Middle Managers in the Public Sector: Evidence from Policing*

Elisa Facchetti†, Arianna Ornaghi‡, Eddy Zou§

October 28, 2026

Abstract

We study the role of middle managers in shaping public-sector performance, focusing on Borough Commanders who lead territorial divisions within the London Metropolitan Police Service. We reconstruct their careers and assignments and link them to administrative data on police outcomes and workforce. Exploiting manager mobility across divisions, we find that individual managers account for approximately one quarter of the variance in police productivity, measured by crime cleared per worker. More effective managers increase clearances across offense types, raise both police-issued sanctions and court convictions, and expand proactive enforcement without increasing frontline staffing. These gains emerge rapidly after managers arrive and largely disappear after they leave. Effective managers also shape organizational responses to policy shocks: they mitigate increases in violent crime following police station closures and limit deterioration in response times following body-worn camera adoption. Finally, residents exposed to more effective managers report greater trust in the police and feel safer.

*We thank Charles Angelucci, Sarah Eichmeyer, Imran Rasul, Pedro Souza, Guo Xu, and all participants at the Bocconi FIBRE workshop, RFBerlin Labor 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

In large public organizations, budgets, incentives, and policies are often determined centrally, but many decisions about how services are delivered are left to local managers. These managers exercise discretion over how fixed resources are organized and deployed, and shape how central decisions are implemented (Aghion and Tirole, 1997; Dixit, 2002). Consistent with this, a growing literature documents that individual managers are an important driver of performance in public organizations (Fenizia, 2022; Muñoz and Otero, 2025). Less is known, however, about *how* managers affect performance.

We study this question in the context of the Metropolitan Police Service (MPS) of London. We focus on Borough Commanders, senior officers who lead territorial police divisions and sit between central leadership and frontline teams. Borough Commanders face precisely the tension between central constraints and local discretion described above. Divisional budgets and staffing levels are set centrally, and commanders lack the incentive tools that private sector managers typically use to motivate workers, such as performance pay or control over hiring and dismissal. At the same time, they retain considerable discretion over operational decisions that are known to shape police outcomes, such as how to handle low-level offenses or how to allocate resources within their division (Blanes i Vidal and Kirchmaier, 2018; Adda et al., 2014; Draca et al., 2011).

Three features of the MPS make it particularly well suited for studying how managers matter. First, we observe an unusually rich set of outcomes that commanders can influence: from crime and clearances, which we follow to final resolution, to public perceptions of the police. We can therefore document not only whether managers matter, but what margins of police activity they change. Second, commanders turn over frequently, and because their decisions affect policing output over short horizons, we can trace both how quickly managers' effects emerge after a commander arrives and whether they persist after the commander leaves. This speaks to whether managerial contributions become embedded in the organization or remain tied to the individual. Third, the MPS periodically introduces policy changes, such as police station closures and the rollout of body-worn cameras, which commanders must implement locally. This allows us to study how managers of different effectiveness adapt to the same change in their operating environment.

To study middle managers in the MPS, we assemble a novel dataset covering the universe of Borough Commanders from 2005 to 2019, combining Freedom of Information requests with archival web sources. Over this period, the MPS employed 191 Borough Commanders across 32 territorial divisions. We link these records 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, whether a police-issued sanction or a court decision. Our analysis focuses

on this period, during which 102 commanders held posts, with an average tenure of approximately two years.

We use these data to implement 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. This model exploits the rotation of Borough Commanders across territorial divisions to separate managerial contributions from persistent differences across the areas they oversee. Our main objects of interest are manager fixed effects, which capture the portable component of performance that follows commanders across assignments. The outcome we focus on is (log) *clearances per worker*, defined as the number of cleared incidents relative to the division's total workforce. Clearances capture the ability of the police to solve (i.e., clear) an incident and are a standard measure of police performance used in the economics of crime literature.

Identification 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 manager-division match quality. The institutional setting, where appointments and transfers are governed by a centralized posting panel, generates movements across divisions that are largely outside individual managers' control. In line with this, event-study estimates around transitions show no differential pre-trends. We also find symmetric changes in performance when divisions switch from higher- to lower-performing managers and vice versa, supporting the interpretation of managerial effectiveness as portable across divisions.

The model also poses two estimation 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. We identify one large connected set containing 93 managers and 29 territorial divisions, which forms the basis of our analysis. Second, the relatively small number of observations per manager introduces sampling error in the estimated fixed effects. We apply Empirical Bayes shrinkage to reduce noise in individual manager effects and, separately, correct for limited mobility bias in the variance decomposition by accounting for correlated estimation errors in manager and division effects.

Individual managers explain a substantial share of variation in police productivity. Under our preferred covariance-shrinkage correction (Best et al., 2023), manager effects account for approximately one quarter of the variance in log clearances per worker. This large contribution is consistent with the extensive operational discretion that commanders enjoy, which leaves considerable scope for individual differences to matter (Rasul and Rogger, 2018; Bandiera et al., 2021). We also find a sizable negative covariance between manager and division effects, indicating that more effective commanders tend to be assigned to divisions with lower underlying performance. This negative assortative matching, which mirrors findings in other public organizations (Fenizia, 2022; Limodio, 2021;

James, 2024), is consistent with the posting panel prioritizing divisions facing greater operational needs.

Differences in managerial effectiveness translate into sizable changes in performance: a one-standard-deviation increase in managerial effectiveness corresponds to approximately 12.6% higher clearances per worker. The effects of new managers materialize quickly: following managerial transitions, performance begins to rise (or fall) within three months of the arrival of a more (or less) effective commander, with no differential pre-trends. The effects are also short-lived. Conditional on the quality of the incoming commander, only about 40% of the outgoing manager's effect remains, on average, in the first six months after departure, with estimates approaching complete disappearance toward the end of the event window. This suggests that much of managerial value remains tied to the individual rather than becoming durable organizational capital.

We use our detailed incident-level data to unpack these productivity gains. In a stacked event-study design around managerial transitions, we find that, relative to transitions with no change in managerial effectiveness, a one-standard-deviation larger change in managerial effectiveness increases total clearances by 7%, while total workforce remains essentially unchanged. This suggests that effective managers make better use of existing resources, rather than relying on additional inputs. The increase in clearances is broad-based and reflects improvements along two margins: clearances increase for property and violent offenses, pointing to more effective investigations, and for drug offenses, which are largely detected through police-initiated activity, consistent with more proactive policing. These gains do not reflect changes in underlying crime, as we find no detectable effect on reported property and violent crime. Importantly, the improvements are not confined to outcomes under police discretion: court convictions rise alongside police-issued sanctions, suggesting that the additional clearances hold up in court rather than reflecting low-quality arrests. Finally, examining the geographic distribution of these gains within a division, we find that increases in property and violent clearances are larger in neighborhoods with higher pre-transition crime levels, while drug clearances increase evenly within the division.

Beyond routine operations, we examine whether commanders shape how divisions respond to centrally imposed policies. We study two policy episodes: the closure of police stations across London (Facchetti, 2025) and the staggered rollout of body-worn cameras. Following station closures, the increase in violent crime in divisions led by more effective commanders is only about a quarter as large as in divisions led by less effective commanders. During the rollout of body-worn cameras, we find that divisions led by less effective commanders experience a sharp increase in police response times, while this deterioration is offset in divisions led by more effective commanders. These findings together suggest that managerial effectiveness matters not only for day-to-day operations, but also for how central policies are implemented on the ground.

Finally, we ask whether managerial effectiveness is reflected in residents' perceptions of the police. This is particularly important in the context of policing, where public trust shapes crime reporting and cooperation with law enforcement. Using individual-level data from the London Public Attitude Survey, we find that transitions to more effective commanders improve residents' assessments of local police performance, increase trust in the police, and raise perceived safety after dark. These improvements occur despite no detectable changes in reported crime or self-reported victimization. Our findings show that managerial effectiveness is reflected not only in administrative measures of performance, but also in how citizens evaluate the public services they receive.

Taken together, these results help characterize how effective managers affect performance. Effective managers improve performance primarily by making better use of existing resources, rather than through additional inputs. They also help their divisions absorb change, cushioning adverse shocks and limiting the disruption from new technologies. However, they do not permanently transform their divisions: performance rises with their arrival and falls with their departure. Effective managers thus add value through ongoing operational decisions, and much of this value remains tied to the individual rather than becoming durable organizational capital.

This paper contributes to two main strands of literature.

First, we contribute to the growing literature on managers in the public sector. With the exception of Janke et al. (2019), who find little evidence that CEOs affect production in English public hospitals, this literature documents meaningful differences in organizational performance attributable to individual managers (Fenizia, 2022; Lavy et al., 2023; Muñoz and Prem, 2024; James, 2024; Muñoz and Otero, 2025).[1] Several of these studies point to changes in who works in the organization as a channel for managers' effects: more effective managers induce older workers to retire (Fenizia, 2022), retain high value added teachers (Muñoz and Prem, 2024), or replace school principals (Lavy et al., 2023), while a reform that improved the selection of hospital CEOs reduced doctor turnover (Muñoz and Otero, 2025). This channel operates even where formal personnel tools are limited (Fenizia, 2022), as they are for the commanders we study. Yet in our setting, commanders raise output with an essentially unchanged workforce. The richness of our data further allows us to trace this increase across distinct margins of police activity, showing that effective managers raise both proactive policing and investigative effectiveness. Like Fenizia (2022) and James (2024), we find no trade-off between quantity and quality, as also reflected in public perceptions of the police.

In studying how commanders raise productivity, we build on a large private sector literature investigating different potential channels through which managers might mat-

[1]More broadly, our paper relates to work on individual bureaucrats as a source of state capacity. This work estimates how much individual bureaucrats matter (Best et al., 2023; Baltrunaite et al., 2025) and studies how their allocation (Limodio, 2021), appointment (Xu, 2018), eligibility requirements (Baltrunaite et al., 2025), and career incentives (Bertrand et al., 2020) shape performance.

ter. As in the literature described above, several of these papers point to personnel management, with teaching, evaluating, retaining, and reallocating workers playing an important role (Lazear et al., 2015; Frederiksen et al., 2020; Hoffman and Tadelis, 2021; Cai and Wang, 2022; Friebel et al., 2022; Minni, 2026). Instead, our results point to managers' ongoing operational decisions as the primary channel, as in Adhvaryu et al. (2026), where good store managers improve pricing and inventory management. In contrast to the view of management as durable organizational capital embodied in practices (Bloom and Van Reenen, 2007), and to evidence that managers' gains persist through workers' careers (Minni, 2026), the productivity gains tied to better commanders largely disappear when they leave. This suggests that they reflect individuals' ongoing decisions rather than practices embedded in the organization (Bertrand and Schoar, 2003). Effective managers also help their divisions absorb central policy changes, paralleling private sector evidence that managers and management practices help organizations absorb shocks and adapt to change (Adhvaryu et al., 2022; Lamorgese et al., 2024; Brandimarti et al., 2026). In our setting, however, divisions operate under the same force-wide rules and common management practices, further underlining the key role of individual managers.

