Middle Managers in the Public Sector: Evidence from Policing*

Elisa Facchetti†, Arianna Ornaghi‡, Eddy Zou§

September 4, 2026

Abstract

How much can middle managers matter when budgets, staffing, and priorities are set centrally? We study this question in the London Metropolitan Police Service, focusing on Borough Commanders, the middle managers who lead territorial divisions. We construct a panel of commanders' careers and link their assignments to monthly data on investigation outcomes, crime, and police workforce. Exploiting manager mobility across divisions in an AKM-style framework, we estimate that managers account for about one quarter of the variance in log clearances per worker. More effective managers raise clearances across offense types, increase both police-issued sanctions and court convictions, and expand proactive enforcement without increasing frontline headcount. These effects emerge quickly after a manager arrives and largely dissipate after departure. Managerial effectiveness also shapes organizational responses to policy change: stronger commanders mitigate the increase in violent crime following police station closures and achieve faster response times following body-worn camera adoption relative to less well-managed divisions.

*We thank Charles Angelucci, Imran Rasul, Pedro Souza, Guo Xu, and all participants to the Bocconi FIBRE workshop, RFBerlin Labour Economists Workshop, Bank of Italy ESPI Conference, the 16th Transatlantic Workshop on the Economics of Crime, the LSE-Chicago Crime Conference, the BSF Public Economics, CESifo Summer Institute Managing the State, and SIOE. The views expressed here are those of the authors alone. We gratefully acknowledge financial support from EIEF.

†University of Rome Tor Vergata, EIEF, IFS, IZA

‡Hertie School, CEPR

§London School of Economics, IFS, STICERD

1 Introduction

Police forces face growing pressure to improve performance amid declining public trust and constrained resources. To address these challenges, the policy and academic debate has largely focused on external levers, such as policing strategies, oversight, and the allocation of resources (Owens and Ba, 2021). Much less attention has been devoted to the internal organization of police forces and, in particular, to the managers responsible for translating centrally determined resources and priorities into local police activity.

The extent to which managers can shape police performance is not obvious a priori. Like other public organizations, police forces have budget and staffing levels that are largely fixed in the short run, and their managers lack the standard incentive levers available in the private sector such as performance pay or control over hiring and dismissal. At the same time, police managers retain considerable operational discretion over how existing resources are organized and deployed, and existing work shows that these operational decisions are important determinants of police output (Blanes i Vidal and Kirchmaier, 2018; Adda et al., 2014; Draca et al., 2011). In this paper, we ask whether managers matter for police performance and, if they do, what effective managers do differently.

We study these questions in the context of the Metropolitan Police Service of London (MPS), focusing on Borough Commanders, the middle managers who lead territorial police divisions. Borough Commanders are senior officers, with more than 25 years of service in the MPS on average, who sit between the central leadership and frontline teams. They operate within centrally determined budgets and strategic priorities, but are responsible for the oversight, resource allocation, and day-to-day operational decisions of one of London's 32 territorial divisions. In other words, they are in charge of making precisely the high-stakes operational decisions that are key for policing outcomes.

To study their impact, we assemble a novel dataset covering the universe of Borough Commanders in the MPS from 2005 to 2019, combining Freedom of Information requests with archival web sources. Over this period, the MPS employed 164 Borough Commanders across 32 territorial divisions. We link the organizational data to the universe of recorded criminal incidents and investigation outcomes between 2012 and 2019, which allows us to measure division-by-month police output and follow cases to their final outcomes. Our empirical analysis focuses on this 2012–2019 window, during which the MPS appointed 103 Borough Commanders, with an average tenure of about 25 months.

We begin by estimating an AKM-style model (Abowd et al., 1999) that decomposes division-level police output into a territorial-division component, a manager component, and a time component. Our main objects of interest are manager fixed effects, which capture the "portable component" of performance ability across police divisions. The outcome we focus on is *clearances per worker*, defined as the number of cleared incidents in a division divided by total division workforce. This value-added-type measure

captures police performance and is the outcome for which managerial fixed effects are most informative in our setting. Relative to specifications with time and division fixed effects only, adding manager fixed effects increases the adjusted R^2 by about 21%, while the corresponding increases for total clearances are much smaller, and for total recorded crime nearly negligible.

Estimating manager and division effects poses several empirical and identification challenges. First, manager and division effects are separately identified only through managers who move across divisions, and only within connected sets of divisions and managers linked by such moves. For 2012-2019, we identify one large connected set containing 93 managers and 29 territorial divisions, which forms the basis of our analysis. Second, with relatively few observations per manager spell, the estimated fixed effects contain sampling error. We address this in two ways: we apply an Empirical Bayes shrinkage procedure to correct the individual effects, and, in the variance decomposition, we correct for the resulting limited mobility bias using several procedures proposed in the literature. Finally, the approach requires that manager assignment be independent of transitory division-specific shocks and that manager effects be additively separable from division effects, ruling out sorting on match quality. We provide diagnostic evidence supporting both conditions. Event-study estimates around transitions show no differential pre-trends, in line with assignments following a centralized process that prioritizes operational needs over individual preferences. In addition, performance changes are approximately symmetric when divisions switch from top to bottom tercile managers and vice versa.

Managers matter for police performance: under our preferred bias-corrected estimate based on the covariance-shrinkage procedure of Best et al. (2023), manager effects account for about one quarter of the total variance in log clearances per worker. We also estimate a sizable negative covariance between manager and division fixed effects, accounting for about 13% of total variance. Consistent with the institutional design of the MPS Posting Panel, which explicitly subordinates individual preferences to divisions' operational needs, this negative assortative matching is consistent with an allocation of managerial talent toward more challenging divisions.

In line with the large share of the variance explained by managers, we document substantial heterogeneity in managerial effectiveness. A one standard deviation increase in manager fixed effects is associated with approximately 12.6% higher (log) clearances per worker. The effects of managerial transitions materialize quickly: clearances per worker show no differential trends in the six months before a new manager comes into the division, but begin to rise (or decline) within three months of a higher-quality manager's arrival. The performance associated with managers is also relatively short-lived. Once we condition on the quality of the incoming manager, only about one third of the outgoing-manager effect remains over the first six months of the successor's spell, with

the estimates approaching complete disappearance toward the end of the event window. Observable manager characteristics and prior career paths have limited power to predict effectiveness, and the dimension of managerial performance captured by our fixed effects is not significantly associated with subsequent promotion.

We next characterize the productivity gains associated with effective managers. Using a stacked event-study design around manager transitions, we estimate that effective managers (defined as being one standard deviation higher in terms of managerial fixed effect) raise total clearances by about 7.4%, while total workforce changes very little. The gains are broad-based: clearances increase for both property and violent offenses and drug offenses, and both police-issued sanctions and court convictions rise. This suggests that effective managers improve both investigative productivity and engage in more proactive policing. Instead, reported property and violent crime is essentially unchanged. These improvements are achieved without expanding resources: front-line officer numbers remain essentially flat, while civilian back-office staff decline by about 3%. We find suggestive evidence that these increases in investigative productivity are larger in wards experiencing higher crime levels before the transition.

We also ask whether managerial effectiveness goes beyond clearances, by investigating whether they shape how divisions respond to changes in their environment. Specifically, we study two policy episodes that affect police divisions in different ways: the closure of more than two thirds of London's police stations (see Facchetti (2025) for more details) and the introduction of body-worn cameras. First, we find that following a police station closure, violent crime rises substantially less in areas led by above-median effectiveness commanders than in those led by below-median commanders, indicating that stronger managers are better able to absorb the loss of local infrastructure. Second, during the staggered rollout across divisions of body-worn cameras, divisions with low-performing managers experience a sharp increase in response times, while this deterioration is offset in divisions led by high-performing managers. These findings show that managerial capacity matters not only for average productivity, but also for how organizations adapt when their operating environment changes.

Taken together, these results help us characterize what effective managers do (and do not) in policing. Effective managers add value through ongoing operational judgment. They improve performance primarily through how existing resources are organized and deployed, rather than through increases in inputs. They also help their divisions absorb change, cushioning adverse shocks and capitalizing on new technologies. What they do not do is durably transform their divisions: performance rises with their arrival and falls with their departure, suggesting that much of managerial value remains tied to the individual rather than becoming durable organizational capital.

This paper contributes to two strands of literature. First, we contribute to the literature on the determinants of the police production function (Chalfin and McCrary, 2017)

by identifying managerial effectiveness as an important input into police performance. Existing work has studied how police outcomes respond to technology (Garicano and Heaton, 2010; Mastrobuoni, 2020), policing strategies, and infrastructure (Adda et al., 2014; Blanes i Vidal and Kirchmaier, 2018; Facchetti, 2025). A smaller literature studies management within police organizations. Kapustin et al. (2022) and Ponce de Leon and Rizzotti (2024) study the effect of management practices and technology-based managerial tools on crime. Closest to our work is Smith (2025), who studies how sergeants affect impact arrests using data from the Dallas Police Department. We distinguish ourselves by studying a different, higher level of the police hierarchy: rather than focusing on frontline supervisors, we study the role of middle managers who oversee entire police divisions. Middle managers explain a much larger share of the variation in policing output (25% rather than 3.4%), suggesting that the strategic decisions taken at more aggregate levels might be particularly consequential for police performance.

Second, we contribute to the growing literature on managers in the public sector, which has documented substantial managerial heterogeneity in public administration (Fenizia, 2022), public procurement (Best et al., 2023), schools (Muñoz and Prem, 2024), and hospitals (Janke et al., 2019; Muñoz and Otero, 2025).[1] Our contribution to this literature is threefold. First, we study middle managers: a layer of the hierarchy where, even though recruitment, pay, and budgets are set centrally, managers retain substantial discretion over day-to-day operations and the allocation of local resources, which differentiates them from frontline managers. Since managerial autonomy is an important determinant of performance in public bureaucracies (Rasul and Rogger, 2018; Bandiera et al., 2021), this discretion may help explain the comparatively large manager effects we estimate. At the same time, relative to studies of top managers such as hospital CEOs or school principals, where each manager effectively comes attached to their own institution, this focus on middle managers means all managers operate under the same central policies, incentives, and organizational culture, so that variation across them is more plausibly attributable to the managers themselves. Finally, the richness of our data allows us to unpack the productivity effect and characterize what effective managers actually do: which clearances rise, whether their quality holds up, and how resources are deployed. It also lets us study how managers respond to shocks to the status quo. We find that effective managers, as identified by our framework, are precisely those whose divisions best absorb adverse shocks and capitalize on new technologies.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting of territorial policing in London and the organizational role of Borough Commanders. Section 3 presents the construction of the middle-manager panel and the linkage

[1] A related literature on manager effects was first established in private-sector organizations (e.g. Bertrand and Schoar, 2003; Bandiera et al., 2020; Metcalfe et al., 2023; Minni, 2026). These settings feature market and performance incentives and greater managerial discretion over hiring, firing, compensation, and worker allocation.

to case-level police records, workforce information, and division outcomes. Section 4 outlines the empirical framework, discusses identification challenges, and presents our estimates of manager fixed effects. Section 5 investigates the mechanisms behind managerial effectiveness, examining patterns of mobility, heterogeneity across offense types, proactive policing, workforce allocation, and how managerial effectiveness shapes divisions' responses to external policy shocks such as police station closures and the rollout of body-worn cameras. Section 6 concludes.

2 Institutional setting

2.1 Territorial policing in the MPS

Policing in England is organized across 39 police forces, each responsible for law enforcement in their area. Each of them operates with considerable autonomy, while funding combines central government grants, allocated by the Home Office and agreed annually in Parliament, and locally raised revenues from council tax precepts set by the police force leadership. We focus our analysis on the Metropolitan Police Service (MPS), that is responsible for law enforcement across Greater London, serving approximately 8.9 million residents.

Within the MPS, territorial policing is delivered through territorial divisions, the primary geographic units for operational management and resource deployment, shown in Figure 1.[2] Each division is responsible for policing a defined area of London and for delivering core functions including neighborhood policing, incident response, and local crime-control activities.

Territorial divisions operate within an organizational structure in which broad strategic priorities and budgets are determined centrally. At the same time, they retain substantial discretion over how policing is organized and delivered locally. In particular, decisions concerning day-to-day deployment, neighborhood priorities, and the implementation of operational strategies are made at the division level.

[2]Throughout most of the study period, these 32 divisions aligned with the boundaries of London's Local Authorities (LAs), also referred to as boroughs, the local government units responsible for delivering key public services (e.g. education, waste collection, social housing). In 2018, the MPS reorganized these divisions into twelve larger Basic Command Units (BCUs). In Section 3, we describe how we account for this change in our data construction.

Figure 1: Police division boundaries in Metropolitan London Police (MPS)



Notes: The figure shows the boundaries of the 32 territorial divisions within the Metropolitan Police Service.

2.2 Borough Commanders

Each territorial division is led by a Borough Commander, typically holding the rank of Chief Superintendent. Reaching this rank generally requires substantial experience within policing, typically fifteen to twenty years or more. Borough Commanders occupy the middle tier of the police hierarchy: they form the first tier of senior operational management between the central leadership of the MPS and the frontline officers and teams responsible for delivering policing within individual neighborhoods.

Borough commanders are responsible for all policing operations within their divisions and manage local budgets and resource deployment. Budget allocation across divisions is determined centrally by the MPS based on local factors such as crime rates, population density, and local community needs. Although budgets and broad strategic priorities are set centrally, commanders retain substantial discretion over how available resources are used. They determine assignment of officers across neighborhoods and specialist units, organize targeted operations such as deployments to crime hot-spots, and allocate funds to community programs, specialist teams, and operational infrastructure (e.g., vehicles, equipment, body-worn cameras). In addition to operational deployment, they make personnel decisions, such as determining staffing levels and recommending promotions, subject to centrally set budgets and vacancy constraints. They are also accountable to senior leadership for divisional performance, but do not receive any performance-related financial compensation. Their role therefore combines substantial operational responsibility with weak formal financial incentives.

Borough Commanders have limited discretion over the divisions to which they are assigned. Promotions within the MPS follow standardized procedures and nationally coordinated timetables. Candidates seeking promotion must apply during fixed application windows, satisfy rank-specific eligibility and minimum-tenure requirements, obtain the

required score in an assessment test, and pass an interview with senior force leadership. Minimum tenure requirements at each rank are summarized in Appendix Table A.1. Officers typically require approximately 15–20 years of service to reach the rank of Borough Commander.

Once promoted, officers are assigned to Borough Commander posts through a centralized, force-wide posting process. Vacancies are reviewed regularly by a Posting Panel, which considers promotion postings alongside voluntary and compulsory transfers. According to information obtained through a Freedom of Information request to the MPS, the primary criterion governing these assignments is operational need. In particular, the Panel considers the circumstances of both the sending and receiving divisions: vacancies in higher-priority or more demanding divisions receive greater priority. Individual preferences may be expressed but are explicitly secondary to organizational needs. Informal transfer arrangements outside the formal process are prohibited, and compulsory postings generally cannot be refused without potential conduct consequences. Expected tenure in a division is approximately two years under the current policy.

As a result, mobility across Borough Commander posts is driven primarily by promotion cycles, vacancies, and organizational needs rather than discretionary sorting into divisions. These institutional features generate an important distinction for our empirical setting: Borough Commanders have limited discretion over the division to which they are assigned. In practice, new postings arise when commanders exit, are promoted, or are transferred, and vacancies are filled through the centralized process rather than through bilateral matching between commanders and divisions.

3 Data

The main dataset links the career histories and monthly division assignments of Borough Commanders to division-level measures of crime, clearances, and workforce constructed from administrative police records. The resulting panel allows us to identify the manager responsible for each division in each month and to examine how division performance changes as managers move across posts. Below, we describe the main data sources and how we combine them.

Middle managers We compile a novel dataset covering the universe of borough commanders in the MPS between 2003 and 2019. The data are collected through Freedom of Information (FOI) requests and scraping of web archives of local police division websites. When information was incomplete or inconsistently reported across sources, we conducted targeted manual checks, such as searches of local news archives, to fill gaps and ensure the accuracy of career histories (see Appendix B for details).

