The Boss in Blue: Supervisors and Police Behavior

Austin V. Smith *

September 17, 2026

Abstract

We know little about the influence of supervisors in high-stakes, discretionary settings such as policing. Leveraging 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 operate through distinct behavioral mechanisms: sergeants who drive low-level arrests disproportionately increase discretionary drug enforcement, while those who increase serious arrests expand officers' 911 call volume. 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 Dan Herbst, Evan Taylor, Ashley Langer, Matthew Ross, and CarlyWill Sloan for their invaluable guidance, support, and mentorship. I would also like to thank Mallory Cash, Spencer Cooper, Price Fishback, Lance Gui, Ami Ichikawa, Cynthia Lum, Daniel Nagin, Juan Pantano, Nayoung Rim, Wint Thu, Mo Xiao, and seminar participants at the University of Arizona, SEA Annual Meeting, ViCE, UCONN, and Providence College 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. 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. I thank Arnold Ventures (25-14011) for financial support. IRB Approval for this project was received from the University of Arizona (STUDY0005458) and Northeastern University (IRB # 25-08-42). Any 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 interpret. 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 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).

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 a substantial share of sergeants rank highly on one dimension while ranking low on the other.[1] These findings suggest that individual police supervisors shape how law enforcement policy is implemented on the ground.

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 using event studies around the time that an officer receives a new sergeant. I find that low-level sergeant effects operate predominantly through drug enforcement. Officers who move to a sergeant with low-level effects one standard deviation larger than their previous sergeant make 47% more drug arrests relative to the mean. Consistent

[1] The distinction between serious and low-level enforcement maps onto a longstanding debate in policing about the merits of broken windows enforcement — the idea that policing minor infractions maintains order and deters more serious crime (Corman and Mocan, 2005). Many sergeants therefore prioritize enforcement patterns consistent with a broken windows philosophy.

with the discretionary nature of these arrests, sergeants who increase low-level enforcement raise arrests from officer-initiated encounters and the likelihood of arrest at 911 calls. These additional enforcement encounters carry collateral costs: a one standard deviation increase in low-level sergeant effects leads to a 12% rise in use-of-force incidents relative to the mean.

By contrast, serious sergeant effects increase only the number of arrests made at 911 calls, not those made through officer-initiated channels. This increase operates through call volume. A one standard deviation increase in serious sergeant effects leads officers to respond to 2.3% more 911 calls per month, with no change in the likelihood of arrest at a given call. A substantial portion of the additional calls come during overtime, consistent with officers exerting effort to supply hours beyond their scheduled shifts. Serious sergeant effects raise arrests across serious offense types, including domestic violence, theft, and DWI. I find no statistically significant effect of sergeant-induced serious enforcement on use of force.

These findings underscore that sergeants shape not only how much enforcement occurs but what form it takes. This has important implications for public service delivery, since each form of enforcement carries its own distinct costs and benefits. I cannot comprehensively measure these costs and benefits, so my evidence does not speak to the welfare implications of sergeants prioritizing one type of enforcement over the other.[2] The evidence does identify first-line supervisors as instrumental in setting the enforcement priorities of police departments, and in determining how the costs and benefits of policing are distributed among the public.

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

[2]Several considerations illustrate the limits of what I observe. On the serious margin, the crime-reduction benefits of the additional arrests are difficult to establish, and I cannot rule out that the additional calls officers answer would otherwise have been answered by other officers. Additional call responses may also trade off against other beneficial activities such as local patrol presence (Weisburd, 2021), and the portion of the effect operating through overtime carries a fiscal cost to the department. On the low-level margin, I measure collateral costs such as use of force, and existing evidence finds that short-term reductions in low-level enforcement do not meaningfully increase serious crime (Cho et al., 2023). However, neither this evidence nor my own can speak to the long-run consequences of sustained changes in low-level enforcement for the communities where it occurs.

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 interpret, 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., 2025), 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.[3] 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. Frake and Harmon (2023) show that officers supervised by sergeants who have accumulated more complaints are themselves more likely to receive complaints. Supervisor misconduct therefore propagates down the chain of command. My paper complements this by estimating the full distribution of sergeant effects on arrests — outcomes with direct implications for civilians' criminal justice involvement. 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 key implication of my approach is that enforcement proclivities vary substantially across

[3] Adger et al. (2025) show that training officers who tend to use force increase use of force in their trainees past the training period, suggesting that changing who trains new recruits can reduce police aggression over time through the *next generation* of officers. My findings establish an analogous result for sergeants: officer behavior responds causally to supervisors, identifying the sergeant tier as a relevant margin for producing immediate behavioral change in the *current* officer workforce—those who have already completed field training.

sergeants and meaningfully shape officer behavior—identifying supervisory enforcement preferences, rather than supervisory racial composition, as a distinct policy-relevant margin. To the extent that departments can identify sergeants with particular enforcement proclivities, this heterogeneity suggests that strategic assignment or promotion decisions could be used to shape frontline outcomes without explicitly targeting supervisor race, which may face greater practical and political constraints. 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.[4]

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.[5] Moreover, the weak correlation I estimate between serious and low-level sergeant effects suggests that policies targeting supervisors may be able to reduce one type of enforcement without undermining enforcement more broadly.

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

[4]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?

[5]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—each unit is assigned to a geographic sector and shift, also known as a watch. I focus on patrol sergeants in the Dallas Police Department (DPD), which assigns sergeants to units across the city's 35 patrol sectors, with separate coverage across three daily watches. Within a given sector-watch, there may be one or two units of officers. Two units within the same sector-watch work overlapping shifts and operate in similar patrol capacities.

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 formal 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.[6] These functions give sergeants leverage to reward or discipline subordinates through commendations and documented infractions (Rim et al., 2024), while hierarchical norms reinforce a culture

[6]Sergeants are also expected to review use-of-force reports and monitor patterns in their officers' searches, citations, and arrests.

of compliance with supervisory authority (King, 2005). Beyond administrative oversight, sergeants can lead by example — actively patrolling alongside officers, responding to calls, and making arrests themselves — shaping expectations about desirable enforcement activity, particularly in ambiguous situations. Ethnographic evidence suggests officers place more trust in such "street sergeants" (Van Maanen, 1984), and survey data show officers are more likely to treat a task as evaluatively relevant when they see their sergeant engage in it (Engel and Worden, 2003). Field activity also allows sergeants to provide real-time guidance by attending their officers' calls, serving simultaneously as a monitoring tool and a managerial signal — strengthening supervisory influence even in the absence of high-powered incentives.[7]

In the DPD, patrol officers regularly receive new sergeants. A sergeant switch occurs either when an officer is reassigned to a new unit or when their current unit receives a new supervisor. Switches can therefore be *officer-driven* or *sergeant-driven*. *Officer-driven* switches occur when officers move to a new unit, inheriting the sergeant who manages it. *Sergeant-driven* switches occur when a sergeant moves to a new unit and inherits its existing officers.

Officers and sergeants have limited control over the timing or nature of switches, as unit 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.[8] Instead, executive command staff fill officer vacancies at their discretion within each division 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 divisions, 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 division, 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 sergeant,

[7]Dallas PD has used automatic vehicle location (AVL) technology to monitor officer locations since the early 2000s (Weisburd et al., 2015). AVL adoption has since spread widely across large and medium-sized agencies (Piza et al., 2025). My findings therefore generalize to similarly equipped departments, though they may be less applicable to smaller or rural agencies with more limited monitoring capabilities.

[8]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.

nor are sergeants allowed to choose their sector or officers. These assignments remain at the discretion of command staff and, based on my interviews with DPD personnel, appear not to follow a consistent rule.[9]

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 division level—a coarser geographic area—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 unit assignments of all responding personnel (including sergeants) on a daily basis.[10] I use these data, along with division assignments and promotion histories, to construct monthly sergeant assignments for patrol officers from June 2014 to July 2019. Since officers may occasionally work with substitute sergeants or be temporarily reallocated, I assign each officer to the unit in which they appear on the most days within a month, and assign each unit the sergeant most frequently observed with that assignment.[11] This sergeant retains administrative responsibility throughout, so the assignment construction captures the regularly assigned supervisor rather than temporary substitutes. To ensure that sergeant switches reflect transitions between regularly assigned supervisors rather than short-term arrangements, I require officer-sergeant spells to last at least two consecutive months and exclude spells with no identifiable sergeant.[12] I also remove officers, sergeants, and other groupings

[9]Lieutenants, the rank directly above sergeant, are assigned at the division (also known as a station) level over each of the three watches, meaning each sergeant who works in the same division and watch at any point in time works under the same lieutenant. Since sergeants are assigned at the finer sector-watch level, most sergeant switches exploit variation within a lieutenant's assignment: 78% of switching events used for identification occur within the same division and watch.

[10]Unit assignments are given by a 4-character alphanumeric that contains information on the assigned sector and watch, as well as an intra-sector-watch unit code, where the unit code distinguishes between units operating within the same sector-watch when applicable. See Supplementary Appendix C for details.

[11]For officers, 87% of their CAD observations match their modal unit assignment. For units, 94% of CAD observations match the modal sergeant.

[12]Supplementary Appendix Figure C.1e shows that including single-month sergeant-officer spells attenu-

(e.g. officer-sergeant pairs) that appear too infrequently to support reliable identification of fixed effects. Full details of assignment construction, sample restrictions and robustness checks to alternative sampling decisions are given in Supplementary Appendix C.

To study trends around officer moves, I construct a separate event study sample. I define an event as two consecutive spells involving the same officer but different sergeants. I focus on the window 5 months before and 5 months after an officer switches sergeants. I require that no other switching events occur during any of the pre- or post-periods, to avoid contaminating periods with behavioral responses to another sergeant transition. I allow the sample to be imbalanced to accommodate moves near the edges of the officer's sample period and moves that contain any time spent in a temporary or non-patrol assignment in pre- or post-switch periods. The transition from one sergeant to the next occurs sometime during the final month the officer is assigned to their previous sergeant, which I denote as month 0. Supplementary Appendix F shows that the main event study results are robust to a balanced sample and a fully imbalanced sample that includes all switches.

