The Boss in Blue: Supervisors and Police Behavior

Austin V. Smith *

August 11, 2025

Abstract

We know little about the influence of supervisors in high-stakes, discretionary settings such as policing. Leveraging frequent rotations of officers between supervisors—sergeants—I estimate causal effects of individual sergeants on arrests. Moving an officer from a 10th to 90th percentile sergeant increases monthly arrests by 42%. Sergeant effects on serious and low-level arrests are weakly correlated and reflect distinct officer behaviors. Sergeants who drive low-level arrests do so mainly via discretionary drug enforcement, disproportionately affecting Black civilians and increasing use of force. These findings position supervisors as critical actors in shaping police behavior and offer new insights for reform.

JEL Codes: M54, J45, K42

*Northeastern University. email: austinvsmith.econ@gmail.com. I am deeply grateful to my committee members, Dan Herbst, Evan Taylor, Ashley Langer, and Juan Pantano, for their invaluable guidance, support, and mentorship. I would also like to thank Jesse Bruhn, Mallory Cash, Spencer Cooper, Price Fishback, Lance Gui, Ami Ichikawa, Andrew Johnston, Cynthia Lum, Daniel Nagin, Nayoung Rim, Matthew Ross, CarlyWill Sloan, Wint Thu, Mo Xiao, and seminar participants at the University of Arizona, SEA Annual Meeting, and ViCE for their insightful feedback. This research benefited greatly from conversations with Major Stephen Bishopp of the Dallas Police Department as well as numerous members of the DPD Public Information Office, whose expertise was essential to understanding many crucial institutional details. This project was supported by Award No. 15PNIJ-24-GG-01582-RESS, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect those of the Department of Justice. IRB Approval for this project was received from the University of Arizona (STUDY0005458). All remaining errors are my own.

1 Introduction

Although it is widely believed that supervisors play a key role in shaping worker behavior, especially in high-stakes environments, much of the empirical literature focuses on routine tasks in private-sector settings where productivity is easily quantified. As a result, we know comparatively little about how supervisors affect complex, discretionary decisions—those that lack clear benchmarks and often carry real human consequences. This gap is particularly salient in policing, where front-line officers routinely exercise discretion in encounters such as stops, arrests, and use of force. Sergeants, as the first-line supervisors in police departments, are charged with mentoring, directing, and evaluating officers. They are also increasingly viewed as levers for reform, with many departments emphasizing their importance in implementing new norms and practices (Santos, 2019). Yet despite their perceived importance, credible evidence on their actual influence remains limited. This paper addresses that gap by asking: Do front-line police supervisors—sergeants—shape officer behavior on the ground?

We have many reasons to expect that supervisors influence frontline outcomes. Descriptive evidence—from both private and public sectors—shows large variation in performance across managers, even within similar organizational settings. Studies have documented manager fixed effects in productivity, worker satisfaction, employee turnover, and innovation (e.g. Bertrand and Schoar, 2003; Bloom et al., 2012; Lazear et al., 2015). This variation is suggestive but not dispositive: causal evidence, while growing, remains comparatively limited. Even in sectors where outputs are more easily measured, identifying supervisor effects is challenging due to endogenous matching, limited turnover, and data constraints. These difficulties are magnified in policing, where discretion is high and outcomes are harder to define. Understanding whether—and how—sergeants influence frontline behavior is critical for ongoing reform efforts, particularly because supervisors represent a relatively small tier of the organization whose leverage may offer a cost-effective path to change. This paper begins to fill that gap by providing new causal evidence on the role of first-line supervisors in shaping officer behavior.

Using administrative data from the Dallas Police Department (DPD), I construct a panel linking the monthly enforcement outcomes of individual officers to their managing sergeants over a five-year period. I exploit frequent sergeant reassignments to estimate individual (Bayes-shrunken) sergeant effects on arrests, using a switching design within a two-way fixed effects framework (Abowd et al., 1999). Causal interpretation of the sergeant effects relies on the assumption that officer rotation across sergeants is uncorrelated with underlying trends or officer-sergeant match effects. In the DPD, reassignments occur due to schedule realignments

initiated by the police chief and vacancies—neither of which officers can directly influence. Consistent with this mechanism, I find no evidence of trend- or match-based sorting. The specifications include location and shift controls, and I provide event-study evidence showing that officer behavior changes sharply and persistently following a sergeant switch, further reinforcing the causal nature of the effects.

I first show that sergeants substantially influence the total number of arrests made by their officers. Moving an officer from a sergeant in the 10th percentile of the sergeant effects distribution to one in the 90th percentile increases arrests by 1.6 per month (42% relative to the mean). Using bias-corrected estimates of the variance components, I find that sergeant effects account for 3.4% of the variation in arrests across officer-sergeant spells——roughly half the magnitude of managerial effects in a rote, assembly-line setting (Adhvaryu et al., 2024). However, I show that sergeant variation has a greater impact than officer variation on the total number of arrests within the department, since sergeants supervise an average of 6.33 officers, which amplifies their influence.[1]

Having shown that sergeants are important determinants of officer behavior, I next ask whether "effective" sergeants share common enforcement objectives by estimating their effects separately on arrests for serious (violent/property) and low-level (victimless) crimes. In principle, sergeants may treat arrests as a measurable form of productivity (Bratton and Murad, 2018), motivating officers to increase enforcement effort across all crime types. However, arrests are not always socially productive, and there is ongoing national debate among lawmakers, police officials, and the public over which types of crime are worth the costs to enforce (Lehman, 2024). Thus, sergeants may differ in their views on which forms of enforcement are most valuable, which could generate trade-offs between enforcing low-level and serious offenses.

Estimating the correlation between sergeants' effects on serious and low-level arrests, I find substantial heterogeneity in sergeants' enforcement effects across crime types. The relationship between serious and low-level effects is nearly flat (correlation = 0.11), and many sergeants lie at opposite extremes of the serious and low-level effects distributions. These findings suggest that individual police supervisors shape how law enforcement policy is implemented on the ground. Using sergeant rotations across geographic and temporal assignments, I find no evidence that either type of sergeant effect is associated with crime reductions—implying that these managerial styles reflect individual preferences rather than differences in public safety effectiveness.

[1]I estimate the effect of replacing sergeants at each percentile of the sergeant effects distribution with a median sergeant for one month. Assuming each supervises the average number of officers, I find that 90% of these replacements would induce a larger change in arrests (in magnitude) than replacing an officer at the same percentile with a median officer.

To characterize how managerial differences more broadly shape police interactions with the public, I estimate the effects of serious and low-level sergeant styles on a wide range of officers' on-the-job actions. I find that low-level sergeant effects operate predominantly through drug enforcement. A one standard deviation increase in low-level effects leads to a 54% rise in drug arrests relative to the mean. Ninety percent of these additional arrests are for simple possession, and they disproportionately impact Black civilians. Consistent with the discretionary nature of these arrests, I find that officer-initiated interactions account for more than half of the sergeant-induced increase in low-level arrests. Moreover, low-level sergeant effects significantly raise the incidence of violent encounters with civilians: a one standard deviation increase leads to a 15% rise in officer use-of-force incidents relative to the mean.

By contrast, serious arrests appear to be induced through greater 911 response effort. A one standard deviation increase in serious sergeant effects leads officers to respond to 3.6% more 911 calls per month. This increased emphasis on call activity results in more arrests for offenses that directly harm others—specifically domestic violence, theft, and driving while intoxicated (DWI). Officers supervised by these sergeants also have more civilian interactions and are therefore more likely to use force, but the impact of serious sergeant effects on use of force is only one-fourth as large as that of low-level sergeant effects.

Finally, I take advantage of the fact that my setting allows for direct observation of *sergeant* behaviors to study the mechanisms through which they influence officer decisions. Sergeants can choose to actively patrol the streets or remain in the station and advise officers over the radio. I show that sergeants who incentivize low-level arrests are more likely to respond to 911 calls themselves and to make low-level arrests independently of their subordinates, suggesting they lead by example by modeling their desired enforcement behaviors in the field. Moreover, low-level sergeant effects are also associated with more direct monitoring of subordinates, as these sergeants are more likely to be recorded at their officers' calls. I find no evidence that serious sergeant effects are associated with either mechanism. However, subordinates of sergeants with large serious effects are more likely to take calls and make arrests outside their regular shift hours, suggesting that these sergeants are more likely to approve overtime to enable additional 911 responses. This analysis shows that sergeants use their institutional roles and responsibilities to shape employee behavior despite limited access to high-powered performance incentives and the complex nature of the decisions.

Beyond the economics of public safety, this paper contributes to the broader labor economics literature on the impacts of supervisors on workplace behavior. While this topic has received substantial attention, much of the prior work documenting supervisor or manager

effects is descriptive and cannot isolate causal effects from confounding factors such as endogenous employee matching (e.g. Ichniowski et al., 1997; Bertrand and Schoar, 2003; Bloom and Van Reenen, 2007; Bloom et al., 2012, 2015; Lazear et al., 2015). Recent advances—particularly field experiments and the use of large administrative datasets—have enabled studies showing that supervisors can have meaningful causal impacts on employees in both public (Fenizia, 2022) and private sector (e.g. Bloom et al., 2013; Giorcelli, 2019; Hoffman and Tadelis, 2021; Adhvaryu et al., 2023) settings. However, these studies typically focus on routinized job tasks and quantifiable productivity measures. We know less about whether, and how, supervisors shape complex, discretionary aspects of worker behavior. This knowledge gap is particularly critical for public service professions—such as law enforcement, teaching, and foreign service—where discretion is high, outputs are difficult to observe, and wage structures are rigid (Wilson, 1991). By estimating the causal impacts of police sergeants, I provide novel evidence on the role of supervisors in a high-stakes setting where workers exercise substantial discretion.

More directly, this paper also contributes to the growing literature on organizational drivers of police behavior. Prior studies have shown that police enforcement is responsive to collective bargaining (Mas, 2006), fiscal pressure (Makowsky and Stratmann, 2009), public access to complaint records (Rivera and Ba, 2022), field training (Adger et al., 2022), and police academy peers (Rivera, 2025). By focusing on the role of sergeants in generating enforcement outcomes, this paper highlights first-line supervision as a critical source of incentives within police departments.[2] My findings build on recent work examining the impacts of police managers in other specific contexts. Several papers analyze police executives—chiefs and commanders, who shape departmental policy at a broad level (Mummolo, 2018; Bacher-Hicks and De La Campa, 2020a; Kapustin et al., 2022). In contrast, my study focuses on first-line supervisors who affect day-to-day enforcement effort. Rim et al. (2024) show that white sergeants are less likely to nominate Black officers for awards conditional on performance. My study differs in that it examines how sergeants affect officer behavior, rather than subjective evaluations. Gudgeon et al. (2023) use the quasi-random rotation of lieutenants (supervisors one rank above sergeant) between pre-assigned off-duty days to study a related question: how do minority supervisors affect arrest decisions? While their focus is on supervisor race as a specific causal pathway, I estimate the full extent of heterogeneity in supervisor effects, even within racial groups. A unique conclusion of my approach

[2]Adger et al. (2022) study the impacts of field training officers, a specific type of supervisor in charge of mentoring new officers. While their findings provide the basis for a powerful training intervention that can reduce police aggression over time by changing how the *next generation* of officers are trained, my findings suggest that interventions targeted at sergeants can have powerful immediate effects through the *current generation* of officers who have already undergone (potentially aggressive) field training.

is that changing the enforcement proclivities of *existing* supervisors can substantially affect officers, which may be more feasible than changing the racial composition of supervisory staff. Additionally, a distinctive feature of my empirical design is that it quantifies the importance of supervisors *relative to officers*, helping to adjudicate between individual- and supervisory-level explanations of police behavior.[3]

Finally, my findings contribute to the public policy debate on police reform. While Americans broadly agree on the need for changes in law enforcement (CBS News, 2023), there is little consensus about where reform efforts should be focused. Many commonly proposed reforms target front-line officers, advocating for policies such as increasing minority recruitment (Ba et al., 2021b) or improving officer training (Dube et al., 2025). My results suggest that interventions aimed at first-line supervisors may be especially effective and durable in changing officer behavior.[4] Moreover, the weak correlation I estimate between serious and low-level sergeant effects suggests that policies targeting supervisors may be able to reduce low-level arrests without undermining enforcement of serious crimes. Because low-level arrests impose significant costs on arrestees and their communities (e.g. Bacher-Hicks and De La Campa, 2020b; Agan et al., 2023), and I find no evidence that sergeant-induced enforcement changes reduce crime, such policies may yield meaningful social benefits.

The rest of the paper proceeds as follows. Section 2 describes the job functions of police sergeants and the assignment process within the DPD. Section 3 details the administrative data and sample construction. Section 4 presents the empirical strategy and discusses identification concerns. Section 5 reports the main results and diagnostic checks. Section 6 analyzes mechanisms. Section 7 concludes.

2 The Role of Police Sergeants

Sergeants represent the first level of formal management within police departments. In each branch of the organization, officers are grouped into units led by a sergeant, who serves as their immediate supervisor. In the patrol division—responsible for general crime control

[3]My findings also contribute to a long-standing debate in criminology over whether sergeants can meaningfully influence officer behavior (Van Maanen, 1984; Brown, 1988). A large body of fieldwork has documented correlations between sergeant conduct and officer decisions (Engel, 2000, 2001, 2002; Engel and Worden, 2003; Johnson, 2011, 2015a,b; Ingram et al., 2014). Instead of examining the association between specific features of a sergeant and the outcomes of their officers, I leverage a quasi-experimental research design to answer a different empirical question with direct policy implications: how would an officer's arrests change if they were assigned a different sergeant?

[4]Durability is a concern for several officer training programs. Studies of diversity training (Mello et al., 2023) and procedural justice training (Owens et al., 2018) find initial effectiveness, but note that the effects often diminish over time. My event study results suggest that sergeants provide a stable set of incentives that generate persistent changes in officer behavior throughout the duration of their supervisory relationship.

and 911 response—units are assigned based on geographic sector and shift. I focus on patrol sergeants in the Dallas Police Department (DPD), which assigns one or more sergeants to each of the city's 35 patrol sectors, with separate coverage across three daily shifts, or watches.[5]

Sergeants are responsible for overseeing the day-to-day conduct of the officers under their command. Formally, their role is to ensure compliance with departmental policy and to monitor performance. In practice, however, sergeants exercise substantial discretion in how they interpret and carry out this responsibility. This discretion leads to variation in supervisory style—a pattern that has been documented in qualitative fieldwork (e.g. Engel, 2001).

For instance, a former DPD sergeant described his role primarily as providing field support rather than directing specific enforcement activity. Others, he noted, are more prescriptive—explicitly encouraging officers to prioritize certain types of offenses. This variation in leadership philosophy is consistent with long-standing ethnographic evidence. One officer, quoted in Van Maanen (1984), captured this heterogeneity as follows:

> "Now you take Sergeant Johnson. He was a drunk-hunter. That guy wanted all the drunks off the street, and you knew that if you brought in a couple of drunks a week, you and he would get along just fine. Sergeant Moss, now, is a different cat... What he wants are those vice pinches. Sergeant Gorden wanted tickets, and he'd hound [you] for a ticket a night. So you see, it all depends on who you're working for. Each guy's a little different."

Beyond crime-specific enforcement priorities, sergeants also differ in how actively they manage their officers' general productivity. Some value proactive field presence and responsiveness to calls; others are more hands-off, focusing on administrative tasks and intervening only when necessary.

While sergeants are not typically present during their officers' interactions with civilians, they have access to a range of administrative and informal tools to shape officer behavior. Formally, sergeants write performance evaluations, approve overtime and scheduling requests, and provide recommendations for internal transfers to desirable positions (e.g., investigative or tactical units). They are also expected to review use-of-force reports and monitor patterns in their officers' searches, citations, and arrests. These oversight functions give sergeants leverage to reward or discipline subordinates through both formal commendations and documented infractions (Rim et al., 2024). In addition, hierarchical norms within police departments reinforce a culture of compliance with supervisory authority (King, 2005).

[5] In especially large or high-crime sectors, multiple sergeants may be assigned. Each manages a separate unit of officers.

Beyond administrative oversight, sergeants can influence officer behavior through direct modeling, or leading by example. Those who choose to actively patrol alongside their officers—responding to calls and making arrests themselves—may shape expectations about desirable enforcement activity. Officers often interpret these behaviors as implicit guidance, particularly in ambiguous situations such as low-level offenses. For example, a sergeant who emphasizes drug enforcement may reinforce that priority by making drug arrests personally. Ethnographic evidence suggests that officers place more trust in such "street sergeants," believing they better understand the demands of patrol work (Van Maanen, 1984). Survey data support this mechanism: officers are more likely to believe a specific task will be used in their evaluations if they see their sergeant engage in it (Engel and Worden, 2003).

Field activity also allows sergeants to overcome limits to their monitoring capacity. They can assign themselves to calls their officers are handling and provide guidance in real time. While officers are expected to seek supervisory input when situations are uncertain, they may be more inclined to do so when they know their sergeant is willing to appear on the scene. In this way, patrol engagement serves both as a monitoring tool and a managerial signal—strengthening supervisory influence even in the absence of high-powered incentives.

In the DPD, patrol officers change sergeants frequently and have limited control over the timing or nature of these changes. A sergeant switch occurs either when an officer is reassigned to a new unit or when their current unit receives a new supervisor. These reassignments occur for one of two reasons. First, they may be driven by vacancies arising from promotions, retirements, deaths, or transfers into specialized units. Unlike many other large police departments, Dallas does not allow officers to bid for or select into vacant positions.[6] Instead, executive command staff fill officer vacancies at their discretion within each station and watch. When sergeant vacancies arise, other sergeants may express interest and interview for the position, but the final decision is again made by executive staff.

Second, officers may receive a new sergeant through department-wide schedule realignments. Once per year, DPD leadership reassesses staffing needs across patrol stations, watches, and days of the week. If major changes are needed, the Chief of Police can initiate a Patrol Bid—a process in which a designated group of officers and/or sergeants selects their preferred station, watch, and days off in descending order of tenure. However, the bid is not held on a regular schedule (occurring in only 3 of the 5 years in my sample), and eligibility is announced just two weeks in advance. As a result, officers have limited ability to sort based on trends — a key assumption for my identification strategy, discussed in Section 4. Crucially, even during the Patrol Bid, officers are not allowed to choose their sector or

[6]See Ba et al. (2021a) for a discussion of vacancy bidding in the Chicago Police Department and its implications for officer sorting between high- and low-crime districts.

sergeant. These assignments remain at the discretion of command staff and, based on my interviews with DPD personnel, appear not to follow a consistent rule.

3 Data

This project uses several administrative datasets obtained through FOIA requests from the Dallas Police Department (DPD) and the Dallas County District Attorney's Office, covering the period from June 2014 to July 2019. I combine data on police incidents, personnel records, officer activity, and court outcomes to construct a monthly panel that links officer enforcement activity to their assigned sergeants. I focus on sergeant assignments for patrol officers, whose primary responsibilities include responding to civilian-initiated 911 calls, patrolling assigned beats, and addressing crimes observed *on-view*. Patrol sergeants are assigned to a unit of officers within a particular patrol sector and watch.

DPD maintains assignment data only at the patrol station level—a coarser geographic unit—meaning it does not keep direct records of sergeant assignments. However, the Computer Aided Dispatch (CAD) system, which logs officer responses to incidents, records the sector and watch assignments of all responding personnel (including sergeants) on a daily basis (see Supplementary Appendix C for further details). I use these data, along with station assignments and promotion histories, to construct monthly sergeant assignments for patrol officers from June 2014 to July 2019. Specifically, I assign each officer to the sector-watch in which they appear on the most days within a month, and assign each sector-watch the sergeant most frequently observed with that assignment. This construction yields a panel of 2,067 officers, 388 sergeants, 15,355 officer-sergeant spells, and 61,166 officer-month observations.

I am interested in the effects of an officer's regularly assigned sergeant, who evaluates officer performance and works with the officer on most of their workdays. In practice, officers may not be assigned to the same sergeant every day. During a sergeant's off days, their duties are filled by a rotational substitute. Officers may also be temporarily reallocated to a different sector-watch based on manpower needs. In both cases, the regularly assigned sergeant still retains administrative responsibility for the officer. To the extent that officers receive advice and instruction from multiple sergeants within a month, my assignment method captures the influence of the sergeant to whom they are most frequently exposed.

To ensure that my estimates reflect the influence of regularly assigned sergeants, I impose two sample restrictions. First, I require that officer-sergeant spells last at least two consecutive sample months. This reduces the risk of assignment errors stemming from officers working temporary assignments with more activity than their permanent one, in which

case arrests might be misattributed to the wrong sergeant. If a sergeant assignment was incorrectly recorded despite no actual change, these errors would attenuate the variance in sergeant effects, since any observed change in officer behavior would reflect noise. This restriction eliminates 5,747 spells, 19% of which are single-month spells with no assigned sergeant. Second, I exclude the remaining 866 spells in which a sergeant cannot be identified. In Supplementary Appendix Figure A.1, I show that these sample restrictions do not meaningfully affect my estimates.

To facilitate identification of sergeant and officer fixed effects, I remove any officers and sergeants who appear only together, any officer/sector-watch and sergeant/sector-watch pairs that appear only together, and any officers, sergeants, sector-watches, or day-off groups that appear only once in the data. I also require that officers appear in at least 5 separate months. These restrictions eliminate 310 officer-supervisor spells, yielding an analysis sample of 1,805 officers, 347 supervisors, 8,432 officer-supervisor spells, and 49,923 officer-month observations.

To study trends around officer moves, I construct a separate balanced event study sample. I define an event as two consecutive spells involving the same officer but different sergeants. Within this sample, I require that pre-switch spells last at least 5 months and post-switch spells at least 4 months. Because switches are measured at the monthly level, the transition occurs sometime during the final month the officer is assigned to their previous sergeant. This switching month is excluded from the count of months prior to the switch.

I supplement the panel of sergeant assignments with officer activity data from several sources. I use the universe of arrest reports to count the number of arrests made by each officer in every month of the sample. Each arrest is matched to its listed charges at the time of apprehension and categorized as either serious or low-level, using the definitions in Rivera (2025). Serious arrests include index crimes (murder, rape, robbery, aggravated assault, theft, burglary, and arson), which are tracked by the FBI due to their high social costs. They also include several non-index but serious offenses: simple assault, any form of domestic violence, sexual assault, fraud, and driving while intoxicated (DWI).[7] All other arrests are classified as low-level. These primarily consist of outstanding warrants,[8] disorderly conduct, and drug possession, which together account for 81% of low-level arrests. Low-level arrests also include public order offenses with no clear victim, such as vagrancy, liquor violations, and prostitution. Because arrests may contain multiple charges, I classify each arrest based

[7]The only difference between my classification and that of Rivera (2025) is the inclusion of DWI as a serious crime.

[8]While I cannot determine the offense associated with each warrant, national data suggest that most are issued for non-violent crimes and ordinance violations, such as unpaid traffic tickets (Slocum et al., 2021).

on the most severe charge.[9]

I link each arrest to court outcomes using records obtained from the Dallas County District Attorney's Office and classify its conviction status.[10] A conviction occurs if the arrest is matched to a court case that does not result in a dismissal. Convictions thus include plea bargains and those rendered by a judge or jury. If a charge does not match to court data, I consider it dismissed. Conviction is defined at the arrest level, such that an arrest results in a conviction if the arrestee was convicted on any of the related charges.

I extract 911 calls from CAD data to separately evaluate civilian-initiated and proactive police encounters.[11] An arrest is considered officer-initiated if it does not originate from a 911 call. Additionally, I merge use of force reports and civilian complaints to the involved officers and the month of occurrence. I link officers and sergeants to internal personnel records containing demographic information, tenure and promotion history, shift, day-off group, and bureau assignments. I also link each sergeant promoted in 2012 or later (accounting for 58% of sergeants) to their score on the promotional civil service exam. I analyze the association between exam scores and sergeant effects in Supplementary Appendix E.

3.1 Summary Statistics

Summary statistics for the full, unrestricted data, the analysis sample, and the balanced event study sample are presented in Table 1. The analysis sample closely resembles the unrestricted data, suggesting that estimates of sergeant effects are unlikely to be biased by sample selection. Officers in the event study sample exhibit slightly lower arrest activity compared to those in the unrestricted data and analysis sample. One likely explanation is that the event study sample requires officers to have successive, stable patrol assignments, which excludes officers who prefer making arrests and are more likely to transfer to specialized teams, such as gang or narcotics enforcement, where arrest counts are typically higher. Since all of my analyses include officer fixed effects, these sample differences should not materially affect my findings.

[9] To ensure my results are not driven by this classification decision, I also consider a more traditional partition of arrests into index and non-index categories. Using this alternative classification does not meaningfully change my findings (see Supplementary Appendix Figure A.11).

[10] Specifically, I use the name of the arrestee and the offense date to match an arrest to a case within the universe of cases disposed in Dallas County from 2014 to 2020. I first match arrests to all court cases with the same offense date, then use Jaro-Winkler distance to calculate the similarity of the first and last names of the matched defendants. If an arrest has a matching case with both names perfectly matching (i.e., Jaro-Winkler score equal to 1), I keep only that case. For all other arrests, I keep a match if the Jaro-Winkler score is 0.9 or higher. This matching technique, similar to that used by Adger et al. (2022), allows for minor spelling errors in arrest reports while remaining conservative in name similarity requirements.

[11] I use the cleaning procedure described by Online Appendix A4 in Weisburst (2024) to isolate 911 calls in CAD.

