (even within periods), selection issues, and lack plausible identification in older cohorts. As a robustness check in Section 5.11, I attempt to minimize these issues while using a single stage regression by interacting period fixed effects with cohort share minority (see Appendix A.6 for further discussion).

[24]For example, it allows me to control for officer tenure flexibly despite the staggered introduction of cohorts, making officers with 3 years of observations comparable to those with 1 year of observations. Also, about three-fourths of the over $9,000$ officers in the panel data started before 2008, and they have higher tenures (allowing them to select out of the panel) and do not have reliable cohort/period assignment data (see Appendix A.1). Excluding these officers would significantly reduce my observations, and thus would significantly reduce number of viable assignment groups ($\gamma^k_{bsw_t}$) able to be estimated leading to selection bias.

4.2 Results

The recovered distributions of main sample officer fixed effects indicate differences across race and exposure to diversity.[25] Figure 1 presents graphical evidence for heterogeneity in officer enforcement being related to officer race and cohort diversity. The distribution of officer arrest propensities (fixed effects) has a long right tail, as in Weisburst (2020). Panel A displays the distributions of fixed effects for arresting Blacks for white and minority officers in the main sample. White officers tend to have higher fixed effects, i.e. a higher individual propensity to arrest Blacks, relative to minority (non-white) officers. This conforms with existing research on white officers policing Blacks more aggressively.[26] Panel B displays the distribution of fixed effects for white officers in the main sample split by cohort share minority. Clearly, white officers in cohorts with more minorities tend to have lower fixed effects—lower individual propensities to arrest Blacks—consistent with the previous results on average arrests. As shown in Figure 2, this negative relationship is present across cohorts extending back to 2006.

4.2.1 Effect on Officer Arrest Propensity

Table 5 displays the effect of cohort diversity on officer fixed effects for arrests of Blacks and arrests of Blacks for low and high severity crimes for the main sample cohorts.[27] Columns (1) and (2) focus on the share of an officer's cohort members that are minorities as the variable of interest, while Columns (3)-(5) disaggregate Cohort Share Minority into Cohort Share Black, Hispanic, and Other (Asian/Native American/Alaskan Native/Pacific Islander).

Consistent with the previous results, the coefficient for Cohort Share Minority is statistically significant ($p < 0.1$ and $p < 0.05$) and negative, both with and without officer controls (Columns (1) and (2)), providing further support for the random assignments of cohorts. The estimate of the effect of minority peers on officer propensity for arresting Blacks in Column (2) is -0.167. This means that an officer assigned to a cohort at the 1st quartile of share minority (45.1% minority) moving to a cohort at the 3rd quartile percentile of share minority (56% minority) will decrease their future propensity to arrest Blacks by -0.018—meaning they will make 0.018 fewer Black arrests per shift across shift locations, types, and

[25] I solely discuss the fixed effects for officers in the main sample in this section, so "officer" or "recruit" both refer to officers in the main sample as sworn/full officers after their probationary period.

[26] See Goncalves and Mello (2018), Hoekstra and Sloan (2020), Weisburst (2020), and Ba et al. (2021).

[27] As in the previous section, I focus on Black arrests as they make up the vast majority of new officer arrests. Table B.7 compares the effect of different forms of racial diversity (minorities broken into Black, Hispanic, and other non-whites) on propensities to arrest Hispanics, whites, and non-Blacks from the first stage. Shares of Black and Hispanic cohort members appear to decrease officer propensity to making all non-Black arrests—though the estimates are noisy, the coefficients are all negative.

timings. This is a 12.3% decrease relative to the average Black arrests per shift for officers in the main sample—similar to the estimate in Table 3. This effect size is comparable to the 12% decline in arrest likelihood of treated officers following a supervisory meeting relating to procedural justice, as studied in Owens et al. (2018).

As shown in Columns (3)-(5), Cohort Share Black and Cohort Share Hispanic are driving the effect of minority peers on Black arrest propensity, with Black peers having an approximately 10% larger effect than Hispanic peers—though this difference is not statistically significant. Moving from a cohort at the first quartile of share Black (9.38%) to a cohort at the third quartile of share Black (16.13%), holding all other cohort characteristics equal (except share white), will cause the average new officer to make 7% fewer arrests of Blacks. Similarly, moving from a cohort at the first quartile of share Hispanic (27.91%) to the third (36.67%), will cause the average recruit to make 8% fewer arrests of Blacks. Alternatively, a 10pp increase in exposure to Black or Hispanic peers (10th to 90th percentile and 30th to 80th percentile, respectively) will cause a 0.2 and 0.19 standard deviation decrease in Black arrest propensity, respectively. (Columns (3)-(5) of this table are reproduced with bootstrap (noisier) and wild-bootstrap (similar), standard errors in Table B.12.)

The coefficient for share Black in Column (3) is not statistically significant, and this is due to the different effects of diversity on low- and high-severity arrests (as Columns (4) and (5) highlight). Officers in cohorts with 10pp more Black (Hispanic) peers have their propensity to arrest Blacks for non-index crimes reduced by -0.024 (-0.018), which is a 0.37 (0.28) standard deviation decrease. By comparison, Column (5) shows that Black and Hispanic peers have a positive and statistically significant effect on officer propensity to arrest Blacks for serious crimes. A 10pp increase in Black (Hispanic) peers increases an officer's propensity to arrest Blacks for index crimes by 0.35 (0.15) standard deviations.

Officers being Black or Hispanic is, as expected, negatively associated with a propensity for Black arrests. Black and Hispanic new officers, all else equal, are 0.44 and 0.28 standard deviations lower than white officers in the distribution of Black arrest propensities. The tendency for Black and Hispanic officers to display a lower propensity for enforcement supports existing studies, such as Weisburst (2020), Goncalves and Mello (2018), Hoekstra and Sloan (2020), and Ba et al. (2021).[28]

[28]Black officers display, on average, lower propensities to make arrests of all racial groups, though the effect is strongest with respect to arrestees who are Black, then Hispanic, and smallest for white. Notably, Hispanic officers have lower fixed effects for arresting Blacks and whites relative to white officers, but Hispanic officers display no significant difference with white officers with respect to Hispanic arrests, consistent with the findings in Ba et al. (2021). See Table B.7.

4.3 Effect on Officer Arrest Quality Propensity

As noted earlier, a unique feature of my data is that I can observe the outcomes of arrests in court, which enables me to measure the quality of arrests. To study the effect of peer diversity on arrest quality, I distinguishing between high- and low-quality arrests as arrests which result in guilty and non-guilty outcomes in court, respectively.[29] Much like arrest quantity, officers are heterogeneous in their arrest quality, and officers that have high propensities to make guilty (high-quality) arrests do not necessarily have high propensities to make non-guilty (low-quality) arrests. I estimate equation (4) on propensities to make non-guilty and guilty arrests separately. By comparing the effect of peer diversity on the propensity to make high-quality arrests with its effect on the propensity to make low-quality arrests, I can infer an effect of peer diversity on arrest quality separate from the influence of assignments, as crimes committed in some locations and times may be easier to arrest and prosecute than others. In Table 6, Columns (1)-(2) and (3)-(4) display these results for Black non-index and Black index arrests, respectively.

Black and Hispanic peers significantly decrease an officer's propensity to make both guilty and non-guilty arrests of Blacks for non-index crimes (Columns (1) and (2)); however, their effects on non-guilty are both about 5 times large than the effect on guilty. Compared to the non-guilty/guilty ratio for non-index arrests of Blacks (3.05), this implies the reduction in the propensity to make non-index arrests of Blacks caused by exposure to Black and Hispanic peers causes an increase in arrest quality, as the reduction of low-quality arrests is proportionately larger than that of high-quality ones. Alternatively, a 10pp increase in exposure to Black (Hispanic) peers causes a 0.22 (0.16) standard deviation decrease in officer propensity to make guilty arrests for non-index crimes, while it causes a 0.43 (0.32) standard deviation decrease in officer propensity to make non-guilty arrests of Blacks for non-index crimes.

For index crimes (Columns (3) and (4)), Black and Hispanic peers have small, positive, and statistically significant effects on guilty and non-guilty arrests. As the average non-guilty/guilty ratio for index Black arrests is 2.83, an increase in exposure to Black peers will

[29] As previously described, a guilty outcome means the arrest was associated with any guilty finding, and a non-guilty outcome means the arrest was not associated with any Cook County court case, resulted in only dropped charges, or had charges only found not guilty in court. Cases which had no final disposition or closed date in the data set are considered incomplete/open and are neither guilty nor non-guilty. While police officers do influence the initial charges against the arrestee and provide evidence and testimony to prosecutors and defendants, their time in front of a jury or judge is limited particularly given the frequency of plea deals. Furthermore, while officer observables may influence credibility in the eyes of judges and juries, altering their ability to make guilty arrests, an officer's cohort diversity is not observable and far removed. The largest influence an officer has on the outcome of a court case is making the arrest and the initial charges (and the evidence supporting those charges), both of which are captured in my design.

increase non-guilty arrests by 2.37 relative to guilty arrests, marking a slight improvement in arrest quality for index arrests of Blacks. Hispanic peers also have a statistically significant effect, though about one-third the size of the effect of Black peers, and marks a slight decrease in arrest quality for index crimes.

I provide further evidence of the effect of diversity on arrest quality by recovering officer fixed effects for the probability of making guilty arrests. The previous arrest quality results were based off of a comparison of two point estimates, and thus indirect. In contrast, this approach provides a direct measure of an officer's arrest quality. Specifically, I estimate:

$$\frac{Arrest_{it}^{k,guilty}}{Arrest_{it}^{k,total}}|_{Arrest_{it}^{k,total}>0} = \theta_i^k + \gamma_{bsw_t}^k + \beta^k V_{it} + \epsilon_{it}^k \tag{5}$$

where $Arrest_{it}^{k,total}$ is the total number of arrests of type k which were either guilty ($Arrest_{it}^{k,guilty}$) or non-guilty made by officer i during shift time and date t, estimated on the shifts where the officer made at least 1 arrest of type k.[30] While this provides a direct measure of an officer's propensity to make high-quality arrests, it can only be estimated using shifts where an officer made at least 1 arrest of type k. This means fixed effects can only be recovered for a subset of officers with sufficiently many non-zero arrest shifts leading to some selection bias.

The results using this direct measure of individual officer arrest quality generally confirm the main analysis. As shown in Figure 3, there is significant heterogeneity in officer arrest quality, and this heterogeneity is related to cohort diversity. Table B.8 displays regression results for the effect of cohort diversity on the arrest quality metric; higher shares of Black and Hispanic peers increase officer likelihood of making a guilty arrest for both index and non-index arrests of Blacks. I conclude that diverse peers increase arrest quality for non-index crimes, and arrest quality of index crimes may also be positively impacted.

5 Robustness

5.1 Shrinkage Estimates

The individual fixed effects used in the main analysis are based upon finite observations of officers. This means that each estimated fixed effect will have some error associated with it, and it is crucial that finite-sample error is not driving the results. A common procedure to correct for this when using individual fixed effects, popularized by the teacher value added literature (Chetty, Friedman, and Rockoff (2014)), is to do an empirical Bayes shrinkage

[30] This excludes arrests with still pending findings.

procedure (based on Morris (1983)). The idea is to shrink estimates (officer fixed effects) toward a prior mean based on how noisy the estimate is (high noise leads to a larger reduction in the estimate). Here, I will construct a shrunken estimate for each officer fixed effect ($\hat{\theta}_i$) based on how noisy it is ($Var[\hat{\theta}_i] = se(\theta_i)^2$) relative to the variance in the distribution of all fixed effects ($Var[\hat{\theta}] = \frac{1}{N}\sum_i^N \hat{\theta}_i^{\,2}$): $\hat{\theta}_i^{shrunken} = \hat{\theta}_i * \frac{Var[\hat{\theta}]}{Var[\theta_i]+Var[\hat{\theta}]}$.[31] As expected, due to the relatively large number of observations per officer in my sample (over 200 observations per main sample officer), the shrunken fixed effects are similar to those of the main results. Table B.9 replicates Table 5 with shrunken fixed effects as the dependant variables; the point estimates are very similar, and the standard errors are, in general, slightly smaller.

5.2 Discrete Outcomes in First Stage

Arrests in a shift are count data by nature, and the distribution of their frequency, as expected, fits a Poisson distribution. As such, I re-estimate the first stage using a Poisson regression.[32] This model is potentially more reflective of the true data generating process, and environment and individual officer fixed effects likely contribute to arrests in a non-linear fashion. However, unlike the linear model, the estimates are not directly interpretable and fewer individual fixed effects can be recovered.[33] I use the recovered fixed effects in my second stage, and, as shown in Column (1) of Table B.10, the results are qualitatively similar to those of the main results, though not directly comparable due to the non-linearity of the model.

While the Poisson first stage is designed to more closely match the distribution in the data, a second concern may be that the results are driven by the skewed nature of the arrest data: most shifts have no arrests at all while very few have many. To test this, I simply the data by re-estimating equation (3) as a linear probability model (LPM) with the dependent variable being if any arrest of type k was made by officer i during their shift. The results are displayed in Column (2) of Table B.10, and the estimates are very similar to the main results but slightly smaller in magnitude, as expected. These tests indicate that the results

[31] This is derived from the posterior mean of normal distribution with prior mean being zero being $\theta_i^n = \bar{\theta}_i \frac{\sigma^2}{\sigma^2 + \frac{\sigma_i^2}{n}}$, where θ_i is drawn from a $N(0, \sigma^2)$ and each observation of $\theta_i^t = \theta_i + \epsilon_i^t$ where $\epsilon_i^t \sim N(0, \sigma_i^2)$. $\bar{\theta}_i$ can be seen as the fixed effect estimate, $\frac{\sigma_i^2}{n}$ is the $se(\bar{\theta}_i)^2$, and σ^2 as the variance across estimates of θ_i's. A key advantage of the 'lfe' R package is its ability to estimate standard errors for fixed effects via bootstrapping.

[32] Specifically,

$$\mathbb{E}[Arrest_{it}^k | V_{it}, \Theta_i^k, \Gamma_{bsw_t}^k] = exp(\Theta_i^k + \Gamma_{bsw_t}^k + \eta^k V_{it}) \tag{6}$$

where Θ_i^k the officer fixed effect.

[33] The estimation is performed using the R package 'fixest' and an algorithm used to efficiently estimate fixed effects in maximum likelihood models (Bergé (2018)). As the data is not overly dispersed, a negative binomial regression is not necessary.

were not driven by either the reliance on a linear model in the first stage nor the skewed distribution of arrests per shift.

5.3 Endogenous Assignment

There may be a concern that new officers are not only assigned based on race and the racial composition of their cohort, but also based on their preferences and unobservables with respect to policing as well. I repeat my analysis on main sample officers during their probationary periods—which alleviates this issue as they have no actual policing experience upon entering this period. Column (3) in Table B.10 display these results, with effects qualitatively similar to the main results. Second, the potential for non-Black officers to be sent to different assignments due to higher shares of Blacks in their cohorts may negatively bias my results (increasing magnitude) as high cohort diversity may lead to low fixed effects solely due to assignments (though the high dimensional working environment controls attempt to solve this issue). To ensure that this potential bias is not driving my results, I study the subset of new officers exposed to very high crime areas (whose average district crime rate is above the 75th percentile of violent crime per capita). As shown in Column (4) in Table B.10, the 'high crime' new officers display similar results as the whole sample, with slightly larger effect sizes.

