

The Impact of Racial and Ethnic Diversity in Policing*

Bocar Ba[†] Dean Knox[‡] Jonathan Mummolo[§] Roman Rivera[¶]

February 15, 2020

Abstract

Racial and ethnic diversification is among the oldest proposed police reforms, but severe data constraints have forced scholars to rely almost exclusively on problematic cross-agency comparisons to evaluate its effects. We assess the impact of diversification using millions of observations of Chicago police officers' behavior, merged with data on their socio-demographic traits. Compared with white officers facing the same working conditions, we show that black officers make substantially fewer discretionary stops (e.g. for "suspicious behavior"), fewer arrests for petty crimes, and use force less often. These reductions stem primarily from a reduced focus on black civilians, a pattern concordant with efforts to end abusive policing and mass incarceration in African American communities. However, we also find evidence that coarse identity groupings mask large variation in police behavior. Most notably, we find Hispanic officers who cannot speak Spanish—the native language of many co-ethnic communities—stop Hispanic residents at the highest rates. Taken together, these results show the substantial impact of diversity on police treatment of minority communities, and emphasize the need to consider multiple facets of police officers when crafting personnel-driven reforms.

*Data for this project were gathered in collaboration with Invisible Institute, with the help of Sam Stecklow and Emma Herman. We thank Rachel Ryley, the Lucy Parsons Lab, and L2, Inc. for generously sharing additional data, and Ted Enamorado for research assistance.

[†]Postdoctoral associate, University of Pennsylvania, Law School, bocarba@law.upenn.edu.

[‡]Assistant Professor of Politics, Princeton University, dcknox@princeton.edu.

[§]Assistant Professor of Politics and Public Affairs, Princeton University, jmummolo@princeton.edu.

[¶]PhD student, Dept. of Economics at Columbia University, r.g.rivera@columbia.edu.

Racial disparities in the volume and nature of police-civilian interactions (Gelman, Fagan and Kiss, 2007; Knox, Lowe and Mummolo, Forthcoming; Lerman and Weaver, 2014) have for decades fueled persistent allegations of racial bias in policing. Central to these critiques are the fact that throughout most of U.S. history—from the formation of the first organized security patrols on slave plantations (Reichel, 1988; Turner, Giacopassi and Vandiver, 2006), to the concentrated implementation of “Broken Windows” policing in communities of color in recent decades (Wilson and Kelling, 1982)—many police forces have been comprised of nearly all white officers (Forman Jr., 2017). In turn, one of the most often proposed police reforms aimed at reducing racial inequities in policing has been to increase racial diversity within law enforcement agencies (DOJ/EEOC, 2016).

But despite a large body of research in economics (Antonovics and Knight, 2009), criminology (Cao and Cullen, 1996; Shusta and Wong, 1995), law (Forman Jr., 2017), and political science (Garner, Harvey and Johnson, 2019), the efficacy of this reform strategy remains contested. In some cases, scholars have concluded diversity has no detectable impact on police violence (Smith, 2003). Other prominent studies conclude the reform changes the racial composition of arrested civilians (Levitt, 2012; McCrary, 2007). In an exhaustive review of the empirical literature on racial diversity in policing, one legal scholar concluded: “[t]he fairest summary of the evidence is probably that we simply do not know” (Sklansky, 2005, 1225).

The central obstacle to evaluating this reform has been a lack of sufficiently fine-grained data on police officer deployment and behavior. A comprehensive assessment requires analysts to test whether minority police officers behave differently from white officers *when faced with otherwise identical circumstances*. Without the ability to make such careful comparisons, virtually all prominent work in this area has relied on cross-agency comparisons (e.g. Legewie and Fagan, 2016; Donohue III and Levitt, 2001; Smith, 2003). This approach cannot distinguish whether apparent differences in officer behavior arise due to selection (e.g., diverse agencies differ on unmeasured traits) or ecological fallacies (Johnson et al., 2019) (e.g., white and nonwhite officers in a jurisdictions are assigned to work in different circumstances, which we demonstrate is in fact true). In addition, even the handful of studies that have leveraged plausibly exogenous variation in agency racial composition (McCrary, 2007; Garner, Harvey and Johnson, 2019; Harvey and Mattia, 2019)¹ mask important details of police-civilian interactions and the contexts in which

¹McCrary (2007), Garner, Harvey and Johnson (2019) and Harvey and Mattia (2019) employ difference-in-differences designs and leverage the timing of court-ordered affirmative action mandates on law enforce-

they unfold. Most prior work also lacks basic information on police and civilians, much of which is relevant to evaluating this policy. For example, skeptics of diversification argue that minorities choosing to apply to police academies are a self-selected group that may not mirror the socio-demographic backgrounds and political views of the broader minority population. This critique is in line with much prior work on selection in bureaucratic staffing (Clinton and Lewis, 2008; Forman Jr., 2017; Lewis, 2010; Linos and Riesch, 2019; Wilson, 1989). In other words, regardless of an officer’s race and ethnicity, it may be the case that “blue is blue.” If so, diversity-based hiring reforms may not deliver meaningful change.

In this paper, we draw on a host of newly collected data sets that together allow us to overcome these longstanding limitations. Our data, which track more than 33,000 Chicago police officers over time, was assembled through years of open records requests and lawsuits and allows for the most credible and fine-grained assessment of the impact of officer diversity on police behavior to date. It includes timestamped and geolocated records of officers’ decisions to stop, arrest, and use force, as well as civilian complaints that arise from officer behavior. After aggressively pruning these data to maximize the validity of our analyses, we construct a panel of 2.9 million detailed observations of events linked to individual officers and specific patrol assignments. Our data further includes information on officers’ demographics, language skills, places of residence, political preferences/participation, and career progressions. Most crucial for the validity of our policy evaluation, we leverage data on officers’ patrol assignments—i.e., the specific days and shifts that officers were assigned to patrol specific beats, or narrow sets of city blocks. This feature allows us to compare officers of different racial and ethnic groups who are working under the same conditions *even when they do not record any enforcement activities*. This means we can effectively control for environmental factors including differential exposure to civilians across officers, overcoming a fundamental limitation in prior work (Knox, Lowe and Mummolo, Forthcoming; Knox and Mummolo, 2020). In short, our data allow us to credibly approximate, for the first time, whether and how white and nonwhite officers perform their jobs differently.

We present five broad conclusions, several of which undermine prior studies that relied on coarsely aggregated data at the city or agency level:

- (1) Minority officers are assigned to very different districts, and even within districts,

ment agencies to estimate agency-level effects of officer diversity on various outcomes including crime and crime victimization.

receive vastly different geographic and temporal patrol assignments. Without accounting for these differences in working conditions, there is no way to meaningfully characterize the differences between minority and white officer behavior.

(2) Like Chicago city residents, white, black and Hispanic officers tend to live in areas of the city with high shares of co-racial and co-ethnic civilians, a pattern which suggests low levels of inter-group contact when off-duty. Black officers also have similar political preferences to black civilians, while Hispanic and white officers are substantially more politically conservative than their civilian counterparts.

(3) Black officers make roughly 31% fewer stops and 22% fewer arrests, and they use force 35% less often, relative to white officers working in the same places and times. Examining a wide range of officer activity, we show this is mostly driven by a decrease in discretionary activity (e.g. a 53% decrease in loitering stops and 30% fewer drug arrests), rather than lower enforcement of severe criminal behavior (only a 10% decrease in violent-crime arrests). Moreover, relatively higher levels of enforcement by white officers fall primarily on black civilians. Black officers stop and use force against black civilians at a rate that is roughly 40% lower than white officers.

(4) Hispanic officers behave comparably to white officers in the aggregate. But there is massive variation in behavior among Hispanic officers. Those who can speak Spanish make significantly fewer stops and arrests than officers who cannot speak Spanish, and are therefore less likely to have cultural ties to immigrant communities. In fact non-Spanish-speaking officers stop Hispanic civilians at even higher rates than white officers.

(5) Despite clear evidence that black officers make fewer stops and arrests and use force less often than white officers, we find civilians lodge complaints against black and white officers working in the same places and times in roughly equal number overall. However, black officers (relative to white officers) and Spanish-speaking Hispanic officers (relative to non-Spanish-speaking Hispanic officers) receive substantially more complaints for minor infractions, a category that includes failure to file paperwork, while appearing comparable in numerous other categories. These patterns suggest the possibility of bias on the part of civilians when lodging complaints against some groups of minority officers. However, due to data limitations on civilian complaints which we outline below, we present these results as suggestive, requiring further investigation.

Overall, our results suggest that diversification has a substantial impact on the treatment of minority civilians. However, our findings also suggest the need to consider multiple facets of police officers when crafting personnel-driven reforms. Like the civilians

they police, officers are complex, not easily reducible to racial or ethnic categories. Scholars and reformers must take stock of other factors that may affect how officers relate to residents, including political preferences and language skills, to develop effective policy.

All data and interactive replication materials are publicly available at [LINK]. We encourage readers to probe and extend our analyses.

1 Data

Over a period of three years, we submitted open records requests to the Chicago Police Department (CPD) and the city's Department of Human Resources. The resulting records include the name, race, gender, birth year, salary, language skill, unit assignments and appointment date of each officer, as well as behavioral outcomes including arrests, uses of force, and associated complaints (Ba, 2019; Rim, Rivera and Ba, 2019). We merged these data with records on stops of civilians and officer beat assignments, proprietary records on voter turnout and party registration,² and U.S. census data. SI A.1 describes these data sets and our pre-processing procedures in detail.

We begin with 4.2 million stops in the period of interest, including stops by officers on special duty, stops that occurred outside their assigned shift times, or other idiosyncratic circumstances that make principled comparisons difficult. After aggressive pruning to maximize the validity of our comparisons, we retain 1.3 million stop records linked to individual officers and specific patrol assignments. Similarly, after merging and pruning, we retain records of nearly 250,000 arrests and 8,000 uses of force. In this period, we match roughly 6,000 records of civilian complaints (these complaint records are often incomplete, a feature we discuss further below). Table 1 contains aggregate information on four behaviors tabulated by officer race: stops, arrests, uses of force, and civilian complaints from Jan. 2012 to Dec. 2015, the period that informs our behavioral analysis. Due to the small sample sizes associated with officer groups including Asian Americans and Native Americans, our analysis is limited to three racial and ethnic groups: white, black and Hispanic officers, (together accounting for roughly 97% of officers for whom shift assignment data is available).

Figure 1 illustrates the dataset's various attributes by highlighting a small slice of its

²We are grateful to Lucy Parsons Labs for publicly releasing data on civilian stops, Rachel Ryley for generously sharing data on beat assignments, and L2, Inc. for allowing access to its national voter file database.

Table 1: Summary of data on officer behavior (counts), 2012-2015

	Black officers	Hispanic officers	White officers
Stops	253,609	356,541	729,078
Arrests	47,406	65,587	132,285
Uses of force	1,355	2,081	4,514
Civilian complaints	1,661	1,345	2,922

temporal and geographic coverage. The figure maps activity during a four-month window in the CPD’s Wentworth District (District 2), a highly segregated 7.5 square mile territory on Chicago’s South Side that is 96% black and consistently ranks among the city’s most violent districts in per-capita crime rates. The district is comprised of 15 beats (small collections of city blocks assigned for patrols),³ which are shaded according to their racial composition. Points distinguish between geolocated stops, arrests, uses of force and civilian complaints during this period.

The figure also describes the behavior of four CPD officers working in District 2 during this time, renamed to maintain anonymity. For example, one officer is female, Hispanic, speaks both English and Spanish, has no political party affiliation, and was born in 1965; another is a white male Republican who does not speak Spanish, born in 1981. The figure shows the specific beats to which officers were assigned over time, each officer’s behavior while working in these beats, the political party affiliation of each officer (Democrat, Republican, or unaffiliated), and whether they turned out to vote in the 2012 November general election. The figure also highlights particular events associated with these officers. For example, one summary details an incident that occurred on Oct. 8, 2012 at 7:13 p.m. in which an officer used force against a 20-year old unarmed black male. The event occurred outdoors, during clear weather in a vacant lot, and the civilian was not injured during the encounter. The officer, who was not in uniform during the event, was also not injured. The precise location of the event is denoted on the map in an encircled red “X.”

With data this granular, we are not only able to describe the makeup of the police force relative to the city population on a host of dimensions, providing a rich context absent in prior work. We can also make the most valid comparisons to date between the behavior of white and nonwhite officers facing common working conditions. To accomplish this, in our behavioral analysis, we restrict our data to the years 2012-2015, the years covered

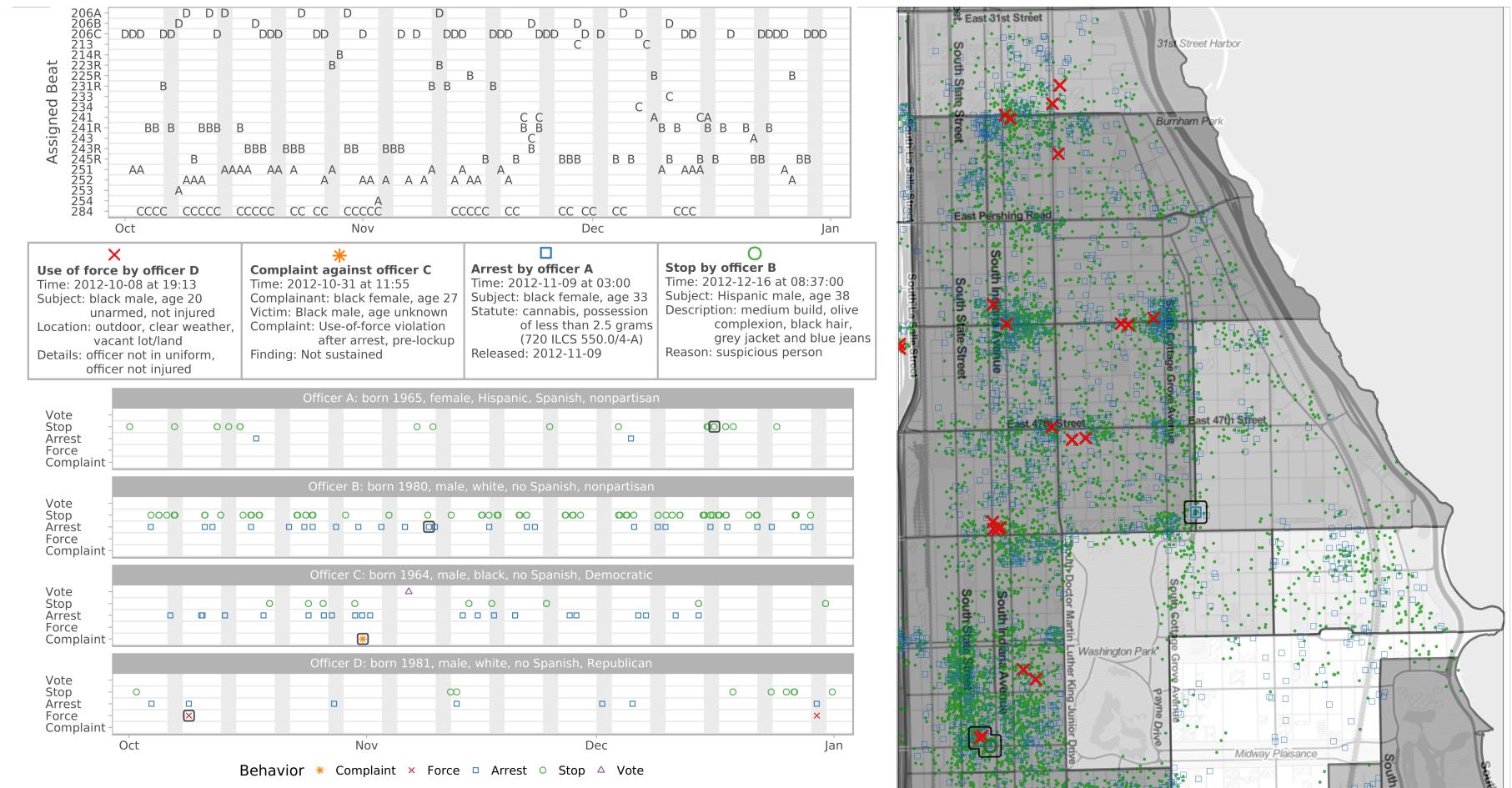
³Based on Chicago Police Department 2008 Annual Report.

by our patrol assignment data. We further restrict analysis to the 87% of officers employed at the rank “police officer” (pay grade D|1) to isolate officers who work primarily on city streets, rather than behind desks or in atypical environments (e.g. excluding officers at rank “helicopter pilot” or “superintendent’s chief of staff”).⁴

In the next section, we further detail our analytic strategy.

⁴This is one of the few officer level covariates we condition on in our analysis. We do this because it represents a potential violation of our key assumption: that white and nonwhite officers common external circumstances within month-day-shift-beat units. We detail this rationale in the following section.

Figure 1: Detailed view of the data. The right panel figure maps police activity in a single police district (Wentworth, CPD District 2), with green circles, blue squares, and red crosses respectively indicating the locations of stops, arrests, and uses of force. Polygons represent geographic beats and are shaded by their proportion of minority residents. Lower left panels chart the behavior of four anonymized officers over a four-month period, with panel headers indicating age, gender, ethnicity/race, language ability, and political orientation. In addition to mapped events, these panels indicate general election turnout and civilian-filed complaints with purple triangles and orange asterisks, respectively. Encircled incidents are described further in the left middle panel, which report civilian and incident specifics. Finally, the top left panel indicates how the four selected officers are assigned to patrol beats over dates and times, with vertical gray bars indicating weekends.



2 Empirical Strategy

Our main goal is to assess the impact of officer diversity on the treatment of civilians. Specifically, we seek to test whether officers of various racial and ethnic groups perform their jobs differently, on average, when faced with the same circumstances. However, previous policing studies have lacked information on the context in which police-civilian interactions occur, as well as insight into how officers relate to civilians on socio-demographic dimensions.

Our analysis therefore proceeds in two main parts. The first is a descriptive analysis of the concordance between several features of officers and the civilians they police. Beyond shared racial and ethnic identity, these include place of residence, political preferences and political participation. Many reformers have asserted the importance of additional factors that may condition the impact of racial diversity, such as the frequency with which officers interact with co-racial and ethnic civilians, whether officers reside in the communities they patrol (Goodyear, 2014) and whether officers share common goals with respect to police reform, among other attributes (Lockwood and Prohaska, 2015). Indeed, there are strong reasons to suspect that racial diversification alone will not bring about meaningful change depending on the prevalence and distribution of these other features. Hiring conservative black officers who believe policing needs no substantial reform (Eckhouse, 2019; Federico and Holmes, 2005; Weaver, 2007; Weitzer and Tuch, 2004), or Hispanic officers who cannot speak the native language of many immigrant communities, for example, may do little to transform the nature of police-civilian interactions. Our analysis thus begins with a rich description of how officers relate to the civilians they police on a host of dimensions.

Following this descriptive assessment, we proceed to our main analysis, presenting comparisons of the behavior of officers belonging to various racial and ethnic groups who are working under virtually identical conditions. Unless white and nonwhite officers behave differently on the job, it is unlikely that diversification will yield tangible benefits for policed populations. The ideal experiment to evaluate this question would be to randomize white and nonwhite officer into patrol assignments. Given this randomization, white and nonwhite officers would encounter the same on-average conditions, and observed behavior across officer groups could inform a valid causal assessment of the policy reform of interest. (For a formal discussion of our quantity of interest, see SI Section B.)

While we cannot directly implement this experiment, our data allow us to recover a

close approximation. We assemble a vast panel data set, where each row contains information on a particular officer’s activities during a particular shift, including contextual information on those activities.⁵ In each of these 2.9 million patrol assignments, we compute the volume of stops, arrests, uses of force, and civilian complaints. We then compare officers of different racial and ethnic groups working in the same unique month, day group (weekend vs. weekday), shift number (one of three nominal eight-hour watches each day), and beat (a small collection of city blocks, averaging 0.82 square miles citywide)—extremely narrow slices of time and space that we label MDSBs for short. Because we are comparing officers working in these narrow units, we can credibly assume that observed differences in officer behavior are not due to disparities in external conditions. That is, we assume that white and nonwhite officers within MDSBs face the same set of enforcement opportunities, and are faced with: the same types of civilian activity; common neighborhood attributes such as infrastructure and architecture; and common time-varying conditions such as seasons and lighting. We term this the “common circumstances” (CC) assumption. For ease of interpretation, we present differences estimated using ordinary least squares regression with MDSB fixed effects, though our results are robust to several other estimators (see SI Figures A6-A9 for these additional results). All statistical inferences are based on officer-level block bootstrap confidence intervals that are robust to unobserved officer-specific peculiarities. Our behavioral analysis organizes patrol assignments into nearly 207,000 groups defined by MDSBs, of which roughly 121,000 contain assigned officers of different ethno-racial groups (i.e., vary on the explanatory variable of interest and thus contribute to our estimates).