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 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).[13] All other arrests are classified as low-level. These primarily consist of outstanding warrants,[14] disorderly conduct, and drug possession, which together account for 52% 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 on the most severe charge.[15] I also link each arrest charged with a felony or misdemeanor to court outcomes using data from the Dallas District Attorney's

ates sergeant effect estimates by around 25%. Baseline and alternative-sample effects have a correlation of 0.91, and regressing alternative-sample on baseline effects yields a slope of 0.76. The two samples produce similar rankings of sergeants, but including short-term relationships generates smaller sergeant effect estimates. This suggests that the documented impacts are driven in part by incentives associated with enduring professional relationships rather than exclusively by short-term changes in direct instruction.

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

[14] 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).

[15] 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.12).

office. I discuss this linking procedure in Supplementary Appendix C.

I extract 911 calls from CAD data to separately evaluate civilian-initiated and proactive police encounters. An arrest is considered officer-initiated if it does not originate from a 911 call. I aggregate 911 calls at two levels: the number of calls made by civilians within a sector-watch-month, and the number of calls each officer responds to within a month.

At each level, 911 calls have two plausible interpretations. First, they measure crime conditions, since calls reflect civilian requests for police service. Therefore, I may observe changes to 911 calls at either level because crime rates have changed. Second, they measure police activity. At the area level, civilians may be more likely to report crimes when trust in law enforcement is higher (Ang et al., 2024). At the officer level, officers can influence how many calls they answer. This could reflect effort, if officers volunteer for calls or work additional overtime hours to respond to them. It could also reflect availability, if officers are less engaged in other activities and dispatchers assign them more calls.[16]

I adjudicate between these interpretations throughout the paper. At the sector-watch level, I use 911 calls primarily as a measure of crime, preferring them over crime reports because proactive police activity directly generates crime reports. As robustness checks, I report results using a subset of the most serious calls: those likely to involve violence (robbery, assault, major disturbances) and those likely to involve costly property crimes (theft, burglary, auto theft). These have a relatively constant reporting rate and therefore reflect actual crime rather than officer behaviors that affect civilians' willingness to call the police. As an additional check, I report results using crime reports for index crimes, which are less likely to be generated by officer activity due to their severity. Report data are obtained from Dallas PD's National Incident Based Reporting System (NIBRS) database. At the officer level, I treat 911 call volume as a combined measure of self-selection behaviors and crime conditions. In Section 6, I provide several pieces of evidence suggesting that sergeant-induced changes in officer call volume are driven by changes in officer behavior rather than crime conditions.

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 G.

[16]This level of officer influence over call assignment is a distinctive feature of Dallas's 911 response system, as documented by Weisburst (2024) and Adger et al. (2025). Dallas dispatchers retain final authority over call assignment. The mechanisms described above give officers several margins to influence which calls they are assigned.

3.1 Summary Statistics

Summary statistics for the full, unrestricted data, the analysis sample, and the event study sample are presented in Appendix Table B.1. The event study sample contains roughly half of the switches that are leveraged for identification in the analysis sample, and captures at least one switch for 98% of sergeants and 81% of officers in the analysis sample.

Officers are highly mobile and sergeants supervise many officers. The average officer works with just under four unique sergeants, and the average sergeant manages over 20 officers. This density is essential for my empirical strategy, since sergeant fixed effects are identified only within groups of officers and sergeants connected by movement across assignments (Abowd et al., 2002). All observations in my data fall within one connected set.[17]

The analysis sample contains 5,798 switching events. Switches occur either because an officer moves to a different unit (officer-driven) or because a sergeant moves to a different unit and inherits its existing officers (sergeant-driven). Of the switches in the analysis sample, 62.7% are officer-driven and 37.3% are sergeant-driven. Both types arise from the vacancies and schedule realignments described in Section 2. I exploit variation from both to estimate sergeant effects. The two switch types may nonetheless raise distinct identification concerns and generate differential behavioral responses, which I discuss in Section 4.

Patrol officers make 3.8 arrests per month on average, three-fourths of which are for low-level crimes, with substantial variation across officer-months (standard deviation = 3.64). Supplementary Appendix Figure A.2a plots the distribution of average arrests per officer-month across sergeants. Officers working under sergeants in the right tail average over six arrests per month, while those in the left tail average one or fewer. These averages do not identify sergeants' causal effect, however, since they confound officer discretion with sergeant influence. Indeed, arrest behavior varies even more widely across officers than across sergeants (Supplementary Appendix Figure A.2b). Disentangling the two requires observing how officer behavior changes when officers switch supervisors, which is the central idea behind the empirical strategy described in the next section.

[17]The sample contains 102 sector-watches and 183 units in total — reflecting the fact that some sector-watches contain two units. The average of 4.05 sergeants per unit reflects turnover within units over time. The higher average of 6.95 sergeants per sector-watch reflects the presence of multiple sergeants managing units within the same sector-watch at the same time.

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) (AKM), 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 control vector x_{it} includes sector-watch fixed effects to account for spatial and temporal variation in crime.[18] Because sergeant and sector-watch assignments overlap, identifying both sets of fixed effects separately requires that each sector-watch be managed by multiple sergeants over time. This condition is satisfied: the average sector-watch is associated with 6.95 sergeants over the sample period, reflecting both turnover within units and the presence of multiple units within the same sector-watch. 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 switching between sergeants. A sergeant's estimated effect reflects changes in the arrest behavior of officers who switch into or out of their supervision, estimated using within-officer variation relative to officers working in the same patrol location, shift, day-off group, and with similar tenure. As noted in Section 3, equation 1 combines within-officer variation from two types of moves: officers moving between units (*officer-driven switches*) and sergeants moving between units (*sergeant-driven switches*). Using both switch types increases power to detect sergeant effects. It is nonetheless worth considering whether each generates the same behavioral response and raises the same identification concerns. On the behavioral side, officers who have worked together in a unit for a long time may respond differently to receiving a new supervisor than officers who are themselves new to a unit. On the identification side, officer-driven switches can coincide

[18] I include sector-watch rather than unit fixed effects because two units within the same sector-watch face similar crime conditions and perform similar functions, making the sector-watch the relevant environmental control. Replacing sector-watch fixed effects with unit fixed effects produces sergeant effect estimates that are highly correlated with the baseline (Supplementary Appendix Figure A.5).

with sector-watch changes and may be susceptible to sector-watch crime trends not absorbed by sector-watch fixed effects.

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.[19] 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 identification. First, sergeant assignments must be uncorrelated with unobserved trends in officer behavior or local crime. For instance, if high-arrest sergeants are more likely to fill vacancies when crime is rising, officers moving to them would make more arrests for reasons unrelated to managerial influence. Second, sergeant switches must not coincide with unobserved shocks to officer behavior. A specific concern is departmental policy changes — such as hot-spot policing initiatives (Weisburd and Eck, 2004) — that overlap with officer moves and could be mistaken for sergeant effects. Third, identification assumes officers are not matched to sergeants based on idiosyncratic comparative advantage in making arrests, which would generate match-specific effects 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. Both officer-driven and sergeant-driven switches arise from the same two mechanisms — Patrol Bids and vacancy-induced moves — meaning the constraints on sorting apply equally to both switch types. While officers and sergeants may select their preferred division 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 within patrol divisions. 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.,

[19]By the same token, the model also permits sorting (for officers and sergeants) based on time-invariant features of sector-watches or day-off groups, which may arise from the schedule realignment process.

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, and plot officer arrest trajectories separately by the tercile of the sergeant they move to and from, residualizing arrests by officer fixed effects and the control vector using within-supervisor variation, following Chetty et al. (2014).[20] For this figure, I use the sample of switches that are balanced 2 months prior to the move and 2 months after the move. To limit the amount of information on the graph, I only report moves between or within the top and bottom terciles. The figure reveals four patterns supporting the identifying assumptions. First, officer arrests shift sharply and persistently following a change to a different sergeant tercile. Second, pre-move trends do not vary systematically with the direction of the switch. Third, moves in opposite directions produce roughly symmetric effects, inconsistent with match-quality sorting (Card et al., 2013).[21] Finally, 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, but consistent estimation requires observations to tend to infinity within each officer-sergeant pairing, so raw fixed effects may be estimated with error. 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) and frequently applied in the policing literature to estimate officer fixed effects (e.g. Rozema and Schanzenbach, 2019; Weisburst, 2024; Adger et al., 2025). Specifically, following Best et al. (2023), I bootstrap equation 1 to decompose the variance in sergeant fixed effects into signal variance, σ^2_ψ, and sampling error variance, σ^2_ϵ, then multiply each raw fixed effect by the shrinkage factor, $\frac{\hat{\sigma}^2_\psi}{\hat{\sigma}^2_\psi+\hat{\sigma}^2_\epsilon}$.[22] 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. 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

[20]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.

[21]In Supplementary Appendix Figure A.3, I show this symmetry holds across all opposite-direction tercile transitions.

[22]For the bootstrap, 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 $\hat{\sigma}^2_\epsilon$.

the same Empirical Bayes procedure. I focus on the two-dimensional relationship between a sergeant's serious and low-level effects, $Cor(\psi_j^S, \psi_j^L)$, 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.

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

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

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 first 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.[23] 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

[23]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).

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 3 sergeants and 183 officers from my sample.

Rather than relying on ex-post bias corrections for the noisily-estimated fixed effects, one could instead estimate the variance terms directly using a random effects model, which specifies the officer and sergeant effects as random variables drawn from population-level Normal distributions.[24] This more parsimonious model apportions variability between officers and sergeants in a way that minimizes mean squared error, conditional on its distributional assumptions. Following Beuermann et al. (2023), it can also be used to directly estimate the correlation between serious and low-level sergeant effects while accounting for measurement error in the components. The normality assumption may be particularly restrictive in the police setting, where agent propensities often exhibit fat tails due to the outsize influence of a small number of officers in generating a disproportionate share of arrests (Weisburst, 2024), discrimination (Goncalves and Mello, 2021), fines (Gonçalves and Mello, 2023), and misconduct (Rozema and Schanzenbach, 2019). The distributions of sergeant and officer mean arrests in Supplementary Figures A.2a and A.2b confirm that this concern applies in my setting. Imposing normality on fat-tailed distributions understates variance. I therefore prefer the fixed effects approach with post-estimation bias correction, which recovers the variance using the sampling variation in the estimated effects rather than relying on a fully parametric likelihood. I report random effects estimates of both the variance components and the cross-outcome correlation using restricted maximum likelihood (REML) as a conservative robustness check in Supplementary Appendix Tables B.2 and B.3. Both specifications include officer tenure covariates and sector-watch and day-off group fixed effects.

