Detective Quality and Investigative Output

Matthew B. Ross* CarlyWill Sloan† Austin V. Smith‡

October 28, 2026

Abstract

We examine how detective quality affects investigative outcomes using over 160,000 case observations from Austin, Texas. Candidates for promotion to detective are ranked by a total score combining a written examination covering criminal law, departmental procedures, and interrogation with points for seniority and education. Our difference-in-differences design exploits vacancy timing under a system that promotes officers in rank order and nearly always assigns units the highest-ranked available candidate. A one-standard-deviation increase in promotion score raises arrest probability by 2 percentage points, or 13 percent relative to the sample mean. Exceptional clearances decline, particularly those in which the victim refuses to cooperate, while overall clearance changes little. The arrest effects extend to cases led by detectives who were already in the unit before the promotion. These findings show that the qualifications rewarded by an existing promotion system predict investigative performance, with contributions extending beyond incoming detectives' own caseloads. The results contribute to research on worker selection and public-sector productivity by providing causal evidence on how qualifications measured before promotion translate into organizational output.

*Department of Public Policy, 310 Renaissance Park, Northeastern University, Boston, MA 02115. ma.ross@northeastern.edu.

†Department of Social Sciences, 607 Cullum Road, United States Military Academy at West Point, New York 10996. carlywill.sloan@westpoint.edu. The views expressed are my own and do not necessarily reflect the official policy or position of the United States Military Academy, the Department of the Army, the Department of Defense, or the U.S. Government.

‡Department of Public Policy, 310 Renaissance Park, Northeastern University, Boston, MA 02115. au.smith@northeastern.edu.

This research was supported by a generous grant from Arnold Ventures. The funder had no role in the analysis or interpretation of the findings and did not review or approve the manuscript.

1 Introduction

Hiring decisions shape organizational performance through the capabilities workers bring to their new roles and their effects on colleagues. Observable qualifications and prior performance provide imperfect signals of a worker's potential contributions to an organization. Choosing the right candidate may be especially important when workers enter roles with little formal preparation or directly relevant experience (Lazear, 2009; Benson et al., 2019). In settings where employees exercise substantial discretion over decisions that are difficult to standardize or monitor, unobserved qualities such as knowledge, judgment, and motivation can be especially consequential (Prendergast, 2007). The choice of whom to hire can also affect colleagues' productivity in team settings (Arcidiacono et al., 2017; Sandvik et al., 2020). These considerations make selection consequential across many sectors, particularly in settings characterized by limited formal preparation, substantial discretion, or team production. Yet despite the prevalence of these features across many settings, there is limited causal evidence on how hiring and promotion criteria shape individual performance and team productivity.

Criminal investigations offer a consequential public-sector setting for studying this question. Detectives develop evidence, interview victims and witnesses, identify suspects, and prepare cases for potential prosecution. They retain responsibility for individual cases while working within units where supervisors allocate caseloads and colleagues can provide formal and informal assistance. Detectives typically bring prior experience in patrol work, which requires a different mix of skills and may provide limited preparation for investigative responsibilities. Yet newly appointed detectives are not consistently required to receive formal investigative training (Kingshott et al., 2015; Lum et al., 2024; Snider, 2025). Their work requires legal and procedural knowledge as well as substantial discretion over which leads to pursue, how to obtain information, and where to direct investigative effort. These decisions help determine which reported crimes result in arrests and prosecution and how investigative services are distributed across victims and communities (Goldstein, 2022).

In this paper, we examine whether the qualifications used to select detectives predict their contributions to investigative output. We link promotion records and civil-service eligibility lists to administrative records from the Austin Police Department covering 2010–2020, yielding over 160,000 case observations surrounding 54 promotion events across 25 investigative units. We measure incoming detective quality using the total promotion score, which is determined before promotion and combines performance on a written examination covering criminal law, departmental procedures, and interrogation with additional points for seniority and education. Promotion exam performance may reflect a combination of factors,

including preexisting subject-matter expertise, intrinsic motivation, effort devoted to preparation, and underlying cognitive ability. To help characterize the attributes captured by the department's promotion scores, we also examine officers' earlier statewide licensing exam scores and measures of prior patrol performance, which are positively and negatively correlated with promotion scores, respectively. Using promotion scores as our primary measure of incoming detective quality, our main analysis estimates how receiving a higher-scoring detective affects investigative output both for the unit as a whole and for cases assigned to incumbent detectives.

To identify the causal effect of incoming detective quality on unit performance, we develop an empirical strategy that exploits a unique feature of Austin's promotion system that constrains which detective an investigative unit receives when filling a vacancy. In Austin, promotion examinations establish ranked eligibility lists that remain valid for two years (City of Austin and Austin Police Association, 2013). Specifically, officers are promoted to detective in order of their promotion score as vacancies arise, and units nearly always receive the highest-ranked available candidate. Administrative records indicate that vacancies arise primarily for exogenous reasons, such as an incumbent detective's retirement or promotion to a senior leadership position, typically outside the criminal investigations bureau.

We assess potential strategic timing by examining whether receiving units' case composition, caseloads, and staffing vary systematically with promotion timing relative to each list cycle. We find little evidence of systematic differences, consistent with limited sorting on these observed characteristics. We therefore use vacancy timing relative to these cycles as a source of plausibly exogenous variation in incoming detective quality: earlier vacancies generally bring higher-scoring detectives, while later vacancies bring candidates farther down the list (Figure 1). We then implement a continuous difference-in-differences event study comparing changes in unit-level arrests and clearances around the arrival of detectives with different promotion scores. Identification requires that units receiving higher- and lower-scoring detectives would have followed similar outcome trends absent promotion and experienced similar gains had they received detectives of the same quality. We find little evidence of differential pre-promotion trends or sorting on observed unit characteristics. The arrest estimates remain similar after matching promotion events on pre-promotion characteristics and are unchanged when we account for common calendar-time shocks.

We find that receiving higher-scoring detectives increases the likelihood that investigations result in arrest. Because incoming detectives take on cases gradually, we average effects over months 2–9 after promotion. Over this period, a one-point increase in incoming detectives' average total promotion score raises arrest probability by 0.42 percentage points, or 2.4 percent relative to the sample mean of 17.45 percent. A 5.50-point increase, equivalent

to one standard deviation of total promotion scores among promoted officers, corresponds to a 13.2 percent increase in arrest probability relative to the mean. The effects emerge gradually in the first three months following a promotion and persist through month nine. This increase in arrests is accompanied by a decline in exceptional clearances: cases meeting clearance requirements but prevented from proceeding to arrest or prosecution by specified circumstances. A one-standard-deviation increase in total promotion score reduces exceptional clearance probability by 2.26 percentage points, or 33.2 percent relative to its sample mean of 6.8 percent. Both the arrest increase and exceptional-clearance decline remain statistically significant under wild-cluster-bootstrap inference. The estimated change in overall clearance is small and statistically insignificant, suggesting that the principal effect of receiving higher-scoring detectives is on how investigations are resolved.

Investigations can be cleared by arrest, cleared exceptionally, or remain uncleared; examining the reasons for exceptional clearance helps interpret the patterns in our main results. A case may be exceptionally cleared when circumstances beyond law enforcement's control– such as the offender's death, denial of extradition, or a victim's refusal to cooperate with prosecution– prevent an arrest despite sufficient evidence and knowledge of the offender's identity (Federal Bureau of Investigation, 2020). We find that a one-point increase in promotion score reduces the probability of exceptional clearance explicitly attributed to victim refusal to cooperate by 0.34 percentage points, while the estimated decline in clearances attributed to the district attorney's decision not to prosecute is smaller and statistically insignificant. We also find evidence of an increase in the total number of suspects identified, concentrated among those who are ultimately arrested. Together, these patterns suggest that higher-scoring detectives are more effective at securing victim cooperation and developing suspects that ultimately lead to an arrest.

Finally, we examine how incoming detective quality affects colleagues' investigations and whether its effects differ across racial and ethnic groups. Among cases assigned to incumbent detectives in the same unit, a one-point increase in incoming detectives' average total promotion score raises arrest probability by 0.47 percentage points, similar to the unit-wide estimate. Advice, assistance, and shared investigative knowledge could contribute to these gains, as could changes in workloads and case allocation that our administrative data cannot fully capture. We also use demographic information on suspects, arrestees, and victims to distinguish who is identified as a suspect or arrested from whose victimization cases result in arrest. Differences across racial and ethnic groups could reflect variation in detectives' cultural competencies and ability to engage with victims. The increase in arrests is concentrated among Hispanic and White arrestees, and we reject equality of the effects across arrestee groups. Estimated arrest effects for cases involving Black, Hispanic, and White victims are

all positive, although their precision varies. We cannot reject equality across victim groups, so these estimates do not establish changes in racial disparities in investigative services.

Our findings contribute to the economics of crime and policing by providing causal evidence on a substantial but understudied component of police production: criminal investigations. A large empirical literature examines police staffing and patrol deployment, as well as how officers' characteristics, training, and peers relate to enforcement behavior.[1] Much less is known about how detectives affect the resolution of reported crimes, despite the substantial resources devoted to investigations. Cook and Ludwig (2018) identify detective effectiveness and witness cooperation as priorities for research on criminal investigations. In 2020, local police departments serving at least 100,000 residents employed roughly four patrol officers for every detective or investigator (Goodison, 2022). Existing research shows that information technology can improve investigative productivity and that investigative priorities can generate unequal service provision (Mastrobuoni, 2020; Goldstein, 2022). We add causal evidence on how the qualifications used to select detectives affect investigative output. Linking case outcomes to arrestee and victim demographics further allows us to distinguish who is arrested from whose reported victimization results in arrest, connecting personnel decisions to both the distribution of enforcement and the provision of investigative services.

The paper also contributes to a broader literature on worker selection and organizational production. Studies of hiring tests and promotion document the value and limitations of the signals organizations use to select workers (Hoffman et al., 2018; Benson et al., 2019), while research on public-service recruitment shows that changing who enters can affect service delivery (Ashraf et al., 2020). We evaluate the qualifications rewarded by an existing promotion system in a setting where workers retain responsibility for individual assignments while contributing to a shared production process. Examining effects on both unit output and incumbent detectives' cases connects this selection question to research on productivity spillovers and workplace knowledge exchange (Arcidiacono et al., 2017; Sandvik et al., 2020). Our findings show that qualifications measured before promotion predict contributions beyond the incoming worker's own assigned tasks. Evaluating selection criteria against individual performance alone can therefore miss an important component of their organizational consequences.

Finally, our findings inform a longstanding debate in criminology and policing policy over how much detectives can influence investigative outcomes beyond the circumstances

[1] For police staffing and deployment, see (Evans and Owens, 2007; Draca et al., 2011; Chalfin and McCrary, 2018; Mello, 2019; Weisburd, 2021; Chalfin et al., 2022). For police hiring and officer characteristics, training, and peers, see (McCrary, 2007; Owens and Ba, 2021; Ba et al., 2021; Hoekstra and Sloan, 2022; Rivera, 2025; Dube et al., 2025; Ba et al., 2025; Rivera, 2026; Adger et al., 2026).

of the reported crime. A canonical series of RAND studies in the 1970s emphasized the importance of initial information and patrol officers' responses, finding little evidence that traditional follow-up investigation improved clearance rates (Chaiken, 1975; Greenwood and Petersilia, 1975; Greenwood et al., 1975; Chaiken et al., 1977). These findings helped establish persistent skepticism about detectives' contributions to case resolution. Subsequent research documents relationships between clearance rates and investigative resources, organizational practices, and detective experience (Braga et al., 2019; Lum et al., 2024; Goodison, 2021), although causal evaluations of investigative activities remain uncommon (Prince et al., 2021). We add causal evidence that receiving higher-scoring detectives increases arrests and reduces exceptional clearances, with effects extending to cases assigned to incumbent colleagues. These findings identify detective selection as a consequential policy margin: the qualifications of an incoming detective can affect investigative outcomes throughout the receiving unit. They also show that overall clearance rates can obscure substantive differences in how detectives can meaningfully contribute to case resolution.

2 Institutional Background

2.1 Detective Work and the Organization of Investigations

Detectives conduct follow-up investigations to determine whether a reported crime can be resolved through an arrest and potential prosecution. Their responsibilities include reviewing initial reports, contacting victims and witnesses, interviewing suspects, collecting and evaluating evidence, and preparing cases for prosecution. APD's General Orders describe these tasks as extending from the development of investigative leads to assistance with prosecution (Austin Police Department, 2025, secs. 403.2–403.3).[2] These responsibilities require detectives to combine knowledge of criminal law and departmental procedures with judgment about which leads to pursue and how to obtain information from the people involved. APD organizes investigations into specialized and regional units. Our sample includes units focused on child abuse, robbery, burglary, auto theft, human trafficking, sex crimes, domestic violence, aggravated assault, narcotics, and financial crimes, among others, as well as regional detective units (Table A2). Cases are routed to units and assigned to individual detectives by supervisors (Austin Police Department, 2025, sec. 403.3). This structure makes the receiving unit a relevant level of analysis: incoming detectives contribute to a shared caseload, and supervisors can adjust assignments across detectives.

[2]The 2025 General Orders postdate our study period, but we have no reason to believe that these investigative responsibilities and procedures differed substantially during the years we study.

APD assigned cases based on workload, case type and complexity, and perceived solvability (Police Executive Research Forum, 2012, pp. 36–37). Individual detectives' outcomes may therefore reflect case allocation as well as investigative performance, motivating our analysis of unit-level outcomes. Incoming detectives may affect unit outcomes through their own investigations, assistance to colleagues, or changes in case allocation. Their share of unit cases rises gradually after promotion (Figure A1 (b)), motivating our focus on average effects beginning in month two. We also examine incumbent detectives' cases to assess whether improvements extend beyond cases handled by incoming detectives. Such improvements may arise because incoming detectives help colleagues investigate cases or because supervisors reallocate cases, changing the difficulty of incumbents' caseloads. Controlling for observed case characteristics helps assess whether changes in case composition explain these improvements.

Whether working on their own cases or assisting colleagues, detectives can influence case outcomes by identifying suspects, securing victim and witness cooperation, and gathering evidence sufficient for arrest and prosecution. We examine arrests and exceptional clearances separately because they represent distinct investigative outcomes: arrests result in enforcement action, whereas exceptional clearances occur when investigators establish grounds for arrest but cannot proceed for reasons beyond their control. Better detectives may overcome barriers that would otherwise lead to exceptional clearance, increasing arrests without changing total clearance. Combining these outcomes could therefore obscure improvements in investigative performance. We also examine suspect identification and specific reasons for exceptional clearance to understand how detective quality affects investigative progress and the barriers to arrest.

