Administrative Science Quarterly

Punishing Misconduct in Moral Gray Zones: Evidence from the Chicago Police Department

Journal:	*Administrative Science Quarterly*
Manuscript ID	ASQ-25-0692
Manuscript Type:	Original Articles
Methodology Keywords:	Archival data, Regression analysis
Keywords:	Organizational theory, Corporate wrongdoing, Grand challenges, wicked problems, social innovation, Legitimacy
Abstract:	Police misconduct is a prevalent problem in the United States, and many have called for more punishment to address it. This idea aligns with longstanding sociological theory suggesting that punishment functions as a widespread deterrent by making clear the costs of certain behaviors. However, we argue that this logic only holds when the threat of punishment is clear and predictable, and many settings lack this clarity. In this paper, we investigate the Chicago Police Department (CPD), whose historical reluctance to consistently punish misconduct places its police officers in a moral gray zone: certain behaviors are technically prohibited but are, in practice, implicitly tolerated. We theorize that in such contexts, punishment will be seen as unjust and ultimately backfire, producing more misconduct. We test our theory with longitudinal data on all misconduct and punishments within the CPD between 1980 – 2017. Our findings show that after the CPD punishes an offending officer for misconduct, misconduct increases among their peers, and more severe punishment amplifies subsequent misconduct. This study offers new insights into research on misconduct in moral gray zones, examines the broader consequences of punishment in organizations, and highlights a major challenge in addressing police misconduct.

Please review ASQ's Data and Methods Transparency policy and guidelines:

I confirm I have read the policy and guidelines and agree to comply.

Does this manuscript use any qualitative data?

No, this manuscript does not use any qualitative data.

Does this manuscript use any quantitative data and/or code?

Yes, this manuscript uses quantitative data and/or code, and the author(s) will publicly share all data and/or code (and, if applicable, experimental stimuli or survey instruments) as required by ASQ's Data and Methods Transparency Policy and Guidelines.

Explain any constraints on data/code sharing and describe your alternative transparency plan; ensure that your explanation contains no information that identifies you or your institution. If no explanation is needed, please type N/A*

N/A

* Alternative approaches to transparency might include sharing a subset of the data, providing a detailed enough description of the data collection steps to enable other researchers to construct a comparable dataset, or supplying precise instructions for accessing licensed data, among other possible strategies.

INTRODUCTION

Police misconduct in the United States is one of the country's most persistent and consequential social problems. Every year, more than one million people report excessive use of force, resulting in roughly 1,000 deaths and 75,000 major injuries that require hospital treatment (Center for Disease Control, 2024). As a result, public trust in the police has declined to its lowest point in years (Pew Research Center, 2022). Among the many solutions likely required to address this complex problem, scholars, practitioners, and even the U.S. Department of Justice have called for more stringent punishment of the officers who engage in misconduct as a way to deter others in the organization from engaging in similar behaviors in the future (Rozema & Schanzenbach, 2023; Stephens, 2011; U.S. Department of Justice, 2017).

At first glance, this seems like a reasonable suggestion that builds on longstanding sociological theories of social deviance. For example, scholars have long argued that it is the punishment of rules, in addition to the rules themselves, that help people understand and internalize what is accepted or tolerated (Durkheim, 1895; Becker, 1968). Indeed, punishing a behavior reestablishes clear expectations as to where the line is between right and wrong as well as clarifies the costs actors will incur if they cross that line. Empirical work on misconduct and punishment largely supports this deterrence effect. Criminal justice scholars, for example, have demonstrated that punishing offenders reduces crimes committed by others in the community (Zimring, Hawkins, & Vorenberg, 1973; Schnake, 1986; Lee, 2017) and organizational theorists have shown that punishment deters others from engaging in similar behaviors (Palmer & Yenkey, 2015; Yiu, Xu, & Wan, 2014).

Upon closer examination, however, this work has focused almost entirely on settings where punishment is already a clear and predictable threat. For example, Yiu and colleagues

(2014) examined Chinese firms operating within a regulatory environment where punishment for financial fraud was expected, consistently enforced, and clearly communicated. Similarly, Palmer and Yenkey (2015) studied professional cycling during a period of systemic anti-doping reform, which established clear and consistent enforcement of penalties with severe consequences. In these studies, the threat of punishment not only helped clarify the line between right and wrong but also created straightforward costs for would-be offenders if they crossed that line. Under these conditions, it makes sense that punishment functions as an effective deterrent (Stafford & Warr, 1993; Bandura, 1977), and that more severe punishment would only amplify this effect (Nagin, 1998; Tittle, 1980). Yet not all settings have such a clear and predictable threat of punishment.

In fact, many settings are moral gray zones (Anteby, 2008), contexts where "the existence of guidelines but absence of enforcement" means that while certain behaviors are technically prohibited, they are tacitly permitted in practice (Zhang et al., 2023:229). Moral gray zones are prevalent in society. Indeed, politicians frequently violate formal policies without the expectation of being punished (Graffin, Bundy, Porac, Wade, & Quinn, 2013) and physicians consistently overprescribe dangerous drugs without fear of punishment (Zhang, Mohliver, & King, 2023). We suggest that police organizations similarly operate in a moral gray zone. Although police officers are rigorously trained and have knowledge of the formal rules surrounding appropriate officer conduct, informal norms or expectations to subvert those rules are regularly passed down to subordinate officers (Frake & Harmon, 2024), with formal punishment for engaging in misconduct being notoriously rare (Invisible Institute, 2018). In this study, we propose that punishment, if used in a moral gray zone such as this one, may have a different outcome than expected based upon traditional sociological work, and that new theory may be needed.

To build our argument, we integrate a growing body of work on misconduct in moral gray zones (Anteby, 2008; Zhang et al., 2023) with research on justice (Colquitt & Zipay, 2015; Conlon, Meyer, & Nowakowski, 2005) to theorize that punishing misconduct in these settings not only will fail to deter others, but will in fact produce even more misconduct. Specifically, we argue that because punishment in moral gray zones is historically absent or inconsistently applied, when punishment *is* used, it will be seen as arbitrary and unjust rather than a credible commitment to enforce the rules (Sheppard, Lewicki, & Minton, 1992). Building on work showing that perceptions of injustice can lead to resentment and moral outrage (Bies, 1987), prompting self-regulation failures (Mawritz, Greenbaum, Butts, & Graham, 2017) and even behavioral retaliation (Skarlicki & Folger, 1997), we argue that peers of the punished officer will engage in more (not less) misconduct after the punishment is levied, and that more severe punishments will further amplify this backlash effect. Taken together, this argument suggests that when punishment is used so inconsistently as to not be considered a serious threat, it may be difficult for organizations to start employing it as an effective deterrent.

We test our hypotheses using longitudinal data from the Chicago Police Department (CPD) between 1980 and 2017, which includes all misconduct complaints filed against police officers as well as any corresponding punishments imposed by the CPD. To identify the effects of punishment on the punished officer's peers, we exploit a lottery used by the CPD that assigns incoming officers to police academy cohorts. We leverage this quasi-random assignment to identify the peers of our focal officers who are later punished (Chappell & Lanza-Kaduce, 2010). Our findings reveal that when the CPD punishes a focal officer for misconduct, there is a subsequent increase in misconduct among their training cohort peers. The effects are amplified when the CPD imposes a more severe punishment. Supplementary analyses also reveal that the

types of misconduct peers engage in vary, from minor insubordination to excessive use of force, suggesting a more generalized reaction to injustice (Barclay, Skarlicki, & Pugh, 2005; Bies & Tripp, 2002).

PUNISHING MISCONDUCT IN MORAL GRAY ZONES

Punishment in Organizations

Classic sociological work views punishment as a key mechanism that enforces behavioral norms and rules for the broader population (Durkheim, 1895; Becker, 1963; Merton, 1938). Indeed, punishment operates as an effective deterrent against misconduct because it helps delineate which behaviors are acceptable (Weber, 1978; Meier & Johnson, 1977) and communicates to observers the expected costs should they violate established rules (Foucault, 1975; Goffman, 1961; Matsueda, Kreager, & Huizinga, 2006). Underlying the idea that punishment operates as an effective deterrent is thus a rational choice framework where individuals weigh the expected utility of a given behavior against the certainty and severity of its costs (Becker, 1968; Paternoster, 1987). As a result, when punishment is visible, it increases the perceived costs of rule-breaking and reduces the likelihood that others will engage in similar behaviors in the future (Bandura, 1977; Andenaes, 1974).

A large body of empirical work in criminal justice supports this argument and demonstrates that punishment generally operates as a deterrence mechanism for observers (Zimring et al., 1973; Stafford & Warr, 1993). Indeed, witnessing punishment has been shown to help peers learn what the appropriate rules are (Lee, 2017; Schnake, 1986) as well as decrease the likelihood that they violate those rules in the future (Klepper & Nagin, 1989; Simpson, 2002). Moreover, because the severity of punishment increases the costs of engaging in misconduct (Blumstein, Cohen, & Nagin, 1978; Pratt & Cullen, 2005), scholars have also found

that more severe punishment operates as a greater deterrent (Ross, 1984). Taken together, this research suggests that punishment increases the expected costs of misconduct and, thus, effectively deters others from engaging in similar behaviors (Nagin, 1998; Wilson & Herrnstein, 1998).

