

Misconduct, Metrics, and Motivation: Incentive-Driven Changes in Policing Behavior

Yuki Otsu^{*†}

The University of Tokyo

December 18, 2025

Abstract

Because managers are responsible for running their organizations in a sound manner, they may be held accountable for supervisory failures when subordinates commit misconduct, which can in turn hinder their chances of promotion. This study reveals that such promotion incentives for public officials can distort organizational performance. Using data from the Japanese police, the analysis investigates how misconduct committed by individual officers affects station-level performance indicators, such as the number of arrests and the clearance rate. The results show that when a disciplinary action within a police station is publicly announced, the police station's performance improves during that year. This improvement is not driven by changes in the behavior of the disciplined officers or citizens, but rather by the promotion incentives of the police station chief.

JEL classification: D73, H83, J45, K42, M54

Keywords: misconduct, disciplinary actions, promotion, trust, policing

^{*}Center for Spatial Information Science, The University of Tokyo. 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8568 Japan. Email: y.otsu@csis.u-tokyo.ac.jp

[†]I am grateful to Keisuke Kawata and Takuma Kamada for their helpful comments and suggestions, and Yurina Anzai for her work as a research assistant. I would also like to thank the participants of ARSC, the Japanese Public Choice Society Meeting, the WEAI Annual Meeting, and a seminar at Osaka University and Tohoku University for their discussions. This research is supported by JSPS KAKENHI Grant Number 23K12488. The author declares that he has no relevant material or financial interests that relate to the research described in this paper. Earlier versions of this paper were circulated under the title "Does police misconduct increase crime?"

1 Introduction

Police organizations operate under intense public scrutiny, in which even minor misconduct can pose significant reputational threats. These incidents, while often not legally serious, may nonetheless damage public perception and institutional credibility. In such moments, police departments must decide how to respond—not just in terms of internal discipline, but also in their broader operational behavior. This study investigates how police organizations respond to routine misconduct in a strategic manner, focusing on performance-oriented behaviors driven by internal incentive structures.

While previous research emphasizes how public trust in the police can influence reporting, cooperation, and crime rates, this paper argues that organizational responses to misconduct may be more deeply rooted in internal career incentives, particularly the promotion prospects of police leaders. Misconduct can undermine the station chief’s perceived managerial competence, thereby reducing their chances of promotion. In turn, this creates a strong incentive to improve performance indicators quickly and visibly, such as increasing cases cleared or clearance rates. Alternatively, police organizations might act to restore public trust by improving engagement. This study assesses the impact of misconduct on police organizations’ activities and the relative strength of these two mechanisms (internal incentives and public trust) as causes.

Japan provides a unique empirical context for examining these dynamics. Unlike the U.S., where police research is more prevalent, Japanese police officers typically resign from their positions when they are disciplined, especially if the case is publicly reported, even if it is very minor. This institutional norm allows the researcher to separate the organizational consequences of misconduct from the continued existence and behavioral changes of the sanctioned individual. Thus, this analysis can capture how the broader police organization, not simply individual police officers, adjusts its behavior after a misconduct is publicized.

Using annual data from 96 police stations under the Tokyo Metropolitan Police Department (TMPD) between 2013 and 2023, this paper employs a staggered difference-

in-differences design, applying the estimator proposed by de Chaisemartin and D’Haultfœuille (2024) to account for dynamic treatment effects. The timing of disciplinary actions reported in major newspapers is treated as a quasi-random shock to each police station. The analysis reveals that, in the year misconduct is reported, theft-related cases cleared, especially for burglary, significantly increase, while the number of crimes recognized by police, particularly non-burglary theft, declines. Together, these shifts temporarily improve clearance rates. However, these effects are highly concentrated on minor offenses and do not persist in the long run.

Importantly, additional evidence shows that these behavioral changes are not driven by changes in citizen reporting or offender behavior, but rather by internal police decisions. Moreover, misconduct has a measurable negative effect on the promotion likelihood of station chiefs, which can be partially mitigated by short-term improvements in performance. These findings support the conclusion that internal incentives, specifically, promotion-related considerations—are the primary driver of the observed organizational response, more so than changes in public trust.

Prior studies on police misconduct have largely focused on the individual-level consequences of high-profile incidents, such as violent encounters or systemic abuse, often in the context of U.S. cities. In contrast, this paper shifts the analytical lens to police-station-level management and routine infractions, providing new empirical evidence on how promotion-related incentives shape organizational behavior. Rather than being driven by shifts in public trust, the observed changes in policing appear to reflect internal motivations to protect or improve career outcomes. These findings underscore the importance of designing performance metrics that align with substantive improvements in public safety, rather than encouraging superficial adjustments that primarily serve reputational or promotional concerns.

1.1 Literature

This paper contributes to the literature by providing empirical evidence on the impact of routine police misconduct in Japan, a context less explored compared to studies

based in the U.S.

The first strand of literature concerns the impact of disciplinary actions on police behavior, particularly at the individual level. Because disciplinary actions are usually taken after receiving complaints from citizens, prior research has focused on these complaints (Rozema and Schanzenbach, 2019; Jordan and Kim, 2025; Prendergast, 2001, 2021). Rozema and Schanzenbach (2019) used data from Chicago to analyze how officers' behavior changes after being accused by a citizen. What they found is that investigations and disciplinary actions lead to sustained reductions in misconduct by the same individuals. Focusing on individuals, this paper suggests that disciplinary actions may reduce misconduct. Jordan and Kim (2025) analyzed the impact of the assignment of complaint investigators on police behavior. They found that high-investigation supervisors cause investigated police officers to make more arrests and receive fewer complaints at normal times. Unlike previous studies, this study analyzes the impact of disciplinary action on the organization to which the individual belongs. In particular, the analysis is unique in that it uses Japanese data, where most disciplined officers resign their positions,¹ allows us to exclude the portion of the impact of disciplinary actions on individuals and then analyze the impact on the organization.

This study also contributes to the literature on the effects of incentives for public officials. Dewatripont et al. (1999) applied a career concerns model developed by Holmström (1999) to public servants' incentives and demonstrated that incentives to exert effort are stronger when agents are assigned narrow and clear tasks. Regarding police, the theoretical framework provided by Adamson and Rentschler (2025) describes the trade-off between crime detection and investigation, positing that stricter penalties for false arrests incentivize more careful investigations at the expense of detection efforts. Although measuring public sector outcomes presents inherent challenges (Tirole, 1994), recent studies such as Jia (2024) have examined the effects of promotion-based incentives in public sector. Kim (2022), using data from the Chicago Police Department, analyzed how promotion incentives affect the performance of individual police

¹See Appendix A for more details.

officers. Similarly, Baicker and Jacobson (2007) analyzed how financial incentives affect both local governments and law enforcement agencies. In relation to data manipulation, Xu et al. (2025) investigated the link between promotion incentives for local bureaucrats in China and the manipulation of GDP figures. Building on this line of research, the present study analyzes how promotion incentives for public officials affect the performance of police organizations, and examines the potential for data manipulation.

The third strand of literature examines the impact of media coverage on police behavior. There is substantial evidence that media coverage can influence police actions (Mastrorocco and Ornaghi, 2025; Rackstraw, 2023). For example, Mastrorocco and Ornaghi (2025) show that a decrease in crime reporting due to the takeover of local television stations led to a decrease in violent crime clearance rates. Rackstraw (2023) concludes that while arrests increase during the filming of TV shows that closely follow the police, the increase is concentrated on minor crimes that would not normally result in arrests, and does not help to substantially improve public safety. In this paper, too, the increase in cases cleared is concentrated on petty crimes such as theft, and the breakdown of the increase comes from an increase in cases cleared for burglaries and non-burglary thefts, suggesting that, as in the previous study, the potential to improve public safety is limited.

The impact of media coverage on someone may be due to being monitored; Conover et al. (2023) shows that the existence of a CCTV camera increases issuing traffic violations and improves the performance of the police. Their results suggest that it may be enough if someone sees police activity without TV media coverage. Unlike these studies, this paper focuses on misconduct as the trigger of public scrutiny, and highlights how media-induced pressure operates through internal performance incentives.

The fourth strand of literature to which this paper contributes is the effect of public trust in police on their activity. The significance of trust has been extensively studied, as highlighted by Jackson and Bradford (2010). In the US, there are studies such as the Ferguson Effect and the Minneapolis Effect, which examine the effects on police activi-

ties and others after an incident occurs that shakes trust in police (Mikdash and Zaiour, 2022; Cheng and Long, 2022; Campbell, 2024; Ang et al., 2025). Although high-profile incidents show a dramatic decrease in police activity (Devi and Fryer, 2020), they also reveal that not all police violence causes similar results. However, these studies often rely on high-profile incidents and primarily U.S. data, leaving open the question of how police respond to routine misconduct attracting limited public attention. This study suggests that, in such routine incidents, organizational behavior is driven more by internal incentive structures than by changes in public trust.

In summary, this study contributes to several strands of literature by providing empirical evidence that promotion-related incentives, rather than public trust, are the primary driver of behavioral responses to routine misconduct. By focusing on a low-profile, high-frequency setting in Japan, it offers a novel perspective on how performance metrics interact with internal accountability mechanisms in public-sector organizations.

2 Potential organizational responses

Why could disciplinary actions related to police misconduct be associated with crime rates, especially officially recorded crimes? Previous studies on the impact of minor misconduct have used data from the U.S. to analyze the impact on the individual who committed the misconduct, but few have analyzed the impact on the organization to which he or she belongs (Owens and Ba, 2021; Mastrorocco and Ornaghi, 2025; Ang et al., 2025). However, the Japanese police, the focus of this study, rarely continue to be in office after the person who caused the reported level of misconduct. Therefore, by analyzing the impact of scandals in Japan, it is possible to examine the impact on the organization to which the person belongs, excluding the impact on the person himself or herself.

Why, then, would there be an impact even though the person who committed the misconduct is not present? Based on previous studies, the following three factors

may be relevant: promotional incentives, citizen trust, and political feedback (Owens and Ba, 2021; Mastrorocco and Ornaghi, 2025; Ang et al., 2025). These factors may provide incentives for the organization to improve its performance (i.e., increasing arrest counts, raising the clearance rate, and reducing offenses known to the police).

2.1 Promotional incentives

First, instances of misconduct are likely to affect the prospects for promotion of superiors, especially the chiefs of the police stations in which the incidents occur. In the Japanese police system, police chiefs are held organizationally responsible for incidents under their supervision, even when they are not personally involved. Public disclosure of misconduct by a subordinate can harm the chief’s reputation, as it signals a lapse in oversight. This reputational risk creates strong incentives for chiefs to actively restore performance metrics in the aftermath of a scandal. If performance within their jurisdiction significantly improves, the chief is likely to be regarded as a competent leader, thereby increasing their chances of promotion. Conversely, if officers under their supervision are involved in misconduct, the chief may be held accountable for insufficient oversight. In rare but serious cases, such as corruption or abuse of power, the chief may be directly subject to disciplinary action. More commonly, however, no explicit sanction is imposed. Nevertheless, such incidents can still cast a shadow over future promotions, creating a strong incentive for chiefs to both improve policing performance and prevent further misconduct.

To better understand how internal incentives operate in the wake of a misconduct shock, I present a stylized one-period career concerns model based on Holmström (1999). Consider a simplified one-period model in which an agent (e.g., a police chief) of unobservable management ability $\theta \in \mathbb{R}$ chooses an effort level $a \in \mathbb{R}_+$ to influence an observable performance outcome $y \in \mathbb{R}$ (e.g., clearance rate). Performance is generated as:

$$y = \theta a + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2),$$

where ε is a noise term. The agent derives utility from the principal's posterior expectation (i.e. the police headquarter's posterior) of θ , which determines future career rewards, net of a convex effort cost:

$$U(a) = \mathbb{E}_y[\theta | y] - \frac{1}{2}ca^2.$$

Following Dewatripont et al. (1999), the principal observes y and updates beliefs about θ using Bayesian inference. Suppose the prior distribution of θ is normal: $\theta \sim \mathcal{N}(\mu, \tau^2)$. Then the posterior expectation is:

$$\mathbb{E}[\theta | y] = \omega \cdot \frac{y}{a} + (1 - \omega)\mu, \quad \text{where } \omega = \frac{\tau^2 a^2}{\tau^2 a^2 + \sigma^2}.$$

The agent's optimal effort (a^*) solves:

$$a^* = \underset{a}{\operatorname{argmax}} \left\{ \omega(a) \cdot \mu a + (1 - \omega(a)) \cdot \mu - \frac{1}{2}ca^2 \right\}.$$

The agent's unobservable ability (θ) reflects the perceived managerial competence of a police station chief, as assessed by the upper ranks of TMPD during promotion decisions. The effort level (a) translates into actions that improve measurable outcomes such as cases cleared or the clearance rate. These indicators are regularly monitored and often cited in internal evaluations. The performance outcome (y), therefore, can be interpreted as the realized crime-solving statistics attributed to the organization under the chief's command.