2.2 Detective Promotion, Scores, and Unit Assignment

Entry into APD's corporal/detective rank occurs through an internal promotion process in which existing police officers take a civil-service examination and are ranked on an eligibility list. Candidates must have completed four continuous years of service as an APD police officer, and examination questions are drawn from a published list of source materials (City of Austin Firefighters', Police Officers', and Emergency Medical Services Personnel's Civil Service Commission, 2019).[3] Because the examination serves the combined corporal/detective rank, we restrict our analysis to promotions into investigative units rather than patrol-focused

[3]The examination notice and General Orders cited in this section concern the January 2020 examination and departmental policies issued in 2025, respectively. These documents postdate the promotion lists used in our analysis. We have no reason to believe that the institutional features described here changed materially over this interval, although we cannot verify every provision against contemporaneous documents.

corporal assignments.

The examination covers departmental policies and procedures (35 percent), criminal law and procedure (25 percent), interrogation techniques (20 percent), leadership (10 percent), and the agreement between the City and the police association (10 percent) (City of Austin Firefighters', Police Officers', and Emergency Medical Services Personnel's Civil Service Commission, 2019). These materials address arrests, search warrants, follow-up investigations, and case clearance, connecting examination performance to knowledge officers may use when developing cases for prosecution. Candidates are ranked by their total promotion score, which adds seniority and education points to the written examination score. The 2011–2017 eligibility lists report all three components separately: examination scores range from 70 to 99 among listed candidates, seniority points range from approximately 2 to 15, and education points take values of 0, 0.5, 1, or 2. The total score therefore combines tested knowledge with credit for experience and educational attainment, with seniority contributing substantially more additional points than education. We use total score as our preferred measure because it determines promotion-list positions while preserving score differences that ranks discard.

Beyond these formal components, pre-promotion characteristics help describe the officers the system ranks more highly. Figure A1, panel c, relates total scores to officer characteristics and prior patrol behavior. Higher-scoring officers have more experience, more education points, and higher scores on the Texas Commission on Law Enforcement (TCOLE) licensing examination. The first two relationships partly reflect the points explicitly awarded for these characteristics; the third suggests that promotion scores also capture skills or preparation associated with success on an earlier professional examination. Higher scores are negatively associated with use of force, dispatched calls, and self-initiated incidents during the preceding six months, while associations with index, non-index, and public-order arrests are small and have confidence intervals that include zero. These descriptive relationships may reflect differences in assignments, experience, or policing behavior. Although the score does not directly measure investigative skills such as building rapport with victims or evaluating evidence, it provides a measure determined before promotion whose relationship with subsequent investigative output we can assess. Separate analyses examine how examination performance, seniority points, and education points each relate to investigative outcomes.

While scores determine candidates' ordering, vacancies determine when positions become available. Although we do not observe why specific positions became vacant in our analysis sample, later administrative records attribute approximately 44% of reported openings to retirement, 51% to promotion of the incumbent—the officer previously holding the position—and 5% to resignations, deaths, terminations, military leave, or police-association

assignments. APD's General Orders explicitly link promotion timing to these events (Austin Police Department, 2025, sec. 918.4.1). Candidates are promoted in rank order, and units nearly always receive the highest-ranked available promotee. Figure 1 illustrates how this process links vacancy timing to incoming detective quality: units filling vacancies earlier in an eligibility-list cycle generally receive higher-scoring candidates, while later vacancies are filled by candidates farther down the list.

This ordering constrains the available candidates, but promotion into rank and assignment to a particular unit remain distinct decisions. APD's General Orders allow candidates approaching promotion to apply for posted vacancies and provide for placement in an open assignment if a candidate has not been selected for a position (Austin Police Department, 2025, sec. 918.4.2).

Our empirical design uses the interaction between vacancy timing and eligibility-list position to compare changes in investigative outcomes after units receive detectives with different total promotion scores. The credibility of this comparison depends on whether the resulting differences in incoming scores are unrelated to units' potential investigative outcomes. We assess this condition using receiving-unit characteristics by promotion-list timing (Figure 2), balance tests, and pre-promotion trends in the empirical-design section. This approach asks whether the characteristics rewarded by APD's existing promotion system predict detectives' contributions to investigative output.

3 Data and Sample Construction

We construct a sample of promotion events into investigative units using the eligibility lists from the 2011, 2013, 2015, and 2017 Austin PD corporal/detective exams. Each list reports the officer's total exam score, its individual components, and the officer's rank on the list. We link the eligibility lists to promotion records to recover each promoted officer's promotion date and assigned unit. Exam-takers can be promoted either to detective or to corporal, a patrol-focused rank. We drop promotions to corporal, which we identify from the assigned unit. We observe 173 detective promotions across the lists in our data (Table A1).

Two features of the promotion data require prediction. First, our promotion data begin partway through the 2011 promotion cycle, so we do not observe promotion dates or assigned units for the first 33 officers on the 2011 list. We predict their promotion dates using a quadratic model of days to promotion as a function of list position, estimated on the promotions for which we observe dates. Second, the promotion data record assigned units in an open-text field, which we hand-code to match the unit codes in the case data. Some entries cannot be matched to a unit code. For these promotees, and for the 33 officers

missing from the 2011 data, we assign the unit in which the promotee takes the most cases, as either lead or supplemental detective, in the first 6 months after promotion. Errors in these predictions misattribute the timing or unit of a promotion. We expect such errors to attenuate our estimates, since a promotion should have no effect before it occurs or in units it did not involve. Units sometimes promote multiple officers in the same month. We treat these promotions as a single event and assign it the average promotion score of the promoted officers.

We link promotion events to the follow-up investigations (cases) that the investigative unit handles around the promotion. The case data cover the universe of 798,709 NIBRS-classified cases handled by Austin PD from 2010 to 2020 (Table A1).[4] For each case, the data report the type of crime, all assigned detectives and supplementary workers, and the investigative unit of each assigned worker. We restrict each case to its lead detective. We link each case to offense reports, arrest records, person files with details on victims and suspects, and dispatch records. We also link promotees to personnel records from the Texas Commission on Law Enforcement (TCOLE). These records report each officer's hire date, educational background, and score on the licensing exam required to become a police officer in Texas.

Our event-study sample consists of cases handled by investigative units in the 9 months before and after a promotion event in the unit. We restrict the sample so that no event window contains another promotion into the same unit. We drop events with another promotion event in the pre-period. We keep events with another promotion in the post-period, but remove cases in the months on or after that promotion. This rule retains more events at the cost of an unbalanced panel. Figure A5 shows that our results are similar in a balanced panel that requires no other promotion events in the 9 months before or after each event. The sample is at the case-by-event-by-relative-month level. Table A1 reports how each restriction affects the number of events and case observations.

We measure investigative outcomes using whether a case results in an arrest, an exceptional clearance, or any clearance. Arrests capture whether an investigation progresses toward prosecution, while exceptional clearances identify cases in which police have identified the offender, established probable cause, and know the offender's location, but circumstances prevent arrest or prosecution. Overall clearance includes both clearance by arrest and exceptional clearance. Examining these outcomes together distinguishes improvements in the likelihood of clearing a case from changes in how cases are resolved. Detectives may

[4]We exclude cases without an offense code in the National Incident Based Reporting System (NIBRS). Based on our conversations with APD, these cases do not reflect investigations of a crime. Instead, they typically record general investigative activity not tied to a specific crime, such as suspicious activity reported by a patrol officer that may later be connected to a crime.

influence these outcomes by developing leads, collecting evidence, locating suspects, and preparing cases that support arrest and prosecution.

To better understand how detective quality affects case resolution, we also examine recorded reasons for exceptional clearance and the number of suspects identified. We distinguish exceptional clearances attributed to victim refusal from those attributed to declined prosecution. Detectives may help sustain victim participation or develop evidence that allows a case to proceed, although recorded reasons for exceptional clearance do not directly measure these efforts. We further distinguish identified suspects who are arrested from those who are not, allowing us to assess whether improvements involve identifying additional suspects or developing investigations that lead to arrest. Finally, we examine arrest outcomes by arrestee race and ethnicity and separately by victim race and ethnicity to distinguish who experiences enforcement from whose victimization cases progress toward prosecution.

3.1 Summary Statistics

Table 1 summarizes the analysis sample, which includes 54 promotion events drawn from four promotion lists across 25 investigative units. The sample contains 160,248 case-by-event observations and 427 unique lead detectives. Before promotion, units have an average of 9.62 detectives assigned cases, with a standard deviation of 4.43. Detectives handle an average of 190 cases per year, although caseloads vary substantially across detective-years, with a standard deviation of 321 cases. Promotion events almost always add one detective, with an average of 1.11 promotees per event.

Panel B summarizes the characteristics of incoming detectives. The average total promotion score is 96.12, consisting of 85.67 raw examination points, 9.59 seniority points, and 0.87 education points. Total scores have a standard deviation of 5.50 points. Raw examination scores exhibit greater dispersion than either seniority or education points, with standard deviations of 6.37, 2.88, and 0.48, respectively.

Panel C reports case outcomes using the same weights as the regressions. Overall, 17.5% of cases result in an arrest and 26.1% are cleared, including 6.8% cleared exceptionally. Exceptional clearance reasons include victim refusal in 4.8% of cases and declined prosecution in 1.9%. Investigations identify an average of 0.74 suspects per case, including both arrested and non-arrested suspects.

4 Empirical Design

We next examine whether differences in incoming detectives' promotion scores translate into differences in investigative outcomes, using the timing of promotions and variation in scores across promotion events. The continuous event-study specification is:

$$Y_{itc} = \sum_{k \neq -1} \beta_k \left(1[\text{RelMonth}_{it} = k] \times \text{AvgScore}_i\right) + \alpha_i + \gamma_t + \delta_{CaseType(c)} + \varepsilon_{itc}, \tag{1}$$

where AvgScore_i is the average total promotion score of newly promoted detectives within promotion event i. Total scores combine raw examination scores, seniority points, and education points; higher scores indicate higher-scoring promotees. The specification includes promotion event (α_i), relative month (γ_t), and NIBRS case type ($\delta_{CaseType(c)}$) fixed effects. We report estimates with and without NIBRS case type fixed effects and obtain similar results, indicating that the findings do not depend on their inclusion. Standard errors are clustered by investigative unit to allow outcomes to be correlated across cases and promotion events within the same unit. Because the sample contains only 25 investigative units (clusters), we report wild-cluster-bootstrap p-values based on 9,999 replications to improve inference with few clusters. Regressions use analytic weights equal to the number of promotees in the event divided by the average number of detectives in the unit during the pre-period. These weights approximate the share of the unit's caseload that the new promotees are expected to take over. Promotees who take over a larger share should have a larger effect on unit-level outcomes, so these events receive more weight.

Our main interest is the coefficients β_k, which capture the dynamic impact of higher average total promotion scores on investigation outcomes. The first key assumption required to interpret β_k as the causal impact of average total promotion score is parallel trends. That is, cases assigned to units that receive promotees with higher and lower average total promotion scores would have followed similar trends in outcomes had the promotion never occurred. We evaluate the credibility of this assumption by testing whether the coefficients on the average total promotion score and relative month indicators are 0 for pre-promotion months ($k < 0$).

Our continuous event study compares units receiving promotees with different average total promotion scores. Identification requires that these units would have experienced the same *treatment effects* had they received promotees of equal quality, ruling out sorting based on potential gains (Callaway et al., 2025). For example, if lighter caseloads or greater staffing allow an incoming detective to be more effective, units with these conditions must not sys-

tematically receive higher-scoring promotees. Otherwise, the estimated score gradient could partly reflect differences in the conditions under which detectives work. APD's promotion process supports this assumption: exam-takers are promoted in rank order, while unit vacancies arise independently of promotion exam timing.

We provide three pieces of empirical evidence to assess potential sorting. First, because units filling vacancies earlier in an eligibility list generally receive higher-scoring detectives, we examine whether receiving-unit characteristics vary with promotion timing. Figure 2 plots differences relative to the month the list opened. Panel (a) examines the shares of cases involving violent crime, property crime, cases originating from 911 calls, and overnight offenses. Panel (b) examines the shares of incumbent officers who are minority or female, while panel (c) examines neighborhood minority population shares, income per capita, and unemployment at the crime location. These differences are expressed in standard deviation units. Panel (d) reports differences in monthly caseloads and the number of active detectives. Neighborhood characteristics, caseloads, and staffing show no clear relationship with promotion timing. Case composition and officer demographics fluctuate, but provide little evidence that earlier promotions systematically enter units with lighter caseloads, greater staffing, or more advantaged neighborhoods.

Second, we examine whether incoming detectives' average total promotion scores are associated with the pre-promotion characteristics of receiving units. Table 2 reports case-level and unit-level regressions using the nine months before promotion. Case-level regressions use the main analysis weights, assessing balance in the mix of cases represented in the analysis. Unit-level regressions average characteristics within promotion events and give each event equal weight, preventing larger caseloads from dominating the comparison. Both approaches include promotion-list fixed effects and cluster standard errors by investigative unit. Neither approach reveals statistically significant relationships between incoming scores and case composition or neighborhood characteristics at conventional levels. At the unit level, scores are also unrelated to pre-promotion caseloads and staffing. The only relationship significant at the 10% level is the share of cases assigned to female detectives in the case-level specification: a one-point increase in average promotion score is associated with a 0.66 percentage-point increase in this share ($p = 0.089$). This relationship is smaller and statistically insignificant when events receive equal weight. Overall, the estimates provide little evidence of systematic sorting on these observed characteristics, although they do not rule out sorting on unobserved determinants of investigative productivity.

Third, following Chen et al. (2026), we use propensity-score matching to assess whether our results persist when comparing promotion events with similar observable characteristics. We divide events into higher- and lower-scoring groups based on the median average total

promotion score and match events using pre-promotion case composition, neighborhood demographics, staffing, and caseloads. Matching improves observable balance (Table A3), and reestimating the continuous event study in the matched sample yields similar effects on arrests (Figure A2, panel b). Appendix A describes the matching procedure in greater detail.

Taken together, these exercises provide evidence against systematic sorting of higher-quality promotees into particular types of investigative units. Although balance on observables and propensity-score matching cannot rule out selection on unobserved characteristics or potential gains, the results support our interpretation that promotion timing generates plausibly exogenous variation in the quality of incoming detectives.