Second, 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) examine the effects of management practices and technology-based managerial tools on crime. Closer to our work, Smith (2025) studies how frontline sergeants, who directly supervise local teams of officers, shape officers' arrest behavior. We instead focus on middle managers who oversee entire territorial divisions and coordinate operations across multiple frontline teams. Our findings show that managerial effectiveness shapes not only enforcement activity, but also the ability of police organizations to quickly adapt to changes in their operating environment. Finally, a few papers recently have emphasized the role of police-citizen interactions in shaping attitudes toward law enforcement (Desmond et al., 2016; Blattman et al., 2021; Ang et al., 2025; Facchetti, 2025). We complement them by showing that managerial effectiveness is also reflected in residents' trust in the police and perceived safety, even in the absence of detectable changes in crime.

The remainder of the paper is as follows. Section 2 describes the institutional setting, and Section 3 introduces the data. Section 4 presents the empirical framework and estimates of managerial effectiveness. Section 5 examines how effective managers shape police performance and organizational responses to policy changes. Section 6 investigates the effects on residents' perceptions of policing. Section 7 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 organized into 32 divisions, the primary geographical units for operational management and resource deployment during most of our study period (Figure 1).[2] Territorial divisions operate within an organizational structure in which broad strategic priorities and budgets are determined centrally. At the same time, they are responsible for delivering core functions including neighborhood policing, incident response, and local crime-control activities, 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.

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



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

[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.

2.2 Borough Commanders

Each territorial division is led by a Borough Commander, typically holding the rank of Chief Superintendent. Borough Commanders occupy the middle tier of the police hierarchy, between the central leadership of the MPS and the frontline officers and teams responsible for delivering policing within individual neighborhoods. They are senior officers with substantial policing experience, typically having served for at least fifteen to twenty years before reaching this rank.

Borough commanders oversee 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. Despite this, commanders retain substantial discretion over how available resources are used. They assign 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 also make personnel decisions, such as 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. Officers typically require approximately 15–20 years of service to reach the rank of Borough Commander (Appendix Table A.1).

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.

These institutional features imply that Borough Commanders have limited discretion over the divisions to which they are assigned. Managerial mobility is driven primarily by promotion cycles, vacancies, and organizational needs, with postings determined through a centralized process rather than by individual preferences. Thus, while commanders exercise substantial discretion over day-to-day operations, they have limited control over where they serve. This distinction provides the institutional basis for our empirical strategy, which exploits managerial mobility across divisions to estimate individual contributions to police performance.

3 Data

We combine the career histories and monthly division assignments of Borough Commanders to division-level crime, clearances, and workforce constructed from administrative data. 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. We complement these records with individual-level survey data on residents' experiences and perceptions of policing. 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 commander, 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 reconstruct the prior positions, their subsequent appointments, 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 use 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]

We aggregate incidents and investigation outcomes to the division-by-month level, geo-coding each incident and assigning them to one of London's 32 original territorial divisions. During 2017–2018, most divisions were consolidated into 12 larger Basic Command Units (BCUs). We retain the original geographical boundaries throughout the analysis to construct a consistent division-level panel over time.[4]

Our main measure of police performance is *clearances*, which capture the police 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.

The detailed investigation outcomes allow us to distinguish between cases resolved through police–issued sanctions and those resulting in court convictions. Police-issued sanctions include community resolutions, formal cautions, drug-possession warnings, and penalty notices, which the police can administer without a court decision. These cover 6% of all investigation outcomes, and represent the set of investigative outcomes directly determined by police actions and discretion. Convictions instead capture cases resulting in a guilty verdict and a court-imposed sentence. They constitute 80% of all court sentences and include crimes sanctioned to imprisonment, fines, or other sentences by the court, excluding acquittals and discharges This distinction allows us to investigate whether differences in managerial effectiveness are associated with broader investigative outcomes rather than solely with discretionary police sanctions.

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,

[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 territorial divisions were progressively merged into 12 BCUs. Our geographical measures continue to refer to the pre-merger divisions.

[5]We follow the same definition given by the Home Office, that 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.

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 strength of four categories: front-line officers, 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; 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%.

Public Attitude Survey. We complement administrative police records with individual-level data from the London Public Attitude Survey (PAS), conducted on behalf of the Mayor's Office for Policing and Crime (MOPAC). This cross-sectional survey is fielded annually via face-to-face interviews to around 13,000 adults aged 16 and older, representative of Greater London, excluding the City of London. We use the survey to examine whether improvements in police performance are accompanied by changes in residents' experiences and perceptions of policing. The survey collects residents' assessments of local police performance, trust in the MPS, perceived safety, and experiences of crime and policing. We combine 16 quarterly survey waves from 2015Q2 to 2019Q1, covering approximately 51,000 respondents before restricting the sample to the manager connected set. We link respondents to police managers using their borough of residence and survey quarter, obtaining a working sample of approximately 46,000 respondents across the 29 divisions in the manager connected set. We apply survey weights, and standardize each outcome within the working sample to facilitate comparisons across outcomes. Appendix B3 provides further details on variable construction and sample restrictions.

3.1 Sample selection and descriptive evidence

Our main empirical analysis covers 2012–2019, the period for which incident-level police records are available. During this period, we observe 102 middle managers (Table 1), with an average of 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 held before their commander spell, 40% previously held a similar command role, 26% came from another territorial-policing role, and 35% from specialist, headquarters, or other non-territorial positions. Almost all of these prior positions (97%) were within London. After their final observed spell, 39% of managers move laterally to another policing position, 16% 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 on average around 550 police officers, 30 civilian staff, and 40 PCSOs. The final working sample contains around 2,800 division-month observations.

To motivate our analysis, Figure 2 documents variation in police performance across territorial divisions.Panel (a) plots the raw distribution of crime and clearances per capita, while Panel (b) residualizes both measures with respect to division and time fixed effects and workforce size. 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 these factors, substantial variation in clearance outcomes remains. For example, the standard deviation of residualized clearances per 1,000 residents is approximately one-fifth of the mean clearance rate. This unexplained variation in police effectiveness across territorial divisions motivates our investigation of whether individual commanders can account for part of these differences.

4 Measuring managerial effectiveness in policing

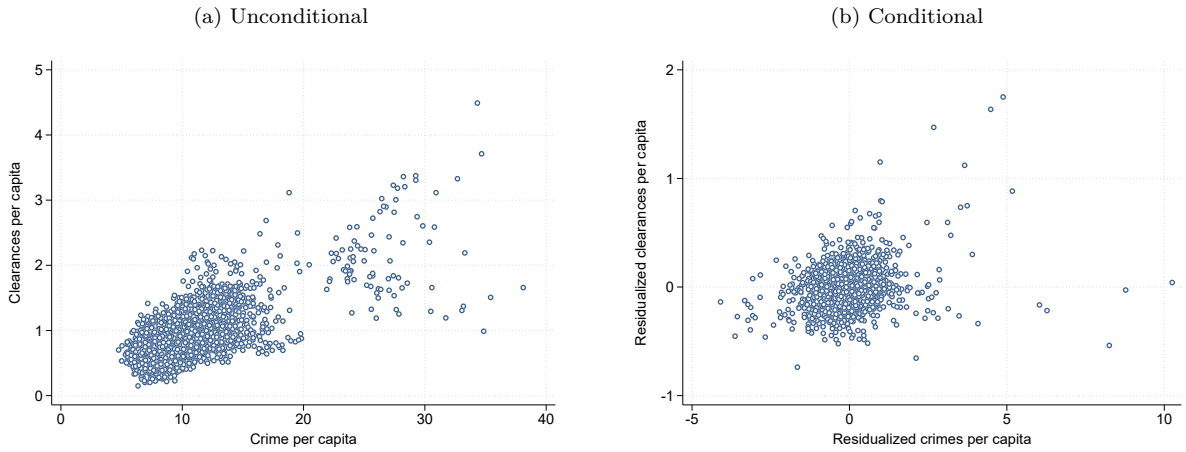
This section presents our empirical strategy for estimating the contribution of middle managers to police performance. We begin by introducing an AKM-style framework that separates persistent differences across police divisions from a manager-specific component of performance. We then discuss the estimation 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)	Observations (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	1.38	102
Average spell length (months)	23.84	102
Temporary/acting appointment	0.19	97
Position before first spell:		
- Previous command responsibility	0.40	81
- Territorial role, excluding command	0.26	81
- Specialist/HQ/non-territorial role	0.35	81
- Previous position in London	0.97	74
Position after final spell:		
- Promotion	0.16	99
- Lateral move	0.40	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	238.36	2784
- Convictions	92.07	2784
- Police-issued	116.94	2784
Offenses:		
- Total	2,620.15	2784
- Violent	545.53	2784
- Property	1,160.88	2784
- Drug	99.57	2784
- Other	151.60	2784
Workforce:		
- Total	621.76	2784
- Police officers	546.76	2784
- Staff	33.28	2784
- PCSO	41.72	2784
Clearances per worker	0.38	2784

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 102 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.

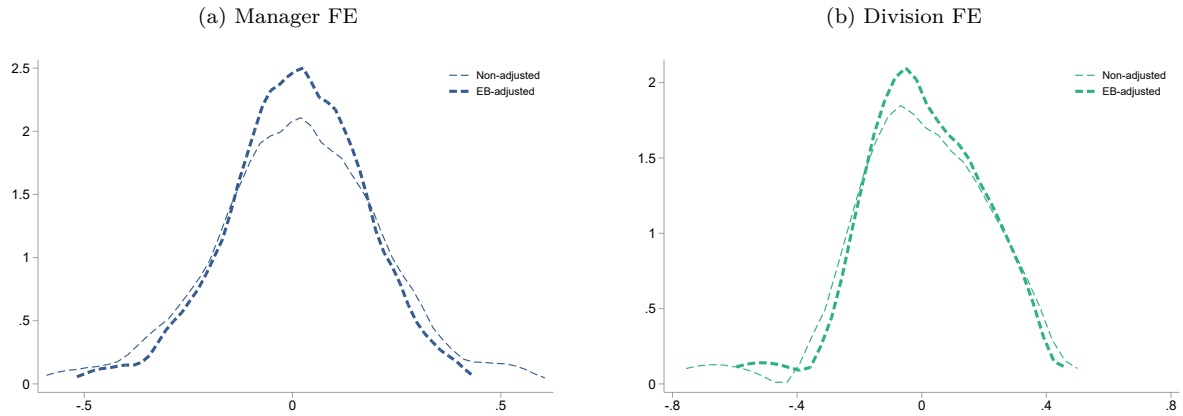
4.3 Identification

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 on manager–division-specific match effects, ($\mu_{d,m}$), that capture complementarities between particular managers and divisions. We examine each concern in turn.