Importantly, because we have data on micro-level assignments, we are able to retain officers in our panel analysis in every time period *whether or not they engaged in any enforcement activity*. Here, our analysis differs from most prior studies of police administrative data, which only rely on instances in which officers recorded an activity (e.g. a stop or arrest; see discussion in Knox, Lowe and Mummolo (Forthcoming)). Without these patrol assignment records, inaction is invisible to the analyst, and officers who do not record any enforcement activity simply vanish from the data. To our knowledge, no

⁵In analyzing officer behavior and associated complaints, we take patrol assignments—the level at which commanding officers typically exercise control—as the basic unit of analysis. In SI Section C.5, we examine the possibility that differences in clock-in/out times may be one mechanism behind observed differences in behavior between officer racial groups. While some statistically significant differences can be observed (roughly 0.1% disparities in patrol time), these gaps are two orders of magnitude smaller than the differences observed in stops, arrests, uses of force, and civilian complaints.

previous study has been able to leverage patrol assignments in this fashion.⁶

We stress that our goal is to estimate the first-order impact of a policy intervention—the counterfactual substitution of an officer of differing race into a patrol assignment. (This conceptual manipulation differs substantially from the imagined scenario of changing an officer’s race or ethnicity while holding their other attributes fixed, an arguably impossible counterfactual (Greiner and Rubin, 2011; Holland, 1986).) The policy intervention we examine may have second-order effects as well. For example, the presence of additional nonwhite officers could affect department culture in ways that eventually alter the nature of policing. Such effects are far more difficult to evaluate rigorously and beyond the scope of our analysis.

For full transparency, we highlight a number of possible threats to the validity of our analysis given our analytic goal. Confounding factors in this scenario include all variables that correlate with officer race in ways that violate the CC assumption. An example would be if black and white officers were assigned to the same beat and shift, but black officers were ordered to stay in their patrol cars the entire time while white officers were allowed to freely roam the beat. Another violation, given that we control for day group but not day of week, would occur if white and nonwhite officers were systematically assigned to work particular days of the week in ways that correlated with criminal activity. However, because we are not seeking to identify the effect of race *per se*, other correlates of officer race which do not violate the CC assumption do not obstruct our ability to evaluate this counterfactual. Examples of these innocuous correlates include: (1) black and white officers possessing different levels of education which in turn lead to differential enforcement; or (2) black and white officers choosing to focus on different corners of their beats once assigned in ways that influence policing outcomes. These facets of officer race would represent different mechanisms through which the policy intervention of interest affects police-civilian interactions, but would not bias causal estimates relating to minority officer deployment. (For a related discussion of conceptualizing race as a causal variable, see Knox, Lowe and Mummolo (Forthcoming) and Sen and Wasow (2016)).

This analysis, while a substantial advance over prior work, has important limitations. First, we measure relative differences in officer behavior, which do not necessarily imply racial bias. For example, if we were to find that white and nonwhite officers use force at

⁶In some cases, our data indicate that two officers were involved in stopping a civilian. Throughout, such instances enter as separate rows in our officer-shift data, since in such cases each officer made a stop. However, to gauge robustness, Si Section C.10 presents a reanalysis of stops after restricting the data to the first reporting officer, which yields essentially identical results.

equal rates against nonwhite civilians under identical circumstances, that may indicate no racial bias. However, it may also indicate a high but equal level of bias across officers. A second limitation is that we cannot identify specific mechanisms connecting officer race to behavior. The behavioral differences we observe could be due to differences in how officers engage civilians, differences in how civilians behave toward certain officer groups, or some combination of the two. The effects we demonstrate still inform the efficacy of the policy of racial diversification, but the mechanisms behind these effects deserve further study, a point to which we return in the conclusion.

In addition, while our analysis tests for tangible effects of diversification, we emphasize this policy reform may yield additional intangible benefits, namely, increased trust in government among minorities. A long literature has argued that trust in police is a crucial element of public safety, compelling residents to more readily cooperate with law enforcement (e.g. Tyler and Fagan, 2008) and thereby potentially increasing the chances that officers solve crimes. If racial diversity affects levels of public trust, the cost-benefit calculus when considering this policy would be further complicated.

Finally, our analysis is confined to a single city, a feature that affords us unusually detailed data at the expense of geographic scope. Given the substantial progress our data allow us to make in quantifying the impact of this policy, we view this as a worthwhile trade. These data also emanate from a setting that is ideal for investigating the role of officer diversity in shaping police-civilian interactions: both Chicago and its police department are racially diverse, and the city has come under fire for a range of aggressive police tactics and allegations of misconduct, making it an important case for the study of a widely proposed police reform (Rivera and Ba, 2019). Our analysis also provides a template for future work to be conducted in other jurisdictions.

3 Racial Representation in Chicago

The CPD, which currently employs about 12,000 sworn officers, exhibits substantial levels of racial and ethnic diversity, making it a useful test case for the efficacy of this policy. However, this has not always been the case: the department has slowly and steadily diversified over time. Figure 2 visualizes this process. The figure's left panel plots the share of CPD employees belonging to each racial and ethnic group since 1970, as well as the corresponding shares of the city population belonging to those groups (see SI Section A.2 for details on the coding of race and ethnicity). The right panel compares the share of

new CPD recruits belonging to racial and ethnic groups to the city's population over time. White officers have historically made up a disproportionate share of the CPD, but their representation has been steadily declining, dropping from 78% (61%) to 54% (36%) of the agency (city) between 1970 and 2010. These declines in white representation have been offset by gains by minority groups, with sharp rates of increase for Hispanics in particular in both the CPD and among city residents. At present, the CPD is roughly 24% black and 21% Hispanic. The figure also shows that blacks have been declining as a share of city residents since 1980, which has likely contributed to stagnant growth in the CPD's share of black officers. In recent years, only about 15% of new CPD recruits have been black, down from a high during this period of nearly 35% in the late 1970s. While the agency does not perfectly mirror the racial composition of the city, it does contain substantial shares of nonwhite officers—a necessary feature for our evaluation.

4 Race and geographic assignment

Given Chicago's high levels of racial segregation⁷ among residents, it is unclear the extent to which officers interact with co-racial and co-ethnic civilians while on duty. If such interactions were infrequent, the impact of diversification could be limited. To assess this, Figure 3 plots the the composition of police districts (25 sectors of the city, each covering 9.25 square miles, on average) against the composition of officers assigned to work in those districts. In general, white-dominated districts have virtually no minority officers assigned, and districts with sizeable minority populations tend to have more officers of the corresponding race. However, a number of districts dominated by black residents nonetheless have sizeable contingents of white officers. For example, Wentworth (CPD District 2, depicted in Figure 1) is 95% black, but 24% of officers assigned there are white. The disparity is even starker in Austin (CPD District 15), where a 93% black resident population is policed by a unit that is 55% white. (See SI Section A.3 for details, and SI Figure A1 for additional analyses.)

Further, among officers assigned to a particular district, considerable variation exists in the exact patrol assignments that officers receive. To show this, we examine each unit individually, tabulating officer race and shift time assignments (first, second and third watch, respectively corresponding to the nominal duty periods of midnight to 8 a.m., 8 a.m. to 4 p.m., and 4 p.m. to midnight). SI Figure A2 depicts the frequency of each shift period,

⁷See http://www.censusscope.org/us/print_rank_dissimilarity_white_black.html.

Figure 2: **Composition of CPD officers and city residents over time.** The left panel depicts the proportion of active CPD officers belonging to a racial/ethnic group in January of each year. The right panel shows the corresponding proportion for incoming recruits, aggregated by five-year windows (e.g., 1970–1974, 1975–1980) to reduce noise from relatively small recruitment cohorts. In both plots, decennial Census proportions for each racial/ethnic group for the city of Chicago, tabulated by the National Historical Geographic Information System, are provided for reference.

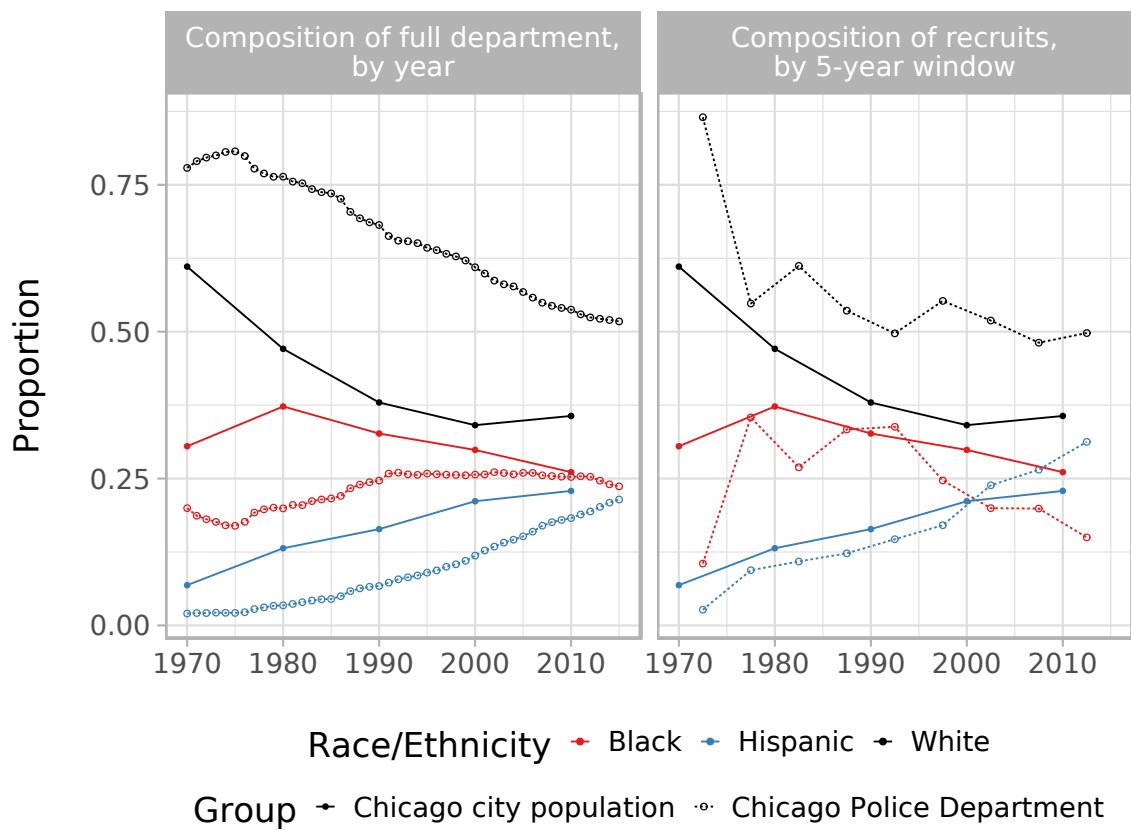
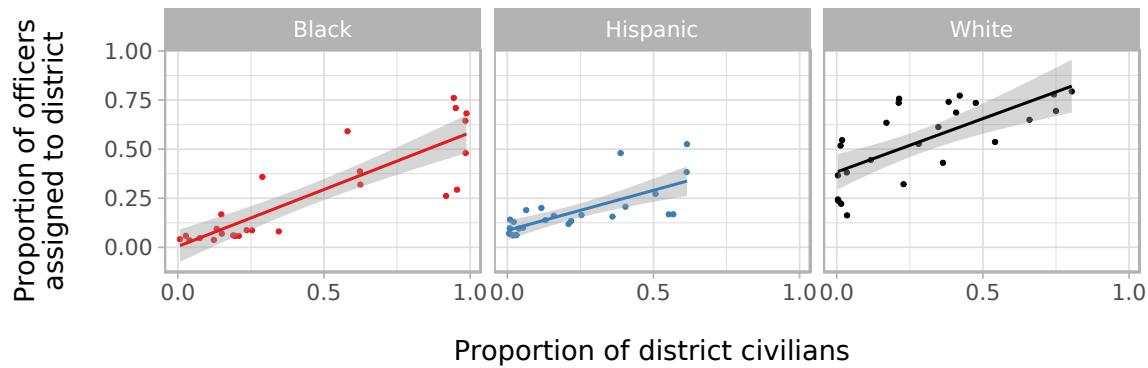


Figure 3: Racial and ethnic composition of officers' assigned districts. In each panel, points corresponding to police districts indicate the proportion of civilians of a given racial/ethnic group residing in 2010 Census data, together with the corresponding share of officers assigned to that district in January 2010 from the same racial/ethnic group.



showing that the pattern of assignments differs dramatically by officer race. For example, white officers in Wentworth (District 2) almost exclusively serve from 4 p.m. to midnight, whereas black officers are more likely to be assigned to mid-day shifts. Similarly, SI Figure A3 shows that relative to white officers, black officers are far more frequently deployed to assigned beat 202—roughly corresponding to a patrol area in the district’s southwest corner (depicted in Figure 1) that has extremely high police activity and a high concentration of black residents. These results undermine analyses in a wide array of previous studies that aggregate at high levels of geography (for example, controlling for district or unit assignment) and which assume that officers face homogeneous conditions within these crude groupings.

These patterns not only establish that officers tend to police co-racial residents, but also underscore a central difficulty in testing whether white and nonwhite officers perform their duties differently. Namely, white and nonwhite officers work in different environments on average, especially with regard to local racial composition. This means any inferred differences in officer behavior that rely on data aggregated to large geographic units, such as districts or agencies—the analytic strategy in most prior work—may simply be artifacts of the divergent environments they police. Our data allow us to overcome this obstacle to inference.

5 Do Officers Live Where They Work?

Our data contain the home addresses of CPD employees for the year 2004, which allows us to assess whether officers reside in or near the areas of the city where they work.⁸ Home location may serve as a rough proxy for all manner of attributes, such that officers living in their assigned district may have more in common with the residents they police. While CPD officers are required to live within the city limits, Chicago's extreme racial and economic segregation means that living in a different part of town often implies a wholly different lived experience. To the extent that officers reside in the same areas they patrol, some reformers have asserted they may have greater familiarity with residents, inspire greater levels of trust, and exhibit greater levels of empathy toward the civilians they serve (Goodyear, 2014; Gramlich, 2017). And by assessing the correspondence between the racial composition of areas where officers live and work, we can gain insight into whether officers are likely coming into contact with various civilian groups when off duty, i.e. in a less adversarial context.

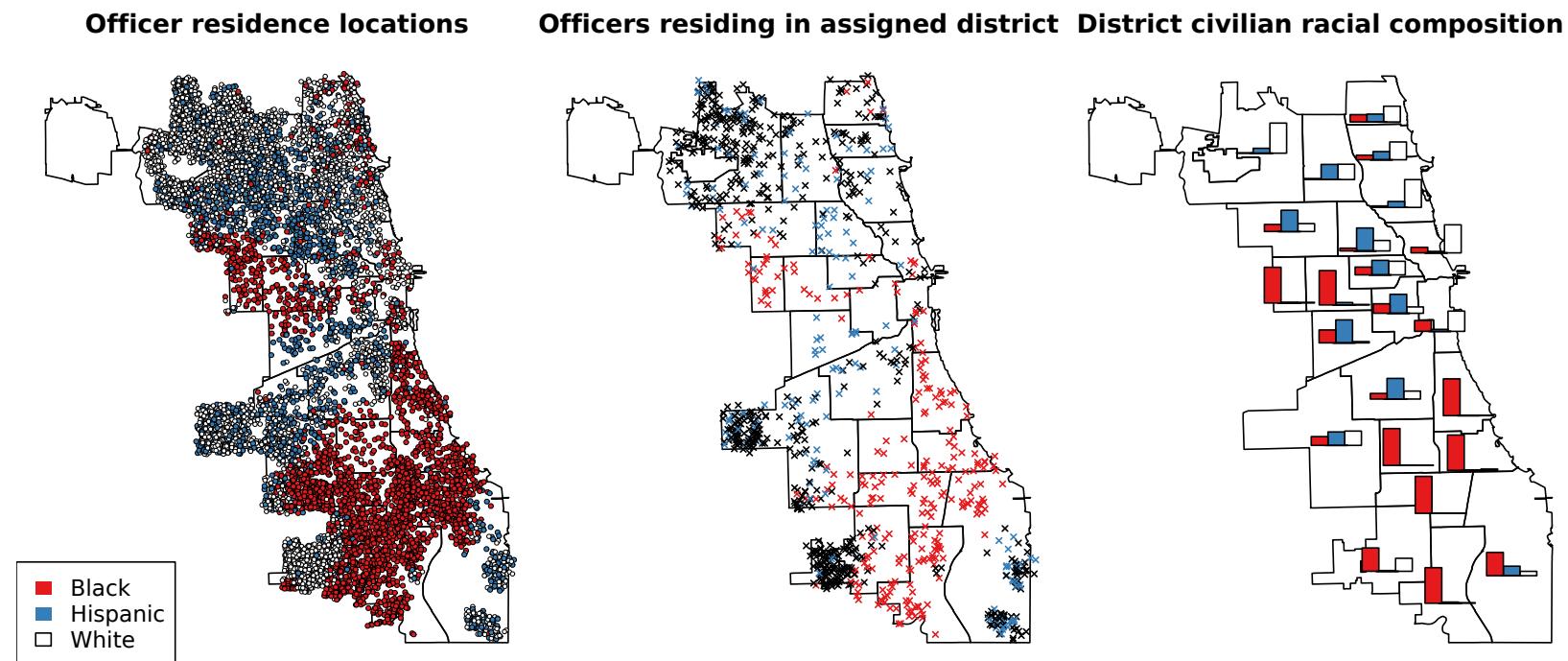
Figure 4 shows that much like city residents, Chicago Police officers are racially segregated. The left panel of the figure depicts the residences of 13,635 officers for whom geocoded 2004 addresses were available, perturbed randomly by 500 meters to maintain privacy. The center panel shows only those officers assigned to work in their home district in 2004. The right panel shows the racial composition of each district for reference. As the figure makes clear, officers tend to live among co-racial and co-ethnic civilians, a pattern echoed in Table 2.

Table 2: Racial composition of officers' home districts, averaged over officers. Civilian demographics of the districts in which officers live, averaged over officers in each racial group. Boxes indicate co-ethnic/co-racial proportions.

	Proportion Black Residents	Proportion Hispanic Residents	Proportion White Residents
Officer White	0.219	0.207	0.525
Officer Hispanic	0.207	0.324	0.420
Officer Black	0.630	0.153	0.199

⁸In addition, we analyze local demographics here at the district level rather than the beat level, because the geographic boundaries of beats change over time, sometimes in unobservable ways. However, this does not compromise our behavioral analysis. Within a given month, beat boundaries remain fixed, so we can be sure officers compared within beats in a given month are patrolling to same geographic areas.

Figure 4: **Police residences by officer race.** The left panel shows the rough locations of all CPD officer residences in 2004, perturbed to maintain privacy. The center panel shows those officers whose homes are located in their assigned district. The right panel shows the local racial composition of each district.