For each manager, we record their start and end dates in post, their assigned territorial division, and their demographic and professional characteristics, including gender, ethnicity, and length of service in policing. We also document the roles they previously held, their subsequent appointments (within or outside the MPS), and retirement dates when applicable. By doing so, we reconstruct the full organizational chart of the MPS over the study period and build the complete career profiles of the middle managers, including their tenure patterns, internal mobility, and transitions in and out of the force. Between 2003 and 2019, the MPS employed 191 middle managers across the 32 police divisions.

Police records. We employ the universe of criminal incidents recorded by the MPS between January 2012 and December 2019. For each incident, we observe the monthly date, offense type, geographical coordinates, and the outcome of the investigation, which captures the action taken by the police or the courts after a crime is reported. The raw dataset contains roughly 8 million records, 75% of them have a valid investigation outcome.[3] All incidents are geo-located and mapped to their corresponding police division.

We construct monthly division-level measures by aggregating reported incidents and their associated outcomes to the division-by-month level. As discussed in Section 2.1, these divisions correspond to the 32 pre-2018 territorial units of the MPS.[4] This approach maintains a consistent panel of 32 divisions.

We adopt a standard measure of police performance, that is the ability to solve (i.e., clear) an incident, conditional on an investigation taking place.[5] Clearance rates are widely used as indicators of police effectiveness in the economics of crime literature (e.g. Thaler, 1977; Mas, 2006; Garicano and Heaton, 2010; Blanes i Vidal and Kirchmaier, 2018; Mastrobuoni, 2020; Facchetti, 2025). On average, 12% of incidents in our sample are cleared, with substantial variation across offense types reflecting differences in severity, the difficulty of identifying a suspect, and the amount of evidence required (Home Office, 2016). For example, clearance rates are approximately 6% for thefts, 13% for violence and sexual offenses, and as high as 80% for drug-related offenses, which are largely recorded when officers detect them through proactive policing, rather than through victim reports.

Thanks to the richness of the police investigation data, we can follow each incident to

[3]We map the 16 offense categories provided in the raw police records into 9 consistent offense types to ensure comparability across years (ASB, burglary, criminal damage, drugs, other crime, theft, robbery, vehicle crime, and violence and sexual offenses). Police records distinguish 24 investigation outcome types (Appendix Table A.2). 95% of incidents without a valid outcome are anti-social-behavior (ASB) incidents, which are never investigated by the police.

[4]Between January 2017 and October 2018, 29 of the 32 divisions were progressively merged into 12 larger Basic Command Units (BCUs). For consistency over time, we retain the pre-merger geography and treat the original 32 divisions as fixed units throughout the sample.

[5]The Home Office defines a crime as cleared when: (i) a notifiable offense has been committed and recorded; (ii) a suspect has been identified; (iii) sufficient evidence is available and the victim is informed of the detection; and (iv) the suspect is charged, cautioned, or otherwise held legally accountable.

its final investigation outcome. We therefore distinguish two outcomes. First, we classify an incident as *police-issued* when it results in one of the outcomes that enable the police to close a case without court involvement, such as a community resolution, a formal caution, a drugs possession warning, or a penalty notice. These cover 6% of all investigation outcomes, and represent the set of investigative outcomes directly determined by police actions and discretion. We also use convictions as a measure of court punishment, which constitute 80% of all court sentences and include crimes sanctioned to imprisonment, fines, or other sentences by the court, excluding acquittals and discharges.

Police workforce. We supplement the incident data with detailed information on police workforce strength at the division-by-month level. To construct the workforce series, we combine multiple administrative sources. Workforce data from April 2008 to July 2017 were obtained from the Mayor's Office for Policing and Crime (MOPAC), while data from August 2017 to July 2019 were released in response to a Freedom of Information request submitted to MOPAC. For each division-period, we observe the full-time equivalent (FTE) strength of four personnel categories: front-line officers, who are sworn constables with full operational policing powers; Police Community Support Officers (PCSOs), uniformed staff with limited enforcement powers whose duties include addressing ASB, handling minor offenses, supporting crime-prevention activities, and traffic control; and civilian staff, who perform administrative, back-office, and support functions (e.g., HR, finance). The MPS also reports the number of Special Constables, who are volunteer officers with powers similar to regular constables. Because Special Constables serve highly variable hours and are not part of the permanent workforce, we exclude them from the FTE counts. Between 2008 and 2019, front-line police officers account for 82% of the division-level workforce, PCSOs for 10% and civilian staff the remaining 8%.

3.1 Sample selection and descriptive evidence

We restrict the analysis period to 2012-2019, the years for which police records are available. During this period, we observe 102 middle managers (Table 1). These are senior officers, with, on average, 27 years of policing experience. Managers hold 1.36 command spells on average, each lasting approximately 24 months, and 23% serve in more than one territorial division during our sample. Among managers for whom demographic information is available, 20% are female and 89% are White. Approximately 13% of observed commands are temporary appointments.

Among managers for whom we observe the position preceding their first command, 35% previously held a position with similar command responsibility, 31% came from another territorial-policing role, and 34% from specialist, headquarters, or other non-territorial positions. Almost all of these prior positions (95%) were within London. After

their final observed command, 39% of managers move laterally to another policing position, 17% are promoted, and 34% retire, with the remainder leaving following disciplinary processes or moving outside policing.

Table 1 also reports descriptive statistics at the division level. On average, a police division experiences just over 2,600 crimes per month, of which roughly 240 are cleared and about 117 are cleared with a police action and the remaining through a court decision. Divisions employ around 550 police officers, 30 civilian staff, and 40 PCSOs on average. The final working sample contains around 2,800 division-month observations.

Motivating evidence We begin by documenting variation in police performance across territorial divisions in Figure 2. The x-axis plots clearances per capita, while the y-axis plots crime rates, for all offenses. Each dot represents a division-date observation. Panel (a) shows the raw, unconditional distributions, while Panel (b) plots the residuals once absorbing police division and date fixed effects and controlling for the total police workforce. This specification accounts for time-invariant differences across divisions and common shocks affecting all divisions in a given period, while also adjusting for workforce-driven variation related to other determinants of crime and policing.

Even after accounting for division and year fixed effects and local police strength, substantial variation remains in both crime and clearance outcomes. Given that many time-varying, division-specific factors influencing crime are outside the control of the police, it is particularly noteworthy that we observe large residual variation in police clearance rates. For example, the standard deviation of the clearance rate residuals (.17 per 1,000 residents) is around one-fifth the size of the mean clearance rate (.97 per 1,000 residents). This persistent unexplained variation in police effectiveness across territorial divisions raises the question of whether differences in police managers help account for part of this pattern. We tackle this question in the next section.

4 Measuring managerial effectiveness in policing

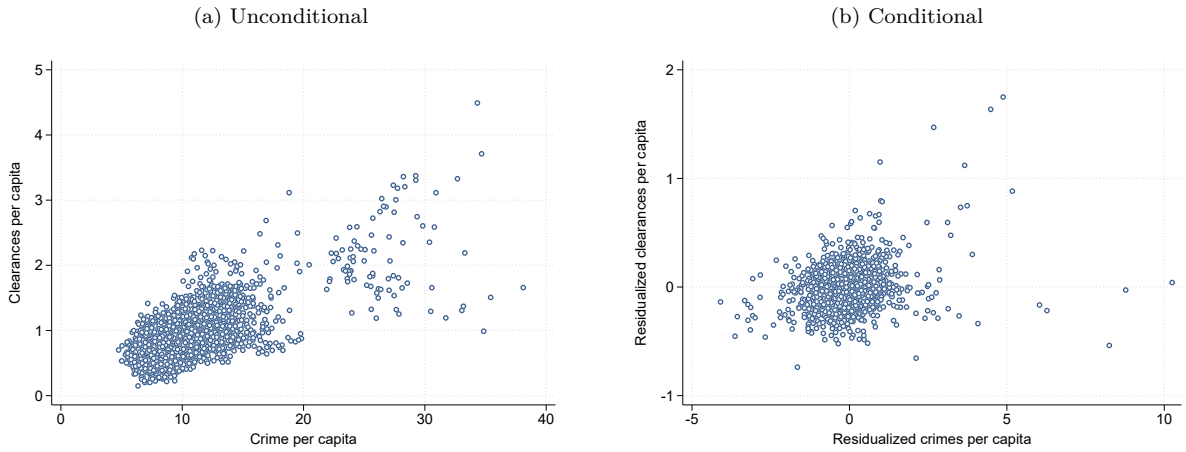
This section presents our empirical strategy for estimating the contribution of middle managers to police performance. We begin by introducing an AKM-style framework that separates persistent differences across police divisions from a manager-specific component of performance. We then discuss the estimation challenges involved in recovering manager fixed effects and the identification assumptions required to interpret them as a portable component of managerial effectiveness across divisions. The recovered estimates will form the basis for the subsequent empirical analyses to characterize who effective managers in policing are and what they do.

Table 1: Descriptive statistics of middle managers and police divisions between 2012 and 2019

	Mean (1)	N (2)
Panel A. Manager characteristics		
Mover	0.23	102
Female	0.20	100
White	0.89	100
Years in policing	26.69	85
Age	47.38	47
Average number of spells	1.36	102
Average spell length (months)	23.91	102
Position before first observed command:		
- Previous command responsibility	0.35	80
- Territorial role, excluding command	0.31	80
- Specialist/HQ/non-territorial role	0.34	80
- Previous position in London	0.95	80
Position after final observed command:		
- Promotion	0.17	99
- Lateral move	0.39	99
- Retirement	0.34	99
- Disciplinary exit	0.05	99
- Left policing for outside employment	0.03	99
- Still in command at end of sample	0.01	99
Post-spell territorial status:		
- Territorial policing	0.16	98
- Outside territorial policing	0.84	98
Panel B. Division characteristics		
Clearances:		
- Total	238.81	2774
- Convictions	92.27	2774
- Police-issued	117.13	2774
Offenses:		
- Total	2,622.47	2774
- Violent	545.85	2774
- Property	1,161.95	2774
- Drug	99.73	2774
- Other	151.75	2774
Workforce:		
- Total	622.56	2774
- Police officers	547.40	2774
- Staff	33.37	2774
- PCSO	41.79	2774
Clearances per worker	0.38	2774

Notes: This table presents descriptive statistics for middle managers and police divisions in the Metropolitan Police Service (MPS) between 2012 and 2019. The columns report the mean and the number of non-missing observations. Panel A reports manager-level characteristics for the 103 middle managers observed in the sample. Age and years of policing experience are measured at the beginning of each manager's first observed command spell. Previous-position categories refer to the position held before the manager's first observed command spell, while subsequent-position categories refer to the position or outcome observed after the manager's final observed command spell. Previous command responsibility includes Borough Commander and comparable command posts; territorial roles include other operational positions within territorial policing; specialist/HQ/non-territorial roles comprise positions outside the territorial command structure. Panel B reports descriptive statistics for the division-by-month panel, covering 32 London police divisions and approximately 2,800 division-month observations. The City of London is excluded because it is policed by the City of London Police rather than the MPS.

Figure 2: Variation in crime and police performance across police divisions



Notes: This figure plots the relationship between crime rates and clearance rates in London police divisions. Panel (a) shows the unconditional relationship between crime rates and clearances per capita. Panel (b) displays the relationship after residualizing both variables with respect to division fixed effects, date fixed effects, and workforce size. Each point represents a division-month observation.

4.1 Empirical framework

Following Abowd et al. (1999), we decompose police performance in each territorial division into three components: a territorial division specific component, a manager component, and a common time component. We estimate a log-linear model:

$$log(y_{d,t}) = \mu_d + \tau_t + \mu_{m(d,t)} + \epsilon_{d,t}, \tag{1}$$

where μ_d is the police division fixed effect, τ_t is the time period fixed effect (in most specifications we use monthly frequency), and $\mu_{m(d,t)}$ is the fixed effect of the manager associated with police division d in time period t. The manager fixed effect captures the component of performance that systematically follows a manager across division assignments and is therefore our measure of managerial effectiveness.

The main police performance measure is *clearances per worker*, defined as the number of cleared incidents in a division divided by its total workforce across all functions (excluding Special Constables). This value-added-type measure parallels approaches used in other public-sector settings in which outputs are not priced, such as hospital outcomes or student achievement (Chandra et al., 2016; Muñoz and Prem, 2024; Muñoz and Otero, 2025). Clearances are particularly suited to this setting because they capture an outcome over which police managers have direct influence through decisions over investigations, personnel deployment, and operational priorities. Crime levels, by contrast, also depend on factors largely outside police control, including offender behavior, victim reporting, and local socio-economic conditions.

Table 2 compares the explanatory power of alternative fixed-effect specifications. A

Table 2: Comparison of Adjusted R2 on police division output

	(Log) clearance per worker			
	(1)	(2)	(3)	(4)
N	2523	2523	2523	2523
Time FE	Yes	Yes	Yes	Yes
Division FE	No	Yes	Yes	No
Manager FE	No	No	Yes	No
Manag-by-Division FE	No	No	No	Yes
R sq.	0.457	0.659	0.798	0.805
Adj. R sq.	0.438	0.643	0.780	0.786
RMSE	0.197	0.157	0.123	0.122
P-value			0.000	
SD Manager FE			0.192	
SD Division FE			0.212	
Corr. manager/division FE			-0.711	

Notes: This table shows the explanatory power of different fixed effects specifications for log clearance rates. All models include time fixed effects. Column (1) shows time FE only Column (2) adds division FE. Column (3) adds manager FE. Column (4) includes manager-division interactions, with marginal additional explanatory power.

model containing only time fixed effects explains 44% of the variation in log clearances per worker, while adding division fixed effects raises the adjusted R^2 from 0.438 to 0.643. Adding manager fixed effects further increases the adjusted R^2 to 0.780, an increase of approximately 21% relative to the specification with division and time fixed effects, and we reject the null that the manager effects are jointly zero (p-value= 0.00). This increase, that is larger in magnitude than the increases typically found in other settings (e.g. Bertrand and Schoar, 2003; Fenizia, 2022; Muñoz and Otero, 2025), indicates that in the context of policing, middle managers explain a meaningful share of performance variation across police divisions. By comparison, the incremental explanatory power of manager fixed effects is substantially smaller for log total clearances (5%) and almost negligible for log recorded crime (Appendix Table A.3). These patterns are consistent with managers directly influencing police productivity, while crime levels are shaped primarily by external factors. They motivate our focus on clearances per worker as the main outcome for measuring managerial effectiveness.

4.2 Estimation challenges

Recovering manager and division effects from Equation (1) raises a few estimation challenges, including mobility requirements, connected sets, and measurement error.

First, manager and division fixed effects are separately identified only if managers move across divisions. In the absence of mobility, manager and division effects would be perfectly collinear and thus not separately estimable. Furthermore, fixed effects can only be compared within *connected sets*, that is, sets of managers and divisions linked through managers' mobility. Different connected sets normalize the fixed effects relative to different benchmarks, thus making comparison of fixed effects' magnitudes across different connected sets problematic. During our 2012–2019 estimation period, we identify one

large connected set containing 93 middle managers and 29 territorial divisions. Within this set, we observe 114 managerial transitions, with 23 managers moving at least once and each division experiencing at least one managerial change. All 29 divisions experience at least one managerial change. We focus the analysis on this largest connected set.[6]

A second estimation challenge arises because we estimate a large number of division and manager fixed effects from a relatively small number of division-month observations. As a result, the estimated fixed effects may contain substantial measurement error. To address this issue, we follow Chandra et al. (2016) and Metcalfe et al. (2023) and apply an Empirical Bayes (EB) correction. The procedure shrinks each estimated fixed effect toward the mean in proportion to its sampling variance, using the standard error from Equation (1) as the estimate of the latter. Noisier estimates therefore receive greater shrinkage, while more precisely estimated effects are affected less.