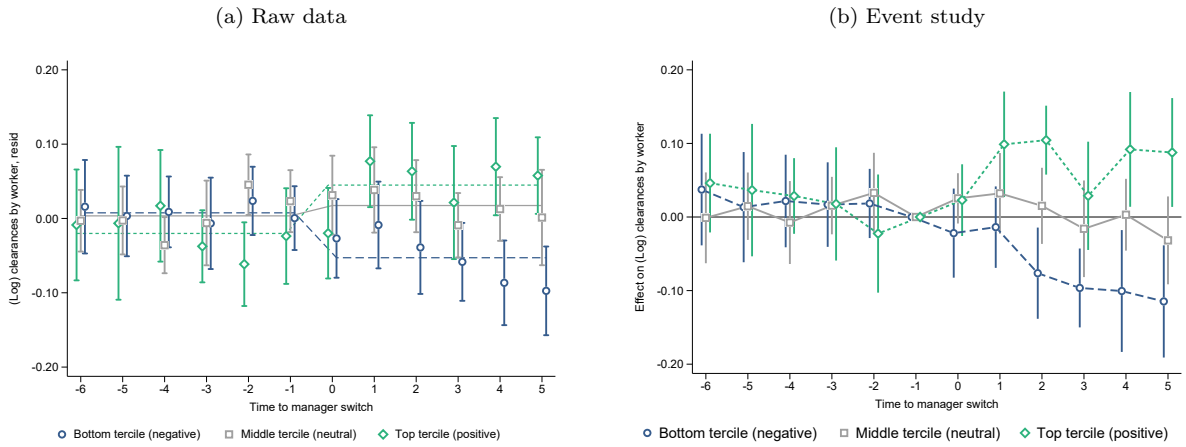
Sorting on time-varying shocks A potential threat to identification arises if manager mobility is systematically related to time-varying division-specific shocks. For instance, if managers on an upward performance trend are more likely to transfer into divisions that are already experiencing improvements 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 effectiveness 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

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.

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 69 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 12-month window around each 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_{st}^{k} 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–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

[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.

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 contribute to territorial police productivity?

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. We decompose the variance of Equation (1) as follows (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 this exercise is limited mobility bias (Andrews et al., 2008). As discussed in Section 4.2, manager and division effects are estimated jointly using a relatively small number of managers who move across divisions. Finite-sample bias can therefore overstate the true dispersion of both manager and division effects and mechanically bias their estimated covariance downward. While Empirical Bayes shrinkage reduces noise in each set of fixed effects separately, 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. We estimate this matrix by resampling the AKM residuals with replacement 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.For comparison, we also report the unadjusted AKM estimates, the finite-sample correction proposed by Andrews et al. (2008), and estimates based on individual Empirical Bayes shrinkage.

Table 3 reports the resulting variance decomposition. Under our preferred covariance-shrinkage procedure, the variance of the manager component is 0.019, corresponding to approximately 27% of the total variance in log clearances per worker, about 70% as large as the division-specific component. The estimated manager share is larger under the alternative procedures, ranging from 37% with individual Empirical Bayes shrinkage to around 50% under the raw AKM and Andrews estimates. Even under our most conservative correction, persistent differences across Borough Commanders account for a quantitatively important share of variation in police productivity.

We assess the robustness of the decomposition using two additional approaches, reported in Appendix Table A.6. First, we implement a split-sample 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 et al. (2008) corrected estimate. Second, following Adhvaryu et al. (2023), we conduct Monte Carlo simulations to quantify the finite-sample bias given 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 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 variance decomposition also reveals substantial negative assortative matching between managers and divisions. Under our preferred correction, the covariance between manager and division effects is -0.011, approximately 16% of the total variance in log clearances per worker, indicating that commanders with higher estimated effectiveness tend to serve in divisions with lower persistent productivity. Figure A.6 illustrates this pattern: high-performing managers are disproportionately observed in divisions in the bottom tercile of the division-effect distribution, while lower-performing managers are more frequently assigned to higher-performing divisions. This negative assortative match-

ing 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.[8] This pattern is consistent with an organizational policy that allocates stronger managerial capacity toward divisions facing greater operational needs.[9]

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%]	[61.2%]	
EB Shrinkage	0.069	0.025	0.034	-0.021
	[100.0]	[36.6%]	[48.5%]	
	[100.0]		[37.8%]	

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

The estimated manager effects imply economically meaningful differences in police productivity. A one-standard-deviation increase in managerial effectiveness corresponds to approximately 12.6% higher clearances per worker, while moving from a commander at the 10th percentile to one at the 90th percentile of the manager-effect distribution is associated with an increase of approximately 26%.[10] As shown in Section 4.3, these differences materialize rapidly following managerial transitions: divisions receiving a more effective commander experience improvements in productivity within three months of arrival, while transitions to less effective commanders are followed by declines in performance. We now examine whether these contributions persist after a commander leaves

[8] This negative assortative matching does not imply systematic changes in manager or division quality at each transition. Appendix Figures A.8 and A.9 show that the distribution of changes in managerial quality is approximately symmetric around zero and incoming commanders are, on average, similar in effectiveness to those they replace, while transitions between the highest- and lowest-performing managers are also relatively rare.

[9] While this feature is not surprising in the public sector (e.g. Limodio, 2021), similar evidence has been reported also in private sector contexts Metcalfe et al. (2023); Adhvaryu et al. (2024).

[10] Observable characteristics and prior career paths do not predict managerial effectiveness, which is also not significantly associated with subsequent promotion (Appendix Figure A.7 and Table A.5).

the division.

To investigate persistence, we trace how the association between division performance and the outgoing commander's estimated effectiveness evolves around managerial transitions. In doing so, we account for the incoming manager quality to avoid attributing to the outgoing manager differences in productivity associated with the manager who replaces them. We therefore estimate the relationship between performance and outgoing-manager effectiveness both with and without conditioning on the incoming manager's fixed effect.

Figure 5 presents the results. When incoming-manager effectiveness is omitted, the association between division performance and the outgoing manager's fixed effect changes little after departure, suggesting substantial persistence. However, once we condition on the effectiveness of the incoming commander the outgoing-manager contribution declines sharply. On average, only about 38% of the outgoing-manager effect remains during the first six months of the successor's tenure, with estimates approaching complete disappearance toward the end of the event window. These findings suggest that the productivity differences associated with individual commanders are relatively short-lived. Performance responds rapidly to managerial changes, but much of the contribution of an effective commander dissipates after departure. This pattern is consistent with managerial effectiveness depending importantly on the commander's ongoing presence, rather than being driven by durable organizational changes.

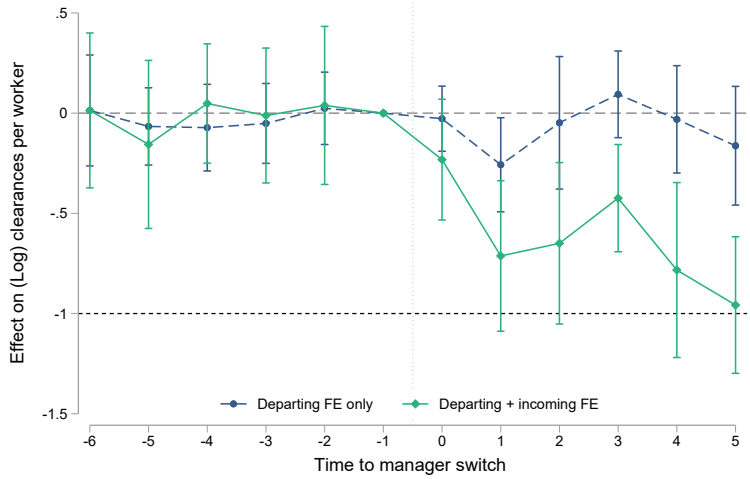
4.6 Robustness

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

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 event-study 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 spans 6 months before and after a manager switch, with the month immediately preceding the transition, $k = -1$, omitted. y_{dst} is log clearances per worker. The figure plots $\widehat{\alpha}_k^D$, which capture changes in departing manager's fixed effects relative to $k = -1$. The green line reports estimates controlling for the incoming manager's fixed effect, while the blue line reports estimates excluding this control. Standard errors are clustered at the police-division level; vertical bars report 95% confidence intervals.

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.

Finally, we explore the dynamics of manager transitions over longer time horizons. Appendix Figure A.5 reports estimates from the stacked specification using a continuous measure of the change in managerial effectiveness (see Section 5.1), using our baseline post-transition period of six months as well as extended periods of twelve and eighteen months. The effect of an incoming manager who is one standard deviation more effective builds up gradually over the first five months and then remains stable through the end of the first year, potentially growing further after that. Since effects over the first six months are nearly identical across the three samples, the longer-run patterns are unlikely to reflect selection into transitions with longer-staying managers.

5 How effective managers shape police performance

Having established substantial heterogeneity in managerial effectiveness among Borough Commanders, in this section we examine what changes within a police division when a more effective commander takes charge. We first investigate the sources of the productivity gains, distinguishing changes in investigative output and enforcement activity from changes in the resources available to divisions. We then examine whether these gains differ across neighborhoods with different policing needs. Finally, we study whether managerial effectiveness also shapes how divisions respond to changes in their operating environment.

5.1 Unpacking the productivity increase

We begin by examining what drives the productivity gains associated with more effective managers. In particular, we ask whether these gains reflect higher investigative output, changes in enforcement activity, or adjustments in the resources available to police divisions. We also investigate whether the improvements are concentrated in particular types of offenses or neighborhoods.

To do so, we use the stacked event-study design introduced in Section 4.3, and replace the tercile indicators with the continuous change in managerial effectiveness between the incoming and outgoing manager. We estimate the following specification:

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

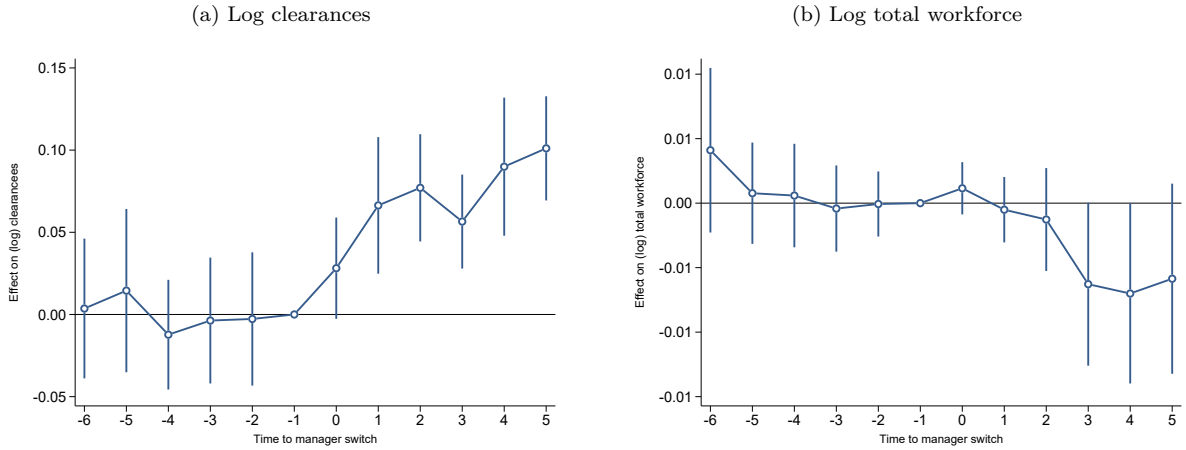
where $log(y_{sdt})$ is (log) outcome for division d measured in month t appearing in stack s, D_{st}^k denotes event time relative to the manager transition in stack s, $Treated_{sd}$ identifies the division experiencing the transition, and ϕ_{sd} and λ_{st} are division-by-stack and date-by-stack fixed 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 larger change in managerial effectiveness, corresponding to a change of approximately 0.11 log points in the manager fixed effect.[11]

We first distinguish changes in police output from changes in workforce size. Figure 6 shows that the effect is driven primarily by clearances, while the respective average

[11] 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. We also verify that the results are not driven by the 2017–2018 BCU reorganization. Appendix Table A.8 shows that estimates are robust to the inclusion of division fixed effects interacted with an indicator for the post-merger periods.

effect sizes are reported in Appendix Tables A.7 and A.9. 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.