Although organization theorists have given comparatively less attention to punishment and its consequences, as recently noted by Greve and Teh (2016), the several studies that have examined this topic find support for this deterrence effect as well. Palmer and Yenkey (2015), for example, studied cyclists and their likelihood of engaging in illegal doping. They found that when a cyclist was punished for doping, their peers were significantly less likely to engage in illegal doping too. Similarly, Yiu and colleagues (2014) studied financial fraud in China and showed that firms were more deterred from engaging in fraud after observing their industry peers being punished, especially when they shared similar characteristics to the punished firm.

It is notable, however, that this work has largely investigated settings where punishment has historically been a clear and predictable threat, with consistent enforcement and a credible expectation that misconduct will be punished. When this is the case, we agree that it makes sense that punishment operates as a deterrent by establishing a real cost in observers' calculus of whether to engage in similar behaviors. However, a growing body of scholarship has started to show that not all settings embody such clear expectations around punishment and, instead, operate as moral gray zones.

Moral Gray Zones

Moral gray zones are settings where there is a disconnect between formal rules and consequences should those rules be violated. While formal regulations can exist and may even be widely known in moral gray zones, those who violate these regulations are rarely punished

(Zhang et al., 2023). This creates an environment in which certain behaviors, despite being technically forbidden, are tacitly permitted in practice and thus tend to persist because no real consequences tend to follow (Palmer, 2012; Ashforth & Anand, 2003; Vaughan, 1999; Anteby, 2008).

A growing body of work has started to empirically investigate moral gray zones and how misconduct tends to spread within them. For example, Earle, Spicer, and Peter (2010) studied the diffusion of the technically illegal practice of delaying wages to employees, arguing that its spread was driven in part by the fact that no one was ever punished. Pozner, Stimmler, and Hirsch (2010) similarly explored the spread of predatory lending during the 2008 financial crisis. They demonstrate that even though banks had formal policies prohibiting these practices, they nevertheless became normalized because regulatory penalties were rare and inconsequential. More recently, Zhang and colleagues (2023) showed that similar patterns emerge in healthcare settings, where physicians routinely overprescribe dangerous benzodiazepines because doctors have little to no fear of being punished for engaging in these technically prohibited behaviors.

Similar to these settings, we argue that police organizations also operate in a moral gray zone. Indeed, with low rates of punishment over the past several decades, the so-called “blue wall of silence” has encouraged officers to protect each other from reports of misconduct and prevent punishment (Benoit & Parry, 2021; Kalven, 2016; Rosenfield, Howey, & Mitchell, 2024). As a result, the majority of officers do not face punishment if they engage in misconduct (Torres & Curiel, 2023), even if it is documented on video or through multiple witness accounts (Greenhouse, 2020). Because of this, scholars have documented that police officers develop shared understandings around certain behaviors that, while technically prohibited, are acceptable

in practice (Crank, 2014; Sierra-Arévalo, 2019), and that these shared understandings have persisted over decades (Frake & Harmon, 2024).

Given the persistence of police misconduct, the societally relevant—and increasingly urgent—question has been how to mitigate it. Among the many proposed solutions, scholars (Rozema & Schanzenbach, 2023), practitioners (Stephens, 2011), and the Department of Justice (U.S. Department of Justice, 2017) alike have consistently called for more stringent punishment of offending officers. This recommendation seems logical and stems from traditional sociological theories on social deviance. As the argument goes, punishing police officers more should help reestablish and clarify the line between right and wrong as well as make clearer the costs of engaging in misconduct, thereby deterring other officers from engaging in those behaviors in the future. However, we argue that this logic falls apart in moral gray zones and that new theory may be needed.

Punishment in Moral Gray Zones

Our primary argument is that in the context of the moral gray zone of police organizations, punishment is more likely to be seen as arbitrary and unjust by peer officers, thereby prompting them to engage in more (not less) misconduct. Our theory also implies that more severe punishment, rather than increasing the costs of engaging in misconduct and deterring peer officers more, should instead amplify this sense of injustice and increase in subsequent peer misconduct.

To develop this argument, we turn to theories of justice (Colquitt & Zipay, 2015; Conlon et al., 2005) to explain why punishment in police organizations might be viewed by peers as unjust, prompting anger and even moral outrage. When police misconduct is widespread but its punishment has historically been inconsistent, the sudden use of punishment can appear

selective, arbitrary, and inconsistent (Emerson, 1981). Indeed, peer officers are likely to observe that the punished behavior is one that they and others have engaged in too, perhaps repeatedly and without consequence (Soltes, 2016; Dyck, Morse, & Zingales, 2021). This perceived arbitrariness can challenge the legitimacy of the police organization's authority (Sheppard et al., 1992), leading peer officers to focus less on the behaviors being punished and more on the injustice of this selective punishment (Greenberg, 1990; Folger, Folger, & Cropanzano, 1998). Thus, rather than clarifying and reestablishing norms, punishment in moral gray zones can actually give rise to the belief that punishment is arbitrary and that formal rules are meaningless (Kane & White, 2009; MacCoun, 1993). This sense of injustice among peers is likely to give rise to severe negative emotions, such as anger, moral outrage, and resentment (Cropanzano, Bowen, & Gilliland, 2007; Mitchell, Vogel, & Folger, 2015; Skarlicki & Kulik, 2005).

We argue that peers' sense of injustice and rise of negative emotions, in turn, will increase the likelihood that it will produce negative downstream consequences in their own behavior. In particular, scholars have shown that negative emotions can hinder one's ability to self-regulate and engage in socially appropriate behaviors (de Ridder, Lensvelt-Mulders, Finkenauer, Stok, & Baumeister, 2012; Mawritz et al., 2017), especially in professions where one's everyday job responsibilities require substantial self-control and restraint (e.g., police officers, service jobs, etc.) (Elfenbein, 2007). Moreover, the moral outrage and anger stemming from a sense of injustice can also lead peers to retaliate against the organization or those around them, intentionally "punishing the punisher" through harmful behaviors to exact revenge (Brebels, De Cremer, & Sedikides, 2008; Mitchell et al., 2015). Indeed, when individuals perceive that their organization has unfairly treated one of their peers, they are more likely to

engage in rule-breaking behavior (Skarlicki & Folger, 1997), even when such behaviors carry personal risk (O'Gorman, Wilson, & Miller, 2005).

This leads to our first prediction: punishing an officer for engaging in misconduct will not deter their peers from engaging in similar behaviors, but instead, will lead to a subsequent increase in their misconduct.

> **Hypothesis 1 (H1):** Punishing a police officer for misconduct will lead to a subsequent increase in misconduct among the punished officer's peers.

It follows from our theory that more severe punishment levied by police organizations should only exacerbate this outcome. In particular, if peers view the focal officer's punishment as unjust, then we should expect more severe punishment to be seen as *even more* unjust, further amplifying their subsequent engagement in misconduct. This prediction diverges from existing work that shows that increasing the severity of punishment increases a punishment's deterrence effect (Palmer & Yenkey, 2015; Yiu et al., 2014). Indeed, this work has argued that more severe punishments increase the salience and perceived costs of engaging in such behaviors, thereby deterring them even more (Nagin, 1998; Tittle, 1980). However, we suggest that in moral gray zones, more severe punishment will only further violate the principle of proportionality, exacerbating feelings of injustice (Adams, 1965; Thibaut & Walker, 1975), and thus amplify (rather than reduce) misconduct among peer officers (Bies & Tripp, 2002). This leads to our second prediction.

> **Hypothesis 2 (H2):** The positive relationship between officer punishment and subsequent peer misconduct will get stronger as the severity of the punishment increases.

EMPIRICAL SETTING AND DATA

We test our hypotheses in the Chicago Police Department (CPD), the second-largest police department in the US (Police Accountability Task Force, 2016). This is a setting where misconduct is an acute problem that can result in injury and even death for members of the broader public. Moreover, misconduct is financially and reputationally costly (Thomson-DeVeaux, Bronner, & Sharma, 2021), making the prevention or reduction of misconduct especially important. Police departments like the CPD employ a variety of approaches to reduce misconduct, including training programs, increased monitoring (e.g., body-worn cameras), and, in theory, punishment. However, punishment is rare. Of the quarter million complaints filed against CPD officers from 1988 – 2023, only eight percent resulted in punishment of any kind, and less than one percent resulted in a suspension or termination from 2000 – 2016 (Invisible Institute, 2018). Finally, this setting is particularly useful for studying the consequences of punishment because it offers access to detailed data on all misconduct complaints filed against all officers and any subsequent punishment levied. This level of granular detail allows us to capture increases in peer misconduct following an imposed punishment.

Punishing Misconduct

Following prior work (Frake & Harmon, 2024; Stroube, 2021), we first define misconduct using complaints filed by residents of Chicago that allege that a police officer's behaviors "transgress[ed] a line separating right from wrong" (Greve, Palmer, & Pozner, 2010: 56). Complaints encompass a wide array of misconduct behaviors, from minor insubordination to excessive use of force. The CPD is made aware of an officer's misconduct when a complaint is filed, which can occur by phone, mail, in person, or online. All misconduct complaints must be accompanied by a sworn affidavit from the complainant that the accusation is indeed factual.