4.3 Event Study Design

To visualize how officer behavior changes around a sergeant move, I implement a movers event study, following Finkelstein et al. (2016), Molitor (2018), Fenizia (2022), and Frakes and Gruber (2025). Specifically, I compare officer movers in the event study sample who move to sergeants with higher or lower arrest effects than their previous sergeant. For each switching event e, uniquely determined by officer i and the switch month T, define $\Delta\hat{\psi}_e$ as the

[24]The random effects model assumes that the *potential* officer and sergeant effects are uncorrelated at the population level, but allows for *observed* officer and sergeant effects to be correlated in small samples (Jackson, 2013).

difference between the Bayes-shrunken sergeant effect for officer i's new and old sergeant.[25] The event study regression takes the following form:

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

Here, k indexes months relative to T, with $k \in [-5, 5]$ and $k = -1$ as the reference month. D_{et}^k is an indicator for being k months relative to the switch, and π_0^k captures dynamics common across all switches. I include baseline controls (tenure, sector-watch fixed effects, and day-off group fixed effects) and an event fixed effect α_e to control for differences in baseline arrest rates prior to the switch.

The parameters of interest are the π_1^k's, which capture the month-k difference in y relative to the month just before the switch, scaled by the difference in sergeant effects between the new and original sergeant, $\Delta\hat{\psi}_e$. For the baseline sergeant effects, π_1^k is interpreted as the difference in y that arises from switching to a sergeant who induces one more arrest than your previous sergeant. This event study set-up mirrors the AKM model in equation 1, as the π_1^k's are identified from officers moving between sergeants with different impacts on arrests. The identifying assumption of this model is also similar: absent the sergeant switch, arrest behavior would have evolved similarly for officers moving to higher- and lower-enforcement sergeants. I test this assumption directly using the pre-switch event study coefficients in Section 5.

In order to summarize the average change in officer behavior across all post-switch periods, I also report estimates from a difference-in-differences (DD) version of equation 4, replacing the post-switch event-time indicators in equation 4 with an indicator variable for the event-time being after the switch:

$$y_{et} = \alpha_e + \sum_{k < -1} (\gamma_0^k D_{et}^k + \gamma_1^k D_{et}^k (\Delta\hat{\psi}_e)) + \pi_0 Post_{et} + \pi_1 (Post_{et} \times \Delta\hat{\psi}_e) + x'_{et}\beta + \epsilon_{et} \tag{5}$$

Here, $Post_{et}$ is an indicator for $k > 0$, and I estimate the regression excluding the switch month where $k = 0$. By including individual pre-period interaction terms as in equation 4, the coefficient of interest, π_1, captures the average post-switch difference in y relative to period -1 for a 1-arrest increase in the effect of an officer's managing sergeant.

To separately study dynamics in officer behavior that are associated with serious and low-level sergeant effects, I estimate a version of equation 4 that includes interaction terms

[25]By using the Bayes-shrunken estimates of the sergeant effects, I avoid bias induced by using estimated objects as regressors (Walters, 2024).

for each sergeant effect type:

$$y_{et} = \alpha_e + \sum_{k \neq -1} [\pi_0^k D_{et}^k + \pi_1^{k,L} D_{et}^k (\Delta \hat{\psi}_e^L) + \pi_2^{k,S} D_{et}^k (\Delta \hat{\psi}_e^S)] + x'_{et}\beta + \epsilon_{et}. \tag{6}$$

To facilitate direct comparisons, $\Delta\hat{\psi}_e^L$ and $\Delta\hat{\psi}_e^S$ are differences in *standardized* sergeant effects, divided by their relevant standard deviations, so that $\pi_1^{k,L}$ and $\pi_2^{k,S}$ capture changes in officer behavior attributable to a one standard deviation increase in the low-level and serious sergeant effects, respectively. For each event study of this type, I estimate an analogous DD specification that replaces both the pre-period interaction terms and the $Post_{et} \times \Delta\hat{\psi}_e$ term in equation 5 with separate interaction terms for each sergeant effect type — $\Delta\hat{\psi}_e^L$ and $\Delta\hat{\psi}_e^S$. The coefficients on $Post_{et} \times \Delta\hat{\psi}_e^L$ and $Post_{et} \times \Delta\hat{\psi}_e^S$ thus capture the averages of the post-switch event study coefficients, and are reported alongside the event study results.

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.[26] 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.4), sergeants supervise multiple officers, amplifying their aggregate influence. To illustrate, replacing a 90th percentile sergeant who

[26]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.

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. Supplementary Appendix Figure A.6 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 1, 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. However, this estimate may be 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.

Supplementary Appendix Table B.2 compares variance decomposition estimates between the fixed effects and random effects approaches. The random effects estimates are substantially smaller — the implied variance, standard deviation, and 90th-to-10th percentile difference in sergeant effects are 40%, 23%, and 42% smaller, respectively — consistent with the normality assumption understating variance in the presence of fat tails, as discussed in Section 4. Even so, sergeant effects remain economically meaningful: a one-standard-deviation increase in an officer's sergeant raises arrests by 12% relative to the mean, and moving from the 10th to 90th percentile raises arrests by 25%. The Best Linear Unbiased Predictors (BLUPs) from the random effects model are highly correlated with the Bayes-shrunken fixed effects (Supplementary Appendix Figure A.8). The two approaches agree on the relative ranking of sergeants even where they differ on the magnitude of variation.

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% 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.[27]

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.

Event Study Results—I plot event study coefficients in Figure 3, and report the DD coefficient in the top right corner of the figure. There is no evidence of heterogeneous trends in arrest behavior prior to officers changing sergeants, lending support to the identification

[27] My results can also be compared to recent research in the economics of public safety literature to assess the importance of sergeants within the production of criminal justice. Weisburst (2024) shows that a 1SD increase in officer arrest propensity at 911 calls leads to a 40% increase in the likelihood of an arrest. This is consistent with my finding that officer variation explains more variation in officer-level enforcement than sergeant variation. The replacement simulations suggest, however, that sergeants may matter more when aggregated over all of the officers they manage. Bacher-Hicks and De La Campa (2020a) find that a 1SD increase in precinct commander effects on street stops raises stops by 7% relative to the mean. My estimates suggest that a 1SD increase in sergeant effects has slightly more than double the effect on officer arrests. This is consistent with commanders having substantial control of large-scale policy decisions, while sergeants are uniquely positioned to impact enforcement decisions through direct interactions with officers.

assumptions. I am unable to reject that the pre-trend coefficients are jointly 0 (F-stat p-value = 0.367). Following a switch, an officer's arrests immediately increase and remain elevated at similar levels throughout the panel period, in line with the insights from the nonparametric event study in Figure 1.[28] This suggests that the sergeant effects captured by equation 1 reflect a set of incentives provided to officers which are both salient (given their immediacy) and persistent (given their stability).[29]

Supplementary Appendix Figure A.13 plots event study coefficients separately by switch type. Both types generate immediate and persistent increases in officer arrests that are not driven by pre-existing trends. The first post-switch coefficient is smaller for sergeant-driven moves than officer-driven moves (0.56 vs. 0.91), as is the DD estimate (0.81 vs. 1.08). Arrests converge across switch types in subsequent post-switch months. An F-test does not reject equality of post-switch coefficients across switch types (p = 0.22). The sergeant-driven results are also reassuring that sergeant effects are not confounded by changes to an officer's assignment, since officers retain the same sector-watch assignment before and after a sergeant-driven switch.

5.2 Diagnostic Checks

As discussed in Section 4, there are three prominent sources of endogenous mobility that could bias the sergeant effect estimates. The event study results in the previous section provide direct evidence against trend-based sorting in officer behavior: the flat pre-switch coefficients indicate that officer arrests do not trend differentially with the direction of the sergeant switch prior to the move.

However, this test 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, requesting to fill officer vacancies when crime is rising. In such cases, officers would make more arrests after the move not due to the new sergeant's influence, but because they are assigned to a

[28]The size of the effect after moving is statistically indistinguishable from 1 in all periods, which is reassuring since 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.

[29]There may be concern that these estimates are contaminated by effects from other periods, given the staggered timing of switches and potential heterogeneity of effects across switchers (Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2020). To address this, I run the event study using an estimator robust to staggered timing and treatment-effect heterogeneity. I binarize $\Delta\hat{\psi}_e$, replacing it with an indicator for whether it falls in the top tercile of switches, meaning switches to a sergeant who induces substantially more arrests than the officer's previous one. I then estimate this modified specification with standard two-way fixed effects and with the interaction-weighted estimator of Sun and Abraham (2021). Supplementary Appendix Figure A.9 presents results, which are similar across the two estimates. This indicates staggered timing is unlikely to bias the main results.

location with increasing demand for enforcement.

To assess this possibility, I test whether changes in crime predict changes in sergeant assignments. Specifically, for each sector-watch-month, I compute the average change in sergeant fixed effects across all officers who move into (or out of) that sector-watch during that month, and regress this average on logged 911 call trends in the relevant sector-watch. I report estimates in Supplementary Appendix Table B.6, separately for officer moves into and out of the sector-watch. Since officers switching sergeants may shift enforcement tendencies in either direction, I report the mean absolute value of the outcomes as a measure of the typical reassignment magnitude, abstracting from sign. F-tests on the pre-period 911 call coefficients suggest that crime trends do not jointly predict the direction or magnitude of changes in sergeant effects, and the largest point estimate across all specifications implies that a 10% increase in 911 calls is associated with a change of no more than 6% of the average magnitude of sergeant reassignments, mitigating concerns that crime trends confound the sergeant effects. I replicate this analysis using violent 911 calls, property crime 911 calls, and index crime reports as alternative crime measures in Supplementary Appendix Tables B.7, B.8, and B.9, respectively. The results are similar across all measures: pre-trends are not jointly statistically significant and individual coefficients are small relative to the average magnitude of sergeant reassignments. The lack of pre-trends is therefore unlikely to reflect differences in civilians' propensity to report crimes across measures.