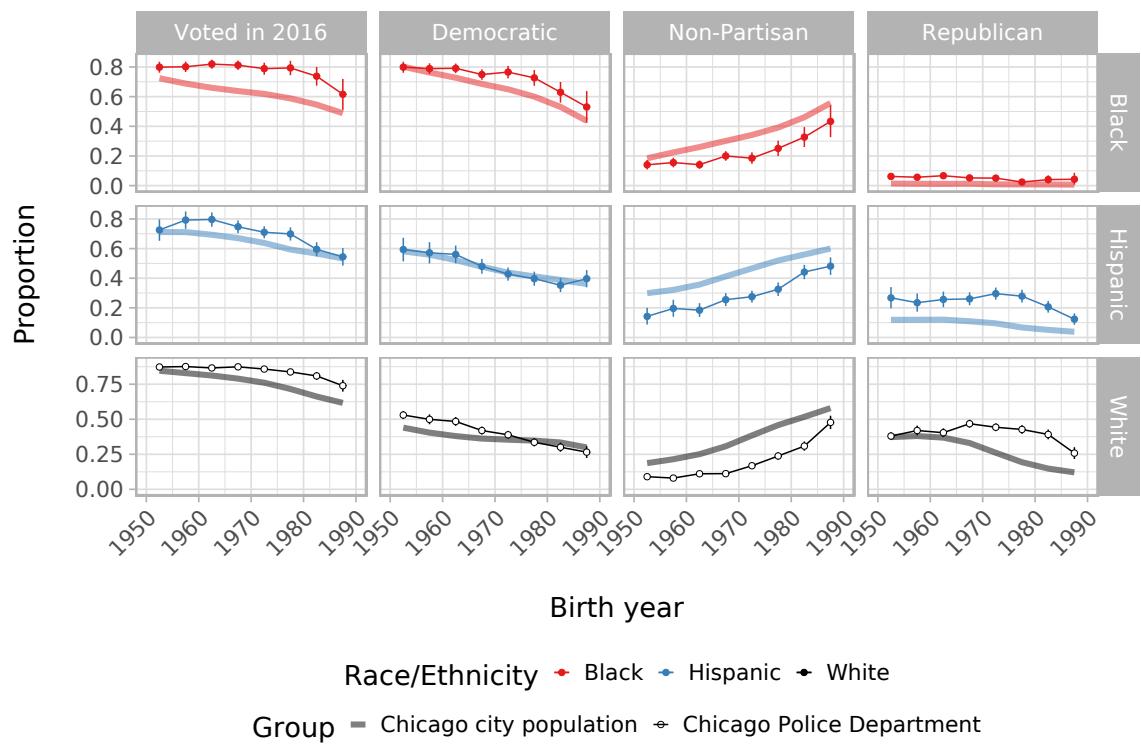
6 Officers' Political Preferences

Another contextual factor that may contribute to the effectiveness of diversification is concordance between the political preferences of officers and residents within racial and ethnic groups. Democrats and Republicans hold markedly different views on criminal justice matters (Eckhouse, 2019; Federico and Holmes, 2005; Weaver, 2007; Weitzer and Tuch, 2004). If nonwhite officers skew conservative, there may be little reason to expect that racial diversification would change the nature of policing. To investigate this, we merge our officer data with a commercial voter database, following Enamorado, Fifield and Imai (2017), to obtain information on officers' political activity and partisan leanings; details of the procedure are given in SI Section A.4.⁹

Figure 5 displays the share of each racial and ethnic group affiliating with the Democratic and Republican parties—as well as the shares affiliating with neither party—by birth year among both city residents and CPD officers. This tabulation, based on contemporaneous voter turnout data, allows us to assess the degree to which CPD officers' current political preferences resemble those of Chicago residents, as well as the degree to which this correspondence varies across cohorts. Turning first to black officers, we see that compared to city residents, officers are slightly more likely to be Democrats, slightly less likely to be nonpartisan, and slightly more likely to be Republican. Though there is some variation in the magnitude of these differences across cohorts, in the aggregate, 7% of black officers affiliate with the Republican party, compared to 1% of black city residents. But Hispanic and white officers diverge sharply from minority populations: while the share of Democratic officers is comparable to the share of Democratic civilians, officers are substantially less (more) likely to be nonpartisan (Republican). In the aggregate, 22% of Hispanic officers and 40% of white officers affiliate with the Republican party, compared with 8% and 28% among co-racial and co-ethnic residents, respectively. In general, black officers appear fairly similar to co-racial residents in terms of political preferences, while Hispanic and white officers appear markedly more conservative. To the extent political preferences predict styles of policing, this pattern suggests that black civilians may experience improved treatment by black officers relative to other groups of officers. In fact, we find precisely this pattern in our behavioral analysis below.

⁹Voters in Chicago are not required to register with a party to vote in party primaries, but can only participate in one party's primary in each election. L2 codes party affiliation based on a proprietary process, and we use their coding here. As a robustness check, we also code party ID based on which party primary voters participated in during the 2016 cycle and obtain highly similar results; see SI Figure A5.

Figure 5: Police more likely to be Republican and to vote than co-Racial and co-ethnic civilians. Cohort political participation and composition of CPD officers and Chicago residents based on current political orientation. Officers are aggregated to five-year windows on birth year (1950–1954, 1955–1960, etc.). For turnout in previous elections from 2000–2016, see SI Section C.3 and SI Figure A4.



In addition, Figure 5 shows police across nearly all cohorts are far more likely to vote than city residents. On average, white, Hispanic, and black citizens' (officers') voting rates in 2016 were 74% (84%), 61% (68%) and 60% (77%), respectively. Given longstanding correlations between turnout and strength of partisanship (Campbell et al., 1960), this result suggests that not only do many officers affiliate with the Republican party at higher rates, they are also likely to be stronger partisans across the board.

7 Racial Diversity and Police Behavior

We now turn to assessing whether racial diversity among police officers affects the nature and volume of police-civilian interactions. Figures 6-8 display the average differences in the number of stops, arrests and uses of force associated with black and Hispanic officers relative to white officers working in the same MDSBs. Turning first to black officers, we see that when faced with the same conditions, black officers make roughly 0.16 fewer stops, make 0.020 fewer arrests, and use force 0.0011 fewer times per shift, on average, than white officers facing the same circumstances—that is, counterparts assigned to the same month, the same part of week, the same shift time, and the same beat. These gaps respectively represent reductions of 31%, 22%, and 41% over typical stop, arrest, and use-of-force volume for white officers (see SI Table A1 for detailed tables of average behavior by officer race).

Importantly, Figure 7 also shows that these disparities are not uniform across situations. Rather, they are driven by a reduced focus on engaging black civilians (41% fewer stops, 26% fewer arrests) and a broad class of enforcement activities that can be thought of as relatively discretionary, including stops of civilians for “suspicious behavior” (33% less) or arrests for drug-related crimes (30% less). However, when it comes to enforcing violent crime, black officer behavior looks similar to that of white officers, with black officers making only 10% fewer arrests for violent crimes than whites. In other words, when it comes to policing the most serious offenses, black officers are comparable to white officers. Black officers also deploy force 35% less often than their white counterparts. This difference is driven largely by reduced uses of force against black civilians: decreased use of force against this group accounts for 83% of the overall disparity between white and black officers. (SI Section C.6 and SI Tables A2–A6 report detailed analyses for the results presented here; SI Section C.7 and SI Figures A6–8 show that results are virtually identical using a wide range of alternative estimators.) This pattern of results is remarkably in line

with the hopes of proponents of racial diversification, who have asserted the policy could reduce abusive policing and mass incarceration in black communities. Our results imply that if the CPD were to hire additional similarly situated black officers from the same pool of recruits from which current officers are drawn, and deployed them in the same manner as current officers, then black civilians would be subject to significantly lower levels of detainment for petty crime and police violence.

Results differ in important ways for Hispanic officers. As with black officers, Hispanic officers facing the same working conditions exhibit fewer stops, arrests and uses of force than white officers, but the magnitudes of these differences are far more modest. Strikingly, these negative differences are driven by less engagement with black civilians, while Hispanic officers exhibit nearly the same volume of enforcement activity against Hispanic civilians as white officers on average. These estimates imply that Hispanic officers can be expected to produce 0.032 fewer stops (6% reduction relative to white officers) and 0.00034 fewer uses of force (11% reduction) per shift.

In addition to stops, arrests and uses of force, Figure 9 displays average differences in complaints filed by civilians against different groups of officers. These results differ somewhat from our analyses of the previous three behaviors. Despite engaging civilians at a lower rate, black officers receive roughly the same number of complaints as white officers working in the same places and under the same conditions, with two notable exceptions: black officers are significantly less likely to receive a complaint related to improper search or arrest of a civilian (0.22 fewer complaints per 1,000 shifts, a 36% reduction), and they are estimated to receive a large but statistically insignificant number of additional minor procedural complaints (e.g. relating to improper filing of paperwork or failure to provide service). Hispanic officers, on the other hand, receive 0.16 fewer complaints per 1,000 shifts relative to comparable white officers. This gap is primarily driven by a decrease in complaints (1) from black civilians (0.10 fewer complaints per 1,000 shifts, a 19% reduction) and (2) relating to improper use of force (0.09 fewer complaints per 1,000 shifts, a reduction of 21%).

We caution that complaint records are incomplete, in part due to extensive missingness and procedural hurdles that impede their filing (Ba, 2019). For instance, of complaints filed between January 2011 and July 2014, investigations were closed for 38% because the investigator was not able to link the allegation to a specific accused officer. In addition, complaints may incorporate elements of civilian perception and bias (e.g., filing of petty or retaliatory complaints). We note that our analysis of complaints necessarily differs from

Figure 6: **Stops.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian stops conducted by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Corresponding regression tables are reported in SI Section C.6, and results from alternative estimators are shown in SI Section C.7.

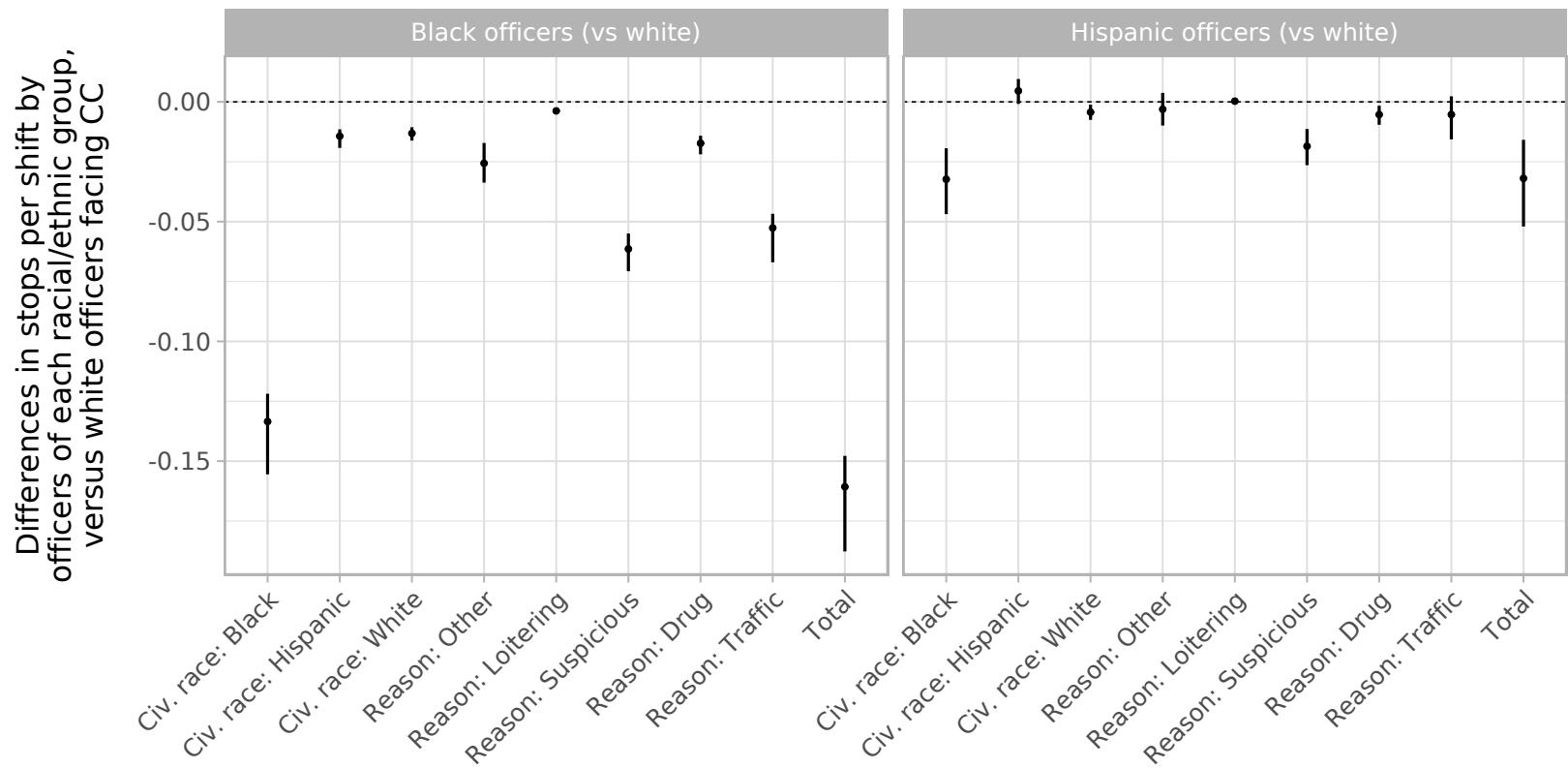


Figure 7: **Arrests.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of arrests conducted by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Corresponding regression tables are reported in SI Section C.6, and results from alternative estimators are shown in SI Section C.7. Results are qualitatively similar when holding officer gender constant; see SI Figures A13–A14 for details.

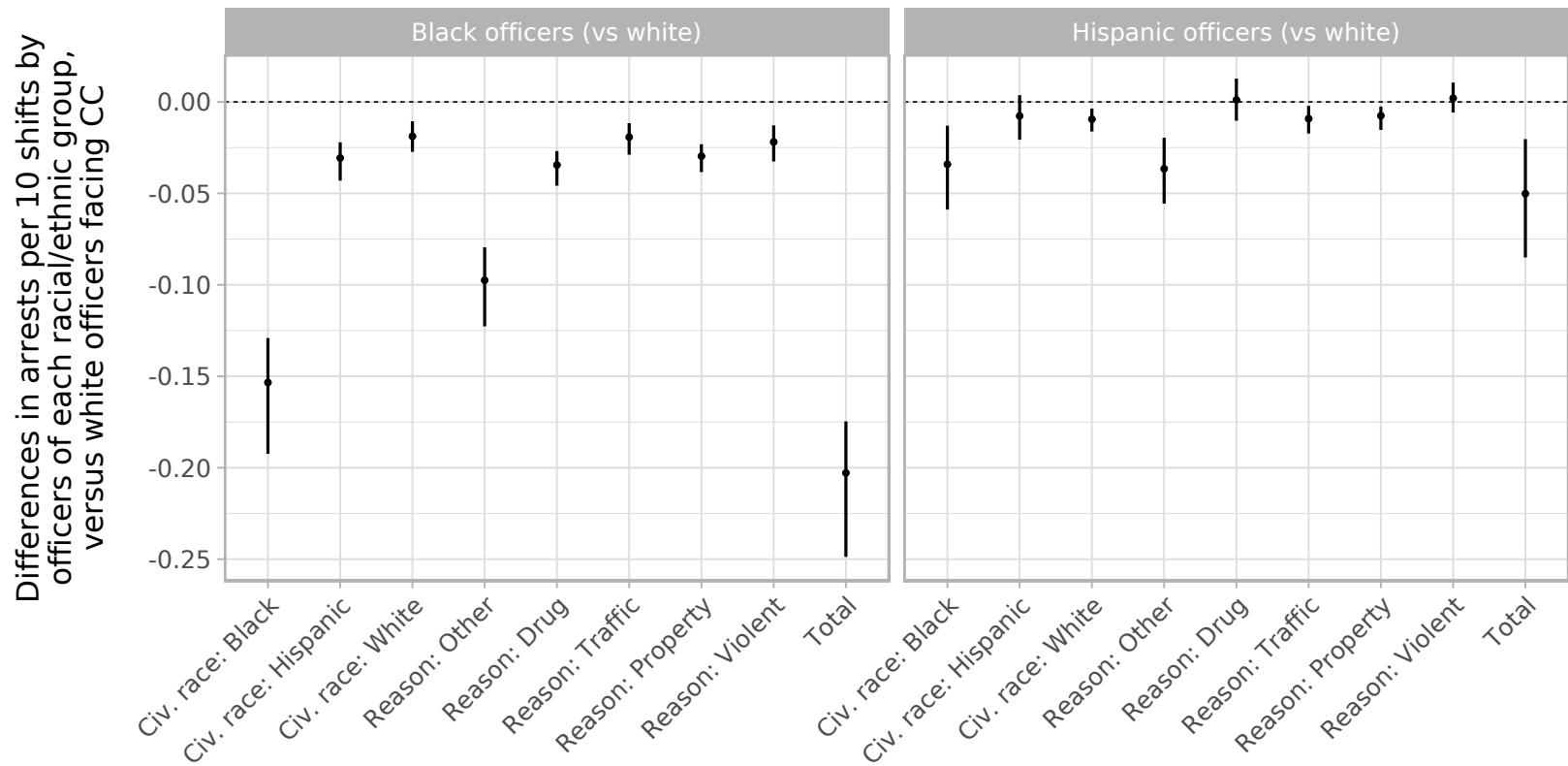


Figure 8: **Uses of Force.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of uses of force by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Corresponding regression tables are reported in SI Section C.6, and results from alternative estimators are shown in SI Section C.7.

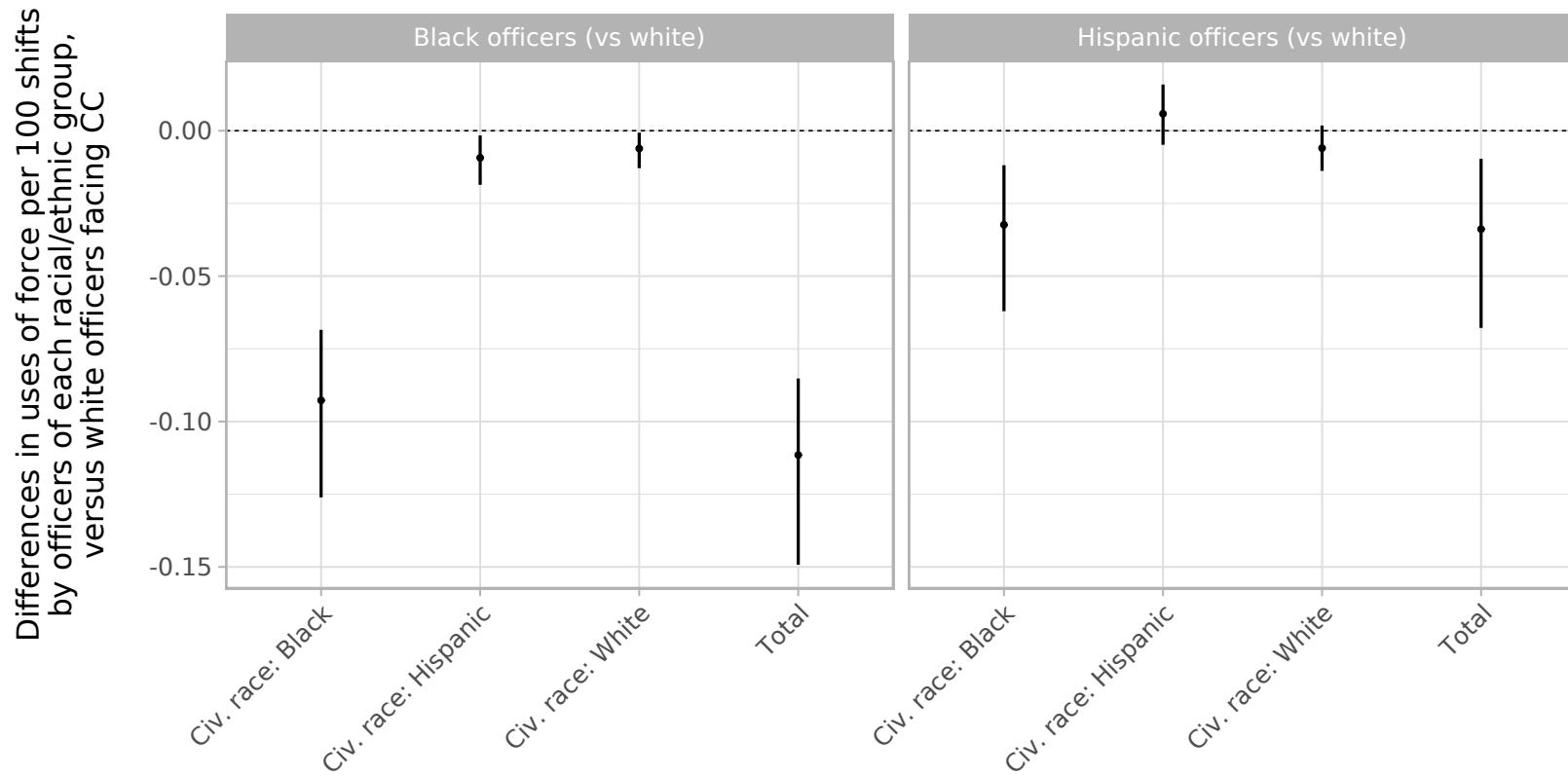
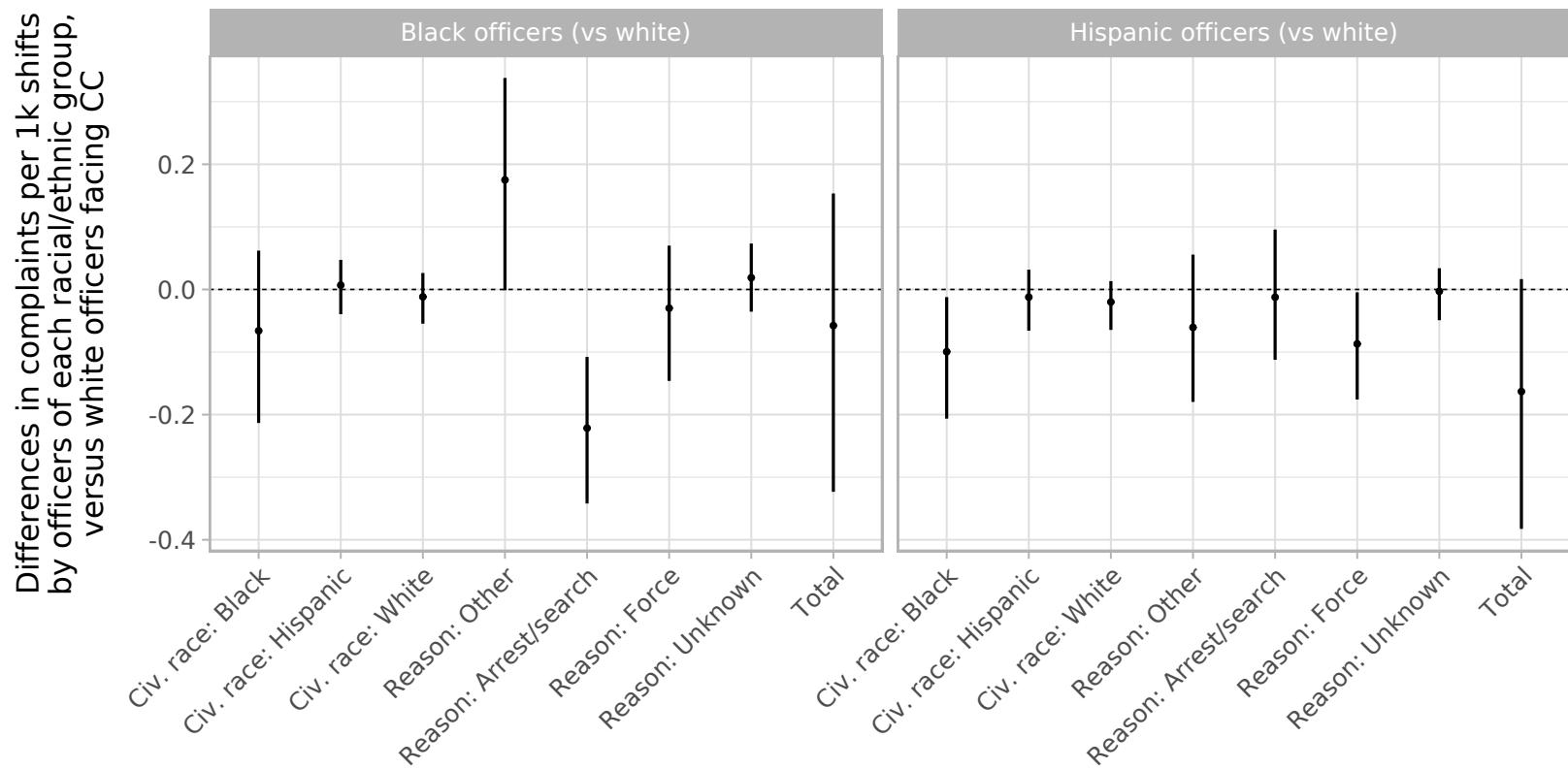