To define and measure the punishment of misconduct, we rely on the detailed and well-documented process the CPD goes through upon receipt of a complaint. When a misconduct complaint is filed, a civilian agency within the CPD notifies the accused officer(s) and begins an investigatory process.[1] Investigators are responsible for collecting evidence, conducting interviews, and assessing preliminary guilt. If investigators find sufficient evidence to justify disciplinary action, they will designate the misconduct complaint as "sustained" and provide the CPD Superintendent a disciplinary recommendation. The Superintendent is not bound to follow the recommendation and has final authority to issue a punishment that is more or less severe than the independent recommendation. There are four disciplinary actions the Superintendent can impose: 1) violation noted (no disciplinary action), 2) reprimand (written and/or verbal), 3) suspension, and 4) termination. In this study, we operationalize punishment as any disciplinary action imposed by the Superintendent that results in at least a one-day suspension (which includes termination), as these punishments are readily observable to other officers.

Peers: The Police Academy

To evaluate the effect of punishing a police officer on their peers' subsequent misconduct, we must identify the punished officer's peer group. There are several approaches to defining a peer group, such as examining officers of the same rank or officers in the same unit. However, these peer-group definitions present serious threats to identifying a causal effect of punishment due to endogenous peer group formation. For example, if misconduct-prone officers select into a unit with others that are also prone to misconduct, then any observed correlation between an officer's punishment and a peer's subsequent misconduct may be driven by selection.

[1] The name of this agency has changed over the years. Prior to 2007, it was the Office of Professional Standards (OPS); from 2007 – 2017, it was the Independent Police Review Authority (IPRA); and ever since it has been the Civilian Office of Police Accountability (COPA).

Moreover, if peers are subject to the same environment (e.g., rank or unit), then unobserved factors (e.g., policies of the unit commander, unit culture, or local crime rate) may simultaneously affect the probability of an officer being punished and their peer receiving a misconduct complaint.

To address the selection problems that stem from endogenous peer-group formation, we follow prior studies (Holz, Rivera, & Ba, 2019; Frake & Harmon, 2024) that define an officer's peers as those who were assigned, based on a random lottery, to the same police academy cohort upon first entering the CPD. This process begins with the CPD making a call for applicants. If applicants pass an entrance exam and meet the minimum age and education requirements, they are assigned a random lottery number. As vacancies within the police force become available, the CPD invites qualified applicants—in order of their randomly assigned lottery number—to join a police academy training cohort. Because these training cohort members are assigned based on a random lottery number, academy-cohort peer assignments are as good as randomly assigned.[2] Figure 1 plots the size of each cohort between 1980 and 2017.

[Insert Figure 1 here]

In addition to this identification strategy diminishing selection concerns, the police academy is where recruits develop longstanding relationships with their peers. The academy consists of roughly six months of classroom training followed by six to twelve months of field assignment. During this period, recruits develop strong bonds with each other due to the extraordinary amount of time together, not only during academy hours, but outside of the academy as well, and instructors often reinforce tie formation by suggesting to academy

[2] Exceptions to this random assignment is that children of those who have died in active service and military veterans and are given priority.

members to break away from former ties and embrace work-peers as friends (Chappell & Lanza-Kaduce, 2010). These ties encourage an "intense sense of loyalty" (Doreian & Conti, 2017: 84) and are formative and long-lasting, with cohort-mates typically remaining close throughout the tenure of their career on the force (Van Maanen, 1973).

Data and Sample

Our data are sourced from the Invisible Institute.[3] The data include organizational unit assignment, police academy start date, misconduct complaints, as well as demographic, rank, and salary data on nearly all sworn members of the Chicago Police Department.[4] Our sample consists of officers who entered the police department between 1980 and 2017, as CPD does not have complete records of misconduct complaints before 1980. Following Holz and colleagues (2019), our unit of analysis is the officer-week, which is a small enough window to allow us to detect short-term behavioral reactions to punishment. This is consistent with prior work showing that reactions to injustice tend to manifest quickly or impulsively, whether intentional or not (Barclay et al., 2005; Baumeister, Vohs, & Tice, 2007; Bies & Tripp, 2002; Heatherton & Wagner, 2011). Our final sample consists of about 7.9 million officer-week observations of 10,976 officers who were members of 230 distinct police academy cohorts. In our sample, roughly 30% of officers receive a misconduct complaint. See Table 1 for summary statistics.[5]

[Insert Table 1 here]

Empirical Approach

Main Effect of Punishment on Peer Misconduct. We define the punished officers' peers as their police academy cohort members. To avoid simultaneity, spillover effects, and omitted

[3] See https://invisible.institute/police-data.
[4] Details for undercover officers and misconduct complaints involving minors are redacted.
[5] Our data can be found at https://github.com/invinst/chicago-police-data/tree/master/data.

variable bias that results from officers and peers being exposed to the same environment (Manski, 1993; Angrist, 2014), we exclude (1) observations while officers are still in the police academy and (2) officers who work in the same geographic unit as the punished officer. Using this definition of peers, our empirical approach thus compares officer *i* in unit *j*, whose peer from the police academy just received a punishment, with officers in the same unit (*j*) and week (*t*), but who did not have a peer from their police academy receive a punishment. Formally, we test Hypothesis 1 by estimating the effect of an officer's punishment on their peer's misconduct using the following regression equation:

$$PeerMisconduct_{itjk} = \beta_0 + \beta_1 Punishment_{ik,-j} + \lambda_i + \delta_{jt} + \varepsilon_{itjk} \quad (1)$$

where $PeerMisconduct_{itjk}$ is the number of misconduct complaints peer *i* received in week *t*. A concern with using complaints to measure misconduct is some police misconduct may go unreported, while other complaints might not be viewed by others, including the CPD, as misconduct (Stroube, 2021; Walker & Bumphus, 1992). We acknowledge measurement error on the dependent variable, but we see little reason for concern that this error will significantly bias our results, since it is unlikely to be correlated with the timing of the officer's punishment, conditional on individual and unit-week fixed effects.

The term λ_i represents officer fixed effects, which account for time-invariant differences in officers' propensity to engage in or receive a misconduct complaint, such as prior experience, personality, and demographic characteristics. The term δ_{jt} represents unit-week fixed effects, which accounts for unit-week level shocks that may influence the propensity for officers to engage in misconduct or for the public to report misconduct.

The treatment, $Punishment_{ik,-j}$, is an indicator equal to one if a member of peer *i*'s academy cohort, *k*, received a punishment (suspension or termination) in the previous week, but

is not working in the same unit as the punished officer (denoted by *-j*). We only include the first time an officer was punished to avoid using the same observation to estimate both pre-treatment and post-treatment effects. Thus, the coefficient β_1 estimates the effect of being a peer of a punished officer relative to other officers in the same unit and week who were not peers of a punished officer. The terms β_0 and ε_{itjk} represent the intercept and a zero-mean error term, respectively. In our primary estimates, we use OLS regression with standard errors clustered at the cohort level because that is the level at which treatment (i.e., being a peer of the punished officer) is assigned (Abadie, Athey, Imbens, & Wooldridge, 2017; Bertrand, Duflo, & Mullainathan, 2004). Figure 4 demonstrates that our main results are robust to alternative specifications.

The key identifying (parallel trends) assumption for Equation 1 is that changes in the average misconduct complaints of officers in the same unit and week are equal to those that treated officers (peers of a punished officer) would have experienced in the absence of treatment. There are several plausible threats to this assumption. First, this assumption will be violated if there is treatment leakage, or if control units are also affected by the punishments of the focal officer. To minimize this violation, we exclude all observations in which the punished officer and their peers are in the same unit. Despite this, it is possible that information about a non-cohort member being punished in a different unit affects officers in the control group (i.e., unit members who were not in the academy cohort with the punished officer). Under the assumption that the effect of punishment on non-academy-cohort members is weaker and in the same direction as on academy cohort peers, our results will be biased towards zero. Second, there could also be anticipation effects, whereby treated officers (i.e., peers) receive and react to information about a punishment before the punishment is issued. Assuming the anticipatory

effects are in the same direction as the effect of a punishment, our estimates will again be biased towards zero. Thus, we interpret our estimates as a lower bound.

Moreover, we empirically investigate the likelihood of anticipation effects and the plausibility of the parallel trends assumption by estimating the effect of punishment in the periods *preceding* that punishment. The results in Figure 2 demonstrate that the pre-treatment estimates are small and not significantly different from zero. Additionally, there is no visible trend in the pre-treatment estimates, suggesting that the parallel trends assumption is reasonable. This alleviates concerns that the peers of officers who were punished in a given week differ from those who were not peers of punished officers that week.