We also perform additional exercises to assess other potential threats to identification. Restricting the sample to events observed throughout the full nine-month pre- and post-promotion window (i.e., a balanced sample) addresses changes in event composition over relative time. Unweighted estimates test whether our preferred weights drive the results by giving each case equal weight. Alternative weights based on pre-promotion caseload examine sensitivity to how we measure exposure to incoming detectives. To account for common calendar-time shocks, we add units that receive no promotion during the event window and include stack-by-unit and stack-by-calendar-month fixed effects. As reported in a later section, the main findings remain similar across these exercises, suggesting that sample composition, weighting, and common calendar-time shocks do not explain the results.

Additional specifications examine whether our findings depend on how detective quality is measured or how it enters the model. Replacing average total promotion score with average exam rank tests sensitivity to using candidates' relative positions rather than the size of score differences. Using raw exam scores removes seniority and education points, allowing us to assess whether the results extend to examination performance alone. Our baseline specification also imposes a linear relationship between average total promotion score and investigative outcomes. Grouping promotion scores into bins relaxes this restriction and allows effects to differ across quality groups. We estimate these specifications both with and without units that receive no promotion to distinguish sensitivity to functional form from sensitivity to the comparison group. The results support the same broad conclusion across these specifications: higher-scoring incoming detectives improve investigative outcomes.

5 Results

5.1 Detective Quality and Case Outcomes

5.1.1 Arrests and Case Resolution

We begin by examining whether higher-scoring incoming detectives affect the likelihood that investigations result in arrest, exceptional clearance, or any clearance. These outcomes distinguish whether cases are resolved from how they are resolved. The balance and matching exercises discussed above provide little evidence that observable sorting drives the results. Figure 3 also shows little evidence of differential pre-promotion trends in arrests or exceptional clearances.

Panel a of Figure 3 shows that higher-scoring incoming detectives increase arrest probability. The effects are small during months 0–1, grow around month three, and remain positive through month nine. Panel A of Table 3 summarizes these estimates: averaging across months 2–9, a one-point increase in incoming detectives' average total promotion score raises arrest probability by 0.42 percentage points, or 2.4 percent relative to the sample mean of 17.45 percent. The standard deviation of the average promotion score across events is 5.5 points. A one-standard-deviation increase therefore raises arrest probability by 2.3 percentage points, or 13.2 percent relative to the sample mean. The effect is statistically significant using wild-cluster-bootstrap inference ($p = 0.017$). Excluding crime-type fixed effects yields a similar increase of 0.36 percentage points ($p = 0.011$).

Panel b of Figure 3 shows a corresponding decline in exceptional clearances that becomes more pronounced after promotion. Over months 2–9, a one-point increase in average promotion score reduces exceptional clearance by 0.41 percentage points, approximately 6 percent relative to the sample mean of 6.8 percent ($p = 0.038$). The estimate without crime-type fixed effects is nearly identical at -0.42 percentage points ($p = 0.048$). In contrast, Table 3 shows little change in overall clearance: an increase of 0.09 percentage points with crime-type fixed effects ($p = 0.489$) and 0.04 percentage points without them (*data&*). Together, these findings indicate that incoming detective quality affects how investigations are resolved, with more arrests and fewer exceptional clearances but little detectable change in overall clearance. They are consistent with higher-scoring detectives helping units overcome barriers that would otherwise prevent an arrest.

5.1.2 Victim Cooperation and Suspect Identification

Having shown that higher-scoring incoming detectives increase arrests and reduce exceptional clearances, we next examine which recorded reasons for exceptional clearance account for

this decline. We focus on cases exceptionally cleared because the victim refused to cooperate or prosecution was declined. Separating these outcomes provides evidence on which barriers to arrest may respond to detective quality.

Figure 4 shows little evidence of differential pre-promotion trends for either outcome. Consistent with this pattern, Panel B of Table 3 cannot reject that the average pre-promotion coefficient equals zero for victim refusal ($p = 0.983$) or declined prosecution ($p = 0.651$). Following promotion, cases in units receiving higher-scoring detectives become less likely to be exceptionally cleared because the victim refused to cooperate. Averaging across months 2–9, a one-point increase in incoming detectives' average total promotion score reduces exceptional clearances attributed to victim refusal by 0.34 percentage points ($p = 0.040$), a 7 percent decline relative to the sample mean of 4.82 percent. The estimated effect on declined prosecution is smaller and statistically insignificant: a decrease of 0.07 percentage points relative to a mean of 1.90 percent ($p = 0.378$). These estimates use crime-type fixed effects; estimates without them are similar.

The decline in exceptional clearances is therefore largely associated with fewer cases exceptionally cleared because of victim refusal. Higher-scoring detectives may be more effective at maintaining contact with victims and securing their participation. However, the data record reasons for exceptional clearance rather than directly measuring cooperation. The results could also reflect stronger evidence that reduces reliance on victim participation or differences in how reasons for exceptional clearance are recorded. We therefore interpret this pattern as consistent with improved victim cooperation, without treating it as definitive evidence of that mechanism.

Next, we examine effects on total suspects identified, non-arrested suspects, and arrested suspects. Unlike the main arrest outcome, which indicates whether a case results in any arrest, these measures count suspects in each category. Figure 5 shows their dynamics, while Panel C of Table 3 summarizes the estimates. None of the average pre-promotion coefficients is statistically distinguishable from zero.

Over months 2–9, a one-point increase in average promotion score increases total suspects identified by 0.0040 per case, approximately 0.5 percent relative to the mean of 0.7402 ($p = 0.067$). Non-arrested suspects decline by 0.0012 per case, although this estimate is imprecise ($p = 0.709$). In contrast, arrested suspects increase by 0.0051 per case, or 2.0 percent relative to the mean of 0.2537 ($p = 0.011$). Without crime-type fixed effects, the increase in arrested suspects remains statistically significant at 0.0046 per case ($p = 0.004$), while the increase in total suspects is smaller and statistically insignificant ($p = 0.340$).

Together, these results provide stronger evidence that higher-scoring incoming detectives increase arrests than that they expand the total number of suspects identified. Combined

with fewer exceptional clearances attributed to victim refusal, the findings are consistent with improvements in developing investigations that support arrest.

5.1.3 Effects on Incumbent Detectives' Cases

Having shown that higher-scoring incoming detectives improve unit-wide arrest outcomes, we next examine whether these improvements extend to cases handled by incumbent detectives. Our outcome is the probability of arrest among cases assigned to detectives already working in the receiving unit. This restriction allows us to assess whether the effects extend beyond investigations handled directly by incoming detectives.

Figure A4 presents the event-study estimates. The pre-promotion coefficients are generally small and show no clear trend, although the estimate for month -2 is positive and larger than the other pre-period estimates. Following promotion, the coefficients become more positive around month three and remain elevated through month nine. Averaging across months 2–9, a one-point increase in incoming detectives' average total promotion score raises arrest probability among incumbent cases by 0.47 percentage points, a 2.8 percent increase relative to the mean of 16.9 percent. The figure reports that this average effect is statistically significant ($p < 0.001$), indicating that improvements extend beyond cases handled directly by incoming detectives.

This pattern is consistent with incoming detectives improving their colleagues' productivity through advice, assistance, or shared investigative knowledge. Changes in case allocation or workloads may also contribute, and the estimates do not distinguish among these channels. Regardless of the mechanism, the results show that the benefits of higher-scoring incoming detectives extend beyond their own investigations to cases handled by their colleagues.

5.2 Detective Quality and Racial Disparities

The effects of incoming detective quality may differ across racial and ethnic groups along two dimensions: who is arrested and whose victimization cases result in an arrest. The first captures the distribution of enforcement, while the second captures the distribution of investigative services. These outcomes need not move together because the race or ethnicity of an arrested suspect may differ from that of the victim. Figure A8 presents event-study estimates for arrests of Black, Hispanic, and White suspects. Each indicator equals one if a case results in at least one arrest of a suspect in the specified group and is measured across all cases, rather than conditional on an arrest occurring. The pre-promotion coefficients fluctuate without a clear, sustained trend. Following promotion, arrests of Hispanic and White suspects increase, with the effects becoming more apparent around month three and

remaining positive through month nine. There is no comparable increase in arrests of Black suspects.

Averaging across months 2–9, a one-point increase in incoming detectives' average total promotion score raises the probability of arresting a Hispanic person by 0.25 percentage points ($p = 0.002$), or 3.8 percent relative to the sample mean of 6.60 percent. The corresponding increase for White arrestees is 0.23 percentage points ($p < 0.001$), or 4.5 percent relative to the mean of 5.13 percent. The estimated effect for Black arrestees is a statistically insignificant decline of 0.08 percentage points ($p = 0.311$). We reject equality of the average effects across the three groups ($p = 0.011$), indicating that the increase in arrests is unevenly distributed across arrestee race and ethnicity.

Next, we examine arrest probability within cases involving Black, Hispanic, or White victims. These estimates capture whether investigations of each group's victimization are more likely to result in an arrest, regardless of the arrestee's race or ethnicity. Figure A9 shows pre-promotion fluctuations without a consistent trend and generally positive estimates following promotion, although the timing and precision vary across groups. Over months 2–9, a one-point increase in average promotion score raises arrest probability by 0.25 percentage points among cases involving Black victims, 0.35 percentage points among cases involving Hispanic victims, and 0.55 percentage points among cases involving White victims. Relative to the respective sample means of 19.18, 20.51, and 11.44 percent, these estimates represent increases of 1.3, 1.7, and 4.8 percent. The average effect is statistically significant for White victims ($p = 0.014$), marginally significant for Hispanic victims ($p = 0.079$), and less precisely estimated for Black victims ($p = 0.124$).

The largest point estimate occurs among cases involving White victims, which have the lowest sample arrest rate. However, we cannot reject equality of the effects across victim groups ($p = 0.52$), so the estimates do not establish that higher-scoring detectives change racial disparities in investigative services. Taken together, the results show that additional arrests primarily involve Hispanic and White arrestees, while the positive victim-group estimates suggest that improvements in investigative outcomes extend across racial and ethnic groups.

5.3 Assessing Potential Threats to Identification

Our main results indicate that higher-quality incoming detectives increase arrests and reduce exceptional clearances, particularly those attributed to victim refusal. We now examine whether these findings depend on sample construction, weighting, or modeling choices.

Event composition and weighting. Our baseline sample excludes cases occurring on or after another promotion into the same unit, resulting in an unbalanced event window. To assess whether changes in event composition over relative time affect the estimates, we restrict the sample to events observed throughout the full nine-month pre- and post-promotion window. This restriction holds event composition fixed but reduces the sample from 54 to 34 events, sacrificing precision and focusing on units with more widely spaced promotions. Panel a of Figure A5 shows similar dynamics, although the average arrest effect over months 2–9 declines from 0.42 to 0.23 percentage points per promotion-score point. We cannot reject equality of these estimates ($p = 0.188$). Thus, the positive pattern persists in the balanced sample, although the estimated magnitude is smaller.

Beyond which events enter the sample, the estimates may depend on how much weight each event receives. Panel b of Figure A5 compares the baseline with estimates giving each case equal weight. The average effect declines from 0.42 to 0.31 percentage points, consistent with giving relatively more influence to events in which incoming detectives represent a smaller share of unit staffing. If their influence on unit-wide outcomes increases with that share, these events should exhibit smaller effects. Panel c of Figure A5 instead weights observations by promotees relative to pre-promotion caseload and yields an effect of 0.37 percentage points. These weights measure exposure through case quantity rather than the personnel shock and may place less emphasis on opportunities for incoming detectives to influence their colleagues. A smaller estimate is therefore consistent with effects operating partly through peers. Neither alternative differs significantly from the baseline, with equality-test p-values of 0.106 and 0.694, respectively. Overall, the positive arrest effect persists across weighting rules, with smaller estimates consistent with differences in how each approach weights exposure to incoming detectives.

Calendar-time shocks. Another concern is that promotion timing coincides with broader changes in policing or investigative outcomes. Our baseline specification includes promotion-event, relative-month, and NIBRS case-type fixed effects. These account for persistent differences across promotion events, common patterns over event time, and differences across crime categories, but do not absorb shocks occurring in a particular calendar month. For example, higher-scoring promotions could coincide with a department-wide enforcement initiative, a change in arrest policy, or seasonal conditions that make cases easier to resolve. The estimated score gradient could then partly reflect these changes rather than incoming detective quality.

To assess this possibility, we add comparison units that receive no promotion during the event window and include stack-by-unit and stack-by-calendar-month fixed effects. This

specification compares outcomes within units while absorbing calendar-month shocks shared by promoted and comparison units in each stack. Figure A6 shows that the average arrest effect remains essentially unchanged at 0.42 percentage points per promotion-score point and remains statistically significant ($p = 0.001$). We cannot reject equality with the baseline estimate ($p = 0.955$). Although this exercise does not absorb shocks unique to receiving units, the stability of the estimates argues against department-wide policy changes, common seasonality, or other shared calendar-time shocks as explanations for the increase in arrests.

Quality measurement and functional form. The total promotion score rewards examination performance, experience, and education, each of which could contribute to investigative productivity. To distinguish their contributions, Figure A3 reports a specification that includes separate event-time interactions for all three components, standardized so their magnitudes are comparable. Averaging across months 2–9, a one-standard-deviation increase in raw exam score raises arrest probability by 2.51 percentage points ($p = 0.001$). The corresponding estimates for seniority and education points are also positive, at 0.97 and 0.76 percentage points, respectively, but less precisely estimated ($p = 0.199$ and $p = 0.331$). A test of equality across the three effects yields $p = 0.057$. Examination performance therefore emerges as the strongest component of the score, while the positive estimates for seniority and education leave room for experience and educational attainment to contribute independently.

A separate question is whether the continuous specification masks differences in the returns to quality across the score distribution. Panel a of Figure A7 replaces the continuous score with groups approximately one standard deviation wide. Relative to the lowest-score group, average arrest effects over months 2–9 rise across successively higher groups: 1.4, 3.6, and 6.9 percentage points. Panels b–d incorporate non-promoted comparison units and compare score quartiles. Relative to the first quartile, the estimated effects are 0.89 percentage points for the second quartile ($p = 0.482$), 2.55 percentage points for the third ($p = 0.096$), and 6.79 percentage points for the fourth ($p < 0.001$). Both approaches show progressively larger improvements among higher-scoring promotees, with the clearest gains at the top of the distribution. Thus, the conclusion that higher-scoring detectives improve arrest outcomes persists when effects can vary flexibly across score groups.