5.4 2006-2007 Cohorts

Another concern may be that my results are only applicable to the set of cohorts I use in my analysis, that 2012-2014 was a unique time in Chicago for new officers, or that the results depend on observing the officers at the beginning of their careers. The breadth of the data set allow me to perform a robustness check using a sample of oldest cohorts in my sample, specifically officers starting between 2006-2007.[34] Column (5) in Table B.10 presents the results of equation (4) on the sample of oldest cohorts. I find effects that are qualitatively similar to the main sample. Furthermore, Figure 2 visually illustrates that more diverse cohorts are associated with lower arrest propensities across periods.

5.5 Restricted First Stage

In the main sample for the first stage regression, I do not filter out officers for having additional information codes (e.g., indicating injury, training, union business, etc.) during

[34]However, because the arrest data begins in 2010, I can only estimate their fixed effects at least 3 years after they started at the academy, meaning some officers may have been reassigned to positions outside of my analysis.

their shifts as this information is not available for my full sample of assignment data. I re-estimate the first stage on a restricted set of shifts (those with no additional information code indicating an absence and only the regular shifts, 1-3) and control for environment more granularly (using full beat codes instead of beat numbers). This corresponds closely to the sample of analysis in Ba et al. (2021). Column (6) in Table B.10 present the results of equation (4) on the fixed effects recovered from this restricted first stage, and they are similar to those of the main sample.

5.6 Selection into the Academy

While eligible applicants are permitted to enter the academy when their lottery number is drawn (and passing further tests), it is not required. A potential recruit may have moved, found a different job, or decided against joining the CPD between the time of the test and when their number is called. This means the composition of cohorts may be influenced by selection into the academy due to different start dates. I test this by estimating equation (4) on the subset of cohorts which started within 5 months of the initial cohort (for the main sample this means only cohorts starting between July and December of 2012). Column (7) in Table B.10 displays these results. The results for non-index arrests are similar to those in the main results, and the effects on index arrests are small but negative. Figure B.1 displays how the coefficients of interest change as more cohorts are added in the main and 2006-2007 cohort samples (the 2008-2011 cohorts are too few for this exercise), with propensity to arrest Blacks as the outcome; the coefficients are qualitatively similar as more cohorts are included.

5.7 First Arresting Officer

As multiple officers can be listed on a single arrest, this means some arrests are double counted in my analysis. To check the robustness of my results against this issue, I re-estimate the first stage only counting arrests for the first arresting officer. Column (8) in Table B.10 displays the results of equation (4) on these recovered fixed effects, and they are similar to the main results but smaller, as expected.

5.8 Excluding Crime from First Stage

The level of crime during an officer's shift may be partially determined by the officer's arresting activity, or the officer's interactions with civilians may encourage or discourage the reporting of crime. This may cause my first stage estimates to be biased as shift-specific unobservables that determine arrests would not be orthogonal to crime rates. As a robustness check, I remove controls for crime from the first stage estimation. Column (9) in Table B.10 display the results of equation (4) on these recovered fixed effects; they are qualitatively similar and larger in magnitude.

5.9 Car Patrols and Assignment Type

While the assignment fixed effects control for significant amounts of heterogeneity in assignment, they do not control for the type of assignment an officer has, e.g. in car, on foot, or on bike patrol (though his is often determined by their beat and shift). To ensure this heterogeneity is not contaminating my results, I re-estimate equation (3) with more precise assignment fixed effects $\gamma^k_{absw_t}$, where a indicates the assignment type (in car, on foot, or on bike patrol). As car patrols make up the vast majority of assignments in the sample (88.58% of assignments), I also re-estimate equation (3) on the sub-sample of car assignments only. The results are displayed in Columns (10) and (11) of Table B.10 for the assignment type interaction and the car-only subsample, and the results are almost identitical to those from main sample.

5.10 Including Training Time Violators

The main sample only contains officers who passed a series of filters discussed previously. A concern may be that these filters would select officers who display different effects from those who do not make it past all filters. To test this, I construct an expanded main sample, which contains any officer in the main sample cohorts for whom a fixed effect could be recovered (thus including those with training time violations). The results are displayed in Column (12) of Table B.10 and are very similar to those of the main sample.

5.11 Combined First and Second Stage

These results rely on the two-stage method previously described. To assure the robustness of these results, I estimate a combined first and second stage regression using all officers in the daily panel data. As previously discussed, these results suffer from weighting issues, which I attempt to reduce by interacting cohort share minority with a period fixed effect—the period denoting when the cohort started, such as 2012-2014 (main sample), 2008-2011, 2006-2007, 2002-2005, etc.. See Appendix A.6 for further discussion of sample used, relevant issues, and results. For the main sample cohorts, I find a statistically significant effect of minorities, with a coefficient of -0.04. This is smaller than the main results but still economically significant; it implies that a 10pp increase in cohort share minority decreases arrests of Blacks by 3%

(compared with a 11% decline found in Section 4.2). As this is the smallest effect I find, I use it as the lower bound. Consistent with the main results, the effect is driven by a decline in low-quality arrests of Blacks for low-level crimes, leading to an increase in arrest quality. The results for index arrests of Blacks are small and insignificant (see Figure B.4).

5.12 Gender Composition

Black officers are more likely to be female than non-Black officers, and there is significant discussion around the role of women in policing. To see if gender differences are solely driving my results, I control for Cohort Share Female and report the results in Table B.11. I find no statistically significant effect of gender composition when controlling for racial composition, and the cohort share coefficient magnitudes do not change significantly relative to Table 5.

6 Mechanisms

The results show that the share of minorities in a cohort influences its members' propensity to make arrests and improves arrest quality. In general, the underlying mechanisms for peer effects of diversity can be divided into two categories based on mechanisms, direct and indirect, with significantly different policy implications.

Direct effects are the most common focus of the diversity and peer effects literature. Studies find that socialization with Blacks and other minorities causes whites to be less prejudiced towards minorities, more supportive of affirmative action, and more likely to have Black roommates or partners in the future (Laar et al. (2005), Boisjoly et al. (2006), Merlino, Steinhardt, and Wren-Lewis (2019), Carrell, Hoekstra, and West (2019)). Naturally, these effects depend on interracial contact, which is not frequent, unless contrived, due to racial homophily in socialization (Marmaros and Sacerdote (2006), Baker, Mayer, and Puller (2011)). In the police academy, more minorities may increase the likelihood of white recruits to befriend or socialize with minorities, altering their perceptions and reducing prejudice, which would reduce their preferences for arresting Blacks for low-level crimes in the future.

Indirect effects, on the other hand, are at play when peers influence the environment, how group members interact, or how they are instructed. Their prevalence is common in the education literature, where, for example, student misbehavior can damage all others' learning environment and reduce teacher effectiveness (Lavy and Schlosser (2011), Carrell, Hoekstra, and Kuka (2018)). The indirect effects of peer race have been found to alter how jurors discuss cases (Sommers (2006)) and may influence how cases are presented and defended in court (Anwar, Bayer, and Hjalmarsson (2012)). One reason for this is that in-group members may self-censor or speak differently in the presence of out-groups, such as women or minorities (Loury (1994)). In the police academy, more minorities may cause their peers to arrest fewer Blacks in the future due to instructors teaching recruits differently, a crowding out or censorship of anti-Black views, or a change in how officers view themselves with respect to race and group-identity (Akerlof and Kranton (2000)).

While data on officer friendships and instruction during the academy are not available, I can infer suggestive evidence for direct and indirect peer effects by their expected effects on white and non-white recruits. If direct effects are driving the results, white officers should be significantly more affected by cohort diversity than non-white officers. If indirect effects are the core mechanism, then both whites and non-whites should be similarly affected by cohort diversity. I test these predictions in Table 7 by interacting an indicator for an officer being white with cohort share of minorities.

Columns (1) and (4) display the results for these coefficients on officer propensity to arrest Blacks for non-index and index crimes respectively. The interaction coefficients can be interpreted as the *additional* impact of Black and Hispanic peers on white officers relative to the impact of Black and Hispanic peers on minority officers. The point estimates of share Black and share Hispanic did not change significantly relative to those in Table 5, and the interaction coefficients are small and not statistically significant.[35] Minority peers impact white and minority officers in a similar way; this is inconsistent with direct effects and consistent with indirect effects.

I find similar insignificant results for the interaction terms in the arrest quality regressions for non-index (Columns (2) and (3)) and index (Columns (5) and (6)) arrests of Blacks. This is not supportive of a potential 'recognition' effect, wherein familiarity with Black peers would make white officers better at recognizing or communicating with Black suspects and thus allow them to make higher quality arrests of Blacks. It may be that diversity's positive effect on arrest quality is simply a result of the decline in arrest propensity, that the effect of diverse peers is to increase an officer's threshold for initiating an arrest or resolving an interaction by arrest.

In conjunction with the results in Table 6, it appears that officers in diverse cohorts simply make fewer arrests for low-level crimes, and those arrests they do make are on average of higher quality. Suppose diverse peers cause the cost of the marginal arrest to increase, resulting in a decrease in arrest propensity. Then officers have a higher threshold for making an arrest. Thus, for an arrest to occur the expected probability of a guilty finding must be higher—assuming officers value the suspect being 'guilty' and the expected guilt of the

[35] While point estimates are small, they are also imprecise, meaning the data may simply not be rich enough to definitely answer this question.

marginal arrest is decreasing in arrests—leading to an increase in average arrest quality and decrease in arrest quantity. While I cannot directly distinguish between officers working less hard and officers being more discerning due to cohort diversity, the fact that cohort diversity has a non-negative, and likely positive, effect on quantity and quality of arrests for serious crimes is suggestive evidence that diverse cohorts do not drive officers to simply reduce effort.

I also find decreasing returns to cohort diversity, meaning that the negative effect of minority peers on future arrest propensities is driven by cohorts with lower minority shares. I find this to be present using cut-off values exploring a 'tipping point' dynamic (Card, Mas, and Rothstein (2008), Anwar, Bayer, and Hjalmarsson (2012), Blair (2017))– discussed in Appendix A.7. The decreasing returns to diversity are also discussed in Appendix A.8 which explores the difference between diversity and representation (which I have used interchangably). Most clearly, the decreasing returns to diversity is visible graphically across all cohort periods (2006-2007, 2008-2011, and 2012-2014) in Figure 2. In each period, there is a clear negative relationship between residualized share minority and arrest propensity while the smoothed relationship shows that this is strongest in low-minority cohorts and the relationship flattens or turns positive at higher levels of minorities. The fact that minority representation matters most when they are less represented is inline with how groups of recruits and instructors may respond to minority presence, e.g. going from 15 to 20 minority recruits in a class of 50 is noticeable while going from 25 to 30 minority recruits is less significant for changing how the group and instructors behave. Naturally, these interactions, friendship formation, and other factors likely shift as minorities become the majority, but these dynamics are not able to be explored with the present data.

Diverse peers influencing how officers learn to police rather than more interracial contact reducing bias is also consistent with the results for the effect of peer diveristy on non-Black arrests. As shown in Table B.7, I find that diverse peers decrease officers' propensity to make non-Black (white, Hispanic, Asian, Native American) arrests as well, though the effect sizes are smaller and noisier due to Blacks being the vast majority of arrestees in Chicago. Overall, minority peers decreases arrests of non-Blacks for both index and non-index crimes, but the reduction is largest for non-index crimes. These results suggest that officers in cohorts with more minorities end up policing both Blacks and non-Blacks differently, which further supports the indirect effect of diversity as officers learn to police differently which spills over into their interactions with non-Blacks as well.

While these results provide evidence against direct peer effects playing a significant role, I cannot determine the actual mechanisms with the data. This provides significant room for future research. Studying officer friendships formed during the academy or how instructors (and officers) behave and communicate depending on the composition of their classrooms would provide a clearer explanation of the present results. For policy, if white-minority friendships are key, assigning white officers to work with minority partners may be beneficial, but this may complicate assignments and community policing efforts. Alternatively, as suggested by the results, minority peers may influence how officers learn to police, which entails a wide range of potential mechanisms such as altered demonstration or field training, reduced normalization of poor policing tactics, or a reduced sense of immunity from misconduct or bad policing among officers. For example, if minority recruits cause instructors to teach differently, this altered teaching may be able to be replicated through supervision, improved curriculum, or training of the instructors themselves. The result that all officers' policing is improved by cohort diversity refocuses the common discussion around race as focusing on the effect of Blacks (minorities) on Whites; peer diversity may have positive effects for all officers.

7 Conclusion

Minority peers in the police academy, specifically Blacks and Hispanics, significantly reduce their fellow officers' long-run propensities to arrest Blacks for less serious crimes. Additionally, peer diversity causes officers to improve the average quality of their arrests, driven by a reduction in the quantity of low-quality arrests. The effect on the quantity and quality of arrests for serious crimes is smaller and less consistent but generally positive—implying a shift in focus from less serious arrests to more serious arrests. The effect of diversity on arrests for low-level offenses is consistent across various specifications and robustness checks.

These results are particularly striking, as they cannot be explained by increased white familiarity with minorities leading to decreased bias. Rather, I find that white and non-white officers are equally impacted by the share of Black and Hispanic peers in the academy. In this way, Black and Hispanic officers improve the policing of all of their peers. Future scholarship should examine the mechanisms by which diverse peers affect individual policing behavior in more depth. Though understudied in economics, such work will also advance the study of how diversity impacts institutions, as this paper suggests that the effects of diversity may be more consequential than previously thought, particularly in areas where race is highly salient.

The policy implications of these findings are far reaching and promising for improving policing. The inclusion of minority officers can result in long-run effects on their peers, reducing over-policing of low-level offenses while not reducing propensities to make arrests for more serious crimes. Importantly, officers' propensities to make high-quality arrests increases with peer diversity. This means increasing departmental diversity through recruitment of more minority officers can result in fewer wasted public resources, fewer individuals put under undue burdens, and fewer separated families.

8 References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics* 115 (3): 715–53. https://doi.org/10.1162/003355300554881.

Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics*, Special Section articles on "What determined the dynamics of labour economics research in the past 25 years? Edited by Joop Hartog and and European Association of Labour Economists 25th Annual Conference, Turin, Italy, 19-21 September 2013 Edited by Michele Pellizzari, 30 (October): 98–108. https://doi.org/10.1016/j.labeco.2014.05.008.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *The Quarterly Journal of Economics* 127 (2): 1017–55. https://www.jstor.org/stable/23252002.

Ba, Bocar A., Dean Knox, Jonathan Mummolo, and Roman Rivera. 2021. "The Role of Officer Race and Gender in Police-Civilian Interactions in Chicago." *Science* 371 (6530): 696–702. https://doi.org/10.1126/science.abd8694.

Baker, Sara, Adalbert Mayer, and Steven L. Puller. 2011. "Do More Diverse Environments Increase the Diversity of Subsequent Interaction? Evidence from Random Dorm Assignment." *Economics Letters* 110 (2): 110–12. https://doi.org/10.1016/j.econlet.2010.09.010.

Bergé, Laurent R. 2018. "Efficient Estimation of Maximum Likelihood Models with Multiple fixed-Effects: The R Package FENmlm." *CREA Discussion Papers*, no. 13: 39.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2013. "Under Pressure? The Effect of Peers on Outcomes of Young Adults." *Journal of Labor Economics* 31 (1): 119–53. https://doi.org/10.1086/666872.

Blair, Peter. 2017. "Outside Options (Now) More Important Than Race in Explaining Tipping Points in US Neighborhoods." Working Paper. Human Capital; Economic Opportunity Working Group.