A misconduct shock can be modeled as a downward revision of the prior mean $\mu' < \mu$, reflecting a reputational loss following a public scandal. The agent's optimal effort response to such a shock depends critically not on whether rewards scale directly with μ , but rather on how sensitive the principal's posterior belief about ability is to observable performance. In particular, when the marginal reputational return to effort, captured by the slope of the mapping from performance to perceived ability, increases as μ declines, agents may have strong incentives to engage in effortful be-

havior aimed at reputational recovery, despite lower prior expectations. Thus, under appropriate informational conditions, a decline in μ may lead to greater effort a^{*2} , not due to a larger expected reward per se, but because the marginal benefit of influencing beliefs increases. While not all career concerns models assign primacy to dynamic reputational feedback, in contexts where posterior evaluations are highly sensitive to performance, such feedback can amplify incentive strength more than static baseline assessments.

2.2 Public trust in the police

Another explanation is the change in public trust in the police (Ang et al., 2025). When disciplinary actions are taken, they may signal the possibility that the police are unreliable, potentially leading to a loss of trust. This decline in trust could affect both residents and the police in the following ways.

First, residents who lose trust in police may be less likely to report crimes or cooperate in investigations. This may result in fewer crimes being recognized by the police, and consequently, a reduction in cases cleared, which in turn may lead to a worsening of the clearance rate. On the other hand, the police may respond by conducting more aggressive investigations to restore public trust (Adamson and Rentschler, 2023). This could lead to an increase in crimes officially recorded if more crimes are uncovered through these efforts and an increase in cases cleared. Although loss of trust could reduce cases cleared due to less public cooperation, it could also lead to an increase if police adopt more aggressive tactics. Therefore, whether arrest rates increase or decrease depends on the relative strength of these opposing factors, and the overall impact of disciplinary actions on crime rates through changes in citizens' trust remains uncertain.

²See Appendix E.

2.3 Political feedback

Lastly, in this context, political feedback may be relevant. Political feedback refers to political pressure on politicians and police leadership. In particular, in the United States, police departments are typically operated at the municipal level and mayors usually appoint the heads of these departments. As a result, police misconduct could potentially raise accountability issues for the appointing mayor, making the impact of political pressure significant and difficult to ignore. In fact, Ornaghi (2019) showed the possibility that using a merit system to reduce pressure from politicians improved police performance.

However, the organizational structure of the Japanese police, the focus of this study, differs significantly from that of the United States. The heads of these prefectural police departments, such as the Superintendent General or the prefectural police chiefs, are appointed by the National Police Agency, a national-level organization, and serve as national civil servants. Consequently, although prefectural governors hold a nominal supervisory role, their actual authority over police departments is extremely limited. Therefore, in the context of Japan, the influence of political feedback is considered to be weak, at least at the prefectural level.

For these reasons, this paper analyzes the impact of misconduct on overall organizational performance through the police chief's promotion incentives and changes in trust in the police.

While misconduct might generate multiple organizational responses, this study focuses on two primary mechanisms: (1) internal incentives, where station chiefs respond to protect or restore their chances of promotion by improving performance metrics; and (2) external trust-based dynamics, where perceived legitimacy shapes both citizen cooperation and police initiative. Accordingly, the following analysis concentrates on the two mechanisms most relevant in the Japanese context. The empirical analysis that follows evaluates the relative strength and timing of these channels.

3 Empirical strategy

3.1 Data and background

This study focuses on the Tokyo Metropolitan Police Department (TMPD), Japan's largest police force, which oversees the entire Tokyo prefecture.³ According to Tokyo Metropolitan Police Department (2024), as of 2024, TMPD had 46,581 employees, of which 43,566 are police officers. This size is comparable to the New York Police Department 46,804 (including 33,307 police officers) and larger than Chicago Police Department 12,370 (11,667) and the Los Angeles Police Department 11,435 (8,820).

The crime data utilized for this analysis are derived from the TMPD's annual records on recognized crimes and cases cleared spanning from 2013 to 2023. The TMPD's jurisdiction includes 102 police stations, of which data from 96 stations (excluding island and airport police stations) were selected for this study. The summary statistics are in Table 1. As for offenses known to the police, non-burglary thefts occupy the majority of the crime total, and more than half of that is vehicle thefts. As for cases cleared, while non-burglary is the most popular type of crime, vehicle thefts account for less than 10 % of the thefts.

In this paper, officially issued disciplinary actions are treated as indicators of police misconduct. I constructed records of disciplinary actions using data collected from newspaper articles. Misconduct reports were extracted from the databases of four major newspaper companies: Asahi, Mainichi, Nikkei, and Yomiuri. Disciplinary action is defined at the police station level and is considered to have occurred if at least one of these newspapers reports it.⁴ The analysis focuses on disciplinary actions taken against officers serving in police stations. Since local police stations are directly involved with residents, disciplinary cases involving personnel working in the head-

³See National Police Agency (2021) for more details about the institutional background.

⁴Note that not all misconduct is published as newspaper articles. To assess the extent of this discrepancy, I obtained the official reports on the disciplinary actions taken in 2022 through a request for disclosure of information. In 2022, 31 officers were disciplined, but only eight were reported in newspaper articles. Although relatively minor misconduct is typically not disclosed, in most cases, TMPD appears to follow official guidelines on what should be made public. The difference between reported and unreported disciplinary actions is further investigated in Appendix A.

quarters or the area headquarters that oversee these police stations are excluded from the analysis.

3.1.1 Geographic variation

Figure 1 illustrates the geographic variation and frequency of reported disciplinary actions between 2013 and 2023. Approximately half (52 out of 96) of the police stations have experienced at least one disciplinary action in 11 years, with a maximum of four of those years (in the Shinjuku police station). Figure 1 shows that the police stations where disciplinary actions were taken are scattered throughout the metropolitan area, without significant geographic bias. Figure 2 plots the variation in timing. During the entire sample period, 2 to 15 police stations report misconduct in a year, and on average, 7.4 police stations in one year report misconduct.

3.1.2 Report timing

This study treats the official announcement of disciplinary actions as an intervention. However, within police departments, it is plausible that incidents are recognized internally prior to their public disclosure. Consequently, there exists the possibility that the timing of such announcements is subject to deliberate control, or that certain incidents may not be disclosed at all. To assess the potential endogeneity of announcement timing, this section examines the relationship between disciplinary cases and the dates on which the underlying incidents occurred. The analysis reveals that, even when considering the longest observation windows, the vast majority of announcements occur within six months of the incident date.⁵ This finding suggests that conducting analyses at the annual level does not pose substantial issues in terms of temporal alignment.

3.1.3 Disciplinary action

Depending on the severity of the disciplinary case, there are several types of disciplinary actions that can be taken against officers. The four types of disciplinary action

⁵See Appendix A for details.

are dismissal, suspension from work, reduction in pay, and reprimand. However, not all of them are subject to media coverage. According to the TMPD guidelines, in principle, public announcements are to be made in the following three cases: (1) disciplinary action for conduct related to the performance of one's duties, or (2) disciplinary action for private conduct, in which the employee is dismissed or suspended (except in cases where there is a need to protect the privacy of those involved, such as sexual offenses), or (3) other disciplinary actions for which it is deemed appropriate to make an announcement in order to ensure public confidence. The guideline also states that, in principle, information is to be announced immediately after the disciplinary action is taken. Mostly, TMPD seemed to follow the guideline. In 2022, all dismissals and suspensions became public except one.⁶

The reasons for disciplinary actions can be broadly divided into two categories: those related to official duties and those unrelated. In Table 2, among the reported disciplinary cases, only about one-third are related to police duties. These cases mostly involve information leaks, concealment of suspects, and similar misconduct. More than 60% of all disciplinary actions concern off-duty conduct, and the majority of these involve legal violations such as sexual offenses, theft, traffic accidents, and traffic violations. These types of cases account for more than half of all disciplinary actions.

Another notable feature of the disciplinary outcomes is that one-third of the officers disciplined were dismissed, and most of the remaining officers resigned voluntarily (Table 3). Consequently, more than 92% of the officers involved in misconduct left their positions immediately after the disciplinary actions were taken.

3.2 Approach

A simple approach to examine the effect of police misconduct is a two-way fixed-effects approach (TWFE) and specifically the following equation can be considered.

$$y_{it} = \sum_{j=-7, j \neq -1}^8 \beta_j \times D_{j(it)} + \delta X_{it} + \eta_i + \zeta_t + \epsilon_{it}.$$

⁶See Appendix A.

$D_{j(it)}$ takes the value one if the officer in the i police station in year t has misconduct j years ahead, and 0 otherwise. Some police stations experience multiple misconduct during the sample period. I treat those observations as follows: the reported year is set as the event date. The lead indicators are relative to the first year the misconduct case is reported in the sample period. The lagged indicators are relative to the most recent report. The outcome variable y_{it} is the number of reported crimes or the number of cases cleared per square kilometer.⁷ η_i and ζ_t are the police station and year fixed effects.

However, estimators based on this approach can be severely biased when there is variation in treatment timing and dynamic treatment effects (Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2020; de Chaisemartin and D'Haultfœuille, 2024; Goodman-Bacon, 2021; Sun and Abraham, 2021). Given that misconduct reports exhibit this staggered nature and are likely to have heterogeneous effects over time, it is important to use methods that can mitigate this potential bias. Since some police stations report misconduct multiple times during the sample period, the method proposed by de Chaisemartin and D'Haultfœuille (2024) (hereafter dCDH) is applied to estimate what they call the actual-versus-status-quo (AVSQ) effect. To estimate the AVSQ effect of treatment for a group g at l period later from the treatment, the dCDH estimator ($DID_{g,l}$) is described as follows:

$$DID_{g,l} = Y_{g,F_g-1+l} - Y_{g,F_g-1} - \frac{1}{N_{F_g-1+l}^g} \sum_{g': D_{g',1}=D_{g,1}, F_{g'} > F_g-1+l} (Y_{g',F_g-1+l} - Y_{g',F_g-1}).$$

where $Y_{g,t}$ is an outcome of the group g in period t , F_g is the first period that the group g is treated, and N_t^g is the number of groups with the same period-one treatment as g and that have maintained the same treatment from period 1 to t . The DID estimator compares $F_g - 1$ to $F_g - 1 + l$ outcome evolution of g to groups with same baseline treatment, and that have kept that treatment from period 1 to $F_g - 1 + l$.⁸

⁷Appendix D presents the results based on a per capita measure.

⁸The sample size used for DID estimation in each period varies from period to period. See Appendix B for the sample size.

4 The impact of misconduct on police performance

This section presents the core empirical findings of this study. Building on the empirical strategy outlined in Section 3, it investigates how police stations in Tokyo change their operational behavior following publicized disciplinary actions, focusing on theft-related indicators that dominate overall crime statistics.⁹ While both promotion incentives and changes in public trust could theoretically influence police behavior, this section does not attempt to disentangle these mechanisms. Instead, it focuses on documenting the patterns of behavioral change, with a detailed analysis of underlying motivations deferred to Section 5.

Figure 3 shows the dCDH estimators for cases cleared. Panel A shows that the number of thefts increases when disciplinary action is taken. In the year when the police station is treated, cases cleared increase by 5 per square km. In addition, the estimates for the periods prior to disciplinary action are not statistically significant at the 5% level and no deviations from the parallel trend assumption are identified.

Keeping other conditions constant, an increase in cases cleared is generally viewed positively as positive police activity. Thus, this result indicates that misconduct, whether motivated by promotion incentives or the restoration of trust, results in positive activity. Notably, theft accounts for the largest proportion of all crimes, making changes in cases cleared for theft a significant contributor to overall changes in numbers. Therefore, especially following media coverage of disciplinary actions, the police may allocate more resources to clearing theft offenders rather than focusing on violent offenders. This shift in resource allocation could explain the observed increase in overall cases cleared.

To check what drives the increase in theft, Panels B and C in Figure 3 decompose the results by burglary and non-burglary theft. Panel B shows an increase in burglary theft in the treatment year, while Panel C shows no change in non-burglary theft before and after treatment. One possible reason why cases cleared have increased only

⁹See the Appendix C for the impact on the other major categories (felony, violent offenses, intellectual offenses, and moral offenses).

for non-burglary thefts and not for burglary thefts can be attributed to the relative ease of clearing. Burglary thefts account for only 5.6% of reported theft cases but constitute 21.9% of all theft cases cleared. By allocating resources to the easier-to-clear burglary thefts rather than the more numerous but harder-to-solve non-burglary thefts, the police are likely able to increase the overall number of cases cleared.

Furthermore, with respect to burglary theft, only year zero shows a statistically significant increase, with little change in the other years. The reason why this increase in the cases cleared is observed only temporarily and not over the long term may lie in the relatively short tenure of police chiefs, which typically lasts less than two years. This result supports the rationale of promotion incentives for police chiefs. The average tenure of chiefs is less than two years, suggesting that they may be attempting to compensate for the misconduct that occurs during their tenure by increasing cases cleared. This point will be examined in detail in Section 5.

Figure 4 presents the dCDH estimators for the number of individuals cleared for offenses, which measures the impact of disciplinary actions on the number of persons cleared for offenses before and after the police misconduct. Although the point estimates show some negative effects on overall theft and non-burglary theft at the treated year, the estimated effects are not statistically significant. Thus, while it cannot be said that the number of persons cleared is decreasing, at least, the number is not increasing.