6 Conclusion

Detective quality matters for how criminal investigations are resolved. Exploiting promotion-list ordering and vacancy timing, this paper shows that higher-scoring incoming detectives

increase arrests and reduce exceptional clearances, with little change in overall clearance. Fewer exceptional clearances attributed to victim refusal and more arrested suspects suggest that these improvements involve developing investigations that support arrest. The gains extend to cases handled by incumbent detectives, demonstrating that incoming detectives contribute beyond their own caseloads. Together, these findings show that personnel quality can improve investigative performance across an entire unit, making promotion decisions consequential for the delivery of criminal justice services.

References

Adger, C., Ross, M. B., and Sloan, C. (2026). The effect of field training officers on police use of force. *American Economic Review*, 116(5):1837–1875.

Arcidiacono, P., Kinsler, J., and Price, J. (2017). Productivity spillovers in team production: Evidence from professional basketball. *Journal of Labor Economics*, 35(1):191–225.

Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. S. (2020). Losing prosociality in the quest for talent? sorting, selection, and productivity in the delivery of public services. *American Economic Review*, 110(5):1355–1394.

Austin Police Department (2025). *General Orders*. Austin Police Department. Edition issued September 5, 2025.

Ba, B., Ge, H., Kaplan, J., Knox, D., Komisarchik, M., Lanzalotto, G., Mariman, R., Mummolo, J., Rivera, R., and Torres, M. (2025). Political diversity in U.S. police agencies. *American Journal of Political Science*, 69(4):1617–1635.

Ba, B. A., Knox, D., Mummolo, J., and Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in Chicago. *Science*, 371(6530):696–702.

Benson, A., Li, D., and Shue, K. (2019). Promotions and the Peter Principle. *The Quarterly Journal of Economics*, 134(4):2085–2134.

Braga, A. A., Turchan, B. S., and Barao, L. M. (2019). The influence of investigative resources on homicide clearances. *Journal of Quantitative Criminology*, 35(2):337–364.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. (2025). Difference-in-differences with a continuous treatment.

Chaiken, J. M. (1975). The criminal investigation process: Volume II: Survey of municipal and county police departments. Technical Report R-1777-DOJ, RAND Corporation, Santa Monica, CA.

Chaiken, J. M., Greenwood, P. W., and Petersilia, J. R. (1977). The criminal investigation process: A summary report. *Policy Analysis*, 3(2):187–217.

Chalfin, A., Hansen, B., Weisburst, E. K., and Williams, Morgan C., J. (2022). Police force size and civilian race. *American Economic Review: Insights*, 4(2):139–158.

Chalfin, A. and McCrary, J. (2018). Are U.S. cities underpoliced? theory and evidence. *Review of Economics and Statistics*, 100(1):167–186.

Chen, H., Chen, Y., and Yang, Q. (2026). Women in the courtroom: Technology and justice. *Review of Economic Studies*, 93(3):1574–1601.

City of Austin and Austin Police Association (2013). Agreement between the city of austin and the austin police association. Effective October 1, 2013–September 30, 2017.

City of Austin Firefighters', Police Officers', and Emergency Medical Services Personnel's Civil Service Commission (2019). Police Corporal/Detective Exam: Notice of Examination with Source Material List. Civil service examination notice. Posted July 5, 2019, for the examination scheduled January 9, 2020.

Cook, P. J. and Ludwig, J. (2018). Policing guns: Why gun violence is not (just) a public health problem. Social Science Research Council, *Items*.

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

Dube, O., MacArthur, S. J., and Shah, A. K. (2025). A cognitive view of policing. *Quarterly Journal of Economics*, 140(1):745–791.

Evans, W. N. and Owens, E. G. (2007). COPS and crime. *Journal of Public Economics*, 91(1-2):181–201.

Federal Bureau of Investigation (2020). Clearances. Crime in the United States, 2019.

Goldstein, R. (2022). Inequality in the provision of police services: Evidence from residential burglary investigations. *The Journal of Law and Economics*, 65(3):487–513.

Goodison, S. E. (2021). The role of detective experience on homicide clearances. *Policing: An International Journal*, 44(4):548–559.

Goodison, S. E. (2022). Local police departments personnel, 2020. Bulletin NCJ 305187, Bureau of Justice Statistics.

Greenwood, P. W., Chaiken, J. M., Petersilia, J. R., Prusoff, L. L., Castro, R. P., Kellen, K., and Wildhorn, S. (1975). The criminal investigation process: Volume III: Observations and analysis. Technical Report R-1778-DOJ, RAND Corporation, Santa Monica, CA.

Greenwood, P. W. and Petersilia, J. R. (1975). The criminal investigation process: Volume I: Summary and policy implications. Technical Report R-1776-DOJ, RAND Corporation, Santa Monica, CA.

Hoekstra, M. and Sloan, C. (2022). Does race matter for police use of force? evidence from 911 calls. *American Economic Review*, 112(3):827–860.

Hoffman, M., Kahn, L. B., and Li, D. (2018). Discretion in hiring. *Quarterly Journal of Economics*, 133(2):765–800.

Kingshott, B. F., Walsh, J. P., and Meesig, R. T. (2015). Are we training our detectives? a survey of large law enforcement agencies regarding investigation training and training needs. *Journal of Applied Security Research*, 10(4):481–509.

Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, 117(5):914–940.

Lum, C., Wellford, C., Scott, T. L., Vovak, H., Scherer, J. A., and Goodier, M. (2024). Differences between high and low performing police agencies in clearing robberies, aggravated assaults, and burglaries: Findings from an eight-agency case study. *Police Quarterly*, 27(2):135–157.

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

McCrary, J. (2007). The effect of court-ordered hiring quotas on the composition and quality of police. *American Economic Review*, 97(1):318–353.

Mello, S. (2019). More COPS, less crime. *Journal of Public Economics*, 172:174–200.

Owens, E. and Ba, B. (2021). The economics of policing and public safety. *Journal of Economic Perspectives*, 35(4):3–28.

Police Executive Research Forum (2012). Austin Police Department Patrol Utilization Study: Final Report. Technical report, Police Executive Research Forum, Washington, DC.

Prendergast, C. (2007). The motivation and bias of bureaucrats. *American Economic Review*, 97(1):180–196.

Prince, H., Lum, C., and Koper, C. S. (2021). Effective police investigative practices: An evidence-assessment of the research. *Policing: An International Journal*, 44(4):683–707.

Rivera, R. (2025). Do peers matter in the police academy? *American Economic Journal: Applied Economics*, 17(2):127–164.

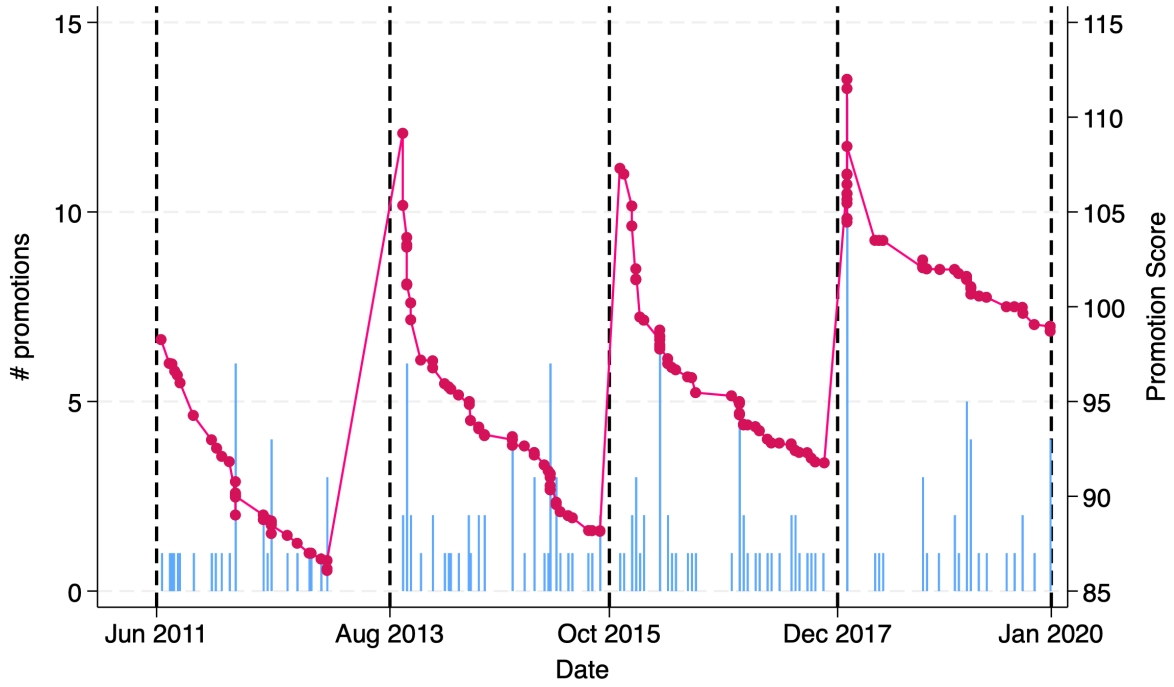
Rivera, R. (2026). What makes a good cop? measuring officer quality and the trade-off between policing and public safety. Working paper.

Sandvik, J. J., Saouma, R. E., Seegert, N. T., and Stanton, C. T. (2020). Workplace knowledge flows. *The Quarterly Journal of Economics*, 135(3):1635–1680.

Snider, J. (2025). Solving crime, protecting communities: A blueprint for safer streets. R Street Policy Study 327, R Street Institute.

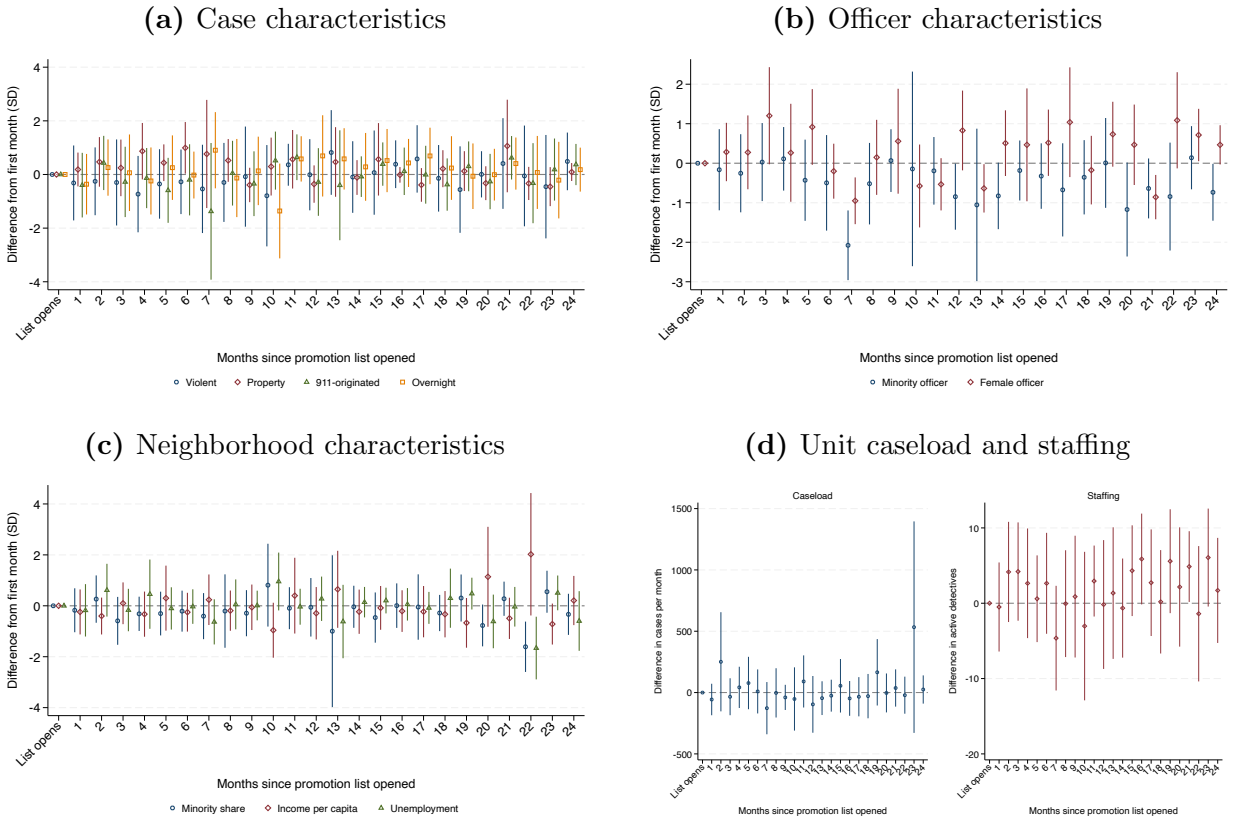
Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, 103(2):280–293.

Figure 1: Promotion Timing, Detective Quality, and Investigative Outcomes



Notes: This figure plots the timing and promotion scores of the 173 detective promotions in our data. Bars show the number of promotions in each week (left axis). The pink line shows each promoted officer's total promotion score, plotted at the officer's promotion date (right axis). The total score combines the written exam score with seniority and education points. Dashed vertical lines mark the dates of the corporal/detective promotion exams. Each exam produces a new eligibility list, and officers are promoted from the list in rank order until the next exam. Promotion dates for the first 33 officers on the 2011 list are predicted, as described in Section 3.