Boisjoly, Johanne, Greg J. Duncan, Michael Kremer, Dan M. Levy, and Jacque Eccles. 2006. "Empathy or Antipathy? The Impact of Diversity." *The American Economic Review* 96 (5): 1890–1905. https://www.jstor.org/stable/30035002.

Burns, Justine, Lucia Corno, and Eliana La Ferrara. 2015. "Interaction, Prejudice and Performance. Evidence from South Africa." *Working Paper*, February, 52.

Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100 (1): 1–40. https://doi.org/10.1086/261805.

Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation*." *The Quarterly Journal of Economics* 123 (1): 177–218. https://doi.org/10.1162/qjec.2008.123.1.177.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review* 108 (11): 3377–3415. https://doi.org/10.1257/aer.20160763.

Carrell, Scott E., Mark Hoekstra, and James E. West. 2019. "The Impact of College Diversity on Behavior Toward Minorities." *American Economic Journal: Economic Policy* 11 (4): 159–82. https://doi.org/10.1257/pol.20170069.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *The Quarterly Journal of Economics* 126 (4): 1593–1660. https://www.jstor.org/stable/41337175.

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. https://doi.org/10.1257/aer.104.9.2593.

Close, Billy R., and Patrick L. Mason. 2007. "Searching for Efficient Enforcement: Officer Characteristics and Racially Biased Policing." *Review of Law and Economics* 3 (2): 263–322. https://heinonline.org/HOL/P?h=hein.journals/rvleco3&i=263.

Coviello, Decio, and Nicola Persico. 2015. "An Economic Analysis of Black-White Disparities in the New York Police Department's Stop-and-Frisk Program." *The Journal of Legal Studies* 44 (2): 315–60. https://www.jstor.org/stable/26457029.

"CPD 2017 FAQ." 2017.

Currie, Janet M., and W. Bentley MacLeod. 2020. "Understanding Doctor Decision Making: The Case of Depression Treatment." *Econometrica* 88 (3): 847–78. https://doi.org/10.3982/ECTA16591.

Donohue III, John J., and Steven D. Levitt. 2001. "The Impact of Race on Policing and Arrests." *The Journal of Law & Economics* 44 (2): 367–94. https://doi.org/10.1086/322810.

"Education and Training Division (ETD) Chicago Police Department." 2020. https://home.chicagopolice.org/about/specialized-units/education-and-training-division-etd/.

"Field Training and Evaluation Program." 2018. http://directives.chicagopolice.org/directives/data/a7a57be2-1294231a-bf312-942c-e1f46fde5fd8c4e8.html?hl=true.

Garner, Maryah, Anna Harvey, and Hunter Johnson. 2019. "Estimating Effects of Affirmative Action in Policing: A Replication and Extension." *International Review of Law and Economics*, November, 105881. https://doi.org/10.1016/j.irle.2019.105881.

Gaure, Simen. 2013a. "Lfe: Linear Group Fixed Effects." *The R Journal* 5 (2): 104. https://doi.org/10.32614/RJ-2013-031.

———. 2013b. "OLS with Multiple High Dimensional Category Variables." *Computational Statistics & Data Analysis* 66 (October): 8–18. https://doi.org/10.1016/j.csda.2013.03.024.

Goncalves, Felipe, and Steven Mello. 2018. "A Few Bad Apples? Racial Bias in Policing," October, 80.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman. 2009. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence*." *The Economic Journal* 119 (540): 1243–69. https://doi.org/10.1111/j.1468-0297.2009.02271.x.

Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics* 1 (4): 34–68. https://doi.org/10.2307/25760181.

Harvey, Anna, and Taylor Mattia. 2019. "Reducing Racial Disparities in Crime Victimization." *Working Paper*, December, 45.

Hoekstra, Mark, and CarlyWill Sloan. 2020. "Does Race Matter for Police Use of Force? Evidence from 911 Calls." Working Paper 26774. National Bureau of Economic Research. https://doi.org/10.3386/w26774.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." Working Paper 7867. National Bureau of Economic Research. https://doi.org/10.3386/w7867.

Kass, John, and Robert Blau. 1991. "POLICE HIRING LOTTERY LATEST DALEY HEADACHE." *Chicago Tribune*, August, 3. https://www.chicagotribune.com/news/ct-xpm-1991-08-02-91032.

Keller, Meg. 2015. "Diversity on the Force: Where Police Don't Mirror Communities." Governing. https://media.governing.com/documents/policediversityreport.pdf.

Laar, Colette Van, Shana Levin, Stacey Sinclair, and Jim Sidanius. 2005. "The Effect of University Roommate Contact on Ethnic Attitudes and Behavior." *Journal of Experimental Social Psychology* 41 (4): 329–45. https://doi.org/10.1016/j.jesp.2004.08.002.

Lavy, Victor, and Analía Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics* 3 (2): 1–33. https://www.jstor.org/stable/41288627.

Linos, Elizabeth. 2018. "More Than Public Service: A Field Experiment on Job Advertisements and Diversity in the Police." *Journal of Public Administration Research and Theory* 28 (1): 67–85. https://doi.org/10.1093/jopart/mux032.

Loury, Glenn C. 1994. "Self-Censorship in Public Discourse: A Theory of 'Political Correctness' and Related Phenomena." *Rationality and Society* 6 (4): 428–61. https://doi.org/10.1177/1043463194006004002.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531–42. https://doi.org/10.2307/2298123.

Marmaros, David, and Bruce Sacerdote. 2006. "How Do Friendships Form?" *The Quarterly Journal of Economics* 121 (1): 79–119. https://www.jstor.org/stable/25098785.

Mas, Alexandre. 2006. "Pay, Reference Points, and Police Performance." *The Quarterly Journal of Economics* 121 (3): 783–821. https://www.jstor.org/stable/25098809.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *The American Economic Review* 99 (1): 112–45. https://www.jstor.org/stable/29730179.

McCrary, Justin. 2007. "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police." *The American Economic Review* 97 (1): 318–53. https://www.jstor.org/stable/30034397.

Merlino, Luca Paolo, Max Friedrich Steinhardt, and Liam Wren-Lewis. 2019. "More Than Just Friends? School Peers and Adult Interracial Relationships." *Journal of Labor Economics* 37 (3): 663–713. https://doi.org/10.1086/702626.

Miller, Amalia R, and Carmit Segal. 2018. "Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence." *The Review of Economic Studies* 86 (5): 2220–47. https://doi.org/10.1093/restud/rdy051.

Morris, Carl N. 1983. "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association* 78 (381): 47–55. https://doi.org/10.2307/2287098.

Owens, Emily, David Weisburd, Karen L. Amendola, and Geoffrey P. Alpert. 2018. "Can You Build a Better Cop?" *Criminology & Public Policy* 17 (1): 41–87. https://doi.org/10.1111/1745-9133.12337.

"Personnel Transfer and Assignment Procedures – (FOP)." 2011. http://directives.chicagopolice.org/directives/data/a7a57be2-12bcf25e-31612-bcf2-5ebc1c9f5d96947f.html?hl=true.

Pettigrew, Thomas F. 1998. "Intergroup Contact Theory." *Annual Review of Psychology* 49 (1): 21. https://doi.org/10.1146/annurev.psych.49.1.65.

Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90 (5): 751–83. https://doi.org/10.1037/0022-3514.90.5.751.

Pritchard, Paige. 2013. "Do You Have What It Takes to Join the Chicago Police Department?" *Chicago Magazine*, August. https://www.chicagomag.com/Chicago-Magazine/The-312/August-2013/CPD/.

Rim, Nayoung, Roman Rivera, Bocar Ba, and Andrea Kiss. 2019. "In-Group Bias and the Police: Evidence from Award Nominations." *Working Paper*, October, 50.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *The Quarterly Journal of Economics* 116 (2): 681–704. https://www.jstor.org/stable/2696476.

Shi, Lan. 2009. "The Limit of Oversight in Policing: Evidence from the 2001 Cincinnati Riot." *Journal of Public Economics* 93 (1-2): 99–113. https://doi.org/10.1016/j.jpubeco.2008.07.007.

Sommers, Samuel R. 2006. "On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations." *Journal of Personality and Social Psychology* 90 (4): 597–612. https://doi.org/10.1037/0022-3514.90.4.597.

Weisburst, Emily K. 2020. "'Whose Help Is on the Way?' The Importance of Individual Police Officers in Law Enforcement Outcomes," July, 66.

West, Jeremy. 2018. "Racial Bias in Police Investigations." *Working Paper*, October, 37.

Table 1: Summary Statistics by Sample

	Main Sample of Recruits	Sample Cohorts	All Officers
	(1)	(2)	(3)
Male	0.806	0.801	0.758
Female	0.194	0.199	0.242
White	0.511	0.504	0.488
Minority	0.489	0.496	0.512
Black	0.138	0.13	0.242
Hispanic	0.307	0.326	0.234
Other Race	0.0437	0.0394	0.0362
Birth Year	1982.7	1982.7	1972.39
Start Age	29.96	30.11	29.19
Cohort Size	61.52	54.24	-
N	962	21	9343

Note: Table compares the average characteristics of main sample officers to the rest of the officers in the panel data. Column (1) contains the pooled average characteristics over all main sample recruits. Column (2) contains the average characteristics of the cohorts of the recruits in (1), including those recruits that do not appear in the main analysis due to attrition. Column (3) contains the average characteristics of all officers in the daily assignment panel data.

Table 2: Summary Statistics of Main Sample Officer Outcomes

	All	High Minority Cohort	Low Minority Cohort
	(1)	(2)	(3)
Total Arrests	0.19 (0.47)	0.17 (0.45)	0.2 (0.48)
White Arrestees	0.01 (0.1)	0.01 (0.1)	0.01 (0.1)
Black Arrestees	0.15 (0.42)	0.14 (0.4)	0.16 (0.43)
Hispanic Arrestees	0.03 (0.18)	0.03 (0.18)	0.03 (0.19)
Total Index Arrests	0.06 (0.27)	0.06 (0.26)	0.06 (0.27)
Guilty Index Arrests	0.02 (0.13)	0.01 (0.13)	0.02 (0.13)
Total Non-Index Arrests	0.12 (0.39)	0.11 (0.37)	0.13 (0.4)
Guilty Non-Index Arrests	0.03 (0.18)	0.02 (0.17)	0.03 (0.18)
Violent Crime Rate in District	15.3 (7.24)	14.9 (7.47)	15.63 (7.04)
Obs	322729	143948	178781
Unique Officers	962	492	470

Note: Table presents the average number of arrests per shift, violent crime rate in average working district, and total observations for main sample recruits as full officers from 2013 to 2016. Columns (2) and (3) divide those recruits by whether or not they were in a cohort with a high (at least 50%) or low (less than 50%) share of minorities. Violent crime rate is determined by the district's average violent crime rate per 10,000 population (2010 Census) in the month of an officer's assignment. Standard deviations are reported in parentheses.

Table 3: Effect of Cohort Diversity on Arrests

	Arrests of Blacks per Shift			
	All		Non-Index	Index
	(1)	(2)	(3)	(4)
Cohort Share Minority	-0.182^{**}	-0.205^{**}	-0.183^{**}	-0.022
	(0.080)	(0.088)	(0.077)	(0.019)
Cohort Mean Age	-0.005	-0.002	-0.002	0.000
	(0.009)	(0.012)	(0.010)	(0.002)
Black		-0.032^{***}	-0.031^{***}	-0.001
		(0.012)	(0.011)	(0.003)
Hispanic		-0.024^{***}	-0.020^{***}	-0.004^{**}
		(0.008)	(0.007)	(0.002)
Other Race		-0.043^{***}	-0.030^{**}	-0.013^{***}
		(0.015)	(0.013)	(0.004)
Male		0.046^{***}	0.037^{***}	0.008^{***}
		(0.008)	(0.007)	(0.002)
Start Age		-0.004^{***}	-0.003^{***}	-0.001^{***}
		(0.001)	(0.001)	(0.000)
Cohort Size		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Intercept	0.374	0.379	0.313	0.066
	(0.249)	(0.360)	(0.303)	(0.071)
R^2	0.012	0.096	0.091	0.037
Num. obs.	962	962	962	962

Note: Table displays the effect of cohort diversity on main sample (2012-2014 cohorts) officers' average arrests per shift of Blacks (Columns (1) and (2)) and of Blacks for index (Column (3)) and non-index (Column (4)) crimes. The parameter estimates are based on the specification in equation (1). Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 4: Effect of Cohort Diversity on Arrests - Extended Sample

	Arrests of Blacks per Shift		Monthly Arrests of Blacks		
	All		All	Non-Index	Index
	(1)	(2)	(3)	(4)	(5)
Cohort Share Minority	-0.150^{*}	-0.137^{**}	-1.000^{***}	-0.869^{**}	-0.130
	(0.082)	(0.064)	(0.350)	(0.361)	(0.166)
Cohort Mean Age	-0.006	-0.013^{*}	-0.002	0.008	-0.009
	(0.008)	(0.007)	(0.044)	(0.046)	(0.017)
Cohort 2014-08 - 2015-02	-0.071^{***}	-0.084^{***}	0.017	0.175	-0.158^{***}
	(0.021)	(0.021)	(0.149)	(0.158)	(0.037)
Cohort 2008 - 2011		0.034^{***}	-0.290	-0.207	-0.082
		(0.012)	(0.207)	(0.169)	(0.074)
Cohort 2006 - 2007			-0.568^{*}	-0.438	-0.130
			(0.339)	(0.324)	(0.095)
Officer Controls	X	X	X	X	X
Tenure and Shift Controls			X	X	X
Month and Unit Fixed Effects			X	X	X
R^2	0.110	0.108	0.399	0.395	0.120
Num. obs.	1252	1667	96064	96064	96064

Note: Table displays the effect of cohort diversity on officers' average Black arrests per shift (Columns (1)-(2)), monthly arrests of Blacks (Column (3)), and monthly arrests of Blacks for non-index crimes (Columns (4)) and index crimes (Column (5)). Columns (1)-(2) iteratively add more cohort periods in addition to the main sample cohorts. The officers are those from the sample cohorts between 2008 and Feb. 2015, and the monthly panel covers 2010 to 2016 with the inclusion of the 2006-2007 cohorts for sufficient variation to control for unit fixed effects. The parameter estimates are based on the specification in equation (1) (Columns (1)-(2)) and equation (2) (Columns (3)-(5)). Officer Controls refers to the controls in Column (2) of Table 3. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at the cohort level for Columns (1)-(2) and unit level for Columns (3)-(5). $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 5: Effect of Cohort Diversity on Arrest Propensity