[Insert Figure 2 here]

Heterogeneous Effects by Punishment Severity. To test Hypothesis 2, which predicts that the effects of punishment are amplified for more severe punishments, we must measure the severity of punishment. One approach would be to use the number of days the officer was suspended. However, this measurement depends on the nature of the underlying misconduct. For example, a five-day suspension for a minor uniform infraction (e.g., failure to tuck a shirt into pants) may be perceived as severe, while the same punishment for a major infraction (e.g., excessive use of force that resulted in a citizen's death) may be perceived as lenient. Thus, it is important that the measure captures the severity of the punishment relative to the severity of the act being punished. To do this, we replace the dichotomous independent variable, $Punishment_{ik\text{-}j,}$ in Equation 1 with a continuous measure of the punishment's severity using the following equation:

$$PunishmentSeverity_{ik,\text{-}j} = \frac{(EnactedSuspension_{ik,-j} - RecommendedSuspension_{ik,-j})}{RecommendedSuspension_{ik,-j}} \quad (2)$$

where $EnactedSuspension_{ik,-j}$ represents the number of days the officer in cohort k was suspended by the CPD Superintendent. $RecommendedSuspension_{ik,-j}$ is equal to the number of days of suspension recommended by the independent investigatory agency. This measure assumes that conditional on the agency recommending a given suspension period, longer suspensions will be perceived by officers as more severe, while shorter suspensions will be perceived as less severe. By dividing by $RecommendedSuspension_{ik,-j}$, we assume that one extra day of *Enacted Suspension* will be perceived as more severe if the *Recommended Suspension* were shorter.

One challenge with the measure of punishment severity presented in Equation 2 is that the suspension length for terminations is undefined. A second challenge is that the operationalization in Equation 2 may be driven by outliers as it enforces a linear relationship when, in fact, the relationship may be non-linear. To address these, we also develop categorical measures of punishment severity as follows:

$$LenientPunishment_{k,-j} = \begin{cases} 1, \textit{ if } EnactedPunishment_{k,-j} = RecommendedPunishment_{k,-j} \\ 0, \textit{ otherwise} \end{cases} \quad (3)$$

$$NeutralPunishment_{k,-j} = \begin{cases} 1, \textit{ if } EnactedPunishment_{k,-j} = RecommendedPunishment_{k,-j} \\ 0, \textit{ otherwise} \end{cases} \quad (4)$$

$$SeverePunishment_{k,-j} = \begin{cases} 1, \textit{ if } EnactedPunishment_{k,-j} = RecommendedPunishment_{k,-j} \\ 0, \textit{ otherwise} \end{cases} \quad (5)$$

Thus, severe (lenient) punishments are defined as those where the enacted punishment was more (less) severe than what was recommended by the investigators. Punishments are

considered neutral if the Superintendent enacted the punishment recommended by the investigators. We classify all terminations as severe punishment.

RESULTS

Main Effects

To test the effect of an officer's punishment on their peers' subsequent misconduct, we estimate Equation 1 in Table 2. Model 1 includes officer and year-week fixed effects. Model 2 introduces unit fixed effects. Model 3 includes unit-week (interacted) fixed effects. The estimates across all models are similar in effect size and significance. The results from Model 3 suggest that an officer's *punishment* increases the likelihood of *peer misconduct* in the following week by about 35% relative to the baseline risk of receiving a complaint. Thus, we find strong support for Hypothesis 1.[6]

[Insert Table 2 here]

Next, we explore how long this main effect on peer misconduct persists. Behavioral reactions to injustice tend to be relatively short-lived (Barclay et al., 2005; Baumeister et al., 2007; Bies & Tripp, 2002; Heatherton & Wagner, 2011). As such, our theory would predict that this effect should dissipate relatively quickly. To explore this, we re-estimate Equation 1, but replace *punishment* with a series of leads and lags five weeks before and after punishment. The results, plotted in Figure 2, demonstrate that the effects of *punishment* on *peer misconduct* is

[6] Recent work has shown that the two-way fixed effects approach to estimating difference-in-differences and event studies with staggered treatments and dynamic effects may result in biased estimates due to the implicit comparisons between later-treated and earlier-treated units (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021; de Chaisemartin & D'Haultfoeuille, 2022; Borusyak, Jaravel, & Spiess, 2022). Though we do not expect these issues to significantly affect our estimates due to the expected short-term duration of the effects, we nevertheless replicate our results using the imputation methods developed by Borusyak et al. (2022) in Models 4 – 6. The results are similar to those in Models 1 through 3 with similar point estimates and levels of significance. Thus, we conclude that the two-way fixed effects estimates are unlikely to be significantly biased. We do not use Borusyak et al. (2022) estimator to test the results of Hypothesis 2 because it does not allow for continuous treatments or interactions.

indeed relatively short-lived. The effect diminishes to insignificance just two weeks after the punishment was imposed and remains small and insignificant thereafter.

Heterogeneous Effects

We now turn to Hypothesis 2, which predicted that the effects of punishment on peer misconduct become stronger for more severe punishments. To test this hypothesis, we re-estimate Equation 1 but replace *Punishment* with a continuous variable, *Punishment Severity*, equal to the severity of the punishment relative to the punishment recommended by investigators (see Equation 2). The results from Models 1-3 provide support for the notion that more severe punishments increase peers' subsequent misconduct more than less severe punishments.

[Insert Table 3 here]

We also tested Hypothesis 2 using the three categorical variables, which represent punishments that are lenient, neutral, and severe (see Equations 3–5), thus allowing us to explore asymmetric effects that cannot be captured by the prior continuous measure. The results in Models 4–6 suggest that relative to non-peers in the same unit and cohort, peers of officers who received a lenient punishment did not receive misconduct complaints at significantly different rates. However, peers of officers who received a neutral or severe punishment had significantly increased rates of misconduct complaints in the following week. The marginal effects are plotted in Figure 3. We note that the difference in effects of lenient and severe punishments are statistically different from one another, but that the estimated effect of a *neutral punishment* is not statistically different from either a *lenient punishment* or *severe punishment.* Taken together, these results support Hypothesis 2.

[Insert Figure 3 here]

Exploring Types of Peer Misconduct

To better understand the nature of this effect, we next explore the types of misconduct peers are engaging in after the punishment of their coworker has been imposed. One way to disaggregate misconduct types is to identify the target of the behaviors. To do so, we use the categorization by the Invisible Institute, which classifies misconduct complaints as either internal or external. Internal complaints relate to misconduct that is targeted toward the CPD, such as being absent or late to duty or insubordination. External complaints are those directed toward the public and include excessive use of force, illegal search and seizure, false arrest, and so on. When rerunning Models 1 and 2 of Table 4, we find that punishment increases both internal and external peer misconduct.

[Insert Table 4 here]

We then further disaggregate complaints into their specific categories to explore whether the increase in peer misconduct post-punishment occurs mostly in specific categories or represents a more generalized increase in misconduct. Models 3-6 disaggregate misconduct complaints into the four most prevalent categories. All show large increases in peer misconduct following a punishment. Taken together, these results suggest a generalized response to punishment among peers, which is in line with prior work showing that reactions to injustice are often generalized in nature as well as direction (Barclay et al., 2005; Bies & Tripp, 2002).

Alternative Explanation

A possible alternative explanation for our results could be that peers who learn of a punishment perceive it as relatively minor, leading them to engage in more misconduct because they view the costs as insignificant. While possible, two pieces of evidence are inconsistent with this interpretation. First, the duration of the main effect is quite short. If the effect was driven by

peers updating their perceived costs of misconduct, we would expect the effect to remain constant over time. Second, punishment severity amplified the main effect. If this effect reflected peers' information updating, we would have expected more severe punishments to *reduce* future misconduct, not amplify it. Taken together, our findings are more consistent with the theory that peers are reacting with a sense of injustice rather than performing a cost-benefit analysis of such behaviors.

Robustness

Finally, we test the robustness of our main results using a variety of alternative specifications, including OLS (with a count-measure of misconduct), linear probability models (LPM) using a dichotomous measure of misconduct, Poisson regression, and staggered DD imputation methods (Borusyak, Jaravel, & Spiess, 2022). We also test robustness with a variety of fixed effects and standard error estimators. Figure 4 plots the coefficients and 95% confidence intervals from each regression. The results demonstrate that all alternative estimations are positive and significant. Estimates that lack year-week fixed effects are significantly larger than our preferred estimate, while Poisson estimates are generally smaller. Our preferred specification falls around the middle of the distribution.

[Insert Figure 4 here]

DISCUSSION

This study examines what happens when organizations attempt to use punishment in moral gray zones. Deterrence theory holds that punishment reduces misconduct by clearly signaling behavioral boundaries and by making violations costly (Durkheim, 1938; Zimring et al., 1973). This logic, however, rests on an important assumption: that punishment is predictable and consistently enforced. In many organizations, this is not the case. Rules may be explicit on

paper, yet enforcement is inconsistent or absent in practice (Anteby, 2008), and this creates a moral gray zone in which technically prohibited behaviors become informally tolerated. Even so, organizations sometimes turn to punishment in an effort to reassert control. Yet we know little about the consequences of wielding punishment in such contexts. By theorizing how punishment operates under these conditions, this study advances emerging research on misconduct in moral gray zones, answers recent calls to examine the broader organizational consequences of punishment, and speaks to an urgent challenge in addressing the persistent social problem of police misconduct.