Figure 9: **Complaints.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian complaints registered against black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Corresponding regression tables are reported in SI Section C.6, and results from alternative estimators are shown in SI Section C.7.

24



our analysis of the previous three behaviors due to these data constraints. In particular, the exact time of the incident which generated a complaint is often unknown. By necessity, if an incident is reported on the date of the accused officer's shift, we assume the incident occurred while the officer was on duty. (If complaint-generating incidents occur while officers are off-duty, after departing their assigned beat, the CC assumption is likely to be violated.) Given these limitations, we advise readers to treat our conclusions regarding civilian complaints as tentative. However, we include these data in our analysis because they offer a unique civilian perspective on the appropriateness of police behavior that is difficult to obtain from officer-generated reports, and suggest several patterns which future research should further examine.

8 Officers' Spanish Language Proficiency

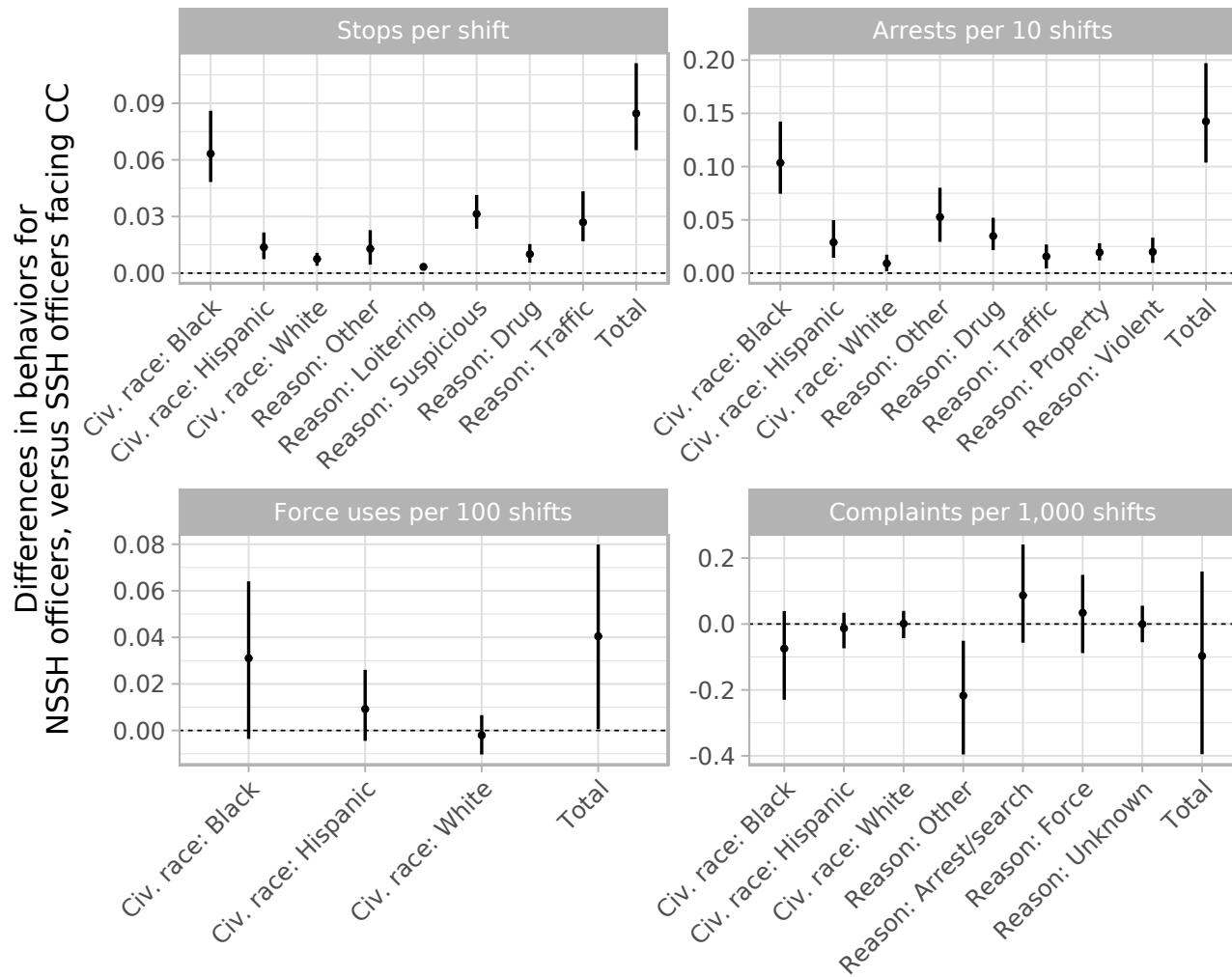
Our initial analysis indicated that white and Hispanic officers behaved comparably toward Hispanic civilians, suggesting hiring and deploying additional similarly situated Hispanic officers would not alter the treatment of Hispanic civilians by police, on average. However, this analysis does not account for the role of language in police-civilian interactions. According to a Pew Research Center study, roughly 77% of Hispanic residents in the Chicago metropolitan area speak Spanish at home as of 2015 (Krogstad and Lopez, 2017). However, personnel data we obtained through open records requests show less than half of Hispanic CPD officers can speak Spanish. This could preclude effective communication with immigrant populations. Moreover, Hispanic officers who cannot speak Spanish may diverge from many Hispanic residents in their cultural backgrounds and lived experiences. If Hispanic officers cannot speak the native language of the co-ethnic civilians they serve, the effects of co-ethnic policing may not materialize.

To investigate this possibility, we use the same analytic approach as in the previous section to compare the behavior of Hispanic officers who can and cannot speak Spanish on their volume of stops, arrests, uses of force and civilian complaints. These comparisons, displayed in Figure 10, reveal important heterogeneity in officer behavior within this varied ethnic group.¹⁰ Relative to Hispanic officers who speak Spanish and are working in the same places and times, Hispanic officers who cannot speak Spanish exhibit substantially higher numbers of stops, arrests and uses of force. For example, non-Spanish-speaking Hispanic officers make roughly 9% more stops and 13% more arrests of Hispanic civilians,

¹⁰SI Figure A10 shows similar differences between Spanish-speaking Hispanic officers and white officers.

relative to Spanish-speaking Hispanic officers facing a comparable pool of civilian activity. Perhaps counterintuitively, non-Spanish-speaking officers stop significantly more Hispanic civilians than even white officers. However, like black officers, Spanish-speaking Hispanic officers receive more minor complaints (those falling into the “other” category).

Figure 10: **Spanish Language Ability.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian stops, arrests, uses of force, and civilian complaints generated by non-Spanish-speaking Hispanic (NSSH) officers per shift, relative to Spanish-speaking Hispanic (SSH) facing common circumstances (CC) officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat.



9 Discussion and Conclusion

A string of fatal encounters between white police officers and unarmed civilians of color has led to renewed calls for law enforcement reforms. Prominent among these proposals is the racial diversification of police agencies, but the effects of this widely studied reform have remained uncertain due to severe data limitations. Using an unusually rich individual-level data set on police activity in Chicago, we provide the most credible assessment to date of the impact of this reform on police-civilian interactions.

Our results reveal a much more nuanced story than previous accounts—which are largely based on agency-level summary statistics (e.g. McCluskey and McCluskey, 2004; Smith, 2003)—have been able to show. Using demographic and political data on officers, we first show that officers are geographically assigned in ways that increase interactions with co-racial and co-ethnic residents. Though the majority of officers do not live in the police districts where they are assigned to work, they do tend to live among co-racial and co-ethnic residents, suggesting lower levels of inter-group contact when off duty. In addition, while the political preferences of black officers are fairly comparable to co-racial city residents, white and Hispanic officers are markedly more conservative than their civilian counterparts—a noteworthy pattern given the correlation between political conservatism and support for punitive criminal justice policies.

After establishing these descriptive portraits of police officers, we report a range of analyses characterizing police behavior and how racial diversity translates into meaningful but sometimes unexpected differences in the lives of the policed. Faced with the same working conditions, black officers are less likely to stop, arrest, and use force against civilians, especially black civilians, relative to white officers. These disparities are driven by a reduced focus on the enforcement of discretionary stops and arrests for petty crimes, including drug offenses, which have long been thought to fuel mass incarceration (Alexander, 2010). In contrast to these drastic differences in the policing of petty crime, black officers' enforcement of violent crime is only slightly lower than that of white officers. But this pattern, which closely comports with the hopes of advocates of racial diversification, is complicated by several additional results. For one, while Hispanic officers display lower levels of enforcement activity than whites overall, their behavior toward Hispanic civilians is broadly comparable to that of white officers. And after accounting for language ability, we find non-Spanish speaking Hispanic officers engage Hispanic civilians far more often than their Spanish-speaking counterparts.

Our analysis of civilian-generated complaints provides suggestive evidence of the role of civilian behavior. While black officers engage in lower levels of enforcement activity than white officers, they receive a roughly equal number of civilian complaints. This pattern could be explained by several mechanisms. On the one hand, it may indicate that black officers engage in misconduct at higher rates, though we find this unlikely given that most of the complaint disparity stems from allegations of minor infractions. Another explanation is that civilians judge the behavior of black officers more harshly than that of white officers, or that civilians believe complaints against white officers are less likely to be taken seriously and therefore opt not to file them in a variety of situations. Our ability to parse these mechanisms with current data is limited, and we encourage future research to investigate these patterns further. Future work should also seek to discern whether the patterns we reveal stem from differential levels of proactive behavior across racial and ethnic groups of officers, as opposed to differential reactions by civilians toward officers of various groups. Doing so would likely require close, unobtrusive observation of police-civilian encounters, perhaps via the analysis of dash or body-worn camera footage.

Finally, while our analysis focuses on evaluating policies related to race and ethnicity, other officer attributes—most notably, officer gender (Lockwood and Prohaska, 2015; Raganella and White, 2004)—that deserve renewed and extended analysis. Our data and research design offer an immediate path for researchers to explore the impact of this and other personnel related reforms. Taken together, our results show that the effects of racial diversification are neither simple nor monolithic. Like the civilians they serve, officers are multidimensional. Crafting effective personnel reforms requires thinking beyond the coarse racial and ethnic categories typically used in diversity initiatives, and consideration of how multiple officer attributes relate police to the civilians they serve.

References

- Alexander, Michelle. 2010. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: The New Press.
- Antonovics, Kate and Brian G. Knight. 2009. “A new look at racial profiling: Evidence from the Boston Police Department.” *The Review of Economics and Statistics* 91(1).
- Ba, Bocar A. 2019. “Going the Extra Mile: the Cost of Complaint Filing, Accountability, and Law Enforcement Outcomes in Chicago.” *Working paper*.
- Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. University of Chicago Press.
- Cao, Liqun, James Frank and Francis T. Cullen. 1996. “Race, community context and confidence in the police.” *American Journal of Police* 15(1):3–22.
- Clinton, Joshua D. and David E. Lewis. 2008. “Expert opinion, agency characteristics, and agency preferences.” *Political Analysis* 16(1):3–20.
- DOJ/EEOC. 2016. “Advancing Diversity in Law Enforcement.” <https://www.justice.gov/crt/case-document/file/900761/download>.
- Donohue III, John J. and Steven D. Levitt. 2001. “The impact of race on policing and arrests.” *The Journal of Law and Economics* 44(2):367–394.
- Eckhouse, Laurel. 2019. “Race, Party, and Representation in Criminal Justice Politics.” *The Journal of Politics* 81(3).
- Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2017. “Using a Probabilistic Model to Assist Merging of Large-scale Administrative Records.” *Working Paper*.
- Federico, Christopher M. and Justin W. Holmes. 2005. “Education and the interface between racial perceptions and criminal justice attitudes.” *Political Psychology* 26:47–75.
- Forman Jr., James. 2017. *Locking up our own: Crime and punishment in Black America*. Farrar, Straus and Giroux.
- 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* p. 105881.

- Gelman, Andrew, Jeffrey Fagan and Alex Kiss. 2007. "An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias." *Journal of the American Statistical Association* 102(429):813–823.
- Goodyear, Sarah. 2014. "Should Cops Have to Live Where They Work?" *CityLab* . <https://www.citylab.com/equity/2014/08/should-cops-have-to-live-where-they-work/378858/>.
- Gramlich, John. 2017. "Black and white officers see many key aspects of policing differently." <https://www.pewresearch.org/fact-tank/2017/01/12/black-and-white-officers-see-many-key-aspects-of-policing-differently/2017>.
- Greiner, James D. and Donald B. Rubin. 2011. "Causal effects of perceived immutable characteristics." *Review of Economics and Statistics* 93:775–785.
- Harvey, Anna and Taylor Mattia. 2019. "Reducing Racial Disparities in Crime Victimization." *Working Paper* . https://s18798.pcdn.co/annaharvey/wp-content/uploads/sites/6417/2019/12/Victimization_Harvey_Mattia.pdf.
- Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American Statistical Association* 81(396):945–60.
- Johnson, David J., Trevor Tress, Nicole Burkel, Carley Taylor and Joseph Cesario. 2019. "Officer characteristics and racial disparities in fatal officer-involved shootings." *Proceedings of the National Academy of Sciences* . Published online ahead of print July 22, 2019. <https://www.pnas.org/content/early/2019/07/16/1903856116>.
- Knox, Dean and Jonathan Mummolo. 2020. "Making Inferences About Racial Disparities in Police Violence." *Proceedings of the National Academy of Sciences* 117:1261–1262.
- Knox, Dean, Will Lowe and Jonathan Mummolo. Forthcoming. "Administrative Records Mask Racially Biased Policing." *American Political Science Review* . https://papers.ssrn.com/sol3/papers.cfm?abstract_id=33363387.
- Krogstad, Jens Manuel and Mark Hugo Lopez. 2017. "Use of Spanish declines among Latinos in major U.S. metros." *Pew Research Center* . <https://www.pewresearch.org/fact-tank/2017/10/31/use-of-spanish-declines-among-latinos-in-major-u-s-metros/>.

- Legewie, Joscha and Jeffrey Fagan. 2016. “Group threat, police officer diversity and the deadly use of police force.” *Columbia Public Law Research Paper*. <http://raceandpolicing.issuelab.org/resources/25209/25209.pdf>.
- Lerman, Amy and Vesla Weaver. 2014. *Arresting Citizenship: The Democratic Consequences of American Crime Control*. University of Chicago Press.
- Levitt, Steven D. 2012. “Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not.” *The Journal of Economic Perspectives* 18(1):163–190.
- Lewis, David E. 2010. “Measurement and public service motivation: New insights, old questions.” *International Public Management Journal* 13(1):46–55.
- Linos, Elizabeth and Nefara Riesch. 2019. “Thick Red Tape and the Thin Blue Line: A Field Study on Reducing Administrative Burden in Police Recruitment.”
- Lockwood, Daniel and Ariane Prohaska. 2015. “Police Officer Gender and Attitudes toward Intimate Partner Violence: How policy can eliminate stereotypes.” *International Journal of Criminal Justice Studies* 10(1):77–90.
- McCluskey, Cynthia Perez and John D. McCluskey. 2004. “Diversity in policing: Latino representation in law enforcement.” *Journal of Ethnicity in Criminal Justice* 2:67–81. <https://www.pewresearch.org/fact-tank/2017/01/12/black-and-white-officers-see-many-key-aspects-of-policing-differently/2017>.
- McCrary, Justin. 2007. “The effect of court-ordered hiring quotas on the composition and quality of police.” *American Economic Review* 97(1):318–353.
- Raganella, Anthony J. and Michael D. White. 2004. “Race, gender, and motivation for becoming a police officer: Implications for building a representative police department.” *Journal of criminal justice* 32(6):501–513.
- Reichel, Philip L. 1988. “Southern slave patrols as a transitional police type.” *American Journal of Police* 7:51–77. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=33363387.
- Rim, Nayoung, Roman G. Rivera and Bocar A. Ba. 2019. “In-group Bias and the Police: Evidence from Award Nominations.” *Working paper*.

- Rivera, Roman G. and Bocar A. Ba. 2019. “The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago.” *Working paper*.
- Rubin, Donald B. 1974. “Estimating causal effects of treatments in randomized and non-randomized studies.” *Journal of Educational Psychology* 66:688–701.
- Rubin, Donald B. 1990. “Formal mode of statistical inference for causal effects.” *Journal of statistical planning and inference* 25(3):279–292.
- Sen, Maya and Omar Wasow. 2016. “Race as a bundle of sticks: Designs that estimate effects of seemingly immutable characteristics.” *Annual Review of Political Science* 19:499–522.
- Shusta, Robert M., Deena R. Levine Phillip R. Harris and Herbert Z. Wong. 1995. *Multicultural law enforcement: Strategies for Peacekeeping in a Diverse Society*. 9 ed. Yale University Press.
- Sklansky, David Alan. 2005. “Not your father’s police department: Making sense of the new demographics of law enforcement.” *J. Crim. L. & Criminology* 96:1209–1244.
- Smith, Brad. 2003. “The Impact of Police Officer Diversity on Police-Caused Homicides.” *Policy Studies* 31(2):147–162.
- Turner, K. B., David Giacopassi and Margaret Vandiver. 2006. “Ignoring the past: Coverage of slavery and slave patrols in criminal justice texts.” *Journal of Criminal Justice Education* 17(1):181–195.
- Tyler, T.R. and J. Fagan. 2008. “Legitimacy and cooperation: Why do people help the police fight crime in their communities?” *Ohio State Journal of Criminal Law* 6(1):231–275.
- Weaver, Vesla M. 2007. “Frontlash: Race and the development of punitive crime policy.” *Studies in American political development* 21:230–265.
- Weitzer, Ronald and Steven A. Tuch. 2004. “Reforming the police: Racial differences in public support for change.” *Criminology* 42(2):391–416.
- Wilson, James Q., ed. 1989. *Bureaucracy: What government agencies do and why they do it*.
- Wilson, James Q. and George L. Kelling. 1982. “Broken windows: the police and neighborhood safety.” *Atlantic Monthly* 249(3):29–38.