A second identification concern is that changes in sergeant assignment may coincide with unobserved shocks that independently affect officer behavior — for instance, a department-wide or assignment-specific policy change that encourages more aggressive enforcement at the same time an officer is reassigned to a high-arrest sergeant. To evaluate this concern, I focus on officer-driven switching events and construct a panel of incumbent officers — those already working under the switching officer's new sergeant throughout the event window. I then estimate an event study model using incumbent officers' monthly arrest counts as the outcome. The specification mirrors equation 4, interacting event-time indicators with the difference in sergeant effects at the time of the switch, and includes incumbent officer-by-switching event fixed effects alongside controls for sector-watch, day-off group, and officer tenure. If policy shocks rather than managerial influence drive the results, incumbent officers' arrests should respond to the switch in the same way as the switching officer's — since they face the same policy environment. The specification also serves as a secondary test for endogenous crime trends: if larger sergeant effects were driven by growing local enforcement demand, incumbent arrests would trend upward prior to a positive switch.[30] In

[30] I am unable to perform a similar test for sergeant-driven switches, since the incumbent officers are themselves affected by the switch in that case. However, sergeant-driven switches produce similar behavioral

Supplementary Appendix Figure A.10, I present the estimates and 95% confidence intervals for the event study coefficients. The estimates are close to zero and statistically insignificant across all months relative to the switching officer's move date, supporting the interpretation that the estimated sergeant effects reflect genuine differences in managerial behavior rather than correlated shocks in officers' decision environments. The absence of pre-switch trends in incumbent arrests also reaffirms that crime trends do not drive the observed changes in officer behavior around officer-driven switches.[31]

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. In Supplementary Appendix D, I conduct additional diagnostic checks testing the additive separability assumption underlying equation 1 and assessing whether the estimated variance in sergeant effects can be distinguished from estimation noise generated by randomly reallocating sergeants in my data. Both checks are supportive of the AKM estimates as valid reflections of sergeants' causal effects on officers. 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.[32] Figure 4 plots each sergeant's estimated serious arrest effect (vertical axis) against their

responses to officer-driven switches and arise from analogous assignment mechanisms, suggesting they are unlikely to be uniquely correlated with policy shocks. Additionally, in Supplementary Appendix H, I show that pre-switch crime trends in a unit's sector-watch do not predict the direction or magnitude of sergeant effect changes when sergeants move between units, providing direct reassurance against crime trend sorting for sergeant-driven moves.

[31] A related concern is that estimated sergeant effects may be conflated with peer effects if switches coincide with changes to an officer's peer group. In Supplementary Appendix D.3, I show this is unlikely: officer-driven switches to a higher-effect sergeant are associated with *reductions* in peer officer arrest effects, including peer attributes as controls in the AKM model leaves sergeant effect estimates unchanged, and results are similar when focusing on sergeant-driven moves, where peer composition does not change.

[32] Supplementary Appendix Tables B.4 and B.5 present variance decompositions for serious and low-level arrests separately, following the same approach as Table 1. Sergeant effects account for 5.91% and 3.58% of the variation in serious and low-level arrests, respectively — magnitudes comparable to the share estimated for total arrests. Event studies around sergeant switches for serious and low-level arrests separately (Supplementary Appendix Figure A.15) show sharp, immediate post-switch increases for each, with F-tests failing to reject that pre-switch coefficients are jointly zero. Sergeant-driven and officer-driven switches produce similar changes in officer behavior along both dimensions of sergeant effects (Supplementary Appendix Figure A.16). I cannot reject that the two switch types produce equal changes for either arrest type.

low-level arrest effect (horizontal axis). Observations are binned to ensure each point represents an equal number of sergeants, and a linear regression line is fitted across all sergeants. 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.[33] This relationship is driven by a small subset of sergeants with large negative effects in both dimensions. Excluding the bottom 5% of low-level arrest sergeants reduces the correlation to 0.07, which is no longer statistically significant.[34] These findings do not support the presence of systematic complementarities or tradeoffs in how sergeants influence different types of enforcement. Instead, sergeants vary independently along the two enforcement dimensions. Over 20% of sergeants fall in the top tercile of one effect and the bottom tercile of the other, while 11% fall in the top tercile of both (see Supplementary Appendix Figure A.11). Sergeants therefore occupy a wide range of positions in the joint distribution of the two effects.[35]

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.14. 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

[33] One might be concerned that the near-zero correlation between serious and low-level sergeant effects reflects differences in the crime profiles of assigned areas rather than enforcement priorities. The sector-watch fixed effects absorb all time-invariant location characteristics, including differences in crime profiles across assignments. Supplementary Appendix Figure A.7 plots sector heat maps of sergeant and officer effects alongside crime measures. Correlations with crime measures are weak (all below 0.25 in magnitude). The serious/low-level distinction therefore reflects enforcement priorities, not neighborhood crime profiles.

[34] Results are similar when categorizing arrests using the FBI's traditional index crime classification; see Figure A.12 in the Supplementary Appendix.

[35] In Supplementary Appendix G, 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.

are policed most intensely.[36]

In Supplementary Appendix H, I examine the impact of serious and low-level sergeant-induced enforcement on crime by exploiting within-unit variation in sergeant effects driven by sergeant rotations between units. In the five months following a new sergeant's arrival, neither sergeant type uniquely reduces crime: joint F-tests do not reject equality of post-switch effects between serious and low-level sergeant effects across all outcomes. Where I observe statistically significant reductions for one sergeant type — suggestive property crime reductions for serious sergeant effects — I cannot reject that low-level effects are similar. The one exception is violent crime reports, where I reject equality at the 1% level, driven by a suggestive increase for low-level effects rather than a reduction for either type. This analysis is subject to two limitations: officers' arrests are not confined to their assigned sector-watch, and sergeant-induced enforcement represents a small share of overall sector-watch enforcement activity — both of which attenuate detectable crime effects.[37] The finding of limited and indistinguishable crime reductions is consistent with broader evidence on enforcement-driven variation in arrests. Cho et al. (2023) find that large reductions in arrests following an officer death — driven by officer behavior rather than changes in crime or policy — do not meaningfully increase index crimes, suggesting that marginal arrests are not substantially deterring or incapacitating those who will commit the most costly crimes.

These findings have important implications for law enforcement policy. Serious and low-level arrests are strongly complementary at the officer level but only weakly related at the sergeant level. Changing the composition of enforcement may therefore be more feasible through sergeant assignment than through officer-targeted policy. For example, replacing a unit's sergeant with one lower in the low-level effects distribution need not come at the cost of fewer serious arrests. Officer-targeted reforms must instead navigate a stronger productivity tradeoff: officers who make more low-level arrests also tend to be more active in serious enforcement. Policies aimed at reducing low-level encounters among officers therefore risk broader reductions in enforcement activity, potentially leading to de-policing effects (e.g.

[36]Supplementary Appendix Table B.3 estimates the correlations between serious and low-level effects for both sergeants and officers jointly using a random effects approach, following Beuermann et al. (2023). The qualitative pattern is preserved: the officer correlation remains strong (0.603) while the sergeant correlation is negligible (-0.017). The implied variances are smaller than the shrunken fixed effects estimates in Tables B.4 and B.5, consistent with the normality assumption understating dispersion in fat-tailed distributions, as discussed in Section 4.

[37]Because a substantial portion of sergeant-induced enforcement occurs outside the assigned sector-watch, marginal enforcement may displace enforcement by officers in location- or time-adjacent units. I cannot rule out this possibility. However, this concern is weaker for low-level arrests. As shown in Section 6, low-level sergeant effects increase the number of arrests made through channels tied to officer discretion, which reflects enforcement actions unlikely to have occurred otherwise. Serious sergeant effects instead increase arrests through the number of calls officers answer, and I cannot rule out that these are calls other officers would otherwise have answered.

Heaton, 2010). These results underscore the managerial leverage sergeants have in shaping enforcement priorities, and suggest that changing supervisory behavior may be a targeted and cost-effective lever for reform.

6 Mechanisms: How Sergeant Effects Shape Officer Behavior

In this section, I consider the behavioral channels through which officers increase arrests in response to their sergeant. I focus on the impact of serious and low-level sergeant effects in order characterize how each style of enforcement more broadly shapes police interactions with the community. I focus the analysis on three questions. First, are sergeant effects concentrated in particular offense sub-types? Second, are additional arrests generated through officer- or civilian-initiated interactions? Third, are sergeant arrest effects associated with outcomes beside arrests, particularly harmful byproducts of policing such as use of force? Throughout, I use the event study specification from equation 6 and report DD estimates summarizing the average post-switch change in officer behavior relative to the last period before the switch, as described in Section 4.3.[38]

Arrest Subtypes—Serious and low-level crime categories obscure substantial heterogeneity in offense type and severity. Serious arrests include both violent and non-violent offenses, while low-level arrests span a range of offenses — such as disruptive behavior (e.g., disorderly conduct), drug use, and outstanding warrants — each with distinct social implications.

I estimate equation 6 on disaggregated sub-categories of serious and low-level arrests. Figure 5 plots estimates for the three most common sub-categories within each arrest type, which together account for over half of all arrests in each category. The left column plots coefficients on low-level sergeant effects for low-level crimes in green; the right column plots coefficients on serious sergeant effects for serious crimes in purple. Supplementary Appendix Figure A.17 presents mean-normalized versions of these coefficients on a common scale to facilitate direct comparisons.

Among low-level crimes, drug arrests exhibit the largest effects in both absolute and relative terms. A one standard deviation increase in a sergeant's low-level effect generates a large, sustained increase in drug arrests immediately following a switch. Event study

[38]Supplementary Appendix Figure A.31 shows that results are qualitatively similar when using the full analysis sample to estimate a version of equation 1 that replaces sergeant effects with their Bayes-shrunken estimates.

point estimates exceed 40% of the mean in each post-switch period, equivalent to 0.12–0.14 additional drug arrests per month. The DD estimate is 0.136, suggesting the post-switch coefficients are 47% larger than the pre-period mean on average. Supplementary Appendix Figure A.18 shows that the increase is driven predominantly by possession rather than sales in volume: possession accounts for roughly 85% of the absolute increase in drug arrests. Possession arrests account for the vast majority of drug arrests, however, and the proportional impact on drug sale arrests exceeds that of possession arrests relative to their respective means. Low-level sergeants also increase warrant and disorderly conduct arrests, though by smaller margins. Event study estimates suggest increases of 11–13% and 16–29% of their respective means. I reject equality of the mean-normalized post-switch coefficients between drug and warrant arrests at the 1% level, and between drug and disorderly conduct arrests at the 10% level (Supplementary Appendix Figure A.17).