Figure 3 compares the distributions of the raw and EB-adjusted manager and division fixed effects. The adjusted and unadjusted estimates are highly correlated (correlation = 0.99 for both the manager and the division fixed effects). The adjustment reduces the dispersion of manager effects, as expected. Division effects are affected even less by the adjustment. The EB-adjusted fixed effects are our preferred measures when characterizing individual managers and divisions and when constructing changes in managerial effectiveness around transitions below. A separate issue arises when using the estimated fixed effects to decompose the variance of police performance. Because sampling errors in estimated manager and division effects are correlated, shrinking each fixed effect independently does not in general recover the variance and covariance of the underlying effects. We address this additional issue when presenting the variance decomposition in Section 4.4.

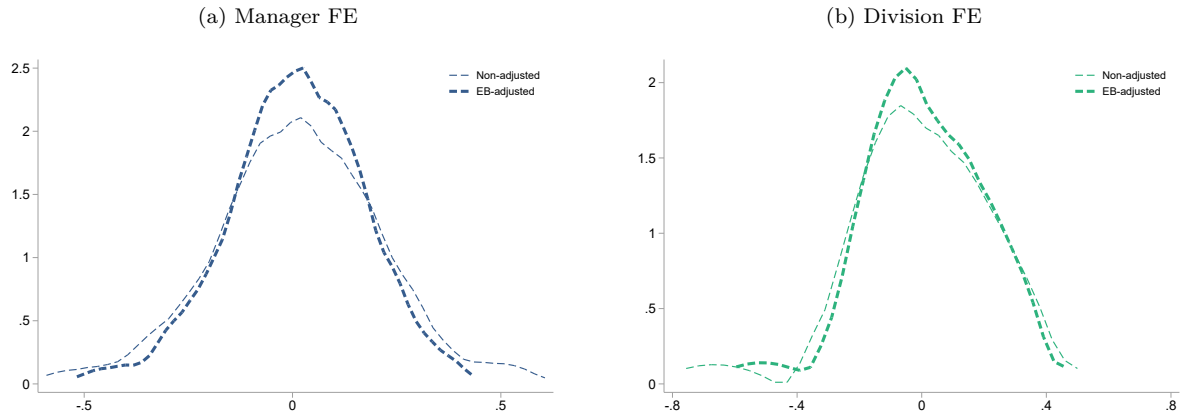
4.3 Identification

The identification of manager and division fixed effects in Equation (1) relies on the assumption that the assignment of managers to divisions is conditionally independent of the error term $\epsilon_{d,t}$. This assumption is often referred to as "exogenous mobility condition" (Card et al., 2013). Importantly, this assumption does not require managers to be randomly assigned across divisions. Managers may be systematically assigned according to permanent division characteristics, (μ_d), or according to their own persistent managerial effectiveness, (μ_m). For example, systematically sorting higher-performing managers into more challenging divisions, or vice-versa, does not violate the identification assumption.

Two potential violations are particularly relevant in this setting (Card et al., 2016).

[6]Appendix Figure A.1 displays the mobility network. A second small connected set contains three divisions (Barnet, Brent, Harrow) and nine managers who do not move across divisions. Because their division effects are collinear with time fixed effects, the design matrix is not full rank and this set is excluded from the analysis.

Figure 3: Distribution of manager and division fixed effects



Notes: This figure shows the distributions of estimated fixed effects from Equation (1). Panel (a) displays manager fixed effects and panel (b) shows division fixed effects. The dashed lines represent baseline estimates while the solid lines show empirical Bayes-adjusted estimates that account for estimation error. Both distributions are centered at zero by construction.

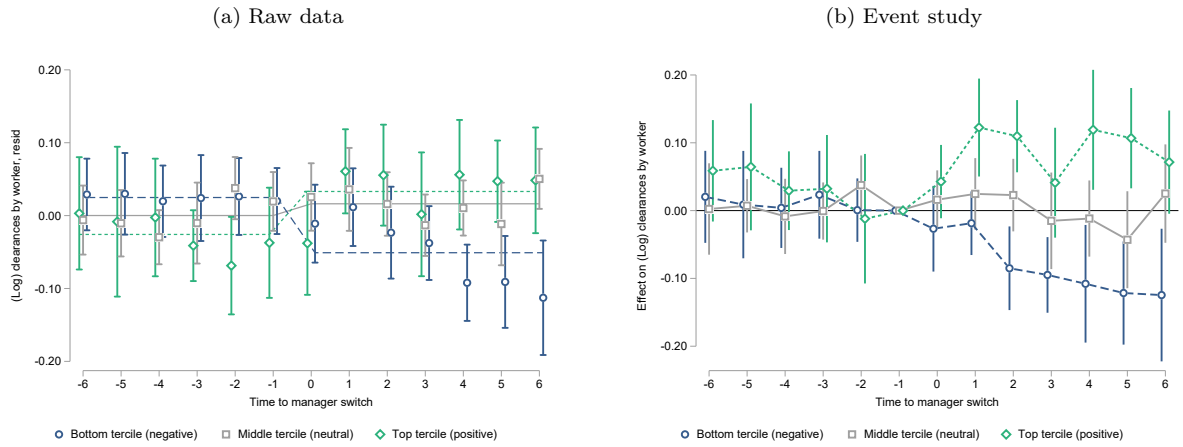
First, managerial mobility may respond to time-varying division-specific shocks. For instance, managers could be more likely to leave divisions experiencing negative shocks and to arrive in divisions that are already improving. Second, the additive structure of Equation (1) may be mis-specified if managerial effectiveness depends substantially on manager–division-specific complementarities. We examine each concern in turn.

Second, the additive structure of Equation (1) may be misspecified if police output also depends on manager–division-specific match effects, ($\mu_{d,m}$), that capture complementarities between particular managers and divisions. In that case, sorting on match quality could make observed mobility endogenous if these match effects are correlated with manager effectiveness. We examine each concern in turn.

Sorting on time-varying shocks A potential threat to identification arises if manager mobility is systematically related to time-varying division-specific shocks. For instance, if managers on an upward performance trend are more likely to transfer into divisions that are already experiencing improvements in clearance rates, we could overestimate the division fixed effects (or vice-versa). The institutional setting provides a first reason why such sorting may be limited. As discussed in Section 2.2, assignments into divisions are made by a centralized panel that explicitly subordinates officer preferences to operational needs and prohibits informal transfer arrangements. Movements of officers are highly regulated and largely constrained by vacancy availability, limiting managers' discretion in choosing divisions. We complement this institutional evidence with tests based on the behavior of division outcomes around managerial transitions.

We first classify transitions according to the change in estimated managerial effective-

Figure 4: Log clearance per worker by manager effectiveness



Notes: This figure shows the evolution of log clearances per worker around manager transitions. Transitions are classified into terciles based on the change in empirical-Bayes-adjusted manager fixed effects between the incoming and outgoing manager. Bottom-tercile transitions correspond to the largest declines in managerial effectiveness, middle-tercile transitions to relatively small changes, and top-tercile transitions to the largest improvements. The terciles are defined considering all transition events that are part of our sample. Panel (a) shows raw residualized clearance rates per worker, computed absorbing date and division fixed effects, by month since manager switch; Panel (b) presents coefficients from stacked event-study specification in Equation (3). 95% confidence intervals are shown with vertical bars.

ness between the incoming and outgoing manager in the division:

$$\Delta ManEff_d = \hat{\mu}_{m(d,t)} - \hat{\mu}_{m(d,t-1)} \tag{2}$$

The manager fixed effects $\hat{\mu}_m$ are the empirical-Bayes-adjusted estimates obtained using log clearances per worker as the outcome described in section 4.2. We divide transitions into three groups based on the distribution of $\Delta ManEff_d$: bottom-tercile transitions, in which managerial effectiveness declines; middle-tercile transitions, involving relatively little change in effectiveness; and top-tercile transitions, corresponding to increases in managerial effectiveness.

To examine whether these transitions are preceded by differential trends in division performance, we implement a stacked event-study design. For each managerial transition s, we construct a separate stack containing the division experiencing the transition and divisions that do not experience a manager change over the corresponding event window. We focus on 63 transitions for which the same outgoing manager is observed throughout the six months preceding the transition and the incoming manager remains in the division for at least six months. We retain a balanced window from six months before to six months after the transition. We append the resulting stacks into a single dataset. A division may therefore serve as a control in more than one stack, and a given division-month observation may appear multiple times.

We estimate the following specification:

$$log(y_{sdt}) = \sum_{k=\neq -1} (\beta_k \text{Bottom}_s + \gamma_k \text{Mid}_s + \delta_k \text{Top}_s) \times D^k_{st} \times Treated_{sd} + \\ + \phi_{sd} + \phi_{st} + \epsilon_{sdt} \tag{3}$$

where D^k_{st} denote a set of relative event-time dummies, each taking the value of 1 if period t is k periods after (or before, if k is negative) the relevant manager transition in stack s. $Treated_{sd}$ is an indicator variable equal to one if division d is the one experiencing a transition, and $Bottom_s$, Mid_s, Top_s denote the tercile of the transition. We omit the period $k = -1$ before the managerial transition. The division-by-stack fixed effects, ϕ_{sd}, absorb permanent differences between divisions within each stack, while the calendar-month-by-stack fixed effects, λ_{st}, absorb shocks common to the treated and control divisions in a given stack and month. Standard errors are clustered at the police-division level, thereby allowing observations from the same division to be correlated across stacks.

The coefficients β_k, γ_k, and δ_k trace the treated-versus-control evolution of the police performance around transitions involving declines, small changes, and improvements in managerial effectiveness, respectively. For example, δ_k measures the change between month k and the month before the transition in a division receiving a top-tercile improvement in managerial effectiveness, relative to the contemporaneous change among control divisions.

Figure 4 shows both residualized raw outcomes (Panel (a)) and the corresponding stacked event-study estimates (Panel (b)). Division performance follows similar trends across the three transition groups prior to the manager change. Following the transition, clearances per worker fall when a division receives a substantially lower-performing manager and rise when it receives a substantially higher-performing manager. Appendix Table A.4 reports formal joint tests of the pre-transition coefficients for several clearance and crime outcomes and fails to reject the null hypothesis. The absence of differential pre-trends provides evidence consistent with mobility not being driven by transitory division-specific shocks. It also implies that while managerial mobility may correlate with persistent division characteristics, it is not systematically related to transitory shocks, supporting the assumption of exogenous mobility.[7]

Sorting on manager-division matches A second concern is that the effect of a manager may depend on the particular division in which they work. Equation (1) imposes additive separability between manager and division effects. A manager's contribution is assumed to be portable across assignments rather than arising from a specific manager–

[7] As an additional check, we restrict attention to manager transitions triggered by retirement episodes of departing managers. Because retirement is more closely related to age and career stage, its timing is plausibly less responsive to short-run changes in division performance. Appendix Figure A.2 shows no evidence of differential pre-trends around these transitions.

division match. As a result, a high-performing manager should have the same effect regardless of the division.

To assess potential violations of this assumption, we implement several diagnostic tests following Card et al. (2013) and Card et al. (2018). First, we estimate a more flexible model that includes manager-by-division fixed effects. This model relaxes the additive structure of Equation (1) and adds idiosyncratic match effects, $\mu_{d,m}$. If match-specific components were important, we would expect this saturated model to display a substantially better fit (proxied by the adjusted R^2 of the model) than the baseline specification. Table 2 column (4) reports the result of this diagnostic check. Adding manager by division fixed effects increases the fit of the model marginally, from 0.780 in the baseline model in column 3 to 0.786 (equal to 0.8%), suggesting that the match effects explain little variation in outcomes. As a complementary test, we divide the estimated manager and division effects into quintiles and compute the mean residual from Equation (1) for each manager–division cell. Appendix Figure A.3 shows that the residuals are small and display no systematic pattern across the cells, providing no evidence of substantial match-specific effects.

Finally, Figure 4 provides a test for the additive separability assumption. In the absence of sorting based on manager-specific productivity, the AKM model predicts that divisions moving from a low- to a high-performing manager will experience the same change in performance (equal in magnitude and opposite in sign) than divisions moving from a high- to a low-performing manager. We observe that the estimated post-transition responses are approximately symmetric across improvements and declines in managerial effectiveness. This symmetry supports the additive specification and, more broadly, the interpretation of the estimated manager fixed effects as portable across divisions.

4.4 Do managers matter in territorial policing?

Having recovered manager and division fixed effects, we now quantify how much of the variation in division-level policing output can be attributed to each component.

To quantify the contribution of managers relative to divisions, we decompose the variance of Equation (1) (Abowd et al., 1999):

$$\begin{aligned} Var(log(y_{d,t})) = Var(\mu_d) + Var(\mu_{m(d,t)}) + Var(\tau_t) + Var(\epsilon_{d,t}) + \\ + 2 \times Cov(\mu_d, \mu_{m(d,t)}) + 2 \times Cov(\mu_d, \tau_t) + 2 \times Cov(\mu_{m(d,t)}, \tau_t). \end{aligned} \quad (4)$$

The main challenge with the variance decomposition exercise comes from limited mobility bias (Andrews et al., 2008). As discussed in Section 4.2, finite-sample bias can arise because a large number of manager and division fixed effects are identified from a relatively small number of managers who move across divisions. This may overstate the true dispersion of both manager and division effects and bias their estimated covari-

ance downward. As manager and division effects are estimated jointly from the same observations, estimation error in the two components is therefore mechanically related: an overestimate of the manager effect can be accompanied by an underestimate of the division effect, and vice versa. Empirical Bayes shrinkage reduces noise in each set of fixed effects separately, but does not account for this correlation in their estimation errors. Independently shrinking the two sets of effects therefore does not fully correct the variance-covariance components.

Our preferred approach therefore follows Best et al. (2023) and applies a covariance-shrinkage correction that jointly adjusts the manager and division fixed effects using an estimate of their full sampling covariance matrix. To do that, we resample the AKM residuals with replacement $B = 100$ times, stratifying within each manager-division match to preserve the match-specific error patterns. As this method captures the full off-diagonal sampling covariance and does not rely on homoskedasticity, the covariance shrinkage addresses the same bias as Andrews et al. (2008) but under weaker assumptions.

Table 3 reports the resulting variance decomposition. For comparison, we present four estimates. The first uses the raw AKM fixed effects. The second applies the finite-sample correction proposed by Andrews et al. (2008), which corrects the variance and covariance components under homoskedasticity.[8] The third uses the Empirical Bayes-adjusted fixed effects introduced in Section 4.2. The final row reports our preferred covariance-shrinkage estimates.

The estimated contribution of managers is substantial under all four approaches, although its magnitude falls once estimation error is accounted for. The raw AKM estimates attribute 53% of the total variance in log clearances per worker to manager fixed effects. The Andrews correction yields a similar share of 49%. This relatively small adjustment is consistent with the mover graph in our setting being well connected, so that the homoskedastic trace correction removes relatively little bias. Scalar Empirical Bayes shrinkage reduces the manager share to 37% by removing noise within the manager fixed-effect vector. However, because it shrinks the manager and division fixed effects separately, it does not account for the covariance in their estimation errors. Our preferred covariance-shrinkage correction uses the full sampling covariance matrix of the estimated manager and division effects. Under this correction, the variance of the manager component is 0.017, corresponding to approximately 25% of the total variance in log clearances per worker, about 70% as large as the division-specific component. Thus, even under our most conservative correction, persistent differences across Borough Commanders account for a quantitatively important share of variation in police productivity.

Appendix Table A.6 reports results from two additional bias-correction methods.

[8]This method is valid only within a single connected set, which is why most empirical papers restrict estimation to their largest connected set, while Best et al. (2023) present an approach that is applicable when multiple connected sets are present.

First, we implement a split-sample variance correction (e.g. Best et al., 2023; Dahis et al., 2025), which randomly partitions the sample into two halves, stratified by manager-division pair so each cell contributes to both, and estimate the AKM model separately on each half. Because the sampling noise in the two halves is independent by construction, the cross-half covariance is an unbiased estimator of the true signal variance of the fixed effects, without requiring homoskedasticity or assumptions on the FE distribution. This approach yields a manager FE share of 48%, close to the Andrews-KSS-corrected estimate of 49% and above our preferred covariance-shrinkage estimate. Second, following Adhvaryu et al. (2023), we conduct a Monte Carlo simulation that quantifies the finite-sample bias of the OLS variance estimator on our specific data structure. Holding the observed manager-division assignment network fixed, we simulate outcomes using the estimated AKM effects as the "true" data-generating effects and re-estimate the AKM model in each simulated sample. If the OLS variance estimates were unbiased, their average across simulations would coincide with the variance of the fixed effects used to generate the data. Across 10,000 simulations, OLS overstates the variance of manager effects by 4% and the variance of division effects by 2.3%, suggesting that finite-sample inflation of the variance components is relatively modest in our setting.