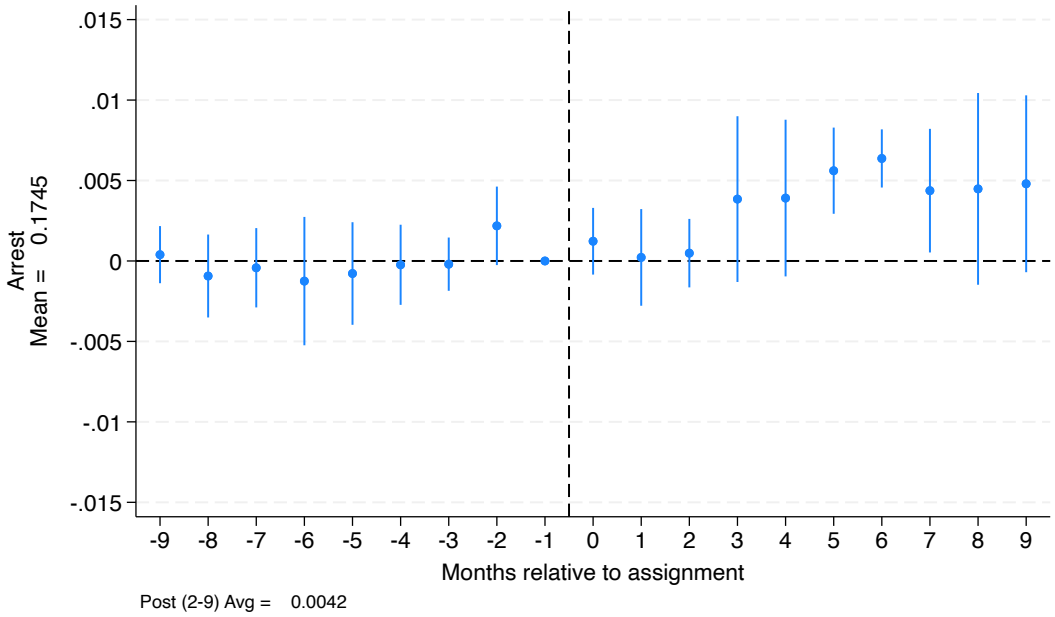
Figure 2: Promotion-List Timing and Receiving-Unit Characteristics



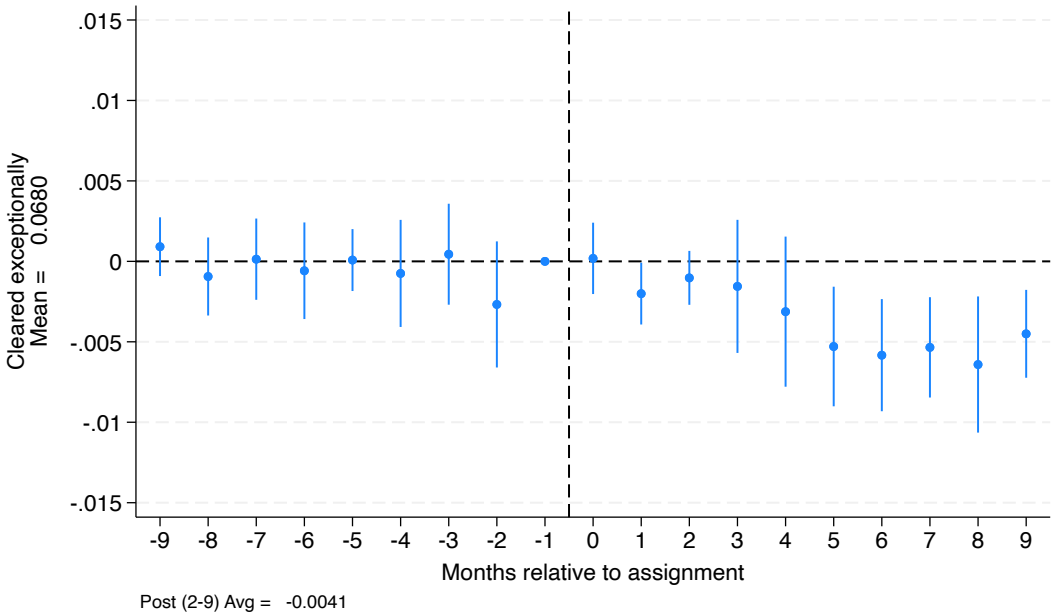
Notes: This figure examines whether the characteristics of units receiving promotees vary with time since the eligibility list opened. The sample includes all 173 detective promotion events on the 2011, 2013, 2015, and 2017 lists. An event is a unit-month with at least one promotion. Each event's characteristics are averaged over the cases its unit handled in the 9 months before the promotion. Each series comes from a separate regression of the characteristic on indicators for months since the list opened and list fixed effects, with standard errors clustered by investigative unit. Coefficients are relative to the list-opening month, and bars show 95 percent confidence intervals. Panels (a)–(c) standardize each characteristic across events. Panel (a) reports the shares of violent, property, 911-originated, and overnight offenses. Panel (b) reports the shares of cases led by minority and female detectives. Panel (c) reports the minority share, income per capita, and unemployment rate of the neighborhoods where cases occurred. Panel (d) reports, in levels, the average number of cases and of active detectives per month.

Figure 3: Effects of Promotee Quality on Arrests and Exceptional Clearances

(a) Arrest



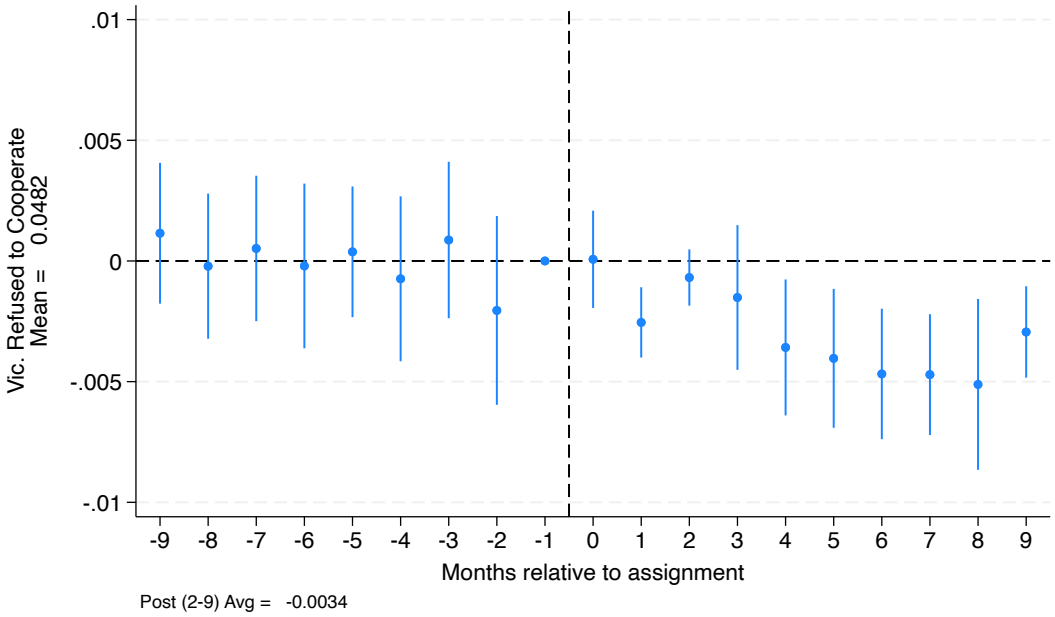
(b) Exceptional clearance



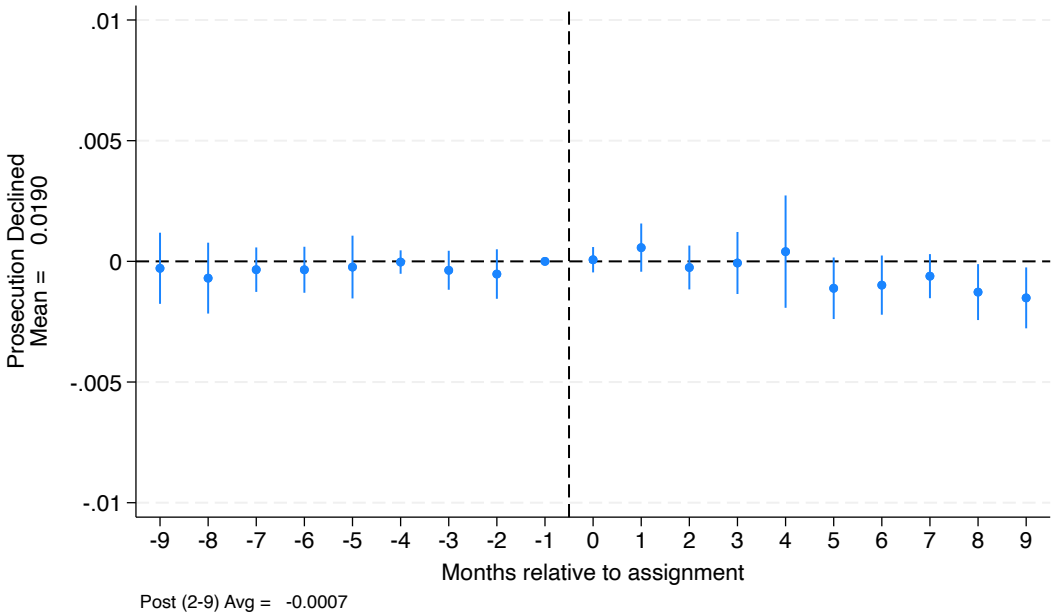
Notes: Each panel plots event-time coefficients for an additional point of incoming detectives' average promotion score; the horizontal axis is months relative to promotion (omitting month -1) and the vertical axis is the change in the indicated case outcome. Regressions include promotion-event, relative-month, and crime-type fixed effects and cluster standard errors at the investigative-unit level. An exceptional clearance occurs when police identify the offender, establish probable cause for arrest and prosecution, and know the offender's exact location, but cannot arrest or prosecute because the offender has died, prosecution is declined for reasons other than lack of probable cause, the offender is in another jurisdiction's custody (including extradition denied), the victim refuses to cooperate, or the offender is a juvenile handled without custody or court referral.

Figure 4: Effects on Reasons for Exceptional Clearance

(a) Victim refused to cooperate



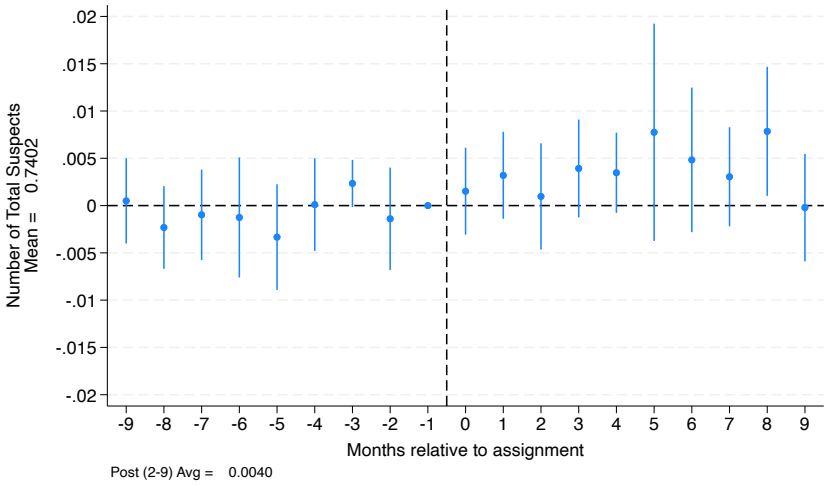
(b) Prosecution declined



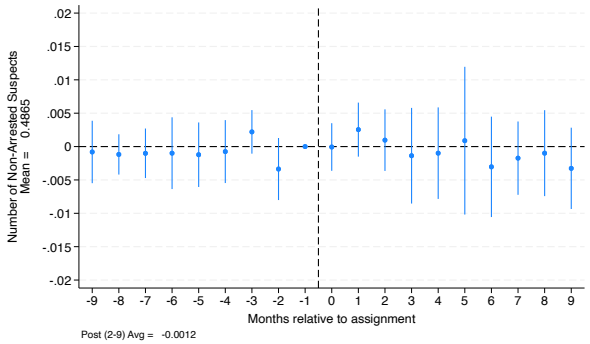
Notes: Each panel plots event-time coefficients for an additional point of incoming detectives' average promotion score. The horizontal axis is months relative to promotion, omitting month -1, and the vertical axis is the change in the probability that a case receives the indicated exceptional-clearance designation. In panel (a), the outcome equals one if the case is exceptionally cleared because the victim refused to cooperate. In panel (b), the outcome equals one if the case is exceptionally cleared because prosecution was declined. Regressions use the analysis sample and include promotion-event, relative-month, and crime-type fixed effects. Standard errors are clustered at the investigative-unit level. Figure 3 defines exceptional clearance.

Figure 5: Additional Investigative Outcomes

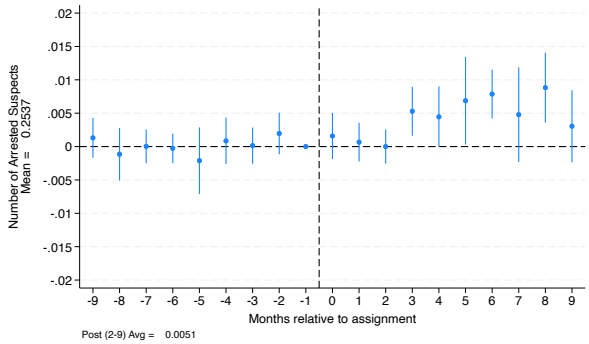
(a) Total suspects identified



(b) Non-arrested suspects



(c) Arrested suspects



Notes: Each panel plots event-time coefficients for an additional point of incoming detectives' average promotion score. The horizontal axis is months relative to promotion, omitting month −1, and the vertical axis is the change in the indicated case-level count. Panel (a) reports the total number of suspects identified in the case. Panels (b) and (c) split this total into suspects who were not arrested and suspects who were arrested. Regressions use the analysis sample and include promotion-event, relative-month, and crime-type fixed effects. Standard errors are clustered at the investigative-unit level.

Table 1: Summary Statistics: Detectives, Promotions, and Investigations

Characteristic	Mean / count	SD
Panel A. Analysis sample and investigative units		
Promotion lists	4	
Promotion events	54	
Units receiving promotees	25	
Unique detectives in analysis sample	427	
Cases in analysis sample	160,248	
Detectives per unit, pre-promotion	9.62	4.43
Cases per detective per year	190.17	321.24
Panel B. Promotee characteristics		
Total promotion score	96.12	5.50
Raw exam score	85.67	6.37
Seniority points	9.59	2.88
Education points	0.87	0.48
Promotees per event	1.11	0.37
Panel C. Case outcomes		
Any arrest	0.175	0.380
Any clearance	0.261	0.439
Exceptional clearance	0.068	0.252
Victim refusal	0.048	0.214
Prosecution declined	0.019	0.137
Number of suspects identified	0.740	1.020

Notes: This table reports counts, means, and standard deviations for the Austin sample. Cases are case-by-event observations. Pre-promotion staffing averages the number of detectives assigned cases over the preceding nine months. Annual caseload is measured across active detective-years in sample units during 2010–2020. Scores are summarized across promotees; staffing and promotees per event are summarized across events. Any clearance includes arrest and exceptional clearance. Suspect counts include arrested and non-arrested suspects. Panel C uses each outcome's estimation sample and weights equal to promotees divided by pre-promotion staffing. Other statistics are unweighted.