	Arrest Propensity				
		Black		Black Non-Index	Black Index
	(1)	(2)	(3)	(4)	(5)
Cohort Share Minority	-0.146^{*}	-0.167^{**}			
	(0.075)	(0.078)			
Cohort Share Black			-0.151	-0.237^{*}	0.086^{***}
			(0.134)	(0.135)	(0.020)
Cohort Share Hispanic			-0.141^{**}	-0.177^{**}	0.036^{***}
			(0.070)	(0.077)	(0.013)
Cohort Share Other			-0.345	-0.342	-0.003
			(0.215)	(0.212)	(0.025)
Black		-0.033^{***}	-0.033^{***}	-0.029^{***}	-0.005^{*}
		(0.008)	(0.008)	(0.007)	(0.002)
Hispanic		-0.021^{***}	-0.021^{***}	-0.017^{***}	-0.004^{**}
		(0.006)	(0.006)	(0.005)	(0.002)
Other Race		-0.024^{**}	-0.024^{**}	-0.020^{**}	-0.005
		(0.010)	(0.010)	(0.008)	(0.004)
Male		0.030^{***}	0.029^{***}	0.024^{***}	0.005^{***}
		(0.006)	(0.006)	(0.004)	(0.002)
Cohort Mean Age	-0.004	-0.001	-0.002	0.000	-0.003^{*}
	(0.007)	(0.009)	(0.007)	(0.008)	(0.001)
Start Age		-0.003^{***}	-0.003^{***}	-0.002^{***}	-0.001^{***}
		(0.001)	(0.001)	(0.001)	(0.000)
Cohort Size		0.000	0.000	0.000	-0.000^{***}
		(0.000)	(0.000)	(0.000)	(0.000)
Intercept	-0.038	-0.060	-0.031	-0.260	0.230^{***}
	(0.182)	(0.269)	(0.218)	(0.227)	(0.044)
R^2	0.017	0.114	0.118	0.129	0.055
Num. obs.	962	962	962	962	962

Note: Table displays the effect of cohort diversity on main sample officers' propensities to arrest Blacks for all (Columns (1)-(3)), low-severity (Column (4)), and high-severity (Column (5)) crimes. The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4). Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 6: Effect of Cohort Diversity on Arrest Quality Propensity

	Black Arrest Propensity			
	Non-Index		Index	
	Guilty	Non-Guilty	Guilty	Non-Guilty
	(1)	(2)	(3)	(4)
Cohort Share Black	-0.04	-0.20^{*}	0.03^{***}	0.06^{***}
	(0.03)	(0.10)	(0.01)	(0.02)
Cohort Share Hispanic	-0.03^{*}	-0.15^{**}	0.01^{*}	0.02^{**}
	(0.02)	(0.06)	(0.00)	(0.01)
Full Controls	X	X	X	X
R^2	0.09	0.13	0.02	0.05
Num. obs.	962	962	962	962

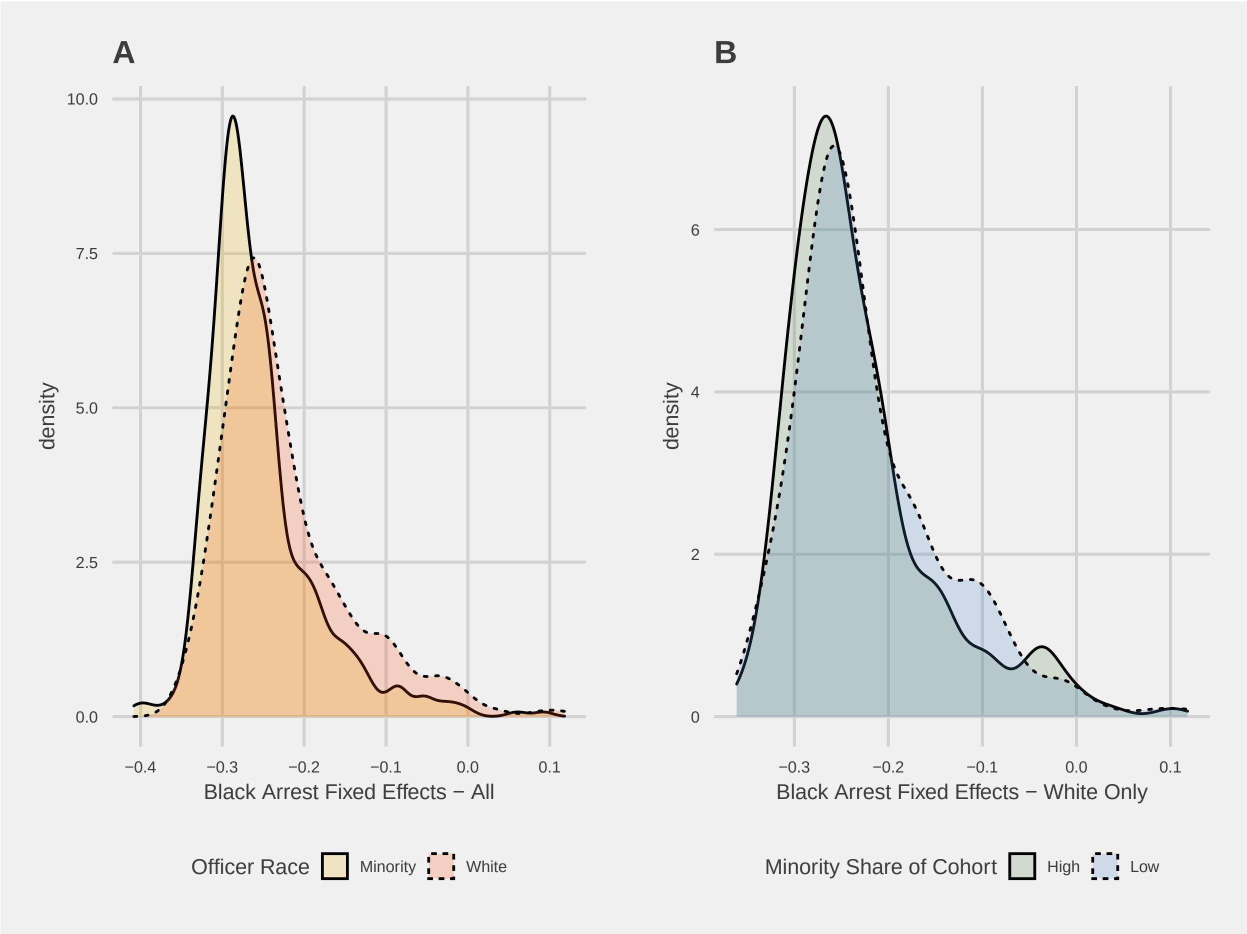
Note: Table displays the effect of cohort diversity on officer propensity to arrest Blacks for index and non-index crimes that resulted in guilty or non-guilty outcomes. The propensities used in Columns (1)-(2) and (3)-(4) is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4), with Full Controls referring to the specification in Column (3) of Table 5. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 7: Effect of Cohort Diversity on Arrest Propensity with Officer Race Interaction

	Propensity to Arrest Blacks					
	Non-Index			Index		
	All	Guilty	Non-Guilty	All	Guilty	Non-Guilty
	(1)	(2)	(3)	(4)	(5)	(6)
Cohort Share Black	-0.25	-0.04	-0.21^{*}	0.07^{**}	0.02^{**}	0.04^{*}
	(0.15)	(0.04)	(0.11)	(0.03)	(0.01)	(0.02)
Cohort Share Hispanic	-0.16^{**}	-0.02	-0.14^{***}	0.03^{**}	0.01^{**}	0.02
	(0.07)	(0.02)	(0.05)	(0.02)	(0.01)	(0.01)
White x Cohort Share Black	0.01	0.00	0.02	0.03	0.00	0.03
	(0.12)	(0.03)	(0.08)	(0.03)	(0.02)	(0.04)
White x Cohort Share Hispanic	-0.03	-0.01	-0.02	0.00	-0.01	0.01
	(0.07)	(0.02)	(0.05)	(0.02)	(0.01)	(0.02)
Full Controls	X	X	X	X	X	X
R^2	0.13	0.10	0.13	0.06	0.02	0.05
Num. obs.	962	962	962	962	962	962

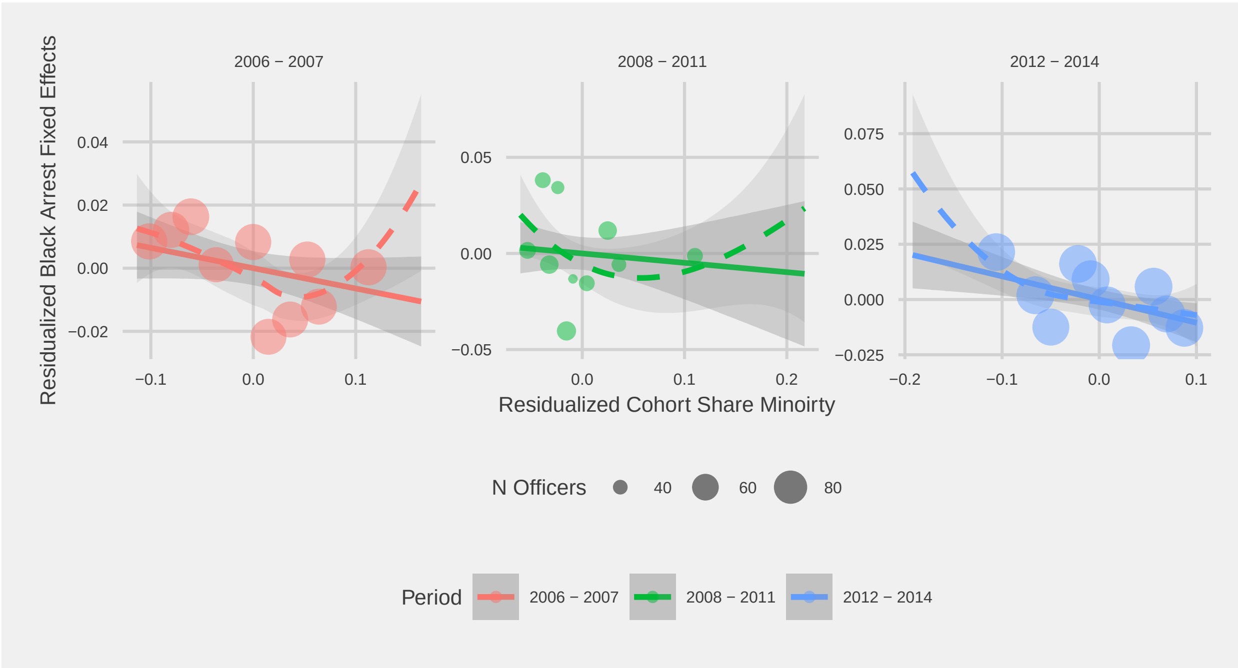
Note: Table displays the effect of cohort diversity on officer propensity to arrest Blacks for non-index and index crimes that resulted in any, guilty, or non-guilty outcomes. The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4), with Full Controls referring to the specification in Column (3) of Table 5, and additional controls for an interaction between officer race being white and cohort share other (Asian/Native American) race. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Figure 1: Distribution of Main Sample Officer Fixed Effects



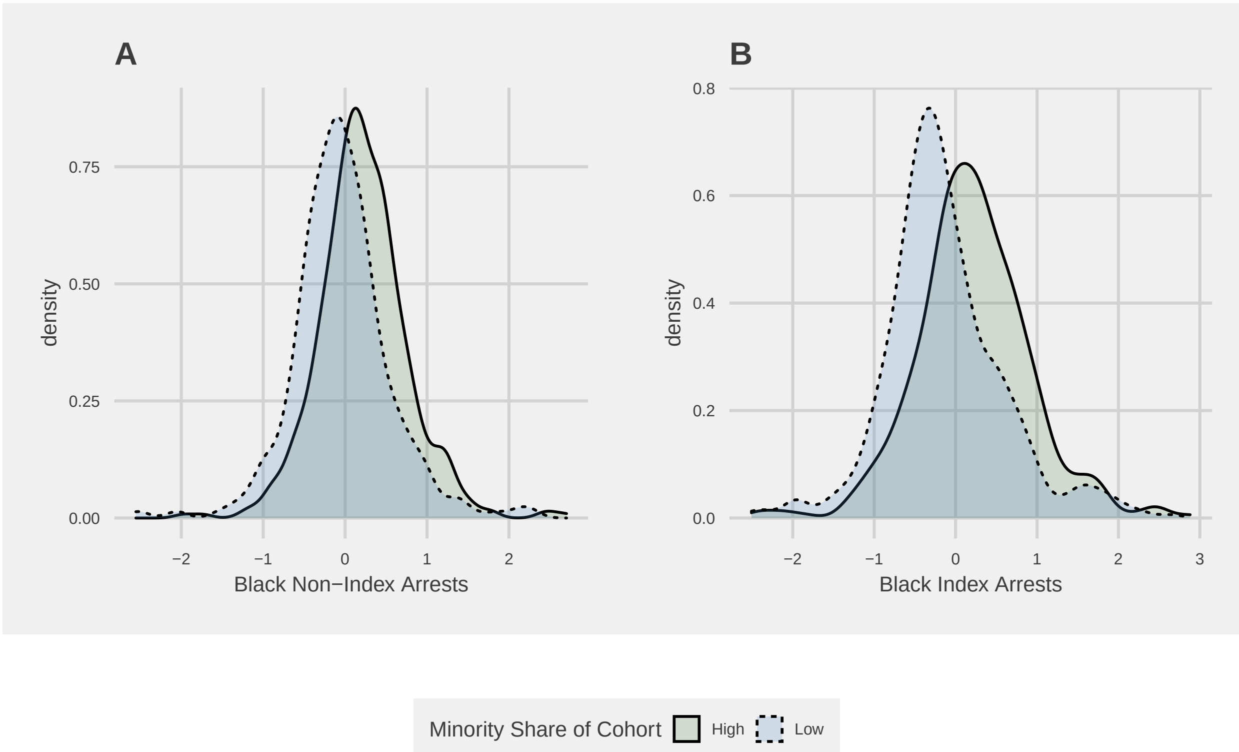
Note: Figure displays the distributions of main sample officer (July 2012 to May 2014 cohorts) fixed effects recovered from estimating equation 3 with arrests of Blacks as the dependent variable. Panel A displays the distributions for white and minority officers separately, and it shows that white officers tend to have higher fixed effects for Black arrests. Panel B displays the distributions of white officers split by on whether they were in a high (at least 50%) or low (below 50%) minority cohort, and it shows that whites in high−minority cohorts tend to have lower fixed effects compared to whites in low−minority cohorts. Displayed fixed effects are generally negative due to the leave−out officer in the first stage having an arrest propensity higher than most main sample officers.

Figure 2: Linear and Non-Linear Effect of Cohort Diversity on Arrest Propensity



Note: Figure displays the relationships between residualized cohort share minority (x−axis) and residualized fixed effects for black arrests (y−axis) for officers in the 2006−2007, 2008−2011, and 2012−2014 (main sample) periods. The solid line is the linear fit for the residuals and the dashed line is the smoothed non−parametric fit of the residuals. Residuals themselves are not shown; the dots are diplayed for simplicity, and they are constructed as, within each period, residualized X's are broken into 10 decile−groups and the y−value is the mean of the residualized Y's within the group−− the size of the dots is determined by the number of officers in each decile group, hence why they are the same size within each period. Residuals are computed using equation 1, with controls being the variables used in Column (2) of Table 5 with the addition of period fixed effects, excluding cohort share minority, and fixed effects for arresting Blacks (recovered from equation (3)) being the dependent variable.

Figure 3: Distributions of Officer Arrest Quality



Note: Figure displays the distributions of new officer (main sample cohorts) arrest quality fixed effects for arrests of Blacks for non−index crimes (Panel A) and index crimes (Panel B), with distinct distributions based on whether or not the officer was in a high (at least 50%) or low (below 50%) minority cohort. Both plots display the distributions of standardized fixed effects recovered from estimating equation (5). It shows that officers in high−minority cohorts tend to have higher individual arrest quality. The fixed effects recovered in Panel B exclude observations more than three standard deviations from the mean.