Supplemental Materials

Table of Contents

A Detailed Description of Data	2
A.1 CPD Data	2
A.2 Coding race	3
A.3 U.S. Census Merge	3
A.4 Voter File Merge	3
A.5 Preprocessing of patrol assignments	4
A.6 Preprocessing of police behavior	5
B Causal Estimand in Behavioral Analysis	5
C Additional Results	7
C.1 Race and District Assignment	7
C.2 Officer Race and Patrol Assignments	9
C.3 Officers' Political Preferences	12
C.4 Officer Behavior: Overall Means	15
C.5 Officer Behavior: Shift Duration	17
C.6 Officer Behavior: Regression Tables	18
C.7 Officer Behavior: Alternative Estimators	23
C.8 Officers' Spanish Language Proficiency	29
C.9 Officer Gender	31
C.10 Robustness Checks: Multiple Stopping Officers	36

A Detailed Description of Data

A.1 CPD Data

The administrative data from CPD used in this study span multiple data sets collected in collaboration with the Invisible Institute, Sam Stecklow, and Emma Herman over the course of three years (2016-2019). We obtained these records from the Chicago Police Department or Chicago Department of Human Resources via Freedom of Information Act (FOIA) or through court ordered releases stemming from requests made by Invisible Institute and Jaime Kalven. CPD provided the following data: rosters of all available current and past officers up to 2018, data on officers allegation of misconduct from 2000 to 2016 (complaints), unit movement data for individual officers from the 1930s to 2016, Tactical Response Reports from 2004 to 2018 (i.e. use of force reports), and arrest data with arresting officers and arrestee demographic information from 2001 to 2017. The Chicago Department of Human Resources provided data on officers' language skills up to 2019 and officers' home address in 2004, 2005 and in early 2019. We supplement our core data with data on "Stop, Question and Frisk" (SQF) activity between 2012-2015, which was publicly shared by the Lucy Parson's Lab. Finally, the Automated Daily Attendance and Assignment sheet data for each police district between 2012 and 2015 was obtained via a FOIA request to the CPD and shared by Rachel Ryley.

These data and others have been used to construct rich profiles of Chicago Police Officers. While no file contains a unique identifier (star numbers change over time, names are common, etc.), we constructed unique officer profiles through a successive merge process described here. Each file contains some identifying information such as of demographic data (birth year, race, gender) or other characteristics (name, start/badge number, appointed date, resignation date, current unit). We used these identifying characteristics to first de-duplicate officers within a file and to then merge to pre-existing officer data with inter-file unique identifiers. The merging process itself is an iterative-pairwise matching method, where the officers in each data set are repeatedly merged on identifying characteristics and any successful 1-to-1 match in a round removes the matched officers from the next round of merging.

The resulting data contains records on 33,000 police officers appointed between March of 1936 to February of 2018. The number of years and officers varies across analyses in our paper due to missing data (for example, assignment data only exists for the years 2012–2015).

A.2 Coding race

We determine race/ethnicity of CPD officers based on demographic data obtained from the CPD through FOIA. The CPD classifies race/ethnicity in 6 mutually exclusive groups: white/Caucasian, Hispanic, Black/African American, Asian/Pacific Islander, Native American/Native Alaskan, and unknown/missing. If an officer has multiple races associated with them across different data sets, we aggregate by most common non-missing races.

For Census and American Community Survey data, we construct corresponding race categories as follows: any Hispanic individual is coded Hispanic; white and black are comprised of individuals who are coded as not Hispanic and white (black) alone.

A.3 U.S. Census Merge

District and beat demographic data was constructed using the 2010 US Census data and the CPD's pre-2012 beat map. The centroid of each census tract was identified, then the demographic information of all the centroids inside a beat were aggregated to determine the beat's population and demographic composition. District demographics were determined by aggregating across all beats within that district. Post-2012 district and beat demographics were constructed based on the pre-2012 beat data discussed previously and using a crosswalk that maps pre-2012 beats to current (2018) beats and their respective districts.

A.4 Voter File Merge

We merged CPD employee data with a nationwide commercial voter file, maintained by L2 Inc., using the Fellegi-Sunter model as implemented in `fastLink` (Enamorado, Fifield and Imai, 2017). The merge was conducted using first, middle, and last name, along with year of birth. Because a direct merge of the CPD roster with a voter file of more than 180 million observations would result in over 6 trillion pairwise comparisons, we employ additional restrictions for computational feasibility. We focus our attention on voter file observations that have a postal code in the greater Chicago area, reducing the number of pairwise comparisons to slightly over 100 million. For string-valued first and last name, we evaluate the quality of a match by Jaro-Winkler similarity, a common string similarity measure widely used for name comparison that employs a character-wise comparison of two strings while placing more emphasis in their first four characters. This similarity is

discretized into three levels (different, similar, and identical/almost identical) by thresholding at 0.85 and 0.94. For year of birth, the absolute value of the difference is thresholded at 1 year (i.e., agreement in year of birth is declared for observations that are at most one year apart). For middle names, we evaluate whether candidate matches have an identical value. Based on these discrete comparisons, the probabilistic model behind `fastLink` assigns a probability of being a match for each pair of records. The intuition is simple: the more that a pair of records agree in the linkage fields, the more likely that pair of records is a match. In total, 21,276 out of 33,645 police officers were matched, for a match rate just above the 63% mark; only pairs with a fitted match probability greater than 0.9 are retained. We then use the resulting mapping to obtain L2’s proprietary predictions of each matched officer’s current political orientation (Republican, Democratic, and non-partisan, with an extremely small number of third-party labels that are not depicted in analysis), along with public records of general election and party primary turnout that are aggregated and maintained by L2.

A.5 Preprocessing of patrol assignments

We restrict analysis to patrol assignments in which black, Hispanic, or white officers serve. Asian/Pacific Islander and Native American/Alaskan Native officers are not examined due to small sample sizes. Within this subset, we further drop non-standard assignments (notably including “station supervisor” and “station security” assignments, as well as special assignments for training, compensatory time, and excused sick leave). Patrol assignments in which officers are indicated as non-present are also dropped. These steps are intended to ensure that officers nominally patrolling a beat are in fact actively circulating in the assigned geographic area, improving the plausibility of the CC assumption. For the same reason, we drop double shifts (patrol assignment slots in which the assigned officer served for more than one shift on the same day) to address the possibility that officers behave differently due to fatigue in these circumstances. We also eliminate officers assigned to non-standard watches (i.e., other than first through third watches). Finally, we drop officers at ranks other than “police officer.” This step eliminates police sergeants, who serve in 8% of beat assignments but make very few stops and arrests, as well as legal officers, helicopter pilots, explosives technicians, and canine handlers.

A.6 Preprocessing of police behavior

Events are merged to the remaining patrol assignments based on officer ID and date. This step discards a large number of events, including those involving officers of higher ranks and incidents occurring on rest days. For stops, arrests, and uses of force, we drop all events that occur outside of the reported patrol start/stop times. A key difference in the handling of complaints, where retrospectively reported incident times are of extremely low quality, is that we do not implement this step (i.e., all complaints about incidents on the date of a recorded patrol assignment are assumed to be occur while the officer is on duty, an assumption that may potentially be violated).

Stops for “dispersal” and “gang and narcotics-related loitering” are coded as loitering stops; those that are “gang / narcotics related” are coded as drug stops; “investigatory stops” and stops of “suspicious persons” are coded as suspicious behavior; and stops under the “Repeat Offender Geographic Urban Enforcement Strategy (ROGUES)” program are combined with the “other” category. For stops, if a single officer is reported as both primary and secondary stopping officer, only one event is retained.

Arrests for municipal code violations and outstanding warrants are categorized as “other.”

Complaints initiated by other officers, which often relate to truancy or job performance rather than treatment of civilians, are discarded. Complaints can have multiple reported victims and complainants. When this is the case, each complaint is treated as a single event, and all victims and complainants are aggregated. When multiple victims are reported, if any victims are black, the civilian race is coded as black. Otherwise, if any Hispanic victims are reported, civilian race is coded as Hispanic. Otherwise, if any white victims are reported, civilian race is coded as white. If and only if victim demographics are unavailable, we fall back on complainant demographics. If multiple complainants are reported, if any are black, the civilian race is coded as black. Otherwise, if any Hispanic complainants are reported, civilian race is coded as Hispanic. Finally, if any white complainants are reported, civilian race is coded as white.

B Causal Estimand in Behavioral Analysis

At a high level, the goal of our analysis is to evaluate the policy effect of a personnel reform that increases the representation of minorities in the CPD by assigning them to

positions that would otherwise be filled by white individuals. The analysis is conducted at the level of the patrol assignment slot. Commanding officers are assumed to have a fixed set of patrol assignments that must be filled, where each slot is associated with a beat (geographic patrol area) and shift time (temporal window). Multiple slots may be available for a particular beat and shift time, but each slot can be filled by only one officer.

We organize beat assignments into groups, indexed by i , based on unique combinations of month (M_i), day group (weekday/weekend, D_i), shift time (first/second/third watch, S_i), and beat (B_i), or unique MDSBs. Patrol assignment slots within a MDSB are indexed by j . For each slot, the realized pattern of officer behavior is denoted $Y_{i,j}(R_{i,j})$, where $R_{i,j}$ is the race of the officer assigned to a particular slot. Our notation implicitly makes the stable unit treatment value assumption (SUTVA, Rubin, 1990), which requires that (1) there do not exist finer gradations of officer identity (i.e., within the broad racial/ethnic categories used) that would result in differing potential officer behavior, and (2) that potential outcomes do not vary depending on the racial/ethnic identity of officers assigned to other slots.¹

The slot-level policy effect is the difference in potential outcomes (Rubin, 1974) $Y_{i,j}(r) - Y_{i,j}(r')$, the change in behavior that would have realized if an officer of race r had been assigned to the patrol assignment slot, rather than another officer of race r' . These counterfactual differences are fundamentally unobservable; instead, we target the average policy effect within the subset of F MDSBs for which policy effects can be feasibly estimated (i.e., for which variation in officer race exists). This quantity is

$$\delta = \frac{1}{F\bar{A}_i} \sum_{i=1}^F \sum_{j=1}^{A_i} Y_{i,j}(r) - Y_{i,j}(r'),$$

where A_i is the number of patrol assignment slots available within MDSB i and \bar{A}_i is the average slot count across MDSBs. This can be rewritten as the weighted average of

¹We explore the validity of this second assumption to the extent possible in SI Section C.10, in which stops made by two officers are reanalyzed. In this section, we re-compute our estimates of differential stopping behavior after excluding the second reporting officer from our analysis; the resulting estimates are highly similar.

MDSB-specific effects, δ_i , with weights given by A_i .

$$\begin{aligned}\delta &= \sum_{i=1}^F \frac{A_i}{\sum_{i'=1}^F A_{i'}} \left(\frac{1}{A_i} \sum_{j=1}^{A_i} Y_{i,j}(r) - Y_{i,j}(r') \right) \\ &= \sum_{i=1}^F \frac{A_i}{\sum_{i'=1}^F A_{i'}} \delta_i.\end{aligned}$$

As we discuss in Section 2, a key identifying assumption is that

$$Y_{i,j}(r), Y_{i,j}(r') \perp\!\!\!\perp R_i | M_i = m, D_i = d, S_i = s, B_i = b.$$

Informally, this requires that minority officers are not selectively assigned to slots within MDSBs, at least in ways that matter for potential officer behavior. (Hypothetically speaking, this independence condition could be achieved even without adjusting for MDSB if white and nonwhite officers were randomly assigned locations and times to patrol.)

Our primary results estimate this quantity with an ordinary least squares (OLS) regression of the form $Y_{i,j} = \alpha_i + \sum_r \beta_r \mathbf{1}(R_{ij} = r)$, where α_i represents a fixed effect for MDSB i . Unbiasedness of this estimator requires the additional assumption that MDSB-specific policy effects are homogeneous, or that $\delta_i = \delta_R$ for all i . It is well known that when this assumption is violated, OLS recovers the weighted average of δ_i s with weights corresponding to the variance of officer race within strata. To allow for the possibility of non-homogeneous policy effects and other departures from our modeling assumptions, we therefore apply a number of alternative estimators, which are described in detail in SI Section C.7. As we show in SI Figures A6–A9, these alternative results are virtually identical to our primary results.

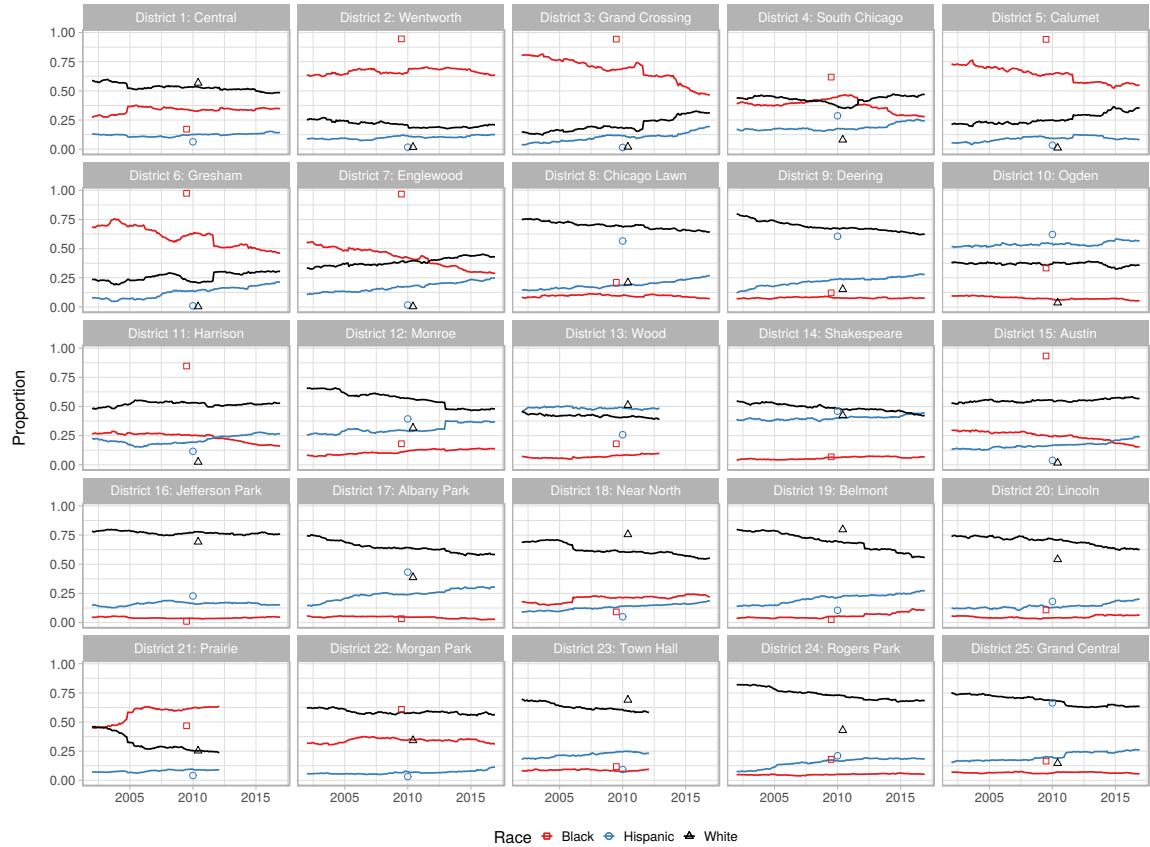
C Additional Results

C.1 Race and District Assignment

The CPD currently subdivides Chicago into 22 policing districts which correspond to CPD units, in which the majority of police officers work. There were 25 districts (numbered 1 - 25) until 2012, at which time 3 smaller districts (ranking 18th, 21st, and 25th in land area²) were eliminated and merged with other districts. Districts 23 and 21 and District 13

²See Chicago Police Department 2008 Annual Report, page 37, <https://home.chicagopolice.org/wp-content/uploads/2014/12/2008-Annual-Report.pdf>

Figure A1: Racial composition of officers' assigned districts. Each panel depicts a geographic police district. Points represent the racial composition of district residents. Lines represent monthly proportions of officers assigned to a district that belong to each racial group. Districts 21, 23, and 13 were eliminated during the observation period.



were eliminated and absorbed into neighboring districts in March and December of 2012, respectively. While District 23 was mostly absorbed by District 19 and most of District 13 was absorbed by District 12, significant parts of District 21 were absorbed by Districts 1, 2, and 9.

Figure A1 displays significant heterogeneity across districts in the racial composition of their officers. There is a correlation between the racial composition of a district's policing unit and that of the population, and the heterogeneity across police units reflects Chicago's racial segregation across geographic areas.

C.2 Officer Race and Patrol Assignments

Among officers assigned to a particular police district, considerable variation exists in the exact patrol assignments that officers receive. We examine each unit individually, tabulating officer race and shift time assignments (first, second and third watch, respectively corresponding to the nominal duty periods of midnight to 8 a.m., 8 a.m. to 4 p.m., and 4 p.m. to midnight). Figure A2 depicts the frequency of each shift period, showing that the pattern of assignments differs dramatically by officer race. For example, white officers in Wentworth (District 2) almost exclusively serve from 4 p.m. to midnight, whereas black officers are more likely to be assigned to mid-day shifts. Similarly, Figure A3 shows that relative to white officers, black officers are far more frequently deployed to assigned beat 202—which roughly corresponds to a patrol area in the district’s southwest corner (depicted in Figure 1) that has extremely high police activity and a high concentration of black residents. These results undermine analyses in a wide array of previous studies that aggregate at high levels of geography (for example, controlling for district or unit assignment) and which assume that officers face homogeneous conditions within these crude groupings.

Figure A2: Assigned shift time by officer race. Each panel depicts officers within a geographic police district. Within a district, each row shows the proportion of shift assignments for black, Hispanic, or white officers to first watch (midnight to 8 a.m.), second watch (8 a.m. to 4 p.m.), and third watch (4 p.m. to midnight). Darker cells indicate a higher proportion of assignments to a shift time, and entries in a row sum to unity. The figure demonstrates that within any particular district, black and Hispanic officers are called to serve at very different times of day ($p < 0.001$ for all within-district statistical tests of independence).