Table 2: Pre-Promotion Characteristics and Incoming Detective Scores

	Case-level regressions				Unit-level regressions			
Characteristic	Mean	Score slope	SE	p-value	Mean	Score slope	SE	p-value
Violent crime	0.0868	0.0070	0.0084	0.4115	0.1555	0.0031	0.0091	0.7389
Property crime	0.4125	0.0093	0.0127	0.4729	0.2272	0.0093	0.0090	0.3111
911-originated	0.6978	-0.0067	0.0072	0.3656	0.7388	0.0060	0.0075	0.4285
Overnight offense	0.2263	0.0014	0.0014	0.2972	0.2264	0.0024	0.0031	0.4353
Nbhd. minority share	0.5977	0.0022	0.0017	0.2170	0.6006	0.0013	0.0022	0.5643
Nbhd. income per capita	34,357	-260	185	0.1731	35,875	97	254	0.7050
Nbhd. unemployment share	0.0553	-0.0000	0.0000	0.9661	0.0558	0.0001	0.0001	0.6518
Minority detective	0.4750	-0.0038	0.0100	0.7094	0.4801	-0.0038	0.0074	0.6150
Female detective	0.2235	0.0066	0.0037	0.0894	0.1700	0.0030	0.0034	0.3957
Cases per month	–	–	–	–	171.5782	-0.0463	4.7353	0.9923
Active detectives per month	–	–	–	–	9.6152	-0.0449	0.1473	0.7629

Notes: Each row reports separate regressions of the indicated pre-promotion characteristic on average total promotion score and promotion-list fixed effects. Slopes represent changes per one-point increase in score. Case-level regressions use cases from months -9 through -1, weighted by promotees divided by average pre-promotion staffing. Unit-level regressions average characteristics within promotion events and give each event equal weight. Caseload and staffing are averages across observed pre-promotion months. Detective demographics describe officers assigned cases. Standard errors are clustered by investigative unit. Nbhd. denotes neighborhood. All shares are measured from zero to one; income is in dollars.

Table 3: Effect of Promotee Quality on Investigative Outcomes

Panel A. Case resolution

	Any clearance		Arrest		Exceptional clearance	
Transition ($k = 0, 1$)	0.0006	0.0009	0.0003	0.0007	-0.0009	-0.0009
	(0.0022)	(0.0017)	(0.0009)	(0.0007)	(0.0011)	(0.0009)
	[0.841]	[0.663]	[0.704]	[0.326]	[0.762]	[0.626]
Main effect ($k = 2, \ldots, 9$)	0.0004	0.0009	0.0036**	0.0042**	-0.0042**	-0.0041**
	(0.0017)	(0.0012)	(0.0008)	(0.0010)	(0.0016)	(0.0014)
	[0.793]	[0.489]	[0.011]	[0.017]	[0.048]	[0.038]
Outcome mean	0.2609	0.2609	0.1745	0.1745	0.0680	0.0680
Pre-period p-value	[0.946]	[0.646]	[0.553]	[0.868]	[0.656]	[0.782]

Panel B. Reasons for exceptional clearance

	Victim refusal		Prosecution declined	
Transition ($k = 0, 1$)	-0.0012	-0.0012	0.0003	0.0003
	(0.0008)	(0.0007)	(0.0003)	(0.0003)
	[0.287]	[0.120]	[0.373]	[0.343]
Main effect ($k = 2, \ldots, 9$)	-0.0034**	-0.0034**	-0.0007	-0.0007
	(0.0012)	(0.0010)	(0.0005)	(0.0005)
	[0.042]	[0.040]	[0.381]	[0.378]
Outcome mean	0.0482	0.0482	0.0190	0.0190
Pre-period p-value	[0.945]	[0.983]	[0.478]	[0.651]

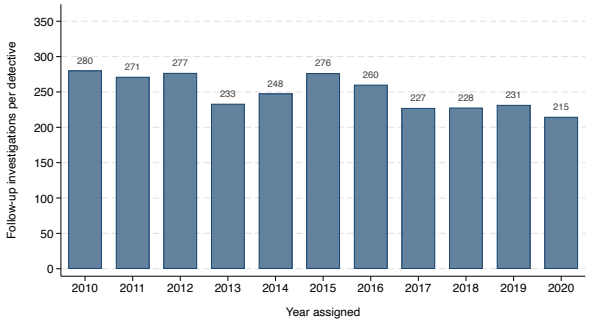
Panel C. Suspect identification and arrest

	Total suspects		Non-arrested suspects		Arrested suspects	
Transition ($k = 0, 1$)	0.0018	0.0024	0.0011	0.0012	0.0008	0.0011
	(0.0025)	(0.0018)	(0.0019)	(0.0015)	(0.0014)	(0.0011)
	[0.494]	[0.209]	[0.595]	[0.432]	[0.592]	[0.324]
Main effect ($k = 2, \ldots, 9$)	0.0035	0.0040*	-0.0011	-0.0012	0.0046***	0.0051**
	(0.0034)	(0.0022)	(0.0033)	(0.0025)	(0.0011)	(0.0013)
	[0.340]	[0.067]	[0.782]	[0.709]	[0.004]	[0.011]
Outcome mean	0.7402	0.7402	0.4865	0.4865	0.2537	0.2537
Pre-period p-value	[0.605]	[0.598]	[0.522]	[0.545]	[0.765]	[0.913]
Promotion events	54	54	54	54	54	54
Cases	160,248	160,248	160,248	160,248	160,248	160,248
Event FE	Yes	Yes	Yes	Yes	Yes	Yes
Rel. month FE	Yes	Yes	Yes	Yes	Yes	Yes
Crime-type FE	No	Yes	No	Yes	No	Yes

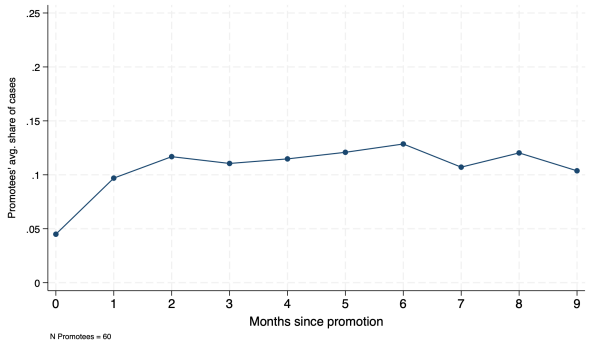
Notes: Each column within a panel reports a separate specification. Entries report averages of event-time coefficients over the indicated months for a one-point increase in incoming detectives' average total promotion score, relative to the omitted month $k = -1$. For each outcome, the first column excludes crime-type fixed effects and the second includes them. The fixed-effect indicators and sample counts at the bottom apply to all panels. Outcomes in Panels A and B are case-level indicators; outcomes in Panel C count suspects per case. Regressions use analytic weights equal to the number of promotees divided by the average number of pre-promotion detectives in the unit. Outcome means use the same weights and each regression's estimation sample. Cases reports the number of case-by-event observations used in estimation. Parentheses contain standard errors clustered by investigative unit, and brackets contain wild-cluster-bootstrap p-values based on 9999 replications. The pre-period p-value tests whether the average coefficient over months $k = -9, \ldots, -2$ equals zero; it is not a joint test of all pre-period coefficients. Significance stars are based on bootstrap p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Figure A1: Detective Workload, Integration, and Raw Exam Correlates

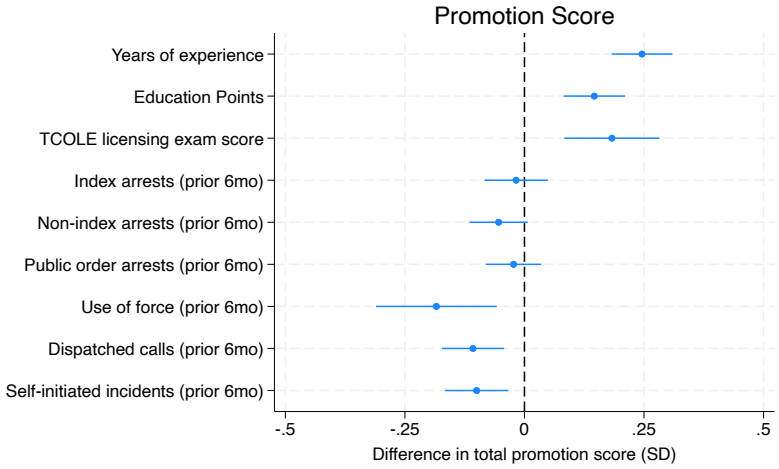
(a) Cases per detective by year



(b) Newly promoted detective case share



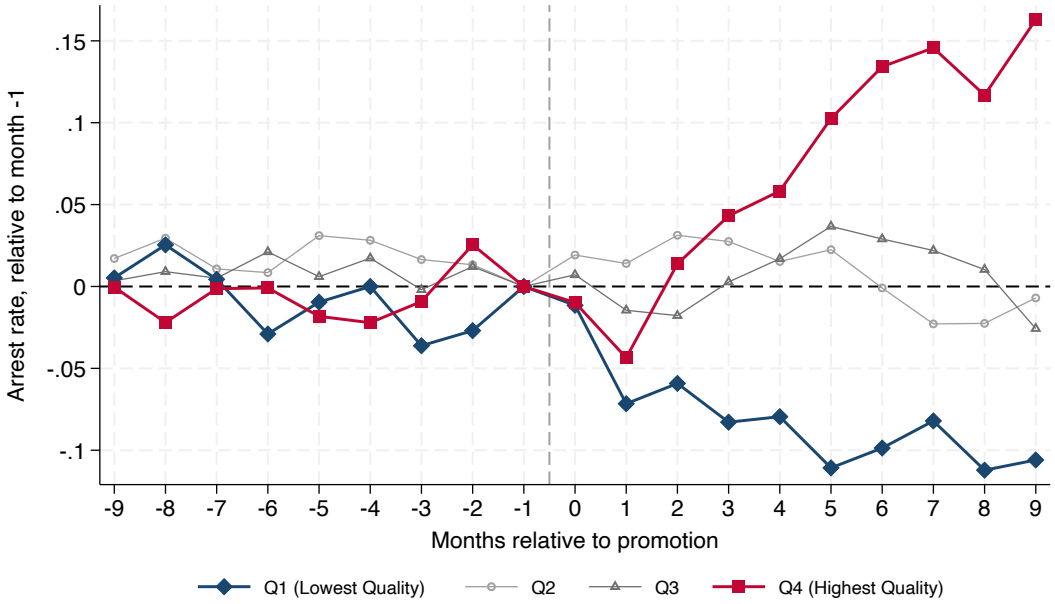
(c) Pre-promotion correlates of total exam score



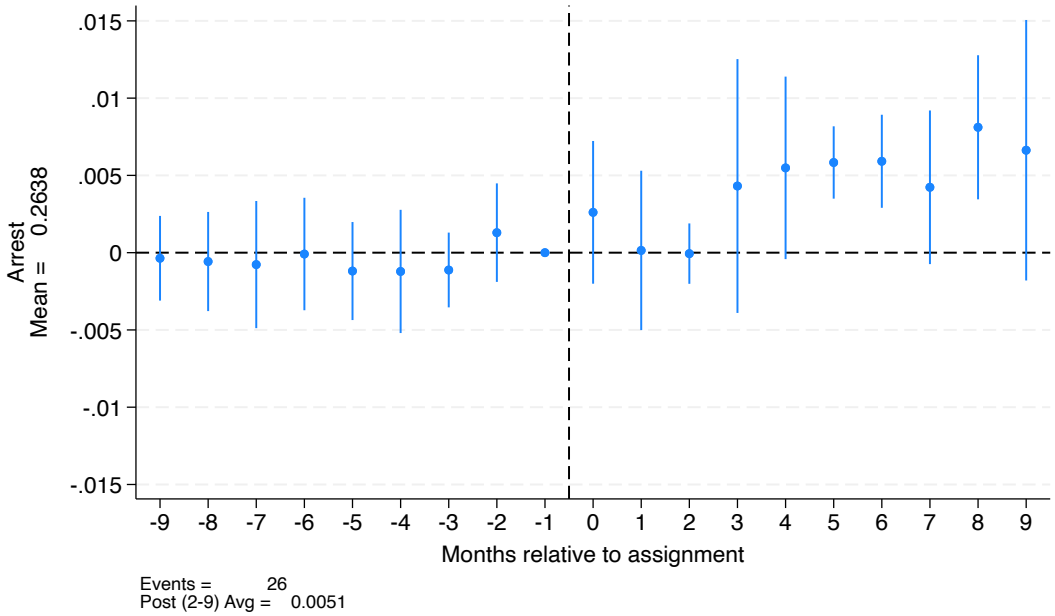
Notes: Panel (a) plots follow-up investigations per detective by calendar year. The sample includes the 787,633 cases handled by units with at least 1,000 cases from 2010 to 2020. Panel (b) plots the share of the unit's assigned cases led by newly promoted detectives in each month after promotion. The sample includes the 60 promotees in the analysis sample described in Table 1, with a share of zero in months when a promotee leads no cases. Panel (c) plots differences in the total promotion exam score, in standard deviations, associated with officers' experience, credentials, and patrol activity in the six months before the exam. Each estimate comes from a separate regression of the standardized score on the characteristic and list fixed effects, with standard errors clustered by officer. The sample includes the 753 of 815 exam scores from the four promotion lists that we can match to officer badge numbers. Estimates for the TCOLE licensing exam score use the 403 scores of officers hired after 2005, when our licensing exam data begin.