Among serious crimes, sergeant effects increase arrests across each of the largest sub-categories. Domestic violence (DV) is the largest driver in absolute terms. Event study estimates indicate 0.16 additional DV arrests in the first post-switch month, a 36% increase relative to the mean, sustained throughout the sample period. DWI and theft arrests exhibit smaller absolute increases, however I cannot reject equality of the percentage-wise increase in arrests across serious sub-categories relative to their respective means, suggesting serious sergeant effects operate broadly across offense types rather than through a single dominant channel.

Supplementary Appendix Figures A.19 and A.20 report results for the opposite sergeant effect type alongside the focal type, as well as results for additional miscellaneous serious and low-level arrests and F-tests of equality of the post-switch coefficients between low-level and serious sergeant effects. The miscellaneous sub-categories exhibit similar patterns to the largest three categories: similar percentage-wise increases across serious crime types, and a meaningful increase for miscellaneous low-level offenses that is smaller than the increase for drug arrests in percentage-wise terms. There is some suggestive evidence of crime-specific tradeoffs that are not captured in the aggregate correlation of serious and low-level effects. DD estimates indicate that serious sergeant effects are associated with reductions in drug arrests and miscellaneous low-level arrests (e.g. public intoxication, weapons offenses, and municipal code violations) of around 8% of the mean, and increases in disorderly conduct and warrant arrests of 5% and 7% of the mean respectively. Only the reduction in miscellaneous low-level offenses is significant at the 5% level.[39] DD estimates for low-level sergeant effects

[39]Supplementary Figure A.18 shows suggestive evidence that the reduction in drug sales for serious sergeant effects is more precise than the reduction in drug arrests overall. The DD coefficient is 47% of the mean and statistically significant at the 5% level. However, drug sales represent a small portion of overall drug arrests.

indicate reductions in DV arrests of 8% and increases in miscellaneous non-violent serious offenses (e.g. burglary, fraud), though only the former is statistically significant. I cannot reject at the 5% level that serious and low-level sergeant effects have equal post-switch impacts on any serious arrest type outside of DV, though I can reject equality of post-switch coefficients across every low-level arrest type except warrants.

Enforcement Channels: Officer-Initiated versus Call-Initiated Arrests— An arrest occurs through one of two channels. Call-initiated arrests occur at 911 calls placed by civilians. Officer-initiated arrests occur during encounters the officer initiates, such as pedestrian or traffic stops. Comparing sergeant effects across the two channels shows how much of the additional enforcement is tied to public demand for police services versus discretionary officer enforcement.

I estimate equation 6 separately for the total number of arrests from each source, plotting coefficients in the left column of Figure 6. Sergeants who induce more serious arrests raise call-initiated arrests. They generate no meaningful change in officer-initiated arrests. Call-initiated arrests increase by 9.5% of the mean in the first post-switch period, sustained throughout the sample period. Sergeants who induce more low-level arrests raise arrests from both sources. Call-initiated arrests increase by 7.6% and officer-initiated arrests by 19% relative to the mean in the first post-switch period, with both effects persisting throughout.[40] I can reject that the effect on officer-initiated arrests is equal between serious and low-level sergeants at the 1% level (p = 0.003). Low-level sergeant effects therefore uniquely raise arrests from officer-initiated, discretionary encounters.

The mechanism through which call-initiated arrests vary is less straightforward than officer-initiated arrests. Sergeants may affect how officers exercise discretion at calls, or the total number of calls officers respond to. I report event study estimates for call arrest probability and calls answered in the right column of Figure 6. Low-level sergeant effects raise call arrest probability by 0.5 pp, or 8.2% relative to the mean. They have no meaningful effect on the number of calls answered. Serious sergeant effects operate on the other margin. Officers answer 1.7 more calls per month, a 2.3% increase relative to the mean. Call arrest probability does not change. I reject equality of the post-switch coefficients across sergeant types at the 5% level for both outcomes.

Call arrest probability could rise for two reasons. Officers may become more willing to arrest at a given call, or they may select into more severe calls. I estimate sergeant impacts on call arrest probability separately by call severity in Supplementary Appendix Figure A.22,

[40] Pre-trends for officer-initiated arrests are not perfectly flat; however, the sustained post-switch increase remains statistically significant and economically meaningful when including an event-specific linear time trend (Supplementary Appendix Figure A.21).

classifying a call as more or less severe based on the arrest rate of its dispatch code. Low-level sergeant effects are positive across both severity categories and strongest for the least severe calls. This pattern is consistent with a change in officer discretion rather than a shift in call composition. Serious sergeant effects have no meaningful impact on call arrest probability at either severity level.

I interpret the serious sergeant-induced increase in call volume as a change in officer behavior, not local crime conditions. Four pieces of evidence support this interpretation. First, the increase is robust to fixed effects that absorb common time-varying shocks to officer call volume at the department, division-by-watch, and sector-by-watch levels, including any changes in crime conditions operating at those levels (Supplementary Appendix Figure A.23). Second, for officer-driven switches, serious sergeant effects do not predict increased call volume for incumbent officers in the switching officer's new unit, whose sergeant does not change (Supplementary Appendix Figure A.24). If crime conditions were driving the increase, incumbents facing the same environment would respond similarly. Third, serious sergeant effects do not increase total 911 call volume within a unit's assigned sector-watch for sergeant-driven switches (Supplementary Appendix H). Fourth, the increase holds when restricting to sergeant-driven switches, where the officer remains in the same assignment (Supplementary Appendix Figure A.25). Because the officer's location and shift are unchanged, the increase cannot reflect a change in the crime conditions they face.

Sergeants may affect call volume by approving overtime. Overtime reflects hours the officer chooses to work beyond their scheduled shift. Therefore additional overtime calls would reflect greater 911 response effort rather than increased availability for dispatch. I classify a call as taken during overtime if it occurs outside the officer's regular shift hours on a day the officer receives an overtime payment, using DPD shift and overtime payment records (Chalfin and Goncalves, 2026).[41] In Supplementary Appendix Figure A.26, I show that overtime calls account for roughly a third of the serious-sergeant-induced increase in call volume.[42] The increase is larger relative to the mean for overtime calls than for on-shift calls: 15% based on the DD estimate versus 1.6%. Effort is therefore an important margin through which serious sergeant effects expand call volume.

I cannot distinguish effort from availability for the remaining increase in calls answered during scheduled shifts. Officers may volunteer for these calls, which would reflect effort.

[41]Requiring an overtime payment on the day of the call distinguishes overtime from a temporary change in the officer's assigned shift. Consistent with this interpretation, I find that serious sergeant effects increase the total number of overtime hours that officers record by 9% following a switch (Supplementary Appendix Figure A.28). Low-level sergeant effects also increase overtime hours, though these additional hours do not translate into more call responses.

[42]Point estimates are less precise for on-shift calls, but I cannot reject equality of the post-switch trends between the two categories (Supplementary Appendix Figure A.27).

They may also become less engaged in low-level activity, leaving them available for more dispatch assignments. I do not observe the reason for call assignment, so I cannot separate these explanations. If officers volunteer for calls that other officers could have taken, the observed gains in call volume may reflect a reallocation of calls between officers rather than an increase in total service provision.[43] As such, the welfare implications of the serious-sergeant induced increase in call volume are unclear. To the extent that the increase operates through additional overtime hours, the additional pay is costly to the department, and the benefits of any net enforcement gains that result would need to be considered alongside their costs.

Taken together, the results suggest low-level sergeant effects alter officers' discretionary behavior, not only by increasing officer-initiated arrests but also by lowering the threshold for arrest during civilian-initiated encounters. Serious sergeant effects, by contrast, raise call-initiated arrests by expanding the number of 911 calls officers answer, in part through additional response effort via overtime hours. This distinction underscores how different managerial styles shape distinct forms of enforcement. Low-level sergeant effects exert influence over the most discretionary policing decisions, while serious sergeant effects increase officer exposure to civilian-initiated interactions.

Secondary Police Outcomes—Beyond arrest counts, sergeants may affect rarer but more directly costly outcomes — use of force and misconduct. Police violence imposes high costs on victims and communities (Ang, 2020), while misconduct erodes public trust and generates costly litigation (Ouss and Rappaport, 2020). I estimate changes in monthly use of force incidents and civilian complaint filings using the event study specification, reporting results in Figure 7.

Moving to a sergeant with larger low-level effects generates additional use of force incidents: the DD estimate is 0.017, roughly 12% of the mean, significant at the 5% level. Event study coefficients suggest an immediate and sustained increase following a switch. For serious sergeant effects, DD results and event study coefficients are near zero or slightly negative, though standard errors are large enough that I cannot reject that switching to a sergeant with larger serious arrest effects has the same impact on use of force. Since most uses of force occur during the course of an arrest (Weisburst, 2019), it is not clear whether additional

[43] If serious sergeant effects cause officers to be more available, they may respond to calls more quickly. In Supplementary Appendix Figure A.29, I estimate the impact of sergeant effects on average monthly response times. I do not find statistically significant changes in response times in either direction for either dimension of sergeant effects. This suggests dispatchers are not matching officers to calls more quickly when the officers work with sergeants who increase serious enforcement. I estimate these models with year-month fixed effects, since a large share of the CAD timestamp data used to construct response times is missing in different periods of the sample. Supplementary Appendix F.3 shows that the main event study results are robust to the inclusion of year-month fixed effects.

incidents reflect more arrests or officers acting more violently conditional on the additional arrests. I therefore estimate effects on use of force *per arrest* in Supplementary Appendix Figure A.30.[44] Event study coefficients for low-level sergeant effects cluster around zero with no clear post-switch trend, and the DD estimate is small (less than 3% of the mean) and insignificant. Low-level sergeant effects therefore increase use of force proportionally with arrests rather than making officers more violent per interaction. For serious sergeant effects, the impact on use of force per arrest is negative across post-switch periods, but DD and event study coefficients are statistically insignificant. For complaints, DD estimates are less than 3.5% of the mean for both sergeant types and show no clear post-switch change. Complaints are rare — fewer than 0.03 per officer-month — so I am likely underpowered to detect small but meaningful changes.