Table 1 also shows that officers are highly mobile, and sergeants supervise a large number of officers within the sample. The average officer works with just under four unique sergeants, and the average sergeant manages over 20 officers. This density in the supervisory network is essential for my empirical strategy, as sergeant fixed effects can only be identified within groups of officers and sergeants connected by movement across assignments (Abowd et al., 2002). In my data, all observations fall within one connected set.

On average, patrol officers in the sample make 3.8 arrests per month, three-fourths of which are for low-level crimes. This proportion is comparable to the national share of misdemeanor arrests, which make up roughly 80% of all arrests according to estimates by Natapoff (2016).[12] There is substantial variation in arrests across officers: the standard deviation is 3.64, nearly as large as the mean. Supplementary Appendix Figure A.3a plots the distribution of average arrests per officer-month across sergeants, revealing notable heterogeneity. Officers working under sergeants in the right tail of the distribution average over six arrests per month, while those in the left tail average one arrest or fewer.

However, average arrest counts by sergeant do not identify the causal effect of sergeants on arrests, since these averages confound officer discretion with sergeant influence. Indeed, officer discretion contributes to even greater variation in arrest behavior across officers (see Supplementary Appendix Figure A.3b). Disentangling officer effects from sergeant effects requires observing how officer behavior changes when they switch supervisors. This is the central idea behind the empirical strategy, described in the next section.

4 Empirical Strategy

4.1 Estimating Sergeant Effects

I estimate the effect of sergeants on their officers' arrest behavior. I follow the two-way fixed effects approach pioneered by Abowd et al. (1999), which has been widely applied to estimate manager effects across a variety of settings (Benson et al., 2019; Frederiksen et al., 2020; Fenizia, 2022; Metcalfe et al., 2023). The model is specified as follows:

$$Arrests_{it} = \theta_i + \psi_{J(i,t)} + x'_{it}\beta + \nu_{it}, \tag{1}$$

where $Arrests_{it}$ is the number of arrests made by officer i in year-month t, θ_i is an officer fixed effect, and $\psi_{J(i,t)}$ is a fixed effect for officer i's sergeant in month t. The time-varying

[12] Low-level crimes, as classified here, are not all misdemeanors, and not all misdemeanors are low-level crimes. For example, possessing personal-use amounts of marijuana is a misdemeanor, while possessing personal-use amounts of cocaine is a felony. However, I classify both as low-level crimes.

control vector x_{it} includes sector-watch fixed effects to account for spatial and temporal variation in crime. Because sergeant and sector-watch assignments overlap, identifying both sets of fixed effects separately requires that each sector-watch in the sample is managed by multiple sergeants. As described in Section 3, the sample construction ensures this condition is met: each sector-watch is managed by an average of 6.95 sergeants. I also include fixed effects for the officer's day-off group to control for schedule changes that coincide with sergeant switches. Finally, I control for a second-degree polynomial in officer tenure to account for changes in arrest behavior over the career cycle, which may be correlated with sergeant assignment due to seniority-based preferences during schedule realignments.

Sergeant fixed effects are identified through officer mobility across sergeants. In particular, a sergeant's estimated effect reflects changes in the arrest behavior of officers who switch into or out of their supervision. By including x_{it}, sergeant effects are estimated using within-officer variation relative to officers working in the same patrol location, shift, day-off group, and with similar tenure. For ψ_j to recover the causal effect of sergeant j, officer mobility across sergeants must be as-good-as random, conditional on officer fixed effects and the included controls. That is, sergeant assignments must be uncorrelated with unobserved, time-varying determinants of officer behavior. Importantly, the model permits non-random sorting between officers and sergeants based on time-invariant characteristics. For example, if officers who prefer making arrests tend to work with sergeants who encourage them, this would not bias the estimates. Following Card et al. (2013), I examine three types of endogenous mobility that could threaten this identification strategy.

First, sergeant assignments must be uncorrelated with unobserved trends in officer behavior or crime levels within an officer's sector. For instance, if sergeants who are more lenient toward low-level arrests are systematically assigned officers whose arrest propensity is increasing over time, the model may incorrectly attribute rising arrest activity to the new sergeant. Similarly, if officers are disproportionately moved to areas with increasing crime or enforcement demand, this could inflate the estimated variation in sergeant effects. Such a pattern might arise, for example, if high-arrest sergeants are more likely to advocate for filling unit vacancies when crime is rising in their assigned sectors.

Second, I require that changes in an officer's sergeant do not coincide with unobserved shocks to the officer's enforcement behavior. In this context, one specific concern is departmental policy changes that overlap with officer moves. For example, hot-spot policing — a common strategy in which resources are concentrated in high-crime areas (Weisburd and Eck, 2004) — could confound identification. If high-arrest sergeants are particularly adept at identifying hot spots and request additional officers for deployment, then I would observe increased arrests among officers moving to these sergeants, but for reasons unrelated

to managerial influence.

Finally, identification assumes that officers are not systematically matched to sergeants based on idiosyncratic match quality. For instance, if command staff matches officers and sergeants based on a comparative advantage in making arrests — a form of positive assortative matching — then match-specific effects, η_{ij}, would be correlated with $\psi_{J(i,t)}$ but omitted from the model.

As described in Section 2, sergeant assignments in the DPD limit officers' ability to sort based on these endogenous factors. While officers and sergeants may select their preferred station and watch through the Patrol Bid, they cannot control the timing of the Bid, which positions are available during their turn, or their eventual sector assignments. Moreover, vacancies — triggered by retirements, promotions, or transfers — are filled at the discretion of command staff, limiting the ability of officers to anticipate or influence the timing or destination of a move. To the extent that contemporaneous policy changes correlate with sergeant switches, the estimated fixed effects will reflect a combination of sergeant influences and other confounding forces. However, departmental policy changes are unlikely to drive either Patrol Bids or vacancy-induced moves. New initiatives are typically implemented by units distinct from regular patrol officers, as was the case with hot-spot policing efforts in Dallas during the sample period (Jang et al., 2012).

To support the identifying assumptions, I plot event studies around officer moves in Figure 1. I split sergeants into terciles based on the average number of arrests made by the officers they supervise during the sample period, and plot officer arrest trajectories separately by the tercile of the sergeant they move to and from. Arrests are residualized by officer fixed effects and the control vector using within-supervisor variation, following Chetty et al. (2014).[13] For this figure, I use the sample of switches that are balanced 2 months prior to the move and 2 months after the move.

Figure 1 reveals several notable patterns that support the identifying assumptions. First, officer arrest behavior shifts sharply and persistently following a change in sergeant, consistent with the fixed effects specification in which the sergeant's effect activates upon reassignment and does not attenuate over time. Second, while arrest behavior fluctuates somewhat prior to a move, these trends do not appear systematically related to the direction of the switch. Third, the figure provides little evidence that match quality is driving mobility. If officers sorted based on comparative advantage with particular sergeants, we would expect asymmetric effects of moves in opposite directions. Instead, the observed effects appear

[13] In practice, this means that I estimate $\hat{\theta}_i$ and $\hat{\beta}$ by estimating equation 1. I then calculate $Arrests_{it} - \hat{\theta}_i - x'_{it}\hat{\beta}$ using these estimates. This is necessary since any sorting pattern of sergeants would lead estimates of $\hat{\theta}_i$ and $\hat{\beta}$ to be contaminated by sergeant effects if the sergeant fixed effects were not included.

roughly symmetric: for example, moves from the 3rd tercile to the 1st are approximately equal and opposite to moves from the 1st to the 3rd.[14] Moreover, moves within the same tercile do not produce systematic changes in arrest behavior, which one would expect if these officers were moving to sergeants with whom they were better (or worse) matches.

Under the identifying assumptions, the fixed effects are unbiased. However, consistent estimation requires the number of observations to tend to infinity *within each officer-sergeant pairing*. As a result, the raw fixed effects are likely to be estimated with error, even if the identifying assumptions hold. This estimation error is more severe for sergeants with few in-sample observations. To reduce this error, I adopt Empirical Bayes shrinkage procedures commonly used in the teacher value-added literature (e.g., Kane and Staiger, 2008; Chetty et al., 2014). Specifically, I bootstrap the estimation of equation 1 in order to obtain estimates of the variance in sergeant fixed effects attributable to the true signal, σ^2_ψ, and the variance attributable to sampling error, σ^2_ϵ.[15] I then multiply each raw fixed effect by the Empirical Bayes shrinkage factor, defined as the ratio of signal variance to total variance, $\frac{\hat{\sigma}^2_\psi}{\hat{\sigma}^2_\psi + \hat{\sigma}^2_\epsilon}$. As the contribution of error variance increases, the shrinkage factor pulls a sergeant's effect toward the mean of the fixed effect distribution, which is 0 by construction (see Supplementary Appendix D for further details). I apply the same procedure to officer fixed effects.

To understand how sergeants' effects relate across different levels of crime severity, I also estimate versions of equation 1 using serious and low-level arrests separately as the dependent variable. The corresponding serious (ψ^S_j) and low-level (ψ^L_j) sergeant effects are shrunk using the same Empirical Bayes procedure. I focus on the two-dimensional relationship between a sergeant's serious and low-level effects, $Cor(\psi^S_j, \psi^L_j)$, which I estimate using the correlation of the shrunken effects. A strong positive relationship indicates complementarities in serious and low-level enforcement—suggesting that some sergeants are generally more effective, but do not shift officers from one enforcement type to another. A strong negative relationship, by contrast, suggests sergeants trade off one type of enforcement for another.

4.2 Decomposing Variation in Arrests

I next estimate the share of observed variation in arrests that can be attributed to variation in sergeant effects.

[14] In the Supplementary Appendix, I provide additional evidence that supports the symmetrical nature of officer moves (see Figures A.4 and A.5).

[15] For the bootstrap, I follow the procedure outlined by Best et al. (2023). I obtain residuals $\hat{\nu}_{it}$ and randomly resample them, stratifying by sergeant-officer pair in order to preserve the match structure of the data. I then re-estimate the sergeant fixed effects. I repeat this process 1000 times and use the distribution of fixed effect estimates for each sergeant to calculate $\hat{\sigma}^2_\psi$ and $\hat{\sigma}^2_\epsilon$.

$$Var(Arrests^*_{it}) = Var(\theta_i) + Var(\psi_{J(i,t)}) + 2Cov(\theta_i, \psi_{J(i,t)}) + Var(\nu_{it}), \tag{2}$$

$$Arrests^*_{it} = y_{it} - x_{it}\hat{\beta}. \tag{3}$$

I focus on variation in pair-level average residualized arrests, since variation within an officer-sergeant pairing is uninformative for estimating sergeant effects. Arrests are residualized using the control variables, where $\hat{\beta}$ is estimated from within-sergeant and within-officer variation using the full model in equation 1 (Chetty et al., 2014).

As with estimates of the fixed effects themselves, variance component estimates may be biased—either due to sampling error or to limited mobility bias arising from insufficient officer movement across sergeants (Andrews et al., 2008). In both cases, the variance of sergeant fixed effects would be biased upward, while the covariance between officer and sergeant fixed effects would be biased downward. While Empirical Bayes shrinkage addresses sampling error in fixed effect estimation, it does not directly address limited mobility bias. To assess this concern, I note that the officer-sergeant mobility network in my data is particularly dense: over 85% of officers in the sample switch sergeants—much higher than the proportion of movers in most related across- and within-firm studies.[16] Moreover, the full sample lies within a single connected set (Abowd et al., 2002). This feature of the data mitigates concerns about limited mobility bias.

Second, I implement estimators designed to directly correct for bias arising from limited mobility. I first apply the Andrews et al. (2008) bias correction, which derives the bias term under the assumption of homoskedastic errors and is widely used in the two-way fixed effects literature (e.g., Fenizia, 2022). I also implement the more recent leave-one-out bias correction proposed by Kline et al. (2020) (hereafter KSS), which allows for unrestricted heteroskedasticity in the error structure. The KSS estimator can only be used on the leave-one-out connected set—that is, the subset of officers and sergeants who remain connected when any single officer is removed. Applying this restriction excludes only 3 sergeants and 3 officers from my sample.

[16]For example, only around 35% of workers moved across firms in the Brazilian data studied by Alvarez et al. (2018), and only around 50% moved between managers within Indian garment factories in Adhvaryu et al. (2024).

5 Results

5.1 The Effects of Sergeants on Arrests

Figure 2 plots the distribution of both raw and shrunken sergeant effects. By construction, the mean of the distribution is zero, so each effect is interpreted as the number of arrests induced per month relative to an average sergeant. As expected, the shrinkage procedure reduces variance and pulls estimates toward the mean. However, even the shrunken distribution shows substantial heterogeneity in enforcement effects. A one standard deviation increase in sergeant effects corresponds to 0.66 additional arrests per month—roughly 17% of the mean. The distribution is roughly symmetric but exhibits a heavy left tail, suggesting a meaningful share of sergeants are associated with especially low levels of enforcement. The implied impact of switching supervisors is substantial: moving an officer from a 10th percentile sergeant to one at the 90th percentile would increase arrests by 1.6 per month, or 42% relative to the mean.

To contextualize these magnitudes, I simulate the effect of replacing high-arrest sergeants with ones at the median.[17] Replacing all sergeants above the 90th percentile with median sergeants would reduce total arrests by 2,380 over the sample period—a 2.26% decline—despite requiring changes to just 35 individuals. While officer effects show greater raw variation (see Supplementary Appendix Figure A.6), sergeants supervise multiple officers, amplifying their aggregate influence. To illustrate, replacing a 90th percentile sergeant who manages the average number of officers (6.33 per month) with a median sergeant would reduce arrests by 2.8 per month. By comparison, replacing a 90th percentile officer with a median officer would reduce arrests by only 2.6 per month.

To generalize this result, I calculate the change in arrests that would result from replacing a sergeant at each percentile of the effects distribution with a median sergeant for one month, assuming they supervise the average number of officers. I conduct the same calculation for each officer in the effects distribution. Figure 3 compares the magnitude of these changes across the two groups. In 90% of cases, replacing a sergeant yields changes in arrests that are at least as large (in absolute value) as replacing an equivalently ranked officer. I interpret these findings as evidence that variation in sergeants is more consequential for aggregate arrest behavior than variation in officers.

I present variance decomposition results in Table 2, which include decompositions using raw fixed effects, Bayes-shrunken fixed effects, and two bias-correction methods. The raw sergeant fixed effects suggest that sergeants explain roughly 5% of the variation in arrests.

[17] To avoid double-counting arrests attributed to multiple officers, I adjust the arrest measure so that each officer receives credit for only half an arrest when two officers are listed on the same report.

However, this estimate is likely upward-biased due to measurement error and limited mobility. When using the Bayes-shrunken effects, the share of variation explained by sergeants falls to 3.43%. Results using the Andrews et al. (2008) and KSS bias-corrections are consistent with the shrinkage-based estimate, confirming that limited mobility is unlikely a major concern in this context. In contrast, officer fixed effects explain a substantially larger share of the variation in arrests—just over 70% of the spell-level variance. Given that arrests are estimated at the officer level, rather than at the level of patrol units, the larger contribution of officers relative to sergeants is not surprising. However, as shown previously, the aggregate influence of sergeants is magnified by the number of officers they supervise—an effect not captured in the variance decomposition. The fourth row of Table 2 reports the covariance between sergeant and officer effects. I find evidence that high-arrest officers tend to sort to low arrest sergeants, however the magnitude of sorting is small and accounts for no less than -1.51% of the total variation across each specification. Consistent with institutional practices that constrain an officer's ability to select specific sergeants, sorting—even on fixed characteristics—appears to be limited.

The variance decompositions in this setting can be indirectly compared to estimates of managerial influence on productivity in other industries. My estimate of the variance in arrests attributable to sergeants—3.4%—is roughly half the size of the corresponding figure in Adhvaryu et al. (2024), who find that line managers in Indian garment factories explain 7.3% of the variance in worker productivity. However, the role of frontline personnel differs sharply: in my setting, officers account for 72.3% of the variance in arrests, compared to just 5.4% of the variance in output explained by workers in the garment factory context. These findings underscore the substantial discretion exercised by police officers relative to workers in more structured production environments.

Lazear et al. (2015) estimate that in a technology-based services firm, a one standard deviation increase in manager effects raises productivity 2.6 times more than a comparable increase in worker effects, assuming the manager oversees an average-sized team. A similar calculation in my setting yields a ratio of 1.44, again suggesting that while sergeants explain less variation in output than managers in more routinized settings, their effects are still economically meaningful and amplified by their leverage over multiple officers.

That sergeants influence police enforcement behavior to this degree is noteworthy for two reasons. First, it shows that supervisors can shape worker behavior even in settings characterized by complexity and high discretion. Second, in the context of ongoing debates about police reform, these findings suggest that managerial interventions—such as replacing high-arrest sergeants or altering their enforcement orientation—could offer a more targeted and scalable approach to changing frontline behavior than broad-based reforms aimed at

officers or recruits.

5.2 Diagnostic Checks

This section conducts diagnostic checks to address concerns regarding the validity of the sergeant effect estimates. To begin, I provide evidence in support of the identifying assumption that officer mobility is exogenous with respect to unobserved determinants of officer arrest behavior. If this assumption is violated, then the estimated sergeant effects may instead capture changes in an officer's broader decision environment. As discussed in Section 4, there are three prominent sources of endogenous mobility that could bias the sergeant effect estimates. I assess each of these identification threats in turn.

First, I consider the potential for endogenous mobility based on trends in officer behavior or local crime conditions. There may be concern that officers are assigned to sergeants based on recent changes in their arrests. For example, an officer may make fewer arrests after attending a mandated training program. If that officer is then reassigned to a sergeant with a lower enforcement orientation, part of the subsequent decline in arrests may be incorrectly attributed to the new sergeant, rather than to the pre-existing downward trend. To test for this, I implement an event study design that examines arrest behavior around the time of sergeant switches:

$$Arrests_{et} = \alpha_e + \sum_{k \neq -1} [\pi_0^k D_{et}^k + \pi_1^k D_{et}^k (\Delta \hat{\psi}_e)] + x'_{et}\beta + \epsilon_{et}. \tag{4}$$

Here, e indexes a switching event, uniquely determined by officer i and the switch month T, and k indexes months relative to T. The variable D_{et}^k is an indicator for an observation being k months relative to the switch. The coefficients π_0^k capture dynamics related to a change in sergeants which are common across all switches. I include baseline model controls (tenure, sector-watch fixed effects, and day-off group fixed effects) to adjust for time trends and an event fixed effect, α_e, in order to control for differences in baseline arrest rates prior to the switch.

The parameters of interest are the π_1^k's, which capture period-specific heterogeneity depending on the size of the change in sergeant effect. I test for endogenous reassignment by examining the pre-move event study coefficients. The event study model also nests a test for general misspecification of the sergeant effects, as equation 1 implies that a sergeant switch results in an instantaneous and non-degrading change in arrests. I estimate equation 4 using the balanced event study sample, so that $k \in [-5, 4]$.

I plot the event study coefficients in Figure 4. Reassuringly, there is no evidence of heterogeneous trends in arrest behavior prior to officers changing sergeants.. An F-test

of joint significance for the pre-move coefficients yields a p-value of 0.8473 (see column 1 of Supplementary Appendix Table B.1). Moreover, following a switch to a high-arrest sergeant, an officer's arrests immediately increase and remain elevated throughout the panel, in line with the insights from the nonparametric event study in Figure 1.[18]

While the previous test examines sorting based on trends in officer behavior, it does not address the possibility that officer reassignments are correlated with trends in local crime. One concern is that sergeants with a preference for aggressive enforcement may respond more strongly to increases in crime in their sectors. In such cases, they might request to fill officer vacancies when crime is rising, resulting in officers making more arrests after the move—not due to the new sergeant's influence, but because they are assigned to a location with increasing demand for enforcement. To assess this possibility, I test whether changes in crime predict changes in sergeant assignments. Specifically, I examine whether average unit-level changes in sergeant fixed effects are associated with trends in crime, measured by the volume of 911 calls in the relevant sector-watch.[19] I report estimates from regressions of changes in the sergeant fixed effects on 911 call trends in Supplementary Appendix Table B.2, separately for officers moving into and out of each sergeant's unit. Joint F-tests on the pre-period 911 call coefficients suggest that crime trends do not predict the direction or magnitude of changes in sergeant effects, mitigating concerns that crime trends confound the sergeant effects.

A second identification concern is that changes in sergeant assignment may coincide with unobserved shocks that independently affect officer behavior. For instance, officers might be reassigned to high-arrest sergeants at the same time a department-wide policy change encourages more aggressive enforcement. In such cases, increases in arrests following a sergeant switch could reflect the policy change, rather than the causal effect of the new sergeant. To test for this possibility, I implement a placebo event study using *incumbent* officers—those who are already working under the new sergeant at the time a switching officer joins the unit. If wider policy shocks are correlated with the timing of the sergeant switch, then these shocks should also affect the arrests of incumbent officers during the same time window.

For each switching event e in which officer i changes from sergeant j to sergeant j', I

[18]The size of the effect after moving is not statistically distinguishable from 1, which is reassuring, as the π_1^k's are interpreted as the change in arrests following a move to a sergeant who induces one more arrest per month than the previous sergeant.

[19]I use 911 calls as a measure of crime rather than crime reports, as crime reporting is endogenous to police activity (Weisburd, 2021). Although civilian willingness to call 911 may also be influenced by policing strategies (Ang et al., 2024), I treat 911 calls as a more appropriate proxy for underlying crime in this context. Unlike crime reports, which are often generated by officer-initiated activity, 911 calls are initiated by civilians and are thus less likely to mechanically respond to differences in officer enforcement.

model the number of arrests made by officers $l \neq i$ managed by sergeant j' 5 months before the switch and 4 months after:

$$Arrests_{let} = \alpha_{le} + \sum_{k \neq -1} [\pi_0^k D_{et}^k + \pi_1^k D_{et}^k (\Delta \hat{\psi}_e)] + x'_{let}\beta + \epsilon_{let}. \tag{5}$$

The model takes a form similar to equation 4. Once again, I am interested in the π_1^k terms, which describe how arrests made by incumbent officers in month k change when the difference in the effects of sergeants j' and j increases by 1. This model also provides a secondary test for endogenous crime trends, as we would expect arrests to increase for incumbent officers prior to a positive switch in sergeant effects if larger sergeant effects are driven by growing demand for arrests. In Supplementary Appendix Figure A.7, I present the estimates and 95% confidence intervals for the event study coefficients. The estimates are close to 0 and statistically insignificant across all months relative to the new officer's switch date. This supports the interpretation that the estimated sergeant effects reflect differences in managerial behavior, rather than correlated shocks in the officers' decision environments.

The third identification concern relates to match-specific error components. If officers sort to sergeants with whom they have a comparative advantage in making arrests, then the model will be misspecified and the fixed effects biased. However, the event study results in Figure 1 provide no evidence that officers and sergeants sort on match quality, suggesting that such sorting is unlikely to bias the estimates.

A closely related concern is the assumption of additively separable officer and sergeant fixed effects. If sergeant effects were officer-specific, then the separate officer and sergeant fixed effects would not be informative and may be a product of statistical noise. I conduct two tests to assess this assumption. First, following Card et al. (2013), I examine the average residuals from equation 1 separately by groups of officer and sergeant effects. Specifically, I divide each officer-month observation into quintiles of officer and sergeant effects. If the additive separability assumption did not hold, then I would expect the model to systematically under- or over-estimate arrests for certain officer-sergeant groups. For example, if aggressive officers perform especially well under disengaged sergeants, we would expect large positive residuals in the cell corresponding to top-quintile officers paired with bottom-quintile sergeants. Supplementary Appendix Figure A.8 demonstrates that the mean residuals do not exhibit any clear patterns that would indicate a violation of the additive separability assumption. Across all officer-sergeant groups, the residuals are relatively small, ranging from -0.1 to 0.18, suggesting that the threat of misspecification is minimal in this setting.

The second test of additive separability compares the explanatory power of the baseline specification to a fully saturated model that contains a fixed effect for each officer-sergeant

pair. I report the R^2 and Adjusted R^2 for these models in Supplementary Appendix Table B.3. The fully saturated model fits better than the baseline, though the increase in Adjusted R^2 of 0.054 suggests that match components play a limited role in this setting. Insofar as match effects do matter, my findings suggest the additively separable model is a useful approximation.

In practice, each sergeant effect is identified from a relatively small number of officer switches—33.4, on average—which raises the concern that the estimated effects may be driven by noise, even after accounting for measurement error. To show that sergeant effects capture meaningful variation in arrests, I estimate a set of 'placebo' sergeant effects by randomly reallocating sergeants to officers, preserving the number of unique officers for each sergeant. For each random assignment, I estimate the variance in arrests attributable to these placebo sergeants. I repeat this procedure 100 times and plot the resulting distribution of placebo variance estimates in Supplementary Appendix Figure A.9, along with the KSS variance estimate from Table 2. To be conservative, I do not apply bias correction to the placebo estimates, which biases them upward and makes the test more stringent. Reassuringly, the placebo estimates are close to 0 and my model variance estimate lies well outside a 95% confidence interval of the sergeant effect variance that would be obtained by chance.

In sum, the findings from this section suggest that the sergeant effects identify meaningful changes in officer behavior that are attributable to supervision.