The decomposition shows a sizable negative relationship between manager and division effects, which survives even after covariance shrinkage. The estimated covariance is -0.009, which corresponds to approximately 13% of the total variance in log clearances per worker. Hence, managers with higher estimated effectiveness tend to be observed in divisions with lower estimated permanent components, and vice versa. Figure A.5 illustrates this pattern directly. We divide managers and divisions into terciles of their respective fixed-effect distributions and plot the frequency of the resulting matches. High-performing managers are relatively unlikely to be paired with high-performing divisions and are instead disproportionately observed in divisions in the bottom of the division-effect distribution. Conversely, lower-performing managers are more frequently assigned to divisions with relatively high division effects. This negative assortative matching aligns with the centralized posting process described in Section 2. Because the MPS Posting Panel prioritizes operational needs over individual preferences, stronger managers may be directed toward divisions facing greater challenges. This pattern is consistent with an organizational policy that allocates stronger managerial capacity toward divisions facing greater operational needs.[9]

[9]While this feature is not surprising in the public sector (e.g. Limodio, 2021), similar evidence has been reported in private-sector contexts too. For instance, Metcalfe et al. (2023) find evidence of negative assortative matching between managers and firms, despite standard matching models predicting positive assortative matching when profit maximization is the main objective.

Table 3: Variance decomposition of log(clearances per worker)

Method	Var(Outcome)	Var(Manager)	Var(Division)	Cov(Manager, Division)
	(1)	(2)	(3)	(4)
Raw AKM	0.069	0.037	0.045	-0.029
	[100.0]	[53.0%]	[64.7%]	
Andrews	0.069	0.034	0.042	-0.027
	[100.0]	[48.9%]	[61.1%]	
EB Shrinkage	0.069	0.025	0.034	-0.020
	[100.0]	[36.5%]	[48.5%]	
Covariance Shrinkage	0.069	0.017	0.024	-0.009
	[100.0]	[24.7%]	[35.3%]	

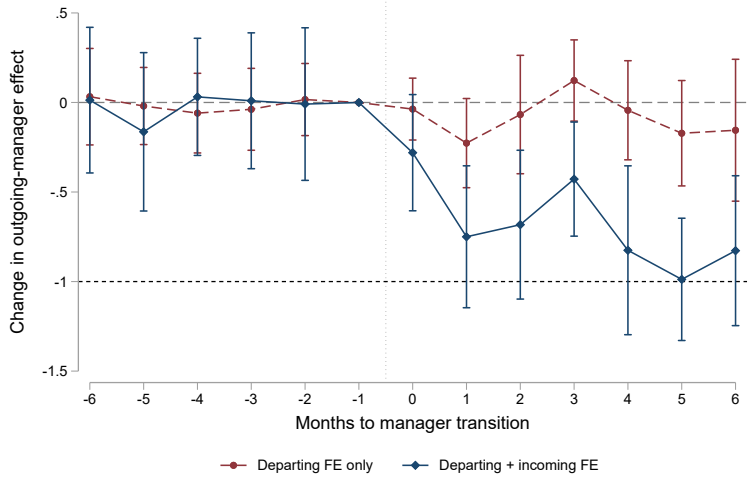
Notes: This table shows the variance-covariance decomposition of log clearance per worker based on the AKM model in Equation (1). Rows correspond to the following four methods: *Raw AKM*, the finite-sample trace correction of Andrews et al. (2008), scalar empirical-Bayes shrinkage, and the covariance shrinkage methods of Best et al. (2023). Brackets report each term as a share of the total variance in column 1.

4.5 Magnitude and persistence of manager effects.

Figure 3 shows the distribution of the estimated manager effects. A one-standard-deviation increase in managerial effectiveness corresponds to approximately 12.6% higher clearances per worker, while moving from a manager at the 10th percentile to one at the 90th percentile of the distribution is associated with an increase of approximately 26%. We next ask how much of the overall variation in police productivity can be attributed to these persistent differences across managers.

We investigate whether the performance change associated with a manager persists after the manager leaves the division. For each transition, we trace how the association between division performance and the outgoing manager's fixed effect evolves following the arrival of a new commander. Figure 5 reports estimates both with and without conditioning on the effectiveness of the incoming manager. The distinction is important: when successor quality is omitted, the outgoing-manager effect changes little after departure, suggesting substantial persistence. However, once incoming-manager quality is taken into account, the outgoing-manager effect declines sharply. Averaged over the first six months of the successor's spell, approximately 34% of the outgoing-manager effect remains, and the estimates suggest complete disappearance toward the end of the event window. This evidence suggests that as manager effects emerge quickly after a new commander arrives, they also dissipate relatively quickly after departure. This pattern provides little evidence that the performance gains are primarily driven by durable organizational changes that persist under subsequent managers.

Figure 5: Persistence of departing-manager effects



Notes: The figure reports the evolution of manager contributions of the departing manager around managerial transitions. We estimate the following equation:

$$y_{dst} = \phi_{ds} + \phi_{st} + \sum_{k \in \mathcal{K}} \alpha_k^D D_{ist}^k \widehat{\mu}_s^D + \sum_{k \in \mathcal{K}} \alpha_k^I D_{ist}^k \widehat{\mu}_s^I + \epsilon_{dst}$$

where $\widehat{\mu}_s^D$ and $\widehat{\mu}_s^I$ denote the fixed effects of the departing and incoming managers in division k associated with transition s. The event window is $\mathcal{K} = \{-6, \ldots, -2, 0, \ldots, 6\}$, with the month immediately preceding the transition, $k = -1$, omitted. y_{dst} is log clearances per worker. The solid blue line plots $\widehat{\alpha}_k^D$, which measures the change in the coefficient on the departing manager's fixed effect relative to $k = -1$. The dashed red line is estimated without controlling for the incoming manager's fixed effect. Standard errors are clustered at the police-division level; vertical bars report 95% confidence intervals.

We next examine whether managers systematically move to better divisions. Figure A.7 documents the distribution of changes in manager and division fixed effects around transitions. Panel (a) plots the differences between the incoming and the outgoing manager fixed effects. The distribution is approximately symmetric around zero, with a mean difference equal to -0.057 SD. This implies that incoming managers are, on average, not different from the ones they replace and is consistent with the pattern observed in the raw data, plotted in Figure A.8, which presents the share of division-manager matches by the relative quality of incoming and outgoing managers. Divisions tend to replace managers with others of similar effectiveness, while transitions from top to bottom-quality managers (and vice versa) are relatively rare. Panel (b) computes instead the changes in division fixed effects for the moving managers. We show that managers tend, on average, to move to divisions with slightly higher estimated division effects (mean difference of 0.008 SD).

4.6 Robustness

We conduct several robustness checks. First, our estimates are insensitive to the definition of the police value-added measure. Our preferred outcome is clearances per worker, defined as total cleared incidents divided by the total division workforce, including police officers, PCSOs, and civilian staff. This corresponds to a value-added interpretation in which all workforce inputs contribute to the production of clearances. Re-estimating Equation (1) using clearances per police officer instead produces virtually identical manager fixed effects, with a correlation of 0.994 across the two specifications.

Second, we validate the estimated manager effects using an alternative data source and sample period. Specifically, we re-estimate Equation (1) using borough-level counts of *sanctioned detections* published by the MPS for 2008–2016. Despite the different data source and only partially overlapping sample period, the manager fixed effects recovered from the two samples are positively correlated: the correlation is 0.57 among the 54 managers for whom fixed effects can be estimated in both samples. Appendix B2 provides details on the construction of these data and the corresponding estimates.

Third, our event-study results are not specific to the stacked design. We re-estimate the transition effects using a conventional two-way fixed-effects event-study specification. Appendix Figure A.4 shows that the resulting patterns around improvements and declines in managerial effectiveness are very similar to those obtained from the stacked design. We retain the stacked specification as our preferred approach because it provides a cleaner comparison around individual transitions and is less sensitive to concerns arising from heterogeneous treatment effects in conventional two-way fixed-effects event studies.

5 Understanding effective police managers

Having established substantial heterogeneity in managerial effectiveness among Borough Commanders, in this section we investigate what distinguishes high-performing managers and how they improve police performance. We begin by asking whether managerial effectiveness is related to observable characteristics or subsequent career outcomes. We then examine what changes within a division when managerial effectiveness improves, before studying whether effective managers shape how divisions respond to policy shocks.

5.1 Characteristics and career paths of effective managers

We first ask whether the differences in managerial effectiveness recovered in Section 4 are systematically related to observable manager characteristics. Figure A.6 reports coefficients from regressions of the estimated manager fixed effects on demographic and career characteristics, including the type of position held before entering command. Managerial

effectiveness is largely unrelated to gender, ethnicity, entry cohort, whether a manager moves across divisions, or tenure as a Borough Commander. We do not find meaningful differences across managers coming from previous command positions either.

We next ask whether managerial effectiveness is reflected in managers' subsequent careers. Table A.5 relates the transition following a completed Borough Commander spell to manager characteristics, managerial effectiveness, and average division performance during the spell. Columns (1)-(3) distinguish lateral moves, promotions, and retirement, while Column (4) captures whether the manager's subsequent position remains within territorial policing. Career stage is an important predictor of these outcomes: managers who entered policing before 1990 are less likely to be promoted and more likely to retire, consistent with older officers approaching the end of their careers. By contrast, we find no significant relationship between either the estimated manager fixed effect or average clearance performance during the spell and subsequent career progression, including promotion. We find some correlation between crime experienced during the spell and subsequent mobility, with managers serving during higher-crime spells less likely to move laterally or remain in territorial policing and more likely to retire. This contrasts with evidence that career incentives tied to promotion can be an important determinant of bureaucratic performance (Bertrand et al., 2020). In our setting, the dimension of managerial effectiveness captured by clearance productivity does not appear to be strongly rewarded through subsequent promotion. This result should be interpreted with caution, as our fixed effects capture one particular dimension of managerial performance, while promotion decisions may also depend on other dimensions of managerial ability that we do not capture in our data.

5.2 Unpacking the productivity increase

We next ask what lies behind the productivity gain associated with effective managers. We first decompose the gain in clearances per worker into its numerator and denominator, then examine which types of clearances and offenses drive it, and finally ask whether it reflects changes in underlying crime. To do so, we study how policing outcomes evolve around transitions that induce a change in managerial quality using a stacked event-study design.

Specifically, we use the stacked sample and corresponding event-study specification introduced in Section 4.3, but replace the tercile indicators with the continuous change in managerial effectiveness between the incoming and outgoing manager. This results in

the following specification:

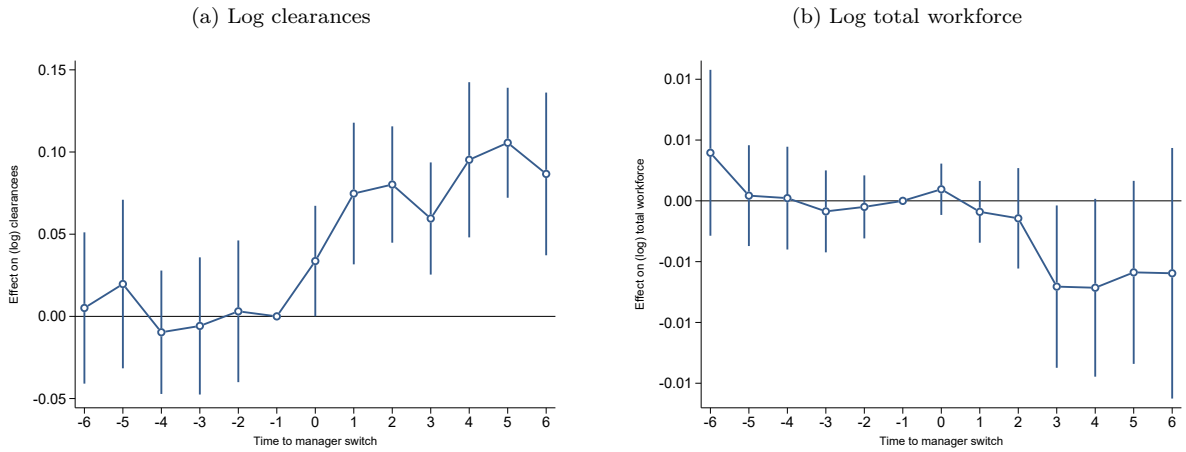
$$log(y_{sdt}) = \sum_{k\neq -1}^{T} \pi_0^k D_{ts}^k \times Treated_{sd} + \sum_{k\neq -1}^{T} \pi_1^k D_{ts}^k \times Treated_{sd} \times \Delta M_s + \phi_{sd} + \phi_{st} + \epsilon_{sdt}, \tag{5}$$

where $log(y_{sdt})$ is (log) outcome for division d measured in month t appearing in stack s, D_{st}^k denotes event time relative to the manager transition in stack s, $Treated_{sd}$ identifies the division experiencing the transition, and ϕ_{sd} and λ_{st} are division-by-stack and date-by-stack fixed effects. Standard errors are clustered at the police-division level. The π_0^k coefficients capture the evolution of the outcome around a transition with no change in managerial effectiveness. π_1^k are our coefficients of interest and capture the differential effect associated with a one-standard-deviation improvement in managerial effectiveness, corresponding to approximately 12.6% in clearances per worker. [10]

We begin by asking whether the increase in productivity associated with more effective managers reflects higher police output or changes in workforce size. Figure 6 shows that the effect is driven primarily by clearances, while the respective average effect sizes are reported in Appendix Tables A.7 and A.8. A one-standard-deviation larger change in managerial effectiveness raises total clearances by 7.4%, while total workforce declines by 0.4% (significant at the 10% level). The increase in clearances per worker therefore arises predominantly from an increase in police output rather than from a contraction in the workforce, consistent with this type of changes being typically slow-paced in large bureaucratic organizations such as a police department.

[10] To summarize average effects, we also estimate a version of this specification that replaces the event-time dummies D_{st}^k with a $Post_{st}$ indicator equal to 1 in all post-transition periods. Given the small number of clusters, the corresponding tables report wild-bootstrap $p-$values for both the neutral-transition effect and the differential effect of managerial quality.

Figure 6: Clearances or workers?



Notes: Panels (a) and (b) plot the estimated event-time coefficients from the stacked differences-in-differences specification in Equation (5). Coefficients are normalized relative to the month immediately preceding the transition ($k = -1$). Standard errors are clustered at the police division level.

We next examine the composition of this additional output. Figure 7 Panel (a) and Appendix Table A.7 columns (3) and (4) show that clearances increase across offense categories, for both property and violent offenses and drug-related offenses. A one-standard-deviation increase in managerial effectiveness raises clearances of property and violent offenses by 5.8% and of drug offenses by 9.9%. Given that drug offenses are rarely reported by victims and are instead largely uncovered through police-initiated activity, this indicates that higher managerial effectiveness increases both the resolution of reported offenses and proactive policing.[11]

The fact that productivity gains are driven by higher clearances raises the concern that more effective managers may simply inflate clearance numbers through low-quality arrests. Figure 7 Panel (c) and Appendix Table A.7 columns (5) and (6) shows that this is not the case: both police-issued sanctions and court convictions increase following transitions to more effective managers, indicating improvements in investigation quality.[12]

Finally, a natural question is whether these changes in police activity also translate into changes in crime through deterrence or incapacitation. Consistent with the earlier finding that manager fixed effects explain much less variation in crime than in clearances, Panel (b) and Appendix Table A.9 show no detectable effect of managerial effectiveness on reported property and violent offenses, while drug offenses increase significantly. Because

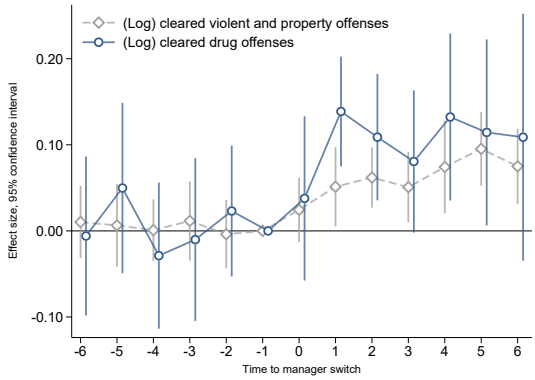
[11] We also verify that the results are not driven by the 2017–2018 BCU reorganization. Appendix Table A.10 shows that estimates are robust to the inclusion of division fixed effects interacted with an indicator for the post-merger periods.