Punishing Misconduct in Moral Gray Zones

While misconduct itself has been widely studied, the organizational consequences of punishing misconduct have received far less attention. As Greve and Teh (2016) note, punishment is generally assumed to follow misconduct—and to be effective in reducing it—so its consequences have often been taken for granted in organizational research. Our study deepens this conversation in several important ways.

First, we demonstrate when punishment can backfire, increasing the very behaviors it is meant to prevent. Classic deterrence theory asserts that punishment works by clearly signaling organizational norms and by making violations costly (Becker, 1968; Merton, 1938). However, the limited organizational research examining punishment operates in settings characterized by stable norms and consistent rule enforcement (e.g., Palmer & Yenkey, 2015; Yiu et al., 2014). Many organizations, however, lack these conditions. We probe punishment in a moral gray zone, where rules are formally clear but enforcement has long been inconsistent. Contrary to deterrence predictions, we find that when an officer is punished for misconduct, misconduct among the punished officer's peers increases. This finding challenges the presumed universality

of deterrence theory and shows that in ambiguous normative environments, punishment can catalyze the behaviors it aims to suppress.

Second, we theorize why punishment backfires in moral gray zones by integrating insights from organizational justice research. Existing work largely assumes that observers weigh the prospective costs of punishment against the benefits of misconduct (Stafford & Warr, 1993; Bandura, 1977; Paternoster, 1987). We argue instead that when punishment is introduced into a moral gray zone, it is experienced as unjust because the punished behaviors, while technically prohibited, have been historically tolerated by the organization. Under these conditions, punishment is interpreted not as norm enforcement but as arbitrary and illegitimate (Sheppard et al., 1992), provoking negative emotional reactions that increase misconduct rather than deter it. Consistent with this account, we find a short-term rise in peer misconduct following punishment, and more severe punishment amplifies this effect. This extends deterrence theory by showing that when normative clarity is weak, affective responses to perceived injustice—rather than rational risk assessments—become primary drivers of behavioral reactions.

Third, our findings also refine research on third-party reactions to punishment. Prior work has shown that observing unfair punishment can lead to disengagement, such as reduced job satisfaction or withdrawal (Aryee & Chay, 2001; Brockner, 2002; Dailey & Kirk, 1992). These consequences still align with a broader deterrence logic because they do not directly increase misconduct. Our findings reveal a more volatile dynamic: peers respond by escalating misconduct. This suggests that perceived injustice can produce generalized behavioral retaliation that extends beyond counterproductive work behaviors (i.e., Skarlicki & Folger, 1997; Skarlicki, Van Jaarsveld, & Walker, 2008; Jones, 2009) to misconduct itself. We also note that policing involves high emotional and physical demands (Ashforth & Humphrey, 1993; Grandey &

Gabriel, 2015), which may intensify negative affect and make emotional regulation more difficult (Wharton, 2009). Future research might directly examine these emotional processes by tracing how perceptions of illegitimacy give rise to anger, resentment, or defiance, and how such affective states evolve into misconduct. Doing so would clarify the mechanisms through which punishment backfires and identify when emotional escalation is most likely to occur.

Finally, this study highlights the broader challenge organizations face when attempting to reassert control in long-standing moral gray zones. Prior research has shown how weak or absent enforcement enables misconduct to spread (e.g., Earle et al., 2010; Zhang et al., 2023). We instead show what happens when organizations belatedly attempt to impose discipline. Rather than signaling renewed commitment to rule enforcement, punishment may confirm the belief that the system is arbitrary. Our findings suggest that once moral gray zones are established, they become self-reinforcing moral orders in which punishment no longer restores norms but instead threatens the shared understandings that have sustained everyday work. Under these conditions, efforts to curb misconduct may paradoxically intensify it.

Taken together, our study thus reframes deterrence as a context-dependent process that relies on a particular moral infrastructure—one in which rules are not only explicit but also consistently enacted and socially legitimate. In the absence of that infrastructure, punishment can lose its normative meaning and provoke the very actions it is intended to prevent. This reorientation positions moral gray zones not simply as environments of weak enforcement, but as distinct organizational moral systems that shape how control efforts are interpreted and enacted. Recognizing this distinction is essential for understanding why misconduct persists—and why attempts to eliminate it often fail—in many contemporary organizations.

Practical Implications

Despite these challenges, our findings point to several possible strategies for managing misconduct in general and for reform efforts within police organizations in particular. First, given how difficult it is to (re)establish control over misconduct in moral gray zones, leaders must be strategic in reintroducing punishment. One approach may involve beginning with more lenient forms of discipline. In our data, such punishments did not provoke the same increase in peer misconduct as more severe sanctions. Whether leniency alone is sufficient to deter future violations remains an open question, however. Another important consideration is avoiding ad hoc or performative punishments that may be perceived as symbolic rather than substantive. Inconsistently applied discipline—especially when it appears to target specific individuals—undermines the legitimacy of enforcement efforts and deepens perceptions of injustice (Sheppard et al., 1992; Kane & White, 2009). Instead, organizations should aim to apply punishment systematically and transparently, allowing members to recognize it as part of a credible and fair accountability regime.

Second, timing and messaging may play a critical role in shaping how punishment is received. For organizations attempting to transition out of a moral gray zone, openly acknowledging past inconsistencies in enforcement and clearly articulating new expectations may help reduce the backlash that punishment can provoke. Transparency around why a punishment is being imposed—and how it fits within broader reform efforts—can decrease the likelihood that it is interpreted as arbitrary or politically motivated (Brockner, 2002; Briscoe & Murphy, 2012). Prior research shows that officers who view punishment procedures as fair are less likely to engage in future misconduct (Wolfe & Piquero, 2011). Our findings are consistent with the idea that enhancing the perceived legitimacy of punishment—through fair procedures,

consistent application, and clear communication—can temper negative emotional responses and limit the spread of misconduct.

Third, our findings raise broader system-wide concerns about how the efficacy of punishment is typically evaluated. Punishment is often assessed based on its effect on the individual who receives it, yet our study shows that its consequences can ripple outward to peers and potentially to the wider organization. We measured peer relationships using shared police academy cohorts—a quasi-random grouping that allowed us to isolate the effect of punishment on close observers. These relationships are meaningful and long-lasting (Doreian & Conti, 2017; Van Maanen, 1973). However, similar dynamics may arise among peers who share stronger bonds of identification, such as unit mates or beat partners. If so, our estimates likely understate the broader organizational effects of punishment. Existing research suggests that peers who share background, identity, or role-based solidarity can react strongly to perceived injustice (Tajfel, 1982; Brockner & Greenberg, 2015; Mollica, 1999; Spencer & Rupp, 2009). Understanding which organizational members are most sensitive to perceived injustice—and under what conditions—is therefore a critical area for future research.

Finally, these dynamics complicate how organizations assess the effectiveness of punishment. Even if punishment successfully deters the focal individual, it may unintentionally increase misconduct among peers. In our context, for the punishment of a single officer to produce a net reduction in misconduct, it would need to prevent that officer from engaging in any misconduct for roughly 16 weeks just to counteract the short-term increase among their peers. Yet prior work suggests that even the direct deterrent effects of punishment are not reliably consistent (Harris & Worden, 2014). Evaluating the organizational impact of punishment, then, requires tracking not only the behavior of the punished individual but also the

networked spillover effects that punishment can generate. Doing so will require richer data, finer-grained measurements of interpersonal identification, and greater attention to how perceived justice travels through organizational relationships.

Conclusion

This study challenges a foundational assumption about how punishment maintains social order. Across both scholarly research and public discourse, punishment is often treated as the primary lever for curbing misconduct—an especially prominent view in current debates over police reform. The underlying logic is long-standing and intuitive: rules matter because violations are met with consequences, and those consequences deter future misconduct and reaffirm collective norms (Durkheim, 1893; Weber, 1922; Foucault, 1975). Our findings reveal a consequential boundary condition to this logic. In moral gray zones—contexts where formal rules exist but have long gone unenforced—punishment is not interpreted as legitimate rule enforcement. Instead, it is filtered through a sense of injustice, provoking backlash that increases misconduct among peers. This dynamic is not unique to policing. Many institutions, from regulatory agencies to healthcare systems, operate in similar environments where informal practices have become entrenched and enforcement is inconsistent. In such settings, importing deterrence-based punishment strategies from high-enforcement environments is unlikely to generate the desired effects. Efforts to reduce misconduct must instead begin by restoring the legitimacy and consistency of enforcement practices themselves. Only once punishment is perceived as fair and credible can it serve its intended role in reinforcing behavioral boundaries. Until then, punishment may not simply fail to deter misconduct—it may deepen it. The broader implication is clear: punishment can stabilize organizations only when the moral infrastructure that gives it meaning is intact.

REFERENCES

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. 2017. ***When Should You Adjust Standard Errors for Clustering?*** w24003. no. w24003, Cambridge, MA: National Bureau of Economic Research.

Adams, J. S. 1965. Inequity In Social Exchange. In L. Berkowitz (Ed.), ***Advances in Experimental Social Psychology***, vol. 2: 267–299. Academic Press.