5.3 Serious and Low-level Crime Enforcement

I next examine the relationship between sergeants' effects on serious and low-level enforcement—two distinct domains that reflect different crime-prevention priorities for the department. Figure 5 plots each sergeant's estimated low-level arrest effect (horizontal axis) against their serious arrest effect (vertical axis), along with a best-fit line. The correlation between the two dimensions is 0.11, statistically significant at the 5% level, indicating only a weakly positive relationship between a sergeant's impact on serious and low-level arrests. This relationship appears to be driven largely by a small subset of sergeants with consistently low effects: excluding the bottom 5% of low-level arrest sergeants reduces the correlation to 0.07, which is no longer statistically significant.[20]

These findings do not support the presence of systematic complementarities in how sergeants influence different types of enforcement. Rather, they reveal that a significant number of sergeants trade off one type of enforcement for another. Indeed, over 20% of sergeants fall into the top tercile of effects for one crime type and the bottom tercile for

[20]Results are similar when categorizing arrests using the FBI's traditional index crime classification; see Figure A.11 in the Supplementary Appendix.

the other (see Figure A.10 in the Supplementary Appendix). While some sergeants appear to specialize—inducing either more serious arrests or more low-level arrests (quadrants II and IV of Figure 5)—others exert broader influence, encouraging both types of enforcement (quadrant I), or suppressing arrest activity across the board (quadrant III). Since police are responsible for enforcing both serious and low-level offenses, this variation implies that front-line supervisors play a key role in determining how police enforcement priorities are operationalized on the ground.[21]

Because police enforcement is highly discretionary, one might interpret the variation in sergeant effects between crime types as simply reflecting the broader discretion available to all frontline officers. As shown in the previous section, officers themselves account for a substantial share of variation in arrests, raising the possibility that officers, rather than their supervisors, are primary drivers of enforcement priorities. To evaluate this possibility, I examine the relationship between serious and low-level *officer* effects in Supplementary Appendix Figure A.12. In contrast to the weak correlation among sergeant effects, officer effects exhibit strong complementarity across crime types: the correlation is 0.56, and remains strong and positive across the full range of low-level arrest effects. This pattern indicates that differences in officer enforcement are more likely driven by overall productivity rather than differing crime priorities. In other words, officers who make more low-level arrests also tend to make more serious arrests, consistent with general differences in effort or activity. This points to a distinctive managerial role for sergeants in shaping which types of crimes are policed most intensely.

Given the heterogeneity in sergeants' priorities over serious and low-level enforcement, an important question is which enforcement focus, if either, is more effective in advancing the police department's objectives. To explore this, I estimate the impact of serious and low-level sergeant effects on crimes reported through 911. I average the serious and low-level effects of all officers and sergeants within each sector-by-watch-by-month-by-year cell.[22] I then use OLS to estimate a model of 911 calls as a linear function of sergeant effects, officer effects, and controls.

[21] In Supplementary Appendix E, I examine whether observable sergeant characteristics are predictive of their enforcement effects. I do not find evidence of significant differences by race, gender, or age. However, I do find evidence that sergeants who score below average on the civil service exam that is used to make promotions tend to have larger low-level arrest effects. These findings align with recent evidence from Dahis et al. (2025) showing that skills measured by civil service exams are reflected in future job performance.

[22] About one-third of sector-watch-month-years have 2 assigned sergeants, since there may be multiple distinct patrol units in a sector-watch. The other two-thirds of sector-watch-month-years have only 1 assigned sergeant.

$$Log(911Calls_{sw,m}) = \alpha_1 \bar{\hat{\psi}}^S_{sw,m} + \alpha_2 \bar{\hat{\psi}}^L_{sw,m} + \alpha_3 \bar{\hat{\theta}}^S_{sw,m} + \alpha_4 \bar{\hat{\theta}}^L_{sw,m} + X_{sw,m} + \epsilon_{sw,m} \quad (6)$$

I estimate specifications that include sector-by-watch fixed effects as well as both sector-by-watch and month-by-year fixed effects. Similar to the two-way fixed effects model, α_1 and α_2 are identified by sergeants who change sector-watch assignments. Unbiasedness of these estimators requires an analogous assumption of as-good-as-random mobility of sergeants between sector-watch assignments, which is supported by the same mechanism used by DPD to assign sergeants to sector-watches as is used to assign officers.[23] Under this assumption, α_1 and α_2 capture the impact of sergeant-induced serious and low-level enforcement on 911 calls. I estimate models for both the total number of 911 calls as well as violent 911 calls, which include shootings, assaults, violent disturbances, and robberies.

Results are presented in Table 3. Across each specification, the impact of serious and low-level sergeant effects are economically small and statistically insignificant. The 95% confidence intervals rule out (total and violent) 911 call reductions larger than 3.4% due to a sergeant who induces one more low-level arrest per month. I can rule out call reductions larger than 6% as a result of sergeants who induce one more serious arrest per month. Scaling these estimates by the standard deviation of the low-level and serious sergeant effects, I am able to rule out 911 call reductions larger than 2.1% and 0.5%, respectively.[24] These results suggest that heterogeneity in sergeants' enforcement priorities does not translate into meaningful differences in crime reduction. Given recent evidence that low-level arrests can entrench civilians in cycles of criminality (Agan et al., 2023), sergeants who prioritize this kind of enforcement without meaningfully improving public safety may harm net social welfare.

These findings also have important implications for law enforcement policy. Since serious and low-level arrests are strongly complementary at the officer level but only weakly related at the sergeant level, shifting enforcement priorities may be more feasible by targeting sergeants rather than officers. To illustrate this, I simulate a hypothetical personnel policy that replaces the top 5% of sergeants with the largest low-level effects with median-effect sergeants for both enforcement types. Although stylized, this intervention would reduce low-level arrests by 1,018 over the five-year sample period while *increasing* serious arrests

[23] An additional concern is that both the officer and sergeant effects are estimated objects, creating the potential for attenuation bias due to measurement error. However, as Walters (2024) shows, using Bayes-shrunken fixed effect estimates as explanatory variables in OLS corrects the bias that would arise from instead using the noisy raw fixed effect estimates. This limits the threat of measurement error in this case.

[24] The standard deviation is 0.64 for the Bayes-shrunken low-level sergeant effects and 0.09 for the Bayes-shrunken serious sergeant effects.

by 10. This pattern reflects the weak correlation between the two dimensions of sergeant effects: high low-level arrest sergeants are equally likely to have high or low serious effects. As a result, reducing aggressive low-level enforcement through sergeant reassignment need not compromise serious crime enforcement. By contrast, officer-targeted reforms must navigate a stronger productivity tradeoff: officers who make more low-level arrests also tend to be more active in serious enforcement. Therefore, policies aimed at reducing low-level arrests among officers risk broader reductions in enforcement activity—potentially leading to de-policing effects (Devi and Fryer, 2020). These results underscore the managerial leverage that sergeants have in shaping enforcement priorities, and suggest that supervisor-level interventions may offer a targeted and cost-effective lever for reform.

6 Mechanisms

6.1 What Officer Behaviors Drive Sergeant Effects?

The previous section provides evidence that sergeants differentially affect their officers' serious and low-level arrest behaviors. However, this evidence cannot clarify how officers generate more (or less) serious and low-level arrests under different sergeants. In this section, I investigate this question by estimating the relationship between sergeant effects and a broad range of officer behaviors connected to serious and low-level arrests. To do so, I estimate regressions of the following form:

$$y_{it}^{c} = \alpha_{L}^{c}\hat{\psi}_{J(i,t)}^{L*} + \alpha_{S}^{c}\hat{\psi}_{J(i,t)}^{S*} + \theta_{i}^{c} + x_{it}'\beta^{c} + \nu_{it}^{c}, \tag{7}$$

where y_{it}^{c} is an outcome of type c for officer i in year-month t. These outcomes include granular types of arrests, as well as secondary enforcement outcomes such as uses of force or civilian complaints. As in the baseline model, I include officer, sector-watch, and day-off group fixed effects, as well as controls for officer tenure. The variables $\hat{\psi}_{J(i,t)}^{L*}$ and $\hat{\psi}_{J(i,t)}^{S*}$ represent the Bayes-shrunken low-level and serious sergeant effects, respectively, each divided by their standard deviations. The α^{c} coefficients capture the change in y_{it}^{c} associated with a one standard deviation increase in the low-level/serious effect of an officer's sergeant. I report estimates of these coefficients scaled by their relevant dependent variable means in Figure 6.[25]

I begin by examining how sergeant effects relate to arrests for specific crimes, focusing on the three most common offenses that comprise the bulk of serious and low-level arrests. Figure 6a shows that a 1SD increase in the serious sergeant effect is associated with sizable

[25]Regression tables for all outcomes are included in the Supplementary Appendix.

increases in arrests for each of the most frequent serious crimes: domestic violence, theft, and DWI. The strongest relationship is for domestic violence arrests, which rise by 37% relative to a mean of 0.45 arrests per month. In contrast, a 1SD increase in low-level sergeant effects *reduces* domestic violence arrests by 3.5% relative to the mean, suggesting that sergeants induce crime-specific trade-offs even though the two dimensions of sergeant effects are uncorrelated in the aggregate. There is no evidence that low-level sergeant effects change enforcement of theft or DWI.

Among low-level offenses, drug arrests exhibit the most pronounced divergence in response to serious versus low-level sergeant effects. A 1SD increase in low-level sergeant effects raises an officer's drug arrests by more than 50% relative to the mean of 0.31 arrests per month. Roughly 90% of this increase is attributable to arrests for simple possession rather than drug distribution (see Supplementary Appendix Table B.8). In contrast, a 1SD increase in serious sergeant effects is associated with a 10% reduction in drug arrests. Both serious and low-level sergeant effects are positively associated with arrests for warrants and disorderly conduct, though the magnitude of the association is consistently larger for the low-level effect.

Prioritizing low-level enforcement has implications for racial disparities, since Black civilians make up a disproportionate share of drug arrests. As shown in Figure 6b, sergeants who encourage greater low-level enforcement disproportionately increase arrests of Black civilians relative to white or Hispanic civilians, even after accounting for baseline racial differences in arrest rates. In contrast, serious sergeant effects increase arrests across all racial groups by similar proportions.

Next, I examine how sergeants influence the channel through which arrests are made—specifically, whether arrests are primarily the result of proactive officer behavior or reactive responses to 911 calls. Officers can make arrests either through self-initiated enforcement—such as traffic stops, checking abandoned buildings, or stopping pedestrians—or through dispatched 911 calls. Figure 6c presents estimates using arrests from these two sources as outcomes. I find that sergeant-induced serious arrests are initiated entirely through 911 calls, while low-level arrests are driven by a combination of calls and officer-initiated interactions. Officer-initiated interactions account for over 60% of the total increase in arrests generated by low-level sergeant effects, which is particularly striking given that officer-initiated arrests only account for 45% of arrests in the sample.

While officer-initiated arrests are clearly discretionary, the mechanism through which 911 call-based arrests vary is less straightforward. Two possibilities could explain increased arrests from calls: exposure to more 911 calls, or a lower threshold for arrest conditional on responding to a call. In Dallas, officers have some autonomy in deciding whether to take

calls outside of their assigned area—either by volunteering for available calls or by working overtime. Thus, part of the sergeant effect may operate by encouraging officers to engage with more calls than they otherwise would.

To test this, I estimate the impact of sergeant effects on officers' call activity as well as the likelihood that their calls generate an arrest. I report results from these estimations in Figure 6d. Both serious and low-level sergeant effects increase the total number of 911 calls answered by officers. However, this effect is more pronounced for serious sergeant effects (3.1% vs. 1.6%, relative to a mean of 73.5 calls per officer-month). Importantly, earlier results showed no evidence that sergeants influence the volume of 911 calls made in their assigned sectors. Thus, these findings suggest that officers working under high-arrest sergeants, particularly those with high serious effects, respond to more 911 calls—not because crime increases, but likely due to sergeant-driven shifts in officer effort or prioritization.

While serious sergeant effects are associated with greater increases in the number of 911 calls answered, low-level sergeant effects appear to have a stronger influence on *how* officers respond once they arrive. A 1SD increase in low-level effects results in a 6.7% increase in the likelihood of an arrest being made, relative to a mean of 0.062. By contrast, serious sergeant effects have no meaningful impact on this probability—their effect is near zero and statistically insignificant. These results indicate that the arrest gains associated with serious sergeant effects are proportional to increased call volume, while those associated with low-level sergeant effects reflect both increased call volume and a higher arrest rate per call.

Taken together, the evidence suggests that serious sergeant effects operate primarily by motivating officers to respond to more calls, increasing enforcement through elevated effort. In contrast, low-level sergeant effects alter officers' discretionary behavior, not only by raising the share of officer-initiated arrests but also by lowering the threshold for arrest during civilian-initiated encounters. This distinction underscores how different managerial styles shape distinct forms of enforcement, with low-level sergeants exerting influence over the most discretionary—and potentially most contentious—policing decisions.

When officers make more discretionary arrests, they may be choosing to enforce offenses they would otherwise overlook. Alternatively, they may be lowering the evidentiary threshold for making arrests—an approach more likely to result in case dismissals and failed prosecutions. While the welfare implications of the former depend on the value of enforcing certain offenses, the latter scenario entails clear costs, as low-quality arrests can burden courts and harm civilians without contributing to public safety. To disentangle these two possibilities, I estimate the impact of serious and low-level sergeant effects on officers' conviction rates. Because officers frequently make no arrests in a given month, I cannot directly estimate equation 7 using conviction rates as the outcome. Instead, following Gudgeon et al. (2023),

I separately estimate the effects of serious and low-level sergeant effects on the number of convicted arrests and on the total number of arrests. I use these estimates to calculate how changes in the serious and low-level sergeant effects affect the ratio of convicted arrests to total arrests, and then compare the new ratios to the ratio of averages. I plot the estimated changes in the conviction rate (for all arrests, as well as serious and low-level arrests individually) along with bootstrapped 95% confidence intervals in Figure A.13 in the Supplementary Appendix.

Both dimensions of sergeant effects are associated with higher overall conviction rates, suggesting that neither leads officers to make systematically weaker arrests that fail in court. Breaking conviction rates down by offense type, the results suggest that the increased overall conviction rates are driven by compositional changes rather than higher quality arrests. Serious sergeant effects are not associated with changes in the serious conviction rate and are actually associated with *lower* conviction rates for low-level arrests. On the other hand, low-level sergeant effects are associated with higher conviction rates for both types of arrests. Serious arrests have a higher conviction rate than low-level arrests, so conviction rates increase when working for sergeants with high serious effects, since serious effects make up a larger share of an officer's arrests. On the other hand, drug arrests have a conviction rate 4.75 times higher than the average low-level arrest. As a result, the disproportionate impact of low-level sergeant effects on drug arrests increases low-level conviction rates for their officers, while the negative impact of serious sergeant effects on drug arrests decreases low-level conviction rates for officers working with sergeants who prioritize the enforcement of serious crimes.

Finally, in Figure 6e, I report the impact of sergeant effects on two costly secondary police outcomes: use of force and complaints. I find that increases in both serious and low-level sergeant effects lead to more uses of force. However, the change is significantly larger for low-level sergeant effects. A one standard deviation increase in the low-level effect leads to 0.02 more uses of force per month, a 14% increase relative to the mean, compared to a change of 0.006 for an equivalent increase in serious effects. The results for complaints are noisy and inconclusive, owing to their rarity, since complaints occur in only 1.4% of all officer-months. The point estimates suggest that low-level sergeant effects increase complaints and serious sergeant effects decrease them, however neither estimate is able to rule out changes in the opposite direction that are close to 10% of the mean. Both serious and low-level sergeant effects lead to more officer activity and more formal interactions with civilians, likely contributing to increased use of force. However, the stark difference in effect sizes suggests that targeted low-level enforcement may lead to violent escalation, potentially disproportionate to the costs of the crimes it aims to address.

Taken together, the findings in this section show that sergeants shape serious and low-level arrests through distinct behavioral mechanisms. Serious arrests increase primarily because sergeants induce officers to respond to more 911 calls. In contrast, low-level arrests are driven by officer-initiated, discretionary encounters, especially those involving drug possession.These behavioral patterns support the interpretation that sergeant effects reflect differing enforcement priorities. As such, the choices of individual supervisors can generate inconsistencies in how law enforcement policies are implemented on the ground. This has important implications for public service delivery: sergeant-level discretion can amplify inequality—as seen in the disproportionate effect of low-level enforcement on Black civilians—and may also undermine service quality by increasing harmful byproducts of policing, such as excessive use of force.

6.2 How Do Sergeants Change Officer Behavior?

While the evidence so far has established that sergeants alter distinct officer behaviors, it is not clear how they manage to do so. Like other public sector supervisors, sergeants lack access to many of the personnel tools available in the private sector. They cannot offer performance pay, terminate employees, or meaningfully influence promotions. They also face limited capacity to directly supervise their employees' actions.[26] In this section, I evaluate two measurable sergeant behaviors that may shape officer conduct: leading by example and direct monitoring. As explained in Section 2, officers may internalize their sergeants' enforcement activities as performance expectations, which can influence transfers into specialized units, commendations, or day-to-day job satisfaction due to strong cultural norms around following orders. I use two proxies for leading by example: a sergeant's own arrest activity and their frequency as a first responder to 911 calls. To distinguish between mechanisms driving serious versus low-level effects, I consider separately a sergeant's serious and low-level arrests.

Although monitoring is limited, sergeants may differ in their willingness to exploit the oversight tools they do have. For instance, some sergeants may more often assign themselves to officers' calls. This allows for real-time supervision and may also make officers more likely to seek guidance in future situations. I proxy for this mechanism using CAD call data to measure sergeant presence at their subordinates' calls.

To estimate the relevance of each mechanism, I regress monthly sergeant behaviors on

[26]For example, Wilson (1991) highlights compliance officers in the U.S. Department of Labor's Wage and Hours Division, attorneys in the Antitrust Division of the U.S. Department of Justice, soldiers during wartime, forest rangers, teachers, and (fittingly) police officers as examples of public employees whose managers cannot closely supervise their daily activities.

their estimated serious and low-level arrest effects. Because behaviors may be shaped by the sergeant's assignment, I leverage within-assignment variation. Specifically, for unit u managed by sergeant j in month t, I estimate models of the following form:

$$y_{jut} = \alpha_L \hat{\psi}_j^L + \alpha_S \hat{\psi}_j^S + \alpha_1 \bar{\hat{\theta}}_{ut}^L + \alpha_2 \bar{\hat{\theta}}_{ut}^S + x_{sw(u)} + \epsilon_{jut}, \tag{8}$$

where y is a behavior of sergeant j in unit u during year-month t. I include sector-watch fixed effects ($x_{sw(u)}$) to account for time and location-specific variation. Since some sergeant behaviors may respond to subordinate needs, I control for the average low-level and serious arrest effects of officers in the unit.

Table 4 reports the results. I find that low-level sergeant effects are associated with leading by example (columns 1-4). Sergeants with large low-level effects make significantly more arrests (column 1), almost entirely for low-level offenses (column 3). They also respond to more 911 calls as first responders (column 4). In addition, these sergeants engage in more direct monitoring: a one standard deviation increase in low-level effects corresponds to a 5.8% increase in responses to subordinate calls, relative to the mean (column 5).

In contrast, serious sergeant effects are not significantly associated with either leading by example or enhanced field monitoring. The corresponding coefficients are small and statistically indistinguishable from zero. One alternative mechanism, however, can be indirectly assessed: granting overtime in response to increased call activity. Given limited restrictions on overtime, sergeants can substantially increase officers' compensation by approving extra hours (Chalfin and Goncalves, 2023). Although I lack direct overtime data, I use shift records to examine calls and arrests made outside regularly scheduled hours. Table B.11 in the Supplementary Appendix shows that nearly 20% of the increase in calls associated with higher serious sergeant effects occurs outside regular shifts. Overtime arrests also rise, primarily for serious crimes. I find similar patterns for low-level effects, which are associated with increased after-hours calls and arrests. While imperfect, these proxies suggest that sergeants may use overtime approval to shape officer behavior.

Overall, these results show that sergeants overcome supervisory constraints through both formal and informal mechanisms. The institutional structure of police departments—particularly around patrol roles and overtime pay—provides supervisors with tools to generate pecuniary and non-pecuniary incentives. Several other plausible mechanisms—such as the use of commendations, transfers, or assignment preferences—remain unobservable in my data. Nonetheless, the evidence presented here reaffirms findings from the broader literature on public sector management: supervisors can leverage organizational features to shape the behavior of their employees (Fenizia, 2022).

7 Conclusion

Leveraging officer movements between sergeants in a large urban police department, this paper provides evidence that first-line police supervisors have meaningful causal effects on complex, discretionary officer enforcement decisions. I show that individual sergeants change both the amount of arrests officers make and how they prioritize enforcement for different criminal behaviors. My findings suggest that sergeants induce serious crime enforcement by increasing the number of 911 calls their officers respond to, and induce low-level enforcement by encouraging proactive, discretionary arrests—typically for minor drug offenses. Despite lacking traditional performance incentives, effective sergeants appear to use both formal and informal job features to influence officer behavior: they lead by example, monitor officer activity in the field, and approve overtime for actions aligned with their enforcement priorities. Sergeant effects on serious and low-level enforcement are not systematically related and do not appear to impact crime, suggesting that civilian experiences with law enforcement are meaningfully shaped by the subjective policy preferences of supervisors.

These findings have broader relevance for organizational settings characterized by a high degree of worker discretion—particularly those in the public sector, where decisions are complex, objectives are multifaceted, outputs are difficult to quantify, and incentives are weak. In law enforcement, I show that first-line supervisors significantly shape how street-level bureaucrats use their discretion. This insight is especially important for police reform, as sergeants may be a powerful lever for change. The weak correlation I find between sergeants' effects on serious versus low-level enforcement—combined with recent work questioning the value of low-level enforcement (Agan et al., 2023; Cho et al., 2023)—suggests that sergeants may be able to scale back low-level arrests without compromising serious crime enforcement. I also show that serious and low-level arrests are more complementary for officers than for sergeants, implying that officer-level reforms must tread carefully to avoid inducing de-policing, which may harm public safety (Devi and Fryer, 2020).

A limitation of my analysis is that I do not directly evaluate the effects of specific sergeant-level reforms, which presents a valuable direction for future research. While several officer-level training programs have shown promise (Owens et al., 2018; Mello et al., 2023), I am not aware of experimental or quasi-experimental evidence on the effects of supervisor training on outcomes such as those studied here. Given the prevalence of in-service trainings for police at all levels, this may be a promising and feasible research opportunity.

Finally, my results suggest that managerial selection is a critical determinant of public service delivery. However, it remains unclear how best to identify managers whose preferences align with an agency's objectives. One common selection mechanism in policing

and other public agencies, especially in developing countries, is competitive examination. I find suggestive evidence that exam performance is negatively associated with a sergeant's propensity to induce low-level arrests, though this result is limited by sparse data and low statistical power. Richer exam data could help unpack this relationship and clarify which components—such as written versus oral exams—best predict future managerial behavior (Dahis et al., 2025). Such evidence could inform exam design and offer lessons for management selection across a range of public sector contexts.

References

Abowd, J. M., Creecy, R. H., and Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data. *Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.*

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67:251–333.

Adger, C., Ross, M., and Sloan, C. (2022). The effect of field training officers on police use of force. *Working Paper*.

Adhvaryu, A., Bassi, V., Nyshadham, A., and Tamayo, J. (2024). No line left behind: Assortative matching inside the firm. *Review of Economics and Statistics*, pages 1–45.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. (2023). Managerial quality and productivity dynamics. *The Review of Economic Studies*, 90:1569–1607.

Agan, A., Doleac, J. L., and Harvey, A. (2023). Misdemeanor prosecution. *Quarterly Journal of Economics*, 138:1453–1505.

Alvarez, J., Benguria, F., Engbom, N., and Moser, C. (2018). Firms and the decline in earnings inequality in brazil. *American Economic Journal: Macroeconomics*, 10:149–189.

Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society Series A: Statistics in Society*, 171:673–697.

Ang, D., Bencsik, P., Bruhn, J., and Derenoncourt, E. (2024). Community engagement with law enforcement after high-profile acts of police violence. *NBER Working Paper 32243*.

Ba, B., Bayer, P., Rim, N., Rivera, R., and Sidibé, M. (2021a). Police officer assignment and neighborhood crime. *NBER Working Paper 29243*.

Ba, B. A., Knox, D., Mummolo, J., and Rivera, R. (2021b). The role of officer race and gender in police-civilian interactions in chicago. *Science*, 371:696–702.

Bacher-Hicks, A. and De La Campa, E. (2020a). The impact of new york city's stop and frisk program on crime: The case of police commanders. *Working Paper*.

Bacher-Hicks, A. and De La Campa, E. (2020b). Social costs of proactive policing: The impact of nyc's stop and frisk program on educational attainment. *Working Paper*.

Benson, A., Li, D., and Shue, K. (2019). Promotions and the peter principle. *Quarterly Journal of Economics*, 134:2085–2134.

Bertrand, M. and Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, 118:1169–1208.

Best, M. C., Hjort, J., and Szakonyi, D. (2023). Individuals and organizations as sources of state effectiveness. *American Economic Review*, 113:2121–2167.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *Quarterly Journal of Economics*, 128:1–51.

Bloom, N., Lemos, R., Sadun, R., and Reenen, J. V. (2015). Does management matter in schools? *Economic Journal*, 125:647–674.

Bloom, N., Sadun, R., and Reenen, J. V. (2012). Americans do it better: Us multinationals and the productivity miracle. *American Economic Review*, 102:167–201.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122:1351–1408.

Bratton, W. and Murad, J. (2018). *Precision Policing*. Manhattan Institute.

Brown, M. K. (1988). *Working the Street: Police Discretion and the Dilemmas of Reform*. Russell Sage Foundation.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality*. *The Quarterly Journal of Economics*, 128:967–1015.

CBS News (2023). Most americans think changes to policing are necessary — cbs news.

Chalfin, A. and Goncalves, F. (2023). Professional motivations in the public sector: Evidence from police officers. *NBER Working Paper 31985*.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104:2633–2679.

Cho, S., Gonçalves, F., and Weisburst, E. (2023). The impact of fear on police behavior and public safety. *NBER Working Paper No. 31392*.