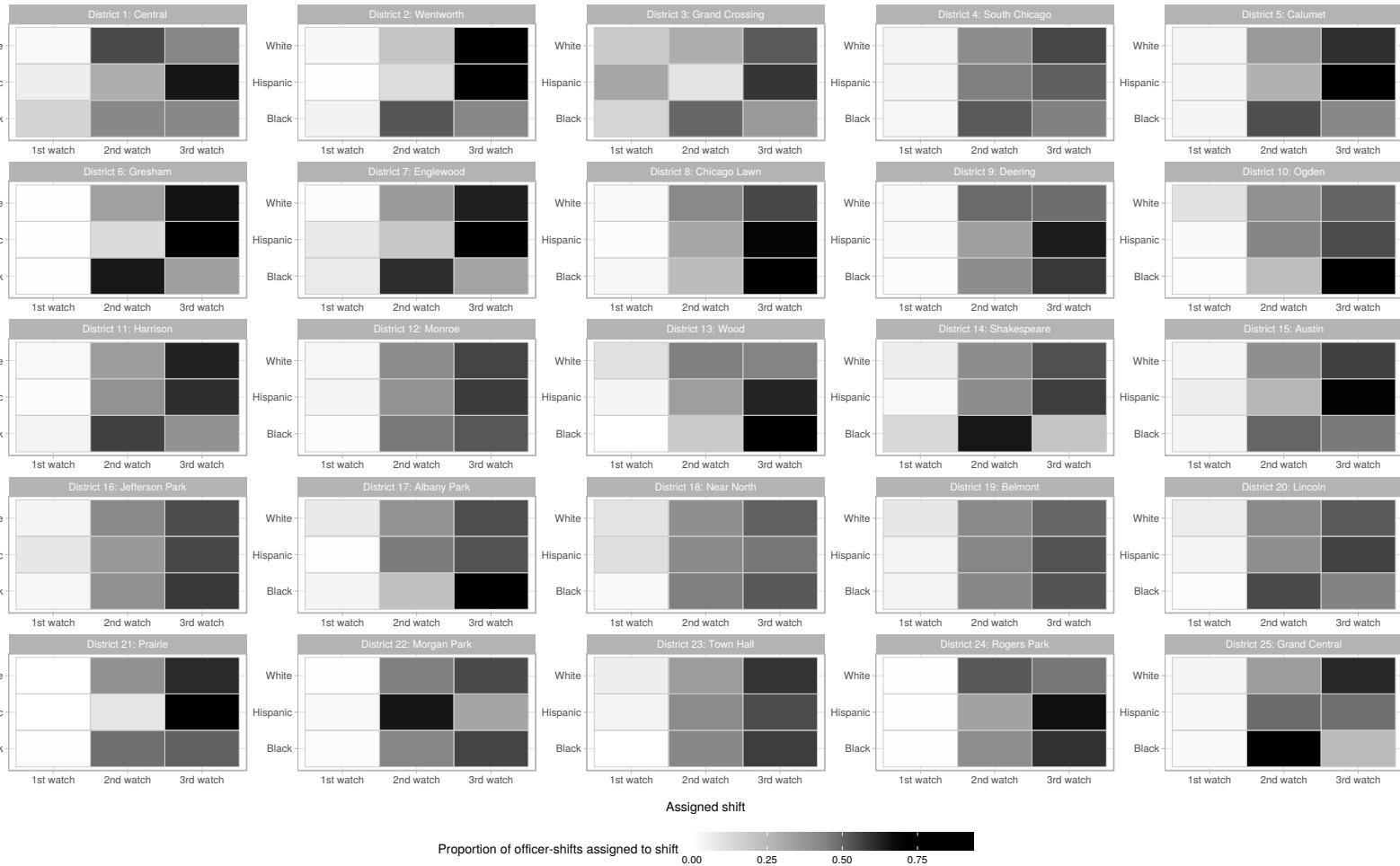
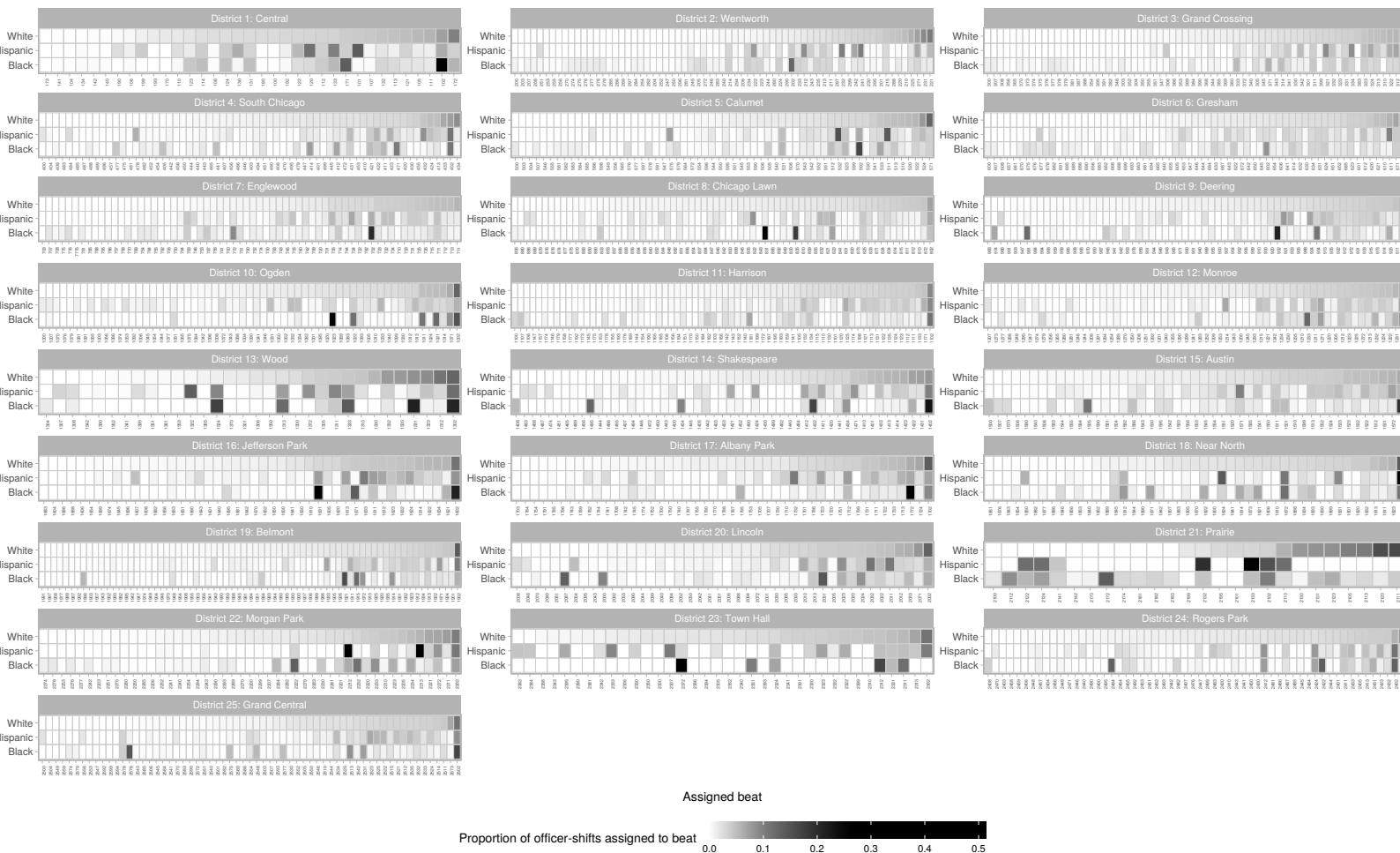


Figure A3: Assigned beat by officer race. Each panel depicts officers within a geographic police district. Within a district, each row shows the proportion of shift assignments for black, Hispanic, or white officers to beats, or geographic patrol areas. Darker cells indicate a higher proportion of assignments to a beat, and entries in a row sum to unity. The figure demonstrates that within any particular district, black and Hispanic officers are called to serve at very different locations ($p < 0.001$ for all within-district statistical tests of independence).



C.3 Officers' Political Preferences

Figure A4: Police are more likely to vote than co-racial and co-ethnic civilians.
 Officer and city resident turnout rates by racial/ethnic groups, cohort, and election.

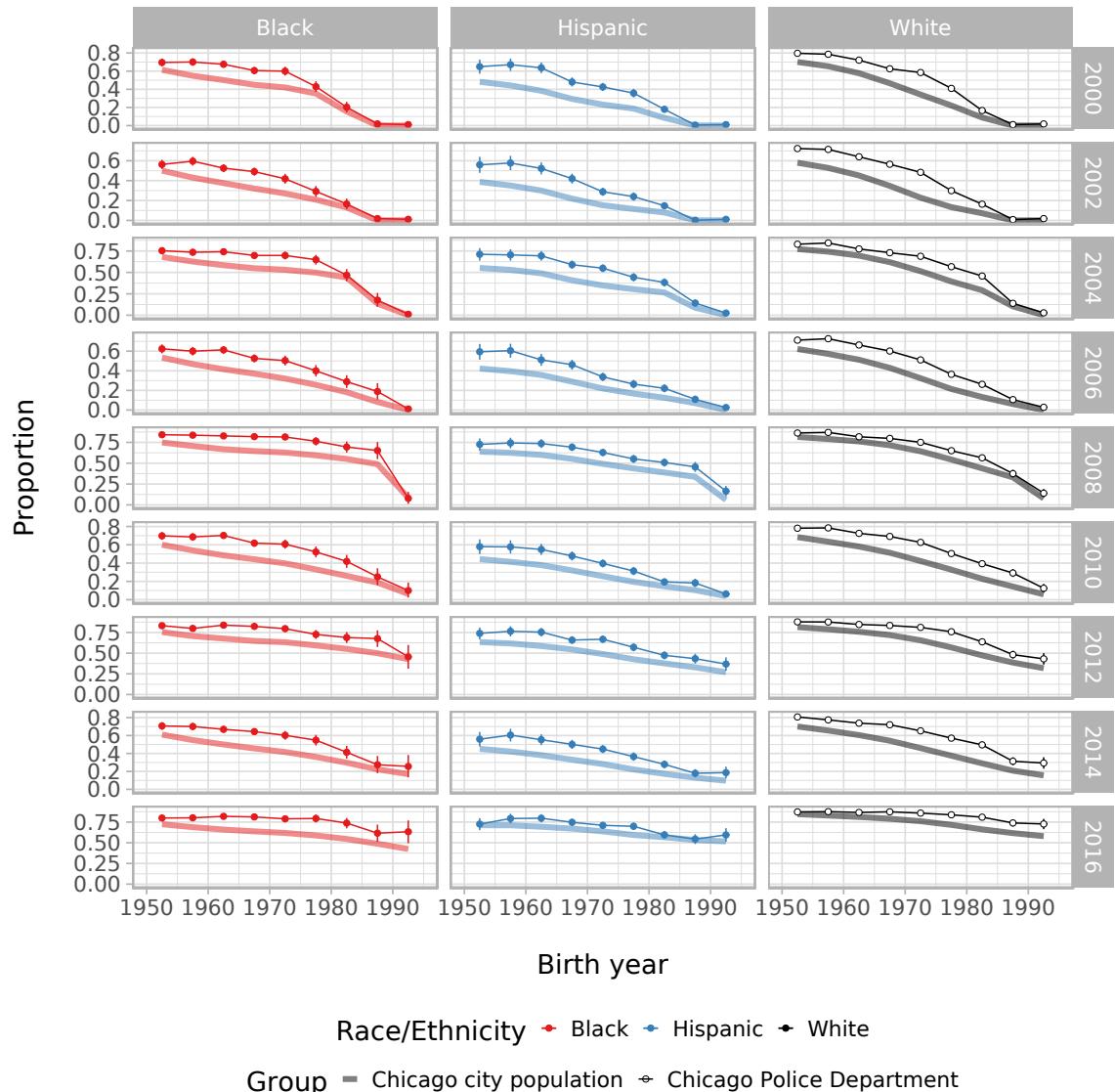
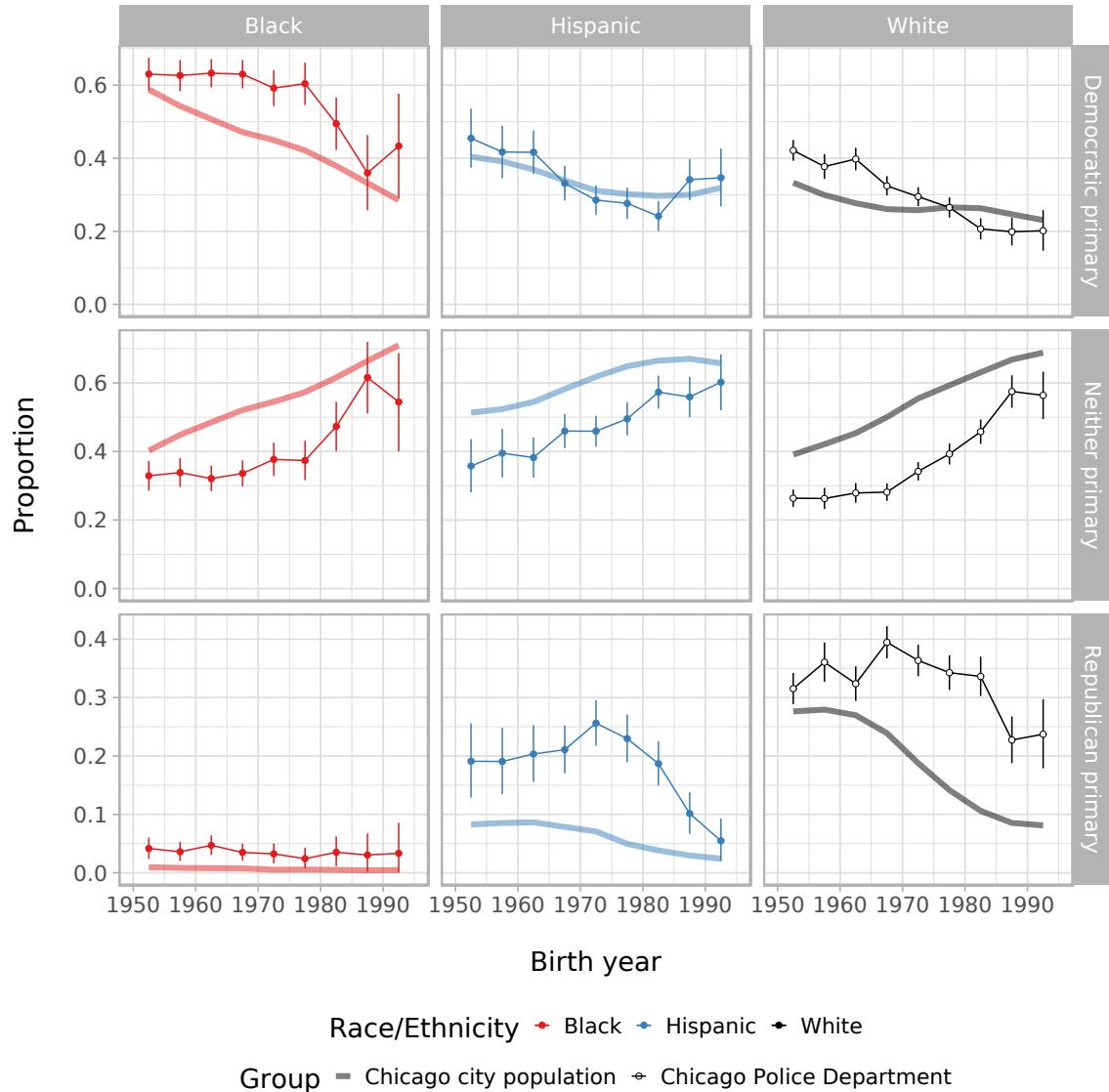


Figure A5: Party identification among officers using 2016 party primary participation as proxy for party ID. Officers aggregated to five-year windows on birth year (1950–1954, 1955–1960, etc.). For yearly bins and previous elections (2000–2016).



C.4 Officer Behavior: Overall Means

Table A1: **Average events per shift.** Mean number of stops, arrests, uses of force, and complaints without adjustment for time or location. Typical behavior is reported for black, Hispanic, and white officers individually, as well as the average pooling three officer races. Records associated with Native American/Alaskan and Asian/Pacific Islander officers are excluded due to small sample sizes. Officer behavior toward Native American/Alaskan and Asian/Pacific Islander civilians is not included for the purposes of computing total and reason-specific events. Note that civilian race is often unknown in complaint records. Values are scaled for ease of interpretation.

Behavior	Mean (Pooled)	Mean (Black off.)	Mean (Hisp. off.)	Mean (White off.)
Stops per shift:				
Civilian race: Black	0.303	0.261	0.309	0.325
Civilian race: Hispanic	0.092	0.021	0.143	0.110
Civilian race: White	0.057	0.021	0.061	0.075
Reason: Other	0.131	0.116	0.135	0.139
Reason: Loitering	0.006	0.003	0.009	0.007
Reason: Suspicious	0.154	0.089	0.173	0.184
Reason: Drug	0.048	0.018	0.063	0.058
Reason: Traffic	0.118	0.081	0.139	0.130
Total	0.458	0.306	0.519	0.517
Arrests per 10 shifts:				
Civilian race: Black	0.564	0.495	0.593	0.591
Civilian race: Hispanic	0.176	0.040	0.252	0.219
Civilian race: White	0.089	0.032	0.098	0.117
Reason: Other	0.310	0.204	0.346	0.354
Reason: Drug	0.094	0.036	0.124	0.113
Reason: Traffic	0.070	0.035	0.078	0.087
Reason: Property	0.152	0.115	0.166	0.168
Reason: Violent	0.211	0.182	0.238	0.216
Total	0.837	0.571	0.952	0.936
Uses of force per 100 shifts:				
Civilian race: Black	0.199	0.145	0.210	0.226
Civilian race: Hispanic	0.041	0.009	0.058	0.050
Civilian race: White	0.026	0.007	0.028	0.036
Total	0.271	0.163	0.302	0.319
Complaints per 1,000 shifts:				
Civilian race: Black	0.542	0.604	0.493	0.530
Civilian race: Hispanic	0.102	0.025	0.147	0.124
Civilian race: White	0.084	0.048	0.081	0.106
Reason: Other	0.988	1.181	0.852	0.941
Reason: Arrest/Search	0.541	0.367	0.596	0.616
Reason: Force	0.386	0.354	0.383	0.407
Reason: Unknown	0.106	0.099	0.120	0.103
Total	2.021	16	2.001	1.951
				2.067

C.5 Officer Behavior: Shift Duration

We consider the possibility that stop, arrest, force, and complaint volume are driven by different amounts of time spent patrolling. Even among officers assigned to a particular shift time (a nominal eight-hour patrol period), minor variation exists in the precise start and end of the officer's duty time. Of the officer-shifts analyzed, 85.5% are 9 hours in duration, with 8.5- and 8-hour shifts making up an additional 7.8% and 5.1%, respectively (percentages are based on rounding shift duration to the nearest 6 minutes). In fixed-effect regression analyses that compare officers within unique MDSB combinations, we estimate that shifts of black officers are 0.0094 hours shorter (roughly 0.1% shorter) than their white counterparts assigned to the same MDSB, and Hispanic officer shift durations are statistically indistinguishable from those of white officers. Because these differences are two orders of magnitude smaller than reported differences in behavior, patrol time disparities are unlikely to be a mechanism driving observed racial gaps in stop, arrest, force, and complaint volume.

C.6 Officer Behavior: Regression Tables

Table A2: Models of Stops, by Civilian Race

	Stops per shift, by civilian race:			
	Black	Hispanic	White	All races
Officer black	-0.133 [-0.156; -0.122]	-0.014 [-0.019; -0.011]	-0.013 [-0.016; -0.011]	-0.161 [-0.188; -0.148]
Officer Hispanic	-0.032 [-0.047; -0.019]	0.005 [-0.001; 0.010]	-0.004 [-0.008; -0.001]	-0.032 [-0.052; -0.016]
MDSB FE	Yes	Yes	Yes	Yes
N	2923458	2923458	2923458	2923458
N _{effective}	2186703	2186703	2186703	2186703
R ²	0.295	0.230	0.183	0.276

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

Table A3: Models of Stops, by Reason for Stop

	Stops per shift, by reason:				
	Loitering	Suspicious	Drug	Traffic	Other
Officer black	-0.004 [-0.005; -0.003]	-0.061 [-0.071; -0.055]	-0.017 [-0.022; -0.014]	-0.053 [-0.067; -0.047]	-0.026 [-0.034; -0.017]
Officer Hispanic	0 [-0.001; 0.002]	-0.019 [-0.026; -0.011]	-0.005 [-0.010; -0.002]	-0.005 [-0.016; 0.002]	-0.003 [-0.010; 0.004]
MDSB FE	Yes	Yes	Yes	Yes	Yes
N	2923458	2923458	2923458	2923458	2923458
N _{effective}	2186703	2186703	2186703	2186703	2186703
R ²	0.170	0.232	0.228	0.268	0.206

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

Table A4: Models of Arrests, by Civilian Race

	Arrests per 10 shifts, by civilian race:			
	Black	Hispanic	White	All races
Officer black	-0.153 [-0.192; -0.129]	-0.031 [-0.043; -0.022]	-0.019 [-0.027; -0.011]	-0.203 [-0.249; -0.175]
Officer Hispanic	-0.034 [-0.059; -0.013]	-0.008 [-0.021; 0.004]	-0.009 [-0.016; -0.004]	-0.050 [-0.085; -0.020]
MDSB FE	Yes	Yes	Yes	Yes
N	2930957	2930957	2930957	2930957
N _{effective}	2193343	2193343	2193343	2193343
R ²	0.183	0.160	0.123	0.185

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

Table A5: Models of Arrests, by Reason for Arrest

	Arrests per 10 shifts, by reason:				
	Traffic	Drug	Property	Violent	Other
Officer black	-0.019 [-0.029; -0.012]	-0.034 [-0.046; -0.027]	-0.030 [-0.038; -0.023]	-0.022 [-0.033; -0.013]	-0.098 [-0.123; -0.079]
Officer Hispanic	-0.009 [-0.017; -0.002]	0.001 [-0.010; 0.013]	-0.008 [-0.015; -0.003]	0.002 [-0.006; 0.011]	-0.037 [-0.056; -0.020]
MDSB FE	Yes	Yes	Yes	Yes	Yes
N	2930957	2930957	2930957	2930957	2930957
N _{effective}	2193343	2193343	2193343	2193343	2193343
R ²	0.161	0.173	0.128	0.114	0.157

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

Table A6: Models of Force, by Civilian Race

	Uses of force per 100 shifts, by civilian race:			
	Black	Hispanic	White	All races
Officer black	-0.093 [-0.126; -0.068]	-0.009 [-0.019; -0.002]	-0.006 [-0.013; -0.001]	-0.112 [-0.149; -0.085]
Officer Hispanic	-0.032 [-0.062; -0.012]	0.006 [-0.005; 0.016]	-0.006 [-0.014; 0.002]	-0.034 [-0.068; -0.010]
MDSB FE	Yes	Yes	Yes	Yes
N	2933232	2933232	2933232	2933232
N _{effective}	2195410	2195410	2195410	2195410
R ²	0.094	0.095	0.096	0.095

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

Table A7: Models of Complaints, by Civilian Race

	Complaints per 1,000 shifts, by civilian race:			
	Black	Hispanic	White	All races
Officer black	-0.066 [-0.213; 0.062]	0.007 [-0.039; 0.047]	-0.012 [-0.055; 0.026]	-0.058 [-0.323; 0.153]
Officer Hispanic	-0.099 [-0.206; -0.012]	-0.012 [-0.066; 0.032]	-0.020 [-0.065; 0.013]	-0.163 [-0.383; 0.017]
MDSB FE	Yes	Yes	Yes	Yes
N	2933278	2933278	2933278	2933278
N _{effective}	2195437	2195437	2195437	2195437
R ²	0.091	0.087	0.089	0.092

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

Table A8: Models of Complaints, by Reason for Complaint

	Complaints per 1,000 shifts, by reason:			
	Arrest/Search	Force	Unknown	Other
Officer black	-0.222 [-0.342; -0.108]	-0.030 [-0.146; 0.070]	0.019 [-0.035; 0.073]	0.175 [-0.001; 0.338]
Officer Hispanic	-0.012 [-0.112; 0.096]	-0.087 [-0.176; -0.005]	-0.003 [-0.049; 0.034]	-0.061 [-0.180; 0.056]
MDSB FE	Yes	Yes	Yes	Yes
N	2933278	2933278	2933278	2933278
N _{effective}	2195437	2195437	2195437	2195437
R ²	0.104	0.087	0.076	0.087

Specification includes fixed effects for each unique combination of month, day group (weekday/weekend), shift time (first/second/third watch), and assigned beat. Cluster-robust 95% confidence intervals are computed by block bootstrap on officers and reported in brackets. N and N_{effective} indicate the number of officer-shifts analyzed and the number of officer-shifts within MDSB containing variation on officer race, respectively.