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%. The distinction is important because property and violent offenses are largely reported by victims, whereas drug offenses are predominantly detected through police-initiated activity. These findings indicate that effective managers increase both the resolution of reported offenses and proactive enforcement.

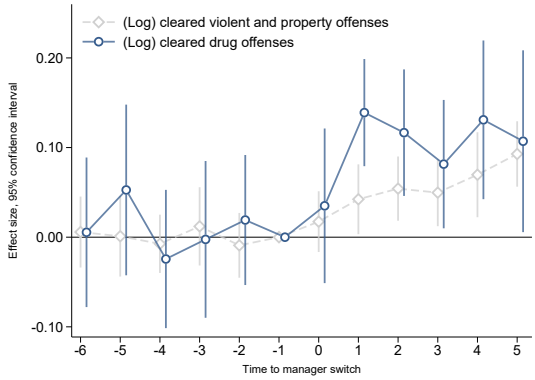
An important question is whether the additional clearances reflect improvements in investigative effectiveness or simply greater use of discretionary police sanctions through low-quality arrests. Figure 7 Panel (c) and Appendix Table A.7 columns (5) and (6) shows that this is not the case: both police-issued sanctions and court convictions increase following transitions to more effective managers, indicating improvements in investigation quality.[12]

[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-

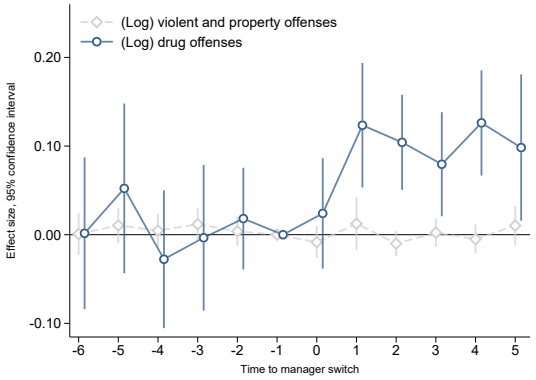
A natural question is whether these changes in police activity also translate into changes in crime through deterrence or incapacitation. Consistent with the earlier finding that manager fixed effects explain much less variation in crime than in clearances, Panel (b) and Appendix Table A.10 show no detectable effect of managerial effectiveness on reported property and violent offenses, while drug offenses increase significantly. Because 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

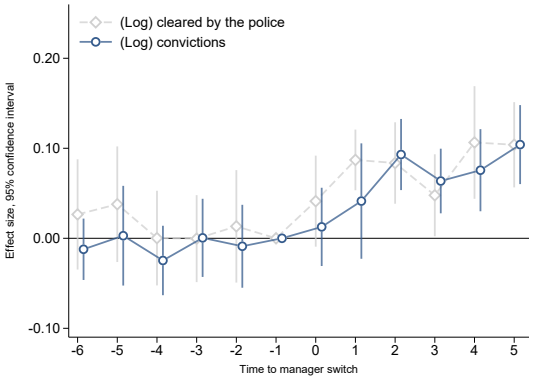
(a) Clearances by offense type



(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.

We next examine the resources available to police divisions. Borough Commanders operate within centrally determined budgets but retain discretion over how personnel are

tasking trade-offs: the offense-specific manager effects are positively correlated, with pairwise correlations of 0.26-0.49 (Appendix Figure A.10), indicating that effective managers tend to perform well across crime types rather than shifting effort between them.

organized and deployed. Appendix Table A.9 and Figure A.11 show that frontline officer numbers remain essentially unchanged following transitions to more effective commanders. The main change occurs among civilian back-office staff, whose numbers decline by approximately 3% on average over the six months following a transition.[13]

Finally, we investigate the geographical distribution of the productivity gains. Although we do not directly observe the deployment of officers across neighborhoods, 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. We apply our stacked differences-in-differences design to ward-level data, comparing wards with above- and below-median property and violent crime during 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 ($p = 0.057$). Drug clearances instead increase in both groups, and the 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. While the evidence does not directly identify changes in personnel deployment, it is consistent with managers directing reactive investigative effort toward hot-spots while extending proactive drug enforcement more broadly across the division.

5.2 Managerial responses to policy shocks

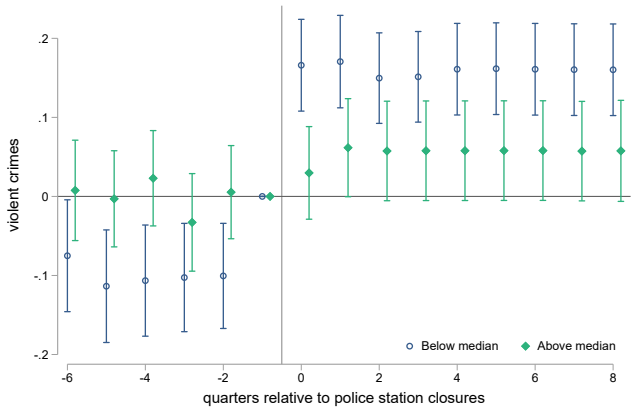
Our analysis so far shows that more effective commanders increase police productivity without expanding frontline resources. We now examine whether managerial effectiveness also shapes how divisions respond to changes in their operating environment. This margin matters because managers and bureaucrats 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 study two policy episodes within the MPS: the closure of police stations, which reduced local policing infrastructure, and the force-wide introduction of body-worn cameras, which changed the technology available to frontline officers. These episodes allow us to examine whether managerial effective-

[13]Estimates show little adjustment in Police Community Support Officers (PCSOs). We also find suggestive evidence of increased use of Special Constables, although the estimates are imprecise. Complementary evidence from annual division-level expenditure data (Appendix Section B4 shows no positive association between changes in managerial effectiveness and total expenditure. These expenditure results are descriptive and cover a shorter sample period.

ness matters not only for standard police operations, but also for how divisions adjust to changes in organizational conditions.

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 B5 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). Second, divisions led by below-median effectiveness commanders see a sharp increase in response time to incidents, of approximately 32%; in divisions led by above-median effectiveness commanders 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)

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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

show, the 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).

We further examine whether BWC adoption affects stop-and-search activity (Appendix Table B4). We find no significant differential effects by managerial effectiveness on either the volume of stops or their outcomes. 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 ("nothing found") rises by 3.1 percentage points, and the share resulting in a drugs warning falls by 2.1 percentage points. The interaction with managerial effectiveness 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.[14]

[14]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

6 Managerial effectiveness and residents' perceptions of policing

Finally, we ask whether the improvements associated with more effective managers are also reflected in residents' assessments of local policing. This question is particularly relevant given that we find no detectable effect of managerial effectiveness on reported property and violent crime. Residents may nevertheless value improvements in investigative performance and police services, even when their exposure to crime remains unchanged.

We use individual-level data from the London Public Attitude Survey (PAS) and apply the stacked differences-in-differences design introduced in Section 4. Because the survey is available only from 2015, the analysis covers 33 eligible managerial transitions. We examine residents' assessments of local police performance, trust in the MPS, perceived safety after dark, and self-reported victimization. All outcomes are standardized to facilitate comparisons across domains.

Appendix Table B2 reports the results, with all outcomes standardized to facilitate comparisons across domains. A one-standard-deviation improvement in managerial effectiveness is associated with a 0.077 standard-deviation increase in residents' assessments of local police performance, a 0.042 standard-deviation increase in trust in the MPS, and a 0.069 standard-deviation increase in perceived safety after dark. By contrast, we find no detectable change in self-reported victimization.

To explore how residents' perceptions may respond to managerial effectiveness, we examine heterogeneity in the effects on trust according to direct contact with the police and reported sources of information about policing. Figure B4 presents these estimates for the subset of PAS waves in 2017-18, alongside the share of respondents in each group. We distinguish residents reporting direct police contact from those without such contact and, within the latter group, examine different sources of information about local policing. We find positive effects on trust among residents without direct police contact, who account for 84% of respondents. Within this group, the effects are particularly pronounced among those receiving information from local policing teams and Safer Neighbourhood Boards, relative to those reporting media-related or local authority sources. These findings suggest that communication about local policing may be one channel through which managerial effectiveness shapes residents' trust, even in the absence of direct interactions with officers. However, because contact and information sources are measured at the time of the survey, these patterns should be interpreted as suggestive rather than causal evidence of this mechanism.

clearance effects; and small composition shifts in stop outcomes that operate at the officer-encounter level.

7 Conclusions

This paper studies the contribution of middle managers to the performance of public organizations, focusing on Borough Commanders in the London Metropolitan Police Service. Despite operating under centrally determined budgets and priorities, individual commanders account for a substantial share of the variation in police productivity. More effective managers increase investigative output and proactive enforcement without expanding frontline resources. These productivity gains do not stem solely from changing how the police handle offenses within their discretion. 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. More effective commanders also improve residents' assessments of the police, trust in law enforcement, and perceived safety. These outcomes are particularly important in policing, where trust and a strong relationship between citizens and the police are essential to the effective functioning of law enforcement.

Our findings also shed light on the role of middle managers in translating centrally determined resources and policies into effective day-to-day operations. Productivity gains emerge rapidly after more effective commanders take charge and largely disappear after they leave. This limited persistence suggests that much of managerial effectiveness is tied to the presence of individual managers, rather than translating into lasting improvements in the organization. At the same time, effective commanders are those who help their divisions quickly adapt and respond to external shocks in operating conditions.

More broadly, our results highlight the importance of middle management in public organizations where resources and policies are largely determined centrally. Improving public-service delivery may therefore require attention not only to resources and policy design, but also to the managers responsible for putting them into practice. An important question left open by our analysis is why some managers are more effective than others, despite operating under similar institutional constraints, receiving standardized training, and facing limited performance incentives. The relative roles of selection, individual skills, and motivation in explaining these differences remain poorly understood. Another promising direction is to study how middle managers interact with other levels of the organizational hierarchy, particularly how they communicate and coordinate with senior leadership and frontline personnel. Although these interactions are difficult to observe, understanding them may be essential to explaining how individual managerial effectiveness translates into broader organizational performance in a public 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.

_ , Parker Howell, Andrea Neyra-Nazarrett, Anant Nyshadham, and Jorge Tamayo, "What Do Good Managers Do and What Do They Leave Behind?," Working Paper 27-005, Harvard Business School 2026.

_ , Vittorio Bassi, Anant Nyshadham, and Jorge Tamayo, "No Line Left Behind: Assortative Matching Inside the Firm," *Review of Economics and Statistics*, 2024, pp. 1–45.

Aghion, Philippe and Jean Tirole, "Formal and real authority in organizations," *Journal of Political Economy*, 1997, *105* (1), 1–29.

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.

Ang, Desmond, Panka Bencsik, Jesse Bruhn, and Ellora Derenoncourt, "Community Engagement with Law Enforcement after High-Profile Acts of Police Violence," *American Economic Review: Insights*, 2025, *7* (1), 124–142.