Dahis, R., Schiavon, L., and Scot, T. (2025). Selecting top bureaucrats: Admission exams and performance in brazil. *Review of Economics and Statistics*, 107:408–425.

Devi, T. and Fryer, R. G. (2020). Policing the police: The impact of "pattern-or-practice" investigations on crime. *NBER Working Paper 27324*.

Dube, O., MacArthur, S. J., and Shah, A. K. (2025). A cognitive view of policing. *The Quarterly Journal of Economics*, 140:745–791.

Engel, R. S. (2000). The effects of supervisory styles on patrol officer behavior. *Police Quarterly*, 3:262–293.

Engel, R. S. (2001). Supervisory styles of patrol sergeants and lieutenants. *Journal of Criminal Justice*, 29:341–355.

Engel, R. S. (2002). Patrol officer supervision in the community policing era. *Journal of Criminal Justice*, 30:51–64.

Engel, R. S. and Worden, R. E. (2003). Police officers' attitudes, behavior, and supervisory influences: An analysis of problem solving. *Criminology*, 41:131–166.

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica*, 90:1063–1084.

Frederiksen, A., Kahn, L. B., and Lange, F. (2020). Supervisors and performance management systems. *Journal of Political Economy*, 128:2123–2187.

Gaure, S. (2013). lfe: Linear Group Fixed Effects. *The R Journal*, 5(2):104–116.

Gaure, S. (2014). Correlation bias correction in two-way fixed-effects linear regression. *Stat*, 3:379–390.

Giorcelli, M. (2019). The long-term effects of management and technology transfers. *American Economic Review*, 109:121–152.

Guarino, C. M., Maxfield, M., Reckase, M. D., Thompson, P. N., and Wooldridge, J. M. (2015). An evaluation of empirical bayes's estimation of value-added teacher performance measures. *Journal of Educational and Behavioral Statistics*, 40:190–222.

Gudgeon, M., Jordan, A., and Kim, T. (2023). Do teams perform differently under black and hispanic leaders? evidence from the chicago police department. *Working Paper*.

Hoffman, M. and Tadelis, S. (2021). People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of Political Economy*, 129:243–285.

Ichniowski, C., Shaw, K., and Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review*, 87:291–313.

Ingram, J. R., Weidner, R. R., III, E. A., and Terrill, W. (2014). Supervisory influences on officers' perceptions of less lethal force policy: A multilevel analysis. *Policing*, 37:355–372.

Jang, H., Lee, C. B., and Hoover, L. T. (2012). Dallas' disruption unit: Efficacy of hot spots deployment. *Policing*, 35:593–614.

Johnson, R. R. (2011). Officer attitudes and management influences on police work productivity. *American Journal of Criminal Justice*, 36:293–306.

Johnson, R. R. (2015a). Leading by example: Supervisor modeling and officer-initiated activities. *Police Quarterly*, 18:223–243.

Johnson, R. R. (2015b). Police organizational commitment: The influence of supervisor feedback and support. *Crime and Delinquency*, 61:1155–1180.

Kane, T. and Staiger, D. (2008). Estimating teacher impacts on student achievement: An experimental evaluation.

Kapustin, M., Neumann, T., and Ludwig, J. (2022). Policing and management. *NBER Working Paper 29851*.

King, W. R. (2005). Toward a better understanding of the hierarchical nature of police organizations: Conception and measurement. *Journal of Criminal Justice*, 33:97–109.

Kline, P., Saggio, R., and Sølvsten, M. (2020). Leave-out estimation of variance components. *Econometrica*, 88:1859–1898.

Lazear, E. P., Shaw, K. L., and Stanton, C. T. (2015). The value of bosses. *Journal of Labor Economics*, 33.

Lehman, C. (2024). Drug policing in the 21st century: Concepts and strategies for policing the new drug crisis. Technical report, Manhattan Institute.

Makowsky, M. D. and Stratmann, T. (2009). Political economy at any speed: What determines traffic citations? *American Economic Review*, 99:509–527.

Mas, A. (2006). Pay, reference points, and police performance*. *Quarterly Journal of Economics*, 121:783–821.

Mello, S., Ross, M., Ross, S., and Johnson, H. (2023). Diversity training and employee behavior: Evidence from the police. *Working Paper*.

Metcalfe, R. D., Sollaci, A. B., and Syverson, C. (2023). Managers and productivity in retail. *NBER Working Paper 31192*.

Morris, C. N. (1983). Parametric empirical bayes inference: Theory and applications. *Journal of the American Statistical Association*, 78:47.

Mummolo, J. (2018). Modern police tactics, police-citizen interactions, and the prospects for reform. *Journal of Politics*, 80:1–15.

Natapoff, A. (2016). *Oxford Handbook Topics in Law*. Oxford University Press.

Owens, E., Weisburd, D., Amendola, K. L., and Alpert, G. P. (2018). Can you build a better cop?: Experimental evidence on supervision, training, and policing in the community. *Criminology and Public Policy*, 17:41–87.

Rim, N., Rivera, R., Kiss, A., and Ba, B. (2024). The black-white recognition gap in award nominations. *Journal of Labor Economics*, 42:1–23.

Rivera, R. (2025). Do peers matter in the police academy? *American Economic Journal: Applied Economics*, 17:127–164.

Rivera, R. G. and Ba, B. (2022). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. *Working Paper*.

Santos, R. (2019). Community policing: A first-line supervisor's perspective. Technical report, Office of Community Oriented Policing Services, United States Department of Justice.

Slocum, L. A., Schaefer, B. P., Torres, L., Huebner, M. M. B., and Hughes, T. (2021). Warrant enforcement in louisville metro and the city of st. louis from 2006 – 2019: A cross-site analysis.

Van Maanen, J. (1984). Making rank: Becoming an american police sergeant. *Urban Life*, 13:155–177.

Walters, C. (2024). Empirical bayes methods in labor economics. *NBER Working Paper 33091*.

Weisburd, D. and Eck, J. E. (2004). What can police do to reduce crime, disorder, and fear? *Annals of the American Academy of Political and Social Science*, 593:42–65.

Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, 103:280–293.

Weisburst, E. K. (2024). Whose help is on the way? *Journal of Human Resources*, 59:1122–1149.

Wilson, J. (1991). *Bureaucracy*. Basic Books, London, England.

Tables and Figures

Table 1: Summary Statistics

	Full Sample	Analysis Sample	Event Study Sample
	(1)	(2)	(3)
1. Number of officers	2,067	1,805	833
2. Number of sergeants	387	347	287
3. Number of officers with >1 sgt.	1,856	1,623	833
4. Number of sergeants with >1 off.	384	344	270
5. Mean number of sergeants per off.	5.21	3.97	2.67
6. Mean number of officers per sgt.	27.7	20.6	7.74
7. Total officer-sergeant spells	15,355	8,432	2,247
8. Total switching events	13,288	5,798	1,277
9. Number of sector-watches	105	102	102
10. Mean number of sergeants per sector-watch	8.48	6.95	4.61
11. Arrests mean	3.81	3.80	3.65
SD	3.65	3.64	3.46
12. Low-level arrests mean	2.88	2.87	2.75
SD	3.03	3.02	2.88
13. Serious arrests mean	0.925	0.923	0.897
SD	1.29	1.29	1.26
14. Drug arrests mean	0.315	0.311	0.276
SD	0.931	0.928	0.885
15. Warrant arrests mean	0.771	0.766	0.739
SD	1.35	1.35	1.29
16. Disorderly conduct arrests mean	0.416	0.409	0.370
SD	0.941	0.921	0.839
17. Proactive arrests mean	1.71	1.70	1.61
SD	2.25	2.24	2.13
18. Convicted arrests mean	0.780	0.777	0.715
SD	1.30	1.30	1.21
19. Use of force mean	0.119	0.118	0.114
SD	0.324	0.322	0.318
20. Complaint mean	0.0139	0.0141	0.0151
SD	0.117	0.118	0.122
Number of observations	61,166	49,923	12,770

Notes: The table reports summary statistics for three samples. The Full Sample is the unrestricted sample of all patrol officers. The Analysis Sample contains all patrol officer months that satisfy the restrictions described in Section 3. The Event Study sample contains all officer-sergeant switching events in which the focal officer is observed with the pre-switch sergeant at least 5 months prior to the switch and the post-switch sergeant at least 4 months after the switch. Serious arrests are defined as arrests for index crimes as well as domestic violence, fraud, simple assault, and DUI. All other arrests are considered low-level. Drug (warrant/disorderly conduct) arrests are any arrests which contain a drug (warrant/disorderly conduct) charge and do not contain any other higher-level (i.e. serious) charges. An arrest is considered to be convicted if the arrest is matched to a court disposition and not dismissed; this includes guilty findings by judge, jury, or plea. Use of force (complaint) is a binary indicator for any use of force (complaint) taking place in a month.

Table 2: Variance Decomposition

	Raw		Shrinkage		Homosk. Bias-Correction		Heterosk. Bias-Correction	
	Component	% Share	Component	% Share	Component	% Share	Component	% Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Var(y^*)$	11.153	100.00%	11.153	100.00%	11.153	100.00%	11.159	100.00%
$Var(\psi)$	0.559	5.01%	0.382	3.43%	0.379	3.40%	0.378	3.39%
$Var(\theta)$	8.906	79.86%	8.002	71.75%	8.097	72.60%	8.068	72.31%
$Cov(\psi, \theta)$	-0.168	-1.51%	-0.157	-1.41%	-0.0592	-0.53%	-0.0599	-0.54%
$Var(\psi + \theta)$	9.129	81.86%	8.071	72.37%	8.357	74.93%	8.33	74.62%
N sergeants	347		347		347		344	
N officers	1805		1805		1805		1802	

Notes: This table presents the variance decompositions described in equation 2. As described in Section 4, y^* is the number of monthly arrests, residualized on sector-watch, day-off group, and a second-degree polynomial of tenure; ψ is the sergeant fixed effect; θ is the officer fixed effect. All statistics are calculated on data aggregated to the officer-supervisor pair. Columns (1) and (2) report results for the raw fixed effects estimates. Columns (3) and (4) use fixed effects that are multiplied by the Bayesian shrinkage factor, constructed as described in Section 4. Columns (5) and (6) use the bias correction method proposed by Andrews et al. (2008) that assumes homoskedastic error terms. This bias correction is implemented using the 'lfe' package in R (Gaure, 2013) and uses simulation methods to calculate the trace of large matrices, as described in Gaure (2014). As such, I report the average of 100 iterations. Columns (7) and (8) implement the Kline et al. (2020) bias correction method that allows for unrestricted heteroskedasticity in the error terms. This method can only be conducted on the leave-out connected set, which is why the number of sergeants and officers decrease. This implementation adapts the Julia package provided by Kline et al. (2020) and developed by Paul Courcera, which can be found at https://github.com/HighDimensionalEconLab/VarianceComponentsHDFE.jl.

Table 3: The Effect of Sergeants on Crime

	Log(911 Calls)		Log(Violent 911 Calls)	
	(1)	(2)	(3)	(4)
Low-level Sergeant Effect	-0.0096	-0.0129	-0.0049	0.0034
	[-0.0313; 0.0122]	[-0.0340; 0.0081]	[-0.0335; 0.0237]	[-0.0176; 0.0244]
Serious Sergeant Effect	0.0010	-0.0088	0.0026	-0.0013
	[-0.0504; 0.0524]	[-0.0606; 0.0431]	[-0.0593; 0.0644]	[-0.0533; 0.0507]
Observations	6,008	6,008	6,008	6,008
Y mean	6.0404	6.0404	4.5238	4.5238
Sector-by-Watch fixed effects	✓	✓	✓	✓
Month-by-Year fixed effects		✓		✓

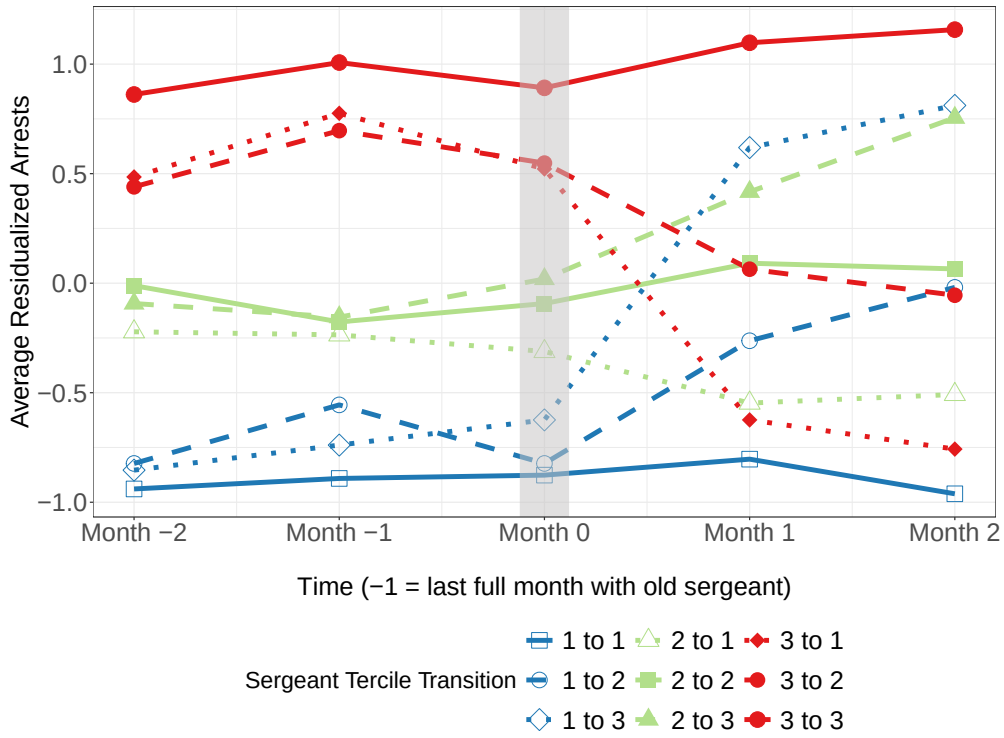
Notes: This table reports the estimated coefficients for the sergeant effects in equation 6. 95% confidence intervals for the coefficients are reported in brackets. The data are aggregated to the sector-by-watch-by-month-by-year level. Low-level (serious) sergeant effects are calculated by averaging over the Bayes-shrunken low-level (serious) sergeant effects for each sergeant assigned to the sector-watch in a given month. Each model includes controls for the average serious and low-level officer effects of officers in that same assignment. Violent 911 calls are calls made for shootings, robberies, assaults, and violent disturbances. Standard errors are clustered at the sector-by-watch level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table 4: Sergeant Effect Mechanisms

	Leading by Example				Monitoring
	Total Arrests	Serious Arrests	Low-Level Arrests	First-Responder Calls	Subordinate Calls
	(1)	(2)	(3)	(4)	(5)
Low-level Sergeant Effect	0.0777***	0.0062	0.0716***	0.4697*	0.3484*
	(0.0249)	(0.0081)	(0.0192)	(0.2749)	(0.2051)
Serious Sergeant Effect	-0.0168	-0.0049	-0.0119	0.0916	0.2218
	(0.0210)	(0.0068)	(0.0160)	(0.2708)	(0.1869)
Controls	✓	✓	✓	✓	✓
Observations	7,983	7,983	7,983	7,983	7,983
R^2	0.08394	0.03646	0.07873	0.13610	0.16558
Y mean	0.31605	0.08130	0.23475	3.9806	5.9782

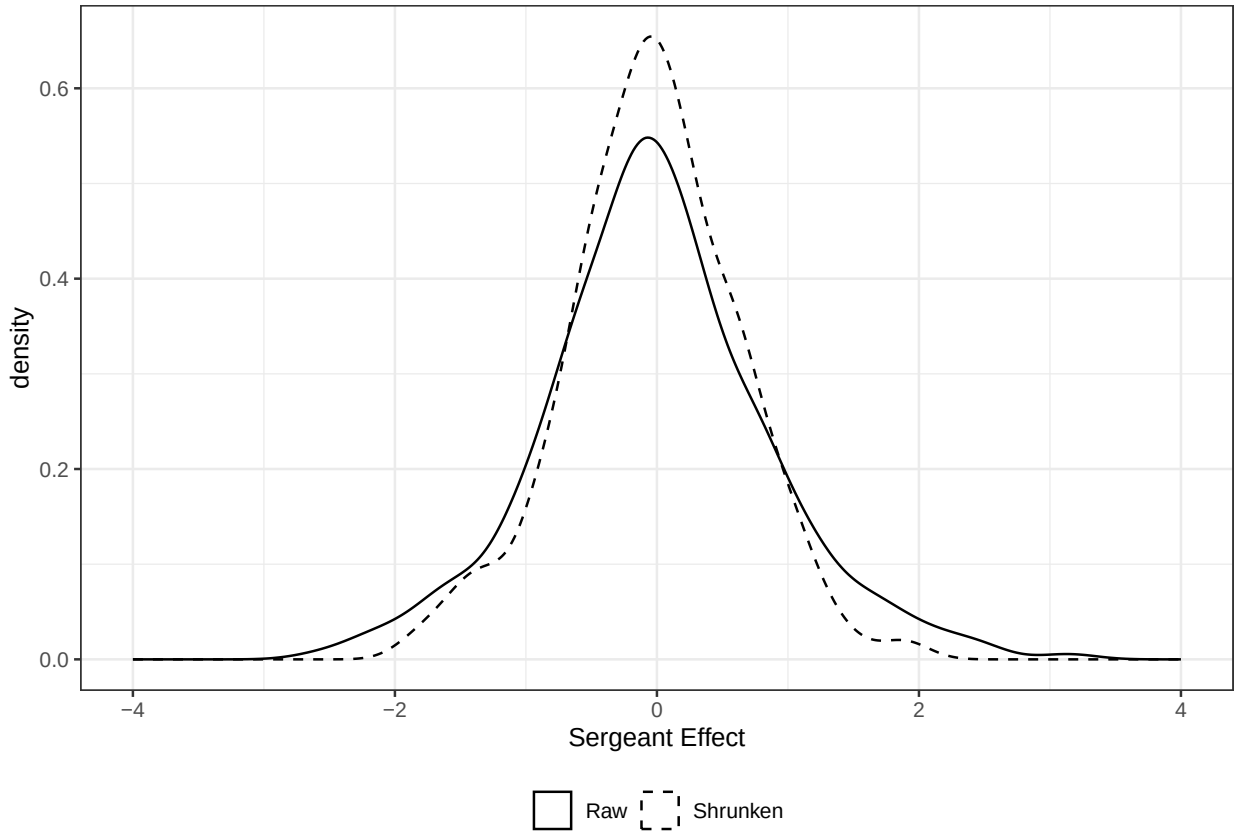
Notes: This table presents results from regressing measures of sergeant behavior on the estimated low-level and serious sergeant effects, as described by equation 8. Data are at the sector-watch by month level. Controls include the average estimated low-level and serious officer arrest effects for officers within the unit and sector-watch fixed effects. The outcome variables in each column are: (1) the number of arrests that the unit's supervisor makes in the month, (2) the number of those arrests which are serious, (3) the number of those arrests which are low-level, (4) the number of calls for service that the sergeant is first to respond to, and (5) the number of calls for service that a sergeant responds to in which their subordinates are also present. Standard errors are clustered at the sergeant level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Figure 1: Event Study Around Sergeant Switch



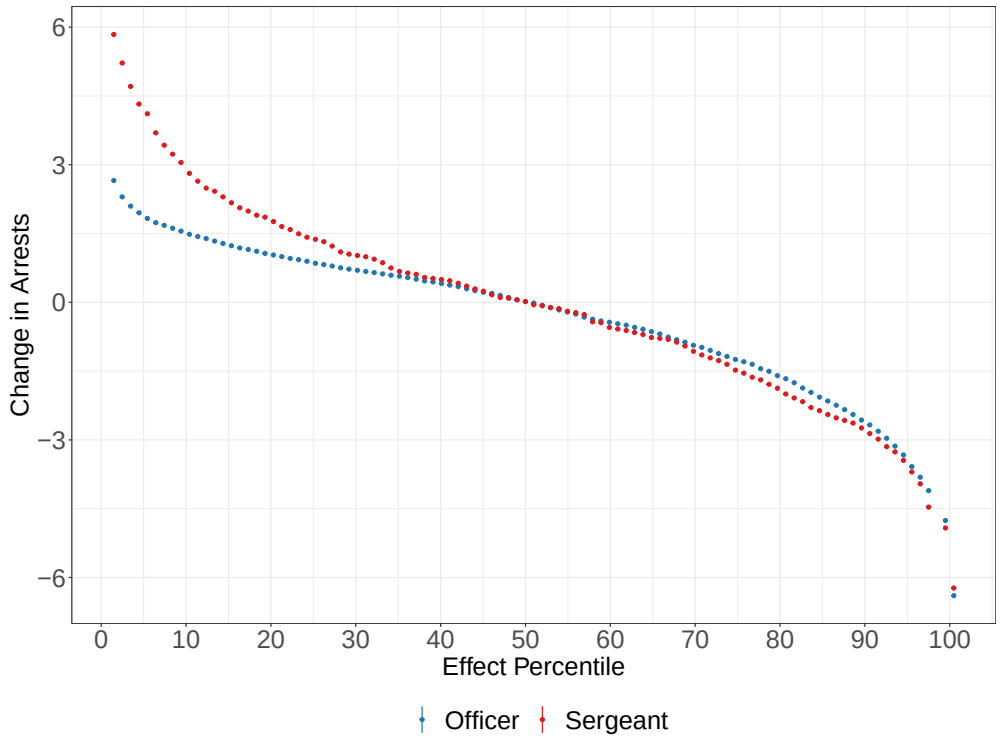
Notes: This figure plots the average number of arrests made by officers in the months around receiving a new sergeant by the magnitude of the sergeant change. In particular, I group sergeants into terciles according to the average number of residual arrests made by their officers throughout the sample. Each line then plots the average residualized arrests made by officers who transition between terciles, where the terciles of the previous and subsequent sergeant are described by "Sergeant Tercile Transition." Arrests are residualized by a second-degree polynomial of officer tenure and officer, sector-watch, and day-off group fixed effects using within-sergeant variation, as described in the text.

Figure 2: Distribution of Sergeant Effects



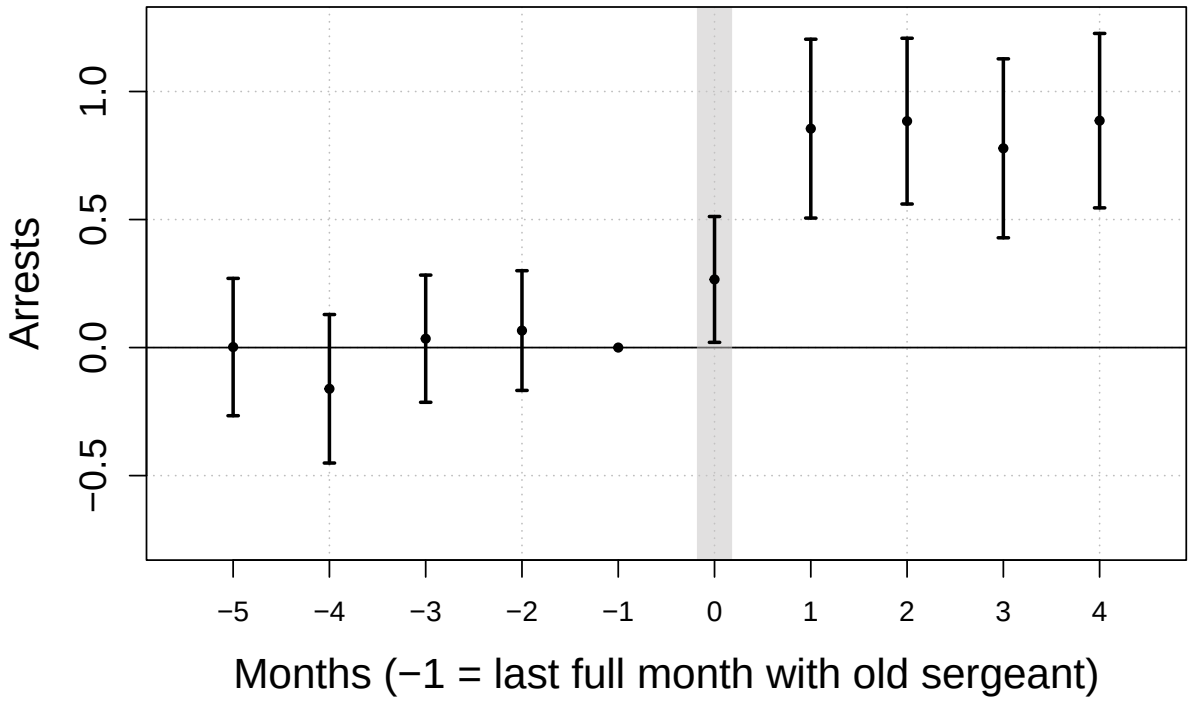
Notes: This figure plots the sergeant effects estimated using the sergeant fixed effects in equation 1. Sergeant effects are interpreted as the number of monthly arrests that an officer makes working under a sergeant, relative to the average sergeant. The solid line represents the raw effects obtained from estimating equation 1 using OLS. The dotted line represents the shrunken effects, which are the raw fixed effects multiplied by the Bayesian shrinkage factor as described in Section 4.