Figure A2: Descriptive Arrest Trends and Matched-Sample Estimates

(a) Raw arrest trajectories by score quartile

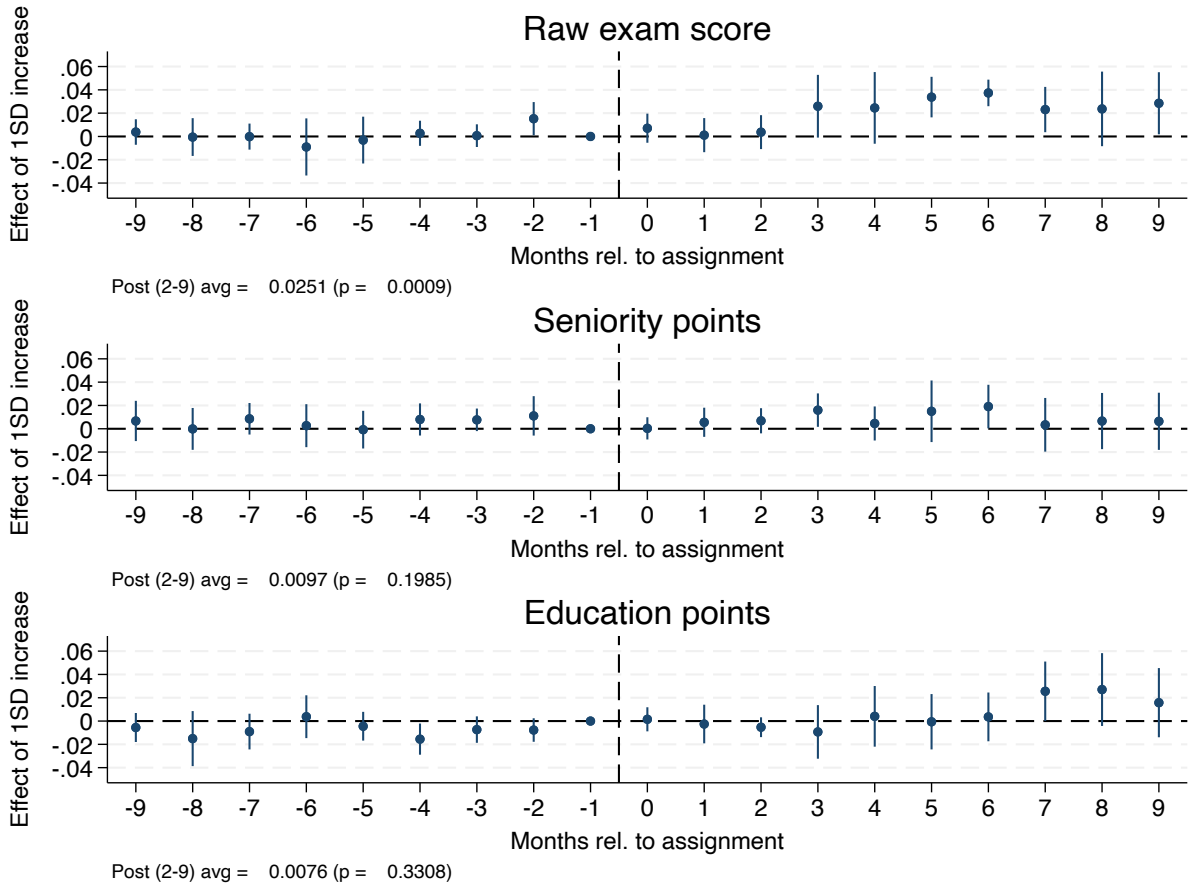


(b) Matched-event arrest study



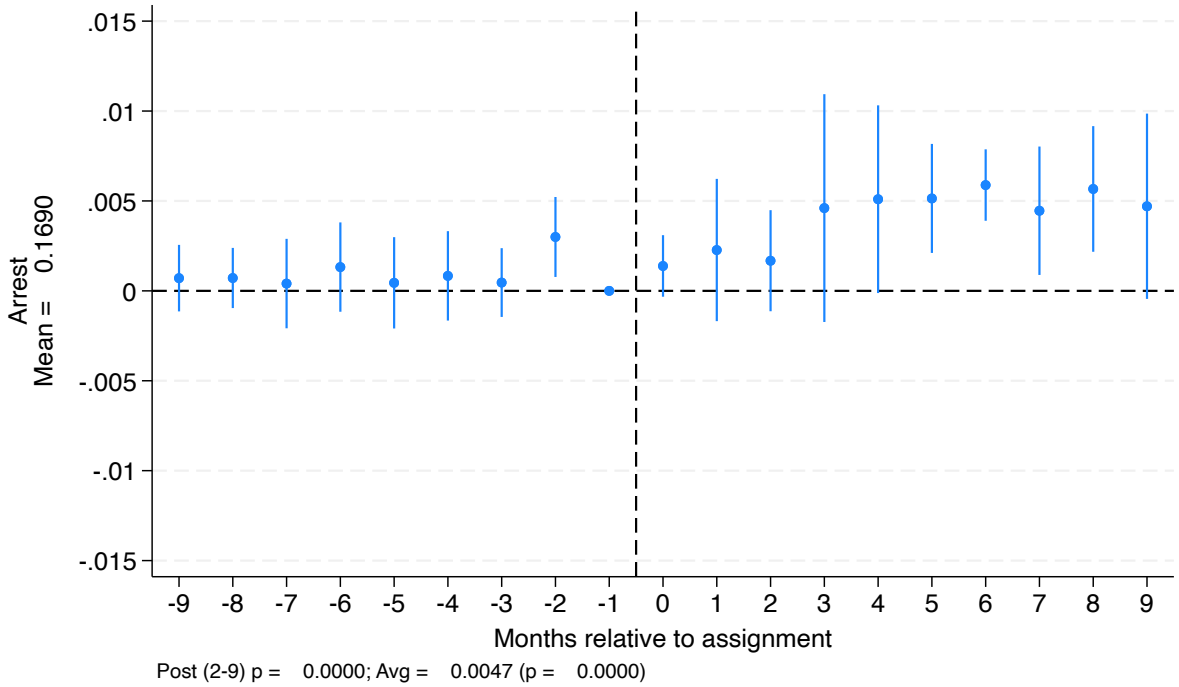
Notes: Panel (a) plots mean arrest rates by event time separately for promotion events in each quartile of incoming detectives' average promotion score. Each series is normalized to zero in month −1 and includes no controls. Panel (b) plots event-time coefficients for an additional point of incoming detectives' average promotion score, estimated on the propensity-score-matched sample of 26 promotion events (13 above and 13 below the median score) described in Appendix A. The horizontal axis is months relative to promotion, omitting month −1. Regressions include promotion-event, relative-month, and crime-type fixed effects. Standard errors are clustered at the investigative-unit level.

Figure A3: Promotion-Score Components and Arrest Effects



Notes: This figure decomposes the effect of promotee quality on arrests into the three components of the total promotion score: the raw exam score, seniority points, and education points. Each component is standardized to mean zero and standard deviation one, so coefficients measure the effect of a one-standard-deviation increase. Event-time interactions for all three components enter a single regression, so each panel reports the effect of one component holding the other two fixed. The horizontal axis is months relative to promotion, omitting month -1. The regression uses the analysis sample and includes promotion-event, relative-month, and crime-type fixed effects. Standard errors are clustered at the investigative-unit level. Each panel reports the average effect over months 2 to 9 and its p-value.

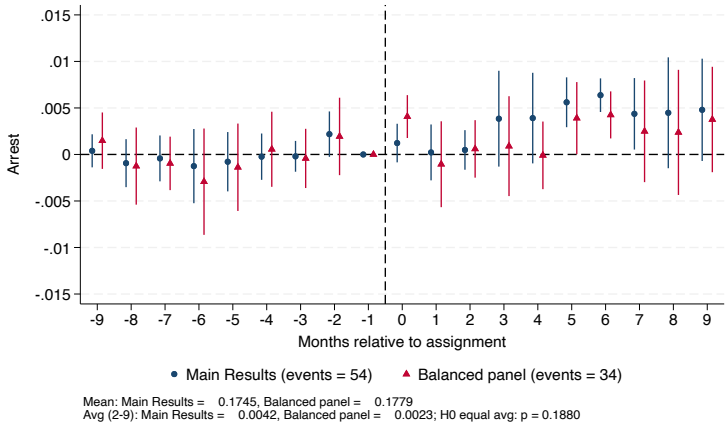
Figure A4: Effects of Promotee Quality on Cases Handled by Incumbent Detectives



Notes: This figure plots event-time coefficients for an additional point of incoming detectives' average promotion score, estimated on cases led by incumbent detectives. Incumbents are detectives who lead at least one case in the unit both before and after the promotion, which excludes the promotees themselves. The outcome is an indicator for whether the case results in an arrest. The horizontal axis is months relative to promotion, omitting month -1. Regressions use the analysis sample and include promotion-event, relative-month, crime-type, and detective fixed effects, so estimates reflect changes in arrest rates within incumbent detectives. Standard errors are clustered at the investigative-unit level, and bars show 95 percent confidence intervals.

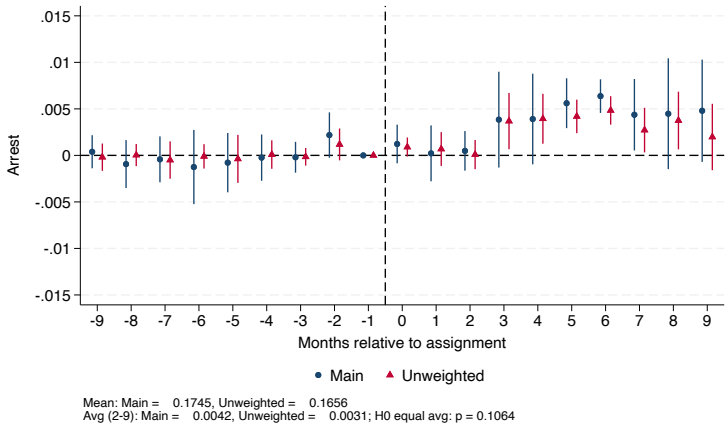
Figure A5: Composition and Weighting Robustness

(a) Balanced event window



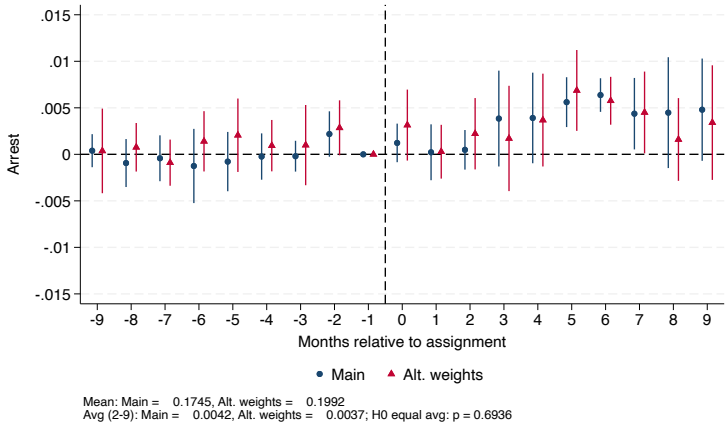
Mean: Main Results = 0.1745, Balanced panel = 0.1779
Avg (2-9): Main Results = 0.0042, Balanced panel = 0.0023; H0 equal avg: p = 0.1880

(b) Unweighted estimator



Mean: Main = 0.1745, Unweighted = 0.1656
Avg (2-9): Main = 0.0042, Unweighted = 0.0031; H0 equal avg: p = 0.1064

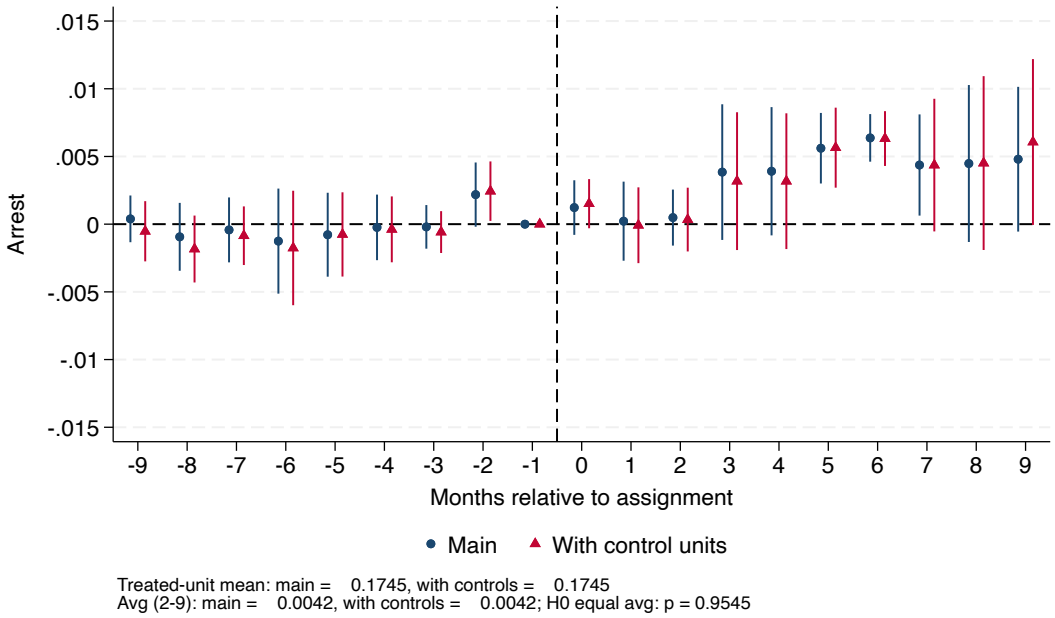
(c) Alternative caseload weights



Mean: Main = 0.1745, Alt. weights = 0.1992
Avg (2-9): Main = 0.0042, Alt. weights = 0.0037; H0 equal avg: p = 0.6936

Notes: Each panel compares the main event-study estimates for arrests with estimates from an alternative sample or weighting rule. All estimates are event-time coefficients for an additional point of incoming detectives' average promotion score, from regressions that include promotion-event, relative-month, and crime-type fixed effects, with standard errors clustered at the investigative-unit level. The horizontal axis is months relative to promotion, omitting month −1. Panel (a) re-estimates the model on a balanced panel. This sample drops events with another promotion into the unit in the 9 months before or after the event, rather than truncating the post-period at the next promotion. Panel (b) estimates the model without weights. Panel (c) replaces the main weights with the number of promotees divided by the unit's average monthly cases per detective in the pre-period.

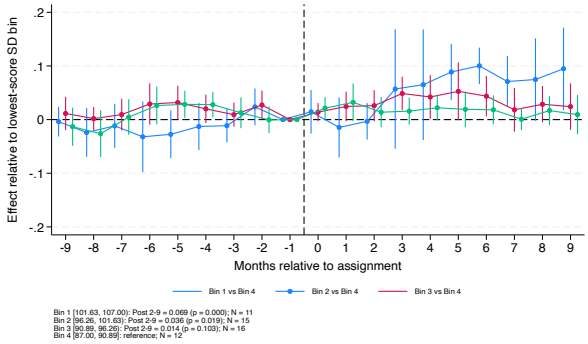
Figure A6: Comparison Units and Calendar-Time Controls



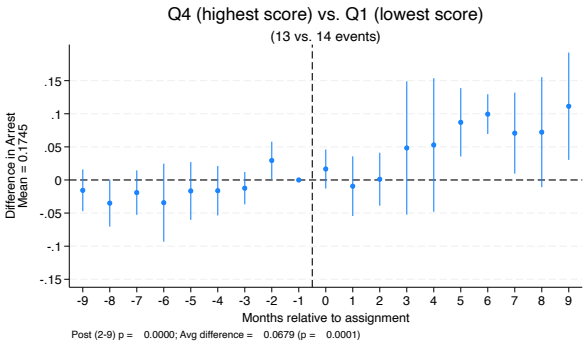
Notes: This figure compares the main event-study estimates for arrests with estimates from a stacked specification that adds control units. For each promotion event, control units are investigative units that do not receive a promotion during the event window. The "With control units" specification includes event-by-unit, event-by-calendar-month, and event-by-crime-type fixed effects, so estimates net out calendar-time shocks common to units within each event. It also includes event-time indicators for promotion, with promotion scores centered on the mean promotee score, so that the plotted coefficients measure the effect of an additional score point relative to a promotion of average quality. Both series plot event-time coefficients per point of incoming detectives' average promotion score. The horizontal axis is months relative to promotion, omitting month −1. Standard errors are clustered at the investigative-unit level. The plot note reports the average effect over months 2 to 9 for each specification and the p-value for the test that the two averages are equal.