Andenaes, J. 1974. ***Punishment and deterrence***. University of Michigan Press Ann Arbor. https://www.ojp.gov/ncjrs/virtual-library/abstracts/punishment-and-deterrence.

Angrist, J. D. 2014. The perils of peer effects. ***Labour Economics***, 30: 98–108.

Anteby, M. 2008. ***Moral Gray Zones: Side Productions, Identity, and Regulation in an Aeronautic Plant***. Princeton University Press. https://doi.org/10.1515/9781400828883.

Aryee, S., & Chay, Y. W. 2001. Workplace justice, citizenship behavior, and turnover intentions in a union context: Examining the mediating role of perceived union support and union instrumentality. ***Journal of Applied Psychology***, 86(1): 154–160.

Ashforth, B. E., & Anand, V. 2003. The Normalization of Corruption in Organizations. ***Research in Organizational Behavior***, 25: 1–52.

Ashforth, B. E., & Humphrey, R. H. 1993. Emotional Labor in Service Roles: The Influence of Identity. ***The Academy of Management Review***, 18(1): 88–115.

Bandura, A. 1977. ***Social Learning Theory***. Englewood Cliffs, NJ: Prentice-Hall. https://www.asecib.ase.ro/mps/Bandura_SocialLearningTheory.pdf.

Barclay, L., Skarlicki, D., & Pugh, S. 2005. Exploring the Role of Emotions in Injustice Perceptions and Retaliation. ***The Journal of Applied Psychology***, 90: 629–43.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. 2007. The Strength Model of Self-Control. ***Current Directions in Psychological Science***, 16(6): 351–355.

Becker, G. S. 1968. Crime and Punishment: An Economic Approach. ***Journal of Political Economy***, 76(2): 169–217.

Becker, H. 1963. Outsiders. ***Studies in the Sociology of Deviance***.

Benoit, J.-P., & Parry, R. 2021. What the blue wall of silence means for police reform. ***London Business School***. https://www.london.edu/think/what-the-blue-wall-of-silence-means-for-police-reform.

Bertrand, M., Duflo, E., & Mullainathan, S. 2004. How much should we trust differences-in-differences estimates? ***Quarterly Journal of Economics***, 27.

Bies, R. J. 1987. The predicament of injustice: The management of moral outrage. ***Research in Organizational Behavior***, 9: 289–319.

Bies, R., & Tripp, T. 2002. Hot flashes, open wounds: Injustice and the tyranny of its emotions. ***Emerging Perspectives on Managing Organizational Justice***, 203–221.

Blumstein, A., Cohen, J., & Nagin, D. 1978. ***Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates***. National Academy of Sciences Washington, DC. https://www.ojp.gov/ncjrs/virtual-library/abstracts/deterrence-and-incapacitation-estimating-effect-criminal-sanctions.

Borusyak, K., Jaravel, X., & Spiess, J. 2022. Revisiting Event Study Designs: Robust and Efficient Estimation. ***arXiv:2108.12419 [Econ]***. http://arxiv.org/abs/2108.12419.

Brebels, L., De Cremer, D., & Sedikides, C. 2008. Retaliation as a response to procedural unfairness: A self-regulatory approach. ***Journal of Personality and Social Psychology***, 95(6): 1511–1525.

Briscoe, F., & Murphy, C. 2012. Sleight of Hand? Practice Opacity, Third-party Responses, and the Interorganizational Diffusion of Controversial Practices. ***Administrative Science Quarterly***, 57(4): 553–584.

Brockner, J. 2002. Making Sense of Procedural Fairness: How High Procedural Fairness Can Reduce or Heighten the Influence of Outcome Favorability. ***Academy of Management Review***, 27(1): 20.

Brockner, J., & Greenberg, J. 2015. The impact of layoffs on survivors: An organizational justice perspective. ***Applied Social Psychology and Organizational Settings***, 6: 45–75.

Brockner, J., Grover, S., Reed, T., DeWitt, R., & O'Malley, M. 1987. Survivors' Reactions to Layoffs: We Get by with a Little Help for Our Friends. ***Administrative Science Quarterly***, 32(4): 526–541.

Callaway, B., & Sant'Anna, P. H. C. 2021. Difference-in-Differences with multiple time periods. ***Journal of Econometrics***, 225(2): 200–230.

Center for Disease Control. 2024. ***Facts and Figures on Injuries Caused by Law Enforcement | Law Enforcement Epidemiology Project | University of Illinois Chicago***. Center for Disease Control. https://policeepi.uic.edu/data-civilian-injuries-law-enforcement/facts-figures-injuries-caused-law-enforcement/.

Chappell, A. T., & Lanza-Kaduce, L. 2010. Police Academy Socialization: Understanding the Lessons Learned in a Paramilitary-Bureaucratic Organization. ***Journal of Contemporary Ethnography***, 39(2): 187–214.

Colquitt, J. A., & Zipay, K. P. 2015. Justice, Fairness, and Employee Reactions. ***Annual Review of Organizational Psychology and Organizational Behavior***, 2(1): 75–99.

Conlon, D. E., Meyer, C. J., & Nowakowski, J. M. 2005. How Does Organizational Justice Affect Performance, Withdrawal, and Counterproductive Behavior? In J. Greenberg & J. A. Colquitt (Eds.), ***Handbook of Organizational Justice*** (0 ed.): 317–343. Psychology Press.

Crank, J. P. 2014. ***Understanding police culture***. Routledge. https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9781315721255&type=googlepdf.

Cropanzano, R., Bowen, D. E., & Gilliland, S. W. 2007. The Management of Organizational Justice. ***Academy of Management Perspectives***, 21(4): 34–48.

Dailey, R. C., & Kirk, D. J. 1992. Distributive and Procedural Justice as Antecedents of Job Dissatisfaction and Intent to Turnover. ***Human Relations***, 45(3): 305–317.

de Chaisemartin, C., & D'Haultfoeuille, X. 2022. Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey. ***NBER Working Paper Series***, 30.

de Ridder, D. T. D., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. 2012. Taking Stock of Self-Control: A Meta-Analysis of How Trait Self-Control Relates to a Wide Range of Behaviors. ***Personality and Social Psychology Review***, 16(1): 76–99.

Doreian, P., & Conti, N. 2017. Creating the thin blue line: Social network evolution within a police academy. ***Social Networks***, 50: 83–97.

Durkheim, E. 1895. The rules of sociological method. ***Social theory re-wired***: 9–14. Routledge.

Durkheim, E. 1938. ***The Rules of Sociological Method... Translated by Sarah A. Solovay and John H. Mueller, and Edited by George EG Catlin***. (S. G. E. G. Catlin & S. A. Solovay, Trans.).

Dyck, I. J. A., Morse, A., & Zingales, L. 2021, February 17. ***How Pervasive is Corporate Fraud?*** Rochester, NY. https://doi.org/10.2139/ssrn.2222608.

Earle, J. S., Spicer, A., & Peter, K. S. 2010. The Normalization of Deviant Organizational Practices: Wage Arrears in Russia, 1991–98. ***The Academy of Management Journal***, 53(2): 218–237.

Elfenbein, H. A. 2007. Emotion in Organizations. ***Academy of Management Annals***, 1(1): 315–386.

Emerson, R. M. 1981. On Last Resorts. ***American Journal of Sociology***, 87(1): 1–22.

Folger, R. G., Folger, R., & Cropanzano, R. 1998. ***Organizational justice and human resource management***, vol. 7. Sage. https://books.google.com/books?hl=en&lr=&id=-lAMyIsLLW0C&oi=fnd&pg=PR9&dq=Folger,+R.,+%26+Cropanzano,+R.+(1998).+Organizational+justice+and+human+resource+management.+Sage+Publications.&ots=TVouOF0LOx&sig=Cs6rg8GG_v2hqBlxegz9QHgCuGo.

Foucault, M. 1975. Discipline and Punish. ***Social Theory Re-Wired*** (3rd ed.). Routledge.

Frake, J., & Harmon, D. 2024. Intergenerational Transmission of Organizational Misconduct: Evidence from the Chicago Police Department. ***Management Science***, 70(6): 3856–3878.

Goffman, E. 1961. ***Asylums***. Anchor Books New York.

Goldman, B. M. 2001. Toward an Understanding of Employment Discrimination Claiming: An Integration of Organizational Justice and Social Information Processing Theories. ***Personnel Psychology***, 54(2): 361–386.

Graffin, S. D., Bundy, J., Porac, J. F., Wade, J. B., & Quinn, D. P. 2013. Falls from Grace and the Hazards of High Status: The 2009 British MP Expense Scandal and Its Impact on Parliamentary Elites. ***Administrative Science Quarterly***, 58(3): 313–345.

Grandey, A., & Gabriel, A. 2015. Emotional Labor at a Crossroads: Where Do We Go from Here? ***Annual Review of Organizational Psychology and Organizational Behavior***, 2.

Greenberg, J. 1990. Organizational Justice: Yesterday, Today, and Tomorrow. ***Journal of Management***.