An increase in cases cleared and a decrease (or no change) in the number of persons cleared imply an increase in cases cleared per person cleared. Therefore, it can be interpreted that police focus more on investigating specific offenders than increasing the number of persons cleared. In addition, since cases cleared have increased, but the number of persons cleared has not, it is possible that the police consider cases cleared more important than the number of persons cleared as an indicator. In either case, increasing cases cleared requires the investment of many resources, such as securing space in the detention centers where suspects are held. Therefore, it is possible to increase cases cleared with fewer resources by pursuing and clearing the additional crimes of those arrested or by arresting criminals who are likely to have committed

more crimes.

Figure 5 shows the dCDH estimators for offenses known to the police. The point estimates for theft cases are consistently negative after treatment, with the largest decline observed in the treatment year. While the effect is not statistically significant at the 5% level, the direction and magnitude suggest a meaningful short-term reduction in recognized theft offenses.

Panels B and C break down the effects into burglary and non-burglary thefts. Since non-burglary theft accounts for the majority of theft cognitions, the results for non-burglary theft are very similar to the overall theft trend, with a significant decrease in the treated year. While neither burglary nor non-burglary thefts show a statistically significant change in the treatment year, burglary thefts exhibit statistically significant declines in the number of offenses known to the police in the first and second years after the disciplinary actions. These findings suggest that police reporting or classification practices may have been affected in the short to medium term after the public disclosure of misconduct. In particular, to increase cases cleared, the police must first recognize offenses that had not yet been officially recorded, as a crime must be registered before it can be cleared. While the effects for non-burglary thefts are not statistically significant, the point estimates show a visible decrease in the treatment year, implying the possibility of discretionary adjustments by the police. Overall, these results indicate that the observed reduction in recorded thefts is likely driven more by internal organizational behavior than by changes in actual crime incidence or citizen reporting.

There are two possible reasons for changes in the number of reported crimes: one is a change in the actual number of crimes occurring, and the other is a change in the proportion of crimes that are recognized. However, incidents of police misconduct or their reporting are unlikely to decrease actual crime rates, even if they could lead to an increase. Therefore, it is reasonable to conclude that a decrease in the proportion of recognized crimes plays a significant role.

Reasons for the decline in the proportion of recognized crimes include a decrease in

citizen reports and under-reporting by the police. However, the supplemental analysis in Section 5 reveals no significant changes in citizen reporting behavior. Additionally, for certain types of crimes—such as non-burglary thefts where cases can be ambiguously classified as thefts or as lost property—the decrease in recognized cases is more likely influenced by police discretion. These crime types are relatively easier for police to manipulate in terms of recognition, suggesting that underreporting by the police has a greater impact.

Combining the two effects on cases cleared and offenses known to the police, Figure 6 shows the effect on the clearance rate (number of cases cleared divided by number of offenses known to the police) for total crime, theft, and non-burglary theft. The clearance rate increases in the first year, but has no effect in the following years. The increase comes from the increase in theft, which shows a statistically significant effect, and the clearance rate for theft increases by 3.0 percentage points. The clearance rate increases only temporarily, due to, as with cases cleared, the incentive of police chiefs to improve performance.

4.1 Spillover effects

The analysis so far has confirmed the temporary effect of the disciplinary action on the police station in question. This section focuses on the spillover effect on surrounding police stations. To confirm the existence of spillover effects, I used the same analytical method developed in Butts (2023) to analyze the effects on adjacent police stations. The method first estimates unit fixed effects and time effects using clean observations and then estimates the treatment effects and spillover effects. Because the treatment effect is assumed to occur only in the treatment year and no dynamic effects are present, all treated units have at least one untreated (clean) observation in the pre-treatment period. Therefore, unit fixed effects are identifiable for every treated unit in the first-stage estimation following Butts (2023).

The results are summarized in Table 4, and, in this estimation model as well, an increase in the number of theft cases cleared was confirmed at the treated police stations.

However, no significant effects were found for the adjacent police stations. These results indicate that the effect of the disciplinary action does not extend to adjacent police stations and is confined to the level of the police station concerned.

5 Mechanism analysis

5.1 Citizen response

The results in the previous section show that scandals may have an impact on the number of reported property crimes. However, the recognition of property crimes often relies on reports from citizens. If citizens lose trust in the police, the number of reported incidents can decrease, leading to a statistical decrease in recognized cases. Therefore, it is difficult to determine whether the change in figures is due to changes in citizens' reporting behavior or shifts in the behavior of the police or offenders. To verify whether scandals influenced citizens' reporting behavior, I analyzed the impact of misconducts on found item reports.

Reporting found property is an individual action that is relatively insulated from the influence of police activities, unlike outcomes stemming from police investigations or public safety measures. Moreover, since it is not the result of a crime, it is not affected by changes in criminal behavior. Therefore, any variation in the number of found property reports can be interpreted as a reflection of changes in citizens' reporting behavior.

Figure 7 analyzes how the occurrence of misconduct affected the filing of found item reports and their content. The results show that the number of found item reports does not appear to have been affected much. The number of reports of found property is not affected in the short term, but may be on a downward trend in the long term. There is also no particular change in the total amount of cash. Rather, there is an increase in the number of items found, which makes it difficult to believe that there has been a decrease in reporting behavior, at least by the general public.

Thus, based on the results using the status of found item reports, it could be in-

ferred that police scandals did not significantly affect actual citizen reporting behavior.

5.2 Career consequences of subordinate misconduct

Does the occurrence of a scandal affect the promotion of a police chief? To examine this mechanism directly, I analyzed whether misconduct affects the promotion outcomes of police station chiefs. Specifically, I conducted a regression analysis in which the outcome variable was whether individuals who served as PS chiefs in TMPD between 2014 and 2023 were promoted within approximately one year after leaving their post.¹⁰

As shown in Table 5, the results indicate that when a subordinate is involved in misconduct, the likelihood of the PS chief being promoted significantly decreases. This suggests that even if the chief is not personally involved, they may still be held accountable from a managerial standpoint, affecting their evaluation.

However, the analysis also reveals that if the station's performance improves more than the annual average for that year, this negative effect of misconduct on promotion likelihood is mitigated. Looking at the magnitude of the coefficients, the positive effect of improving cleared case numbers on promotion is roughly equivalent to the negative effect of misconduct. In other words, strong performance can offset the adverse effect of misconduct on promotion prospects. Furthermore, when the sample is divided by the size of the police station (columns 3 and 4), this pattern appears to be more pronounced among chiefs of smaller police stations. This may be because chiefs of smaller police stations tend to be younger and therefore have higher potential returns from future promotions.

These findings imply that police chiefs have an incentive to improve performance in response to misconduct, aiming to restore their reputation and enhance their chances of promotion. In particular, since the average tenure of a station chief is only 512 days, there is a limited window of time in which to demonstrate results. Therefore, if promotion incentives are the driving force, it is consistent that the performance improvement

¹⁰To measure promotion within approximately one year, the reference period is set to 400 days. This is because transfers of police officers take place twice a year, in spring and fall, and the timing may vary slightly from year to year.

observed in the previous section is concentrated in the same year as the disciplinary action and does not persist in the following year or beyond.

6 Discussion

6.1 Limitation

The first possible issue raised by these results is one of timing. Although the years used in the analysis were the years in which the disciplinary actions were taken, it is difficult to determine whether the timing of the actions preceded the change in cases cleared and reported crimes in the same year, or conversely, whether the change preceded the announcement of the disciplinary actions. In addition, the timing may be endogenous, since the disciplinary action to be taken is known within the police station prior to the press release.

It is difficult to verify from the data used in this study whether timing is endogenous, but even if it is, the impact on the main result would be small. This is because if timing is controlled, it is related to the duration of the investigation period.¹¹ Cases in which the period between the date of occurrence of the case subject to disciplinary action and the disciplinary action exceeds one year are very rare; on average, the period is about six months.¹² Therefore, even if the time to disciplinary action were controlled, the impact on the main results using annual data would be very small.

Another issue is the inability to determine whether it is the effect of the disciplinary action or its coverage in the media. As the analysis in the Appendix shows, disciplinary actions severer than suspension are in principle reported, while pay cuts and reprimands are generally not. Therefore, in this study, the imposition of a heavy discipline action and the reporting are almost synonymous, and it is not possible to distinguish which is the source of the effect.

¹¹There is little difference between the date the disciplinary action is decided and the timing of the newspaper report; most of the reported cases appeared in the newspapers the day after the formal disciplinary action was taken.

¹²See Appendix A.

6.2 Policy implication

In the previous literature analyzing the Ferguson effect, it was found that when police misconduct, such as violent incidents by police officers, was widely reported, subsequent police activity was reduced and the number of arrests decreased (Mikdash and Zaiour, 2022; Cheng and Long, 2022; Campbell, 2024). In contrast, this study finds the opposite: a temporary increase in cases cleared following misconduct.

This divergence can be attributed not only to differences in the severity of incidents but also to the underlying motivational mechanisms. As this study demonstrates, in cases of relatively minor misconduct, the behavioral response is not primarily driven by a loss of public trust, but rather by internal promotion incentives faced by police station chiefs. When a station's reputation is damaged by misconduct, chiefs may seek to offset the negative evaluation by improving visible performance indicators, such as cases cleared or the clearance rate. These short-term performance improvements appear to be a strategic response. They aim to preserve or enhance career prospects, rather than restore community trust.

Moreover, as in the case of Rackstraw (2023), which looked at the impact of the TV show, the indicators of improvement are biased towards those that can be changed relatively easily, such as an increase in cases cleared for burglary thefts. Only the improvement in relatively minor crimes and the short-term nature of the effect suggest that it has contributed little to real improvements in public safety. Furthermore, relying solely on quantitative activity indicators may lead to manipulation of those indicators. Thus, it suggests the need to focus on police performance measures that lead to substantive improvements in public safety and are difficult to manipulate, such as arrest rates for violent offenders, and non-quantitative indicators when assessing performance.

7 Conclusion

This study investigates how police organizations respond to minor incidents of misconduct, using empirical evidence from local police stations in Tokyo. The central finding is that, following public disclosure of disciplinary actions, police activity temporarily intensifies, cases cleared increase, particularly for theft and especially burglary, while the number of reported crimes declines, resulting in a short-term improvement in clearance rates.

Crucially, these behavioral changes are not driven by shifts in public trust. Instead, the evidence points to internal career incentives, specifically, the promotion prospects of station chiefs as the primary mechanism. Supplementary analysis shows that misconduct reduces a chief's likelihood of promotion, but this negative effect can be partially offset by improved performance metrics such as cases cleared or clearance rates. In this sense, the organizational response is better understood as a reputational management strategy aimed at internal audiences, rather than an effort to regain public trust.

This finding challenges the conventional narrative that trust erosion invariably leads to reduced police engagement. Instead, the study demonstrates that internal incentives can lead to increased police activity, even in the wake of reputational damage, particularly when the misconduct is relatively minor and the incentive to signal competence remains high.

More broadly, the results highlight a critical concern in performance management: metric-driven behavior can lead to superficial improvements that do not translate into meaningful gains in public safety. Cases cleared may rise, but often through the selective targeting of crimes that are easier to solve or more visible.

Future research should examine how the design of performance evaluation systems shapes organizational behavior and consider the possible unintended consequences of using narrow indicators for public accountability.

References

- Adamson, J. and Rentschler, L. (2023). Criminal justice from a public choice perspective: an introduction to the special issue. *Public Choice*, 196:223–227.
- Adamson, J. and Rentschler, L. (2025). How policing incentives affect crime, measurement, and justice. *Economic Inquiry*, 63:545–567.
- Ang, D., Bencsik, P., Bruhn, J., and Derenoncourt, E. (2025). Community engagement with law enforcement after high-profile acts of police violence. *American Economic Review: Insights*, 7:124–140.
- Baicker, K. and Jacobson, M. (2007). Finders keepers: Forfeiture laws, policing incentives, and local budgets. *Journal of Public Economics*, 91:2113–2136.
- Butts, K. (2023). Difference-in-differences estimation with spatial spillovers. Working Paper.
- Callaway, B. and Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225:200–230.
- Campbell, T. (2024). Black lives matter’s effect on police lethal use of force. *Journal of Urban Economics*, 141.
- Cheng, C. and Long, W. (2022). The effect of highly publicized police killings on policing: Evidence from large U.S. cities. *Journal of Public Economics*, 206.
- Conover, E., Kraynak, D., and Singh, P. (2023). The effect of traffic cameras on police effort: Evidence from india. *Journal of Development Economics*, 160.
- de Chaisemartin, C. and D’Haultfœuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, pages 1–45.
- de Chaisemartin, C. and D’Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110:2964–2996.

- Devi, T. and Fryer, R. G. (2020). Policing the police: The impact of "Pattern-or-Practice" investigations on crime. NBER Working Paper.
- Dewatripont, M., Jewitt, I., and Tirole, J. (1999). The economics of career concerns, part ii: Application to missions and accountability of government agencies. *Review of Economic Studies*, 66:199–217.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225:254–277.
- Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182.
- Jackson, J. and Bradford, B. (2010). What is trust and confidence in the police? *Policing*, 4:241–248.
- Jia, R. (2024). Pollution for promotion. *The Journal of Law, Economics, and Organization*, 41(3):866–892.
- Jordan, A. and Kim, T. (2025). Strengthening police oversight: The impacts of misconduct investigators on police officer behavior. *Journal of Policy Analysis and Management*, 44(4):1286–1316.
- Kim, T. (2022). Promotion incentives, career decisions, and police performance. SSRN Working Paper.
- Mastorocco, N. and Ornaghi, A. (2025). Who watches the watchmen? local news and police behavior in the united states. *American Economic Journal: Economic Policy*, 17:285–318.
- Mikdash, M. and Zaiour, R. (2022). Does (all) police violence cause de-policing? evidence from george floyd and police shootings in minneapolis. *AEA Papers and Proceedings*, 112:170–173.
- National Police Agency (2021). Police of Japan 2021. Technical report, National Police Agency.