Baltrunaite, Audinga, Enza Maltese, Tommaso Orlando, and Gabriele Rovigatti, "Procurement Managers and Effective Tendering: The Case of Italian Public Works Contracts," CEPR Discussion Paper DP19977, Centre for Economic Policy Research 2025.

Bandiera, Oriana, 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.

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

Blattman, Christopher, Donald P. Green, Daniel Ortega, and Santiago Tobón, "Place-Based Interventions at Scale: The Direct and Spillover Effects of Policing and City Services on Crime," *Journal of the European Economic Association*, 2021, *19* (4), 2022–2051.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

Brandimarti, Eleonora, Virginia Minni, Raffaella Sadun, and Jorge Tamayo, "Reorganizing Work Inside the Firm: Task Change after Technology Adoption," 2026. Working Paper.

Cai, Jing and Shing-Yi Wang, "Improving Management through Worker Evaluations: Evidence from Auto Manufacturing," *The Quarterly Journal of Economics*, 2022, *137* (4), 2459–2497.

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.

Desmond, Matthew, Andrew V. Papachristos, and David S. Kirk, "Police Violence and Citizen Crime Reporting in the Black Community," *American Sociological Review*, 2016, *81* (5), 857–876.

Dixit, Avinash, "Incentives and organizations in the public sector: An interpretative review," *Journal of human resources*, 2002, pp. 696–727.

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.

Frederiksen, Anders, Lisa B. Kahn, and Fabian Lange, "Supervisors and Performance Management Systems," *Journal of Political Economy*, 2020, *128* (6), 2123–2187.

Friebel, Guido, Matthias Heinz, and Nikolay Zubanov, "Middle Managers, Personnel Turnover, and Performance: A Long-Term Field Experiment in a Retail Chain," *Management Science*, 2022, *68* (1), 211–229.

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.*

Hoffman, Mitchell and Steven Tadelis, "People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis," *Journal of Political Economy*, 2021, *129* (1), 243–285.

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

James, Erik, "The Effect of Managers on Public Service Provision: Evidence from Medicaid, SNAP, and TANF," 2024. Working Paper.

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.

Lamorgese, Andrea, Megha Patnaik, Andrea Linarello, and Fabiano Schivardi, "Management Practices and Resilience to Shocks: Evidence from COVID-19," *Management Science*, 2024, *70* (12), 9058–9072.

Lavy, Victor, Genia Rachkovski, and Adi Boiko, "Effects and Mechanisms of CEO Quality in Public Education," *The Economic Journal*, 2023, *133* (655), 2738–2774.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton, "The Value of Bosses," *Journal of Labor Economics*, 2015, *33* (4), 823–861.

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.

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.

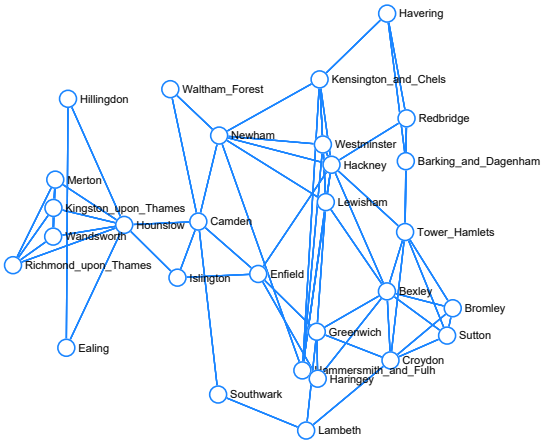
Thaler, Richard, "An econometric analysis of property crime: Interaction between police and criminals," *Journal of Public Economics*, 1977, *8* (1), 37–51.

Xu, Guo, "The costs of patronage: Evidence from the British empire," *American Economic Review*, 2018, *108* (11), 3170–3198.

A Appendix Figures and Tables

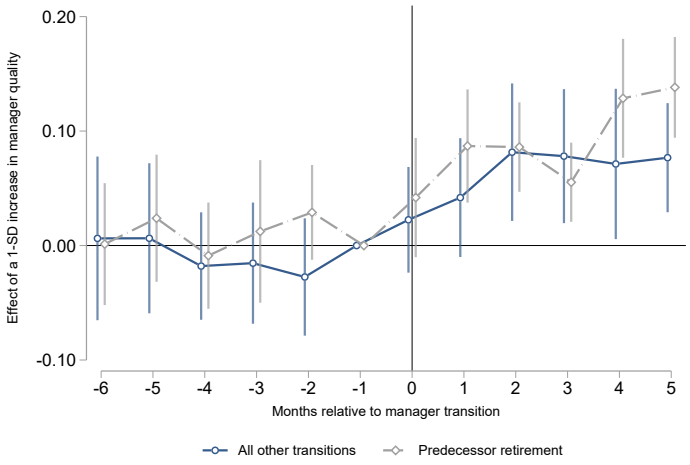
A1 Appendix Figures

Figure A.1: Connected set in MPS



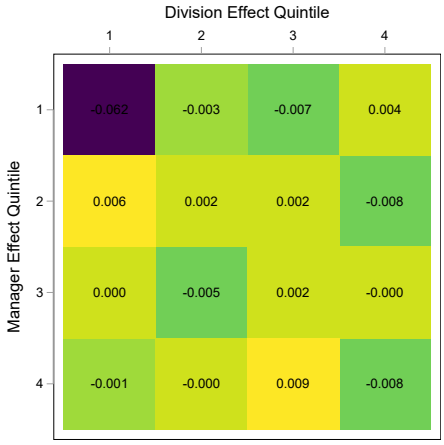
Notes: Each circle represents a police division, and a link between them means the divisions have shared a manager.

Figure A.2: Manager transitions by predecessor retirement



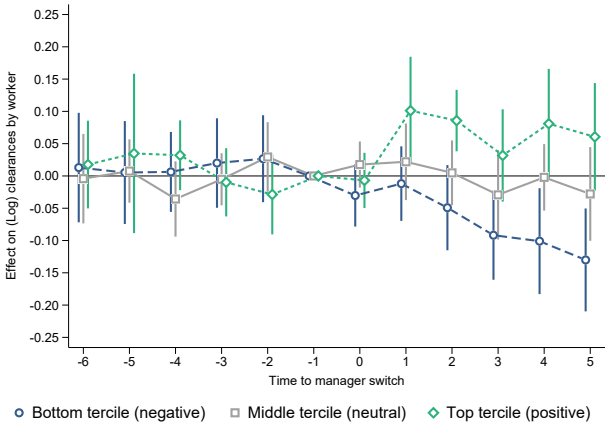
Notes: The figure reports estimates from the stacked event-study specification, allowing the effect of a change in managerial effectiveness to differ according to whether the outgoing manager retires. The figure plots the event-time effect of a one-standard-deviation improvement in manager effectiveness 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



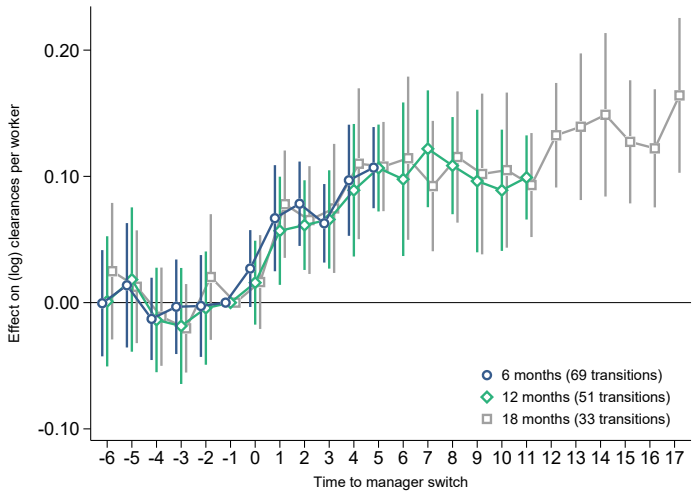
Notes: The figure shows mean residuals of estimated manager and division fixed effects from Equation (1), across quintiles of fixed effects. The figure shows no systematic patterns in the size of residuals across quintiles.

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



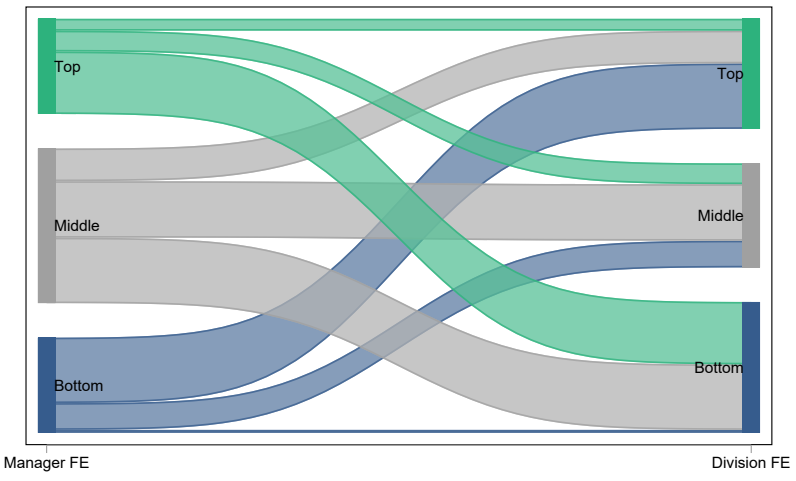
Notes: This figure reports estimates from a TWFE event-study specification around manager transitions. Transitions are classified into terciles based on the change in empirical-Bayes-adjusted manager fixed effects between the incoming and outgoing manager. Bottom-tercile transitions correspond to the 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: Manager transitions for different time horizons



Notes: The figure shows how the effect of manager effectiveness on (log) clearances per worker evolves over longer post-transition horizons. The figure plots the event-time effect of a one-standard-deviation improvement in manager effectiveness, estimated separately on three stacked datasets. Each dataset includes transitions in which the outgoing manager was in post for the full six months before the transition and the incoming manager remained for at least six, twelve, or eighteen months after it, respectively. 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.6: 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.7: Correlates of manager fixed effects

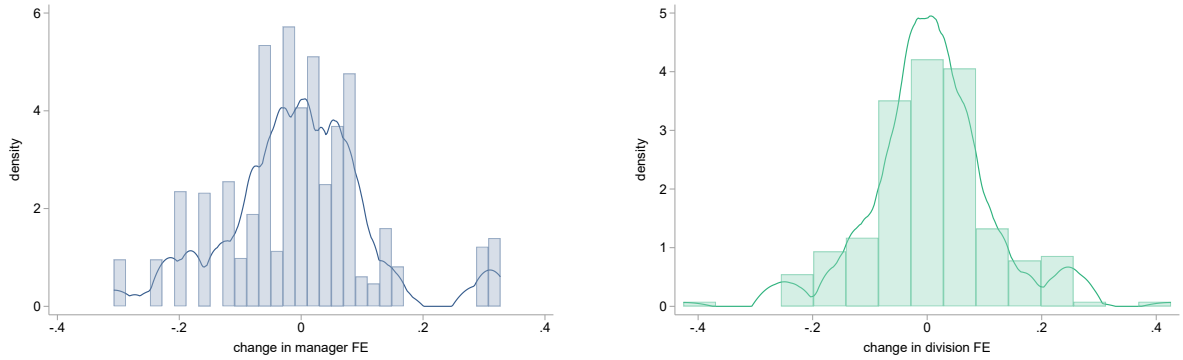


Notes: The figure reports estimates from a multivariate regression of manager fixed effects on manager characteristics and prior-position indicators, included jointly in the same specification. Pre-position categories refer to the role held before the manager's first observed Borough Commander spell.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