Greenhouse, S. 2020, June 18. How Police Unions Enable and Conceal Abuses of Power. ***The New Yorker***. https://www.newyorker.com/news/news-desk/how-police-union-power-helped-increase-abuses.

Greve, H. R., Palmer, D., & Pozner, J. 2010. Organizations Gone Wild: The Causes, Processes, and Consequences of Organizational Misconduct. ***Academy of Management Annals***, 4(1): 53–107.

Greve, H. R., & Teh, D. 2016. Consequences of organizational misconduct: Too much and too little punishment. In D. Palmer, K. Smith-Crowe, & R. Greenwood (Eds.), ***Organizational Wrongdoing***: 370–403. Cambridge: Cambridge University Press.

Harris, C. J., & Worden, R. E. 2014. The Effect of Sanctions on Police Misconduct. ***Crime & Delinquency***, 60(8): 1258–1288.

Heatherton, T. F., & Wagner, D. D. 2011. Cognitive neuroscience of self-regulation failure. ***Trends in Cognitive Sciences***, 15(3): 132–139.

Holz, J., Rivera, R., & Ba, B. A. 2019. Spillover Effects in Police Use of Force. ***SSRN Electronic Journal***. https://doi.org/10.2139/ssrn.3519968.

Invisible Institute. 2018, August 16. Massive Expansion of Police Misconduct Data Tool in Chicago. ***The Intercept***. https://invisible.institute/press-release/massive-expansion-of-police-misconduct-data-tool-in-chicago.

Jones, D. A. 2009. Getting even with one's supervisor and one's organization: Relationships among types of injustice, desires for revenge, and counterproductive work behaviors. ***Journal of Organizational Behavior***, 30(4): 525–542.

Kalven, J. 2016, October 6. Operation Smoke and Mirrors: In the Chicago Police Department, if the Bosses Say it Didn't Happen, it Didn't Happen. ***The Intercept***, 39.

Kane, R. J., & White, M. D. 2009. Bad cops. ***Criminology & Public Policy***, 8(4): 737–769.

Klepper, S., & Nagin, D. 1989. Tax Compliance and Perceptions of the Risks of Detention and Criminal Prosecution Paper. ***Law & Society Review***, 23(2): 209–240.

Lee, H.-W. 2017. Taking Deterrence Seriously: The Wide-Scope Deterrence Theory of Punishment. ***Criminal Justice Ethics***, 36(1): 2–24.

MacCoun, R. J. 1993. Drugs and the law: A psychological analysis of drug prohibition. ***Psychological Bulletin***, 113(3): 497–512.

Manski, C. F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. ***The Review of Economic Studies***, 60(3): 531.

Matsueda, R. L., Kreager, D. A., & Huizinga, D. 2006. Deterring Delinquents: A Rational Choice Model of Theft and Violence. ***American Sociological Review***, 71(1): 95–122.

Mawritz, M. B., Greenbaum, R. L., Butts, M. M., & Graham, K. A. 2017. I Just Can't Control Myself: A Self-Regulation Perspective on the Abuse of Deviant Employees. ***Academy of Management Journal***, 60(4): 1482–1503.

Meier, R. F., & Johnson, W. T. 1977. Deterrence as Social Control: The Legal and Extralegal Production of Conformity. ***American Sociological Review***, 42(2): 292.

Merton, R. K. 1938. Social structure and anomie. ***American Sociological Review***, 3(5): 672–682.

Mitchell, M. S., Vogel, R. M., & Folger, R. 2015. Third parties' reactions to the abusive supervision of coworkers. ***Journal of Applied Psychology***, 100(4): 1040–1055.

Mollica, K. A. 1999. A social identity perspective on organizational justice among layoff survivors. ***Academy of Management Proceedings***, 1999(1): B1–B6.

Nagin, D. S. 1998. Deterrence and Incapacitation. In M. Tonry (Ed.), ***The Handbook of Crime and Punishment***: 24. Oxford University Press, Incorporated.

O'Gorman, R., Wilson, D. S., & Miller, R. R. 2005. Altruistic punishing and helping differ in sensitivity to relatedness, friendship, and future interactions. ***Evolution and Human Behavior***, 26(5): 375–387.

O'Reilly, J., & Aquino, K. 2011. A Model of Third Parties' Morally Motivated Responses to Mistreatment in Organizations. ***Academy of Management Review***, 19.

Palmer, D. 2012. ***Normal Organizational Wrongdoing: A Critical Analysis of Theories of Misconduct in and by Organizations***. OUP Oxford.

Palmer, D., & Yenkey, C. B. 2015. Drugs, Sweat, and Gears: An Organizational Analysis of Performance-Enhancing Drug Use in the 2010 Tour de France. ***Social Forces***, 94(2): 891–922.

Paternoster, R. 1987. The deterrent effect of the perceived certainty and severity of punishment: A review of the evidence and issues. ***Justice Quarterly***, 4(2): 173–217.

Pew Research Center. 2022. ***Americans' Trust in Scientists, Other Groups Declines***. Pew Research Center. https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/.

Police Accountability Task Force. 2016. ***Recommendations for reform: Restoring trust between the Chicago Police and the communities they serve***. Chicago, IL: Police Accountability Task Force.

Pozner, J.-E., Stimmler, M. K., & Hirsch, P. M. 2010. Terminal isomorphism and the self-destructive potential of success: Lessons from subprime mortgage origination and securitization. In M. Lounsbury & P. M. Hirsch (Eds.), ***Research in the Sociology of Organizations***, vol. 30: 183–216. Emerald Group Publishing Limited.

Pratt, T. C., & Cullen, F. T. 2005. Assessing Macro-Level Predictors and Theories of Crime: A Meta-Analysis. ***Crime and Justice***, 32: 373–450.

Rosenfield, N., Howey, B., & Mitchell, J. 2024, May 29. 'Did You Tase Him in the Face!?' Inside 'Goon Squad' Deputies' Group Chat. ***The New York Times***. https://www.nytimes.com/2024/05/29/us/mississippi-sheriffs-goon-squad-whatsapp.html.

Ross, H. L. 1984. ***Deterring the drinking driver: Legal policy and social control***, vol. 1982. Lexington Books Lexington, MA. https://www.ojp.gov/ncjrs/virtual-library/abstracts/deterring-drinking-driver-legal-policy-and-social-control.

Rozema, K., & Schanzenbach, M. 2023. Does Discipline Decrease Police Misconduct? Evidence from Chicago Civilian Allegations. ***American Economic Journal: Applied Economics***, 15(3): 80–116.

Schnake, M. E. 1986. Vicarious punishment in a work setting. ***Journal of Applied Psychology***, 71(2): 343–345.

Sheppard, B. H., Lewicki, R. J., & Minton, J. W. 1992. ***Organizational justice: The search for fairness in the workplace.*** Lexington Books/Macmillan.

Sierra-Arévalo, M. 2019. The commemoration of death, organizational memory, and police culture. ***Criminology***, 57(4): 632–658.

Simpson, S. S. 2002. ***Corporate crime, law, and social control***. https://www.researchgate.net/profile/Sally-Simpson-3/publication/267703139_Corporate_Crime_Law_and_Social_Control/links/55c220ef08aeca747d5dc598/Corporate-Crime-Law-and-Social-Control.pdf.

Skarlicki, D. P., Ellard, J. H., & Kelln, B. R. C. 1998. Third-party perceptions of a layoff: Procedural, derogation, and retributive aspects of justice. ***Journal of Applied Psychology***, 83(1): 119–127.

Skarlicki, D. P., & Folger, R. 1997. Retaliation in the Workplace: The Roles of Distributive, Procedural, and Interactional Justice. ***Journal of Applied Psychology***, 10.

Skarlicki, D. P., & Kulik, C. T. 2005. Third-party reactions to employee (mis)treatment: A justice perspective. ***Research in Organizational Behavior***, 26: 183–229.

Skarlicki, D., Van Jaarsveld, D., & Walker, D. 2008. Getting Even for Customer Mistreatment: The Role of Moral Identity in the Relationship Between Customer Interpersonal Injustice and Employee Sabotage. ***The Journal of Applied Psychology***, 93: 1335–1347.

Soltes, E. 2016. ***Why they do it: Inside the mind of the white-collar criminal***. Public Affairs.

Spencer, S., & Rupp, D. E. 2009. Angry, guilty, and conflicted: Injustice toward coworkers heightens emotional labor through cognitive and emotional mechanisms. ***Journal of Applied Psychology***, 94(2): 429–444.

Stafford, M. C., & Warr, M. 1993. A Reconceptualization of General and Specific Deterrence. ***Journal of Research in Crime and Delinquency***, 30(2): 123–135.

Stephens, D. W. 2011. ***Police Discipline: A Case for Change***: 27. Harvard Kennedy School & National Institute of Justice.

Stroube, B. K. 2021. Using allegations to understand selection bias in organizations: Misconduct in the Chicago Police Department. ***Organizational Behavior and Human Decision Processes***, 166: 149–165.

Sun, L., & Abraham, S. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. ***Journal of Econometrics***, 225(2): 175–199.

Tajfel, H. 1982. Social Psychology of Intergroup Relations. ***Annual Review of Psychology***, 33(1): 1–39.