C.7 Officer Behavior: Alternative Estimators

Our primary analysis of officer behavior uses OLS regression with MDSB fixed effects, of the form $Y_{ij} = \alpha_i + \sum_r \beta_r \mathbf{1}(R_{ij} = r)$, where α_i represents a fixed effect for MDSB i . As we discuss in SI Section B, this estimator will deviate from the desired average policy effect (i.e., the average effect of replacing white officers assigned to a particular patrol assignment slot with a minority officer on resulting stop, arrest, use-of-force, and complaint volume) if MDSB-specific policy effects are highly variable in a way that is associated with the proportion of minority officers that are assigned to MDSBs (in this case, it is well known that OLS recovers the weighted average of MDSB-specific policy effects, where weights are determined by variance of officer race within the MDSB).

To gauge robustness of our results to the violation of this assumption, we present alternative estimates in SI Figures A6–A9 below. The first alternative estimator takes the within-MDSB difference in behavior between average minority- and white-officer patrol assignments, then aggregates these according to the number of patrol assignment slots in each MDSB. Following the notation defined in SI Section B, this estimator can be written as

$$\sum_{i=1}^F \frac{A_i}{\sum_{i'=1}^F A_{i'}} \sum_{j=1}^{A_i} \left(\frac{\sum_{j=1}^{A_i} Y_{ij} \mathbf{1}(D_{ij} = d)}{\sum_{j=1}^{A_i} \mathbf{1}(D_{ij} = d)} - \frac{\sum_{j=1}^{A_i} Y_{ij} \mathbf{1}(D_{ij} = d')}{\sum_{j=1}^{A_i} \mathbf{1}(D_{ij} = d')} \right).$$

To assess the extent to which results are driven by large MDSBs, we further compute the unweighted average of MDSB-specific estimated effects:

$$\frac{1}{F} \sum_{i=1}^F \sum_{j=1}^{A_i} \left(\frac{\sum_{j=1}^{A_i} Y_{ij} \mathbf{1}(D_{ij} = d)}{\sum_{j=1}^{A_i} \mathbf{1}(D_{ij} = d)} - \frac{\sum_{j=1}^{A_i} Y_{ij} \mathbf{1}(D_{ij} = d')}{\sum_{j=1}^{A_i} \mathbf{1}(D_{ij} = d')} \right).$$

Finally, we consider the possibility that observed racial/ethnic differences in officer behavior are driven by differences in experience between minority and white officers. If this were the case, it would undermine the applicability of our results to the effect of a hiring reform that brought in additional minority rookie officers. To examine whether these differences impact our results, we extend the regression specification by adding additional linear and quadratic terms for each officer's length of service.

Figure A6: Stops. Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian stops conducted by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Results presented using four different estimators.

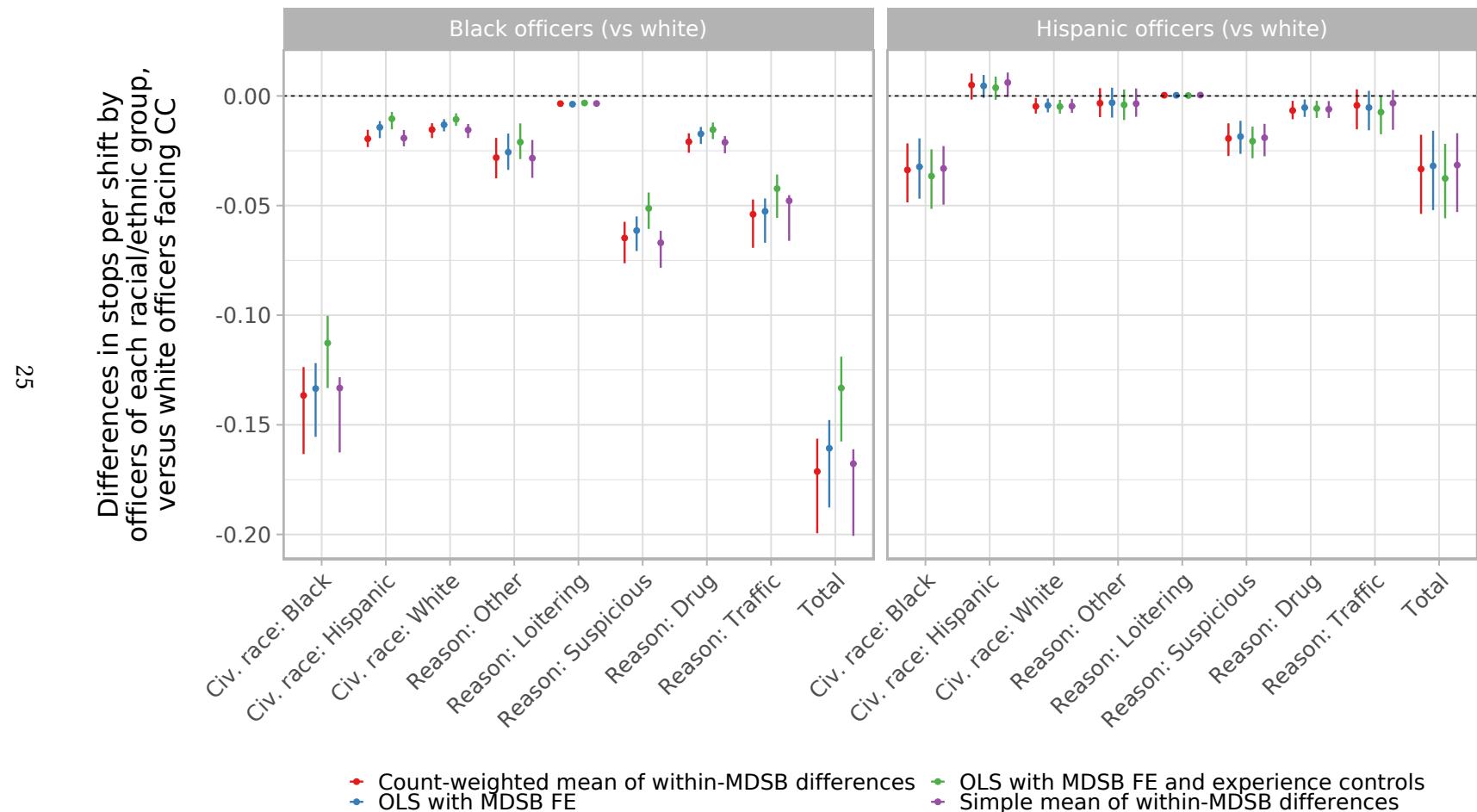


Figure A7: **Arrests.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of arrests conducted by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Results presented using four different estimators.

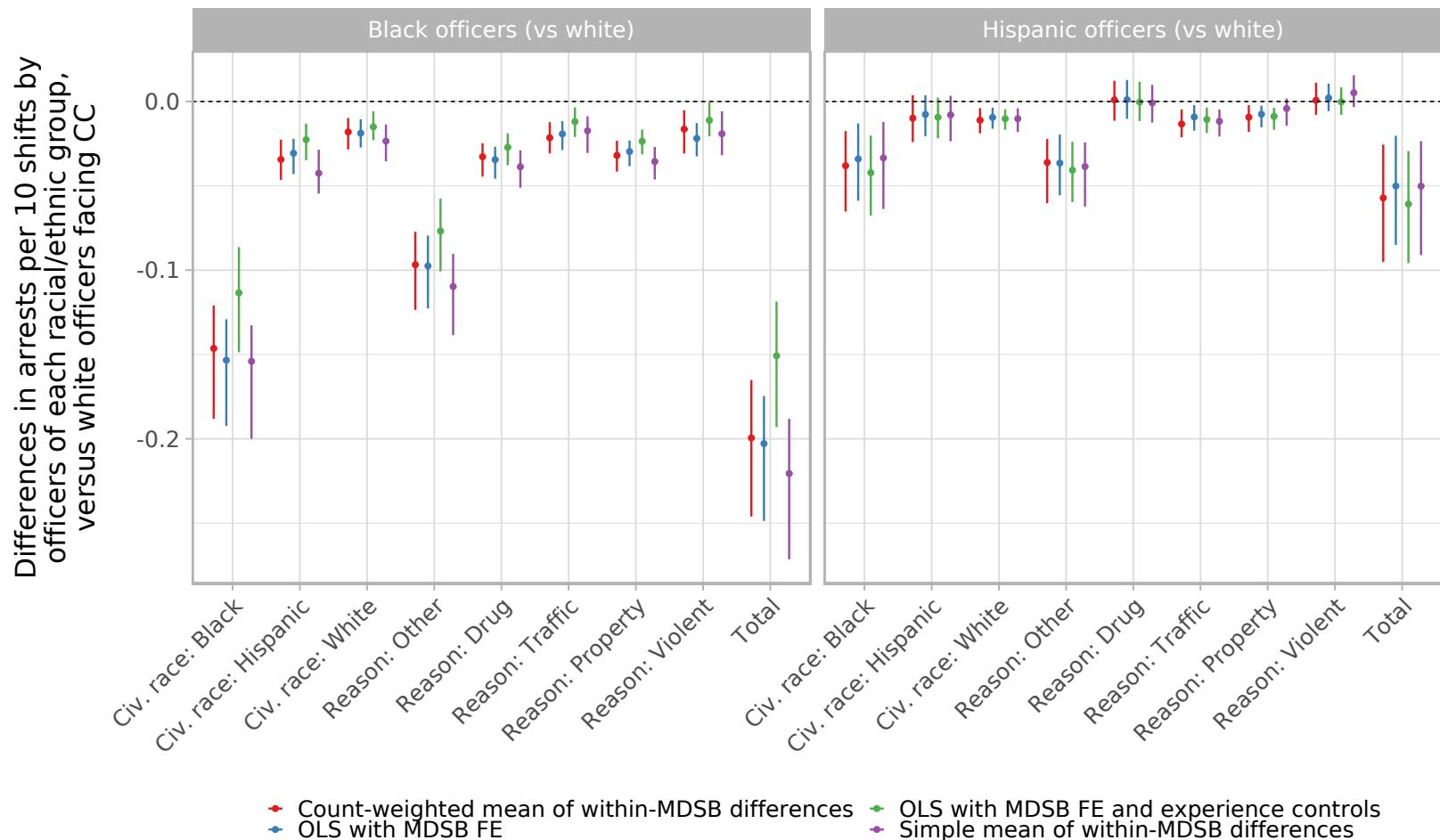


Figure A8: **Uses of Force.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of uses of force by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Results presented using four different estimators.

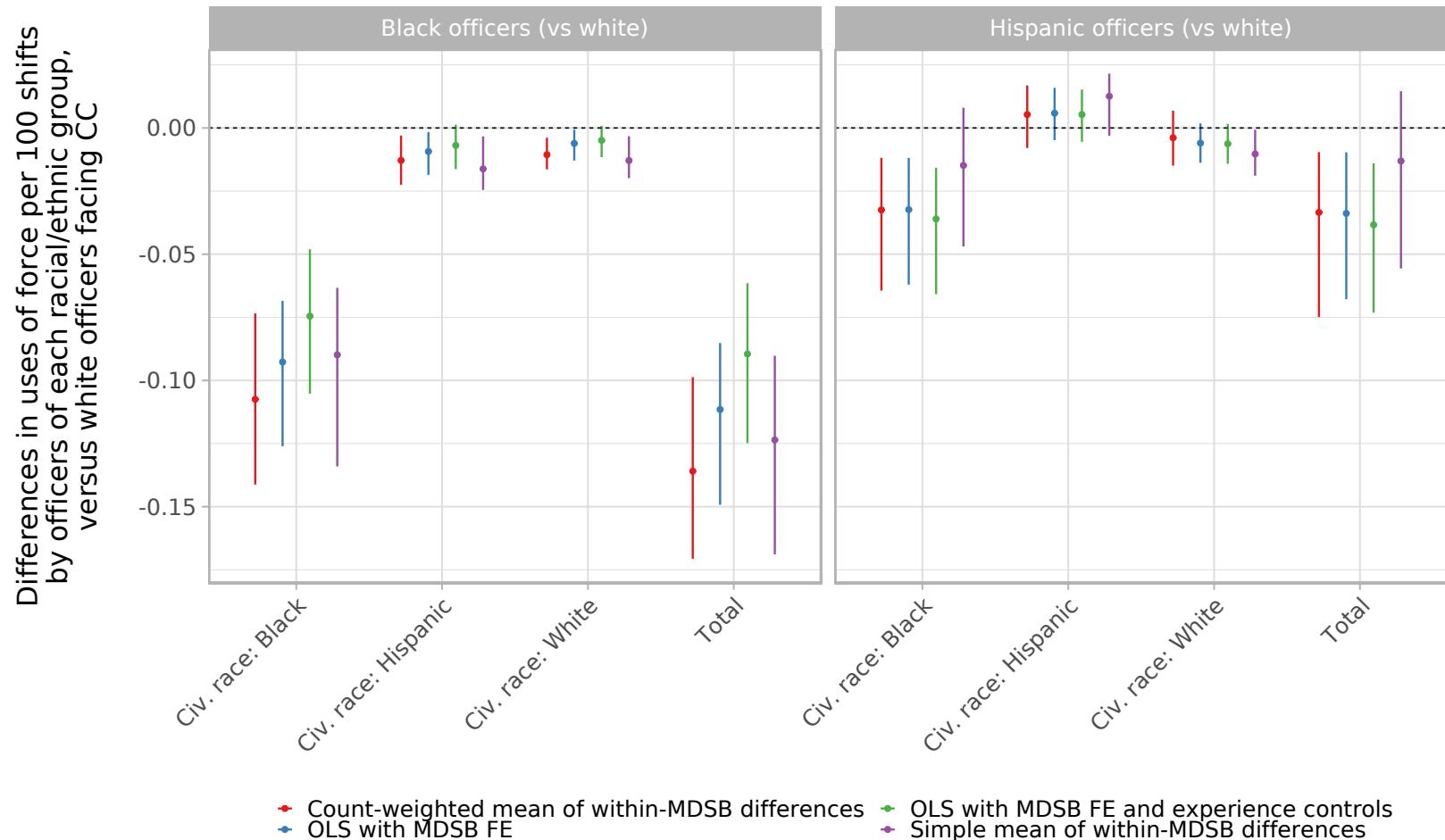
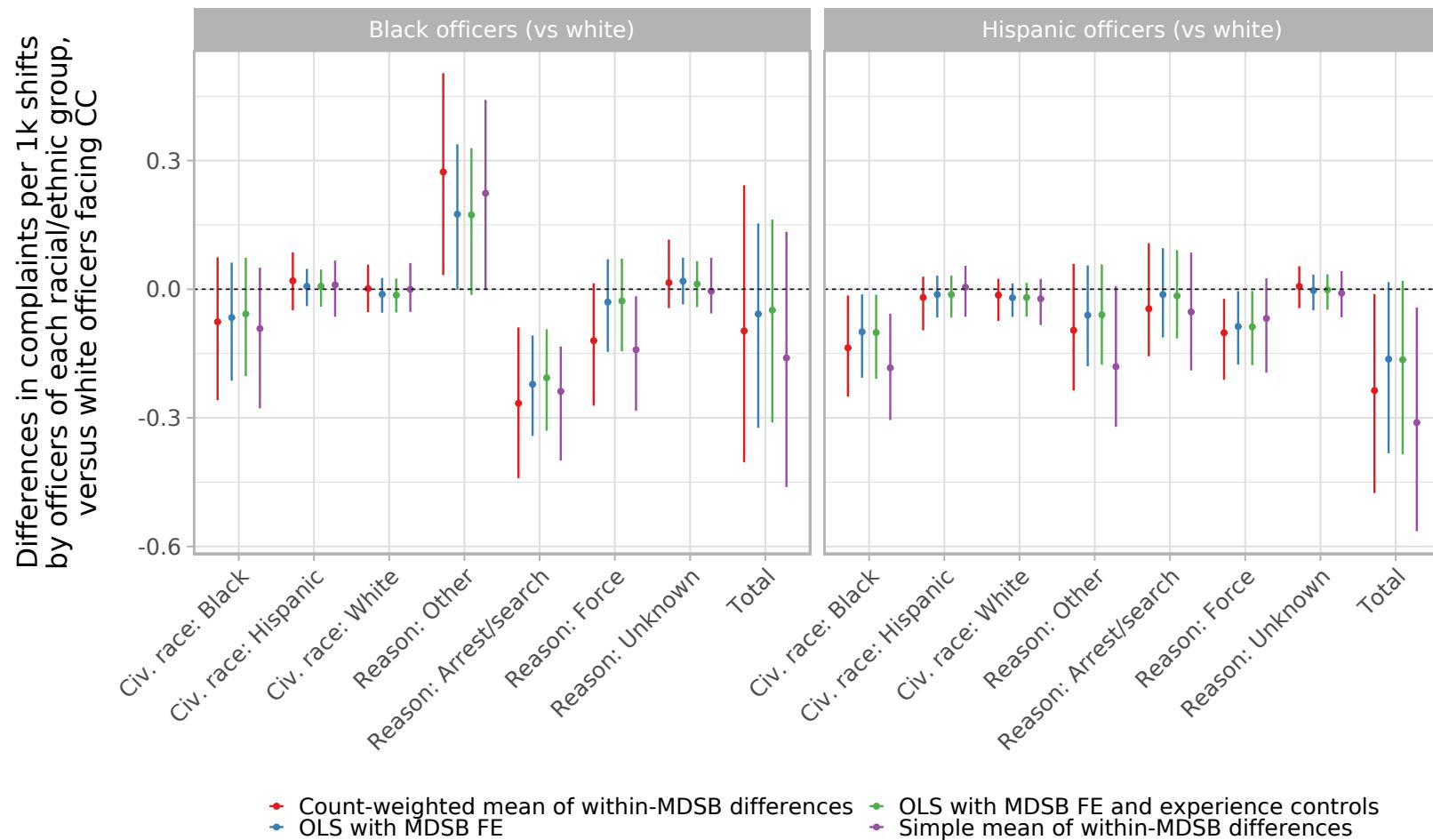
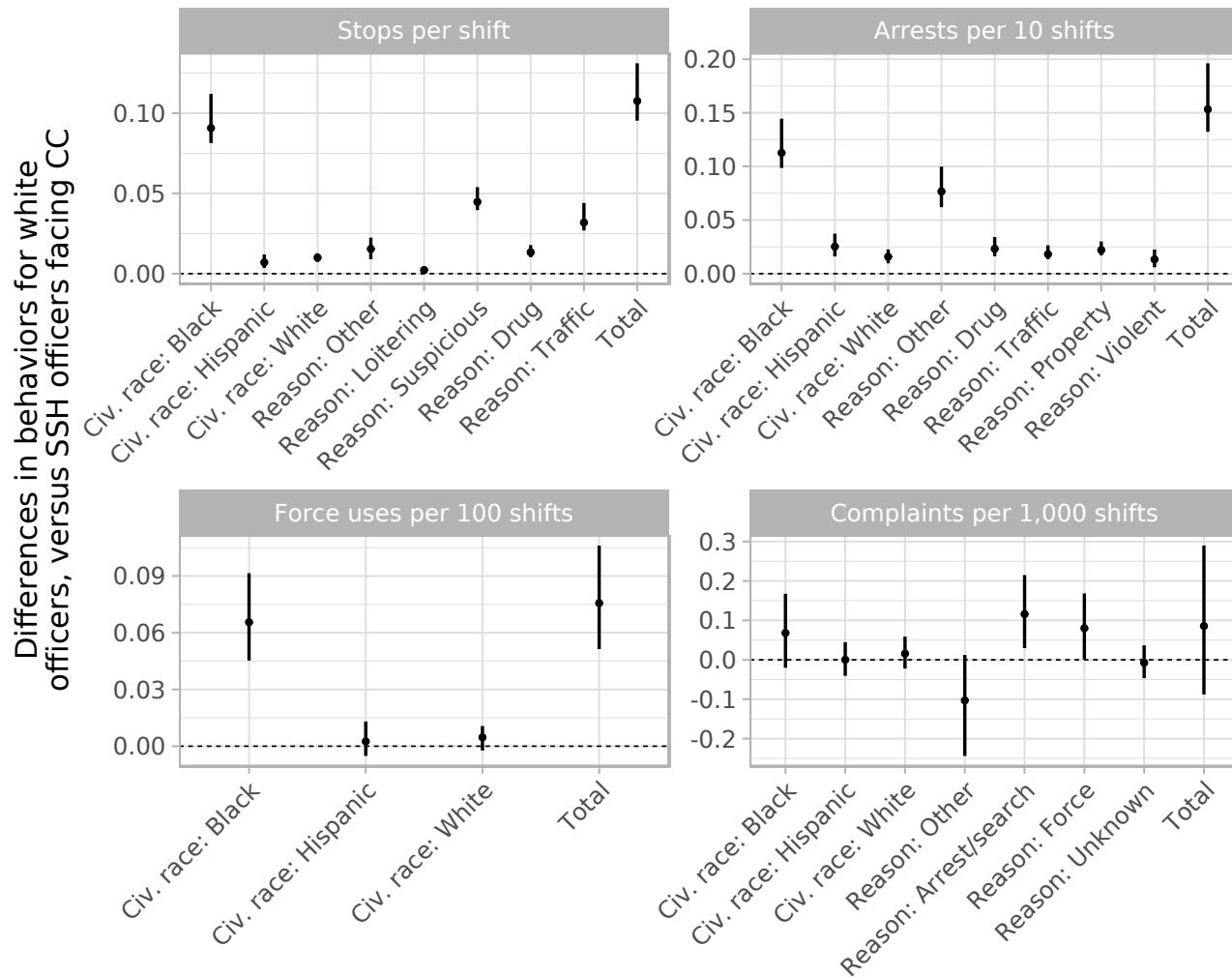