[12] A related concern is that managers may reallocate effort across tasks, improving performance in one dimension while reducing it in others. To assess this, we estimate manager fixed effects separately for clearances per worker in violent, property, and drug-related offenses. We find no evidence of multi-tasking trade-offs: the offense-specific manager effects are positively correlated, with pairwise correlations of 0.26-0.49 (Appendix Figure A.9), indicating that effective managers tend to perform well across crime types rather than shifting effort between them.

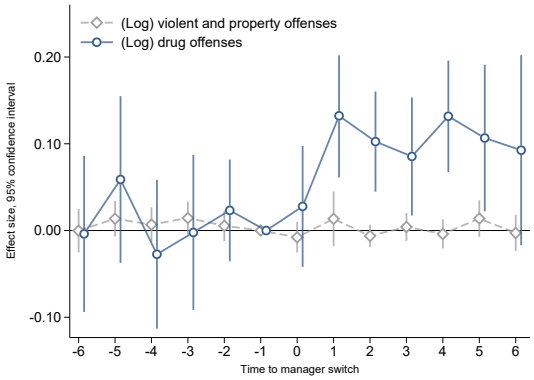
drug offenses are predominantly police-detected rather than victim-reported, this again reflects greater proactive policing rather than a rise in underlying drug offending.

Figure 7: Effects of manager effectiveness on clearances and crime

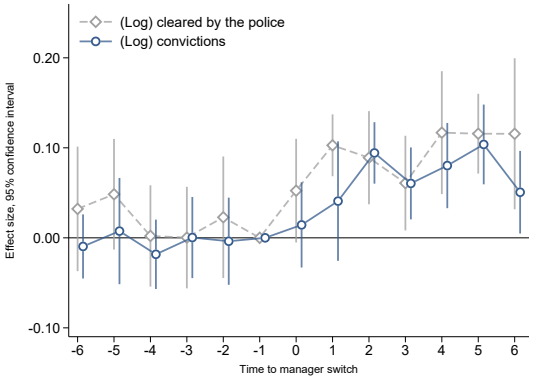
(a) Clearances by offense type



(b) Crime by offense type



(c) Clearance types



Notes: The figure reports estimated event-time coefficients from the stacked differences-in-differences specification in Equation (5). Panel (a) reports effects on clearances of property and violent offenses and of drug offenses, respectively. Panel (b) reports effects on recorded crime for the same offense categories. Panel (c) reports effects on investigation outcomes, specifically police-issued sanctions and court convictions. Coefficients are normalized relative to the month immediately preceding the managerial transition ($k = -1$). Standard errors are clustered at the police-division level.

5.3 Resource levels and allocation

Having unpacked what effective managers achieve, we turn to how they achieve it. One natural channel is resources. Borough Commanders operate within centrally set budgets, but they can lobby for additional officers and budget and retain discretion over how personnel are deployed across functions and priorities. The increase in clearances due to more effective managers could therefore reflect either greater resources or a different allocation of the resources a division already has. We examine both margins in turn.

We begin by further examining workforce composition. While total workforce changes little following improvements in managerial effectiveness, there are meaningful differences

in workforce composition. Appendix Table A.8 shows that front-line police officer numbers remain essentially unchanged, and we find little systematic change in PCSOs. The main adjustment occurs among civilian staff, whose numbers decline by approximately 3% on average in the six months after a transition. Effective managers therefore do not generate higher police output by expanding front-line headcount, but do so relying on fewer back-office personnel. Although the estimates are less precise, we also find some evidence that more effective managers make greater use of Special Constables, volunteer officers who provide operational capacity without expanding total budget.

We next turn to how resources are allocated within a division. Because this allocation is not directly observed, we look for indirect evidence in the spatial pattern of the gains, asking whether the clearance increases under effective managers are concentrated in higher- or lower-crime wards, or spread evenly across the division. Specifically, we apply our stacked differences-in-differences design to ward-level data, estimating it separately for wards with below- and above-median property and violent crime within their division in the six months preceding the transition.

Appendix Table A.11 shows that the increase in total clearances is similar across both types of ward. Disaggregating by offense type, however, reveals distinct spatial patterns. The rise in property and violent clearances is significantly larger in high-crime wards (0.069 vs. 0.033, difference $p = 0.025$). Drug clearances instead increase somewhat more in low-crime wards, but this difference is not statistically significant (0.094 vs. 0.071, difference $p = 0.428$). Effective managers appear to allocate enforcement effort differently across space depending on the type of crime: the gains in property and violent clearances are concentrated in high-crime wards, where these offenses are most prevalent, while drug detections increase more evenly, if anything tilting toward lower-crime wards. This pattern is potentially suggestive of managers directing reactive investigative effort toward established hot-spots while extending proactive drug enforcement more broadly across the division.

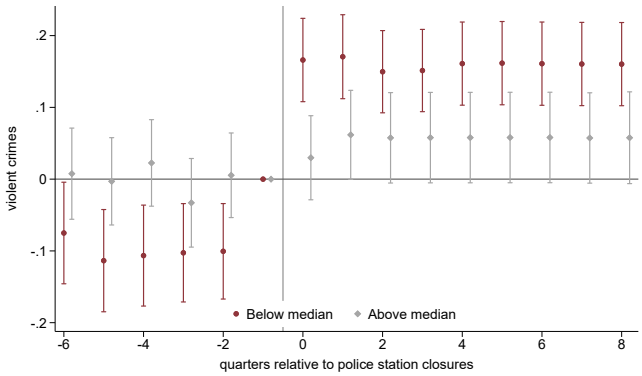
5.4 Managerial responses to policy shocks

The analysis so far has documented how effective managers achieve productivity improvements within the status quo. We now ask whether they also shape how their division responds to external policy changes that alter the conditions they operate in. This margin matters because managers and bureaucrats can be especially consequential when organizations must reallocate resources or implement policy in response to external shocks (Adhvaryu et al., 2022; Best et al., 2023). We examine two policy episodes within the MPS: an austerity-driven wave of police station closures, which cut local police infrastructure while leaving the size of the force unchanged, and the force-wide rollout of body-worn cameras (BWC), which introduced a new monitoring technology to frontline

officers. The two episodes differ in kind, letting us probe two distinct margins of managerial response: the first is a capacity-reducing shock, where an effective manager's role may be to cushion the negative consequences, whereas the second is the adoption of a potentially complementary technology, where the role may be to realize its gains.

Police station closures. Between 2008 and 2017, the Metropolitan Police closed more than 70% of its police stations as part of a broader austerity program, reducing the number of operational stations from approximately 160 to 45 while keeping force-wide police strength constant (Facchetti, 2025). Facchetti (2025) documents that, in census blocks that experienced an increase in distance to the nearest station, violent crime rose by about 9% and clearance rates declined by 2-3% relative to blocks whose station did not close. We use this setting to ask whether manager quality may mitigate the impact of the closures. Adopting the differences-in-differences specification of Facchetti (2025), we compare changes in block-level outcomes in blocks that became farther from a station to blocks that did not, before and after the closures, and split the analysis by whether the borough commander in charge at the time of the closure was above or below the median of the manager fixed-effects distribution. Figure 8 reports the resulting event-study estimates for violent crime. The average post-closure effect is 0.243 for blocks in divisions led by below-median commanders (SE = 0.031), compared with 0.055 for blocks in divisions led by above-median commanders (SE = 0.032). The difference between the two effects is -0.188 and is statistically significant at the 1% level ($p < 0.001$). While blocks under below-median commanders experience the full negative impact of the closure documented in Facchetti (2025), blocks under above-median commanders experience a substantially smaller increase in violent crime. This pattern is consistent with effective managers mitigating the operational disruption from the loss of local infrastructure.

Figure 8: Effect of police station closures on violent crime, by managerial quality



Notes: The figure reports event-study estimates of the effect of police station closures on (log) violent crime at the census-block level, following the differences-in-differences specification of Facchetti (2025). Estimates compare blocks that experienced an increase in distance to the nearest station to blocks whose station did not close, before and after the closure. The sample is split into blocks under borough commanders with below-median (red circles) and above-median (grey diamonds) manager fixed effects in periods before and of the closure, where the median is computed across the borough commanders in the largest connected set of Section 4. The horizontal axis reports time in quarters relative to the closure. Coefficients are normalized to the quarter immediately preceding the closure ($k = -1$). Vertical bars report 95% confidence intervals; standard errors are clustered at the census block level.

Body-worn cameras. The MPS rolled out body-worn cameras (BWC) across all 32 territorial divisions in a staggered fashion between October 2016 and June 2017, with divisions grouped into monthly adoption cohorts. To estimate the effect of BWC adoption and how it varies with managerial quality, we adopt a stacked differences-in-differences design. For each monthly adoption cohort, we construct a stack containing the treated divisions in that cohort and all divisions not yet adopting as of that month, retaining an event window of ± 3 months around adoption. We then estimate a difference-in-differences specification that interacts the post-treatment indicator with an indicator for whether the division's pre-adoption manager fixed effect is above the median across divisions.Appendix B3 provides additional details on the rollout and on the construction of the stacked panel.

Table 4 reports the results across our main outcomes. First, BWC adoption is associated with a meaningful increase in recorded crime, about 7% for total crime and 8.5% for property and violent offenses, in divisions with below-median managerial effectiveness (Columns 1-2). This pattern is consistent with cameras inducing more thorough incident recording (Barbosa et al., 2025), with the formalization effect concentrated where the scope to improve was largest. Second, divisions with below-median manager quality see a sharp increase in response time to incidents, of approximately 32%; in divisions with above-median manager quality this effect is offset by a large differential of -44% ($p < 0.10$). This pattern echoes broader evidence on IT-management complementarities in policing. As Garicano and Heaton (2010) and Mastrobuoni (2020) show, the pro-

Table 4: Effects of body-worn camera adoption on crime, clearances, response time, and complaints

	(Log) crime			(Log) clearances				
	All	Property and Violent	Drugs	Per worker	All	Convicted	Response time	Complaints
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Treated	0.071***	0.085***	-0.087	-0.059	-0.067	-0.062	0.320***	-0.155
	(0.016)	(0.023)	(0.080)	(0.047)	(0.042)	(0.052)	(0.102)	(0.135)
× Above-median manager FE	-0.039	-0.044	0.011	0.022	0.031	0.012	-0.437*	0.053
	(0.023)	(0.031)	(0.082)	(0.054)	(0.050)	(0.068)	(0.230)	(0.155)
Borough-by-Stack FEs	✓	✓	✓	✓	✓	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1407	1407	1407	1407	1407	1407	799	1407
Boroughs (Clusters)	29	29	29	29	29	29	29	29
Stacks	15	15	15	15	15	15	15	15
Mean Dep. Variable	7.689	7.302	4.136	-1.046	5.196	4.253	7.749	3.062

Notes: This table reports estimates from the stacked differences-in-differences specification around the time of BWC adoption (event window of ±3 months). The dependent variable in each column is, respectively, log total crime, log property and violent crime, log drug-related crime, log clearances per worker, log total clearances, log convictions, log response time, and log complaints. *Post* × *Treated* captures the average post-adoption effect for divisions with below-median pre-adoption manager fixed effects; the interaction coefficient captures the differential effect for divisions with above-median pre-adoption manager fixed effects, where the median is computed across treated divisions. All specifications control for the number of cameras issued at the division level, and absorb borough-by-stack and date-by-stack fixed effects. Standard errors clustered at the police division level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

ductivity returns to police technology are substantially larger when complemented by managerial capacity. One possible interpretation is that effective managers better integrate the new technology into dispatch and case-management workflows, while in less well-managed divisions the same technology generates operational friction. Third, we find no evidence of differential impacts on clearances or complaints (Columns 4–6 and 8).

On stop-and-search outcomes (Appendix Table B1), BWC adoption does not significantly change the volume of stops or the arrest rate. If anything, it shifts the composition of encounters in below-median-manager divisions toward lower-quality outcomes: the share of stops that yield no formal action rises by 3.1 percentage points, and the share resulting in a drugs warning falls by 2.1 percentage points. The interaction with manager quality is small and insignificant for all of these outcomes. Taken together, the results suggest that managers act as a complement to the technology where it interacts with operational decisions they directly control (e.g. dispatch, case workflow, response allocation) but their influence does not extend to responses at the officer-encounter level.[13]

[13]Our findings can be related to two recent randomized experiments of body-worn cameras. In an RCT in Brazil, Barbosa et al. (2025) document large effects on the formalization of incident reports, and a reduction in police use of force. The MPS-internal RCT implemented in 2014-15 and evaluated by Grossmith et al. (2015) finds significant reductions in citizen complaints, no overall effect on stop-and-search volumes, and a small reduction in charge rates. Our results align with this evidence in three ways: a formalization effect on recorded crime concentrated in below-median-manager divisions; insignificant clearance effects; and small composition shifts in stop outcomes that operate at the officer-encounter level.

6 Conclusions

Within a policy debate on policing that has focused largely on external inputs, such as police numbers, deployment strategies, or technology adoption, we shift our attention to a different, internal margin: the quality of the middle managers who run police divisions on a day-to-day basis. Building a novel panel of Borough Commanders in the MPS of London between 2012 and 2019, and linking it to the universe of recorded incidents and investigation outcomes, we ask how much of the variation in police output can be attributed to the middle managers in charge, and what these managers actually do to improve it.

We find that middle managers account for approximately one quarter of the variance in log clearances per worker. A one standard deviation improvement in managerial effectiveness raises log clearances per worker by about 13% percent, with essentially no movement in front-line headcount and no detectable response in reported crime. The size of the manager component is at the upper end of the public-sector literature on manager contributions, larger than the share recovered for school principals, hospital managers, or social-security office heads in comparable studies.

Taken together, our results help us understand the role that middle managers play in policing. To begin with, the productivity gains we document do not stem solely from changing how the police handle offenses within their discretion. While the rise in drug detections points to greater proactive policing, clearances for property and violent crime as well as court convictions also increase, suggesting real improvements in investigative capacity. The fact that these gains are concentrated in wards with higher pre-transition crime is consistent with effective managers targeting toward higher-need areas such as crime hot-spots. Importantly, this is not a story of more resources: consistent with the binding budget constraints of public administration, effective managers achieve higher clearances with a smaller workforce, and in particular with fewer back-office staff. In addition, effective managers help their divisions absorb changes to the status quo, both cushioning negative shocks, such as austerity-driven station closures, and realizing the gains from new technologies, such as body-worn cameras. Effective managers are especially valuable in handling exceptions and changes as they arise in the day-to-day operations of their divisions.

However, we find limited evidence that managers make lasting structural changes to how a division operates, for example by permanently altering managerial practices. The timing of the effects is telling in this regard: productivity rises quickly, almost fully within three months, once an effective manager arrives. At the same time, these productivity gains are just as quickly lost once an effective manager leaves: effective managers do not appear to leave a durable imprint on the division beyond the manager's own tenure.

Overall, these results point to the internal organization of public bureaucracies as a

meaningful and largely unexploited lever for improving the performance of a key state function such as law enforcement in a setting where external resources are likely to remain constrained. However, there is potentially space to improve police performance further if these components of managerial quality could be made portable across the organization.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Adda, Jérôme, Brendon McConnell, and Imran Rasul, "Crime and the depenalization of cannabis possession: Evidence from a policing experiment," *Journal of Political Economy*, 2014, *122* (5), 1130–1202.

Adhvaryu, Achyuta, Anant Nyshadham, and Jorge Tamayo, "Managerial quality and productivity dynamics," *The Review of Economic Studies*, 2023, *90* (4), 1569–1607.