- New York City Police Department (2024). Discipline in the NYPD 2023. Technical report, New York City Police Department.
- Ornaghi, A. (2019). Civil service reforms: Evidence from U.S. police departments. Working paper.
- Owens, E. and Ba, B. A. (2021). The economics of policing and public safety. *Journal of Economic Perspectives*, 35:3–28.
- Prendergast, C. (2001). Selection and oversight in the public sector, with the Los Angeles Police Department as an example. Working Paper.
- Prendergast, C. (2021). "Drive and Wave": The response to LAPD police reforms after rampart. Working Paper.
- Rackstraw, E. (2023). When reality tv creates reality: How "copaganda" affects police, communities, and viewers. SSRN Working Paper.
- Rozema, K. and Schanzenbach, M. (2019). Good cop, bad cop: Using civilian allegations to predict police misconduct. *American Economic Journal: Microeconomics*, 11:225–268.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225:175–199.
- Tirole, J. (1994). The internal organization of government. *Oxford Economic Papers*, 46(1):1–29.
- Tokyo Metropolitan Police Department (2024). Keishicho no Toukei (Reiwa 5 nen) (Statistics of the Tokyo Metropolitan Police Department 2023). Technical report, Tokyo Metropolitan Police Department.
- Xu, G., Xu, L. C., Si, R., and Wang, R. (2025). Career incentives, tournament competition, and performance manipulation: Evidence from chinese cities. *The Journal of Law and Economics*, 68:459–498.

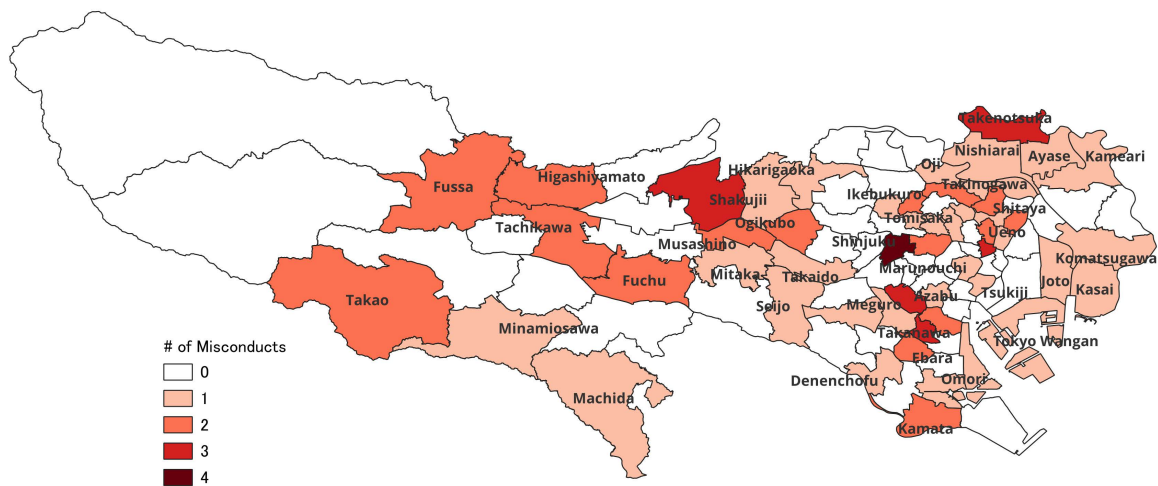
Table 1: Summary statistics

(a) Counts

N=1,056	Offenses known to the police				Cases cleared			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Total	1,205.97	743.24	118	4,815	383.32	200.41	70	1,555
Felony	7.49	6.65	0	69	6.88	5.93	0	72
Violent offenses	81.48	58.31	6	581	61.69	40.71	4	449
Theft	846.44	557.70	77	3,346	201.74	115.80	16	714
Burglary theft	47.49	37.40	1	257	44.28	34.96	1	212
Non-burglary theft	798.95	528.08	71	3,203	157.46	94.23	14	598
Vehicle theft	405.02	318.39	16	1,828	21.28	15.63	0	95
Intellectual offenses	79.40	78.55	8	1,940	36.82	21.57	3	260
Moral offenses	10.63	7.60	0	49	9.04	6.22	0	46

(b) per km^2

N=1,056	Offenses known to the police				Cases cleared			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Total	162.22	155.58	0.679	1,110	54.81	49.88	0.380	344.0
Felony	1.09	1.46	0	14.2	1.00	1.32	0	14.8
Violent offenses	13.06	15.51	0.033	120	9.60	10.64	0.033	92.6
Theft	110.71	108.44	0.434	869	27.34	25.30	0.179	190.0
Intellectual offenses	12.81	32.27	0.043	840	6.10	6.58	0.027	60.1
Moral offenses	1.46	1.68	0	16.4	1.28	1.49	0	14.7



Note: Colored police stations reported at least one disciplinary action during the sample period (2013-2023). The Shinjuku police station reported misconduct four times within the 11-year period, the highest number.

Figure 1: TMPD map

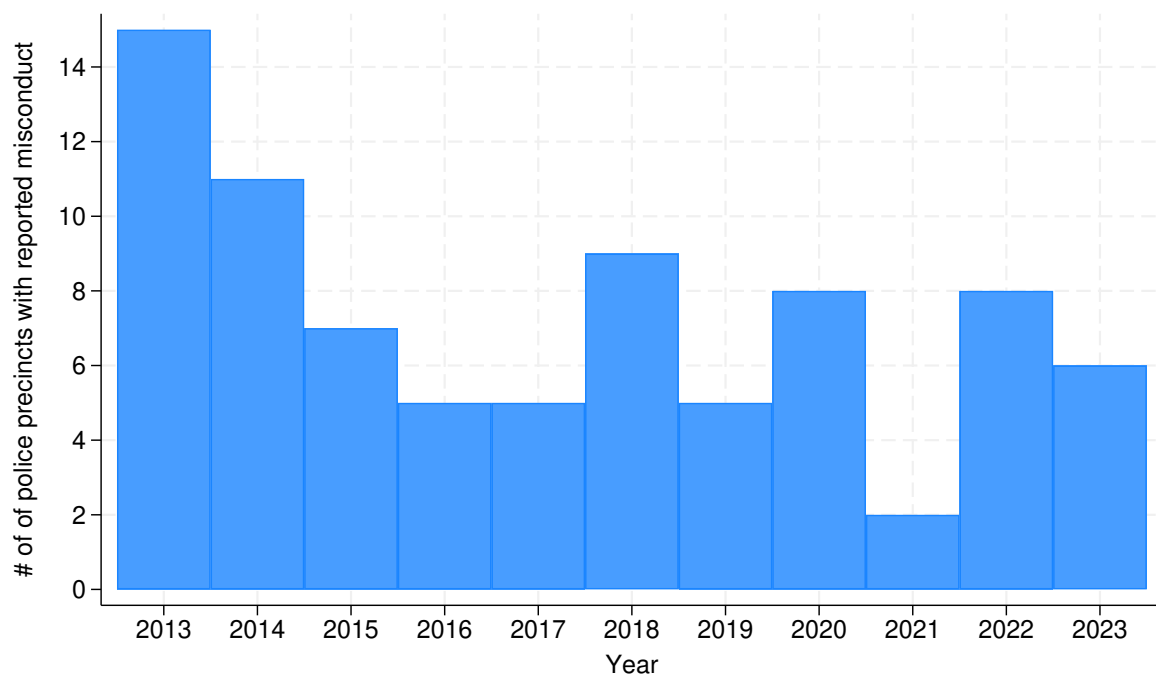
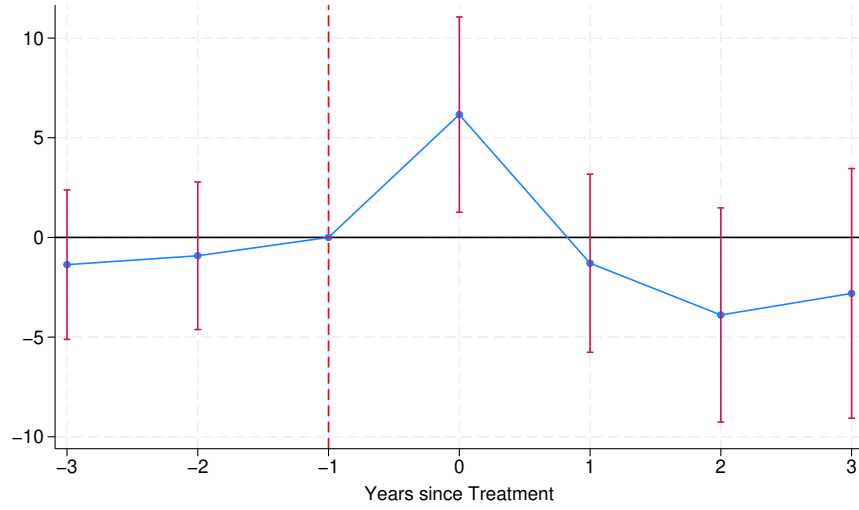
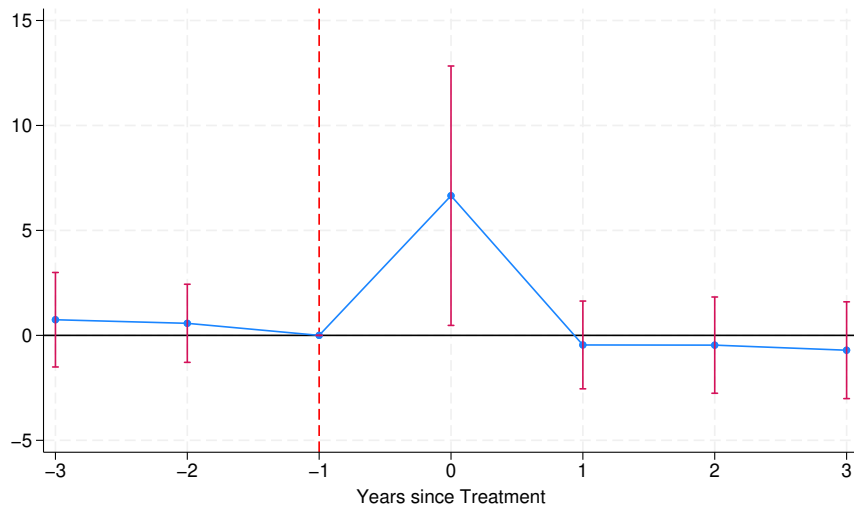