In Supplementary Appendix E, I examine two additional secondary outcomes: arrest quality, measured by conviction rates, and racial disparities in arrests. I find no evidence that sergeant-induced arrests are of lower quality. Low-level sergeant effects are associated with suggestive increases in low-level conviction rates, particularly for drug offenses, which parallels Weisburst (2024)'s finding that officers with higher misdemeanor arrest propensities have higher conviction rates for those arrests. On race, I find no strong evidence that either sergeant type meaningfully shifts the racial distribution of arrests. Black arrestees experience the largest absolute increases under both sergeant types, but this reflects their larger baseline share of arrests. Proportional increases relative to the mean are similar across racial groups.

Sergeant effects on arrests may be correlated with other important police interactions — stops, searches, and citations — which can be antecedents to arrest and generate meaningful costs to civilians (Mello, 2021; Bacher-Hicks and De La Campa, 2020b). Stops and searches may represent important channels through which sergeants increase officer-initiated low-level arrests. I lack detailed data on stops, searches, and traffic enforcement that would allow me to investigate these outcomes.

Robustness to Alternative Samples—In Supplementary Appendix F, I present results that probe the robustness of my primary event study findings to different samples of officer switching events. To ensure that the results are not driven by differences between the select sample of sergeant switches in the event study and the full officer-month panel, I estimate the event studies using all officer-sergeant switches, including those that had other switches in the sample period, but removing any months that are contaminated by the other switch. Qualitative results are similar across all outcomes, and estimates are generally more

[44] I separately estimate event studies for arrests and use of force, add the coefficients to their respective pre-period means, and calculate the implied change in the use-of-force rate relative to the pre-period ratio, with standard errors computed via the Delta method.

precise due to the larger sample size.

To ensure my findings are not driven by compositional changes in my event study sample over time, I also present results using a balanced panel. This panel includes only 16% of all switching events, so estimates are less precise and switching dynamics may be considerably less representative of those that drive the main results. The qualitative results are broadly consistent with the baseline. Point estimates for the impact of low-level sergeant effects on use of force are similar in magnitude, but neither the event study nor DD coefficients reach statistical significance in this smaller sample, likely reflecting limited power given the rarity of use of force incidents.

How do sergeants change behaviors?—In Supplementary Appendix I, I examine a few of the institutional mechanisms through which sergeants may influence officer behavior despite lacking access to many standard personnel tools, such as performance pay and termination authority. Using an event study around sergeants moving between units, with sergeant arrest and call behavior as outcomes, I find evidence that sergeants who induce low-level arrests lead by example: they make significantly more arrests themselves, almost entirely for low-level offenses. Moreover, there is also suggestive evidence that they engage in more direct monitoring by answering calls alongside their subordinates. Serious sergeant effects are not significantly associated with either leading by example or direct monitoring, though estimates are too noisy to rule out effects of similar magnitude to those documented for low-level sergeant effects. Consistent with overtime approval being a key mechanism through which sergeants induce serious arrests, I show that serious sergeant effects increase overtime arrests. Together, these findings suggest that sergeants leverage their institutional roles to generate behavioral changes in officers despite operating in a setting with limited high-powered incentives.

Taken together, the findings in this section show that sergeants who increase serious arrests and those who increase low-level arrests modify distinct channels of officer behavior. Low-level sergeant effects concentrate in the margins where officers exercise the most discretion. Officers make more officer-initiated arrests, they become more likely to arrest at a given civilian call, and the additional low-level arrests are concentrated in drug offenses. They are also involved in more uses of force, in proportion to the additional arrests they make. Serious sergeant effects move none of these discretionary margins. Officers instead answer more calls, and the marginal serious arrests are spread proportionally across serious offense types. These behavioral patterns support the interpretation that sergeant effects reflect differing enforcement priorities, and the choices of individual supervisors can generate inconsistencies in how law enforcement policies are implemented on the ground.

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 who induce serious crime enforcement increase the number of 911 calls their officers respond to, while those who induce low-level enforcement encourage proactive, discretionary arrests, typically for minor drug offenses. Sergeant effects on serious and low-level enforcement are not systematically related, 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 interpret, 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 suggests that sergeants may be able to scale back one type of enforcement without compromising the other. 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 (Heaton, 2010).

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. (2025). The effect of field training officers on police use of force. *American Economic Review*. Forthcoming.

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.

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. (2020). The effects of police violence on inner-city students*. *The Quarterly Journal of Economics*, 136:115–168.

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.

Beuermann, D. W., Jackson, C. K., Navarro-Sola, L., and Pardo, F. (2023). What is a good school, and can parents tell? evidence on the multidimensionality of school output. *The Review of Economic Studies*, 90:65–101.

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.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225:200–230.

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. (2026). Data and code for: Professional motivations in the public sector: Evidence from police officers. Nashville, TN: American Economic Association [publisher], 2026. Ann Arbor, MI: ICPSR [distributor], 2026-07-07.

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*.

Corman, H. and Mocan, N. (2005). Carrots, sticks, and broken windows. *The Journal of Law and Economics*, 48(1):235–266.

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.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110:2964–2996.

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.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of geographic variation in health care: Evidence from patient migration. *Quarterly Journal of Economics*, 131:1681–1726.

Frake, J. and Harmon, D. (2023). Intergenerational transmission of organizational misconduct: Evidence from the chicago police department. *Working Paper.*

Frakes, M. and Gruber, J. (2025). The effect of provider diversity on racial health disparities: Evidence from the military. *Review of Economic Studies.*

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.

Goncalves, F. and Mello, S. (2021). A few bad apples? racial bias in policing. *American Economic Review*, 111:1406–1441.

Gonçalves, F. and Mello, S. (2023). Police discretion and public safety. *NBER Working Paper 31678.*

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.*

Heaton, P. (2010). Understanding the effects of antiprofiling policies. *The Journal of Law and Economics*, 53(1):29–64.

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.

Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*, 95:1096–1116.

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. (2021). Fines and financial wellbeing.

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*.

Molitor, D. (2018). The evolution of physician practice styles: Evidence from cardiologist migration. *American Economic Journal: Economic Policy*, 10:326–356.

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

Ouss, A. and Rappaport, J. (2020). Is police behavior getting worse? data selection and the measurement of policing harms. *The Journal of Legal Studies*, 49:153–198.

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.

Piza, E. L., Connealy, N. T., Reid, S. A., and Palermo, C. M. (2025). Using automated vehicle locator data to classify discretionary police patrol across space. *Journal of Criminal Justice*, 101:102543.

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*.

Rozema, K. and Schanzenbach, M. (2019). Good cop, bad cop: Using civilian allegations to predict police misconduct. *American Economic Journal: Economic Policy*, 11:225–268.

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.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225:175–199.

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, D., Groff, E. R., Jones, G., Cave, B., Amendola, K. L., Yang, S.-M., and Emison, R. F. (2015). The dallas patrol management experiment: can avl technologies be used to harness unallocated patrol time for crime prevention? *Journal of experimental criminology*, 11(3):367–391.

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

Weisburst, E. K. (2019). Police use of force as an extension of arrests: Examining disparities across civilian and officer race. *AEA Papers and Proceedings*, 109:152–156.

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.

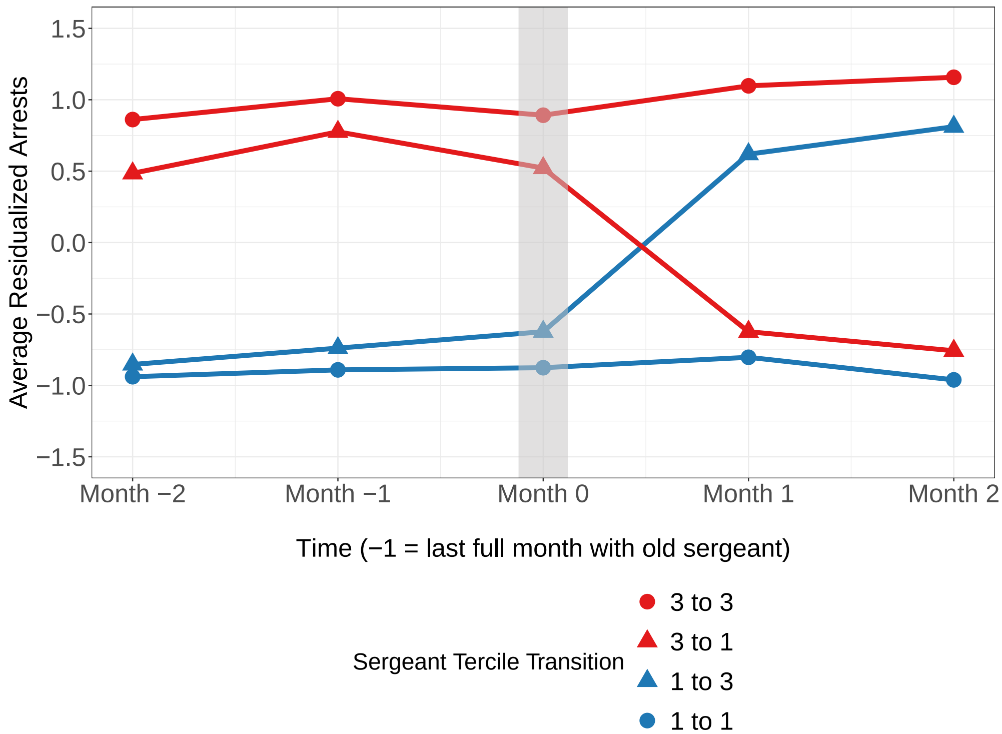
Tables and Figures

Table 1: 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%	11.136	100.00%
$Var(\psi)$	0.559	5.01%	0.382	3.43%	0.379	3.40%	0.379	3.41%
$Var(\theta)$	8.906	79.86%	8.002	71.75%	8.097	72.60%	8.068	72.45%
$Cov(\psi, \theta)$	-0.168	-1.51%	-0.157	-1.41%	-0.0592	-0.53%	-0.0592	-0.53%
$Var(\psi + \theta)$	9.129	81.86%	8.071	72.37%	8.357	74.93%	8.33	74.79%
N sergeants	347		347		347		344	
N officers	1805		1805		1805		1622	