A Appendix A

A.1 Entrance into the CPD Police Academy

In order to become an officer in the CPD, applicants must first meet multiple qualifications before applying to take the entrance exam. For example, by the time of hiring, one must be a US citizen, a resident of Chicago, have sufficient credit hours at a college or university, and meet the age requirement (Pritchard (2013)). Potential applicants meeting these qualifications can apply to take the CPD entrance exam, and they will be notified of the test date and location after the application period ends ("CPD 2017 FAQ" 2017).

Applicants who pass the written exam are then assigned a random lottery number indicating their order of being called into the academy. Random assignment to the academy was not always the case; it was introduced in the early 1990's in an attempt to increase diversity (Kass and Blau (1991)). After an applicant's number is drawn, they must pass a background check, drug screening, and medical, psychological and physical exams (Pritchard (2013)). Upon passing these requirements, potential officers are admitted into the academy.

There are usually tests once every 2 or 3 years (not including makeup exams)—but in 2006 there were three exams issued. Generally, thousands of people take the CPD's written exam and a large portion of them meet the minimum passing score (see Figure B.2). Given the large number of passing applicants, many do not ever have their numbers called before the applicant list is retired. Despite my best efforts, I have not been able to obtain any indication of when the applicant lists are retired (according to the CPD such documentation may not even exist). Also, applicants from a test are likely to be admitted possibly years after they took the test initially, and their entrance into the academy likely occurs while more applicants are taking a new test. This makes identifying which cohorts come from which tests (i.e. the pool from which officers are randomly assigned) difficult.

The main sample (July 2012 - May 2014) cohorts, I believe, are an exception, and these cohorts all came from the same test issued in December of 2010 (see Figure B.2). The December 2010 test was the last test issued before the December 2013 test. The only sizable cohort to enter in 2011 was on October 17, 2011, then about 8 months pass until the first sizable cohort of 2012 started on July 02, 2012, with a total of 7 sizable cohorts starting between July and December of 2012, and then there is continuous cohort intake until May of 2014, when there is a 3 month gap until the next cohort. Given that it takes time for the CPD to draw in passing recruits and give them their multiple examinations, I believe that the main sample cohorts were all drawn from the December 2010 test.

Further supporting this is the change in the composition of cohorts before and after

2012. As shown in Panel A of Figure B.3, the 2011 cohort has a higher share Black than any cohort in the 2012-2014 period, while it is within the range of the 2008-2011 cohorts (likely drawn from the 2006 tests). Then the 2006-2007 cohorts (likely drawn from the 2005 test) have a lower share Black than the 2008-2011 cohorts. Similar patterns emerge when looking at share of the cohort which speaks Spanish (see Panel B of Figure B.3). While I am highly confident in main sample cohorts being drawn from the 2010 test, I am not as confident in the separation of the 2008-2011 and 2006-2007 periods into the 2006 tests and 2005 test, respectively. As I am less confident the further back I move in the data, the main sample cohorts are ideal for my analysis. While I use the 2006-2007 and 2008-2011 cohorts in robustness checks, I cannot be as certain that they are drawn from the same testing pools or from overlapping ones. This is also another reason for using the two-step method, as when combining first and second stage equations from Section 4, as done in Section 5.11 and Appendix A.6, the cohort periods for pre-2008 cohorts are very rough guesses as to which cohorts are possibly from the same test.

I am confident, however, that after May 2014, the cohorts until February 2015 (the last cohort I use in the extended sample) are from the 2013 test. The 2013 test recruits had the new feature that they were permitted to begin the academy at the age of 21, lower than the previous requirement (Pritchard (2013)). As can be seen in Panel C of Figure B.3, the lowest starting age per cohort drops to 21 after the May 2014 cohort. Unfortunately, due to my panel data ending in 2016, I cannot fully exploit the 2013 test cohorts and only use them briefly in robustness checks.

A.2 Random Assignment

Since when a recruit can enter the academy is determined by a random lottery number, cohorts are as-good-as-randomly assigned, and I provide brief empirical evidence for this. Table B.1 displays the results of regressing an officer's cohort characteristics on their individual characteristics. For average cohort start age (Column (1)), I control for years since 2012, as the cohorts start successively and that means the age of officers will change simply due to the passage of time (excluding this results in a much lower R^2). Unsurprisingly, an officer's own age at the start of the academy is positively associated with their cohort's average start age as is years since 2012. Officer gender is positively associated with cohort age, implying female applicants are more likely to select out of joining the CPD if their lottery number is further into the pool. However, the coefficient for being male is very small relative to the intercept (less than 1%). I perform robustness checks against this selection out of the pool of applicants over time in Section 5.6.

Columns (2) and (3) look at the relationship between officer characteristics and their cohort's share minority and male, respectively. As expected, virtually none of the variation in cohort composition is explained by officer characteristics (both R^2's are less than 2%); the only statistically significant coefficient is the corresponding officer characteristic, and it is negative (as an officer being male, for example, means there are fewer males to comprise the rest of their cohort); and these coefficients are economically insignificant, at about 1% of the intercept for race and 0.1% for gender. Furthermore, the p-value of a joint F-test resulting from a multinomial logit (regressing assigned cohort on an officer being male and an officer being a minority) is insignificant (p=0.816).

A.3 Attrition

If attrition is not random and is impacted by diversity of one's cohort, then results in my estimation may be driven by selection bias rather than actual peer effects. In Table B.2, I present results for logistic regressions where each outcome is a form of attrition, and the final column contains a dummy for any attrition form (i.e. whether or not a fixed effect is recovered for the officer) for officers in the main sample cohorts.

Column (1)'s outcome is whether the officer is not in the daily assignment (AA) data (56 recruits). Column (2)'s outcome is whether the officer, conditional on being in the AA data, spent too much or too little time in the academy or probationary period (51 recruits). Column (3)'s outcome is whether the recruit was not in the final AA data, conditional on the previous two restrictions, meaning they were not matched to the salary and rank data as a police officer (68 recruits). Finally, Column (4) pools all recruits in the sample cohorts and looks at any form of attrition, including whether or not fixed effects could be recovered. As displayed across all columns, there is no statistically significant predictor of any form of attrition with respect to cohort composition (neither cohort diversity nor mean age), thus it is unlikely that attrition driven selection is driving my results.

Another form of attrition is attrition from the final sample, e.g. cohort diversity being related to when officers choose to retire or exit the sample. While this may cause some officers to be more represented in the sample than others, the fixed effects recovered for the main sample are based on over 100 observations for almost all officers ($> 93\%$). I test for sample-exiting attrition in Table B.3. Columns (1)-(2) and Column (4) study the relationship between cohort share of Blacks, Hispanics, and other non-whites and officer's number of observations in the assignment data used to estimate fixed effects. For the main sample (2012 to 2014 cohorts), Column (1) shows that there is no economically or statistically significant relationship between cohort diversity and observations– the main factor is the start date of the officer as that directly influences the number of observations in the panel. Column (2) repeats this exercise but on the main sample cohorts which start before 2013 (the sample used in Column (7) of Table B.10) and does not control for start date (as these cohorts started within a short timeframe), and the effects of cohort composition are generally not statistically significant and very noisy. Column (3) shows that cohort diversity has no effect on whether or not the officer exists in the salary and unit history data (which contains officers not in the assignment data) at the end of 2016.

Columns (4)-(7) conduct further checks on the earlier 2006-2007 cohorts. Column (4) shows that cohort shares of minorities have economically small effects. The small but statistically significant relationship may be a result of more experienced officers who have a lower preference for enforcement (more minority peers leading to lower arrest fixed effects) bidding for different assignments or being placed on desk duty more frequently. Columns (5), (6), and (7) show that cohort shares of minorities have no statistically or economically significant effect on an officer's likelihood of being in the salary and unit history data (which includes units and ranks outside of the assignment data), being promoted above D1, or being in a specialized unit (outside of the assignment data).

A.4 Confounding Assignments

Based on Table 2, it would appear that the share of a new officer's arrests made up of arrestees of a specific race is related to that officer's cohort share minority. This phenomena is partially a result of the CPD officer assignment policy—the units (districts) officers are assigned to work in being dependent on departmental demand, the seniority based assignment bidding process (older officers do not want to work in high crime areas), *and* a desire to match officers to racially similar civilians. As a result, it is possible that assignment is influenced by one's cohort's composition, since if the department attempts to place Black officers in districts with Black civilians, having more Black cohort-mates may reduce another officer's probability of being placed in a Black district. In this subsection, I explore the effect of peer diversity on assignment.

The CPD assigns new officers to units based on need and a few other considerations, such as avoiding unreasonable commutes for officers and not placing them too close to their own homes.[36] How need is determined is not clear, and there does not seem to be a significant correlation between a recent decline in the number of officers in a unit and the share of new recruits serving in that unit. In the data there is a clear relationship between the race of officers serving in a district and that of the district's population.

[36]This information is based on conversations with a retired officer.

Table B.4 displays the characteristics of the average unit (district) in which officers work. These simple regressions explain some amount of variation in officer assignments. Based on the results, new officers are much more likely to be placed in high crime areas both during and after their probationary periods—this may be partially explained by those units demanding the most officers and higher seniority officers transferring to less dangerous areas. Gender, which is potentially correlated with underlying officer preferences for enforcement, does not influence assignment. As expected, assignments are influenced by officer race, with a positive relationship between own race and racial composition of one's working environment. For example, being Black increases the share of Black civilians in the average district in which an officer works by almost 30pp. Similarly, there is a clear relationship between officers being Black or Hispanic and their districts having higher crime and lower income.

Focusing on the main sample of new officers (July 2012 - May 2014 cohorts), Table B.5 displays regression estimates predicting the characteristics of the average unit a new officer serves in as a full officer. Notably, having more Black and Hispanic peers in one's academy cohort decreases the Black population share of the average district in which an officer works. As is evidenced from the table, Black officers are more likely to serve in Black districts, thus it is likely the case that having more Black officers in one's cohort crowds out the potential for non-Black officers to be placed in Black districts.

However, this confounding assignment is unlikely to significantly bias estimates because new officers work in high crime areas with higher Black populations regardless of race, as shown in Table B.4. Despite this, new officers in the work in all of the 22 district during the 2013-2016 period and no single district makes up more than 12% of assignments. Furthermore, it is evident from the very small amount of variation explained by observables in Table B.5 that there is likely much CPD-level demand choices made regardless of cohort observables and the actual influence of such observables are economically small. For example, a 5pp increase in cohort share Black (going from the 30th to 70th percentile cohort) only decreases a recruit's average district's Black population share by 8% of the baseline mean and decrease their average district's violent crime by 2% of the baseline mean. Given this, it is unlikely that the influence of cohort composition on new officers' assignments will significantly bias results, and controlling extensively for working environment, as in Section 4, should remove this bias completely.

A.5 Working Peers and Instructors

Another concern is that officers exposed to higher amounts of minorities in their cohorts may end up working with more minorities in the future. If contemporaneous peers influence arresting decisions, then the effect of academy peers may be capturing the selection of future minority peers and their influence. A similar concern is that academy diversity may influence the composition of one's field training officers. To test these, I regress the average composition of officers' contemporaneous peers and field training officers on their cohort composition. Table B.6 displays these results.

In Columns (1) to (4), I use two main peer groupings, officers assigned to the same shift and watch number and the same sector (a subset of a district composed of multiple beats) or same beat, with Columns (1) and (3) being during the officer's probationary period and (2) and (4) being during the officer's time as a full officer. The dependent variable in Column (5) is the share of an officer's field training officers who are white during their probationary period. Relative to the mean white share across each group and officer type (probationary and full), the relationship between an officer's cohort share minority and their future peers' and FTOs' share white is not economically significant, and only one is marginally statistically significant. The effect size with the largest magnitude indicates that a 10pp increase in cohort share minority leads to a 0.02 decrease in the peer share white, which is a 4.41% decrease relative to the mean. The small and noisy, but consistently negative, relationship between cohort share minority and the share of white working peers/FTOs is likely an artifact of this model not controlling for unit assignment: more minority peers slightly crowd out positions in high Black areas which also have slightly fewer white officers. But, as discussed in Appendix A.4, this effect is minor and almost all new officers go to high crime and high share Black districts. Given this, it is unlikely that the effect of future peer or training officer composition is driving the effects, and much of these small and noisy effects can be explained by the weak influence of cohort diversity on unit assignment.

A.6 Combined First and Second Stage

In this section, I provide more detail on the combined first and second stage analysis discussed in Section 5.11. As noted in Section 4.1, the two-step method bypasses many issues. As such, a combined first and second stage regression is plagued by weighting issues as most officers in the 2010 - 2016 panel started their careers long before 2010 and are thus observed more frequently than the newer officers who are the focus of my study (particularly the main sample cohorts). Even within the main sample cohorts, officers starting earlier will contribute more to the coefficient than later officers. Furthermore, even the precise cohorts older officers belong to is not always clear (there are more scattered start dates as one goes back into the data). I attempt to suppress the weighting issues across cohorts by interacting the coefficient of interest, Cohort Share Minority, with a fixed effect for period in which the cohort began. As previously discussed in Appendix A.1, I cannot be certain which cohorts belong to which tests for cohorts other than the main sample, so using periods as proxies for testing groups becomes increasingly inaccurate for older cohorts.

In preparation, I construct cohorts for all officers in the 2010-2016 daily panel making monthly cohorts for them which include all officers who have appointed dates in the same appointed month based on roster data (roster data includes officers who do not appear in the panel for a variety of reasons such as retirement or assignment to other units). Then I compute cohort mean start age, cohort share minority, and cohort size (as was done in the main part of this paper) using the monthly cohorts. Then I group these cohorts into 8 periods based on starting months: pre-1992, 1992-1995, 1996-1999, 2000-2001, 2002-2005, 2006-2007, 2008-2011, and 2012-2014 (July 2012 to May 2014). I drop any officers who are recorded to start before the age of 18 or after the age of 40. Then I merge this cohort data to the daily panel data used to recover fixed effects in the Section 4. In the panel data, I drop officers in cohorts with 20 or less officers, any officers who had 1.5 years or less of tenure, and the cohorts starting after May of 2014—these additional filters account less than 5% drop in observations. Note that the main sample now includes some officers who were previously dropped, such as those who had training time violations or insufficient observations in the panel data.

With this panel, I estimate the equation:

$$Arrest_{it}^k = \pi_0^k X_i + \pi_1^k(\overline{X}_{c(i)} * \alpha_{cp}^k) + \alpha_{cp}^k + \gamma_{bsw_t}^k + \beta^k V_{it} + \epsilon_{it}^k \tag{7}$$

where X_i contains officer race, gender, start age, and cohort size; α_{cp}^k is a fixed effect for the period to which cohort c and officer i belongs; $\overline{X}_{c(i)} = [\overline{Minority}_{c(i)}, \overline{StartAge}_{c(i)}]$ contains the (monthly) cohort share minority and mean start age; $\gamma_{bsw_t}^k$ is the high-dimensional assignment fixed effect, and V_{it} contains the second degree polynomials of watch duration, crime, and tenure—the same as those used in the main analysis in Section 4.

The coefficients of interest are contained in π_1^k, which provides the peer effect of cohort diversity on arrests of type k for each starting period. Standard errors are clustered at the assignment level (i.e. at the bsw_t level).

Figure B.4 displays these coefficients by cohort period. The coefficients for the main sample are smaller than in the main analysis, but display the same pattern with respect to crime severity and arrest quality: the reduction is driven by a decline in arrests of Blacks for low-level crime with a small positive effect on arrests of Blacks for serious (index) crimes; arrest quality increases as the decline in Black arrests for low-level crimes is due to a decline in non-guilty arrests not guilty arrests leading to an increase in quality.