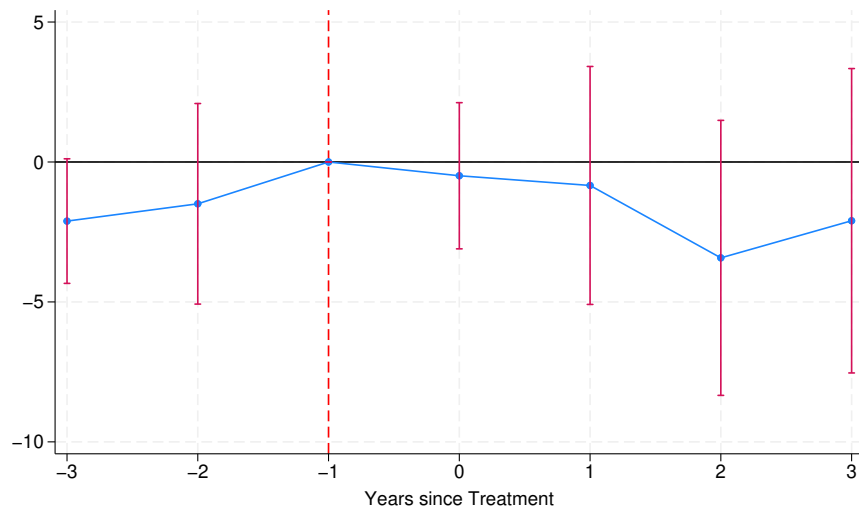
Figure 2: Treatment timing



(a) Theft



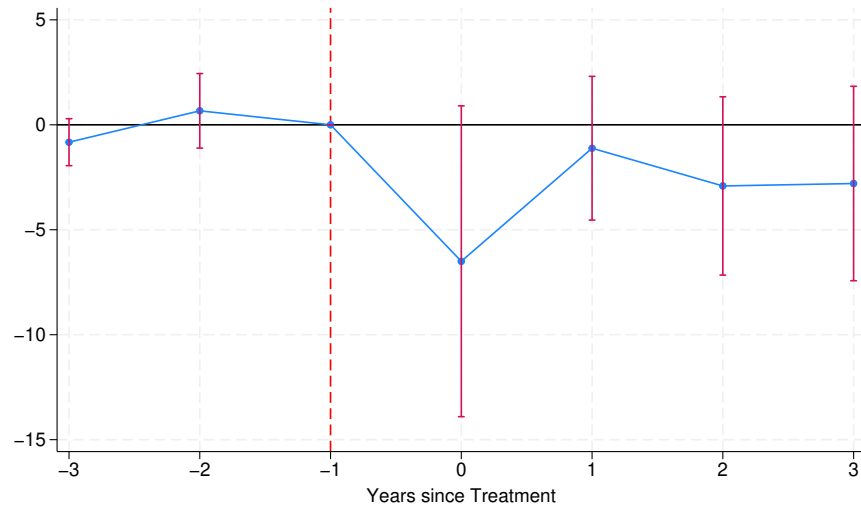
(b) Burglary theft



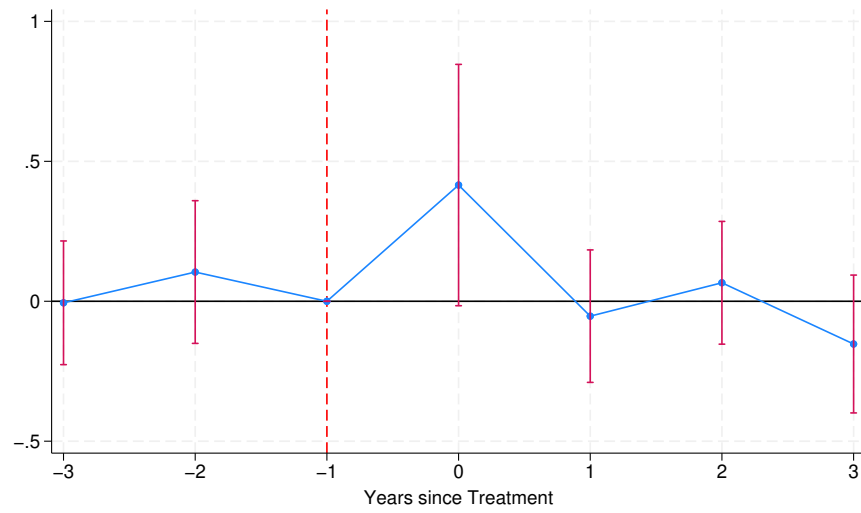
(c) Non-burglary theft

Figure 3: dCDH estimators for specific crime (cases cleared)

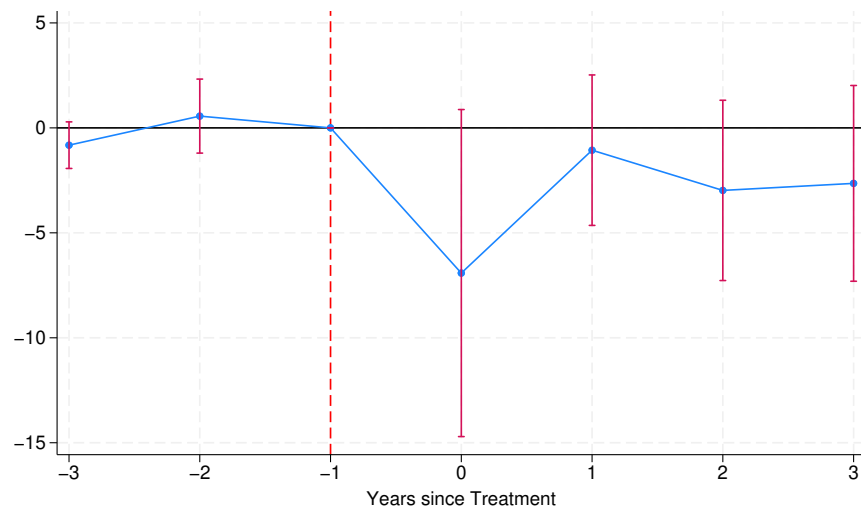
Note: The figure plots the treatment effects relative to the year of the disciplinary action (year 0), with 95% confidence intervals.



(a) Theft

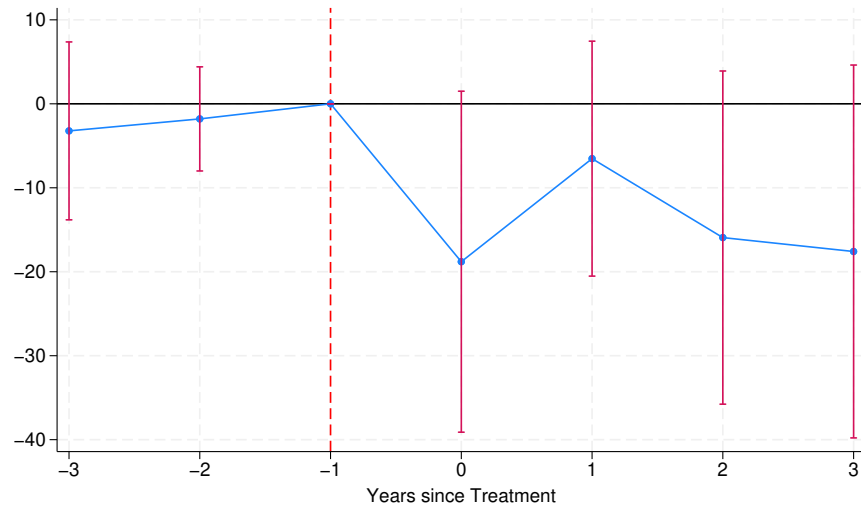


(b) Burglary theft

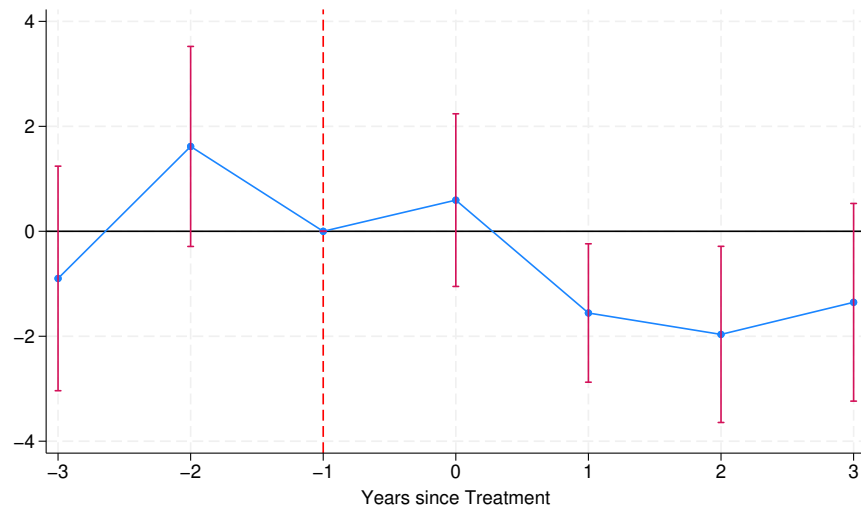


(c) Non-burglary theft

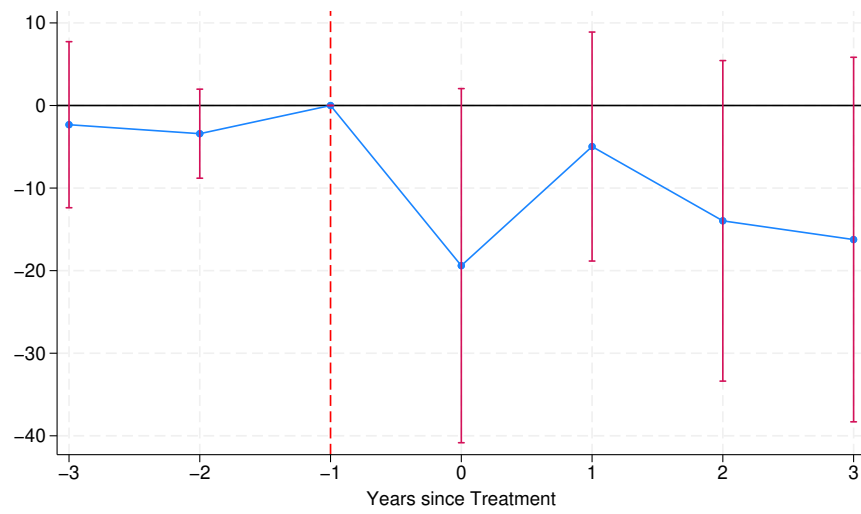
Figure 4: dCDH estimators for specific crime (persons cleared for offenses)



(a) Theft

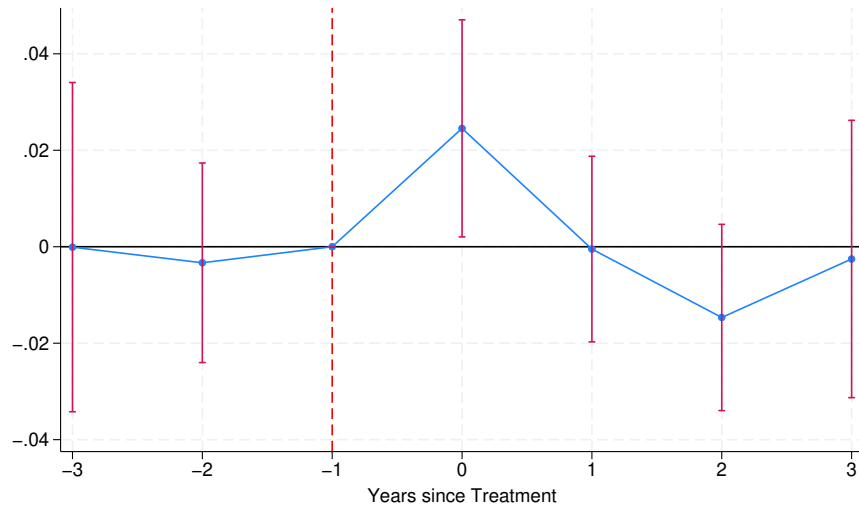


(b) Burglary theft

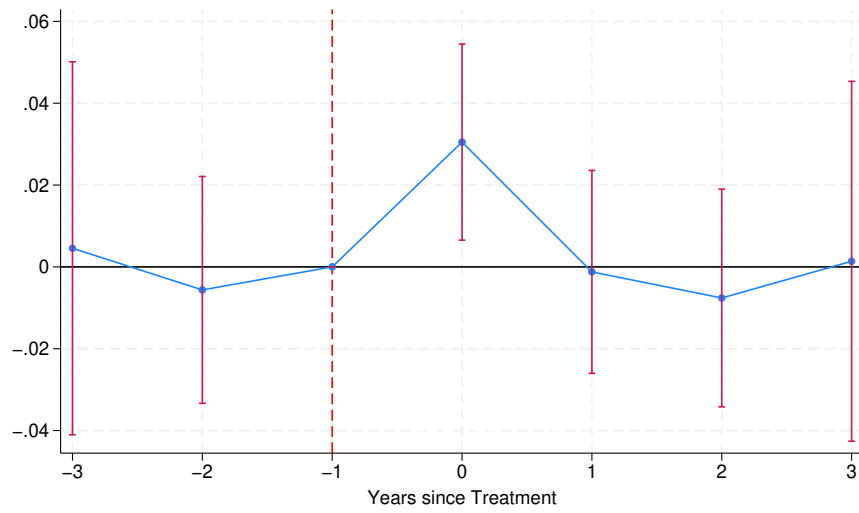


(c) Non-burglary theft

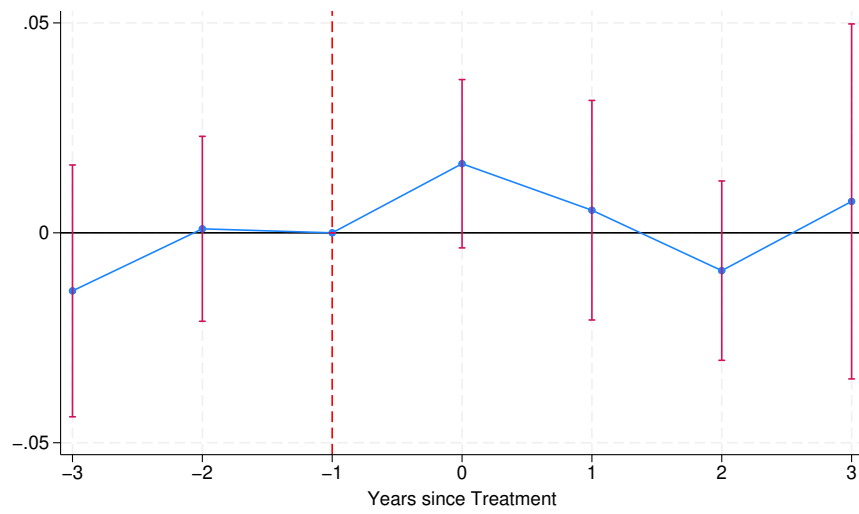
Figure 5: dCDH estimators for specific crime (offenses known to the police)