Figure A.8: Distribution of changes in manager and division fixed effects

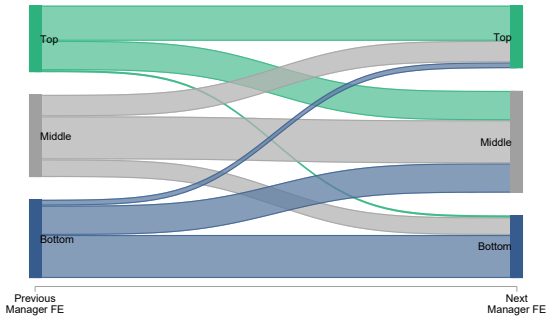
(a) Manager FE

(b) Division FE



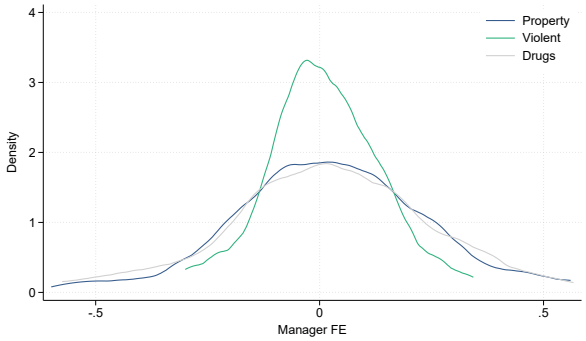
Notes: This figure shows the distribution of changes in manager and division fixed effects corresponding to each transition involving a different manager for the division. The left panel shows the distribution of changes in manager FEs, defined as $\Delta ManEff_d$ in Equation 2. The right panel shows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

Figure A.9: 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.10: Multi-tasking across crime types

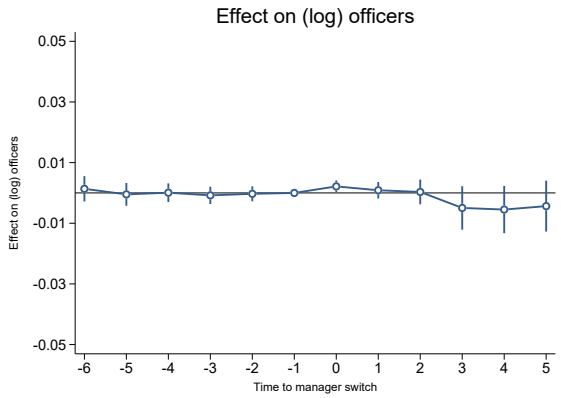


	All	Property	Violent	Drugs
All	1			
Property	0.779	1		
Violent	0.586	0.386	1	
Drugs	0.843	0.486	0.261	1

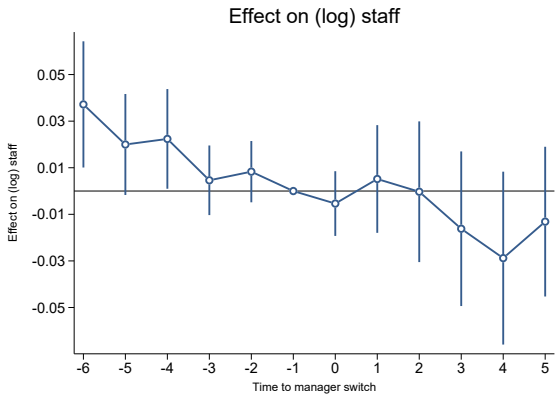
Notes: On the left side we report the distributions of Empirical Bayes-adjusted manager fixed effects estimated separately for (log) clearances per worker in violent, property, and drug-related offenses. On the right side, we report the pairwise correlations between these offense-specific manager effects.

Figure A.11: Effects of manager effectiveness on police workforce allocation

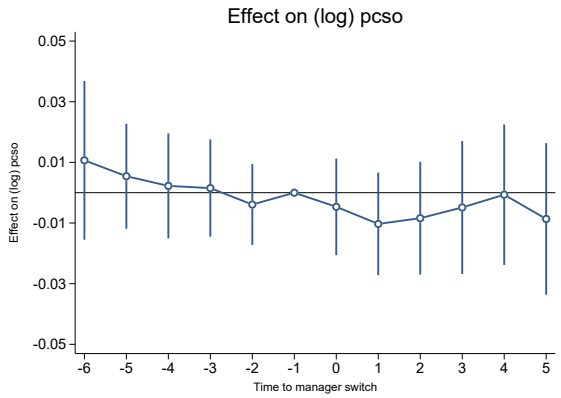
(a) Front-line officers



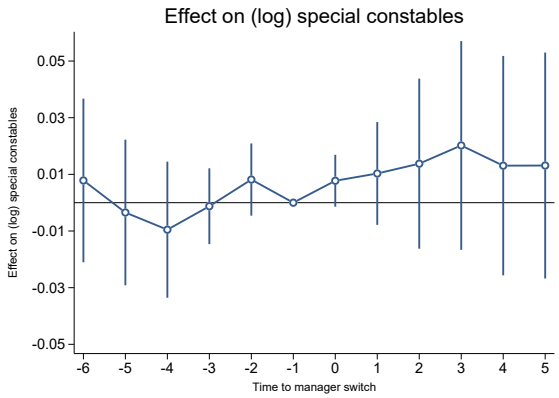
(b) Back-office staff



(c) PCSOs



(d) Special constables



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)
Manager FE (std)	-0.021	0.003	-0.049	-0.027
	(0.051)	(0.029)	(0.046)	(0.048)
Female	0.124	0.052	-0.132	-0.007
	(0.115)	(0.074)	(0.082)	(0.112)
White	0.022	0.075	0.050	-0.032
	(0.136)	(0.083)	(0.103)	(0.135)
Joined pre-1990	-0.203**	-0.147**	0.301***	-0.117
	(0.098)	(0.070)	(0.076)	(0.098)
Tenure on the job (months)	-0.004	-0.000	0.005*	-0.004
	(0.003)	(0.002)	(0.003)	(0.003)
Observations	126	126	126	126
N. managers	90	90	90	90
Baseline mean	0.548	0.119	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. Explanatory variables include manager characteristics, notably the standardized manager fixed effect, and average crime and clearance outcomes during the command spell. All explanatory variables enter jointly in each specification. Robust standard errors are reported in parentheses. The final rows report the number of observations (manager-spells), the number of managers, 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%]	[61.2%]	
EB Shrinkage	0.069	0.025	0.034	-0.021
	[100.0]	[36.6%]	[48.5%]	
Covariance Shrinkage	0.069	0.019	0.026	-0.011
	[100.0]	[27.1%]	[37.8%]	
Split-sample	0.069	0.033	0.041	-0.026
	[100.0]	[48.3%]	[59.4%]	
Monte Carlo corrected	0.069	0.035	0.043	-0.028
	[100.0]	[50.5%]	[62.7%]	

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	-0.016	-0.021*	-0.014	-0.036	-0.015	-0.015
	(0.012)	(0.012)	(0.012)	(0.026)	(0.015)	(0.015)
Post × Treated × Change	0.074***	0.070***	0.054***	0.093***	0.065***	0.072***
	(0.008)	(0.008)	(0.009)	(0.022)	(0.011)	(0.010)
Observations	13680	13680	13680	13680	13679	13680
Divisions (Clusters)	29	29	29	29	29	29
Stacks	69	69	69	69	69	69
Mean Dep. Variable	-0.969	5.438	4.900	4.196	4.710	4.472
P-value Post × Treated	0.158	0.082	0.236	0.166	0.326	0.314
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. Standard errors are clustered at the police division level. $p < 0.10$, $p < 0.05$, $p < 0.01$.

Table A.8: 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	-0.012	-0.018	-0.024	-0.020	-0.017	-0.003
	(0.015)	(0.015)	(0.015)	(0.030)	(0.019)	(0.016)
Post × Treated × Change	0.070***	0.068***	0.050***	0.085***	0.068***	0.056***
	(0.012)	(0.012)	(0.013)	(0.024)	(0.014)	(0.015)
Observations	13673	13673	13673	13673	13672	13673
Divisions (Clusters)	29	29	29	29	29	29
Stacks	69	69	69	69	69	69
Mean Dep. Variable	-0.968	5.438	4.901	4.196	4.710	4.472
P-value Post × Treated	0.420	0.252	0.096	0.512	0.364	0.824
P-value Post × Treated × Change	0.002	0.000	0.014	0.012	0.000	0.002

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. $p < 0.10$, $p < 0.05$, $p < 0.01$.

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

	(Log) workforce				
	All	Officers	PCSOs	Staff	Special Constables
	(1)	(2)	(3)	(4)	(5)
Post × Treated	-0.004	-0.004	0.001	0.006	-0.011
	(0.003)	(0.003)	(0.011)	(0.019)	(0.019)
Post × Treated × Change	-0.004*	-0.002	-0.009	-0.025**	0.013
	(0.002)	(0.003)	(0.010)	(0.010)	(0.015)
Observations	13680	13680	13680	13680	13680
Divisions (Clusters)	29	29	29	29	29
Stacks	69	69	69	69	69
Mean Dep. Variable	6.407	6.285	3.577	3.019	4.535
P-value Post × Treated	0.188	0.240	0.942	0.722	0.522
P-value Post × Treated × Change	0.072	0.522	0.448	0.040	0.500

Notes: This table reports the estimated effects from the stacked differences-in-differences specification described in Equation (5), using log workforce by officer function as the outcome. Standard errors are clustered at the police division level. $p < 0.10$, $p < 0.05$, $p < 0.01$.

Table A.10: Average effect of managerial transitions on log crimes

	(Log) offenses		
	All	Property and Violent	Drugs
	(1)	(2)	(3)
Post × Treated	0.005	0.006	-0.031
	(0.007)	(0.006)	(0.026)
Post × Treated × Change	-0.001	-0.005	0.086***
	(0.006)	(0.006)	(0.016)
Observations	13680	13680	13680
Divisions (Clusters)	29	29	29
Stacks	69	69	69
Mean Dep. Variable	7.810	7.384	4.462
P-value Post × Treated	0.438	0.368	0.226
P-value Post × Treated × Change	0.930	0.406	0.000

Notes: This table reports the estimated effects from the stacked differences-in-differences specification described in Equation (5), using log crime by offense type as the outcome. Standard errors are clustered at the police division level. $p < 0.10$, $p < 0.05$, $p < 0.01$.