Also visible in the figure is the changing effect of diversity by cohort period. There is a trend upward for the effect on minority arrests, though the only economically significant coefficient that is positive is for the cohorts starting in 1992-1995. Given the selection of older officers into the panel (higher tenure allows for officers to transfer out of this data) in addition to other issues such as imprecise start dates and mismatch between tests and periods, the estimates for pre-2008 cohorts should be not be considered causal and should be interpreted with significant caution.

A.7 Tipping Points

There is significant discussion of 'tipping points', cutoff value for some characteristic at which agents alter their decision making, in the literature neighborhood diversity (Card, Mas, and Rothstein (2008), Blair (2017)). In the peer effects literature, for example, Anwar, Bayer, and Hjalmarsson (2012) find having a single Black member of a jury pool eliminates the Black-white gap in convictions. Particularly in the policing environment, we would expect that if peer diversity is influencing how others behave, it would be most influential when minorities are less represented.

To test for tipping points in my environment, I create a cut-off value, c, of Cohort Share Minority (CSM) at 45%, 50% and 55%, then I create a variable Cohort Share Minority Post-Cutoff (CSMPC)=$1\{CSM \geq c\} * (CSM - c), \forall c \in \{0.45, 0.5, 0.55\}$, and I use these as variables of interest in my second stage. Table B.13 displays the results. The effect of cohort share minority before the cutoff is contained in the variable "Cohort Share Minority" and the effect of cohort share minority after the cutoff is computed by adding "Cohort Share Minority" to "CSM Post-Cutoff". For $c = 0.45$, we see the largest initial effect of cohort share minority, and setting $c = 0.5$ still has a large and negative effect, but the magnitude is about 30% smaller than when $c = 0.45$; for $c = 0.55$, the magnitude is even smaller, less than half the size of the coefficient when $c = 0.45$. The slopes postcut off indicate that for $c = 0.45$, cohort share minority has a negative effect post-cutoff, while for $c = 0.5$ and $c = 0.55$ the effect is positive (and noisy). These results indicate that the observed effects of diversity is driven by increasing minority share in low-diversity cohorts.

A.8 Representation and Diversity

Thus far, I have used increased minority representation and diversity interchangably. However, distinguishing between the two is useful in further exploring the mechanisms driving these results. Increasing minority representation is achieved, for example, by moving from a cohort with 50% minority to 70% minority; however diversity is a measure of how diverse cohorts are, i.e. diversity is racial entropy and is maximized when a cohort is equally composed of all racial groups. To distinguish between the two, I substitute my metric of cohort diversity as cohort share minority (Black, Hispanic, and other) with cohort diversity as racial entropy. From Currie and MacLeod (2020), I define the racial entropy of a cohort as $Entropy_{c(i)} = \sum_{t \in \{WO,B,H\}} p_{c(i),t} \frac{1}{log(p_{c(i),t})}$, where $p_{c(i),t)}$ is the share of the cohort of racial group t excluding officer i and B = Black, H=Hispanic, and WO= white or other (other is grouped with white due very low shares and occasional zero shares). I also use the classic Herfindahl–Hirschman Index (HHI) which measures the concentration of one group in a cohort: $HHI_{c(i)} = \sum_{t \in \{WO,B,H\}} p^2_{c(i),t}$.

Table B.14 displays the results for the effect of racial diversity (as entropy and concentration) on officer propensities to arrest Blacks for non-index and index crimes. The entropy results are in Columns (1)-(4), with Column (1)-(2) use three racial groups (Black, Hispanic, white/other) and Columns (3)-(4) use two (minorities and whites), and the HHI results in Columns (5)-(6) use four groups (Black, Hispanic, White, other)– coefficients for HHI are very similar to those of negative entropy if the two groups have close to even shares. The entropy results are noisy, but in the same direction as the main results (negative effect on non-index, positive effect on index). The HHI results are statistically significant and also consistent with the main results: higher concentrations of one group lead to higher arrest propensities for non-index crimes and lower ones for index crimes. However, given that whites are the plurality in all but one cohort, this result is not distinguishable from the stronger effect of minorities on peers in cohorts with fewer minorities.

B Appendix B - Additional Tables and Figures

Table B.1: Balance Regressions

	Cohort Average Start Age	Cohort Share Minority	Cohort Share Male
	(1)	(2)	(3)
Minority	-0.014	-0.005	0.002
	(0.023)	(0.003)	(0.002)
Male	0.092^{***}	0.005	-0.009^{***}
	(0.032)	(0.004)	(0.003)
Start Age	0.008	0.000	0.001^{*}
	(0.007)	(0.000)	(0.000)
Years since 2012	0.413		
	(0.270)		
Intercept	29.143^{***}	0.483^{***}	0.788^{***}
	(0.452)	(0.019)	(0.017)
R^2	0.104	0.003	0.015
Num. obs.	1139	1139	1139

Note: Table displays the effect of officer characteristics on the average characteristics of their cohort (excluding themselves) for officers starting between July 2012 and May 2014. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.2: Attrition from Main Sample

	Not in AA	Training Time Violation	Not in Final AA	Any Attrition
	(1)	(2)	(3)	(4)
Cohort Share Minority	2.25	2.97	−2.58	0.13
	(3.45)	(2.29)	(4.97)	(2.86)
Cohort Mean Age	0.02	0.07	0.84	0.33
	(0.32)	(0.18)	(0.50)	(0.24)
Black	−0.31	0.47	-1.35^{**}	−0.35
	(0.46)	(0.49)	(0.50)	(0.38)
Hispanic	0.38	0.71^{*}	0.07	0.38^{*}
	(0.31)	(0.33)	(0.23)	(0.18)
Other Race	−0.79	1.11	−1.09	−0.07
	(0.93)	(0.62)	(0.76)	(0.45)
Male	−0.10	−0.12	−0.23	−0.23
	(0.37)	(0.40)	(0.22)	(0.20)
Start Age	0.00	−0.02	0.01	−0.00
	(0.02)	(0.05)	(0.03)	(0.02)
Cohort Size	−0.00	-0.02^{**}	−0.03	−0.02
	(0.01)	(0.01)	(0.02)	(0.01)
Intercept	−4.63	−5.11	−25.37	−10.66
	(10.22)	(5.39)	(14.09)	(6.72)
AIC	459.23	416.50	469.72	963.89
Log Likelihood	−220.61	−199.25	−225.86	−472.95
Num. obs.	1139	1083	1032	1139

Note: Table display the logistic regression estimates of cohort and officer observables on officer attrition for various reasons from the main sample (July 2012 - May 2014 cohorts). The dependent variables for Columns (1)-(3) are: (1) whether or not the officer is not matched in the assignment data; (2) whether or not the officer is dropped due to spending too much or too little time in the academy or probationary period; (3) whether or not the officer is not in the final assignment data, meaning they were matched in the salary and unit history data and spent some time as a D1 officer in units 1-25; (4) attrition for any of the listed reasons or if no fixed effects could be recovered due to too few observations. Columns (2)-(3) are estimated on the sample of recruits which were not dropped due to the previous column's reason, and Column (4) is estimated on all initial sample recruits. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table B.3: Attrition out of Sample

	2012-2014 Cohorts			2006-2007 Cohorts			
	N. Obs in Data		Exit Data	N. Obs in Data	Exit Data	Promoted at End	Specialized Unit at End
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cohort Share Black	-48.11	-3159.27	0.15	-252.38^{**}	-0.09	-0.03	-0.09
	(196.42)	(1725.04)	(0.14)	(88.20)	(0.09)	(0.13)	(0.19)
Cohort Share Hispanic	-114.94	-1569.18	0.09	-99.13	-0.09	0.18	0.18
	(136.05)	(904.55)	(0.07)	(153.92)	(0.15)	(0.16)	(0.18)
Cohort Share Other	-91.43	-3036.75^{*}	-0.21	-1223.20^{***}	-0.38	0.16	0.10
	(285.62)	(1477.10)	(0.13)	(233.68)	(0.24)	(0.45)	(0.45)
Cohort Mean Age	16.14	179.80	-0.00	-2.19	0.00	-0.01	0.01
	(10.79)	(102.36)	(0.01)	(14.91)	(0.01)	(0.02)	(0.02)
Black	-35.81^{***}	-98.47^{***}	-0.00	-23.80	0.03^{*}	-0.02	0.02
	(9.14)	(19.15)	(0.01)	(33.94)	(0.02)	(0.02)	(0.03)
Hispanic	-9.61	-34.61	-0.04^{*}	48.53	0.03	-0.00	-0.04^{*}
	(8.48)	(17.62)	(0.02)	(35.76)	(0.02)	(0.02)	(0.02)
Other Race	2.92	-41.03	-0.01	121.80^{*}	0.06^{***}	-0.02	0.00
	(19.03)	(27.20)	(0.02)	(47.55)	(0.01)	(0.04)	(0.04)
Male	64.60^{***}	92.70^{**}	0.01	187.48^{***}	-0.01	0.02	0.00
	(15.05)	(33.24)	(0.01)	(34.09)	(0.01)	(0.02)	(0.02)
Start Age	1.56	5.56	0.00^{*}	10.40^{***}	0.00	-0.01^{***}	-0.01^{***}
	(1.35)	(3.27)	(0.00)	(2.93)	(0.00)	(0.00)	(0.00)
Cohort Size	0.28	6.32	-0.00^{***}	1.13^{*}	0.00	0.00	-0.00
	(0.40)	(4.08)	(0.00)	(0.47)	(0.00)	(0.00)	(0.00)
Start Date	-0.51^{***}			0.02			
	(0.05)			(0.05)			
Intercept	7799.65^{***}	-4592.15	0.99^{***}	367.29	0.84^{**}	0.31	0.19
	(869.64)	(2798.49)	(0.20)	(981.47)	(0.26)	(0.45)	(0.55)
Mean Dep. Var	334.79	409.51	0.97	892.45	0.95	0.07	0.09
R^2	0.43	0.14	0.02	0.07	0.01	0.02	0.03
Adj. R^2	0.42	0.12	0.01	0.06	0.00	0.00	0.02
Num. obs.	964	400	964	905	905	863	863

Note: Table display the linear regression estimates of cohort and officer observables on officer observations and other measures of attrition for the main sample (July 2012 - May 2014 cohorts) and 2006 - 2007 cohorts. The dependent variables are the officer's number of observations (shifts) used to estimate fixed effects in the daily panel data (Columns (1), (2), (4)), whether or not the officer is in the salary and unit history data which contains non-D1 officers and units outside of the assignment data (Columns (3) and (5)), whether the officer has been promoted by the end of 2016 (Column (6)), whether the officer is in a specialized unit at the end of 2016 (Column (7)). Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table B.4: Characteristics of Average Working District

	Violent Crime	Median Income (2010)	Share Black Pop.	Share Hispanic Pop.
	(1)	(2)	(3)	(4)
Probationary	3.07^{***}	-9921.15^{***}	0.19^{***}	-0.02^{***}
	(0.14)	(344.50)	(0.01)	(0.00)
Recruit	5.14^{***}	-12825.76^{***}	0.27^{***}	-0.05^{***}
	(0.15)	(384.64)	(0.01)	(0.01)
Female	-0.23	497.24	-0.01	-0.00
	(0.14)	(401.57)	(0.01)	(0.00)
Black	3.53^{***}	-5747.59^{***}	0.27^{***}	-0.14^{***}
	(0.15)	(463.62)	(0.01)	(0.00)
Hispanic	0.40^{***}	-1926.27^{***}	-0.00	0.04^{***}
	(0.14)	(380.99)	(0.01)	(0.01)
Other Race	-0.90^{***}	1348.46^{*}	-0.06^{***}	-0.00
	(0.32)	(788.83)	(0.02)	(0.01)
Intercept	8.06^{***}	52033.77^{***}	0.32^{***}	0.27^{***}
	(0.10)	(319.23)	(0.01)	(0.00)
R^2	0.16	0.12	0.21	0.10
Num. obs.	9621	9621	9621	9621

Note: Table displays the linear regression estimates of officer characteristics on the average characteristics of the districts in which they work. Population and income are determined based on the 2010 Census and 2014 ACS. Violent crime rates are violent crimes in a month, based on Chicago City Data Portal crime data, per 10,000 population, based on the 2010 Census. The coefficients Recruit and Probationary are indicators for whether or not the officer is a new officer in their post-probationary period or a new officer in their probationary period, respectively. Robust standard errors are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.5: Characteristics of Average District in Post Probationary Period

	Violent Crime	Median Income (2010)	Share Pop. White	Share Pop. Black	Share Pop. Hispanic
	(1)	(2)	(3)	(4)	(5)
Cohort Share Black	-9.69	5288.64	0.10	-0.99^{**}	0.83^{***}
	(8.12)	(8085.86)	(0.10)	(0.40)	(0.30)
Cohort Share Hispanic	-2.43	-550.41	0.02	-0.60^{**}	0.55^{***}
	(4.69)	(4669.32)	(0.06)	(0.25)	(0.19)
Cohort Share Other	10.80	-1651.93	-0.07	0.29	-0.18
	(11.07)	(8800.31)	(0.11)	(0.52)	(0.40)
Cohort Mean Age	-0.18	528.59	0.00	0.02	-0.02
	(0.47)	(491.50)	(0.01)	(0.03)	(0.02)
Black	-0.29	1341.43	0.00	0.07^{**}	-0.07^{***}
	(0.69)	(1106.46)	(0.01)	(0.03)	(0.02)
Hispanic	-0.00	-81.74	-0.01	-0.00	0.01
	(0.53)	(729.79)	(0.01)	(0.03)	(0.02)
Other Race	0.40	-348.69	0.04	-0.05	-0.01
	(1.60)	(1576.91)	(0.03)	(0.06)	(0.04)
Male	0.08	620.10	0.01	0.00	-0.01
	(0.66)	(976.50)	(0.01)	(0.03)	(0.03)
Start Age	-0.12^{**}	79.51	0.00	-0.00^{*}	0.00^{**}
	(0.05)	(88.03)	(0.00)	(0.00)	(0.00)
Cohort Size	-0.04^{**}	33.15^{*}	-0.00	-0.00	0.00
	(0.02)	(19.86)	(0.00)	(0.00)	(0.00)
Intercept	28.17^{*}	12433.65	-0.09	0.61	0.48
	(14.78)	(14919.29)	(0.19)	(0.82)	(0.59)
Mean Outcome	15.416	33781.293	0.079	0.711	0.18
R^2	0.02	0.01	0.01	0.03	0.05
Num. obs.	962	962	962	962	962

Note: Table displays the linear regression estimates of recruit and cohort characteristics on the average characteristics of the districts in which they work after their probationary period. Population and income are determined based on the 2010 Census and 2014 ACS. Violent crime rates are violent crimes in a month, based on Chicago City Data Portal crime data, per 10,000 population, based on the 2010 Census. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.6: Average Characteristics of Peers and Training Officers