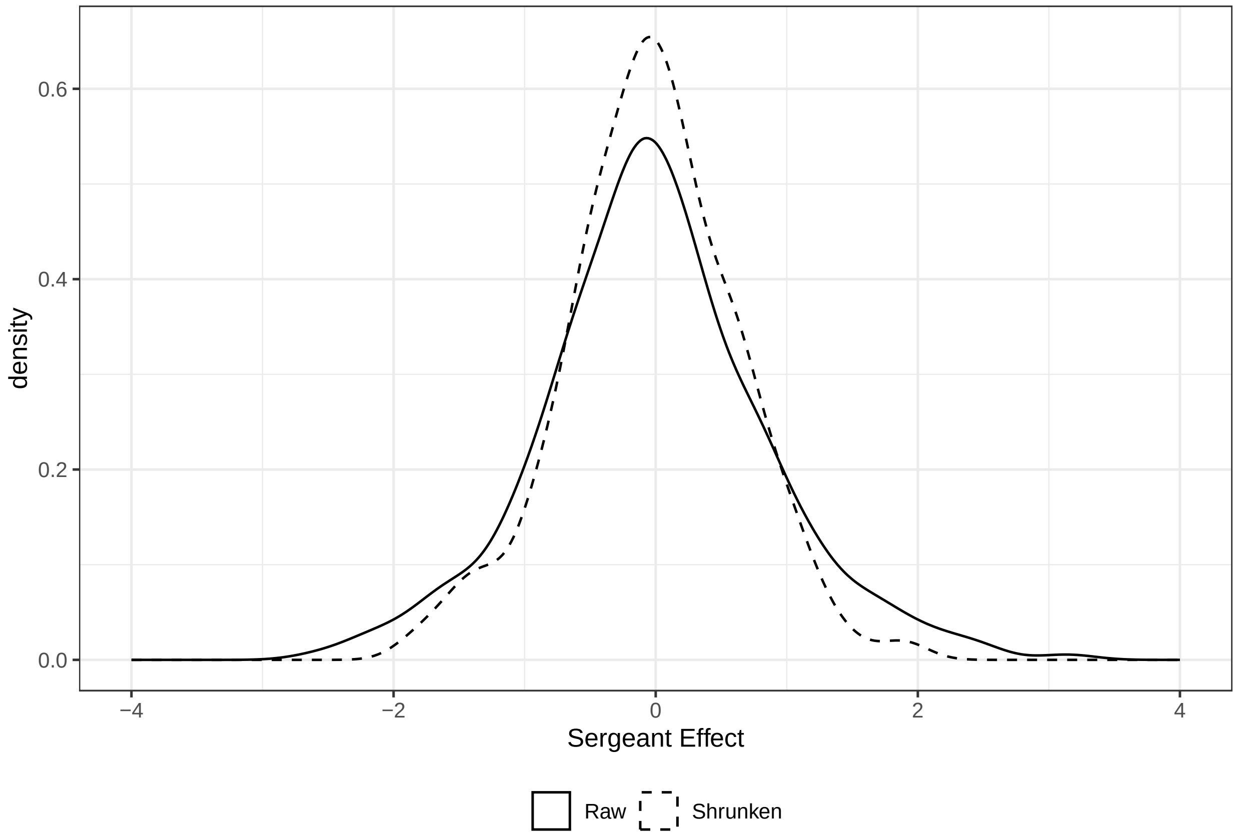
Notes: This table presents the variance decompositions described in equation 2. As described in Section 4, y^* is the number of monthly arrests, residualized on sector-watch, day-off group, and a second-degree polynomial of tenure; ψ is the sergeant fixed effect; θ is the officer fixed effect. All statistics 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. Sergeant and officer fixed effects are estimated using equation 1, which includes controls for a second-degree polynomial of officer tenure, and sector-watch and day-off group fixed effects.

Figure 1: Event Study Around Sergeant Switch



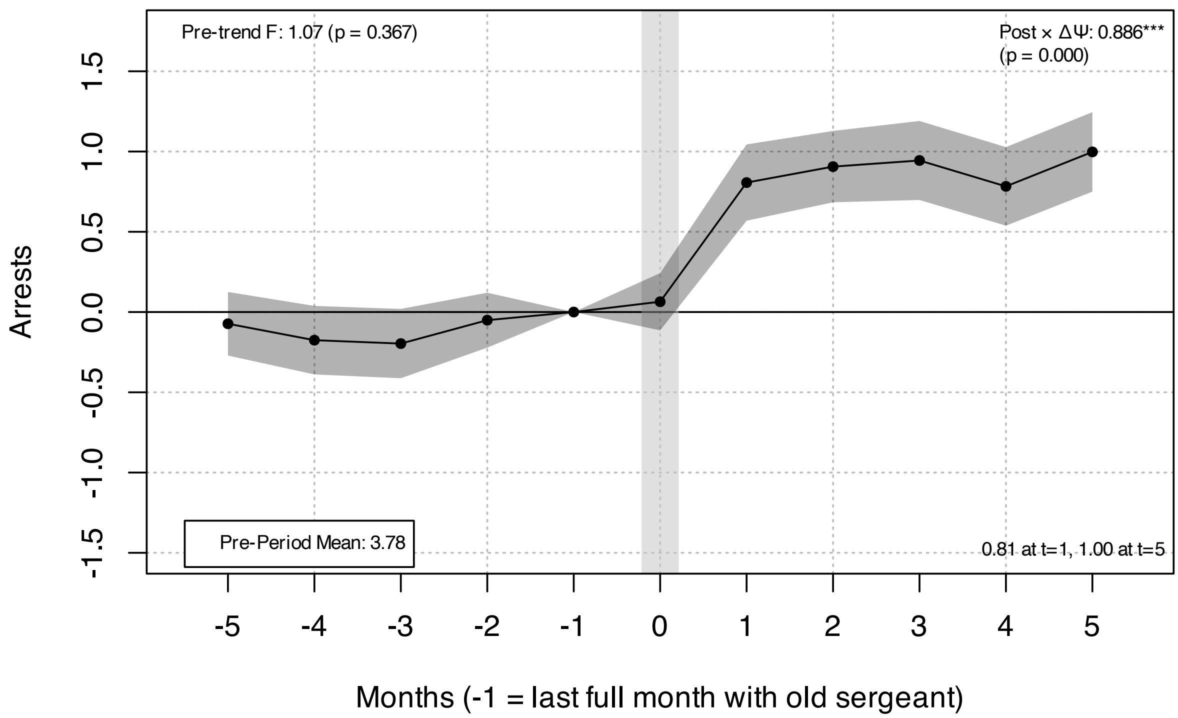
Notes: This figure plots the average number of arrests made by officers in the months around receiving a new sergeant by the magnitude of the sergeant change. In particular, I group sergeants into terciles according to the average number of residual arrests made by their officers throughout the sample. 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." The figure only includes transitions within or between the top (tercile 3) and bottom (tercile 1) terciles. 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



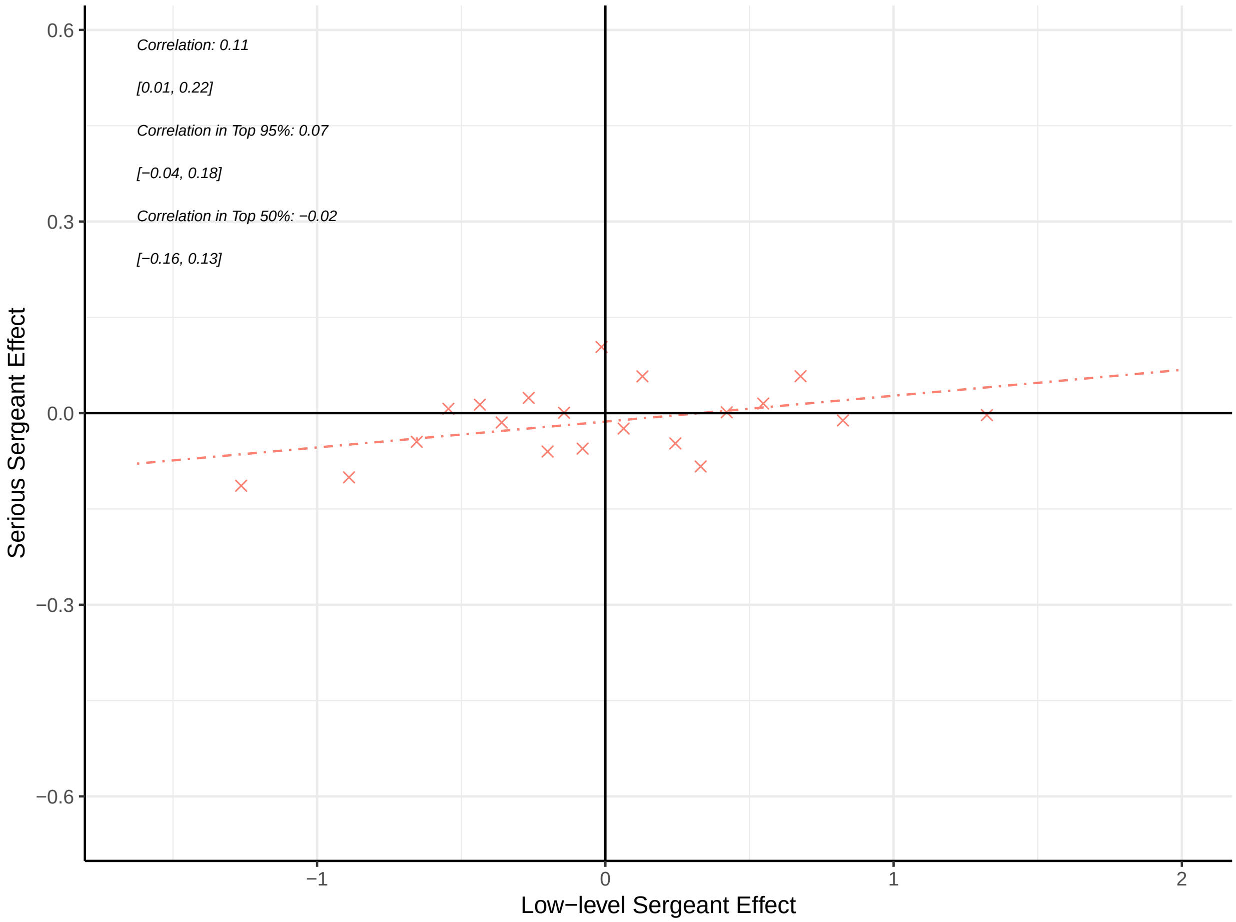
Notes: This figure plots the sergeant effects estimated using the sergeant fixed effects in equation 1, which includes controls for a second-degree polynomial of officer tenure, as well as officer, sector-watch, and day-off group fixed effects. Sergeant effects are interpreted as the number of monthly 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: Event Study: Total Arrests



Notes: Each point estimate (with 95% confidence interval) represents $\hat{\pi}_1^k$ from equation 4 — the effect on officer arrests at period k relative to the switch of moving to a sergeant who induces one more arrest per month. Month 0 is the switch month; month -1, the last full month with the previous sergeant, is the reference period. The pre-trend F-statistic (top left) tests joint equality to zero for periods -5 through -2. The top right reports the $Post \times \Delta\hat{\psi}$ coefficient and p-value from the DD specification in equation 5, which summarizes the average post-switch coefficients. The bottom right lists point estimates at selected months. The model is estimated on the event study sample described in Table B.1, with standard errors clustered at the officer level. Controls include a second-degree polynomial of officer tenure, non-interacted period indicators, and sector-watch, officer-event, and day-off group fixed effects.