_ , Namrata Kala, and Anant Nyshadham, "Management and shocks to worker productivity," *Journal of Political Economy*, 2022, *130* (1), 1–47.

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward, "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*, 2008, *171* (3), 673–697.

Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun, "CEO behavior and firm performance," *Journal of Political Economy*, 2020, *128* (4), 1325–1369.

_ , Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat, "The allocation of authority in organizations: A field experiment with bureaucrats," *The Quarterly Journal of Economics*, 2021, *136* (4), 2195–2242.

Barbosa, Daniel AC, Thiemo Fetzer, Caterina Soto-Vieira, and Pedro CL Souza, "Monitoring Technology: The Impact of Body-Worn Cameras on Citizen-Police Interactions," *Review of Economics and Statistics*, 2025, pp. 1–42.

Bertrand, Marianne and Antoinette Schoar, "Managing with style: The effect of managers on firm policies," *The Quarterly Journal of Economics*, 2003, *118* (4), 1169–1208.

_ , Robin Burgess, Arunish Chawla, and Guo Xu, "The glittering prizes: Career incentives and bureaucrat performance," *The Review of Economic Studies*, 2020, *87* (2), 626–655.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi, "Individuals and organizations as sources of state effectiveness," *American Economic Review*, 2023, *113* (8), 2121–2167.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

_ , _ , Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ , Joerg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Chalfin, Aaron and Justin McCrary, "Criminal deterrence: A review of the literature," *Journal of Economic Literature*, 2017, *55* (1), 5–48.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson, "Health care exceptionalism? Performance and allocation in the US health care sector," *American Economic Review*, 2016, *106* (8), 2110–2144.

Dahis, Ricardo, Laura Schiavon, and Thiago Scot, "Selecting Top Bureaucrats: Admission Exams and Performance in Brazil," *Review of Economics and Statistics*, 2025, *107* (2), 408–425.

de Leon, Maria Mercedes Ponce and Luca Rizzotti, "The Effect of Police Management on Crime and Officers' Behavior: Evidence from Compstat," *Available at SSRN 4799283*, 2024.

Draca, Mirko, Stephen Machin, and Robert Witt, "Panic on the streets of London: Police, crime, and the July 2005 terror attacks," *American Economic Review*, 2011, *101* (5), 2157–2181.

Facchetti, Elisa, "Police infrastructure, police performance, and crime: Evidence from austerity cuts," Technical Report, IFS Working Papers 2025.

Fenizia, Alessandra, "Managers and productivity in the public sector," *Econometrica*, 2022, *90* (3), 1063–1084.

Garicano, Luis and Paul Heaton, "Information technology, organization, and productivity in the public sector: Evidence from police departments," *Journal of Labor Economics*, 2010, *28* (1), 167–201.

Grossmith, Lynne, Catherine Owens, Will Finn, David Mann, Tom Davies, and Laura Baika, "Police, camera, evidence: London's cluster randomised controlled trial of body worn video," *London: College of Policing*, 2015, *10.*

Home Office, *Crime outcomes in England and Wales: Year ending March 2016* 2016.

i Vidal, Jordi Blanes and Tom Kirchmaier, "The effect of police response time on crime clearance rates," *The Review of Economic Studies*, 2018, *85* (2), 855–891.

Janke, Katharina, Carol Propper, and Raffaella Sadun, "The impact of CEOs in the public sector: Evidence from the English NHS," Technical Report, National Bureau of Economic Research 2019.

Kapustin, Max, Terrence Neumann, and Jens Ludwig, "Policing and management," Technical Report, National Bureau of Economic Research 2022.

Limodio, Nicola, "Bureaucrat allocation in the public sector: Evidence from the world bank," *The Economic Journal*, 2021, *131* (639), 3012–3040.

Mas, Alexandre, "Pay, reference points, and police performance," *Quarterly Journal of Economics*, 2006, *121* (3), 783–821.

Mastrobuoni, Giovanni, "Crime is terribly revealing: Information technology and police productivity," *The Review of Economic Studies*, 2020, *87* (6), 2727–2753.

Metcalfe, Robert D, Alexandre B Sollaci, and Chad Syverson, "Managers and productivity in retail," Technical Report, National Bureau of Economic Research 2023.

Minni, Virginia, "Making the invisible hand visible: Managers and the allocation of workers to jobs," *The Quarterly Journal of Economics*, 2026, *141* (3), 1871–1920.

Muñoz, Pablo and Cristóbal Otero, "Managers and public hospital performance," *American Economic Review*, 2025, *115* (11), 4040–4074.

_ **and Mounu Prem**, "Managers' productivity and recruitment in the public sector," *American Economic Journal: Economic Policy*, 2024, *16* (4), 223–253.

Owens, Emily and Bocar Ba, "The economics of policing and public safety," *Journal of Economic Perspectives*, 2021, *35* (4), 3–28.

Rasul, Imran and Daniel Rogger, "Management of bureaucrats and public service delivery: Evidence from the nigerian civil service," *The Economic Journal*, 2018, *128* (608), 413–446.

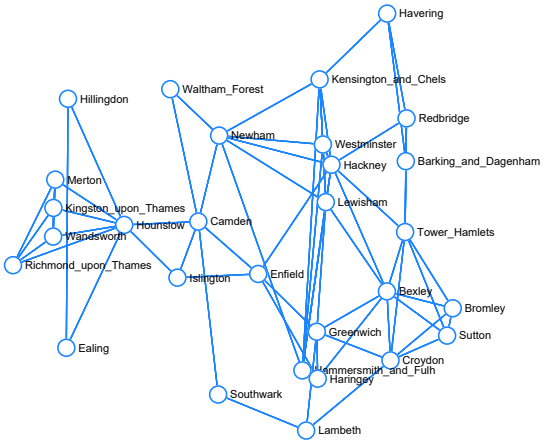
Smith, Austin V, "The Boss in Blue: Supervisors and Police Behavior," 2025.

Thaler, Richard, "An econometric analysis of property crime: Interaction between police and criminals," *Journal of Public Economics*, 1977, *8* (1), 37–51.

A Appendix Figures and Tables

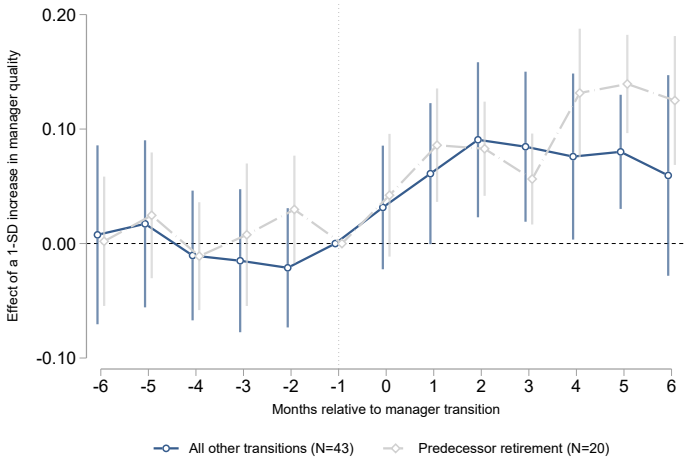
A1 Appendix Figures

Figure A.1: Connected set in MPS



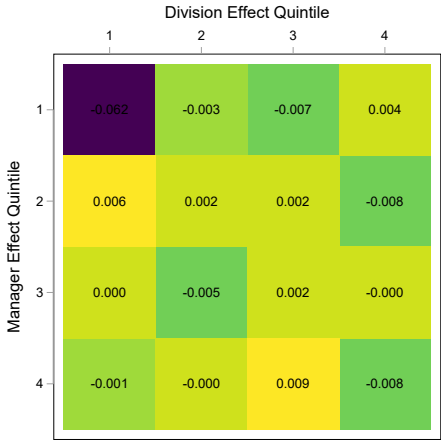
Notes: Each circle represents a police division, and a link between them means the divisions have shared a manager.

Figure A.2: Manager transitions by predecessor retirement



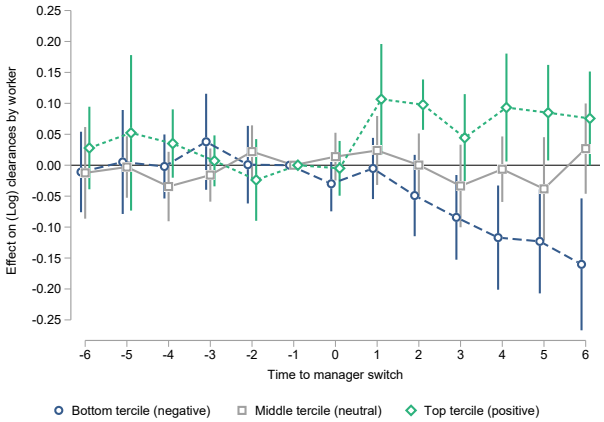
Notes: The figure reports estimates from the stacked event-study specification, allowing the effect of a change in managerial effectiveness to differ according to whether the outgoing manager retires. The figure plots the event-time effect of a one-standard-deviation improvement in manager effectiveness separately for transitions in which the predecessor retires and for all other transitions. The sample contains 20 retirement transitions and 43 other transitions. Coefficients are normalized relative to the month immediately preceding the transition ($k = -1$). All specifications include division-by-stack and date-by-stack fixed effects. Vertical bars report 95% confidence intervals.

Figure A.3: Specification check: mean residuals of fixed effects



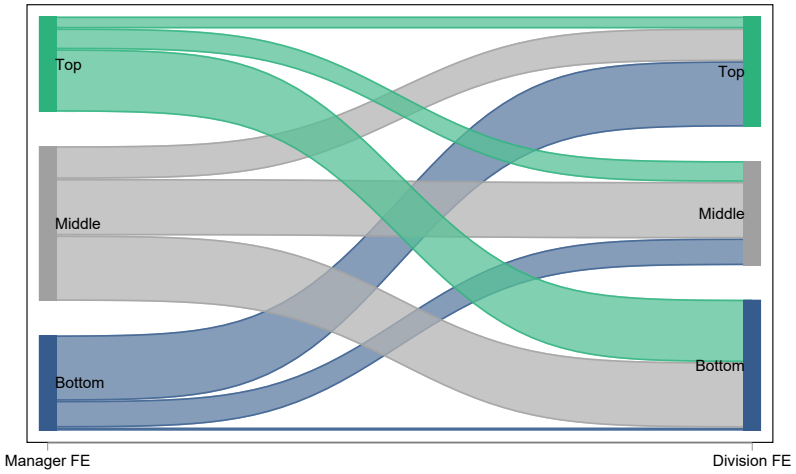
Notes: The figure shows mean residuals of estimated manager and division fixed effects from Equation (1), across quintiles of fixed effects. The figure shows no systematic patterns in the size of residuals across quintiles.

Figure A.4: Event study on log clearances per worker by manager effectiveness



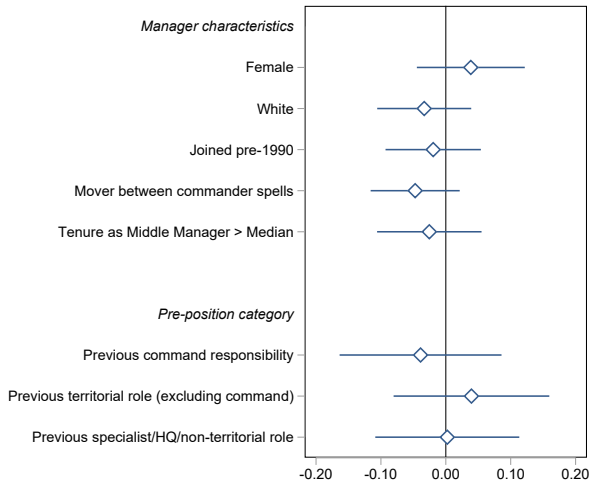
Notes: This figure reports estimates from a TWFE event-study specification around manager transitions. Transitions are classified into terciles based on the change in empirical-Bayes-adjusted manager fixed effects between the incoming and outgoing manager. Bottom-tercile transitions correspond to the largest declines in managerial effectiveness, middle-tercile transitions to relatively small changes, and top-tercile transitions to the largest improvements. The specification includes event-time indicators interacted with transition tercile, division fixed effects, and calendar-month fixed effects. The event window extends from six months before to six months after the transition, and the month immediately preceding the transition, $k = -1$, is omitted. Observations that do not belong to an eligible event window remain in the estimation sample and contribute to the estimation of the division and calendar-month fixed effects, but not to the event-time coefficients. Vertical bars report 95% confidence intervals based on standard errors clustered at the police-division level.

Figure A.5: Match between managers and divisions



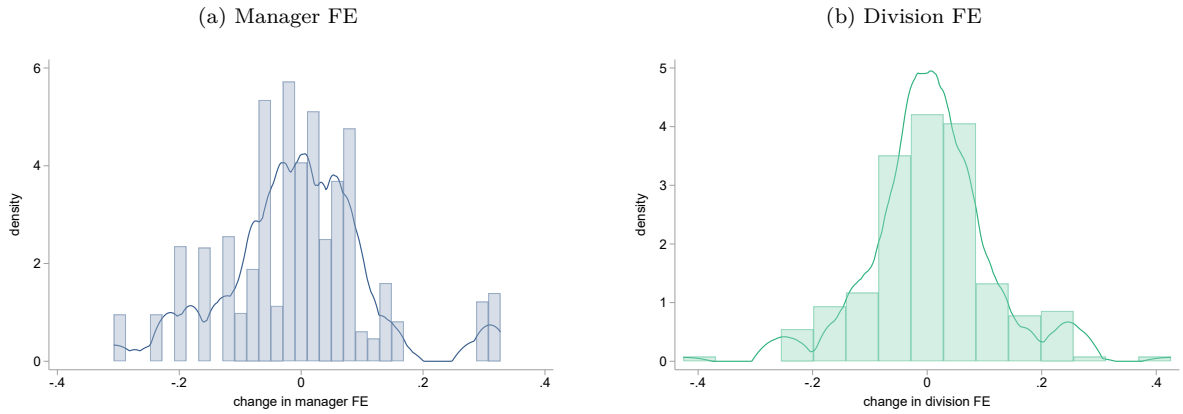
Notes: The figure shows the average fractions of managers with different levels of fixed effects, who are matched/assigned to different police divisions with different fixed effects.

Figure A.6: Correlates of manager fixed effects



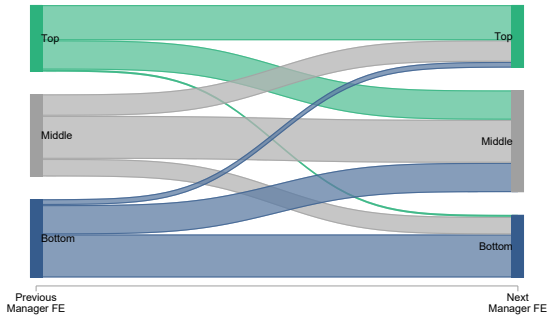
Notes: The figure reports estimates from a multivariate regression of manager fixed effects on manager characteristics and prior-position indicators, included jointly in the same specification. Pre-position categories refer to the role held before the manager's first observed Borough Commander spell. Each observation is a manager. Horizontal bars report 95% confidence intervals based on robust standard errors.

Figure A.7: Distribution of changes in manager and division fixed effects



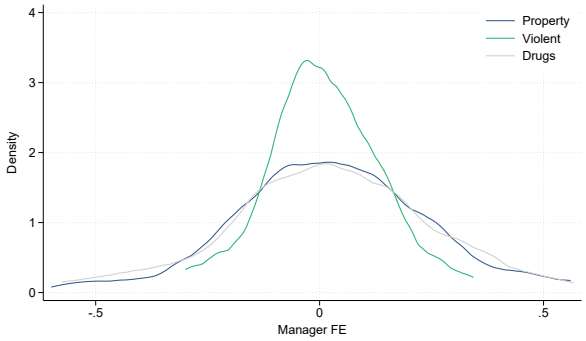
Notes: This figure shows the distribution of changes in manager and division fixed effects corresponding to each transition involving a different manager for the division. The left panel shows the distribution of changes in manager FEs, defined as $\Delta ManEff_d$ in Equation 2. The right panel shows the distribution of changes in the division fixed effects.