(a) Total

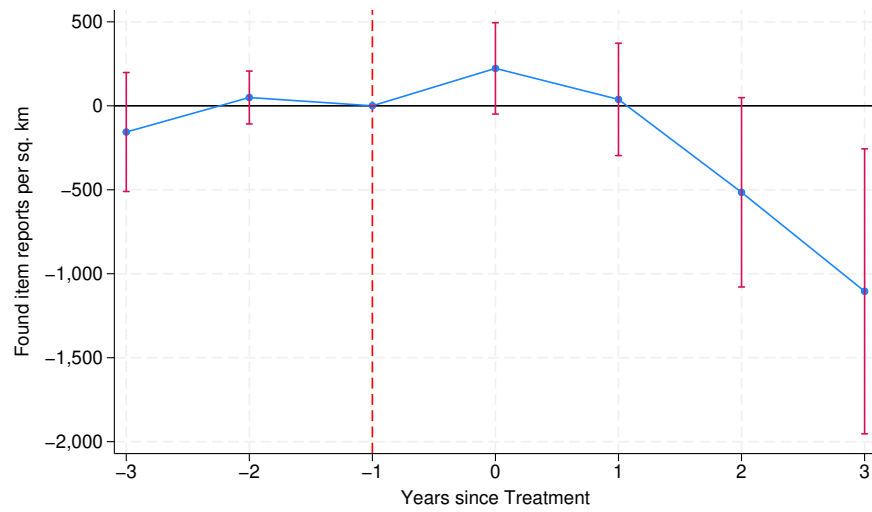


(b) Theft

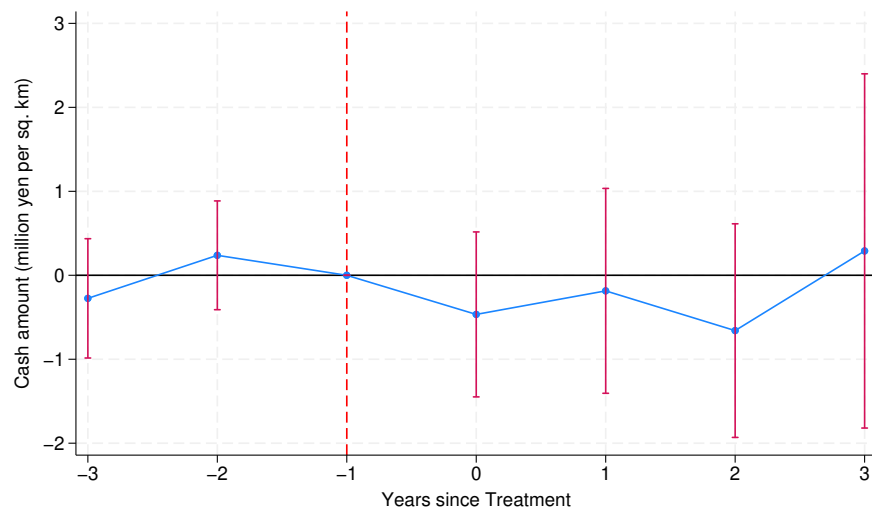


(c) Non-burglary theft

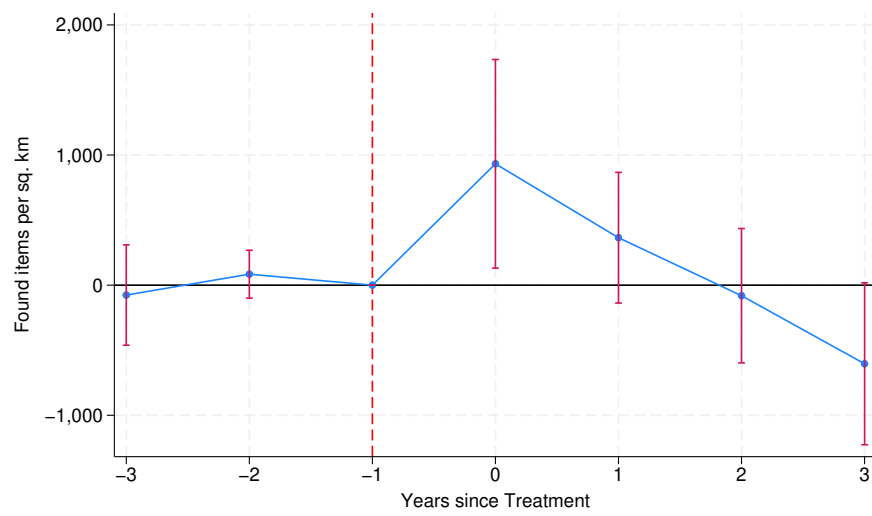
Figure 6: dCDH estimators (clearance rate)



(a) Found item reports



(b) Money



(c) Number of found items

Figure 7: Effects on reporting behavior

Table 2: Reasons for disciplinary action

Reason	Count	%
Job-related		
Bribery, etc.	14	15.6
Abandonment, neglect, or improper duties	11	12.2
Gun accident	4	4.4
Abuse of authority	3	3.3
Violation of work discipline	1	1.1
Private		
Assault, bodily harm, etc.	2	2.2
Indecent assault, molestation, voyeurism, etc.	25	27.8
Theft, fraud, embezzlement, etc.	13	14.4
Traffic accidents and violations	12	13.3
Others	5	5.6
Total	90	100

Table 3: Penalties and resignation

Penalty	Total Person	Left the job Person	%
Dismissal	30	30	100.0
Suspension	33	32	97.0
Pay cut	21	18	85.7
Reprimand	5	2	40.0
Total	89	82	92.1

Note: The determination of whether or not he/she left the position is based on news coverage. One case is excluded because the details of the disciplinary action are not clear.

Table 4: Spillover effects for theft (cases cleared)

$N = 1,056$	Coefficient	S.E.	LB CI	UB CI
Effect (Treated)	13.38	5.32	2.96	23.81
Effect (Adjacent)	1.17	1.77	-2.29	4.63

Note: The estimation follows Butts (2023), and no dynamic treatment effects are assumed. The standard errors are bootstrapped with 1,000 replications.

Table 5: Effects on promotion

	(1)	(2)	(3)	(4)
Promotion within 400 days	Linear	Logit	Linear	Linear
Misconduct	-0.32** (0.13)	-2.65** (1.33)	-0.54*** (0.17)	0.04 (0.11)
HighCCGrowth	-0.02 (0.05)	-0.09 (0.26)	0.01 (0.06)	-0.11 (0.07)
Misconduct $\times$ HighCCGrowth	0.38** (0.18)	3.01** (1.49)	0.60*** (0.22)	-0.06 (0.16)
PS dummies	Y	Y	Y	Y
N	437	353	358	79

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses.

Note: The variable *HighCCGrowth* is a dummy that takes the value 1 if the year-on-year percentage change in cases cleared at a police station exceeds the annual average across other stations. Both regression models include police station fixed effects (PS dummies). Coefficients from the logit model are reported in log-odds units, and the sample size becomes smaller because not all police stations experienced misconduct. Columns 3 and 4 are regressions using only data for chiefs of small and large police stations, respectively. Police officers who retired within approximately one year of the end of their term as chief were not included in the sample.

Appendix

A Disciplinary actions in TMPD

This appendix analyzes the characteristics of disciplinary actions within the Tokyo Metropolitan Police Department (TMPD), including those not publicly reported. To situate these patterns in a broader context, a comparative reference to the NYPD, a similarly sized but institutionally different police force, is also provided.

Although the number of disciplinary actions reported in the four major newspapers was used in this study for analysis, some disciplinary actions are not actually reported in the newspapers. Therefore, this section discusses the characteristics of the reported disciplinary actions based on all data, including those not reported, obtained through the information disclosure request under Japanese law, and clarifies the implications of the main analysis results. Note, however, that the information on disciplinary actions obtained through information disclosure requests is limited to the most recent two years (2022 and 2023) only, and thus does not include the entire period covered by this study.

First, according to data obtained through information disclosure requests, 31 and 17 individuals were disciplined in 2022 and 2023, respectively. As information on the police stations to which they belong is confidential, this number includes officers who belong to the headquarter. The details are summarized in Tables A.1 and A.2.

In Table A.1, it can be seen that the largest number of punishments are based on sexual offense-related issues, while other punishments are mostly based on violations of laws and regulations that are unrelated to work.

Table A.2 shows that the most common disciplinary outcome was a pay cut, which accounts for approximately two-thirds of all disciplinary actions, but is rarely mentioned in news articles. This is probably because the guidelines for the publication of police disciplinary actions stipulate that disciplinary actions heavier than suspension should, in principle, be publicized, while those lighter than suspension do not need to be publicized.

Table A.1: Reasons for disciplinary action

Reason	2022	2023	Total
Indecent assault, molestation, voyeurism, etc.	17	3	20
Theft, fraud, embezzlement, etc.	3	5	8
Traffic accidents and violations	3	3	6
Others	3	0	3
<i>Violation of work discipline</i>	1	1	2
<i>Bribery, etc.</i>	2	0	2
<i>Gun accident</i>	1	1	2
Other violations of laws and regulations	0	1	1
Thoughtless behavior	0	1	1
Assault, bodily harm, etc.	1	0	1
<i>Abandonment, neglect, or improper duties</i>	0	1	1
<i>Abuse of authority</i>	0	1	1

Note: Job-related reasons are italicized.

Nineteen of the disciplinary actions that occurred during the two-year period were reported in the newspaper, seven of which were dismissals. In 10 of the remaining 12 cases, the article states that the disciplined police officers resigned. Thus, in most cases where disciplinary action was reported, the disciplined officer left his/her position.

Lastly, to assess the endogeneity of the date of disciplinary action, I analyzed the discrepancy between the date of the act subject to the disciplinary action and the date of disciplinary action itself. Out of 48 cases, information on the date of the act was obtained for 31 cases. However, some of these cases involved actions that spanned a long period or where the exact date of the act could not be determined. Therefore, based on the available information regarding the date of each act subject to disciplinary action, I calculated the earliest possible date and examined the discrepancy between that date and the date of disciplinary action.

The average time between the act and the date of disciplinary action was 182 days, with the shortest being 7 days and the longest 1,609 days. Furthermore, in 19 out of the 31 cases, disciplinary actions were taken within 70 days. There were three cases where the disciplinary action took more than a year, but all of these involved actions that spanned a long period or cases where the date of the act could only be specified to the year based on the available information. Therefore, it can be considered that cases

Table A.2: Penalties of disciplinary actions

Penalty	2022	2023	Total
Dismissal	4	3	7
covered in the news	4	3	7
left the position	4	3	7
Suspension	5	2	7
covered in the news	5	2	7
left the position	4	2	6
Pay cut	20	11	31
covered in the news	3	1	4
left the position	2	1	3
Reprimand	2	1	3
covered in the news	1	0	1
left the position	1	0	1

Note: The determination of whether or not he/she left the position is based on news coverage.

where more than a year has passed from discovery to disciplinary action are rare. Consequently, even if the date of disciplinary action was determined endogenously, it is likely that it could only be delayed by around 6 to 8 months at most.

To gain a clearer understanding of the situation at the TMPD, it is useful to compare its data with that of the NYPD, an organization of comparable size but with operating under a different disciplinary system.

According to New York City Police Department (2024), 487 disciplinary cases were adjudicated in the NYPD in 2023. Of these, more than half (277) were resolved through settlements. The remaining cases involved either trials or separations from employment. Among the 114 cases that proceeded to trial, 87 resulted in a guilty finding or plea. Of those, 22 led to termination or forced separation, the most severe form of disciplinary action. On average, it takes approximately one to two years for disciplinary measures to be carried out.

In comparison, the number of disciplinary actions taken by the TMPD is significantly lower than that of the NYPD, and the time until such actions are implemented tends to be shorter.

B Estimation

Balanced panel data was used for estimation, but due to heterogeneity in treatment history, the sample sizes used for DID estimates in each period were different. The sample sizes used for DID in each period are shown in Table B.1, and Tables B.2, B.3, and B.4 are the summary of the estimators for Figures 3, 5, and 6.