Figure 4: Relationship between low-level and serious sergeant effects



Notes: This figure plots sergeants' serious effects against their low-level effects. 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 estimated using sergeant fixed effects in models that include a second-degree polynomial of officer tenure and officer, sector-watch, and day-off group fixed effects. Sergeant effects 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. Each point represents an equal-sized bin of sergeants. The dashed line is a linear fit across all sergeants.

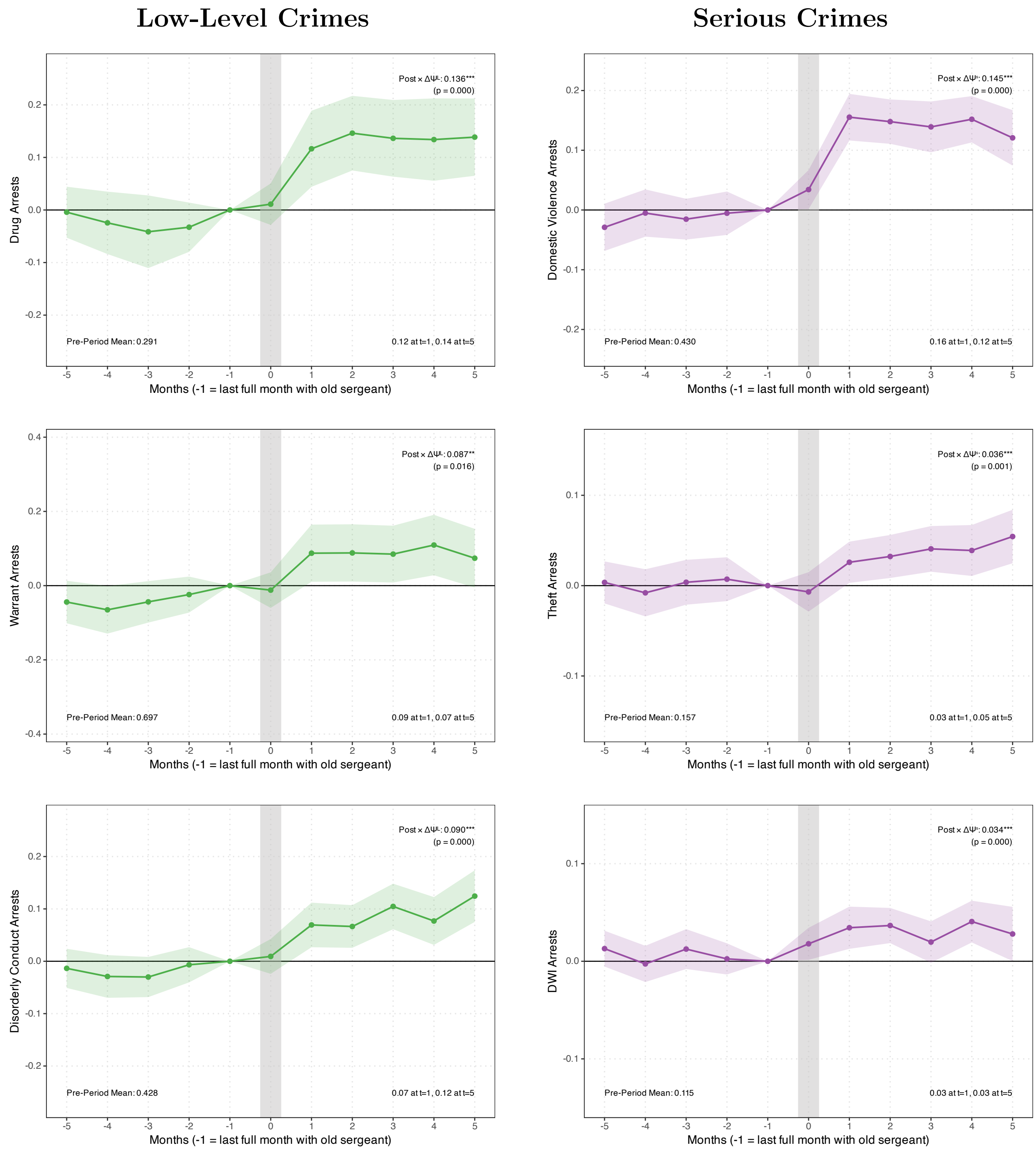


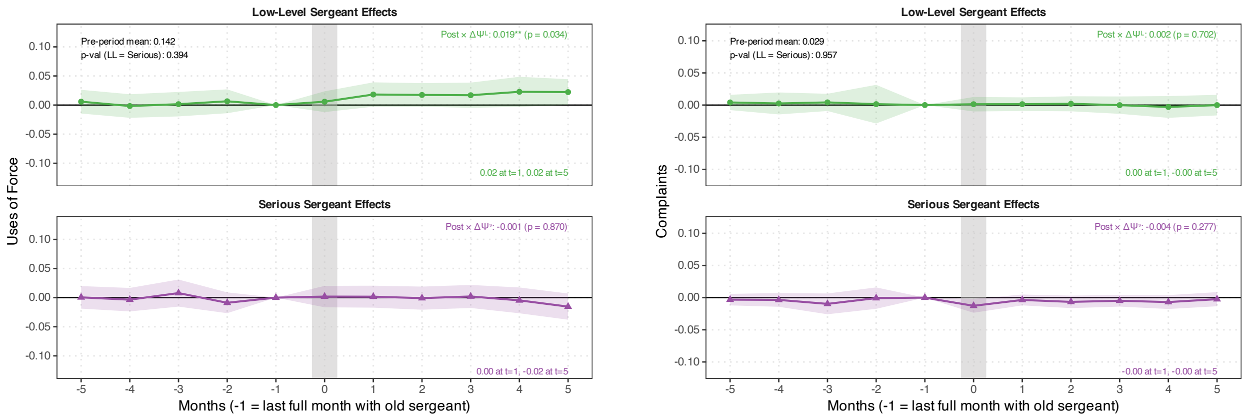
Figure 5: Sergeant Effects by Arrest Sub-Categories

Notes: This figure reports results from officer-level event studies around sergeant switches for the three most common sub-categories of serious and low-level arrests, estimated using equation 6. Each figure is a separate regression. The left column plots point estimates and 95% confidence intervals on the interaction between period indicators and the move-induced change in low-level sergeant effects in regressions with low-level arrest sub-types as outcomes; the right column plots the analogous interaction for serious sergeant effects in regressions using serious arrest sub-types as outcomes. In each panel, the top right reports the associated DD estimate, obtained by replacing the post-switch event-time indicators in equation 6 with a single binary indicator, *Post*, as described in Section 4.3; the bottom left reports the pre-period outcome mean. Standard errors are clustered at the officer level. All regressions include a second-degree polynomial of officer tenure, non-interacted period indicators, and sector-watch, officer-event, and day-off group fixed effects, and are estimated on the event study sample described in Table B.1.

Figure 6: Sergeant Effects on Police Interactions

Notes: This figure reports results from officer-level event studies around sergeant switches for arrests by source (left column) and 911 call activity (right column), estimated using equation 6. Call arrest probability is the share of calls answered within a month that result in arrest. For each outcome, the top panel (green) and bottom panel (purple) plot point estimates and 95% confidence intervals on the interaction between period indicators and the move-induced change in low-level and serious sergeant effects, respectively. The top left of the top panel reports the p-value from an F-test of equality between serious and low-level interaction coefficients across all post-switch periods jointly. In each panel, the top right reports the associated DD estimate, obtained by replacing the post-switch event-time indicators in equation 6 with a single binary indicator, *Post*, as described in Section 4.3; the bottom right lists point estimates at selected months. Standard errors clustered at the officer level. Controls include a second-degree polynomial of officer tenure, non-interacted period indicators, and sector-watch, officer-event, and day-off group fixed effects. Regressions are estimated on the event study sample described in Table B.1.

Figure 7: Sergeant Effects on Secondary Police Outcomes



Notes: This figure reports results from the officer-level event study around sergeant switches for use of force incidents and complaints, estimated using equation 6. For each outcome, the top panel (green) and bottom panel (purple) plot point estimates and 95% confidence intervals on the interaction between period indicators and the move-induced change in low-level and serious sergeant effects, respectively. The top left of the top panel reports the p-value from an F-test of equality between serious and low-level interaction coefficients across all post-switch periods jointly. In each panel, the top right reports the associated DD estimate, obtained by replacing the post-switch event-time indicators in equation 6 with a single binary indicator, *Post*, as described in Section 4.3; the bottom right lists point estimates at selected months. Standard errors clustered at the officer level. Controls include a second-degree polynomial of officer tenure, non-interacted period indicators, and sector-watch, officer-event, and day-off group fixed effects; regressions are estimated on the event study sample described in Table B.1.