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	-0.023	-0.025*	-0.001	-0.025	-0.019	0.000	-0.006	-0.017	-0.017	-0.018
	(0.017)	(0.014)	(0.016)	(0.016)	(0.023)	(0.028)	(0.021)	(0.015)	(0.015)	(0.016)
Post × Treated × Change	0.061***	0.071***	0.034***	0.065***	0.083***	0.072**	0.058***	0.061***	0.039**	0.076***
	(0.012)	(0.012)	(0.011)	(0.014)	(0.016)	(0.028)	(0.012)	(0.016)	(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	148416	142572	148416	142572	148416	142572	148416	142572	148416	142572
Divisions (Clusters)	29	29	29	29	29	29	29	29	29	29
Stacks	69	69	69	69	69	69	69	69	69	69
Mean Dep. Variable	2.242	3.247	1.728	2.679	1.119	1.927	1.545	2.462	1.308	2.243
P-value Post × Treated	0.154	0.092	0.910	0.116	0.450	0.982	0.806	0.268	0.230	0.284
P-value Post × Treated × Change	0.002	0.000	0.024	0.002	0.000	0.020	0.000	0.006	0.034	0.000
P-value Difference		0.420		0.057		0.730		0.844		0.035

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. P-value rows come from a wild-cluster bootstrap procedure. The "p-value difference" row is the p-value of the difference between effects in $< p50$ split and effects in $\geq p50$ split. 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. $p < 0.10$, $p < 0.05$, $p < 0.01$.

B Data Appendix

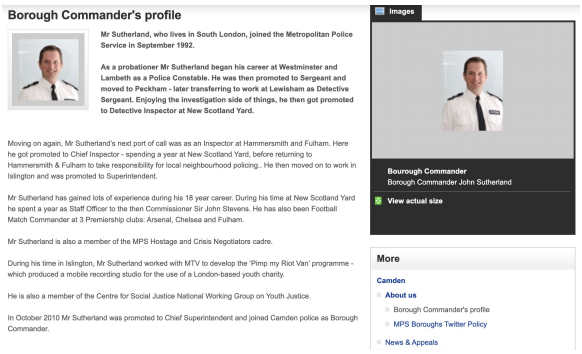
B1 Borough Commanders

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 at the division-manager-date level.

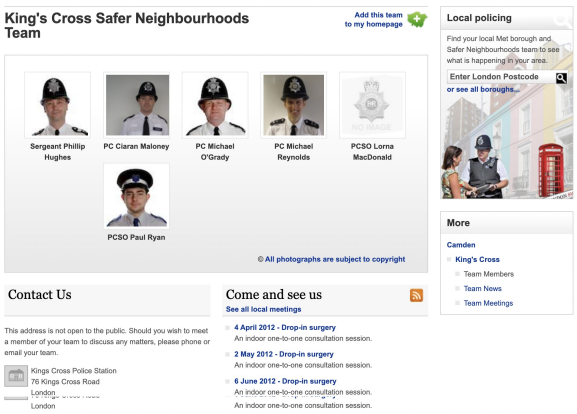
Figure B1 shows an example of the information contained on an archived website for the borough of Camden in 2012.

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.

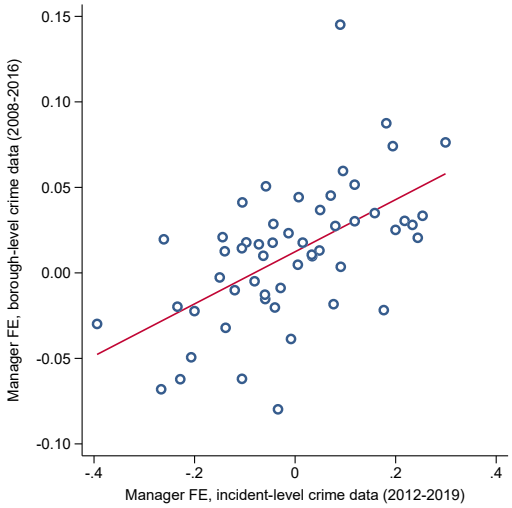
We estimate Equation (1) on the borough-by-month panel, using log SDs per worker as the outcome, with the same connected-set restrictions and empirical Bayes adjustment as in the main analysis. The mover graph in this sample gives 11 connected sets, including a larger set covering 19 divisions, a smaller 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 Public Attitude Survey

Data construction. We use individual-level Public Attitude Survey (PAS) files provided by the UK Data Service, and conducted on behalf of the Mayor's Office for Policing and Crime, combining 16 quarterly waves from 2015Q2 to 2019Q1. Interviews are conducted face-to-face with residents aged 16 and older living in private households in Greater London, excluding the City of London. Residential addresses are sampled within boroughs, and one eligible adult is randomly selected within each household. Continuous fieldwork targets 100 interviews per borough per quarter, yielding approximately 12,800 interviews annually. The pooled data contain 51,295 respondents across the 32 London boroughs.

We match respondents to police divisions using their borough of residence. The shorter time coverage of the PAS restricts the number of managerial transitions that can be used in the analysis, leaving only 33 managerial transitions, each of which defines a separate stack. Each stack contains respondents in the division experiencing the managerial transition and respondents in eligible control divisions. Because the same respondent may serve as a control in multiple transition-specific stacks, respondents can appear more than once in the stacked estimation sample.

The PAS files report respondents' survey quarter and financial year, but do not provide the interview month or exact interview date. We therefore assign treatment status at the

quarterly level. In particular, the quarter in which a managerial transition occurs is classified as post-transition. Among the 33 transitions for which treated respondents are observed in the transition quarter, 16 occur in the second or third month of the quarter, so that the transition quarter is only partially treated in these cases. This timing mismatch introduces measurement error in treatment assignment and is likely to attenuate the estimated effects in the first quarter post-transition.

We use the supplied quarterly individual weights, which account for unequal selection probabilities and differential non-responses. Respondents with zero survey weights do not contribute to weighted estimates.

Variable definition. We construct the following outcomes using the questions below.

1. *Police doing a good job.* "Taking everything into account, how good a job do you think the police in this area are doing?" We code an indicator equal to one for "Excellent" or "Good" and zero for "Fair," "Poor," or "Very poor."

2. *Trust in police.* Respondents report agreement with the statement: "The Metropolitan Police Service is an organization that I can trust." We assign values from 1 to 5 to "Strongly disagree," "Tend to disagree," "Neither agree nor disagree," "Tend to agree," and "Strongly agree," respectively.

3. *Safety after dark.* "How safe do you feel walking alone in this area after dark?" We assign values from 1 to 4 to "Very unsafe," "Fairly unsafe," "Fairly safe," and "Very safe," respectively.

4. *Victimization.* "Have you been a victim of crime or Anti Social Behaviour in the last 12 months?" We code an indicator equal to one for "Yes" and zero for "No." This measure includes antisocial behavior and is not restricted to incidents occurring in the respondent's local area.

We treat "Don't know," refusals, unasked questions, and other non-substantive responses as missing. We standardize every outcome, including binary indicators, using mean and standard deviation computed in the working sample. Standardization takes place before constructing the stacked estimation sample, so respondents' inclusion in multiple stacks does not affect the normalization.

Regressions include categorical controls for age, sex, and ethnicity using the harmonized survey variables. Age groups are 16–24, 25–34, 35–44, 45–54, 55–64, and 65 or older. The recorded sex categories are male, female, and other; ethnicity categories are White British, White Other, Mixed, Asian, Black, and Other. For each control, missing observations are assigned a separate category. We omit education because it is unavailable in the first six quarters of our sample.

Descriptive statistics. Appendix Table B1 reports descriptive statistics for the PAS working sample before construction of the stacked estimation sample. We retain the 29 divisions belonging to the manager connected set. This excludes Barnet, Brent, and Harrow and leaves a working sample of around 46,000 respondents. Around 66% of respondents rate the local police as doing a good or excellent job. Average trust in the MPS is 4.15 on the five-point scale, and average perceived safety after dark is 3.06 on the four-point scale. Approximately 11% of respondents report having experienced crime or antisocial behavior during the previous twelve months.

Table B1: Descriptive table using Public Attitudes Survey (PAS)

	Mean	N
Outcomes		
Police doing a good job	0.660	45,007
Trust in police	4.147	46,346
Reported police presence	3.077	45,623
Insufficient patrols	0.589	45,255
Safety after dark	3.057	42,291
Victimization	0.107	46,453
Demographics		
Age: 16-24	0.132	46,474
Age: 25-34	0.245	46,474
Age: 35-44	0.201	46,474
Age: 45-54	0.163	46,474
Age: 55-64	0.113	46,474
Age: 65 or over	0.143	46,474
Sex: Male	0.490	46,474
Sex: Female	0.508	46,474
Ethnicity: White British	0.445	46,474
Ethnicity: White Other	0.186	46,474
Ethnicity: Mixed	0.031	46,474
Ethnicity: Asian	0.169	46,474
Ethnicity: Black	0.122	46,474
Ethnicity: Other	0.033	46,474

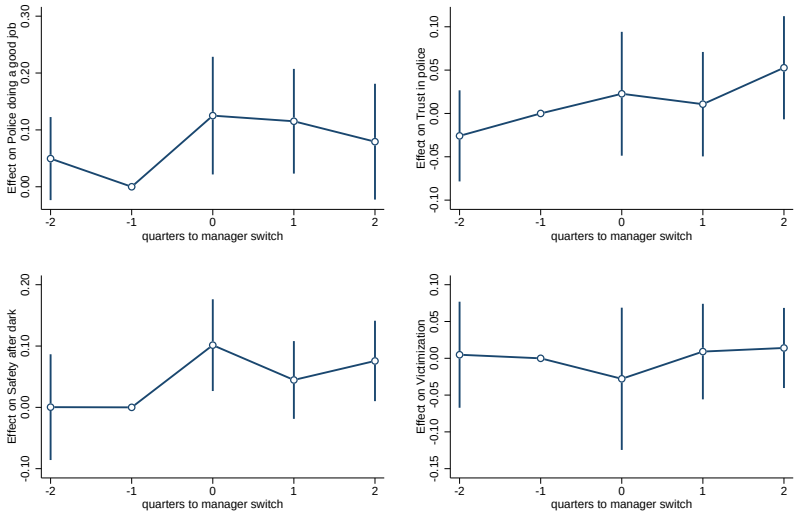
Notes: This table reports survey-weighted means and unweighted respondent counts for the PAS working sample, covering 2015Q2–2019Q1 across the 29 divisions in the manager connected set. Missing demographic categories are omitted. Counts exclude respondents with missing responses.