Figure 3: Impact of a Median Replacement



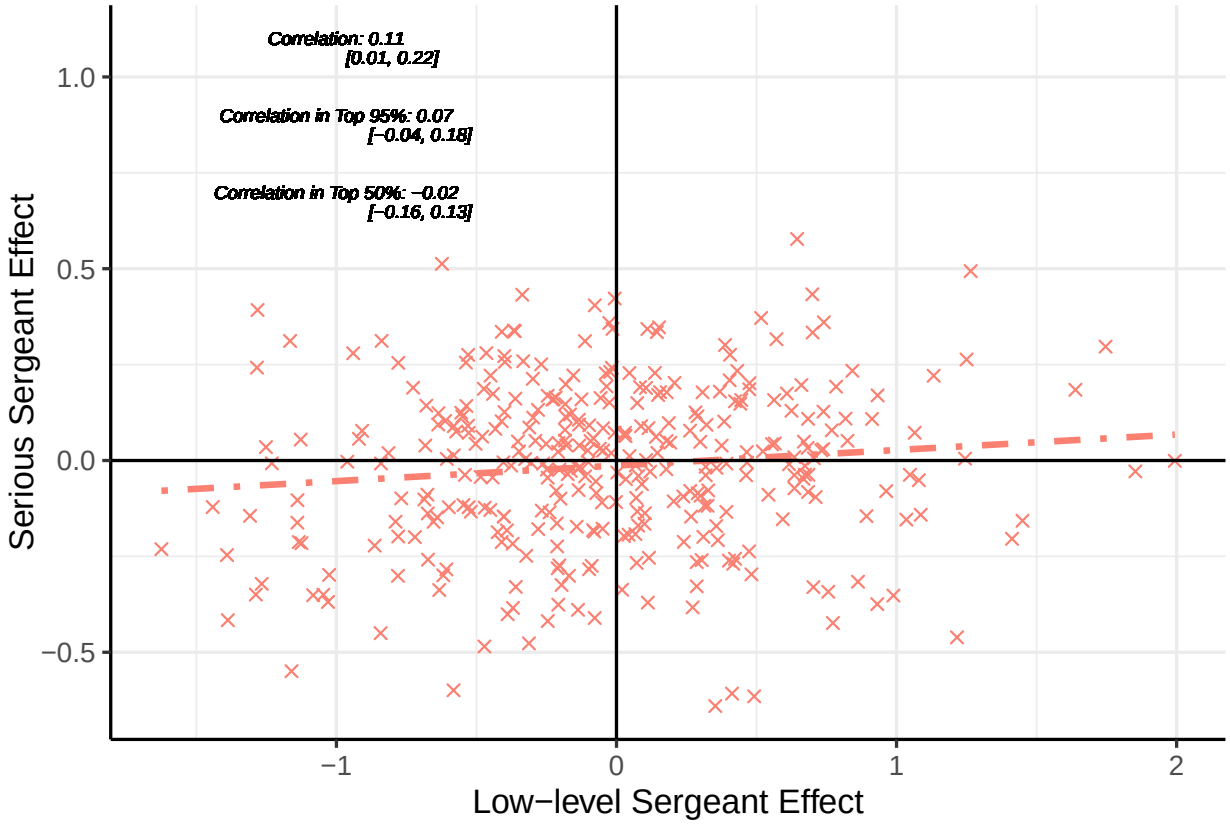
Notes: This figure plots the calculated effect of one-month replacements of sergeants (in red) and officers (in blue) with a median employee from the relevant effects distribution. Each sergeant is placed into their percentile in the effects distribution and the change in arrests that would be produced from replacing them with a median sergeant is calculated by subtracting each sergeant's effect from the median sergeant effect, and multiplying by the average number of officers managed in a month (6.33). I then plot the change in arrests against each percentile by averaging over all sergeants within that percentile. The change in arrests for officers is calculated identically, except I do not multiply by 6.33. For this exercise, sergeant and officer fixed effects are re-estimated using an adjusted arrest measure that only credits officers for half an arrest when there is another officer listed on the arrest report.

Figure 4: Event Study Coefficients



Notes: This figure presents estimates of $\hat{\pi}_1^k$ from the sergeant switching event study model described by equation 4, where k denotes the months around a sergeant switching event. The switch occurs in month 0. Month -1, the last full month an officer spends with their old sergeant, is used as the reference month. The model is estimated using the event study data that are balanced on [-5, 4]. Standard errors are clustered at the officer level.

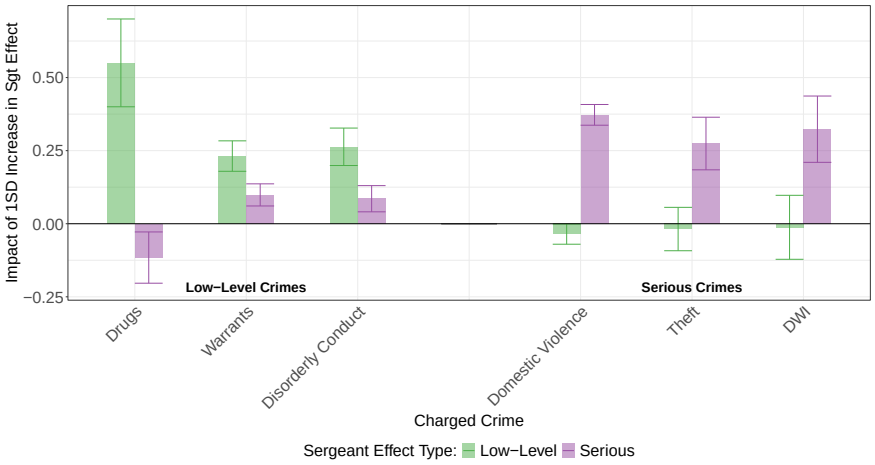
Figure 5: Relationship between low-level and serious sergeant effects



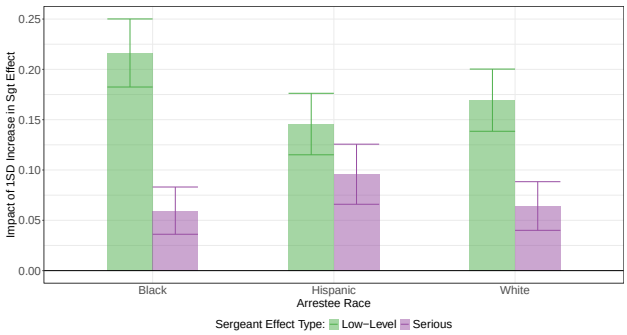
Notes: This figure displays a scatterplot of the relationship between the low-level sergeant effects and serious sergeant effects. Each point represents a sergeant. Low-level (serious) effects describe the sergeant effect on arrests for low-level (serious) crimes, defined as in Section 3. Both types of sergeant effect are shrunken using the Empirical Bayes procedure described in Section 4. Pearson correlations are provided in black text, with 95% confidence intervals for the correlations, calculated using the Fisher z-transformation, displayed below each correlation. Correlation in the Top 95% is calculated by dropping sergeants below the 5th percentile of low-level effects. Correlation in the Top 50% is calculated by dropping sergeants below the 50th percentile of low-level effects. A best fit-line is given by the dashed red line.

Figure 6: Drivers of Sergeant Effects

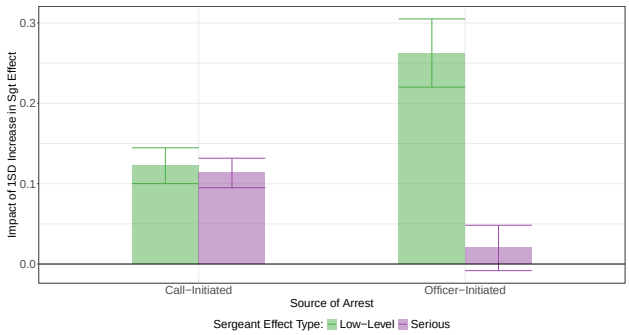
(a) Arrests by Crime Type



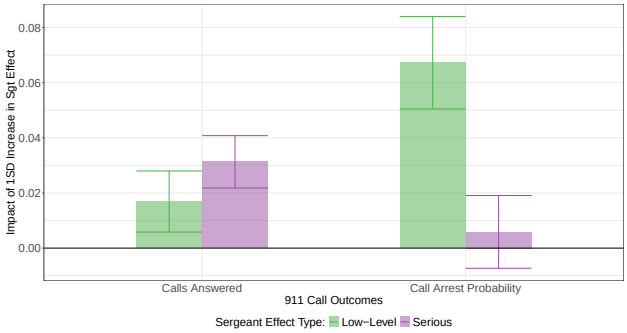
(b) Arrests by Race



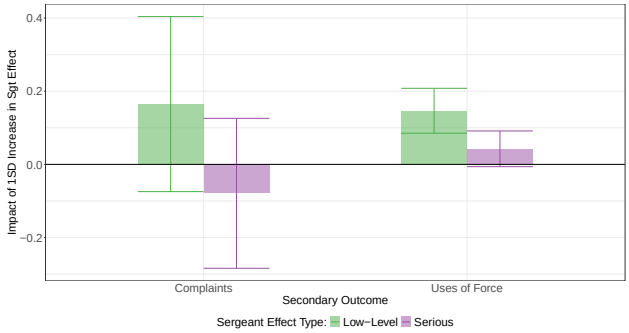
(c) Arrests by Interaction Source



(d) 911 Call Outcomes



(e) Secondary Police Outcomes



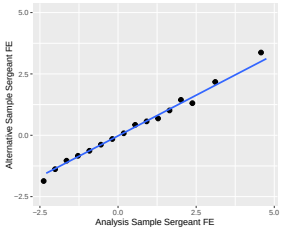
Notes: This figure depicts estimates of coefficients α_L^c and α_S^c from equation 7 for varying police outcomes c, divided by the appropriate outcome mean. The estimates are thus interpreted as the proportion change in the outcome relative to the mean that results from increasing the low-level (serious) sergeant effect by 1 standard deviation. Green lines represent α_L^c estimates and purple lines represent α_S^c estimates. Police outcomes are given on the x-axis of each figure. 95% confidence intervals are depicted for each of the estimates, calculated using standard errors clustered at the officer level. All regressions include officer fixed effects, sector-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Tables containing the regression estimates can be found in the Supplementary Appendix: Table B.4 (a), Table B.9 (b), Table B.5 (c), Table B.6 (d), and Table B.7 (e).

Supplementary Appendix for The Boss in Blue: Supervisors and Police Behavior

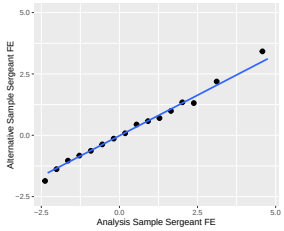
A Additional Figures

Figure A.1: Robustness to Alternative Sampling Decisions

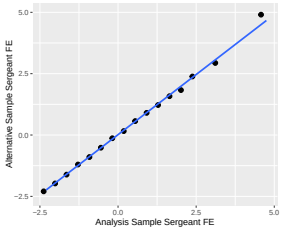
(a) Unrestricted
Correlation = 0.8720



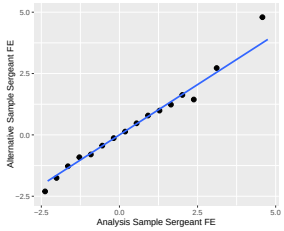
(b) Impute Missing Within Spell
Correlation = 0.8666



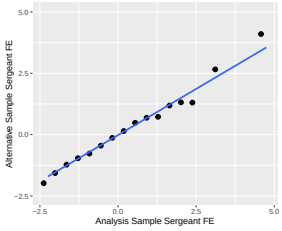
(c) Impute Missing Within Spell, remove everything else
Correlation = 0.9873



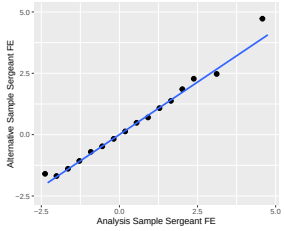
(d) Impute all temporary assignments
Correlation = 0.9362



(e) Keep temporary, remove missing
Correlation = 0.9153

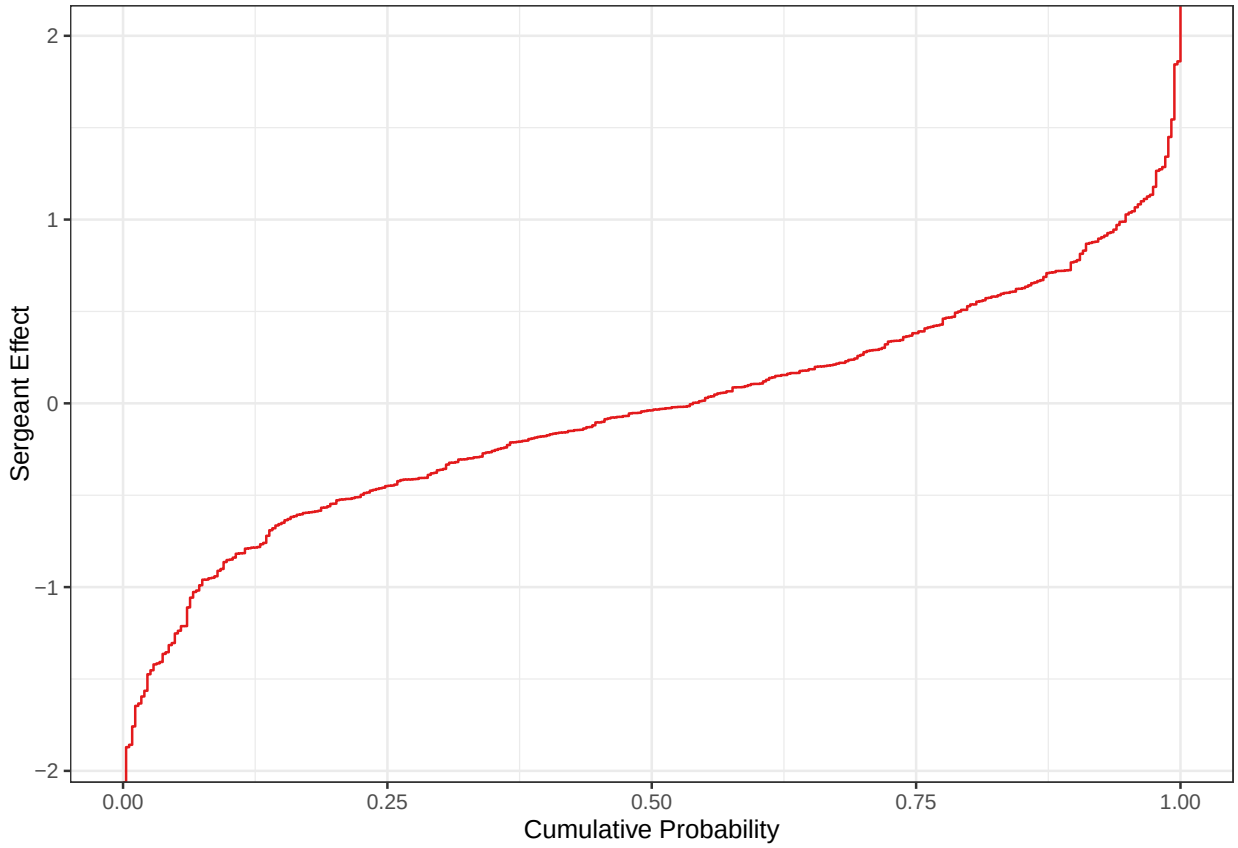


(f) Keep missing, remove temporary
Correlation = 0.9311



Notes: This figure presents the correlation between sergeant fixed effects under different sampling restrictions. (a) makes no sample restrictions, (b) imputes missing months within a sergeant spell and keeps any other observations where the sergeant is unknown, (c) is the same as (b) but other observations with unknown sergeants are removed, (d) imputes temporary one-off assignments with different sergeants using an officer's permanent sergeant, (e) keeps all of the temporary (single month) assignments but removes all months with an unknown sergeant, and (f) keeps months with an unknown sergeant but removes the temporary (single month) assignments. In the specifications that include months with an unknown sergeant, the model includes an 'Unknown Sergeant' fixed effect.

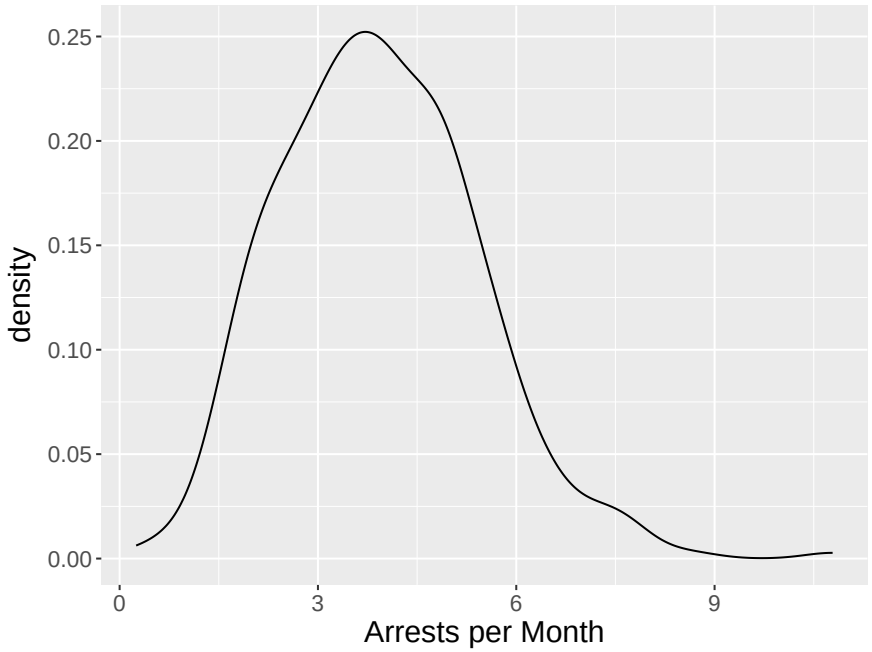
Figure A.2: CDF of Supervisor Fixed Effects



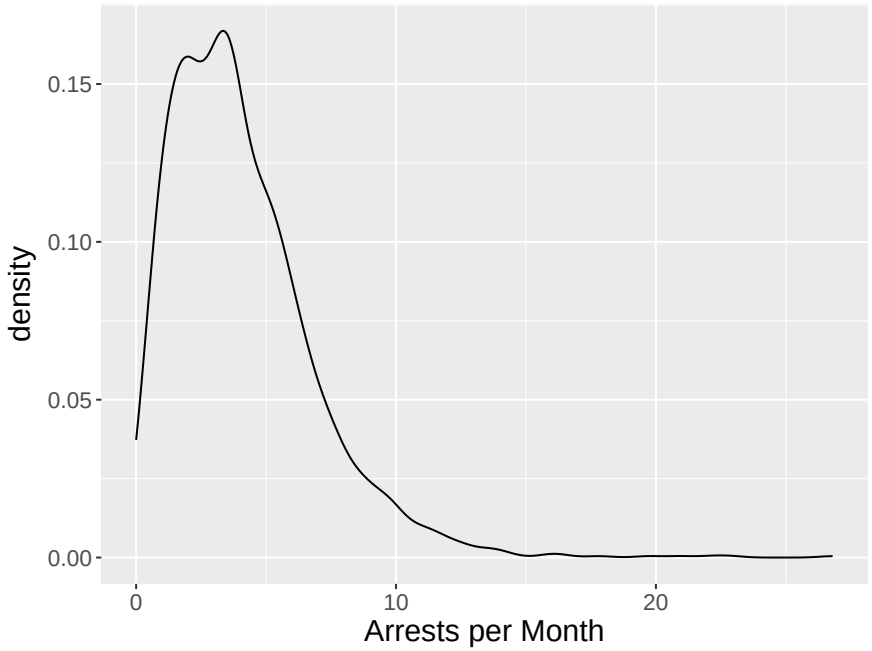
Notes: This figure displays the CDF of the sergeant effects, estimated using the sergeant fixed effects in equation 1 that are multipled by the Bayesian shrinkage factor described in Section 4.

Figure A.3: Distribution of Arrests

(a) Across Sergeants

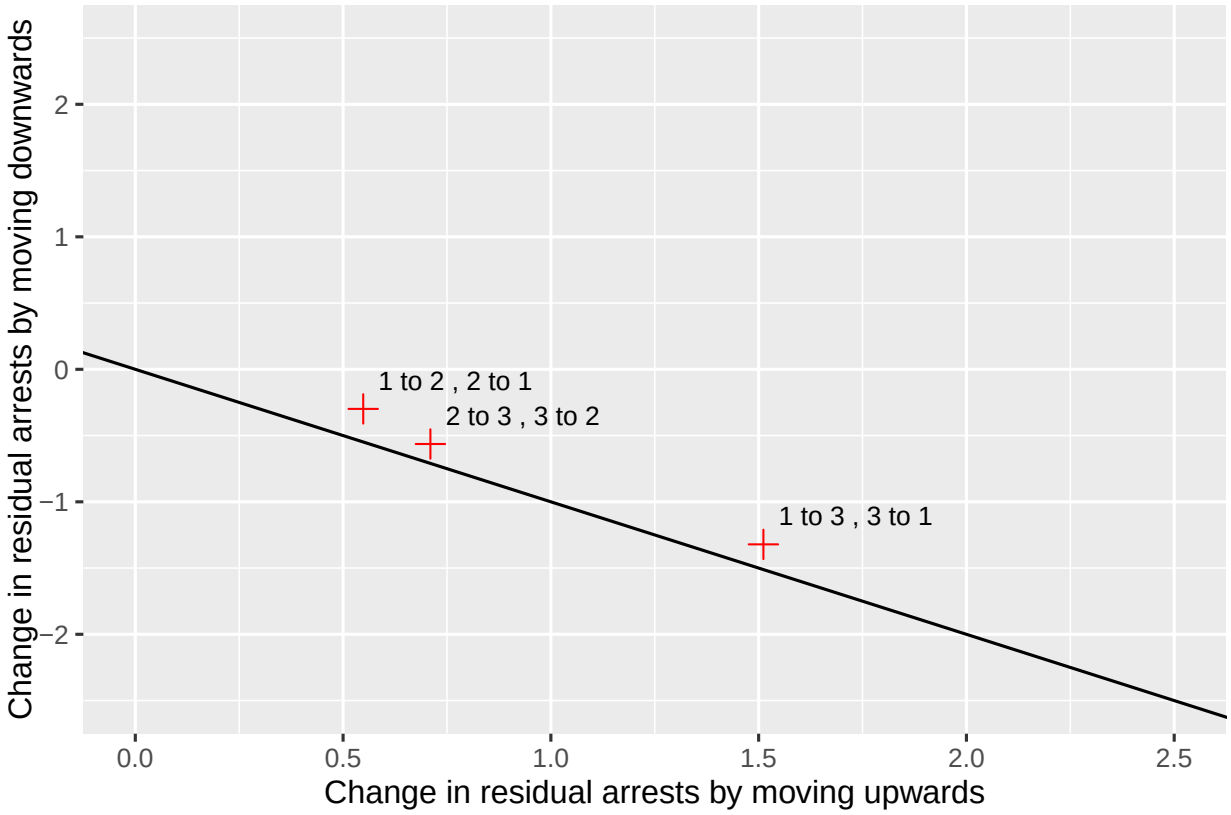


(b) Across Officers



Notes: These figures present empirical distributions for monthly arrests. For a given sergeant, I calculate the average number of monthly arrests made by officers who work for them. A.3a then plots the distribution of this average. Then, for each officer, I calculate the average number of arrests they make in a month across all months they are in the sample. I plot the distribution of this average in A.3b.

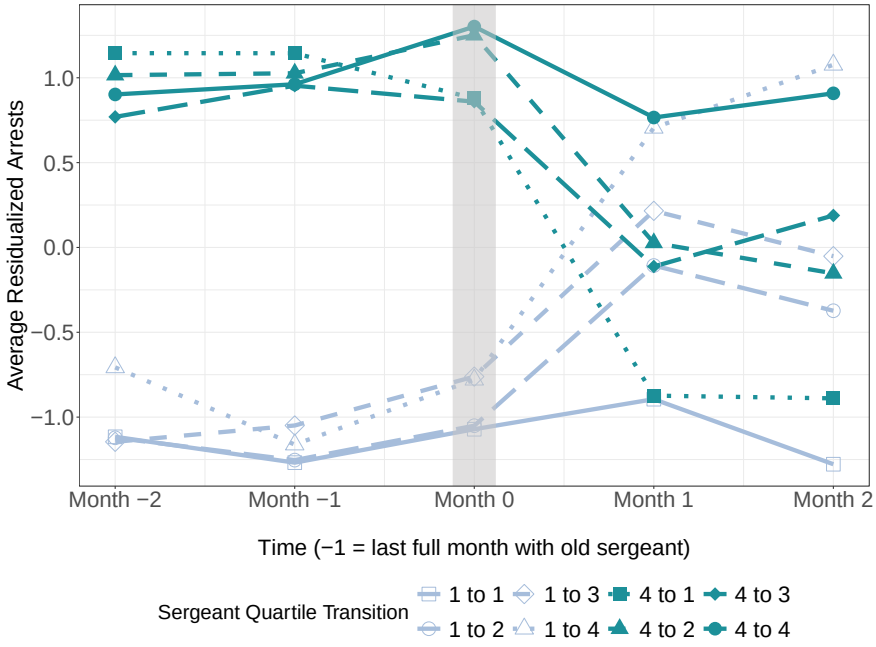
Figure A.4: Symmetry in moves



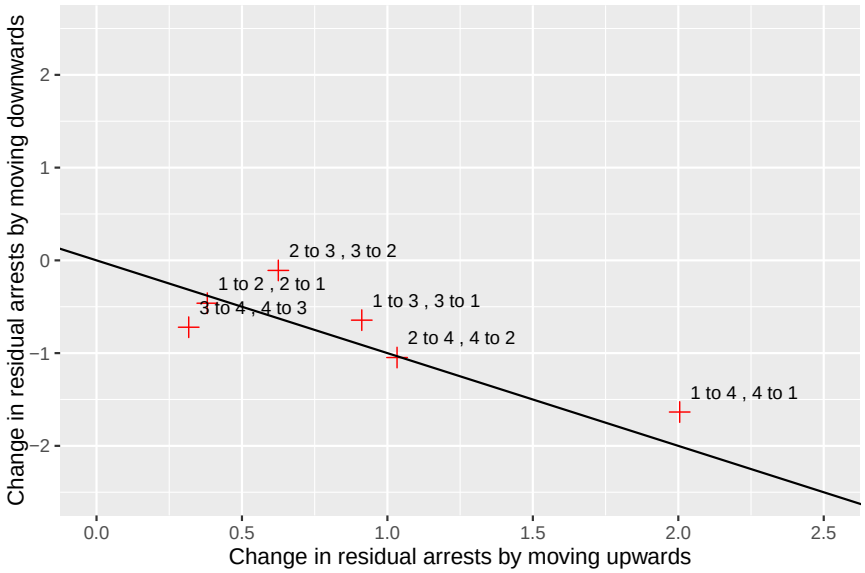
Notes: Each crosshair represents a pair of symmetric moves between sergeants in different terciles of average residualized arrests. Changes in residual arrests are calculated as the average difference between the average number of arrests 2 months after a move and 2 months before a move. Arrests are residualized by officer, sector-watch, and day-off group fixed effects and a second degree polynomial of tenure, as described in Section 4.