	Average Sector Share White		Average Beat Share White		FTO Share White
	Probationary	Full	Probationary	Full	Probationary
	(1)	(2)	(3)	(4)	(5)
Cohort Share Minority	0.00	−0.13	-0.15^{*}	−0.21	0.03
	(0.07)	(0.13)	(0.08)	(0.16)	(0.14)
Cohort Mean Age	0.01^{*}	0.02^{**}	−0.00	0.00	−0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Black	-0.10^{***}	-0.15^{***}	-0.13^{***}	-0.33^{***}	-0.14^{***}
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)
Hispanic	-0.03^{***}	-0.06^{***}	-0.05^{***}	-0.17^{***}	-0.04^{*}
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Other Race	0.05^{***}	−0.03	0.03	-0.13^{***}	0.07
	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)
Male	−0.00	0.05^{***}	0.00	0.08^{***}	−0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)
Start Age	0.00	-0.00^{**}	−0.00	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cohort Size	−0.00	0.00	−0.00	-0.00^{*}	-0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Intercept	0.21	0.15	0.73^{***}	0.62	1.01^{**}
	(0.19)	(0.26)	(0.27)	(0.54)	(0.45)
Mean Dep.	0.48	0.48	0.46	0.48	0.43
R^2	0.08	0.18	0.06	0.24	0.04
Num. obs.	933	962	933	962	933

Note: Table displays the linear regression estimates of main sample officer characteristics on their working peers and training officers during their probationary (odd columns) and post-probationary (even columns) periods. The dependant variable in Columns (1) and (2) is the share of an officer's peers who are white and working the same day, shift, and sector. The dependant variable in Columns (3) and (4) is the share of an officer's peers who are white and working the same day, shift, and beat number. The dependant variable in Column (5) is the share of an officer's field training officers who are white. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.7: Effect of Cohort Diversity on Arrest Propensity by (Non-Black) Arrestee Racial Group

	Hispanic	White	Non-Black	Non-Black Non-Index	Non-Black Index
	(1)	(2)	(3)	(4)	(5)
Cohort Share Black	-0.06^{**}	-0.02	-0.07^{**}	-0.05^{*}	-0.02^{**}
	(0.02)	(0.01)	(0.03)	(0.03)	(0.01)
Cohort Share Hispanic	-0.02	-0.01	-0.03	-0.03	-0.01
	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)
Cohort Share Other	-0.05	-0.05^{**}	-0.08	-0.07	-0.01
	(0.05)	(0.02)	(0.06)	(0.06)	(0.01)
Cohort Mean Age	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black	-0.01^{***}	-0.00^{**}	-0.01^{***}	-0.01^{***}	-0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Hispanic	-0.00	-0.00^{***}	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other Race	-0.00	-0.00	-0.00	-0.00	-0.00^{*}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Male	0.00	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Start Age	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cohort Size	0.00	0.00^{**}	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Intercept	-0.03	-0.01	-0.03	-0.02	-0.01
	(0.05)	(0.02)	(0.06)	(0.06)	(0.02)
R^2	0.03	0.04	0.03	0.03	0.03
Num. obs.	962	962	962	962	962

Note: Table displays the effect of cohort diversity on standardized officer propensity to arrest Hispanics, whites, and all non-Blacks (white, Hispanic, other). The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4). Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.8: Effect of Cohort Diversity on Propensity for Arrest to be Guilty

	Black	
	Non-Index	Index
	(1)	(2)
Cohort Share Black	0.74^{***}	1.15^{***}
	(0.21)	(0.37)
Cohort Share Hispanic	0.50^{***}	0.54^{**}
	(0.18)	(0.25)
Full Controls	X	X
R^2	0.09	0.12
Num. obs.	929	926

Note: Table displays the effect of cohort diversity on officer's arrest quality (share of arrests during shift resulting in guilty outcome) for arrests of Blacks for index and non-index crimes. The officer arrest quality measures used in Columns (1) and (2) are from estimating equation (5) on arrests of Blacks for non-index and index crimes; observations are dropped if the dependent variable is more than 3 standard deviations from the mean. The parameter estimates are based on the specification in equation (4), with Full Controls referring to the specification in Column (3) of Table 5. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.9: Main Results (Table 5) with Shrunken Arrest Propensity

	Shrunken Arrest Propensity				
		Black		Black Non-Index	Black Index
	(1)	(2)	(3)	(4)	(5)
Cohort Share Minority	-0.133^{**}	-0.153^{**}			
	(0.065)	(0.069)			
Cohort Share Black			-0.150	-0.277^{*}	0.052^{***}
			(0.122)	(0.145)	(0.015)
Cohort Share Hispanic			-0.131^{**}	-0.196^{**}	0.021^{**}
			(0.061)	(0.081)	(0.008)
Cohort Share Other			-0.308	-0.358	-0.010
			(0.193)	(0.225)	(0.020)
Black		-0.030^{***}	-0.030^{***}	-0.029^{***}	-0.004^{**}
		(0.007)	(0.007)	(0.007)	(0.002)
Hispanic		-0.018^{***}	-0.018^{***}	-0.017^{***}	-0.003^{**}
		(0.005)	(0.005)	(0.005)	(0.001)
Other Race		-0.022^{**}	-0.022^{**}	-0.019^{**}	-0.004
		(0.009)	(0.009)	(0.008)	(0.003)
Male		0.025^{***}	0.025^{***}	0.023^{***}	0.005^{***}
		(0.005)	(0.005)	(0.004)	(0.002)
Cohort Mean Age	-0.005	-0.001	-0.002	0.002	-0.002
	(0.006)	(0.008)	(0.006)	(0.008)	(0.001)
Start Age		-0.002^{***}	-0.002^{***}	-0.002^{***}	-0.001^{***}
		(0.000)	(0.000)	(0.000)	(0.000)
Cohort Size		0.000	0.000	0.000	-0.000^{**}
		(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.008	-0.037	-0.016	-0.237	0.137^{***}
	(0.165)	(0.237)	(0.192)	(0.239)	(0.033)
R^2	0.019	0.115	0.119	0.138	0.054
Num. obs.	962	962	962	962	962

Note: Table displays the effect of cohort diversity on main sample officers' propensities to arrest Blacks for all (Columns (1)-(3)), low-severity (Column (4)), and high-severity (Column (5)) crimes. The propensity is captured by officers' fixed effects using equation (3), the fixed effects are then shrunken by a factor between (0,1) that is decreasing in standard error of the fixed effect estimate (see Section 5.1). The parameter estimates are based on the specification in equation (4). Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.10: Effect of Cohort Diversity on Arrest Propensity - Alternate Samples and Specifications

	Poisson	LPM	Probation	High Crime	2006-2007 Cohorts	Restricted Sample	First Cohorts	First Arresting PO	No Crime Controls	Car Only	Assignment Type	Expanded Main Sample
						Panel A - Black Non-Index Arrests						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Cohort Share Black	-3.05^{**}	-0.21^{*}	-0.23^{***}	-0.30^{***}	-0.11^{*}	-0.20^{**}	-0.21^{**}	-0.15^{*}	-0.26^{*}	-0.27^{*}	-0.25^{*}	-0.24^{*}
	(1.47)	(0.12)	(0.09)	(0.10)	(0.06)	(0.08)	(0.08)	(0.09)	(0.14)	(0.14)	(0.13)	(0.13)
Cohort Share Hispanic	-1.80^{*}	-0.15^{**}	-0.09^{*}	-0.14^{**}	-0.22^{**}	-0.12^{***}	-0.23^{***}	-0.13^{***}	-0.19^{**}	-0.19^{**}	-0.18^{**}	-0.18^{**}
	(0.92)	(0.06)	(0.05)	(0.06)	(0.09)	(0.04)	(0.06)	(0.05)	(0.08)	(0.08)	(0.08)	(0.08)
R^2	0.13	0.14	0.07	0.13	0.08	0.11	0.16	0.04	0.13	0.14	0.13	0.13
						Panel B - Black Index Arrests						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Cohort Share Black	3.29^{***}	0.06^{***}	0.13^{***}	0.06^{*}	0.03^{**}	0.07^{*}	-0.03^{*}	0.04^{**}	0.10^{***}	0.09^{***}	0.08^{***}	0.07^{***}
	(0.71)	(0.02)	(0.04)	(0.03)	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Cohort Share Hispanic	1.71^{***}	0.02^{**}	0.03	0.06^{**}	0.02	0.02	-0.03^{**}	0.00	0.05^{***}	0.04^{***}	0.04^{***}	0.03^{***}
	(0.54)	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Controls	Full	Full	FTO Demos	Full	Full	Full	No Cohort Size	Full	Full	Full	Full	Full
R^2	0.09	0.05	0.02	0.06	0.07	0.04	0.07	0.01	0.06	0.05	0.05	0.05
Num. obs.	950	962	933	534	905	899	399	962	962	962	962	1011

Note: Table displays the effect of cohort diversity on officer propensity to arrest Blacks for non-index (Panel A) and index crimes (Panel B). The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4) (unless otherwise specified), with controls denoted as Full referring to the specification in Column (3) of Table 5, and additional or removed controls denoted in the table. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level (unless otherwise specified) are in parentheses.

Columns (1),(2): Results from estimating officer fixed effects using equation (6) and modifying equation (3) with the dependant variable (arrests of type k) being whether (1) or not (0) the officer made at least one arrest of type k during their shift. Column (3) : Results are from estimating equation (4) on fixed effects of new officers recovered during their probationary periods only, with additional controls for share of Field Training Officers (FTOs) that were Black, Hispanic, and other (non-white) race.

Column (4): Results use new officers from the main sample whose average district of assignment was at the 75th percentile of violent crime in Chicago. Violent crime rates are violent crimes in a month, based on Chicago City Data Portal crime data, per 10,000 population, based on the 2010 Census.

Column (5): Results use the sample of officers who started between 2006 and 2007.

Column (6): Results from estimating the first stage on a restricted sample of officer assignments, only including assignments in 2012 to 2015, with no additional information codes, and only regular watch assignments (the three main shifts).

Column (7): The sample is the subset of the main sample of recruits who started within 5 months of the first 2012 cohort, which is 7 cohorts. Due to few cohorts, Webb (2014) standard errors are used clustered at the cohort level, and I do not control for cohort size.

Column (8): The fixed effects used as dependent variables were recovered from estimating equation (3) with an arrest only contributing toward the dependent variable if the officer was the first arresting officer.

Column (9): The fixed effects used as dependent variables were recovered from estimating equation (3) excluding crime polynomials from estimation.

Columns (10),(11): The fixed effects from estimating equation (3) on the sample of shifts where officers were in cars (>85Column (12): Sample of estimation includes the main sample officers and those in the main sample cohorts but spent too much or too little time in the academy but were still able to have fixed effects recovered from equation (3).

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

Table B.11: Effect of Cohort Diversity on Arrest Propensity with Share Female

	Propensity to Arrest Blacks		
	All	Non-Index	Index
	(1)	(2)	(3)
Cohort Share Black	-0.109	-0.194	0.085^{***}
	(0.155)	(0.158)	(0.019)
Cohort Share Hispanic	-0.162^{**}	-0.199^{***}	0.037^{***}
	(0.063)	(0.070)	(0.012)
Cohort Share Female	-0.184	-0.191	0.007
	(0.123)	(0.137)	(0.029)
Full Controls	X	X	X
R^2	0.124	0.138	0.055
Num. obs.	962	962	962

Note: Table displays the effect of cohort diversity on main sample officers' propensities to arrest Blacks for all (Columns (1)), low-severity (Column (2)), and high-severity (Column (3)) crimes. The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4), with controls denoted as Full referring to the specification in Column (3) of Table 5. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.12: Main Results (Table 5, Columns (4) and (5)) with Different Standard Errors

	Arrest Propensity			
	Bootstrap		Wild	
	Black Non-Index	Black Index	Black Non-Index	Black Index
	(1)	(2)	(3)	(4)
Cohort Share Black	-0.2374	0.0862^{***}	-0.2374^{*}	0.0862^{***}
	(0.1945)	(0.0328)	(0.1355)	(0.0195)
Cohort Share Hispanic	-0.1773^{*}	0.0362^{**}	-0.1773^{**}	0.0362^{***}
	(0.0990)	(0.0171)	(0.0756)	(0.0125)
Full Controls	X	X	X	X
R^2	0.1287	0.0546	0.1287	0.0546
Num. obs.	962	962	962	962

Note: Table displays the effect of cohort diversity on main sample officers' propensities to arrest Blacks for low-severity (Columns (1) and (3)) and high-severity (Column (2) and (4)) crimes, replicating the results in Table 3. Columns (1) and (2) display standard errors based on bootstrap clustering at the cohort level, and Columns (3) and (4) display standard errors based on wild bootstrap clustering with Rademacher weights at the cohort level. Full Controls refers to the specification in Column (3) of Table 5. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.13: Effect of Cohort Diversity Cut-Offs on Arrest Propensity

	Black Arrest Propensity		
	Cutoff = 0.45	Cutoff = 0.5	Cutoff = 0.55
	(1)	(2)	(3)
Cohort Share Minority	-0.637^{***}	-0.399^{***}	-0.244^{***}
	(0.180)	(0.103)	(0.088)
CS Minority Post-Cutoff	0.584^{**}	0.490^{***}	0.505
	(0.249)	(0.187)	(0.326)
Full Controls	X	X	X
R^2	0.124	0.123	0.118
Num. obs.	962	962	962

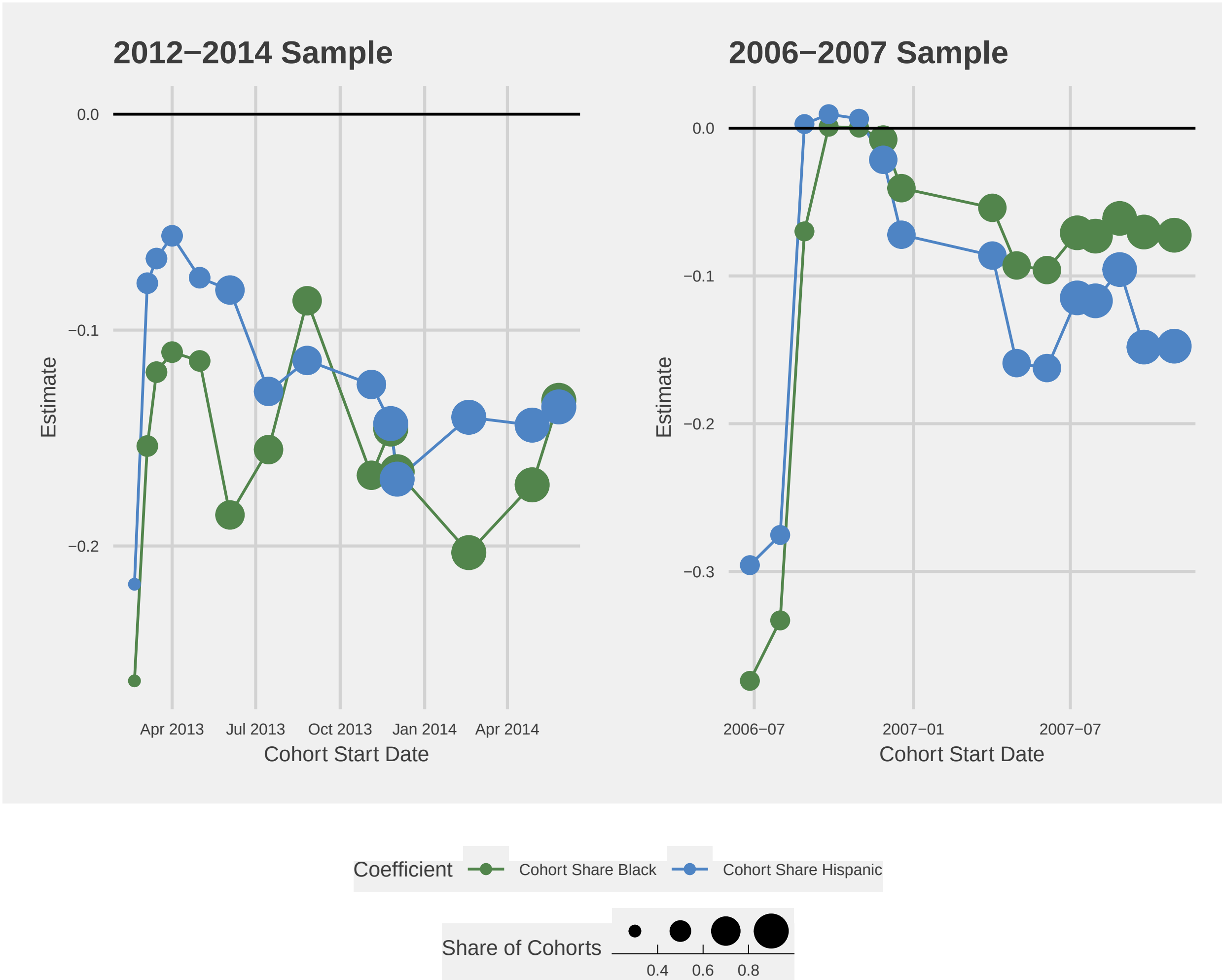
Note: Table displays the effect of cohort share minority on main sample officers' propensities to arrest Blacks based on cutoff points. The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4), with Full Controls referring to the specification in Column (2) of Table 5 with the addition of CS Minority Post-Cutoff. Each column uses a different cutoff value for cohort share minority, 0.45, 0.5, 0.55, which correspond to about the 30th, 50th, and 70th percentiles of cohort diversity main sample officers experience, respectively. Cohort Share Minority is the effect of cohort share minority (CSM), and CS Minority Post-Cutoff (CSMPC) is effect of CSM minus the cutoff if CSM is greater than the cutoff and zero otherwise. For example, if CSM=0.55 and cutoff is 0.45, CS CSMPC = 0.1, and if CSM=0.3 then CSMPC=0. The pre-cutoff effect of diversity is CSM, and the post-cutoff effect is CSM + CSMPC. Cohort shares and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table B.14: Effect of Cohort Racial Entropy and Concentration on Arrest Propensities