Figure A.8: Changes in division quality in manager moves



Notes: The figure shows changes in the assignment of managers (based on their fixed effects) to divisions over time. It shows the proportion of divisions which had different levels, in terciles, of pre-move manager FE and how their subsequent manager FE changed after a given transition.

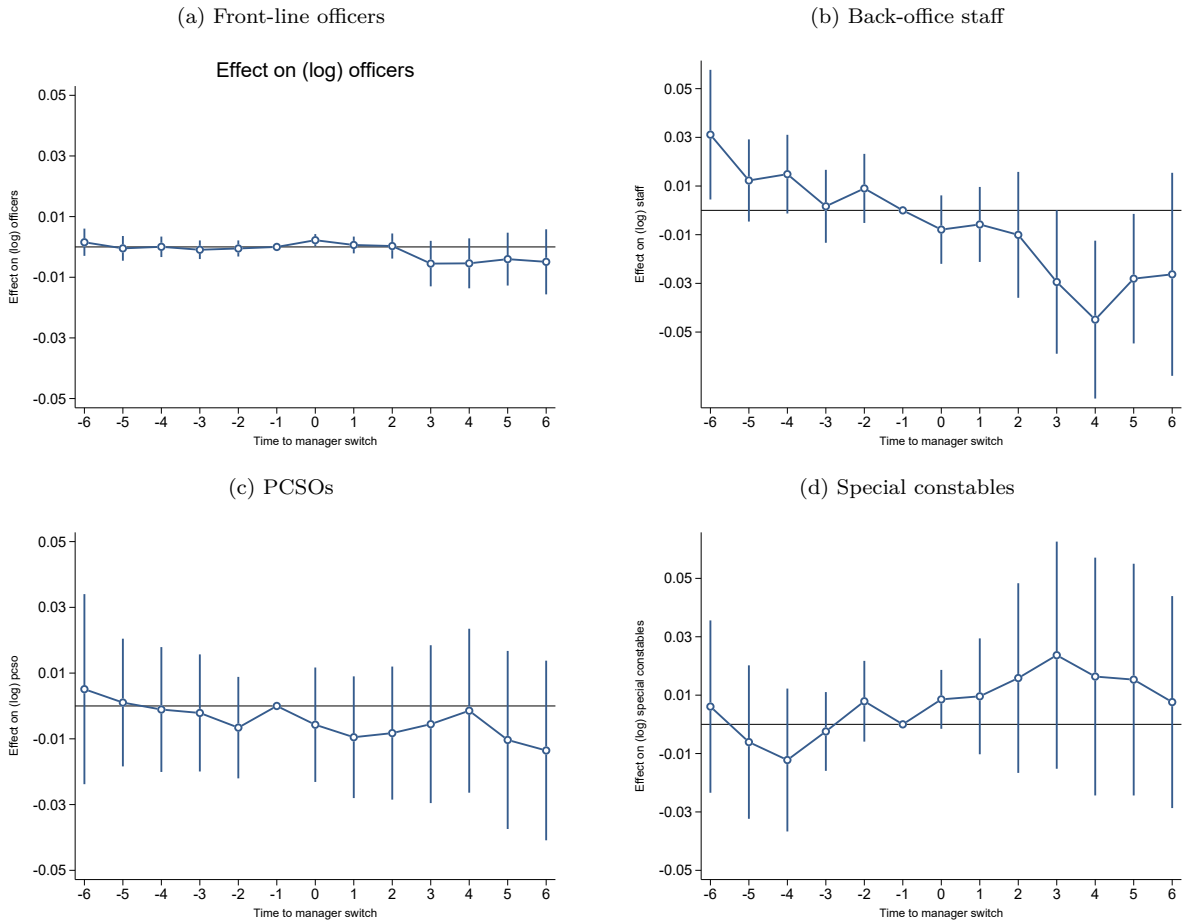
Figure A.9: Multi-tasking across crime types



	All	Property	Violent	Drugs
All	1			
Property	0.779	1		
Violent	0.586	0.386	1	
Drugs	0.843	0.486	0.261	1

Notes: On the left side we report the distributions of Empirical Bayes-adjusted manager fixed effects estimated separately for (log) clearances per worker in violent, property, and drug-related offenses. On the right side, we report the pairwise correlations between these offense-specific manager effects.

Figure A.10: Effects of manager effectiveness on police workforce allocation



Notes: The figure reports event-study estimates from Equation 3, showing the evolution of division-level log workforce before and after a managerial transition. Panel (a) shows front-line officers, Panel (b) covers back-office staff, Panel (c) PCSOs, and panel (d) Special Constables. Vertical bars report 95% confidence intervals.

A2 Appendix Tables

Table A.1: MPS rank hierarchy

Rank Level	Police Rank	Minimum Tenure for Promotion
Top Ranks	Commissioner of Police	
	Deputy Commissioner	No set tenure (extensive senior
	Assistant Commissioners	leadership required)
	Deputy Assistant Commissioners	
Middle Ranks	Chief Superintendents	3+ years as a Superintendent
	Superintendents	5+ years in policing
	Chief Inspectors	2+ years as an Inspector
	Inspectors	2+ years as a Sergeant
Low Ranks	Sergeants	Minimum 2 years as a Constable
	Constables	Entry level

Table A.2: Investigation outcome types in Metropolitan Police Service records

Investigation outcome category	Mean
Awaiting court outcome	0.02
Court case unable to proceed	0.41
Court result unavailable	3.00
Defendant found not guilty	1.21
Defendant sent to Crown Court	0.01
Formal action is not in the public interest	0.00
Investigation complete; no suspect identified	42.63
Local resolution	1.23
Offender deprived of property	0.05
Offender fined	0.85
Offender given a caution	2.40
Offender given a drugs possession warning	1.52
Offender given absolute discharge	0.01
Offender given community sentence	1.01
Offender given conditional discharge	0.47
Offender given penalty notice	0.84
Offender given suspended prison sentence	0.54
Offender ordered to pay compensation	0.07
Offender otherwise dealt with	0.10
Offender sent to prison	1.21
Status update unavailable	30.23
Suspect charged as part of another case	0.16
Unable to prosecute suspect	0.00
Under investigation	12.02
Total	100.00
N	5,956,447

Notes: This table reports the distribution of investigation outcome types recorded by the Metropolitan Police Service (MPS) for incidents with valid investigative outcomes between 2012 and 2019. The 24 categories reflect all possible police and court dispositions recorded at the end of an investigation.

Table A.3: Comparison of Adjusted R2 on police division output

	(Log) clearance				(Log) crime			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N	2523	2523	2523	2523	2523	2523	2523	2523
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Division FE	No	Yes	Yes	No	No	Yes	Yes	No
Manager FE	No	No	Yes	No	No	No	Yes	No
Manag-by-Division FE	No	No	No	Yes	No	No	No	Yes
R sq.	0.309	0.889	0.933	0.935	0.053	0.980	0.985	0.986
Adj. R sq.	0.284	0.884	0.927	0.928	0.020	0.979	0.984	0.985
RMSE	0.383	0.154	0.123	0.121	0.377	0.055	0.048	0.047
P-value			0.000				0.000	

Notes: This table shows the explanatory power of different fixed effects specifications for log total clearances (top panel) and log total crime (bottom panel). All models include time fixed effects. Columns (1) and (5) show time FE only; Columns (2) and (6) add division FE; Columns (3) and (7) add manager FE; Column (4) and (8) include manager-division interactions, with marginal additional explanatory power.

Table A.4: Pre-trends in division outcomes before manager transitions

	(Log) clearances			(Log) crime	
	All	Index	Drugs	Index	Drugs
	(1)	(2)	(3)	(4)	(5)
t = -6	0.030	-0.006	0.076	-0.001	0.085*
	(0.022)	(0.025)	(0.048)	(0.010)	(0.048)
t = -5	0.023	0.019	0.020	-0.013	0.061
	(0.019)	(0.021)	(0.041)	(0.009)	(0.044)
t = -4	0.004	-0.001	0.027	-0.006	0.046
	(0.019)	(0.021)	(0.044)	(0.008)	(0.040)
t = -3	0.014	-0.003	0.036	-0.011	0.066
	(0.017)	(0.022)	(0.041)	(0.010)	(0.039)
t = -2	0.012	0.007	0.010	-0.001	0.037
	(0.018)	(0.023)	(0.026)	(0.007)	(0.025)
F-stat (joint pre=0)	0.58	0.76	0.79	0.72	0.89
p-value	0.716	0.585	0.564	0.615	0.498
Observations	12,779	12,779	12,779	12,779	12,779

Notes: Each column reports coefficients on event-time dummies $t = -6, \ldots, -2$, with $t = -1$ as the omitted reference period, from a stacked event-study specification using only the pre-transition sample. The outcome is indicated in the column header. The final two rows report the joint F-test that all pre-period coefficients are jointly zero, with the associated p-value. Standard errors in parentheses are clustered at the division level.

Table A.5: Correlates of manager exits

	Lateral move	Promotion	Retirement	Territorial policing
	(1)	(2)	(3)	(4)
Panel A. Manager characteristics				
Female	0.147	0.029	-0.132	-0.007
	(0.116)	(0.083)	(0.082)	(0.112)
White	0.026	0.070	0.050	-0.032
	(0.140)	(0.097)	(0.103)	(0.135)
Joined pre-1990	-0.146	-0.204**	0.301***	-0.117
	(0.100)	(0.079)	(0.076)	(0.098)
Tenure on the job (months)	-0.003	-0.001	0.005*	-0.004
	(0.003)	(0.002)	(0.003)	(0.003)
Manager FE (std)	-0.009	-0.008	-0.049	-0.027
	(0.049)	(0.037)	(0.046)	(0.048)
Panel B. Average outcomes during spell				
Average clearance during spell	0.195	-0.134	-0.602	-0.035
	(0.448)	(0.294)	(0.488)	(0.404)
Average crime during spell	-2.240**	0.528	1.411	-2.349**
	(1.038)	(1.024)	(0.897)	(1.087)
N spells	126	126	126	126
Baseline mean	0.484	0.183	0.270	0.357

Notes: The table reports linear probability models of career outcomes following completed Borough Commander spells. The unit of observation is a manager-spell. Columns (1)-(3) indicate whether the spell is followed by, respectively, a lateral move, promotion, or retirement. The remaining 6.3% of spells are disciplinary exits or departures outside policing, and are coded as zero. Column (4) indicates whether the manager's subsequent position is within territorial policing. Panel A reports manager characteristics, including the standardized manager fixed effect. Panel B reports average clearance and crime outcomes during the command spell. All variables shown in Panels A and B enter jointly in each specification. Standard errors, clustered at the manager level, are reported in parentheses. The final rows report the number of completed spells and the mean of the dependent variable. $p < 0.10$, $p < 0.05$, $p < 0.01$.

Table A.6: Additional bias-correction methods of variance decomposition

Method	Var(Outcome)	Var(Manager)	Var(Division)	Cov(Manager, Division)
	(1)	(2)	(3)	(4)
Raw AKM	0.069	0.037	0.045	-0.029
	[100.0]	[53.0%]	[64.7%]	
Andrews	0.069	0.034	0.042	-0.027
	[100.0]	[48.9%]	[61.1%]	
EB Shrinkage	0.069	0.025	0.034	-0.020
	[100.0]	[36.5%]	[48.5%]	
Covariance Shrinkage	0.069	0.017	0.024	-0.009
	[100.0]	[24.7%]	[35.3%]	
Split-sample	0.069	0.033	0.042	-0.024
	[100.0]	[47.6%]	[61.3%]	
Monte Carlo corrected	0.069	0.035	0.044	-0.028
	[100.0]	[50.7%]	[63.1%]	

Notes: This table shows the variance-covariance decomposition of log clearance per worker based on the AKM model in Equation (1). The first four rows reproduce the bias-correction methods in Table 3: *Raw AKM*, the finite-sample trace correction of Andrews et al. (2008), scalar empirical-Bayes shrinkage, and the covariance shrinkage estimator of Best et al. (2023). The last two rows report two additional checks: a split-sample variance correction (e.g., Best et al., 2023; Dahis et al., 2025), and a Monte Carlo bias correction following Adhvaryu et al. (2023). Brackets report each term as a share of the total variance in column 1.

Table A.7: Average effect of managerial transitions on log clearances

	(Log) clearances					
	Per worker	All	Property and Violent	Drugs	Cleared by police	Convicted
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treated	-0.019	-0.022*	-0.021	-0.033	-0.013	-0.023
	(0.013)	(0.013)	(0.013)	(0.029)	(0.017)	(0.015)
Post × Treated × Change	0.079***	0.075***	0.058***	0.099***	0.076***	0.068***
	(0.009)	(0.008)	(0.010)	(0.027)	(0.011)	(0.010)
Division-by-Stack FEs	✓	✓	✓	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓	✓	✓	✓
Observations	12775	12775	12775	12775	12774	12775
Divisions (Clusters)	29	29	29	29	29	29
Stacks	63	63	63	63	63	63
Mean Dep. Variable	-0.958	5.454	4.921	4.206	4.726	4.492
P-value Post × Treated	0.140	0.080	0.110	0.242	0.416	0.112
P-value Post × Treated × Change	0.000	0.000	0.000	0.002	0.000	0.000

Notes: This table reports the estimated effects from the stacked differences-in-differences specification in Equation (5), using log clearance by investigation type as the outcome. The coefficients capture the baseline effect of a manager transition with an average change in managerial effectiveness, and the differential effect associated with the standardized change in manager fixed effects. Standard errors are clustered at the police division level.

Table A.8: Average effect of managerial transitions on log workforce

	(Log) workforce				
	All	Officers	PCSOs	Staff	Special Constables
	(1)	(2)	(3)	(4)	(5)
Post × Treated	-0.003	-0.003	-0.001	0.007	-0.011
	(0.003)	(0.003)	(0.011)	(0.021)	(0.020)
Post × Treated × Change	-0.004*	-0.002	-0.007	-0.033***	0.015
	(0.003)	(0.003)	(0.012)	(0.010)	(0.016)
Division-by-Stack FEs	✓	✓	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓	✓	✓
Observations	12775	12775	12775	12775	12775
Divisions (Clusters)	29	29	29	29	29
Stacks	63	63	63	63	63
Mean Dep. Variable	6.412	6.290	3.585	3.041	4.550
P-value Post × Treated	0.270	0.316	0.930	0.792	0.588
P-value Post × Treated × Change	0.110	0.468	0.630	0.004	0.432

Notes: This table reports the average estimated effects from the stacked differences-in-differences specification described in Equation (5), using log workforce by officer function as the outcome. The coefficients capture the baseline effect of a manager transition with an average change in managerial effectiveness, and the differential effect associated with the standardized change in manager fixed effects. Standard errors are clustered at the police division level.

Table A.9: Average effect of managerial transitions on log crimes

	(Log) offenses		
	All	Property and Violent	Drugs
	(1)	(2)	(3)
Post × Treated	0.008	0.009	-0.029
	(0.006)	(0.006)	(0.028)
Post × Treated × Change	0.000	-0.005	0.089***
	(0.005)	(0.005)	(0.019)
Division-by-Stack FEs	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓
Observations	12775	12775	12775
Divisions (Clusters)	29	29	29
Stacks	63	63	63
Mean Dep. Variable	7.810	7.384	4.468
P-value Post × Treated	0.206	0.174	0.272
P-value Post × Treated × Change	0.974	0.334	0.000

Notes: This table reports the average estimated effects from the stacked differences-in-differences specification described in Equation (5), using log crime by offense type as the outcome. The coefficients capture the baseline effect of a manager transition with an average change in managerial effectiveness, and the differential effect associated with the standardized change in manager fixed effects. Standard errors are clustered at the police division level.