Table B.1: Sample size used for estimation

	N	Switchers
Effect_8	107	16
Effect_7	158	17
Effect_6	212	21
Effect_5	270	26
Effect_4	327	28
Effect_3	390	31
Effect_2	410	31
Effect_1	475	34
Effect_0	561	51
Placebo_1	465	27
Placebo_2	394	21
Placebo_3	329	20
Placebo_4	265	16
Placebo_5	205	11
Placebo_6	150	9

Table B.2: dCDH estimators for specific crime (cases cleared)

	Theft		Burglary		Non-burglary	
	$DID_{g,l}$	S.E.	$DID_{g,l}$	S.E.	$DID_{g,l}$	S.E.
Effect_8	-7.33	4.43	-1.52	1.96	-5.81	3.66
Effect_7	-6.70	3.21	-0.86	1.63	-5.84	3.06
Effect_6	-1.09	3.91	-0.70	1.80	-0.38	2.83
Effect_5	-3.96	4.52	-1.19	1.57	-2.77	3.73
Effect_4	-6.45	4.03	-1.47	1.20	-4.99	3.51
Effect_3	-2.81	3.19	-0.71	1.18	-2.10	2.77
Effect_2	-3.89	2.74	-0.46	1.17	-3.43	2.51
Effect_1	-1.29	2.28	-0.46	1.07	-0.84	2.17
Effect_0	6.16	2.50	6.65	3.15	-0.49	1.33
Placebo_1	-0.92	1.89	0.57	0.95	-1.49	1.83
Placebo_2	-1.37	1.91	0.75	1.15	-2.11	1.14
Placebo_3	0.20	1.85	0.42	1.34	-0.23	1.64
Placebo_4	-1.53	2.30	-1.73	1.24	0.19	1.75
Placebo_5	-1.82	3.22	-1.24	2.01	-0.58	2.15
Placebo_6	-5.18	6.33	0.88	1.27	-6.06	5.63
Mean	27.34		6.09		21.25	

Table B.3: dCDH estimators for specific crime (offenses known to the police)

	Theft		Burglary		Non-burglary	
	$DID_{g,l}$	S.E.	$DID_{g,l}$	S.E.	$DID_{g,l}$	S.E.
Effect_8	-47.39	28.11	-0.52	1.69	-46.87	27.47
Effect_7	-48.19	30.02	-0.49	1.56	-47.69	29.29
Effect_6	-35.76	22.82	-0.43	1.20	-35.34	22.45
Effect_5	-26.04	18.94	-1.18	0.96	-24.86	18.59
Effect_4	-23.71	13.84	-0.97	0.85	-22.74	13.89
Effect_3	-17.59	11.33	-1.35	0.96	-16.23	11.26
Effect_2	-15.93	10.12	-1.96	0.86	-13.97	9.90
Effect_1	-6.53	7.14	-1.56	0.67	-4.97	7.07
Effect_0	-18.80	10.36	0.59	0.84	-19.40	10.94
Placebo_1	-1.80	3.16	1.62	0.97	-3.41	2.75
Placebo_2	-3.23	5.41	-0.90	1.09	-2.33	5.13
Placebo_3	-2.47	2.66	0.68	0.90	-3.15	2.74
Placebo_4	-9.15	4.74	1.27	0.62	-10.41	4.69
Placebo_5	1.73	3.55	-1.81	1.03	3.54	3.40
Placebo_6	1.86	4.59	1.59	0.85	0.27	4.16
Mean	110.71		5.80		104.90	

Table B.4: dCDH estimators for specific crime (clearance rate)

	Total		Theft		Non-burglary	
	$DID_{g,l}$	S.E.	$DID_{g,l}$	S.E.	$DID_{g,l}$	S.E.
Effect_8	0.016	0.023	-0.007	0.024	-0.009	0.022
Effect_7	-0.009	0.022	-0.032	0.024	-0.032	0.022
Effect_6	0.005	0.024	0.007	0.028	0.009	0.020
Effect_5	0.002	0.020	0.009	0.022	0.014	0.024
Effect_4	-0.010	0.015	-0.022	0.018	-0.016	0.019
Effect_3	-0.003	0.015	0.001	0.022	0.007	0.022
Effect_2	-0.015	0.010	-0.008	0.014	-0.009	0.011
Effect_1	0.000	0.010	-0.001	0.013	0.005	0.013
Effect_0	0.025	0.011	0.030	0.012	0.016	0.010
Placebo_1	-0.003	0.011	-0.006	0.014	0.001	0.011
Placebo_2	0.000	0.017	0.005	0.023	-0.014	0.015
Placebo_3	0.013	0.012	0.013	0.016	0.008	0.014
Placebo_4	0.000	0.013	-0.016	0.017	0.002	0.010
Placebo_5	0.003	0.013	-0.005	0.018	-0.010	0.012
Placebo_6	-0.014	0.030	-0.005	0.036	-0.020	0.030
Mean	0.353		0.270		0.224	

C Effects on other crimes

In Figure C.1, unlike theft, cases cleared for almost all other categories of crime does not change before and after the disciplinary action.

Figure C.2 presents the dCDH estimators for the number of individuals cleared for offenses. Unlike cases cleared, the number of persons cleared for offenses decreases in the year of action except for moral offenses. The estimated effects are statistically significant, especially with regard to the decline in felony crimes and theft. However, these effects do not appear to be long-lasting either.

Figure C.3 shows the dCDH estimators for offenses known to the police. The results show that the effect estimates for each crime on the number of known crimes are negative in most cases, but not at a statistically significant level. However, the point estimates for theft cases are coherently negative after treatment, and the effect is statistically significant, especially in the treatment year.

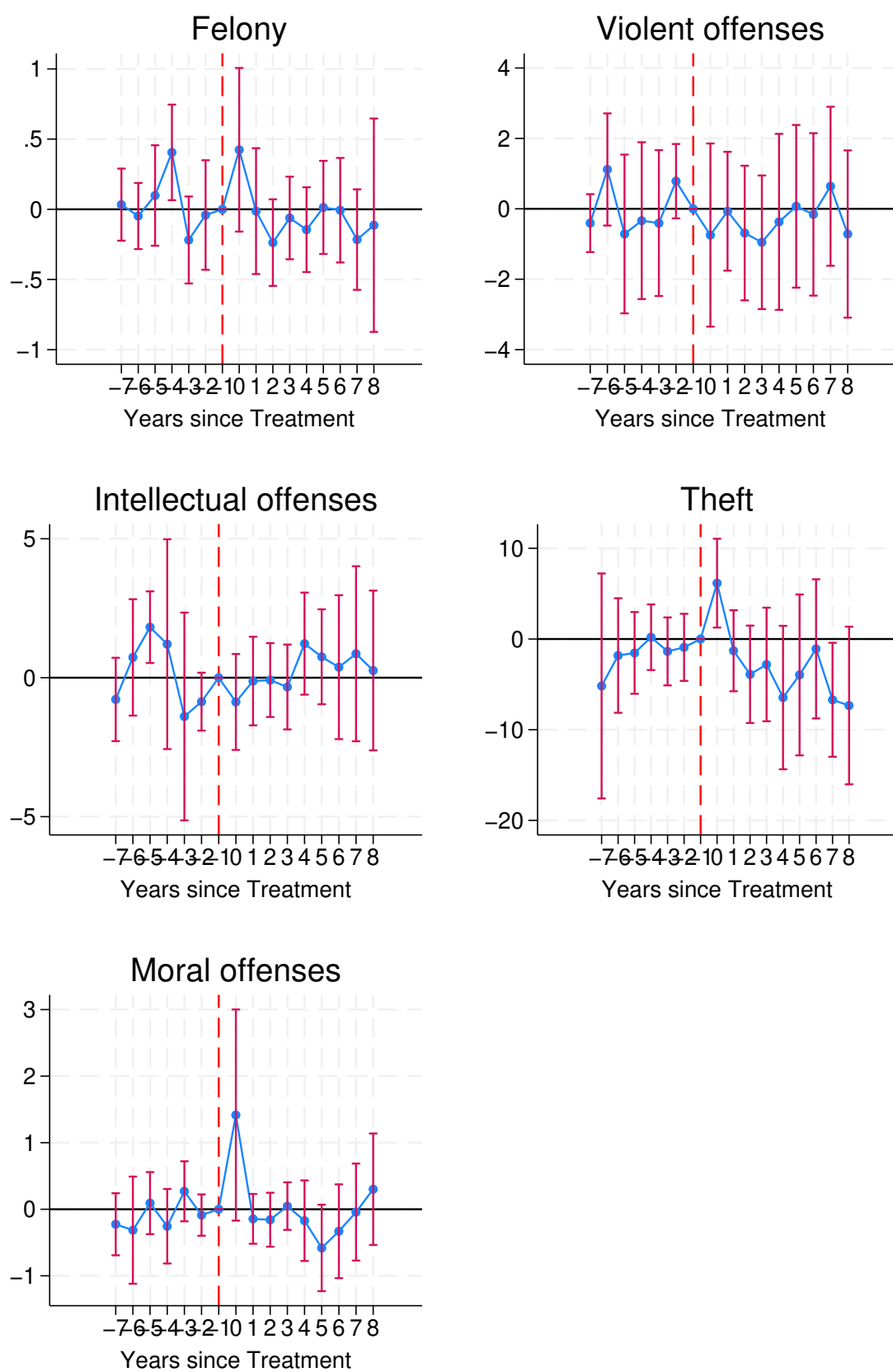


Figure C.1: dCDH estimators (cases cleared)

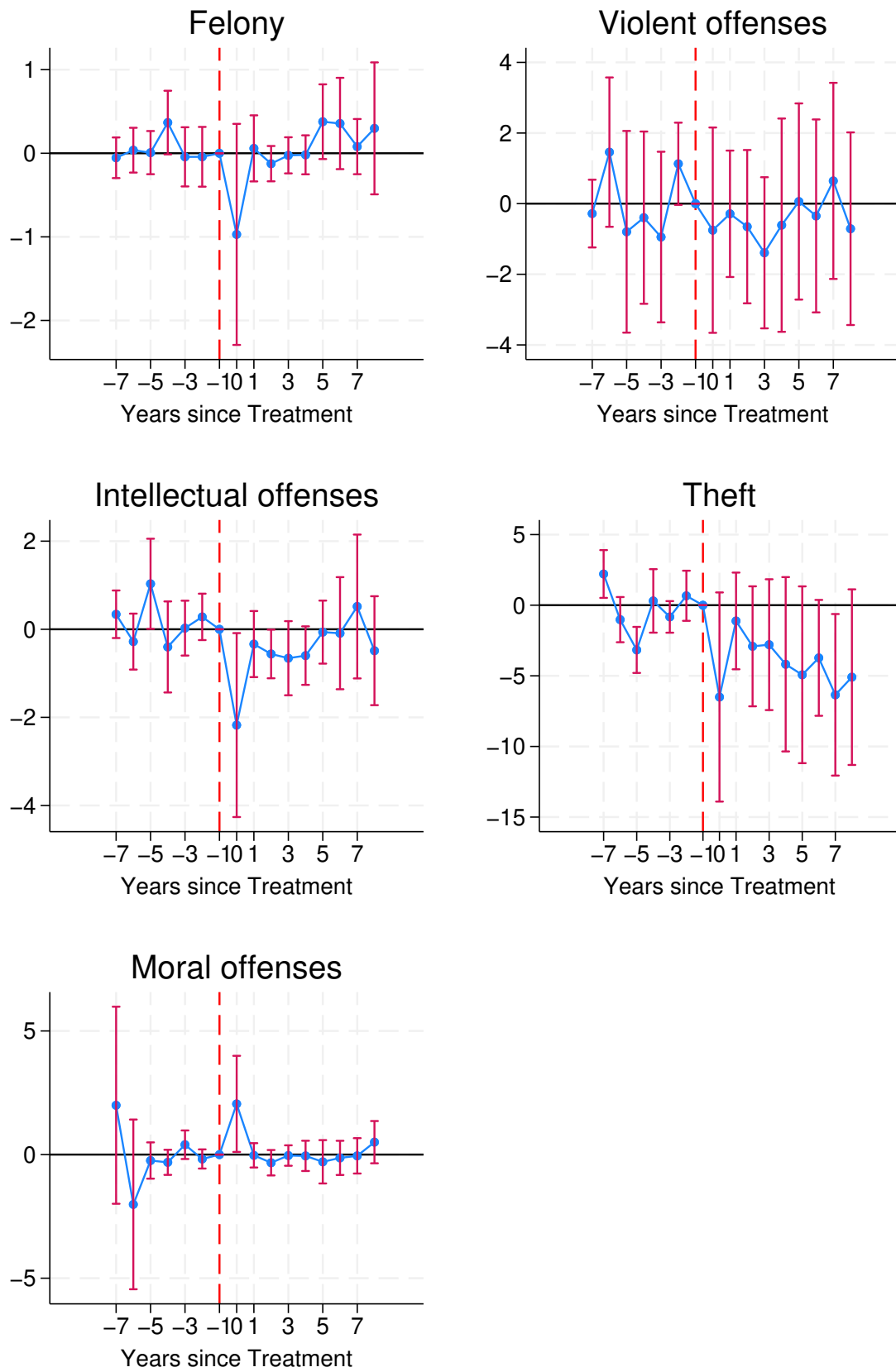


Figure C.2: dCDH estimators (persons cleared for offenses)