Figure A.5: Event study using sergeant quartiles

(a) Event Study Figure

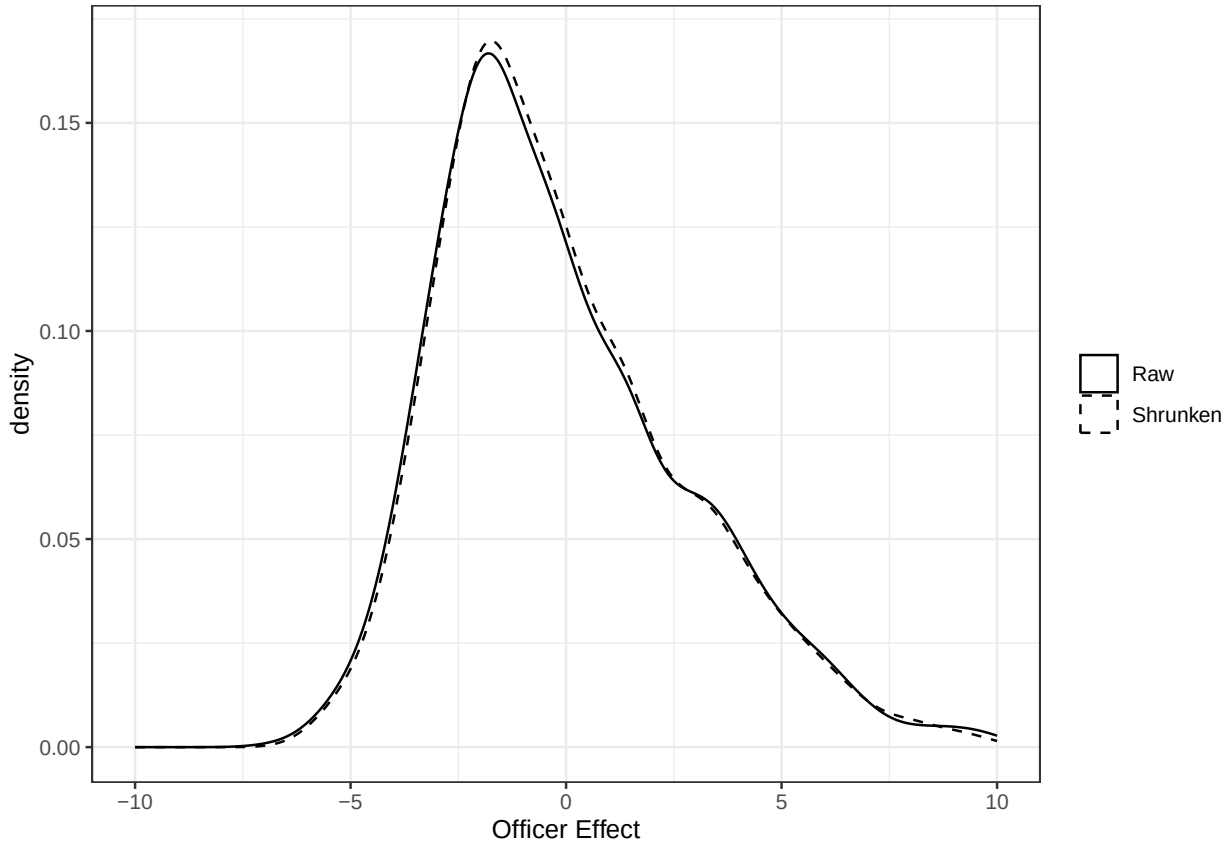


(b) Symmetry Test



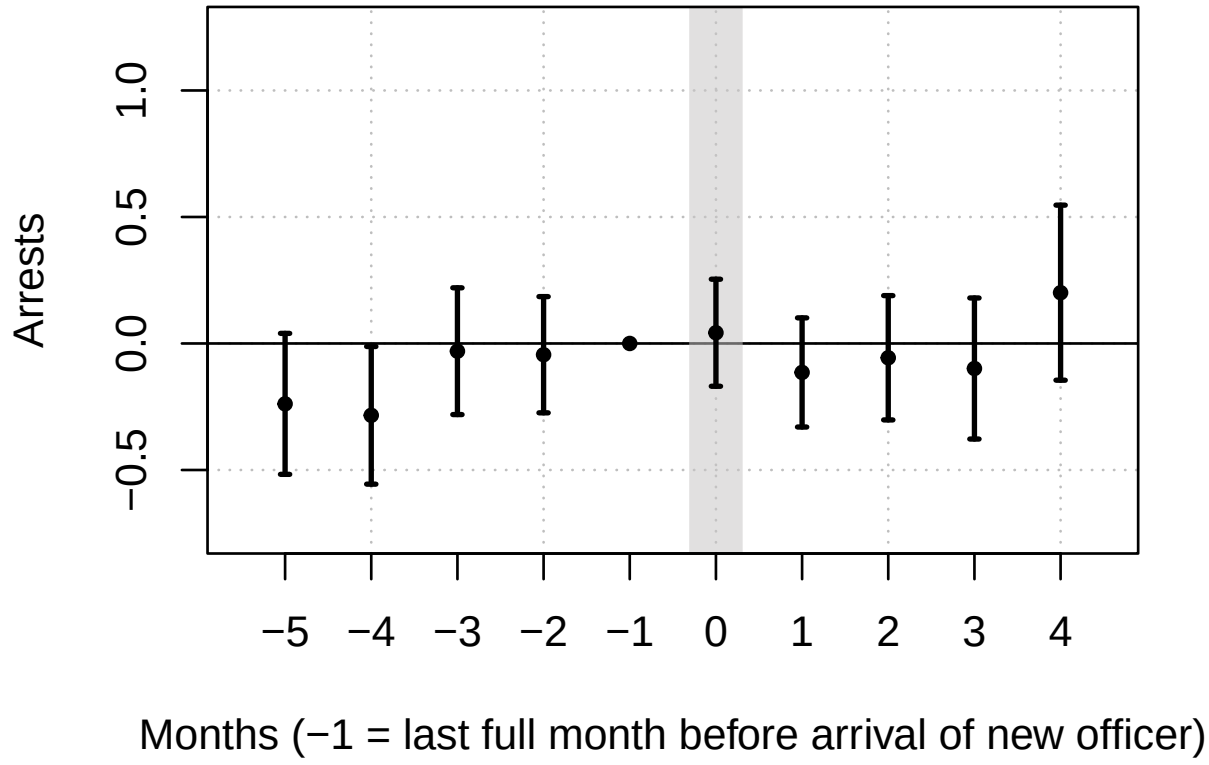
Notes: These figures present the same information as Figures 1 and A.4, instead splitting supervisors into quartiles rather than terciles. To limit the amount of lines in A.5a, I only plot transitions from supervisors in the highest and lowest quartiles. The symmetry test uses all symmetric quartile transitions. In A.5b, the crosshairs align roughly with the -45 degree line, suggesting the presence of symmetry across moves in equal and opposite directions.

Figure A.6: Distribution of Officer Effects



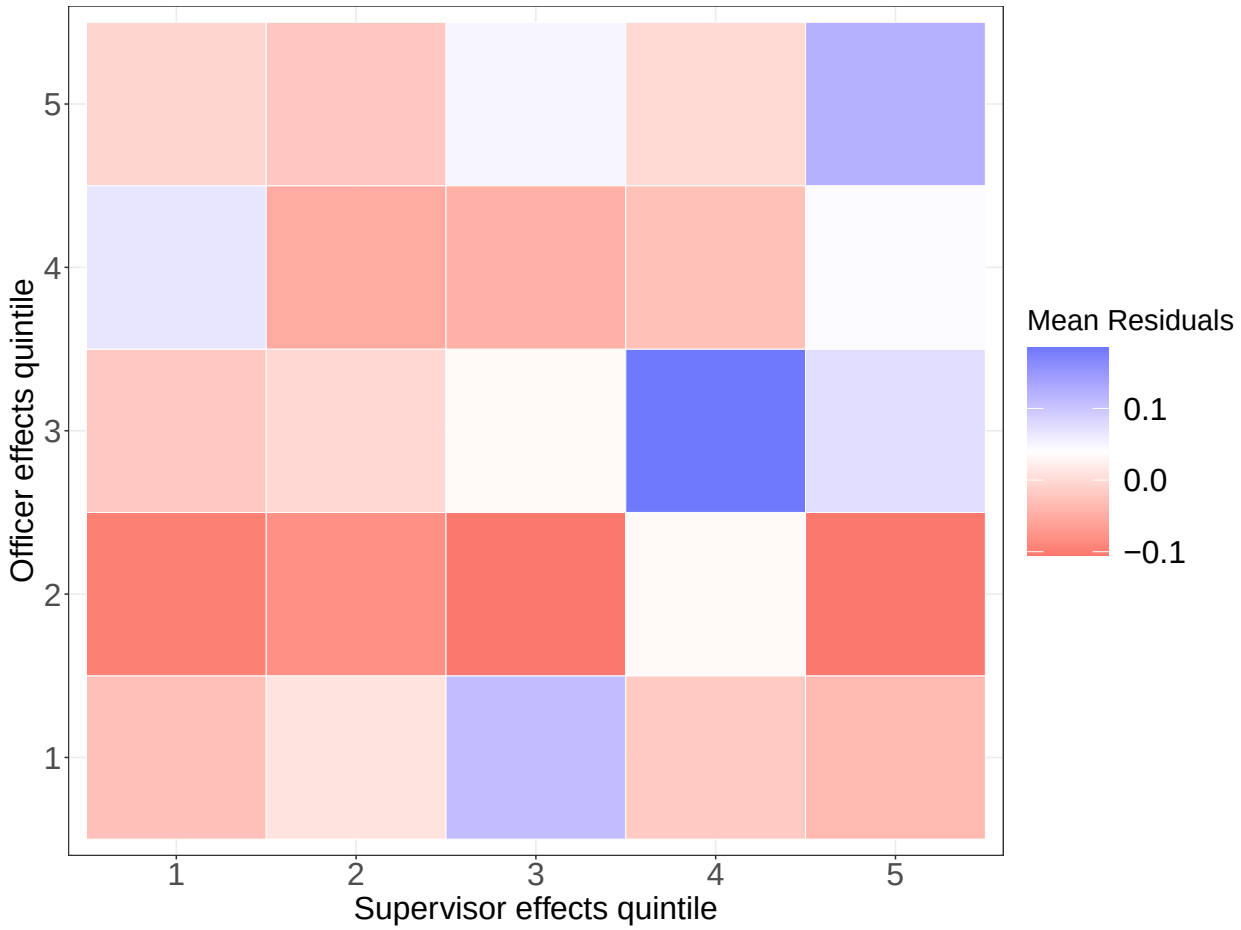
Notes: This figure plots the officer fixed effects estimated using equation 1. The solid line presents the raw fixed effects, while the dotted line presents the raw effects multiplied by the Bayesian shrinkage factor as described in Section 4.

Figure A.7: Arrests Made by Incumbent Officers



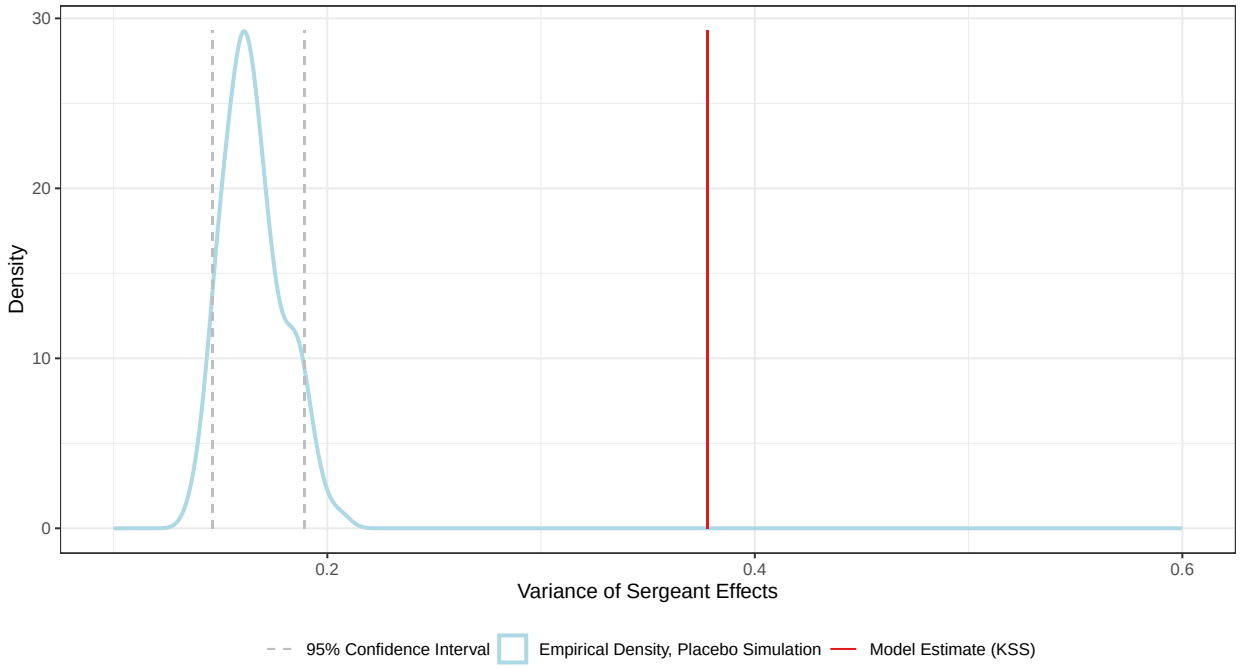
Notes: This figure plots the event study coefficients from equation 5. For an officer switching event e in which officer i switches from sergeant j to sergeant j', incumbent officers are those who work with sergeant j for 5 months before the event and and 4 months after the event. The x-axis indicates months relative to officer i's switch. Standard errors are clustered at the level of the switching officer.

Figure A.8: Residuals by quintile of officer and sergeant arrest effects



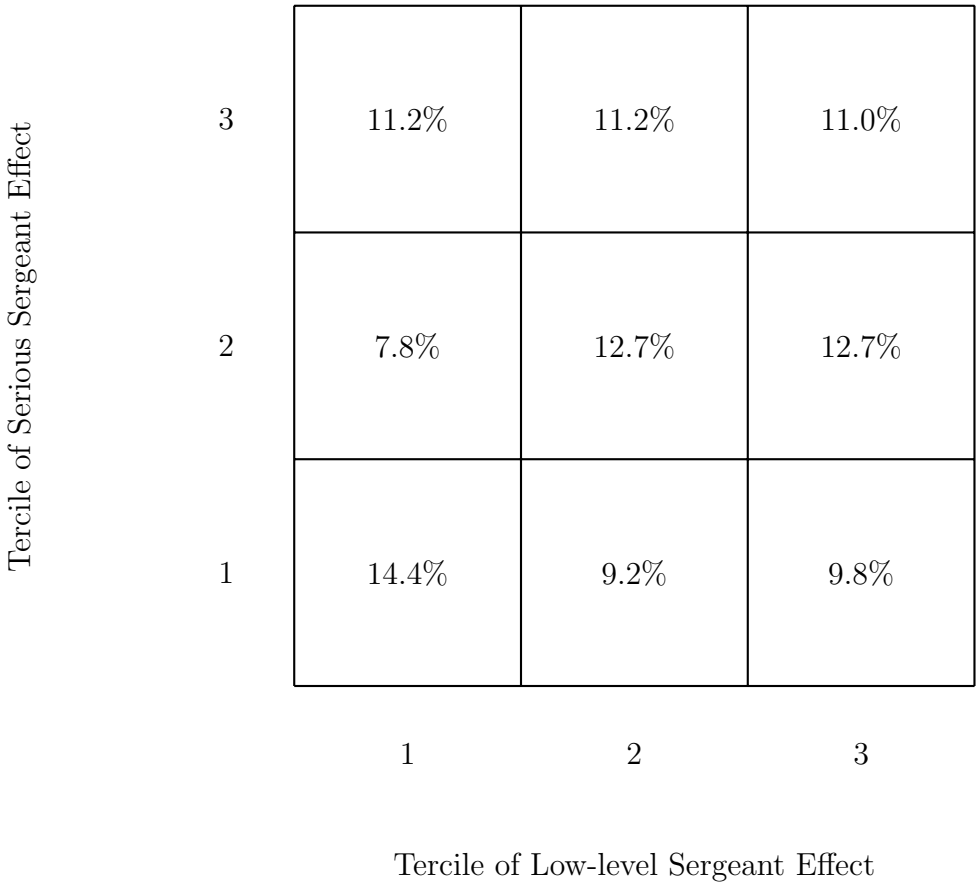
Notes: This figure reports the average residuals by quintiles of officer and sergeant (supervisor) arrest effects. Darker blue indicates more positive residuals and darker red indicates more negative residuals.

Figure A.9: Placebo Test, Sergeants randomly reassigned to officers



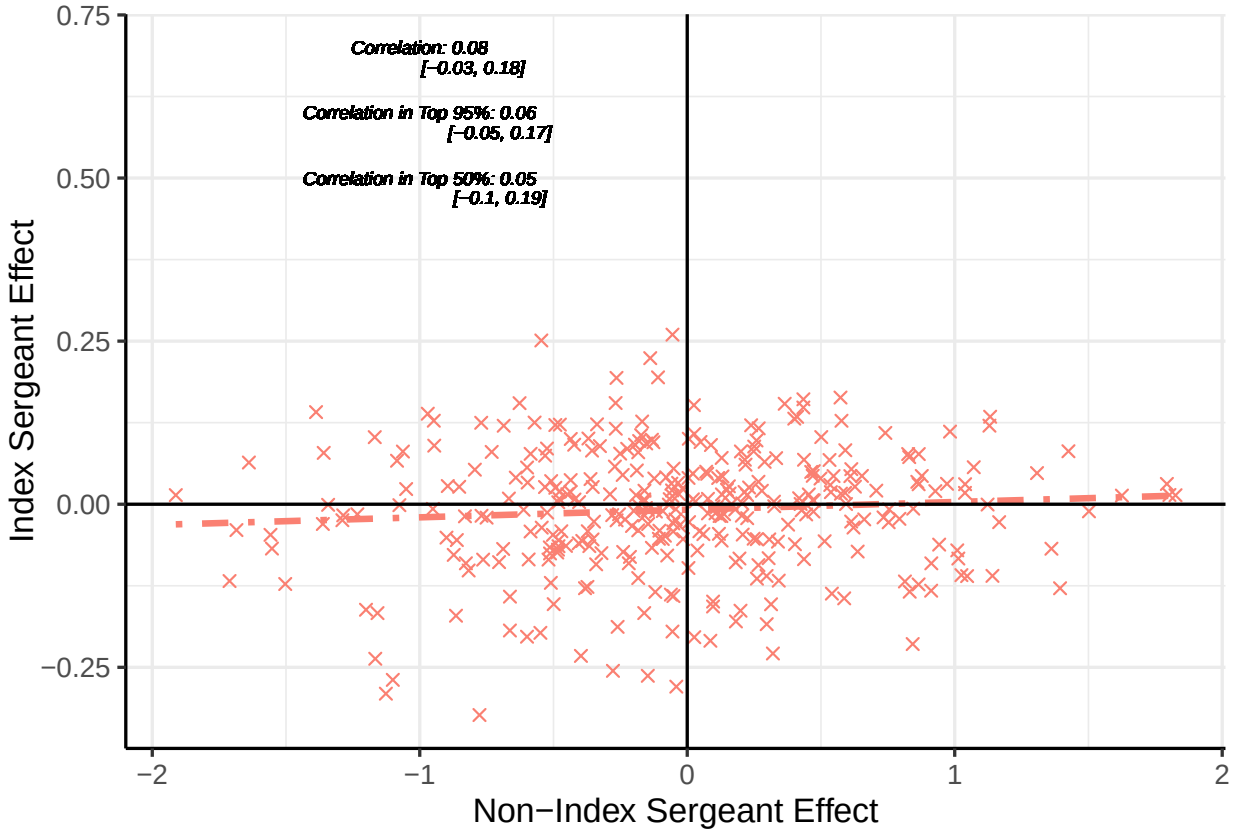
Notes: This figure depicts the results of placebo tests that randomly reallocate sergeants to officers, preserving the number of unique officers managed for each sergeant. For every reallocation, I estimate equation 1 and report the resulting (unadjusted) variance of sergeant effects. The empirical density, in light blue, plots the density of variance estimates for 100 reallocations. The dashed lines denote the 95% confidence interval of the placebo variance estimates. The red vertical line denotes my main estimate of the variance in sergeant effects, adjusted for bias using the KSS method described in Section 4.

Figure A.10: Distribution of Sergeants between Serious and Low-level Effects



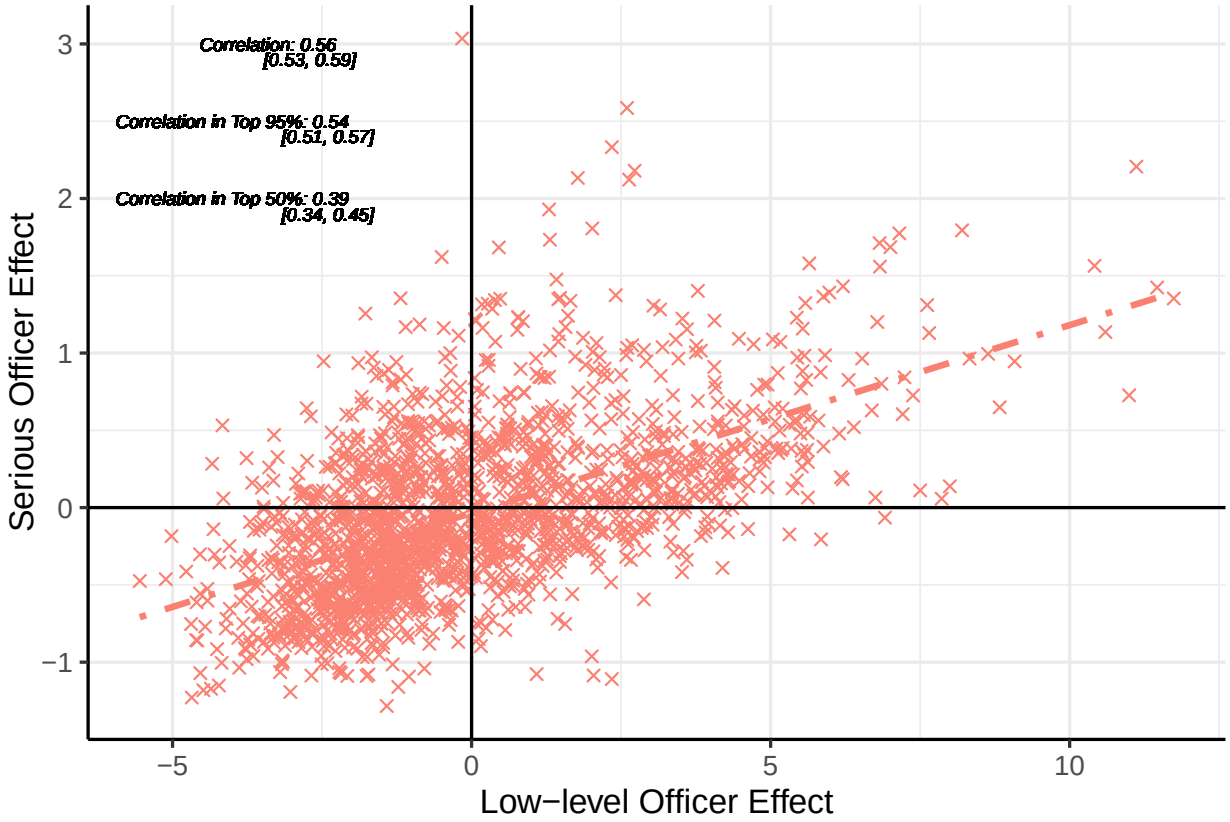
Notes: This figure displays the percentage of sergeants within each tercile grouping of (Bayes-shrunken) low-level and serious sergeant effects.