Table A.10: Average effect of managerial transitions on log clearances post-BCU mergers

	(Log) clearances					
	Per worker	All	Property and Violent	Drugs	Cleared by police	Convicted
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treated	-0.011	-0.016	-0.028*	-0.014	-0.009	-0.009
	(0.016)	(0.016)	(0.016)	(0.032)	(0.020)	(0.017)
Post × Treated × Change	0.074***	0.072***	0.056***	0.084***	0.075***	0.055***
	(0.012)	(0.012)	(0.014)	(0.027)	(0.014)	(0.015)
Division-by-Stack FEs	✓	✓	✓	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓	✓	✓	✓
Observations	12779	12779	12779	12779	12778	12779
Divisions (Clusters)	29	29	29	29	29	29
Stacks	63	63	63	63	63	63
Mean Dep. Variable	-0.958	5.454	4.921	4.206	4.726	4.491

Notes: This table reports the estimated effects from the stacked differences-in-differences specification in Equation (5), augmented with division fixed effects interacted with an indicator for the post-merger period. The mergers affected 29 out of 32 divisions between January 2017 and October 2018. The dependent variable in each column is the log of clearances by investigation type. Standard errors are clustered at the police division level.

Table A.11: Average effect of managerial transitions on log clearances at the ward level, by pre-transition property and violent crime

	(Ihs) clearances									
	All		Property and Violent		Drugs		Cleared by police		Convicted	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post × Treated	-0.024	-0.022	-0.001	-0.027	-0.021	0.006	-0.009	-0.005	-0.014	-0.023
	(0.017)	(0.015)	(0.016)	(0.018)	(0.025)	(0.029)	(0.022)	(0.016)	(0.015)	(0.016)
Post × Treated × Change	0.063***	0.073***	0.033***	0.069***	0.094***	0.071**	0.066***	0.063***	0.035**	0.073***
	(0.013)	(0.012)	(0.011)	(0.014)	(0.017)	(0.029)	(0.014)	(0.015)	(0.015)	(0.014)
Ward-by-Stack FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Split	$< p50$	$\geq p50$	$< p50$	$\geq p50$	$< p50$	$\geq p50$	$< p50$	$\geq p50$	$< p50$	$\geq p50$
Observations	138476	132912	138476	132912	138476	132912	138476	132912	138476	132912
Divisions (Clusters)	29	29	29	29	29	29	29	29	29	29
Stacks	63	63	63	63	63	63	63	63	63	63
Mean Dep. Variable	2.261	3.265	1.749	2.702	1.126	1.935	1.559	2.479	1.325	2.262
P-value Post × Treated	0.152	0.146	0.920	0.146	0.400	0.800	0.668	0.778	0.368	0.182
P-value Post × Treated × Change	0.004	0.004	0.036	0.008	0.000	0.030	0.000	0.012	0.048	0.004
P-value Difference		0.460		0.025		0.428		0.890		0.046

Notes: This table reports the estimated effects from a ward-level version of the stacked differences-in-differences specification in Equation (5), splitting wards depending on the pre-transition level of property and violent crime. Specifically, odd columns report estimates from a sample of wards with property and violent crime below the respective division median in the six months before the transition, while even columns report estimates a sample of above the median wards. The dependent variable in each column is the inverse hyperbolic sine of clearances by offense and investigation type. Standard errors are clustered at the police division level.

B Data Appendix

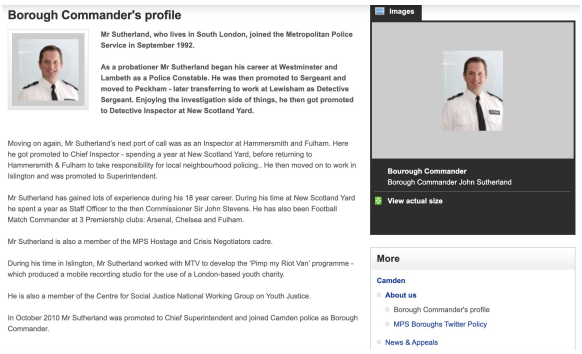
B1 Borough Commanders

We collect information on borough commanders in the MPS through a combination of Freedom of Information (FOI) requests, web-scraping from web archives, and local newspapers. We begin by scraping all archived websites relative to local police divisions and borough commanders from Wayback Machine and National Archive, from 2003 to 2019. Once archived, websites typically remain on the Wayback Machine, allowing for a complete retrospective review of relevant information. We conducted systematic scraping between September 2022 and August 2024.[1] From each archived webpage, we extract information on the name, bio and picture of the middle manager (borough commander) responsible for the police division. From these pages, we reconstruct the full career history of the middle managers, including key dates for appointments, exit and retirement where applicable. We supplement this with information from news reports, and FOI requests to build the final panel dataset of at the division-manager-date level.

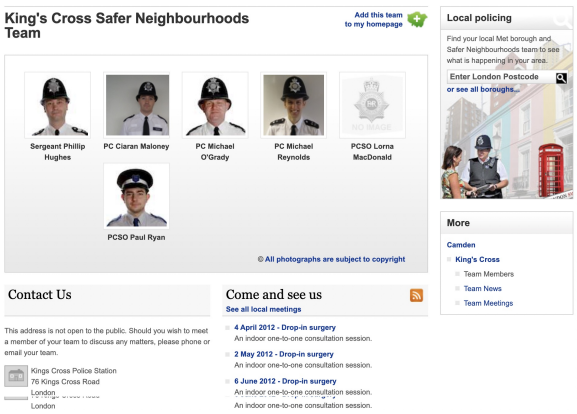
Figure B1 shows an example of the information contained on an archived website for the borough of Camden in 2012.

Figure B1: Example Information from Web Archives

(a) Middle Manager



(b) Safer Neighborhood Team



Notes: Example of web-pages from the borough of Camden in 2012. Source: The National Archives.

B2 Alternative source of police output: sanction-detection data

To complement the incident-level police records used in the main analysis, we collected an entirely separate source of borough-level police output. Specifically, we obtained monthly counts of recorded offenses and sanctioned detections (SDs) covering April 2005 to March 2016 at the borough level from archived versions of the publicly released MPS

[1] The specific websites are http://web.archive.org/ and https://webarchive.nationalarchives.gov.uk/.

tables (http://maps.met.police.uk/tables.htm), retrieved through the UK National Archives.[2] Offenses and SDs are reported in nine categories: violence against the person, sexual offenses, robbery, burglary, criminal damage, drugs, theft, vehicle crime, and other crime. We aggregate counts to total SDs and total offenses, and construct per-worker measures by dividing by total division workforce. After merging with the workforce data described in Section 3, which is available from April 2008 onward, the resulting panel covers April 2008 to March 2016. Compared to the police records, it covers a longer and earlier time window (2008–2016 against 2012–2019 in the main analysis), with only the 2012–2016 sub-window in common.

The two outcomes are designed to capture the same Home Office definition of a crime resolved with a formal sanction, including charges, court summons, adult or youth cautions, Penalty Notice for Disorder, cannabis warning, and offense taken into consideration (TIC) at sentencing. While the main analysis "clearances" are reconstructed by us from incident-level outcomes using the more granular categorization available in the police records data, the published SDs are already-aggregated as borough-by-month counts compiled by the MPS under the Home Office counting rules in force at the time of publication. The sanction-detection framework was the Home Office's official measure of solved crimes until April 2014, when it was replaced by a broader Outcomes Framework. The bulk of our 2008–2016 window falls under the sanction-detection regime.

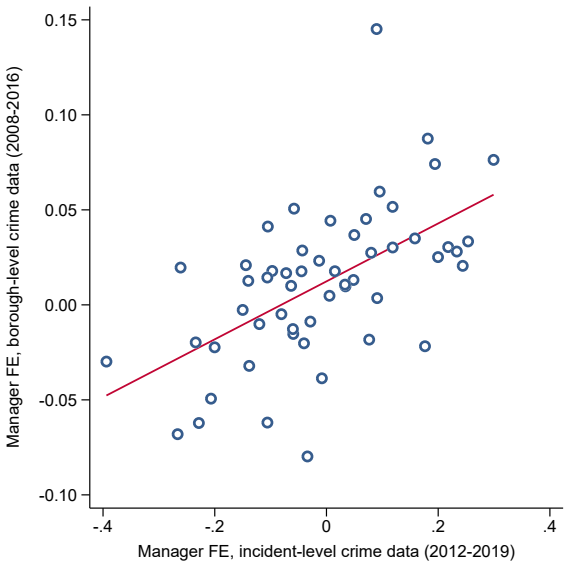
We estimate Equation (1) on the borough-by-month panel, using log SDs per worker as the outcome, with the same connected-set restrictions and empirical Bayes adjustment as in the main analysis. The mover graph in this sample gives 11 connected sets, including a larger set covering 19 divisions, a smaller set covering four divisions, and nine singleton sets containing one division each, in which no manager mobility is observed during the sample window. The singleton sets are excluded. The estimation sample therefore consists of 23 divisions in two connected sets, for a total of 2,185 division-month observations, and we recover fixed effects for 79 middle managers, of whom 54 are also in the main estimation sample. Manager fixed effects are normalized to mean zero within each connected set. Despite covering a longer time window, the alternative sample has fewer divisions and managers in the connected sets than the main sample, This feature reflects the fact that, as discussed in Section 2.2, expected manager tenure was longer prior to 2012 (a three-year minimum, against approximately two years under current policy). As a result, the alternative sample is a lower-power setting than the main one, and we view it as a check on *measurement* rather than on the precision of the recovered effects.

Figure B2 compares the empirical-Bayes-adjusted manager fixed effects recovered from the two samples for the 54 managers in common. The two sets of estimates are positively correlated, with a correlation of 0.57. This correlation is reassuring as it indicates that

[2]One month, March 2015, is missing in the source and is linearly interpolated within borough for each offense and SD type.

managers identified as high- (or low) performing in the main analysis tend to be identified as high- (or low-) performing in the alternative sample.

Figure B2: Correlation between manager fixed effects across data sources



Notes: This figure plots empirical-Bayes-adjusted manager fixed effects recovered from the alternative borough-level SD sample (2008–2016, y-axis) against empirical-Bayes-adjusted manager fixed effects recovered from the main incident-level sample (2012–2019, x-axis), for the 54 managers with FEs estimated in both samples. The correlation across the 54 observations is 0.57.

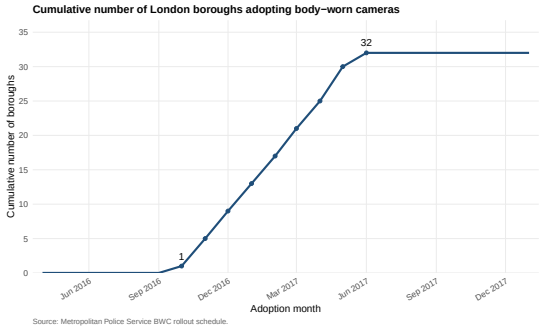
B3 Body-worn cameras

This appendix provides additional details on the rollout of body-worn cameras across MPS and on the stacked differences-in-differences design used in Section 5.4.

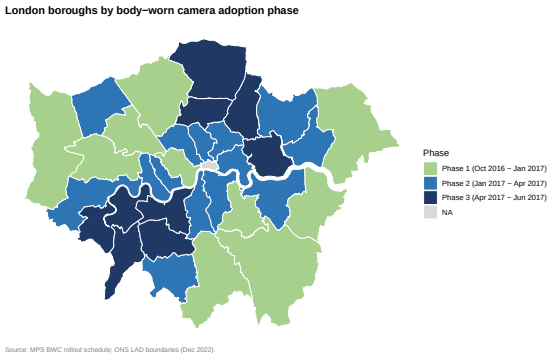
Body-worn cameras were rolled out across the 32 London boroughs in a staggered fashion between October 2016 and June 2017, with divisions grouped into monthly adoption cohorts, as shown in Figure B3. Within each borough, the rollout was coordinated by a central MPS team and completed simultaneously across the borough's police stations over a five-week window. After adoption, the average number of cameras per borough was 449, around .83 per officer. We collected administrative records from the MPS documenting the dates on which BWC equipment was issued to each borough, as well as the stock of cameras per borough.[3]

[3]See the MPS Freedom of Information disclosure on the rollout: https://www.met.police.uk/foi-ai/metropolitan-police/disclosure-2024/january-2024/roll-out-body-worn-cameras-across-london-boroughs/.

Figure B3: Adoption of Body-Worn Cameras Across London



(a) Cumulative Number of Adopting Boroughs



(b) Phases of Adoption

Because all 32 boroughs adopted body-worn cameras within an eight-month window, there is no permanently untreated comparison group. We therefore implement a stacked differences-in-differences design and construct one sub-experiment (*stack*) for each monthly adoption cohort. Stack s contains the boroughs that adopt in cohort s together with all boroughs that have not yet adopted as of the adoption date of cohort s. Within each stack, we retain a symmetric event window of $k \in \{-3, \ldots, +3\}$ months relative to adoption. The panel used for estimation appends all stacks, with each borough $\times$ stack cell treated as a separate unit.

To estimate the average effect of BWC adoption and how it varies with managerial quality, we estimate the following static specification:

$$Y_{ist} = \beta\,(\text{Post}_{ist} \times \text{Treated}_{is}) + \beta^{M}\,(\text{Post}_{ist} \times \text{Treated}_{is} \times \mathbf{1}_{i}^{+,M}) + \alpha_{is} + \lambda_{st} + \varepsilon_{ist}, \quad \text{(B1)}$$

where $\text{Post}_{ist} = \mathbf{1}[k_{ist} \geq 0]$ is an indicator for the post-adoption period; $\text{Treated}_{is} = 1$ if borough i is the treated unit in stack s; $\mathbf{1}_{i}^{+,M} = \mathbf{1}[\hat{\mu}_{i}^{\text{pre}} > \tilde{\mu}]$ is an indicator for whether borough i's pre-adoption manager fixed effect is above the cross-borough median. We compute this considering the distribution of average manager FEs in the pre-adoption window (between $t-1$ and $t-4$). α_{is} is a borough-by-stack fixed effect, and λ_{st} is a date-by-stack fixed effect. The coefficient β captures the average post-adoption effect for below-median-manager divisions, and β^{M} the differential effect for above-median-manager divisions. Standard errors are clustered at the police division level. We find no evidence of correlation between pre-adoption manager fixed effects and adoption timing.

We also obtained publicly available Stop and Search records from January 2014 to March 2018 at the individual stop level with information on borough, date, and the outcome of each encounter. We aggregate these records to the borough$\times$month level, constructing (1) the total value of stop-and-search encounters; (2) the fraction of stops

resulting in arrest, any positive outcome including arrest, summons, or penalty notice, no further action, a drug warning, and a local resolution. Table B1 reports estimates of Equation (B1) on stop-and-search outcomes.

Table B1: Effects of body-worn camera adoption on stop and search

	Stop and search	Arrest rate	Positive outcome	Nothing found	Drugs warning	Local resolution
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Treated	0.013	-0.008	-0.008	0.031**	-0.021**	-0.002
	(0.127)	(0.009)	(0.009)	(0.013)	(0.010)	(0.003)
× Above-median manager FE	0.013	0.003	0.002	-0.001	0.000	-0.001
	(0.111)	(0.011)	(0.011)	(0.017)	(0.011)	(0.003)
Borough-by-Stack FEs	✓	✓	✓	✓	✓	✓
Date-by-Stack FEs	✓	✓	✓	✓	✓	✓
Observations	1036	1036	1036	1036	1036	1036
Boroughs (Clusters)	30	30	30	30	30	30
Stacks	15	15	15	15	15	15
Mean Dep. Variable	5.439	0.210	0.237	0.677	0.079	0.006

Notes: This table reports estimates from the stacked differences-in-differences specification around the time of BWC adoption (event window of ±3 months). The dependent variable in column (1) is the (log) total number of stop-and-search encounters; columns (2)–(6) report the share of stops resulting in, respectively, an arrest, any positive outcome (arrest, summons, or penalty notice), no further action, a drugs warning, or a local resolution. *Post* × *Treated* captures the average post-adoption effect for divisions with below-median pre-adoption manager fixed effects; the interaction coefficient captures the differential effect for divisions with above-median pre-adoption manager fixed effects, where the median is computed across treated divisions. All specifications control for the number of cameras issued at the division level, and absorb borough-by-stack and date-by-stack fixed effects. Standard errors clustered at the police division level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.