Figure A7: Nonlinear Specifications of Promotion-Score Effects

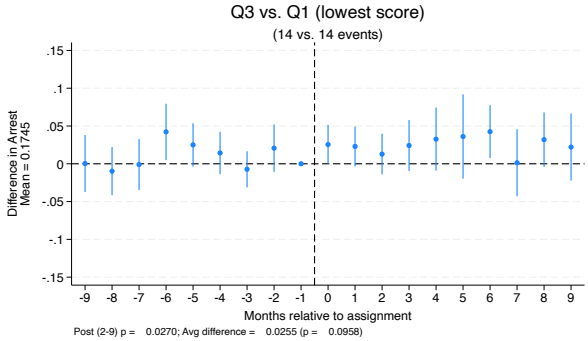
(a) One-standard-deviation score bins



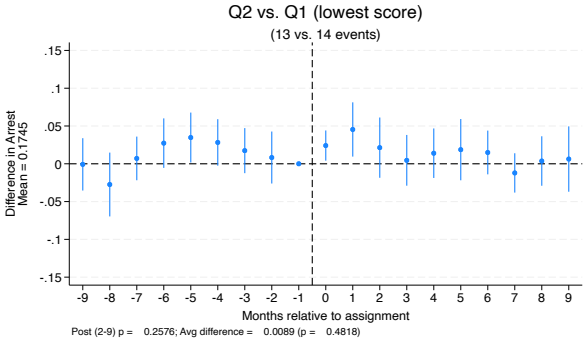
(b) Fourth versus first quartile



(c) Third versus first quartile



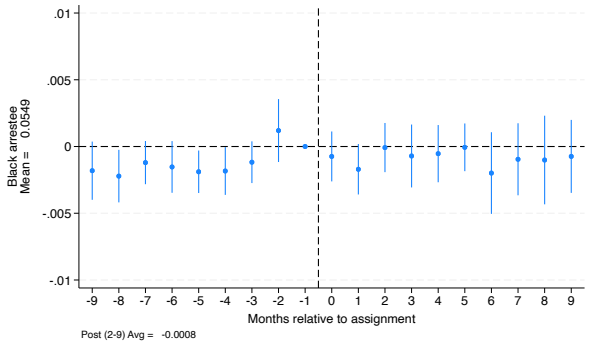
(d) Second versus first quartile



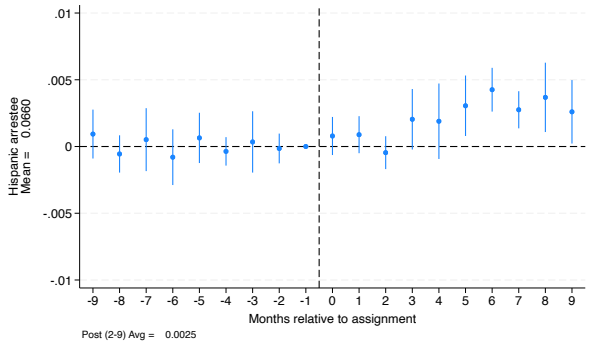
Notes: This figure relaxes the linear specification of promotion-score effects. Panel (a) groups promotion events into bins one standard deviation wide, measured from the highest observed average promotion score, with the standard deviation computed across events in the analysis sample. The specification replaces the score interaction in the main specification with event-time indicators for each bin, so each series plots the effect of promotions in that bin relative to promotions in the lowest-score bin. Panels (b)–(d) use the specification with control units described in Figure A6, replacing the score interaction with event-time indicators for each quartile of promotee score. Each panel plots the difference between the indicated quartile's effect and the first quartile's effect. All panels plot coefficients by months relative to promotion, omitting month −1, with 95 percent confidence intervals based on standard errors clustered at the investigative-unit level.

Figure A8: Arrest Effects by Arrestee Race and Ethnicity

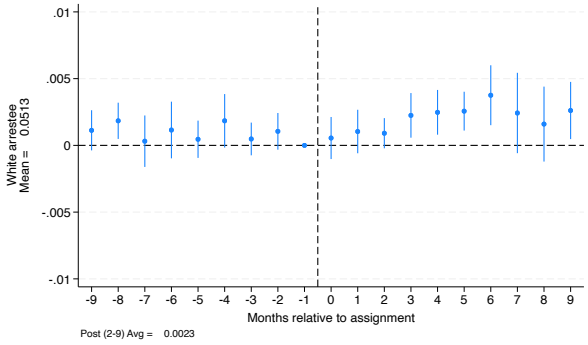
(a) Black arrestees



(b) Hispanic arrestees



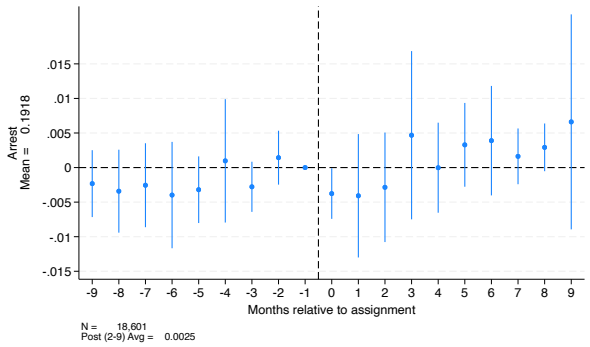
(c) White arrestees



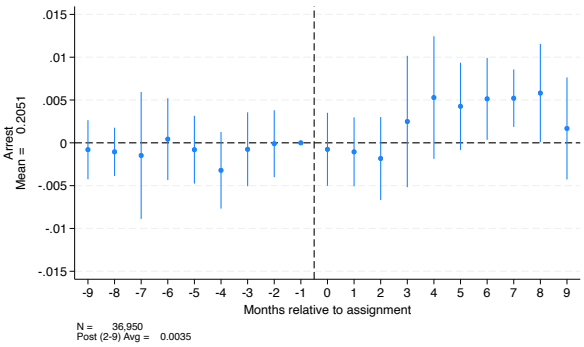
Notes: Each panel plots event-time coefficients for an additional point of incoming detectives' average promotion score. The outcome in each panel equals one if the case results in an arrest and the arrestee is recorded as Black (panel a), Hispanic (panel b), or White (panel c), and zero otherwise, including for cases without an arrest. The three outcomes therefore decompose the overall arrest effect in Figure 3 by arrestee race and ethnicity. Arrestee race and ethnicity come from APD arrest records. The horizontal axis is months relative to promotion, omitting month −1, and the vertical axis is the change in the probability of the indicated outcome. Regressions use the analysis sample and include promotion-event, relative-month, and crime-type fixed effects. Standard errors are clustered at the investigative-unit level, and bars show 95 percent confidence intervals.

Figure A9: Arrest Effects by Victim Race and Ethnicity

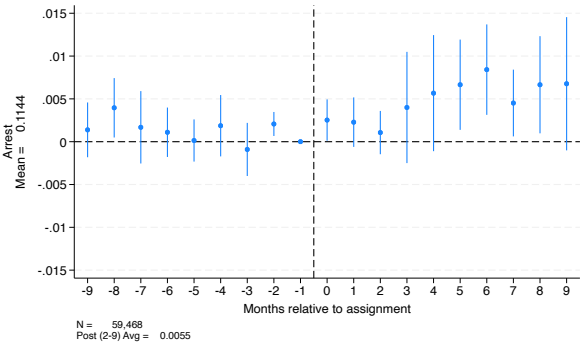
(a) Black victims



(b) Hispanic victims



(c) White victims



Notes: Each panel plots event-time coefficients for an additional point of incoming detectives' average promotion score, estimated separately on cases with at least one victim recorded as Black (panel a), Hispanic (panel b), or White (panel c). Cases with victims in more than one group appear in each corresponding panel, and cases with no recorded victim are excluded. The outcome is an indicator for whether the case results in an arrest. Victim race and ethnicity come from APD person records. The horizontal axis is months relative to promotion, omitting month −1, and the vertical axis is the change in arrest probability. Each regression uses the analysis sample and includes promotion-event, relative-month, and crime-type fixed effects. Standard errors are clustered at the investigative-unit level, and bars show 95 percent confidence intervals.

Table A1: Construction of the Event-Study Sample

Sample construction step	Events	Observations
Panel A. Promotion events		
Detective promotions, 2010–2020	173	–
Combine same-unit, same-month promotions into one event	143	–
Drop events with another unit promotion in the pre-period	76	–
Restrict to events in units with cases in the event window	74	–
Panel B. Case records		
All incident and case records, 2010–2020	–	798,709
Retain cases in event units within −9 to +9 months of promotion	74	193,200
Truncate post-period when another unit promotion occurs	74	169,218
Panel C. Event restrictions based on case data		
Drop events with months lacking any cases	56	160,480
Drop events with fewer than 150 cases	54	160,248
Final event-study sample	54	160,248

Notes: This table reports the number of promotion events and case observations remaining after each sample restriction. Panel A uses the promotion data. Its first row counts individual detective promotions, and later rows count promotion events, defined as investigative-unit-by-month cells with at least one promotion. The pre-period restriction drops events whose unit had another promotion in the 9 months before the event. Panel B uses the case data, restricted to cases with a NIBRS offense code. Observations are case-by-event pairs. The truncation step shortens an event's post-period when its unit has another promotion within 9 months, removing cases assigned after the month of that next promotion. Panel C drops events with at least one month in the event window without any cases, and events with fewer than 150 case observations. The final sample is used in all event-study estimates unless otherwise noted.

Table A2: Promotion Events by Investigative Unit

Unit	Events	Unit	Events
Child Abuse	7	Cyber Crime	1
Auto Theft	4	Domestic Violence 2	1
Burglary	4	Hit & Run	1
Human Trafficking	4	Mental Health Unit	1
Robbery	4	Narcotics 1	1
Sex Crimes 1	4	Commercial Burglaries	1
Domestic Violence 1	3	Regional Detectives 1	1
Narcotics 3	3	Regional Detectives 2	1
Agg. Assault 1	3	Regional Detectives 3	1
Financial Crimes	2	Regional Detectives 4	1
Narcotics 2	2	Narcotics 4	1
Agg. Assault 2	1	Sex Crimes 2	1
		Traffic Crimes	1

Notes: This table reports the number of promotion events in each investigative unit in the final event-study sample of 54 events described in Table A1. A promotion event is an investigative-unit-by-month cell with at least one detective promotion. Numbered units (e.g., Narcotics 1 through 4) are separate units in APD's case data.

Table A3: Pre-Promotion Balance Before and After Matching

	Before matching			After matching		
Characteristic	Mean	Difference	t-stat	Mean	Difference	t-stat
Violent	0.1555	0.1218	1.57	0.1747	0.0987	0.83
Property	0.2272	-0.0024	-0.02	0.1347	-0.0501	-0.40
911-Originated	0.7388	0.0861	1.25	0.8276	0.0022	0.03
Minority population share	0.6006	0.0249	1.35	0.6214	-0.0109	-0.50
Log income per capita	10.2530	-0.0039	-0.08	10.1977	0.0145	0.29
Unemployed population share	0.0558	0.0017	1.33	0.0570	-0.0012	-0.86
Minority officer	0.4801	-0.0221	-0.39	0.5304	-0.0197	-0.27
Female officer	0.1700	-0.0083	-0.24	0.1827	0.0051	0.11
Overnight offense	0.2264	0.0101	0.38	0.2256	0.0087	0.22
Number of unit detectives	13.0556	-0.1111	-0.08	14.0000	-0.3077	-0.17
Number of cases	1544.2037	-334.1852	-0.71	1490.9615	-368.3846	-0.55
Number of events		54			26	

Notes: This table compares pre-promotion characteristics of promotion events with above-median and below-median average promotion scores, before and after propensity-score matching. Characteristics are measured over the pre-promotion months of each event's window. Case characteristics are shares of the unit's cases, neighborhood characteristics are averaged across the unit's cases, and officer characteristics are the shares of cases led by minority and female detectives. The number of unit detectives counts distinct lead detectives, and the number of cases counts distinct cases, over the full pre-period. Mean columns report the mean across all events in each sample. Difference columns report the above-minus-below-median difference, with t-statistics based on heteroskedasticity-robust standard errors. Matching uses a logit propensity score estimated on all listed characteristics and pairs each above-median event with its nearest below-median neighbor, without replacement, on the common support and within a caliper of 0.05.

A Propensity Score Matching

Continuous-treatment difference-in-differences (DID) designs require that units receiving stronger treatments would have experienced the same treatment effects as units receiving weaker treatments had they been exposed to the same treatment dosage (Callaway et al., 2025). If this assumption fails, estimates may reflect selection into treatment intensity.

We assess sensitivity to this concern by reestimating the continuous DID on a propensity-score-matched sample of promotion events, following Chen et al. (2026). We divide the 54 events into groups above and below the median promotee quality. We then estimate a logit model of receiving an above-median-quality promotee using pre-promotion case composition, neighborhood demographics, detective demographics, unit staffing, and caseloads. Each high-quality event is matched without replacement to the nearest low-quality event on the common support, subject to a propensity-score difference of no more than 0.05. The matched sample contains 26 events: 13 high-quality and 13 low-quality. We reestimate the continuous DID in this sample using the full measure of promotee quality.

Table A3 compares pre-promotion characteristics before and after matching. The groups are relatively balanced before matching, and matching reduces the remaining observable differences. Figure A2 shows that the matched-sample estimates are similar to the main estimates. These findings suggest that sorting on the observed matching characteristics does not explain the main results, although they cannot rule out sorting on unobserved characteristics.