Figure A.11: Relationship between non-index and index sergeant effects



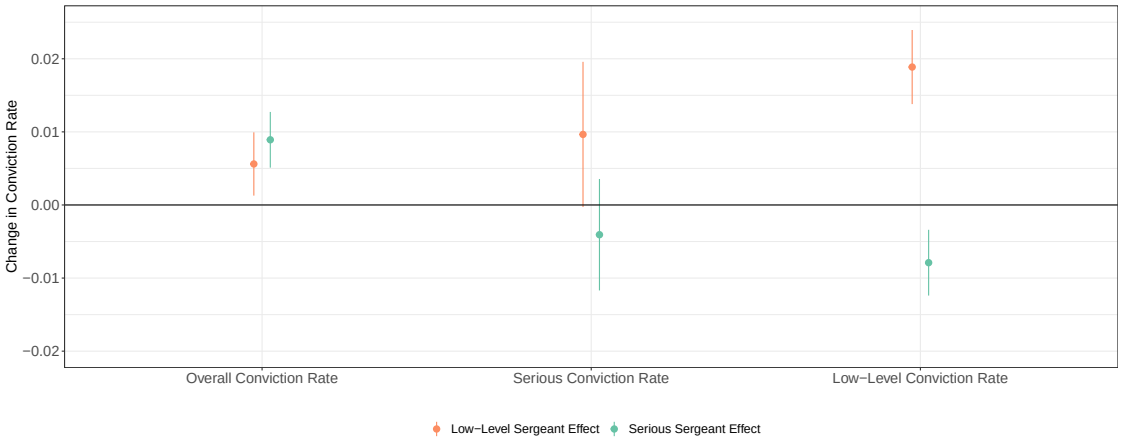
Notes: This figure displays a scatterplot of the relationship between index sergeant effects and non-index sergeant effects, each representing a sergeant's propensity to induce arrests for crimes of either type, estimated using equation the two-way fixed effects model in equation 1. These measures are used as a robustness check on serious and low-level arrest classifications used to generate Figure 5. Each point represents a sergeant. Index crimes encompass murder, rape, aggravated assault, robbery, burglary, theft, and arson. Both types of sergeant effect are shrunken using the Empirical Bayes procedure described in Section 4. Pearson correlations are provided in black text, with 95% confidence intervals for the correlations, calculated using the Fisher z-transformation, displayed below each correlation. Correlation in the Top 95% is calculated by dropping sergeants below the 5th percentile of non-index effects. Correlation in the Top 50% is calculated by dropping sergeants below the 50th percentile of non-index effects. A best fit-line is given by the dashed red line.

Figure A.12: Relationship Between Serious and Low-Level Officer Effects



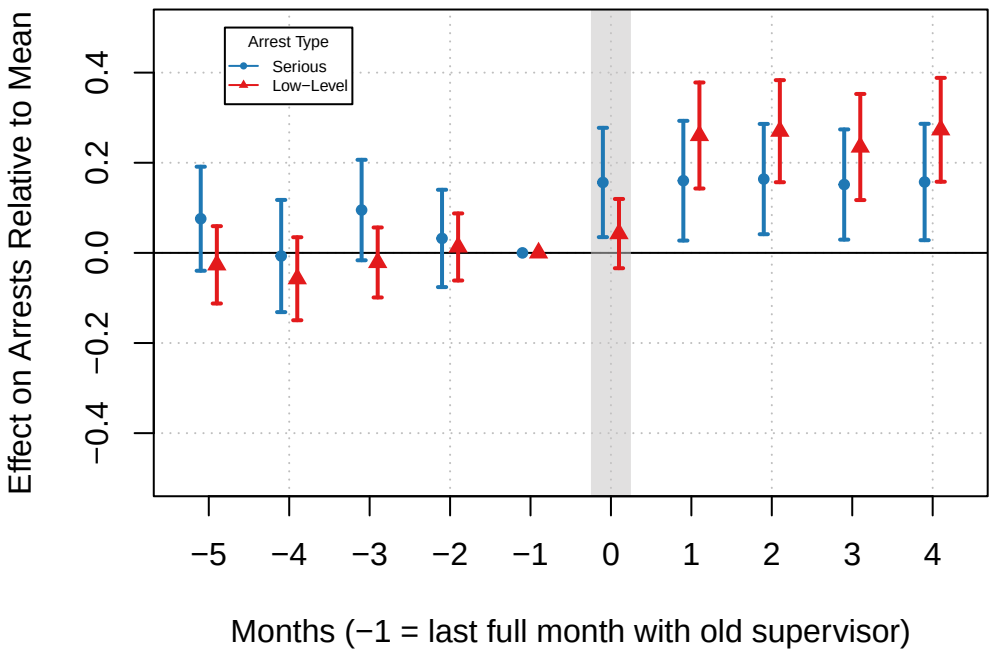
Notes: This figure displays a scatterplot of the relationship between the low-level officer effects and serious officer effects. Each point represents an officer. Low-level (serious) effects describe officer effect on arrests for low-level (serious) crimes, defined as described in Section 3. Both types of officer effect are shrunken using the Empirical Bayes procedure described in Section 4. Pearson correlations are provided in black text, with 95% confidence intervals for the correlations, calculated using the Fisher z-transformation, displayed below each correlation. Correlation in the Top 95% is calculated by dropping officers below the 5th percentile of low-level effects. Correlation in the Top 50% is calculated by dropping officers below the 50th percentile of low-level effects. A best fit-line is given by the dashed red line.

Figure A.13: Impact of Sergeant Effects on Conviction Rates



Notes: This figure plots the change in officer conviction rates that results from increasing the low-level and serious sergeant effects by one standard deviation. I calculate the change in conviction rate in two steps, as described in Section 5. I first regress the number of total and convicted arrests separately on the standardized low-level and serious sergeant effects, along with the model controls as in equation 7. Then, for each of the two sergeant effects, I add the regression coefficients for convicted and total arrests to their respective sample means and take the difference between this ratio and the ratio of the means. For each estimate, I calculate a 95% confidence interval using a bootstrap with 100 resamples. I do this procedure separately for all arrests ('Overall Conviction Rate), serious arrests ('Serious Conviction Rate), and low-level arrests ('Low-Level Conviction Rate').

Figure A.14: Event Study by Arrest Type



Notes: This figure plots event study coefficients for equation 4, separately for models that use serious arrests (defined as index arrests as well as domestic violence, fraud, simple assault, and DUI) and low-level arrests as the dependent variable. Serious arrest results are given in blue and low-level arrest results are given in red. Month -1, the last full month that the officer spends with the old sergeant, is the reference month in all specifications. For each severity level, the effects are divided by the average number of the corresponding category of arrests in month -1. The model is estimated using the event study data that are balanced on [-5, 4].

B Additional Tables

Table B.1: Event-study around sergeant switches

k	Total Arrests	Serious Arrests	Low-Level Arrests
	(1)	(2)	(3)
-5	0.0020	0.0732	-0.0712
	(0.1366)	(0.0568)	(0.1175)
-4	-0.1609	-0.0067	-0.1542
	(0.1477)	(0.0613)	(0.1261)
-3	0.0347	0.0919*	-0.0572
	(0.1265)	(0.0549)	(0.1064)
-2	0.0664	0.0310	0.0354
	(0.1191)	(0.0532)	(0.1020)
0	0.2661**	0.1510**	0.1151
	(0.1252)	(0.0597)	(0.1052)
1	0.8550***	0.1548**	0.7002***
	(0.1779)	(0.0654)	(0.1611)
2	0.8843***	0.1583***	0.7260***
	(0.1649)	(0.0602)	(0.1551)
3	0.7782***	0.1466**	0.6316***
	(0.1780)	(0.0602)	(0.1612)
4	0.8862***	0.1520**	0.7342***
	(0.1735)	(0.0636)	(0.1578)
Observations	12,770	12,770	12,770
Y mean	3.6488	0.89742	2.7514
Pre-trends F stat	0.8473	1.5706	0.6234
p-value	0.8473	0.1791	0.6458

Notes: This table presents the event-study coefficients used to make Figures 4 and A.14. The regressions use switching events that are balanced around 5 sample months prior to the move and 4 sample months after the move. Month -1 is the reference point and the switch occurs at in month 0. The pre-trends F statistic is calculated from an F-test of joint significance of the coefficients for which $k < -1$. Standard errors are clustered at the officer level.

Table B.2: Trends in crime do not predict changes in sergeant effects

	All Months		Months With Movers	
	$E[\Delta\hat{\psi}_{out}]$	$E[\Delta\hat{\psi}_{in}]$	$E[\Delta\hat{\psi}_{out}]$	$E[\Delta\hat{\psi}_{in}]$
	(1)	(2)	(3)	(4)
$Log(911Calls)_{-1}$	-0.0162	0.0014	0.0713	0.1357
	(0.0467)	(0.0483)	(0.1654)	(0.1482)
$Log(911Calls)_{-2}$	-0.0043	0.0173	-0.0571	0.0140
	(0.0459)	(0.0570)	(0.1871)	(0.1764)
$Log(911Calls)_{-3}$	0.0063	-0.0594	-0.0722	-0.3354*
	(0.0552)	(0.0479)	(0.1709)	(0.1770)
$Log(911Calls)_{-4}$	-0.0398	0.0148	-0.0915	0.0855
	(0.0515)	(0.0543)	(0.1877)	(0.1545)
$Log(911Calls)_{-5}$	0.0372	-0.0460	0.2426	-0.0737
	(0.0391)	(0.0473)	(0.1723)	(0.1704)
Observations	5,525	5,525	1,387	1,424
Y mean	0.00738	0.00401	0.02938	0.01555
Joint F p-value	0.91938	0.62787	0.79073	0.38557
Sector-Watch fixed effects	✓	✓	✓	✓

Notes: This table examines the correlation of crime trends with sergeant switches. Regressions are performed at the unit by month level. For each unit and month, I calculate the average change in sergeant effects separately for officers who move in and out of the unit in that month. The table reports results for a regression of the sergeant effect changes on the natural logarithm of 911 calls originating in the unit's sector-watch in each of the 5 months prior to the focal month. Regressions include fixed effects for the sector-watch. Dependent variables are the average change in the sergeant effects for out-movers (columns 1 and 3) and in-movers (columns 2 and 4). Out-movers are officers who leave the unit and in-movers are officers who join the unit in a given month. Columns 1 and 2 use all monthly observations for each unit. Columns 3 and 4 only use the monthly observations in which at least one out-move (column 3) or one in-move (column 4) occurs. Standard errors are clustered at the sector-watch level.

Table B.3: Analysis of Variance

	Arrests				
	(1)	(2)	(3)	(4)	(5)
R^2	0.166379	0.516861	0.527611	0.202556	0.623947
Adjusted R^2	0.164421	0.497526	0.505139	0.195091	0.559736
Controls	✓	✓	✓	✓	✓
Officer FE		✓	✓		
Sergeant FE			✓	✓	
Sergeant-by-Officer FE					✓
Observations	49,923	49,923	49,923	49,923	49,923

Notes: This table reports R^2 and adjusted R^2 for models that vary the included fixed effects. Controls include a second degree polynomial of officer tenure, and sector-watch and day-off group fixed effects.

Table B.4: Sergeant Effects by Crime Type

	Serious Crimes			Low-Level Crimes		
	Domestic Violence	Theft	DWI	Drugs	Warrants	Disorderly Conduct
	(1)	(2)	(3)	(4)	(5)	(6)
Low-level Sergeant Effect	-0.0161**	-0.0025	-0.0013	0.1713***	0.1772***	0.1076***
	(0.0080)	(0.0053)	(0.0061)	(0.0238)	(0.0203)	(0.0134)
Serious Sergeant Effect	0.1687***	0.0382***	0.0352***	-0.0360***	0.0755***	0.0350***
	(0.0082)	(0.0064)	(0.0063)	(0.0139)	(0.0148)	(0.0093)
Baseline Controls	✓	✓	✓	✓	✓	✓
Observations	49,923	49,923	49,923	49,923	49,923	49,923
Y mean	0.45308	0.13927	0.10885	0.31138	0.76578	0.40903

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7, with arrests for the three most frequent serious and low-level crimes as the outcome variables. Serious and low-level crimes are mutually exclusive categories. However, within serious and low-level crimes, an arrest may fall under multiple different criminal charges. The low-level (serious) sergeant effect is given by the standardized Bayes-shrunken sergeant effect on low-level (serious) arrests. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.5: Sergeant Effects on Arrests by Interaction Source

	Officer Initiated Arrests	Call Initiated Arrests
	(1)	(2)
Low-level Sergeant Effect	0.4451***	0.2561***
	(0.0367)	(0.0237)
Serious Sergeant Effect	0.0339	0.2373***
	(0.0244)	(0.0196)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	1.6951	2.0930

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7 using officer-initiated and call-initiated arrests as outcomes. An arrest is call-initiated if it can be linked to a 911 call in CAD. Otherwise, it is considered officer-initiated. The low-level (serious) sergeant effect is given by the standardized Bayes-shrunken sergeant effect on low-level (serious) arrests. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.6: Sergeant Effects on Call Activity

	Calls Answered	Arrest Probability
	(1)	(2)
Low-level Sergeant Effect	1.243***	0.0042***
	(0.4156)	(0.0005)
Serious Sergeant Effect	2.301***	0.0004
	(0.3562)	(0.0004)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	73.542	0.06244

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7 using various measures of call activity as outcome variables. In column 1, calls answered refer to calls in which the officer is recorded as being at the scene in CAD. In column 2, arrest probability at calls is the proportion of calls that an officer answers that result in arrest. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.7: Sergeant Effects and Other Activity

	Use of Force Incidents	Complaints
	(1)	(2)
Low-level Sergeant Effect	0.0203***	0.0040
	(0.0043)	(0.0029)
Serious Sergeant Effect	0.0059*	-0.0019
	(0.0034)	(0.0025)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	0.13813	0.02398

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7 using use of force incidents and complaints as outcome variables. The low-level (serious) sergeant effect is given by the standardized Bayes-shrunken sergeant effect on low-level (serious) arrests. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.8: Drug Arrest Types

	Possession Arrests	Distribution Arrests
	(1)	(2)
Low-level Sergeant Effect	0.1543***	0.0163***
	(0.0210)	(0.0041)
Serious Sergeant Effect	-0.0314**	-0.0051***
	(0.0132)	(0.0018)
Baseline Controls	✓	✓
Observations	49,923	49,923
Y mean	0.28955	0.01839

Notes: The table reports results for a regression of possession (column 1) and distribution (column 2) arrests on low-level and serious sergeant effects, along with baseline model controls from equation 7, which include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. I classify a drug arrest as "possession" if the charge description only mentions possession and not sale or manufacturing. If the charge description mentions sale or manufacturing, then the drug arrest is classified as "distribution," so that the two categories are mutually exclusive. The effect sizes for all drug arrests are 0.1713 for low-level sergeant effects and -0.0361 for serious sergeant effects, taken from Table B.4. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table B.9: Sergeant impacts by race

	Black Arrests	Hispanic Arrests	White Arrests
	(1)	(2)	(3)
Low-level Sergeant Effect	0.4172***	0.1469***	0.1370***
	(0.0333)	(0.0157)	(0.0127)
Serious Sergeant Effect	0.1150***	0.0966***	0.0519***
	(0.0231)	(0.0154)	(0.0100)
Observations	49,923	49,923	49,923
Y mean	1.9291	1.0082	0.80869

Notes: This table reports the estimated coefficients for the sergeant effects in equation 7, using arrests by race as the outcome variables. The baseline controls include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure.. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.10: Sergeant Effects and Conviction Rates

	Difference in Conviction Ratio	Difference in Serious Conviction Ratio	Difference in Low-Level Conviction Ratio
	(1)	(2)	(3)
Low-level Sergeant Effect	0.0056**	0.0097*	0.0189***
	(0.0022)	(0.0051)	(0.0138)
Serious Sergeant Effect	0.0089***	-0.0040	-0.0079***
	(0.0019)	(0.0039)	(0.0230)
Baseline Controls	✓	✓	✓
Observations	49,923	49,923	49,923
Mean ratio	0.2046	0.4282	0.1322

Notes: This table reports the changes in officer conviction rates that results from increasing the low-level and serious sergeant effects by one standard deviation, which are plotted in Figure A.13. I calculate the change in conviction rate in two steps. I first regress the number of total and convicted arrests separately on the standardized low-level and serious sergeant effects, along with the baseline controls, which include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure. Then, for each of the two sergeant effects, I add the regression coefficients for convicted and total arrests to their respective sample means and take the difference between this ratio and the ratio of the means. For each estimate, I calculate a 95% confidence interval using a bootstrap with 100 resamples. I do this procedure separately for all arrests ('Overall Conviction Rate), serious arrests ('Serious Conviction Rate), and low-level arrests ('Low-Level Conviction Rate'). $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table B.11: Sergeant effects and officer overtime activities

	Overtime Calls	Overtime Low-level Arrests	Overtime Serious Arrests
	(1)	(2)	(3)
Low-level Sergeant Effect	0.4808***	0.0312***	-0.0011
	(0.1386)	(0.0057)	(0.0031)
Serious Sergeant Effect	0.4982***	0.0103*	0.0219***
	(0.1238)	(0.0056)	(0.0037)
Baseline Controls	✓	✓	✓
Observations	49,923	49,923	49,923
Y mean	3.7242	0.11327	0.04926

Notes: This table presents results for a regression of overtime calls and arrests on low-level and serious sergeant effects, along with standard model controls from equation 7, which include officer fixed effects, sector-by-watch fixed effects, day-off group fixed effects, and a second-degree polynomial of officer tenure.. I call a call or arrest overtime if it occurs outside of the officer's shift hours listed in the assignments data. Standard errors are clustered at the officer level. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

C Constructing Sergeant Assignments

The Computer Aided Dispatch (CAD) system used by the Dallas Police Department stores assignment indicators for every sworn employee who is assigned to a call. These assignment indicators are known internally as "element numbers." Element numbers are assigned every day to each separate patrol car and describe the watch and beat assignment of the car. Beats are smaller geographic sectors within patrol sectors that individual officers are assigned to patrol.

Watches are described by letters A-F, where A/B/C denote overnight/day/evening watches and D/E/F are variants for day/overnight/evening that allow for multiple units to be assigned to one beat at a time depending on department needs. Beats are given by a 3-digit numeric. Thus, an example of an element number within CAD is A135, which means that the officer is working the overnight shift patrolling beat 135. The first and second digits of the beat code identify the sector in which the beat is located. Returning to the previous example, beat 135 is part of sector 130.

Sergeants are given element numbers that denote the sector and watch to which they are assigned. The sergeant for an officer with the element number A135 has the element number A130. In the case of variant units within a sector, there will be one sergeant in charge of each unit. That is, an officer in the variant overnight unit E135 would have a sergeant with the element number E130. I use this pattern in the element numbers to identify the most common sector-watch assignments for officers and sergeants within each month of the data, as described in Section 3.

The assignments that I construct exclude officer spells in specialty patrol units whose element number does not match a geographic sector. Based on conversations with DPD, these units perform distinct duties from regular patrol officers, as evidenced by the fact that they are not assigned to specific geographic beats. DPD did not provide me with the specific details of the job duties related to these assignments for reasons related to officer safety. But, consistent with a separate and distinct role, officers in these specialty units exhibit significant more variation in arrests than regular patrol officers and sergeants cannot be reliably identified for these units.

D Empirical Bayes Shrinkage

The raw sergeant fixed effects are estimated with error. Suppose that the estimates are given by:

$$\hat{\psi}_j = \psi_j + \epsilon_j, \tag{9}$$

where $\psi_j \sim \mathcal{N}(0, \sigma^2_\psi)$, $\epsilon_j \sim \mathcal{N}(0, \sigma^2_{\epsilon_j})$, and ψ_j and ϵ_j are independently distributed across the population of 347 sergeants. The mean of the sergeant fixed effects is 0 by construction, since the true mean is unidentified in the model. Under these distributional assumptions, we have that

$$\hat{\psi}_j | \psi_j \sim \mathcal{N}(\psi_j, \sigma^2_\epsilon). \tag{10}$$

Hence, each of the fixed effects are unbiased estimates of supervisor j's effect, as is the case under the identifying assumptions laid out in Section 4. As shown by Morris (1983), one can construct a more efficient estimator of ψ_j using the posterior mean of ψ_j conditional on the estimate $\hat{\psi}_j$:

$$E[\psi_j | \hat{\psi}_j] = \lambda_j \hat{\psi}_j, \tag{11}$$

where $\lambda_j = \frac{\sigma^2_\psi}{\sigma^2_\psi + \sigma^2_{\epsilon_j}}$. I estimate the shrinkage factor $\hat{\lambda}_j$ by bootstrapping the estimation of equation 1. For each supervisor j, I obtain bootstrap estimates of the fixed effect $\hat{\psi}^k_j$, where k= 1,...,1000. I estimate the error variance of each $\hat{\psi}_j$ using the sample variance of the bootstrap distribution: $\hat{\sigma}^2_{\epsilon_j} = \frac{1}{k-1} \sum_{k=1}^{1000} (\hat{\psi}^k_j - \bar{\hat{\psi}}^k_j)^2$. I then estimate $\hat{\sigma}^2_\psi$ using the variance estimator proposed by Morris (1983):

$$\hat{\sigma}^2_\psi = \frac{\sum W_j (\hat{\psi}^2_j - \hat{\sigma}^2_j)}{\sum W_j}. \tag{12}$$

For my main estimates, I use weights $W_j = 1$, so that the estimate takes the form:

$$\hat{\sigma}^2_\psi = Var(\hat{\psi}_j) - E_j(\hat{\sigma}^2_j). \tag{13}$$

One can also use the weights proposed by Morris (1983): $\frac{1}{\hat{\psi}^2_j + \hat{\sigma}^2_j}$, which requires estimating $\hat{\sigma}^2_\psi$ iteratively by first plugging in a guess of the across-supervisor variance and calculating as in equation 13 until the values are sufficiently close. In unreported results, I find that using this weighted estimate yields a similar shrinkage factor.

It is also possible to estimate the variance components directly from the regression resid-

uals using a method proposed by Guarino et al. (2015) and implemented in the policing context by Weisburst (2024). In unreported results, I find that this method produces similar, albeit slightly less conservative shrinkage factors.

E Predicting Sergeant Effects

To what extent are sergeant effects mediated by observable characteristics that are determined *before* someone has been promoted to sergeant? I examine variations in sergeant effect distributions across four observable pre-promotion characteristics: race, gender, age at the time of their promotion exam, and the score achieved on the promotion exam.[27] Since I only observe exam scores beginning with the 2012 round of tests, I limit my sample to the 202 sergeants who were promoted from these exams. I call sergeants "older" if they were above the average across all exams and I call them "above-average scorers" if they were above the average score for their particular exam.

In Figure E.1a, I present densities separately by exam score. I observe a notable difference between high and low scorers in the distributions of both serious and low-level effects. For low-level effects, the distribution of above-average scorers is shifted to the left of that for below-average scorers, with a Kolmogorov-Smirnov test indicating that this difference is statistically significant (p-value = 0.034). Thus, individuals who score below average on the promotional exams tend to induce more low-level arrests than those who score above average. These differences are meaningful considering exam score is the primary determinant of promotion. The exams test for knowledge of department procedures and aptitude within relevant supervisory situations. Thus marginal promotees — who are barely promoted by virtue of their low exam score — are more likely to value low-level arrests.

In Figure E.1b, I also find evidence of a statistically significant difference in the distribution of serious effects between high and low exam scorers (Kolmogorv-Smirnov p-value = 0.013). In contrast to low-level effects, the differences in serious effects are not driven by a monotonic shift in one direction. Instead, high scorers exhibit a wider distribution of serious effects than low scorers. The distribution for low scorers is more concentrated around 0. These findings suggest that high scorers are significantly more heterogeneous than low-scorers, at least in terms of their effects on serious enforcement.

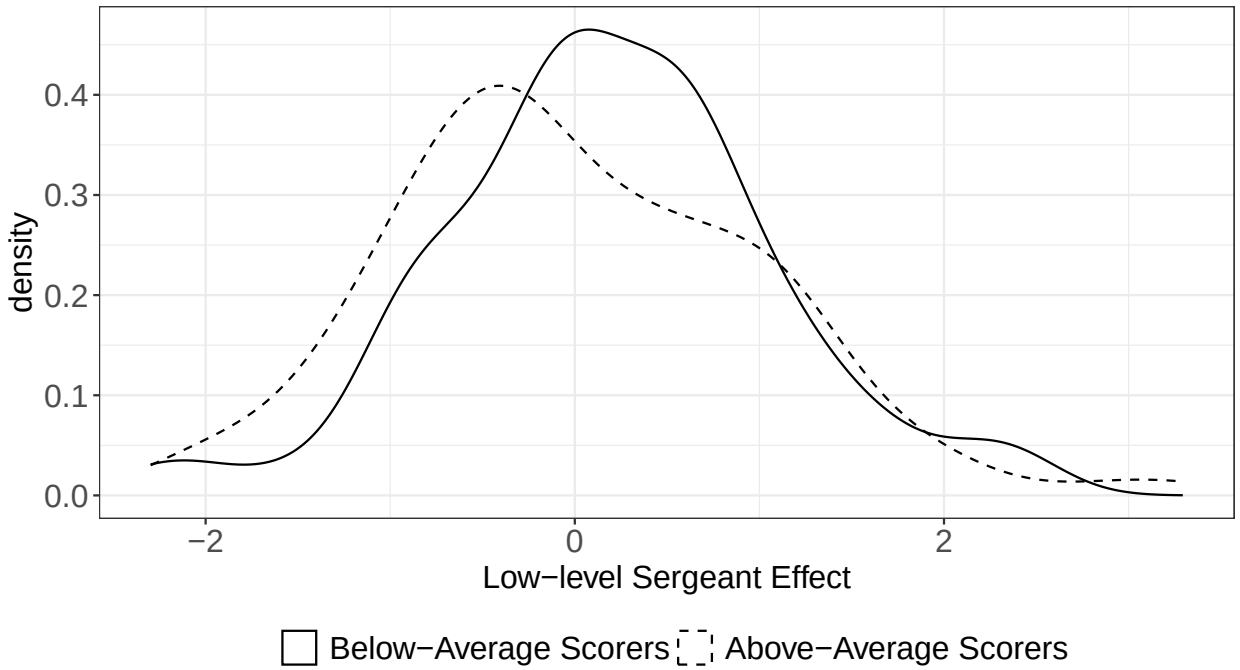
I also analyze the empirical densities of low-level and serious effects separately by race (Figures E.3b, E.3a), gender (Figures E.2b, E.2a), and age groups (Figures E.4b, E.4a). I do not find evidence of significant differences in the distribution of either sergeant effect along each of these three observable dimensions. Kolmogorov-Smirnov tests for the equality of the distributions generate p-values that are well-above standard significance thresholds.