	Propensity to Arrest Blacks					
	Entropy (B,H,WO)		Entropy (M,W)		HHI (B,H,W,O)	
	Non-Index	Index	Non-Index	Index	Non-Index	Index
	(1)	(2)	(3)	(4)	(5)	(6)
Measure	-0.0182	0.0332	-1.1053^{*}	0.0076	0.3410^{**}	-0.0469^{***}
	(0.1273)	(0.0297)	(0.5739)	(0.1045)	(0.1454)	(0.0167)
Full Controls	X	X	X	X	X	X
R^2	0.0941	0.0441	0.1136	0.0411	0.1242	0.0449
Num. obs.	962	962	962	962	962	962

Note: Table displays the effect of cohort racial entropy and HHI on officer propensity to arrest Blacks for non-index and index crimes for main sample officers. The propensity is captured by officers' fixed effects using equation (3). The parameter estimates are based on the specification in equation (4), with Full Controls referring to the specification in Column (3) of Table 5, excluding all racial group cohort shares, and the addition of the variables of interest. The variable of interest ('Measure') is indicated in the header: in Columns (1) and (2) is the racial entropy across cohort shares of Black, Hispanic, and white/other peers; in Columns (3) and (4), it is the racial entropy across cohort shares of minorities and whites; in Columns (5) and (6), it is the HHI caculated across cohort shares of Black, Hispanic, white, and other peers. Cohort shares used for entropy and HHI construction and mean age are computed as the leave-out mean of the officer's cohort's initial composition. Standard errors clustered at cohort level are in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Figure B.1: Coefficient Estimates by Increasing Sample Cohorts

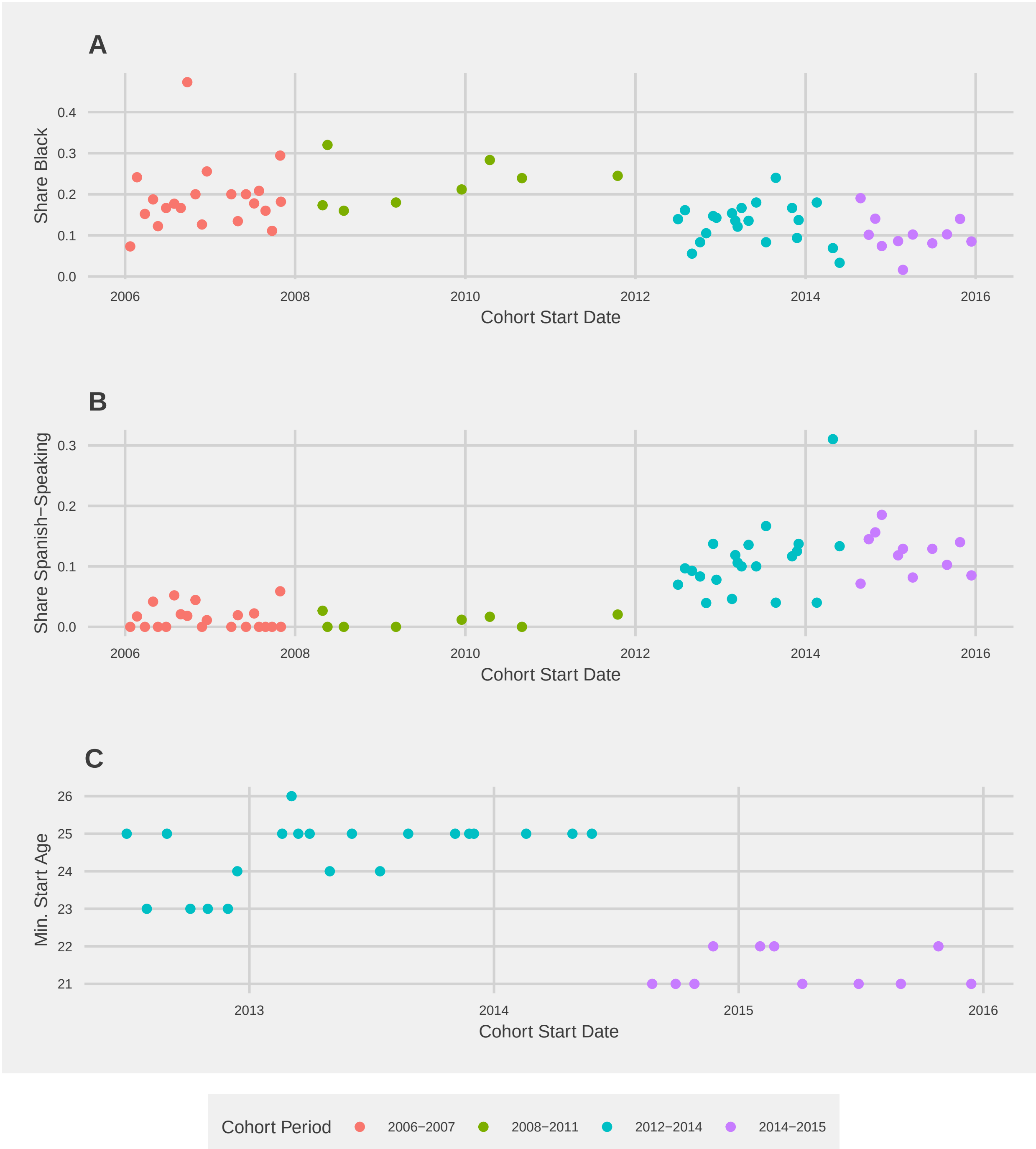


Note: Figure displays coefficients (y−axis) recovered from estimating equation (4) on the sample of cohorts which started on or before each date (x−axis). The dependent variable is officer propensity to arrest Blacks. As the sample size increased (more cohorts are included), the coefficients become more stable.

Figure B.2: CPD Exam Information

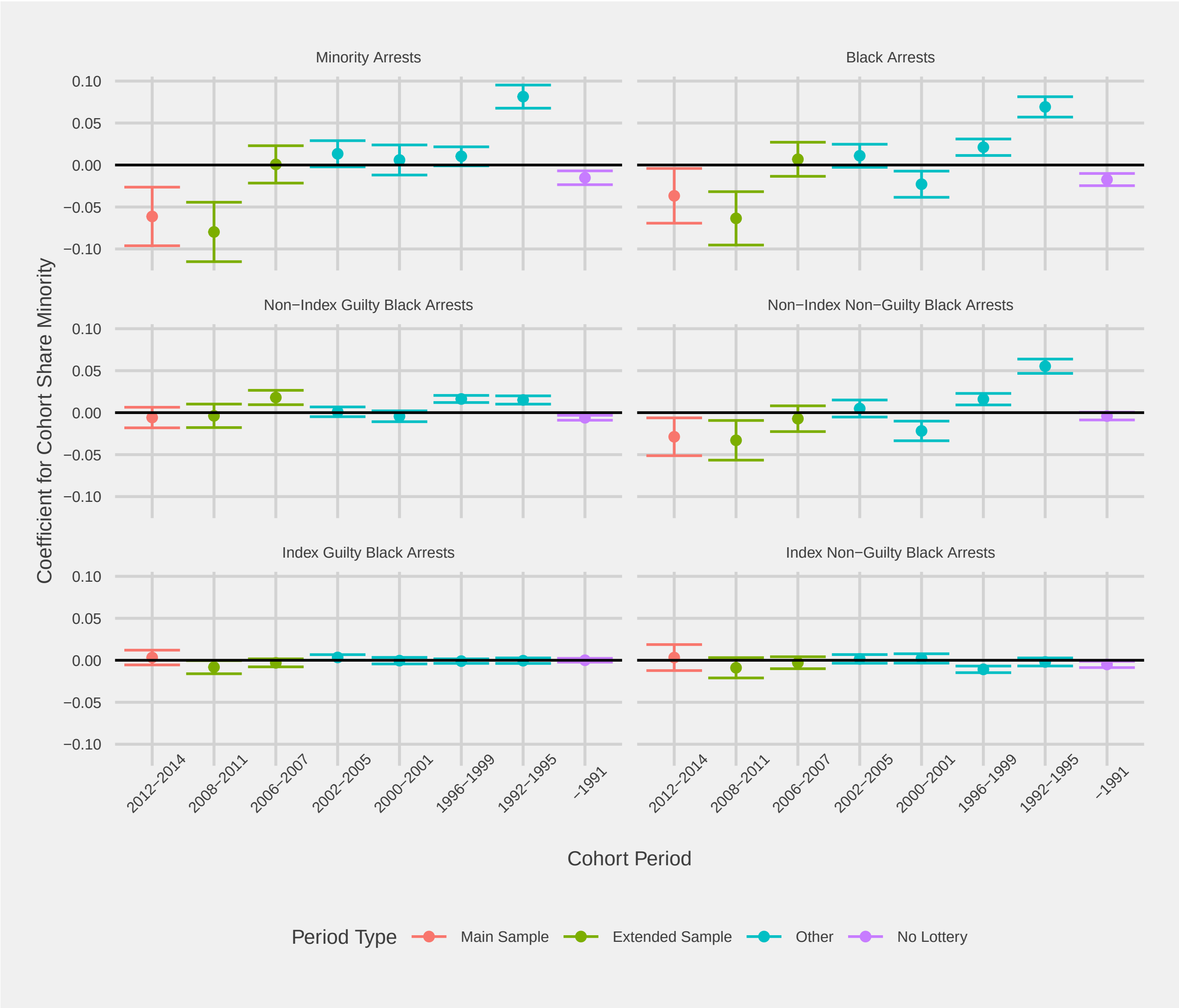
Exam	Date of administration	Attended	Passed	Failed
Police Entrance 1999	3/15/1999; 3/16/1999	3,967	No info available	No info available
Police Entrance 1999	1/5/2000	2,517	No info available	No info available
Police Entrance 2000	7/1/2000	2,053	No info available	No info available
Police Entrance 2000	1/4/2001	1,829	No info available	No info available
Police Entrance 2001	5/19/2001	1,923	No info available	No info available
Police Entrance 2002	1/12/2002	3,150	No info available	No info available
Police Entrance 2003	11/22/2003	3,875	No info available	No info available
Police Entrance 2004	11/20/2004	4,163	No info available	No info available
Police Entrance 2005	2/18/2006; 2/19/2006	4,061	3,338	723
Police Entrance 2006-1	6/4/2006	1,508	1,255	253
Police Entrance 2006-2	8/6/2006	1,025	863	162
Police Entrance 2006-3	11/5/2006	1,795	1,487	308
Police Entrance 2010	12/11/2010	8,621	7,689	932
Police Entrance 2010 make up	makeups: 3/12/2011; 6/11/2011; 9/25/2011; 12/3/2011; 6/2/2013; 12/1/2012; 3/9/2013	No info available	No info available	No info available
Police Entrance 2013	12/14/2013 & military makeups (6/28/2014; 12/7/2014; 6/13/2015; 12/6/2015)	14,788	12,877	1,911
Police Entrance 2016	4/16/2016 & make ups :12/3/2016; 12/4/2016	10,199	9,023	1,176
Police Entrance Spring 2017	4/1/2017-4/2/2017	8,620	7,437	1,183
Police Entrance Winter 2017	12/16/2017,12/17/2017 & makeup: 2/24/2018	7,294	6,418	876
Police Entrance Spring 2018	5/5/2018 & 5/6/2018 & makeup: 6/23/2018	4,273	3,789	484
Police Entrance Winter 2018	12/8/2018	4,433	3,964	469
Police Entrance Winter 2018 make up	3/9/2019	Hasn't occurred	N/A	N/A

Figure B.3: Composition of Cohorts by Start Date



Note: Figure displays the share of cohorts with more than 10 starting members that are Black (Panel A) and speak Spanish (Panel B), and the lowest starting age (Panel C) by the cohort start date, from 2006 to February 2015 for Panel A andnd 2012 to February 2015 for Panel C. Cohort period denotes the time period during which the cohorts started and assumes cohorts in the same period were in the same test pool.

Figure B.4: Effect of Diversity Arrests by Cohort Period



Note: Figure displays coefficients for cohort share minority (y−axis) interacted with cohort−period fixed effects recovered from estimating equation (7) on the full daily assignment panel data from 2010−2016 with cohorts determined by month of start. Error bars indicate 95% confidence intervals, with standard errors clustered at the assignment group level. Each panel contains the type of arrest used as the dependent variable, i.e. the number of that type of arrest an officer made during their shift. The pre−1992 period contains cohorts not subject to random assignment. The periods associated with pre−2012 cohorts do not align with testing pools due to data limitations (which cohorts come from which tests is not known). Officers starting in earlier period can select out of working in the units examined and so the estimates for pre−2008 cohorts may be biased due to selection. The cohort are groupped into periods which belong to four categories (Period Types): Main Sample, the cohort used as the main sample starting from 2012 to 2014; Extended Sample, the cohort used in robustness checks (e.g. in Column (3) of Table 4) which belong to the period 2008−2011 and 2006−2007; Other, which are all other post−1991 cohorts which are not used elsewhere in the paper; No Lottery, which are pre−1991 cohorts whose members were not assigned to lottery numbers (policy changed in 1991). For all Period Types other than Main Sample, the cohorts are broken into periods (e.g. 2008−2011, 2006−2007) which do not correspond to precise test dates−− only the Main Sample cohorts are known to be from the same testing pool (December 2010).