Table B2: Effects of managerial transitions using Public Attitudes Survey (PAS)

	Police doing a good job	Trust in police	Safety after dark	Victimization
	(1)	(2)	(3)	(4)
Post × Treated	0.080**	0.082**	0.051*	0.036
	(0.036)	(0.032)	(0.030)	(0.023)
Post × Treated × Change	0.077**	0.041**	0.069***	0.004
	(0.033)	(0.016)	(0.024)	(0.018)
Division-by-Stack FEs	✓	✓	✓	✓
Quarter-by-Stack FEs	✓	✓	✓	✓
Observations	252484	259965	237980	260452
Mean Dep. Variable	0.008	-0.003	0.014	-0.011
P-value Post × Treated	0.028	0.020	0.100	0.122
P-value Post × Treated × Change	0.042	0.034	0.026	0.822

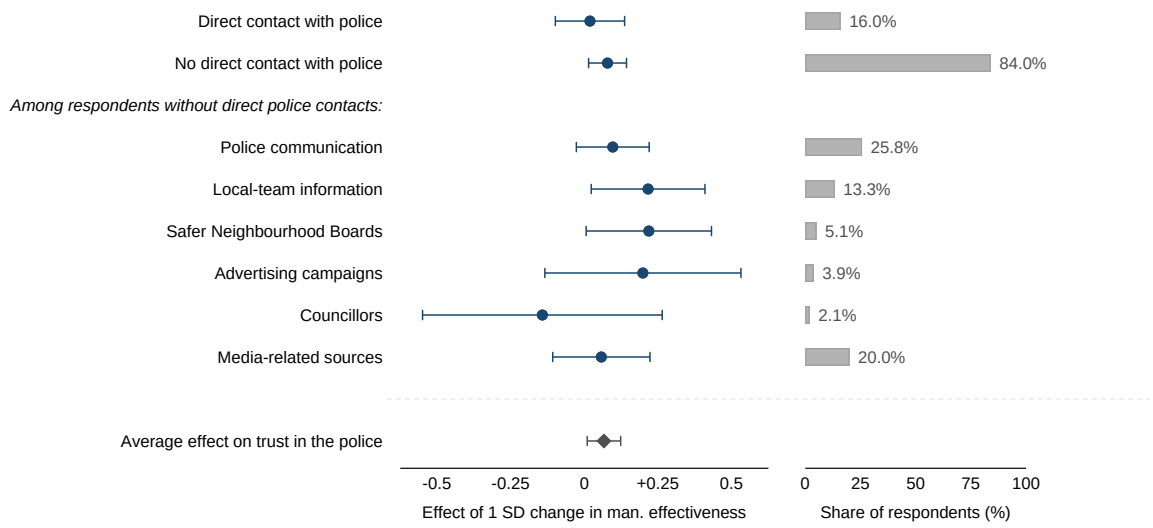
Notes: This table reports the estimated effects from the stacked differences-in-differences specification described in Equation (5), using respondent-level data from the Public Attitude Survey available from 2015 to 2019. Outcomes in Columns (1)-(5) are: an indicator for rating local police performance as good or excellent; trust in the MPS; perceived safety when walking alone locally after dark; and an indicator for having experienced crime or antisocial behavior in the previous 12 months. All outcomes are standardized in the working survey sample before stacking. Regressions use PAS survey design weights and include division-by-stack and quarter-by-stack fixed effects, together with controls for age-band, sex, and ethnicity groups. We report wild bootstrap p-values. Standard errors are clustered at the police division level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B3: Managerial effectiveness and residents' perceptions of policing



Notes: The figure reports event-time coefficients on the interaction between treatment status and the standardized change in managerial effectiveness from the stacked differences-in-differences specification in Equation (5). Outcomes are standardized within the working survey sample. Estimates use respondent-level PAS data and survey weights, with division-by-stack and quarter-by-stack fixed effects and individual controls for age, sex, and ethnicity. Coefficients are normalized relative to the quarter immediately preceding the transition ($k = -1$). Vertical bars indicate 95% confidence intervals based on standard errors clustered at the police-division level.

Figure B4: Managerial effectiveness and trust in police: contact with police and information sources



Notes: The left panel reports the effect of a one-standard-deviation increase in managerial effectiveness on standardized trust in the police, plotting the Post × Treated × Change coefficient in Equation (5). Estimates use survey weights and control for age, sex, and ethnicity; bars indicate 95% confidence intervals with standard errors clustered at the police division level. Direct contact include respondent- or police-initiated contact, station visits for information, or information obtained from police officers. Information-source estimates are restricted to respondents without recorded direct contact. Police communication includes emails/texts and leaflets/newsletters; local-team information refers to information from Safer Neighbourhood Teams or Dedicated Ward Officers. Media sources include newspapers, radio, television, websites, social media, and magazines. The diamond marker reports the average effect across respondents in the common sample period, 2017Q2–2018Q4. The right panel shows weighted shares of distinct respondents: contact shares use all respondents in this period, and source shares use those without recorded direct contact. Sources overlap, so their shares need not sum to 100%.

B4 Borough-Level Police Expenditures Data

In this section, we test whether managerial transitions led to changes in the quantity and composition of local financial resources spent by local police divisions. We use a balanced borough-year level panel of expenditures by categories from 2015/16 to 2019/20 financial years. The data measure total expenditures for each borough split into the following categories: Pay and Overtime for Police Officers, PCSOs, Staff, other Employee related expenditures (e.g. training), premise costs, transport, Supplies & Services, Transfers to/from MPS overall Reserves, and the net total expenditures.

Table B3 presents descriptive statistics about the composition of borough local spending in different categories, split roughly into personnel (officer, PCSO, staff pay and overtime), resource costs (transport, premise), other income (credit items), and transfers to/from the central reserve of the whole MPS. Importantly, officer, staff, and PCSO personnel payments account for an average of over 99% of total borough-level expenditures, consistent with our argument that managers have limited discretion in influencing performance through significant changes in spending on non-personnel items (e.g, technologies).

Table B3: Borough spending by category, 2015-2019

	Borough-year spending (£)					Share of net	
Spending category	P25	Mean	P50	P75	P90	spending (%)	*N*
Officer pay + overtime	26,360,942	48,605,556	37,158,770	72,131,792	83,470,932	95.2	120
Staff pay + overtime	451,327	896,062	577,334	1,425,158	1,795,191	1.8	120
PCSO pay + overtime	1,011,026	1,698,373	1,266,438	2,598,052	3,118,989	3.3	120
Employee-related expenditure	0	4,823	0	650	20,501	0.009	120
Premises costs	0	316	0	0	925	0.0	120
Transport costs	144,489	247,370	193,040	373,418	429,562	0.5	120
Supplies & services	143,198	294,027	217,550	429,241	546,007	0.6	120
Other income (credit)	-974,802	-679,575	-410,889	-142,000	-33,237	-1.3	120
Transfers to/from reserves	-2,408	-5,239	0	0	0	-0.0	120
Net total expenditure	27,815,377	51,061,714	37,904,279	76,998,213	88,465,702	100.0	120

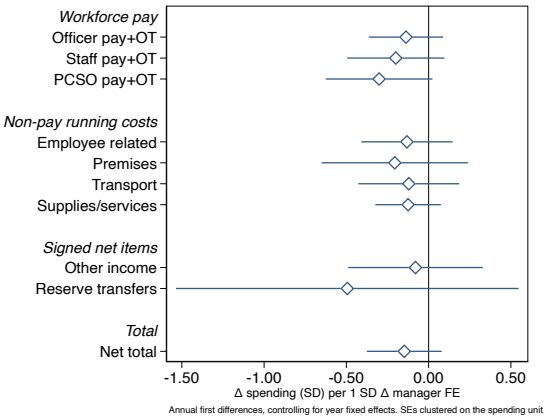
Notes: Borough-by-fiscal-year observations, 2015/16–2019/20. Amounts are in Great British Pounds. The Met's BCU budget statements are presented on a net expenditure basis, which is why two lines are routinely negative. *Other income* is a credit (fees, charges, partnership contributions and recharges to other bodies) netted off gross expenditure, so it is negative by construction. *Transfers to/from earmarked reserves* are signed: positive is a contribution into reserves, negative a draw-down financing in-year spending. For each category, "share of net spending" column is the ratio of the category mean to the mean of net total expenditure over the same sample. Pay and overtime are summed within each workforce group. From 2018 the Met reports at Basic Command Unit rather than borough level; post-2017 values are BCU totals attributed to each constituent borough.

Next, we test for potential impacts of managerial transitions on local spending. Given five annual observations, we are unable to perform event study analysis like our main specifications.

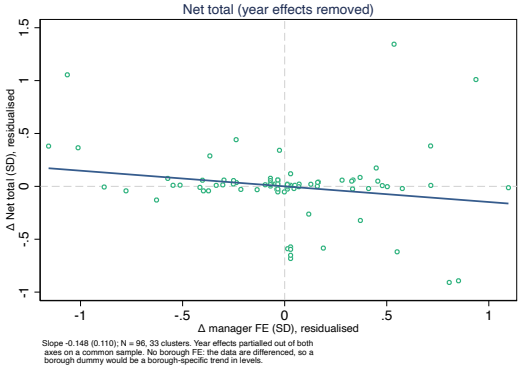
We instead use the First-Differences specification below to test whether changes in managerial fixed effects could influence changes in borough expenditures. The left-hand side outcomes are standardised expenditures (mean 0 standard deviation 1) and right-hand side regressors are standardised manager fixed effects. We control for year fixed effects and cluster standard errors at the borough level.

$$\Delta Spending_{dt} = \mu_t + \mu_1 \Delta \text{Manager FE}_{dt} + \epsilon_{dt} \tag{B1}$$

Figure B5: First-Differences, changes in manager fixed effects on expenditures



(a) First-Difference Coefficient Plots



(b) Δ Net Total Expend. on Δ Manager FE

The left panel of Figure B5 presents estimates of the coefficients for impacts of changes in manager fixed effects on changes in expenditures in different categories, and the right panel shows the (residualised) scatterplots showing the underlying relationship between changes in standardised net total expenditures and changes in standardised manager fixed effects. We consistently find negative and mostly statistically significant relationship between changes in manager fixed effects and expenditures. In particular, the negative coefficient for the transfer category suggests that better managers do not receive more resources (in the form of transfers from central budget), suggesting that changes in resource availability cannot explain our findings on managerial ability.

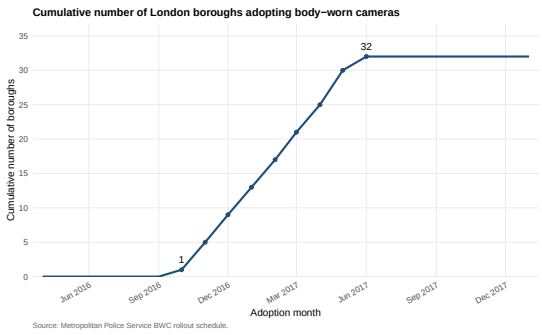
Taken together, the evidence presented strongly suggests that better managers *do not* use more resources (proxied by higher expenditures), if anything, there is a negative and sometimes statistically significant association between changes in manager fixed effects and expenditures. This suggests that better managers, if anything, use fewer resources, consistent with the negative effects of managerial transitions on police workforce documented in our main draft.

B5 Body-worn cameras

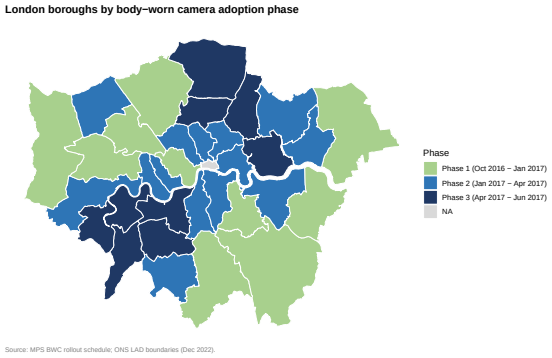
This appendix provides additional details on the rollout of body-worn cameras across MPS and on the stacked differences-in-differences design used in Section 5.2.

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 B6. 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]

Figure B6: 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:

$$Y_{ist} = \beta\,(\text{Post}_{ist} \times \text{Treated}_{is}) + \beta^{M}\,(\text{Post}_{ist} \times \text{Treated}_{is} \times \mathbf{1}_{i}^{+,M}) + \alpha_{is} + \lambda_{st} + \varepsilon_{ist}, \quad \text{(B2)}$$

[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/.

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×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 B4 reports estimates of Equation (B2) on stop-and-search outcomes.

Table B4: 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.