[27]The promotion exam is given once every few years to officers with at least 5 years of experience who want to take part in the promotion process. Exam-takers are ranked according to their performance and, with a few rare exceptions, promoted in order of their exam ranking when openings arise. The exam has both a written and oral component. For this study, I obtained data on the aggregate exam scores for all sergeants promoted during the sample period, accounting for 58% of sergeants in my sample.

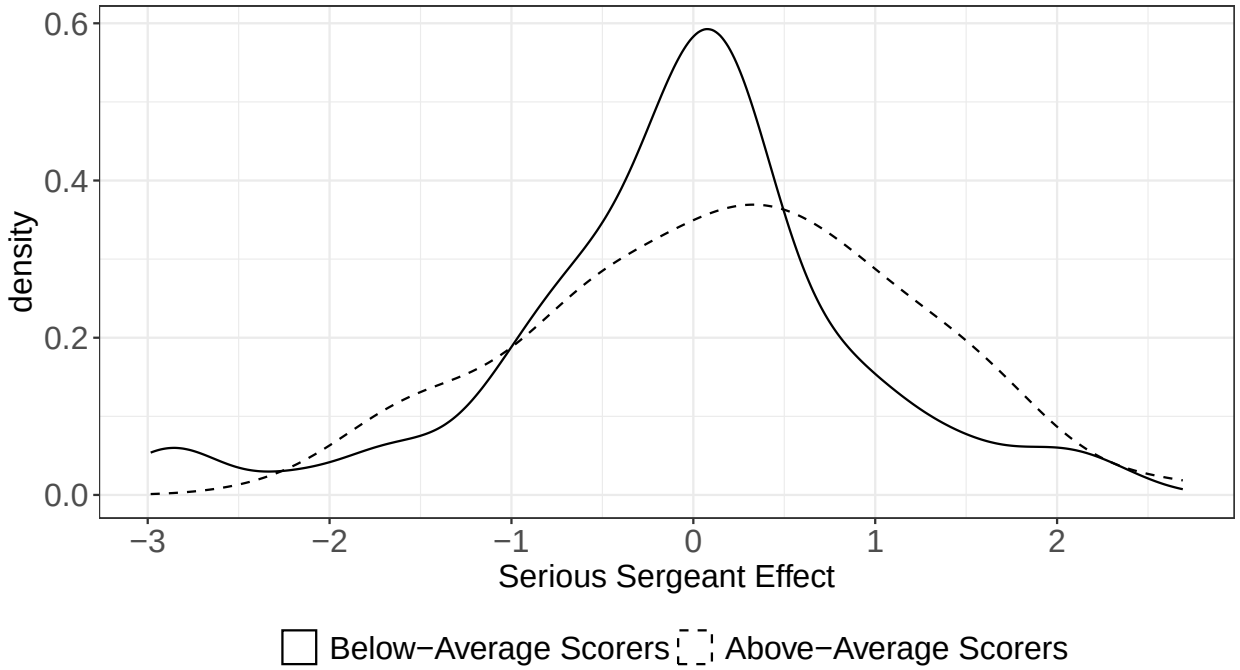
However, I find evidence that low-level officer effects differ by race (Figure E.5), consistent with the previous literature (e.g. Ba et al., 2021b). Thus, it appears that variation in officer enforcement effects do not translate into similar variation in sergeant effects.

Figure E.1: Sergeant Effects Distributions by Promotional Exam Performance

(a) Low-level effect distribution



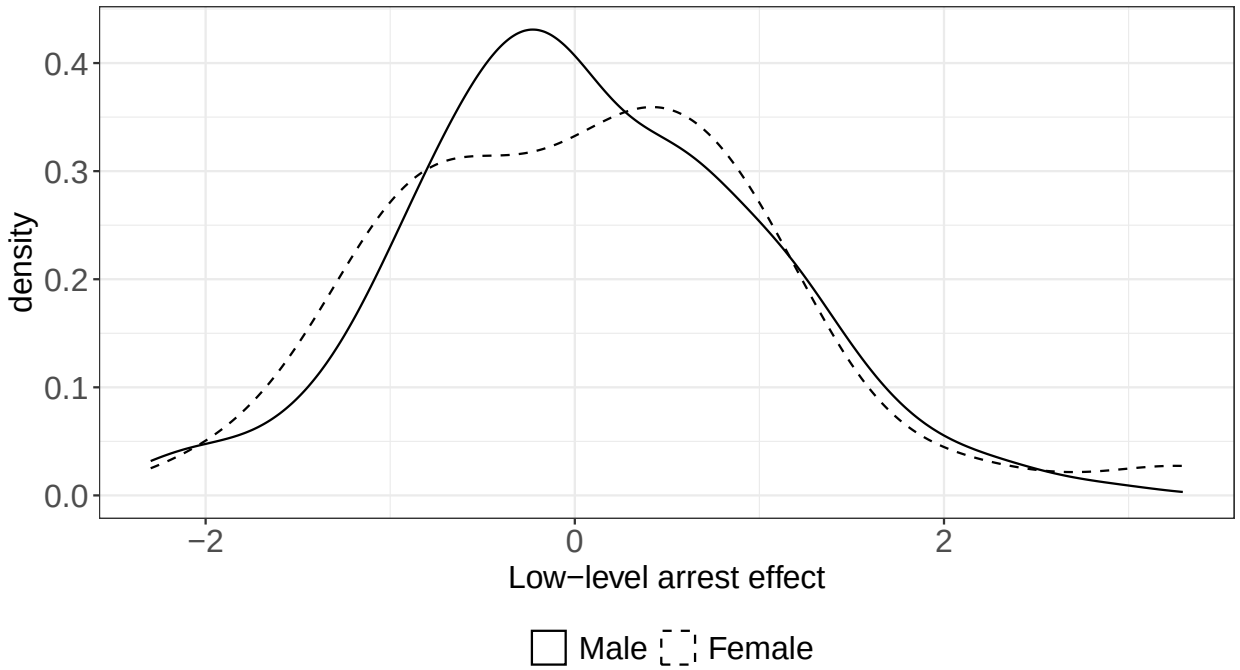
(b) Serious effect distribution



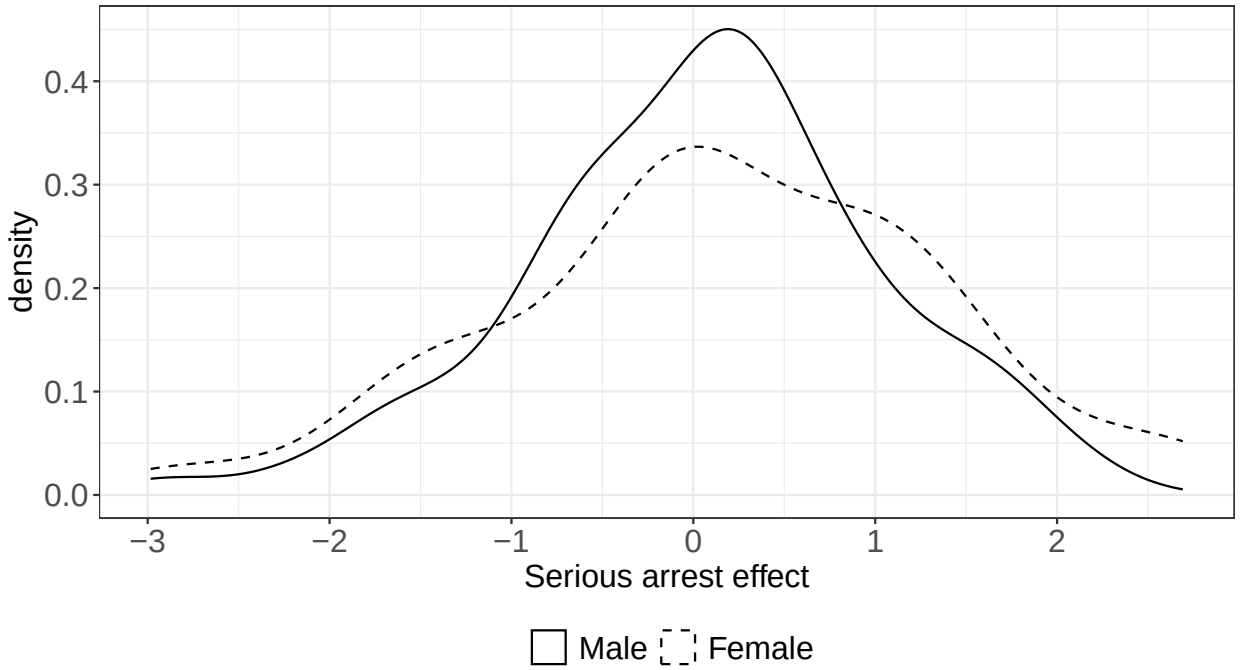
Notes: These figures plot distributions of standardized sergeant effects separately by the sergeant's performance on the promotional exam they took to attain the rank of sergeant. I define above-average scorers as sergeants who score above the average for their exam and below-average scorers as those who score below their exam-specific average. Kolmogorov-Smirnov p-value for low-level effects = 0.0343. Kolmogorov-Smirnov p-value for serious effects = 0.0128.

Figure E.2: Sergeant Effects Distributions by Gender

(a) Low-level effect distribution



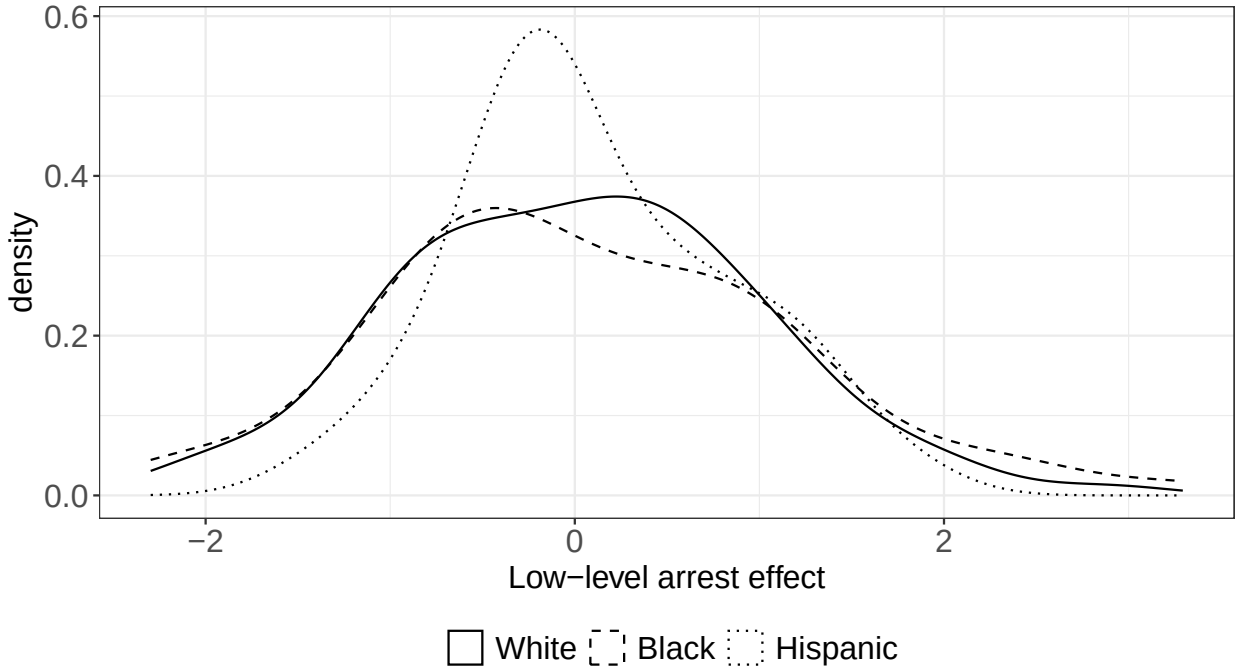
(b) Serious effect distribution



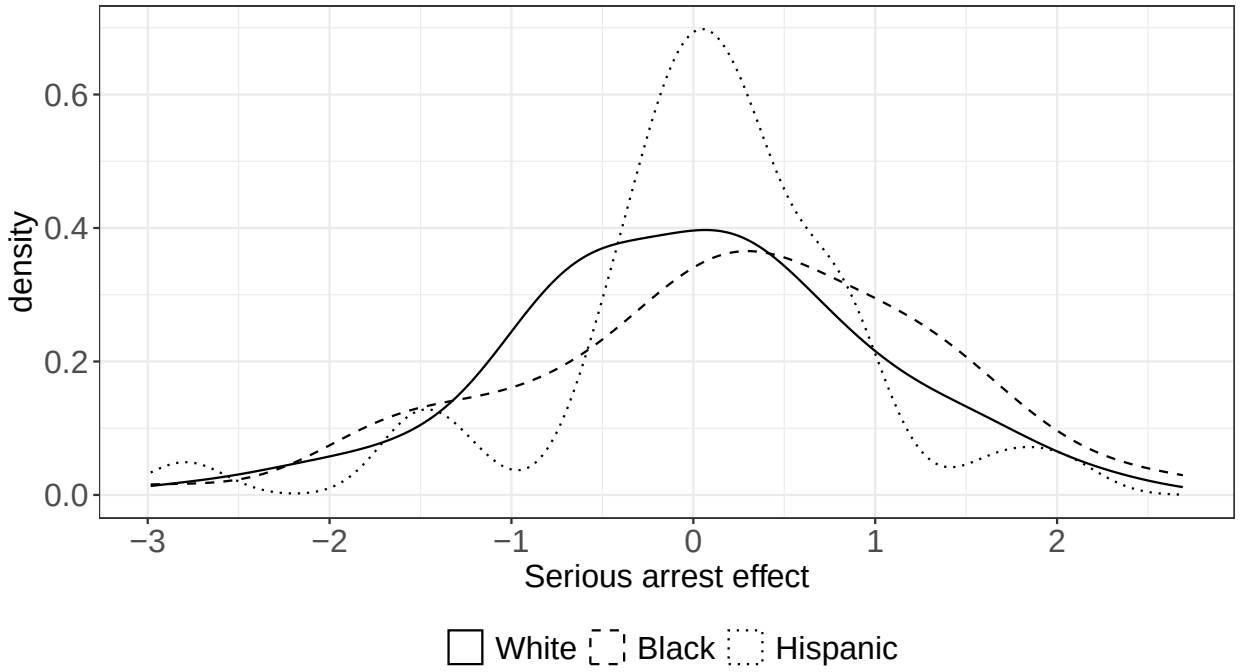
Notes: These figures plot distributions of standardized sergeant effects separately by sergeant gender. Kolmogorov-Smirnov p-value for low-level effects = 0.817. Kolmogorov-Smirnov p-value for serious effects = 0.415.

Figure E.3: Sergeant Effects Distributions by Race

(a) Low-level effect distribution



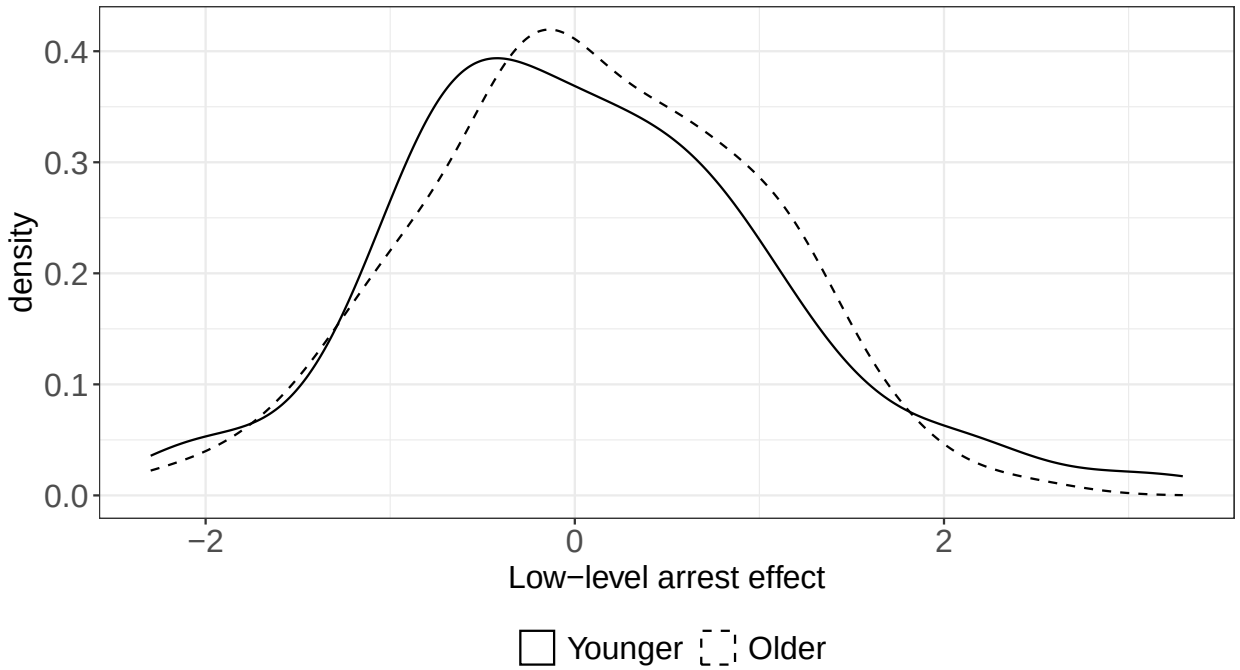
(b) Serious effect distribution



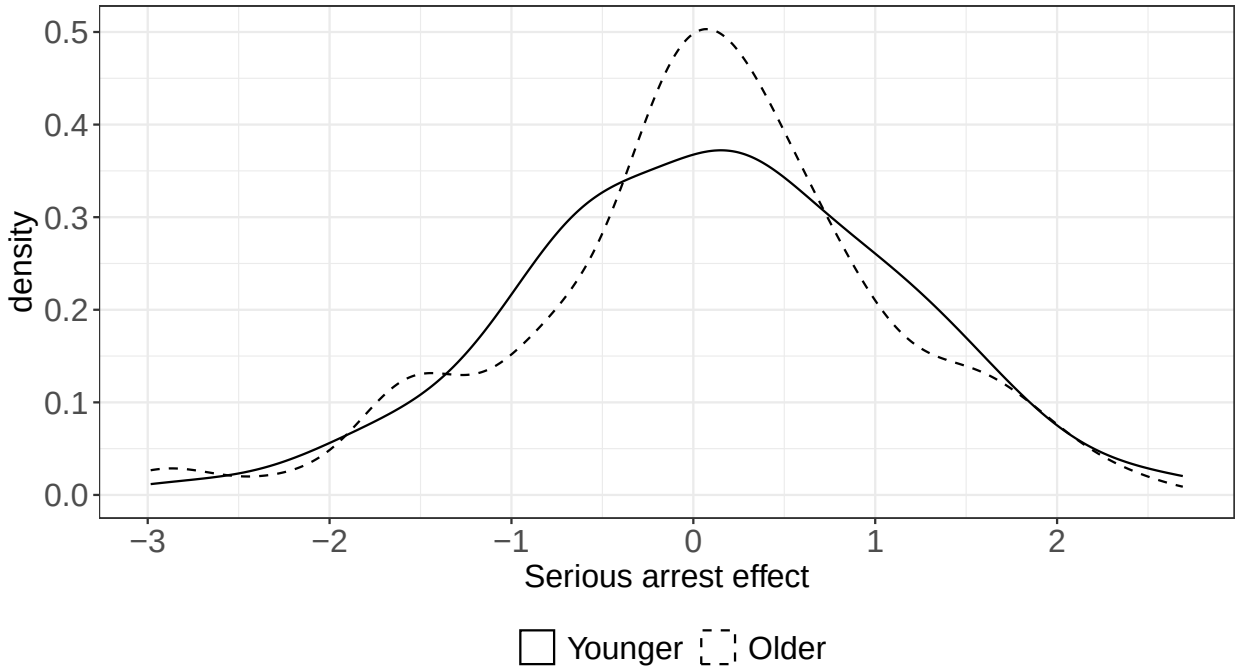
Notes: These figures plot distributions of standardized sergeant effects separately by sergeant race. For low-level effects: Black/Hispanic Kolmogorov-Smirnov p-value = 0.218. Black/White Kolmogorov-Smirnov p-value = 0.241. White/Hispanic Kolmogorov-Smirnov p-value = 0.855. For serious effects: Black/Hispanic Kolmogorov-Smirnov p-value = 0.239. Black/White Kolmogorov-Smirnov p-value = 0.219. White/Hispanic Kolmogorov-Smirnov p-value = 0.130.

Figure E.4: Sergeant Effects Distributions by Age During Promotional Exam

(a) Low-level effect distribution

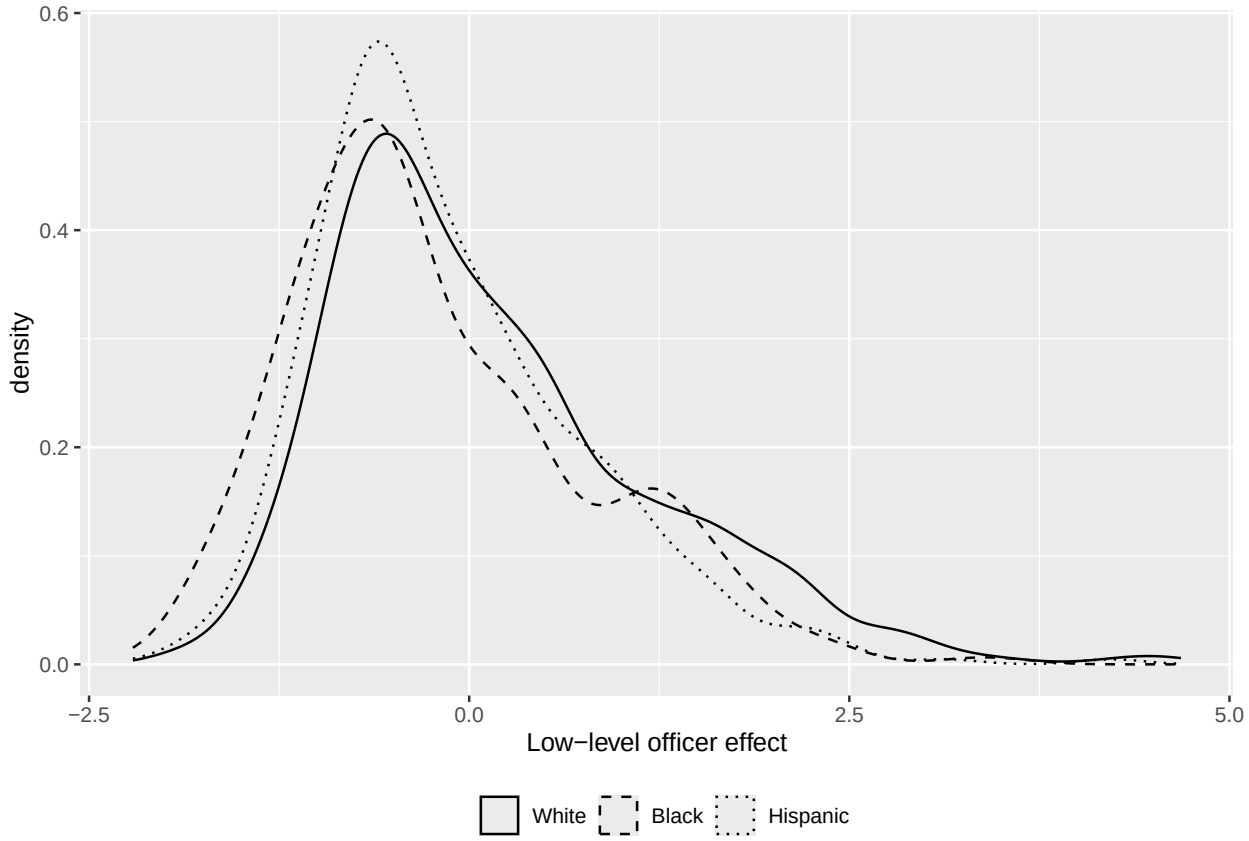


(b) Serious effect distribution



Notes: These figures plot distributions of standardized sergeant effects separately by age at the time of the promotional exam. Older sergeants are those who are above the average age across all exams in my sample, while younger sergeants are below the average. For low-level effects: Kolmogorov-Smirnov p-value = 0.587. For serious effects: Kolmogorov-Smirnov p-value = 0.587.

Figure E.5: Low-level Officer Effects by Race



Notes: This figure plots the distribution of low-level officer effects by officer race. Black/Hispanic Kolmogorov-Smirnov p-value = 0.047. Black/White Kolmogorov-Smirnov p-value = 0.000. White/Hispanic Kolmogorov-Smirnov p-value = 0.003.