Figure A9: **Complaints.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian complaints registered against black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat. Results presented using four different estimators.



C.8 Officers' Spanish Language Proficiency

Figure A10: Comparison of Spanish-speaking Hispanic Officers to White Officers. Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian stops, arrests, uses of force, and civilian complaints generated by white officers per shift, relative to Spanish-speaking officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat.



C.9 Officer Gender

Figure A11: **Stops.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of stops of civilians by black and Hispanic male and female officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat.

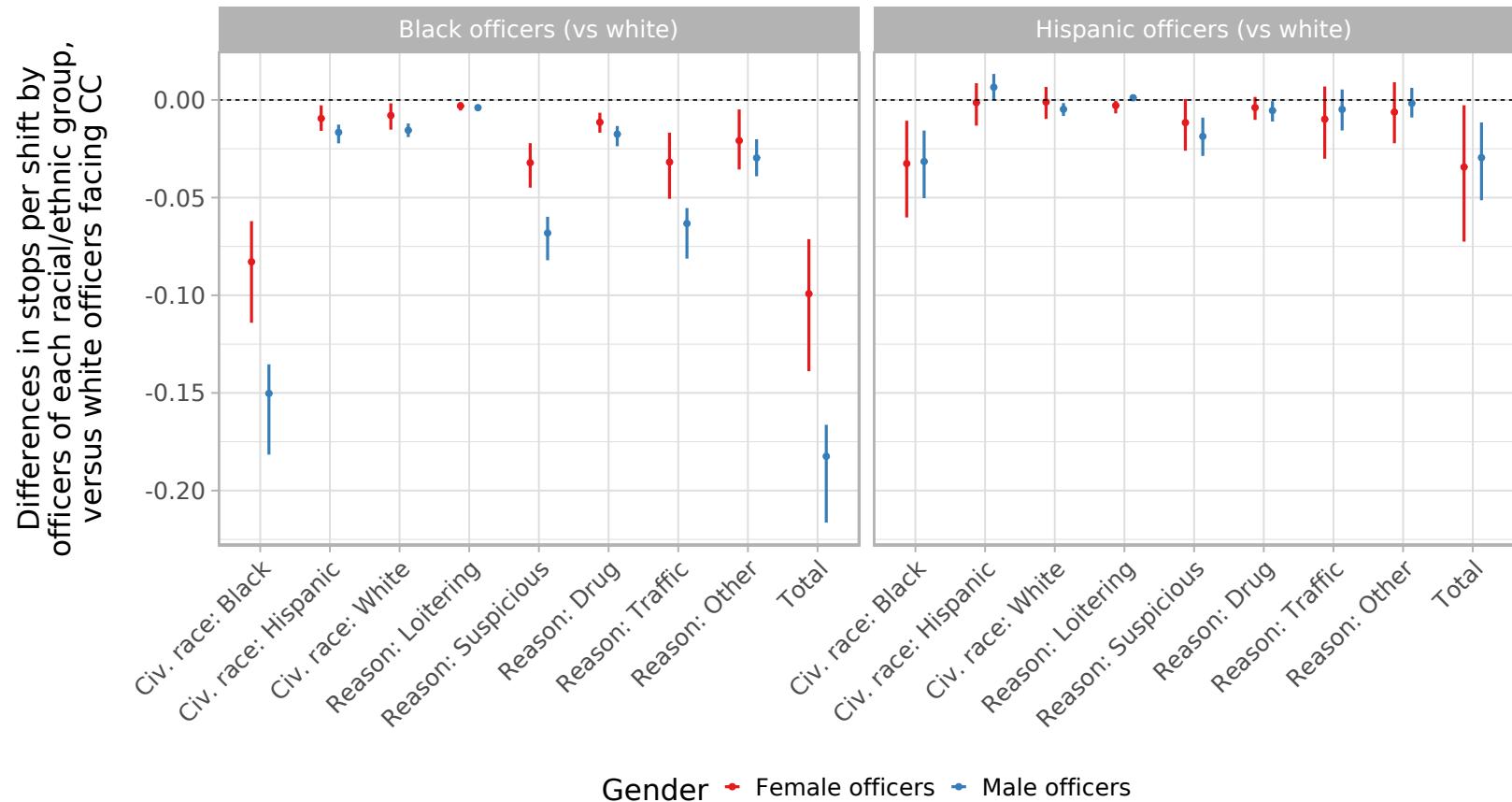


Figure A12: **Arrests.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of arrests by black and Hispanic male and female officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat.

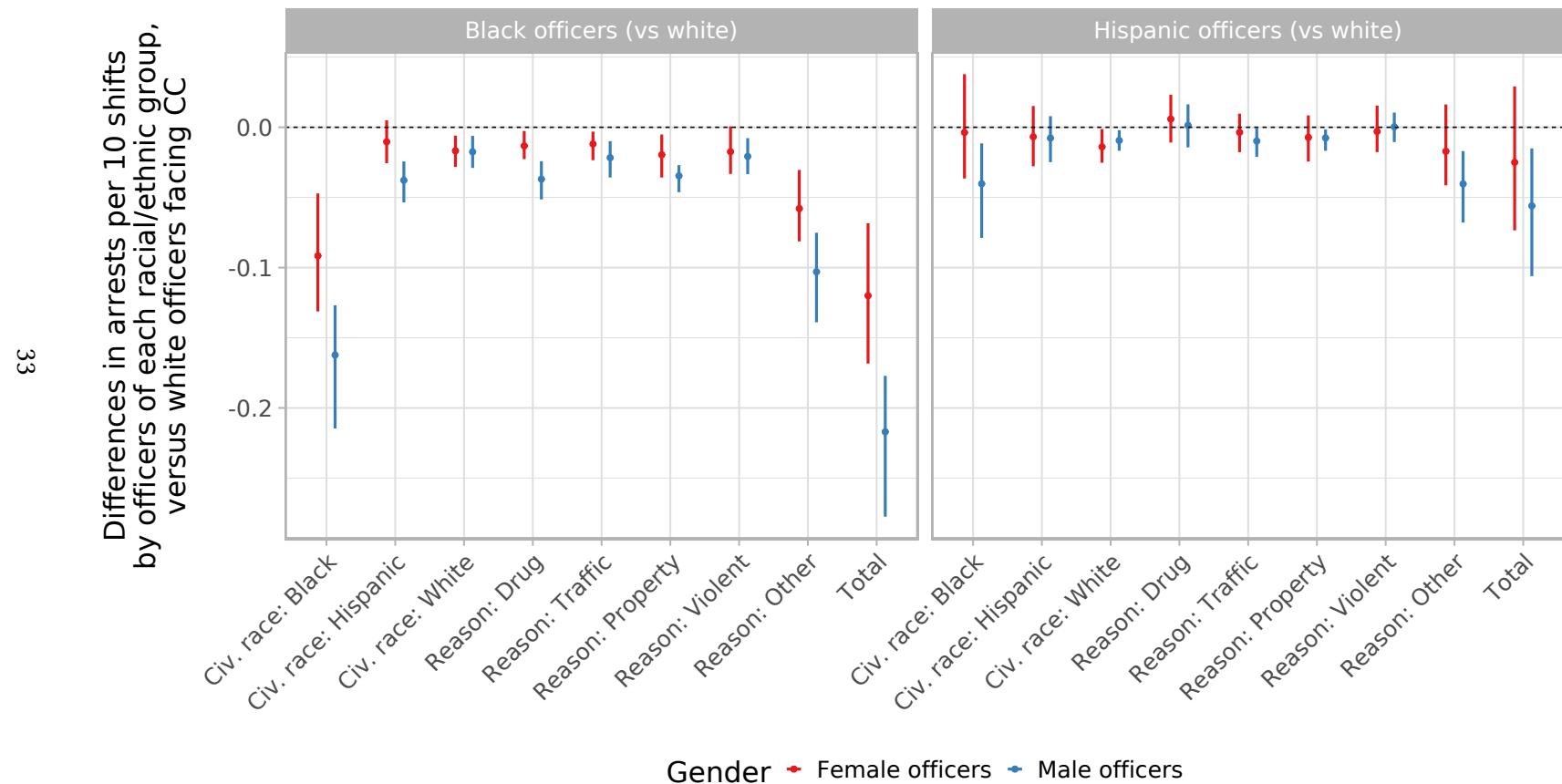


Figure A13: **Uses of force.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of uses of force by black and Hispanic male and female officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat.

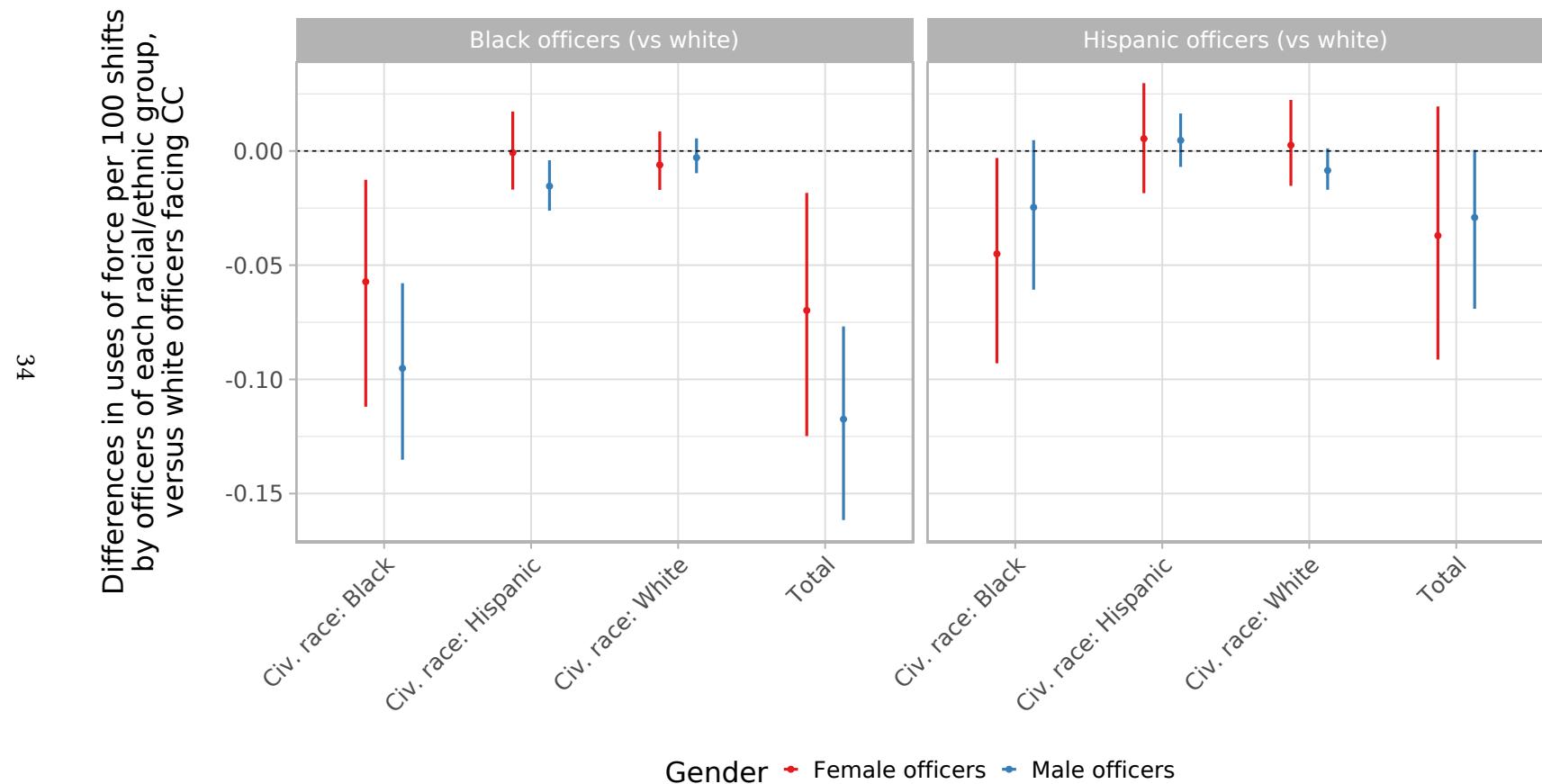
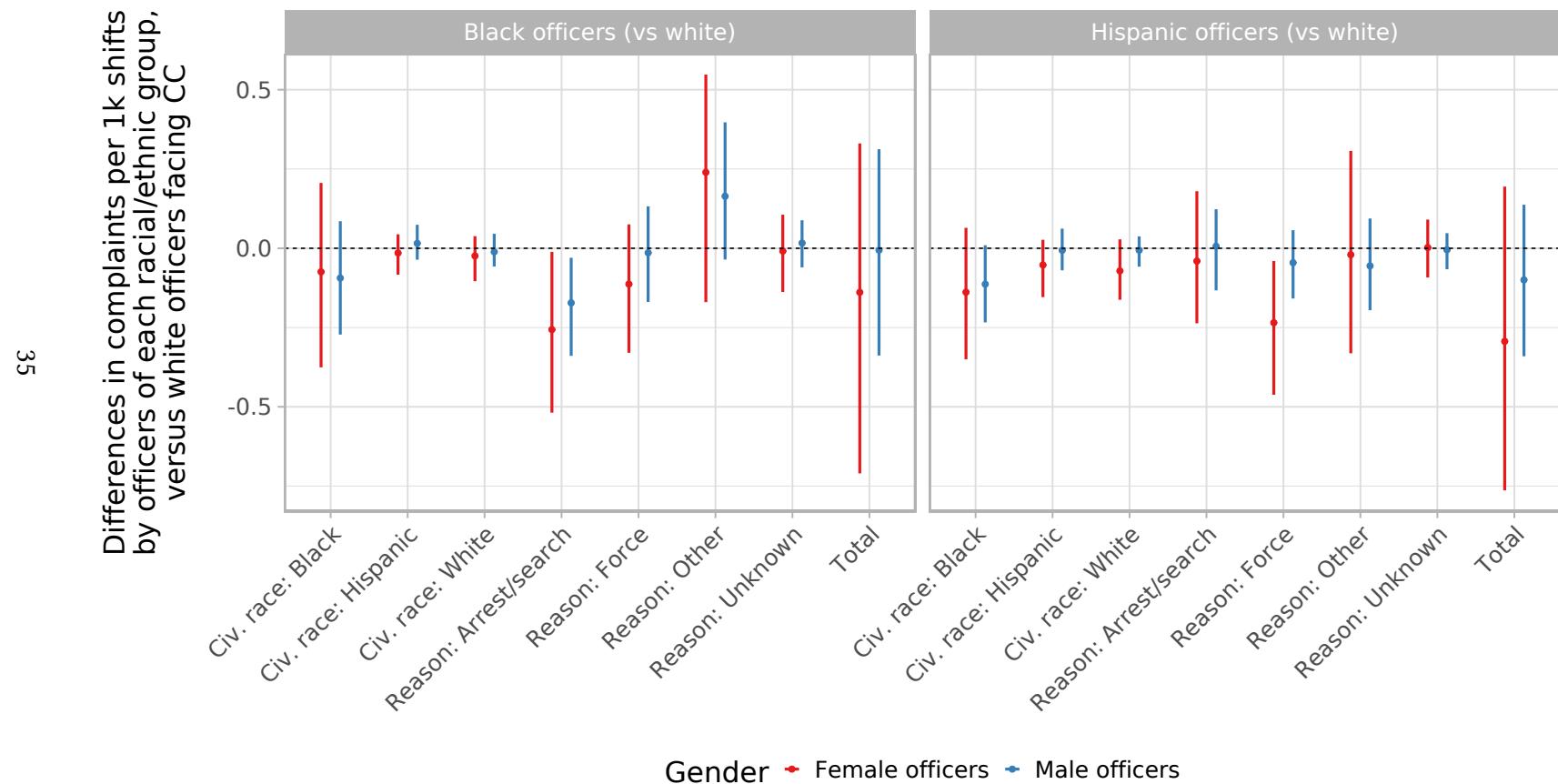


Figure A14: **Complaints.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of complaints registered against by black and Hispanic male and female officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat.



C.10 Robustness Checks: Multiple Stopping Officers

Data on stops of civilians indicate that 87% of stops report two officers as participating in the stop. In our main analysis, we treat a stop by two officers as two incidents in the data, as both officers contribute to the decision to engage a civilian. To gauge the extent to which this decision drives our results, we present an alternative analysis of stops in which we use only data on first reporting officers. As Figure A15 shows, results are essentially unchanged when imposing this restriction.

Figure A15: **Stops.** Points (error bars) depict estimated differences (95% block-bootstrap confidence intervals) in the number of civilian stops conducted by black and Hispanic officers per shift, relative to white officers patrolling in the same month, day group (weekday/weekend), shift time (first/second/third watch), and beat, using only data generated by the first reporting officer.