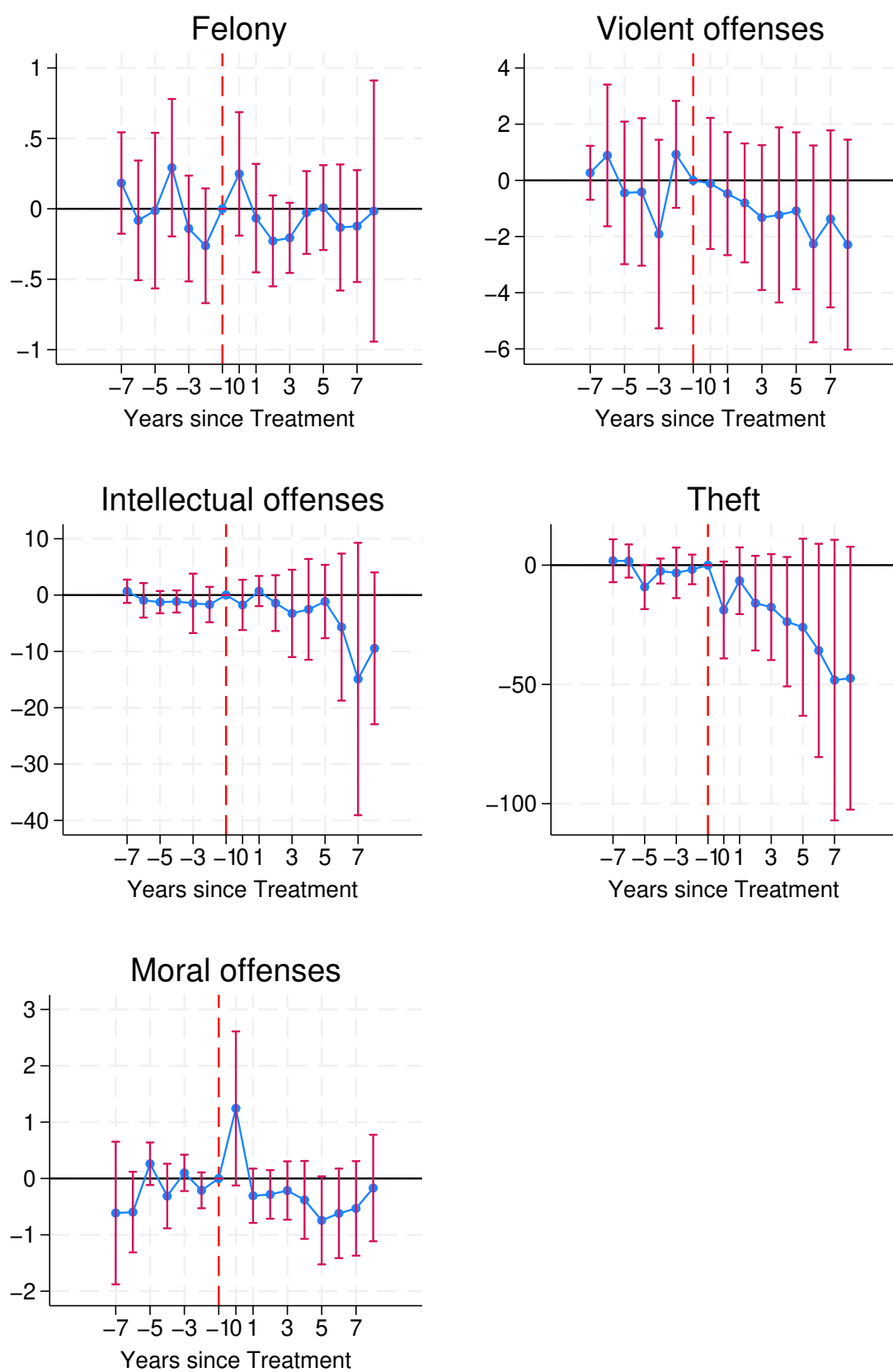


Figure C.3: dCDH estimators (offenses known to the police)

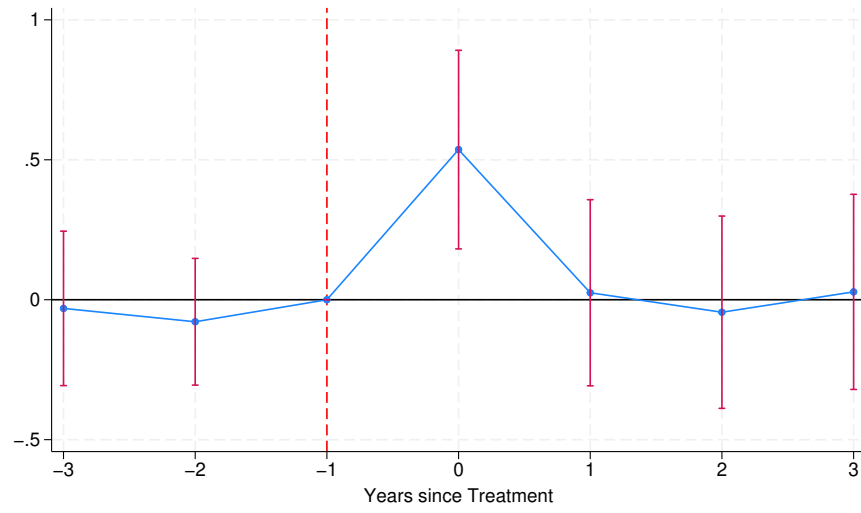
D Robustness checks

This section shows the results where the outcome is measured per capita rather than per area. Note that the Marunouchi police station, which is located in the center of Tokyo, is excluded from the sample because its jurisdiction is dominated by offices and has very few residents, which would result in an extremely high figure when converted to a per capita basis. Figures D.1 and D.2 illustrate the results. As with the per-area analysis, there is a statistically significant increase in cases cleared. A temporary decline is also observed in the number of reported offenses, although this decrease is not statistically significant.

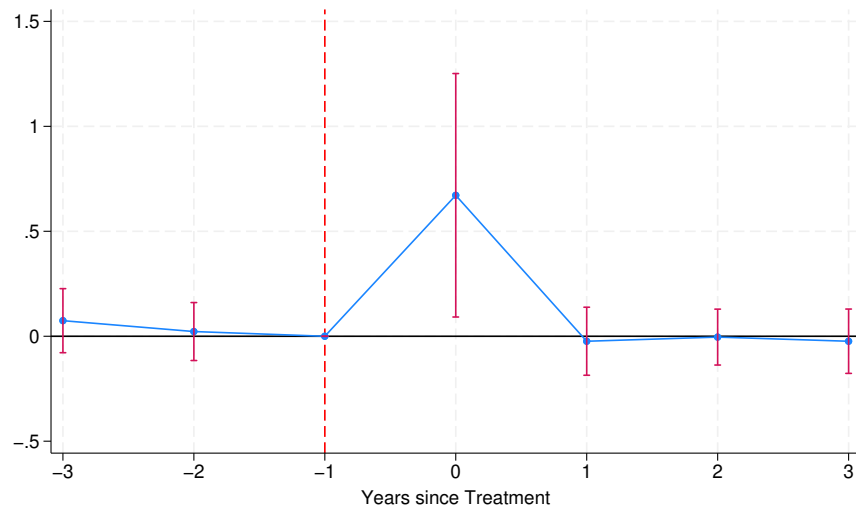
Table B.5: Summary statistics per 1,000 residents

	Offenses known to the police				Cases cleared			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Total	11.80	12.63	2.36	111	3.95	4.04	0.794	32.2
Felony	0.08	0.10	0	0.84	0.07	0.09	0	0.79
Violent offenses	0.94	1.25	0.12	9.16	0.69	0.85	0.083	6.18
Theft	8.03	8.60	1.50	80.8	1.98	2.04	0.324	18.1
Intellectual offenses	0.96	3.06	0.10	80.4	0.44	0.55	0.024	5.33
Moral offenses	0.10	0.12	0	1.19	0.09	0.11	0	1.26

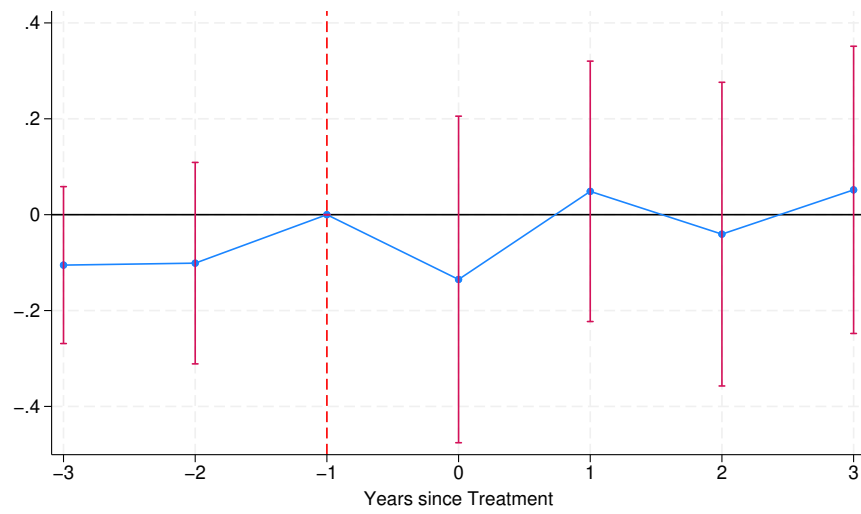
Note: Marunouchi police station is excluded from the sample.



(a) Theft

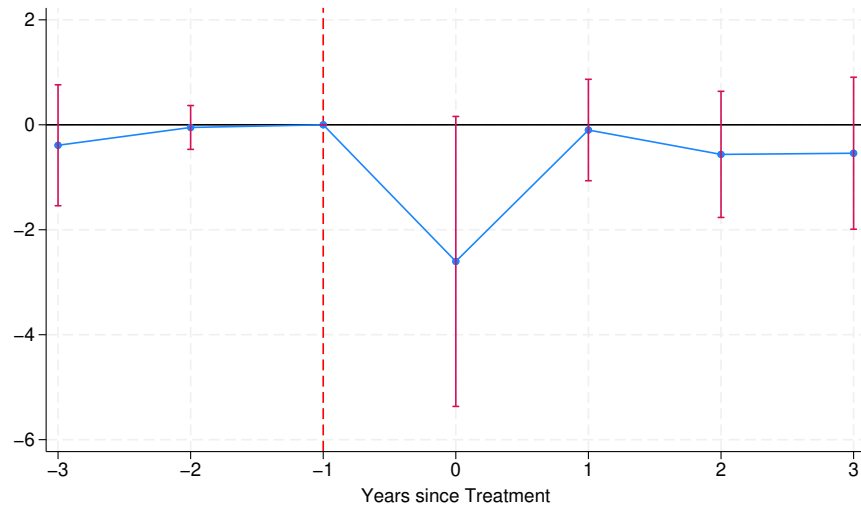


(b) Burglary theft

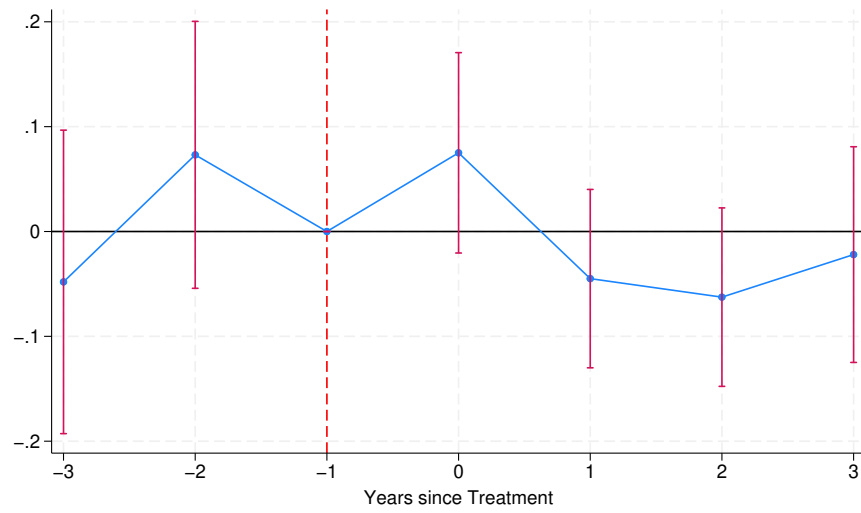


(c) Non-burglary theft

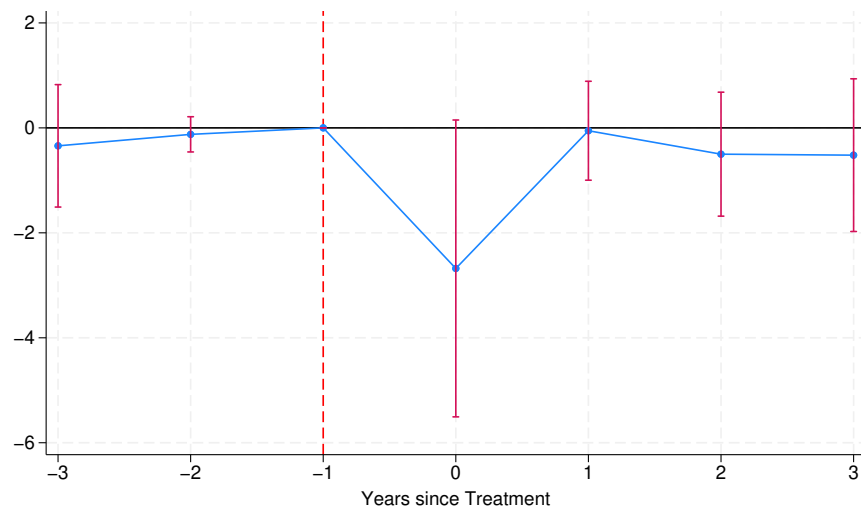
Figure D.1: dCDH estimators for specific crime (cases cleared)



(a) Theft



(b) Burglary theft



(c) Non-burglary theft

Figure D.2: dCDH estimators for specific crime (offenses known to the police)

E Model

Using the implicit function theorem on the first-order condition yields:

$$\frac{\partial a^*}{\partial \mu} = - \frac{\omega'(a^*) (a^* - 1) + \omega(a^*)}{\mu [\omega''(a^*) (a^* - 1) + 2 \omega'(a^*)] - c}.$$

Since $\omega'(a) < 0$ and typically $\omega''(a) < 0$, the denominator is likely negative when c is sufficiently large, implying $\partial a^* / \partial \mu < 0$. Hence, a decline in μ , representing a reputational loss after misconduct, can increase effort a^* , as the marginal reputational return to effort rises.