Thibaut, J., & Walker, L. 1975. ***Procedural Justice: A Psychological Analysis***. Lawrence Erlbaum Associates, Inc.

Thomson-DeVeaux, A., Bronner, L., & Sharma, D. 2021. Cities Spend Millions On Police Misconduct Every Year. Here's Why It's So Difficult to Hold Departments Accountable. ***FiveThirtyEight***.

Tittle, C. R. 1980. ***Sanctions and Social Deviance—The Question of Deterrence***. New York: Praeger. https://www-ojp-gov.proxy.lib.umich.edu/ncjrs/virtual-library/abstracts/sanctions-and-social-deviance-question-deterrence.

Torres, M., & Curiel, F. 2023. Cracking the Blue Wall of Silence: A Necessary Step for Police Reform. ***American Bar Association***. https://www.americanbar.org/groups/gpsolo/resources/magazine/2023-september-october/cracking-blue-wall-silence-necessary-step-police-reform/.

Treviño, L. K. 1992. The Social Effects of Punishment in Organizations: A Justice Perspective. ***Academy of Management Review***, 31.

U.S. Department of Justice. 2017. ***Investigation of the Chicago Police Department***. Washington, DC.

Van Maanen, J. 1973. Observations on the Making of Policemen. ***Human Organization***, 32(4): 407–418.

Vaughan, D. 1999. The Dark Side of Organizations: Mistake, Misconduct, and Disaster. ***Annual Review of Sociology***, 25(1): 271–305.

Walker, S., & Bumphus, V. W. 1992. The effectiveness of civilian review: Observations on recent trends and new issues regarding the civilian review of the police. ***Am. J. Police***, 11: 1.

Weber, M. 1978. 46. Economy and Society. ***Democracy***: 247–251. Columbia University Press.

Wharton, A. S. 2009. The Sociology of Emotional Labor. ***Annual Review of Sociology***, 35(1): 147–165.

Wilson, J. Q., & Herrnstein, R. J. 1998. ***Crime human nature: The definitive study of the causes of crime***. Simon and Schuster. https://books.google.com/books?hl=en&lr=&id=ipsm16OpF2MC&oi=fnd&pg=PA13&dq=Wilson,+J.+Q.,+%26+Herrnstein,+R.+J.+(1985).+Crime+and+human+nature.+Simon+and+Schuster.&ots=epRvkRcZsE&sig=7Wd4GtUKeOYTniNAZgYFx1ssg3k.

Wolfe, S. E., & Piquero, A. R. 2011. Organizational Justice and Police Misconduct. ***Criminal Justice and Behavior***, 38(4): 332–353.

Yiu, D. W., Xu, Y., & Wan, W. P. 2014. The Deterrence Effects of Vicarious Punishments on Corporate Financial Fraud. ***Organization Science***, 25(5): 1549–1571.

Zhang, V. (Shu), Mohliver, A. C., & King, M. 2023. Where Is All the Deviance? Liminal Prescribing and the Social Networks Underlying the Prescription Drug Crisis. ***Administrative Science Quarterly***, 68(1): 228–269.

Zimring, F. E., Hawkins, G., & Vorenberg, J. 1973. ***Deterrence: The legal threat in crime control***, vol. 6511. University of Chicago Press Chicago. https://www.ojp.gov/ncjrs/virtual-library/abstracts/deterrence-legal-threat-crime-control.

Figure 1. Chicago Police Academy Cohort Size (1980-2017)

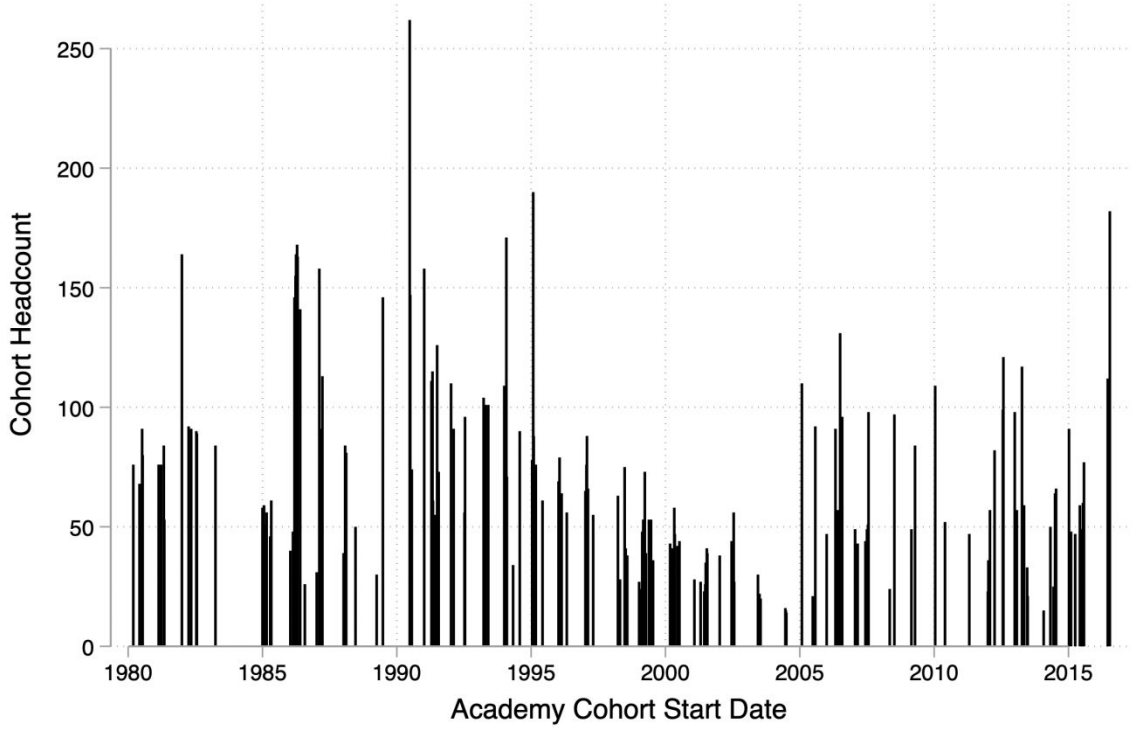
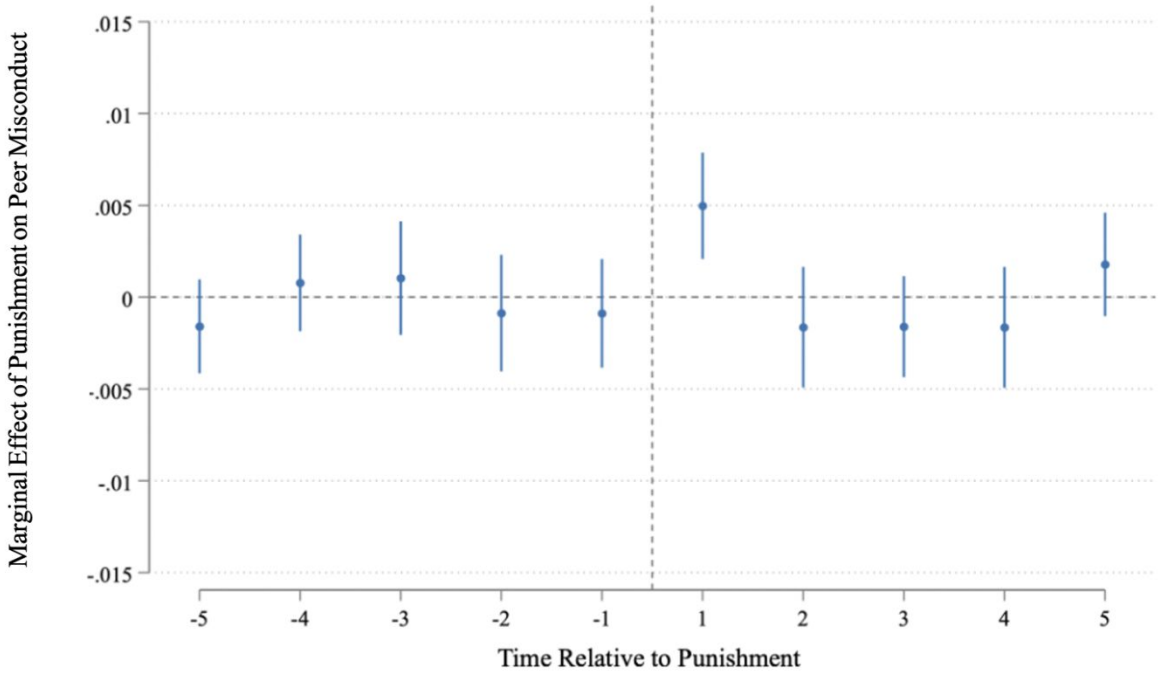


Figure 2. Time Trend of Punishment on Peer Misconduct



Note: Estimates are based on the regression similar to Equation 1, except we use five leads and lags of punishment as the independent variables. Includes officer and week fixed effects. The 90% confidence intervals are based on standard errors clustered at the cohort level.

Figure 3. Heterogeneous Effects of Punishment on Peer Misconduct, by Punishment Severity

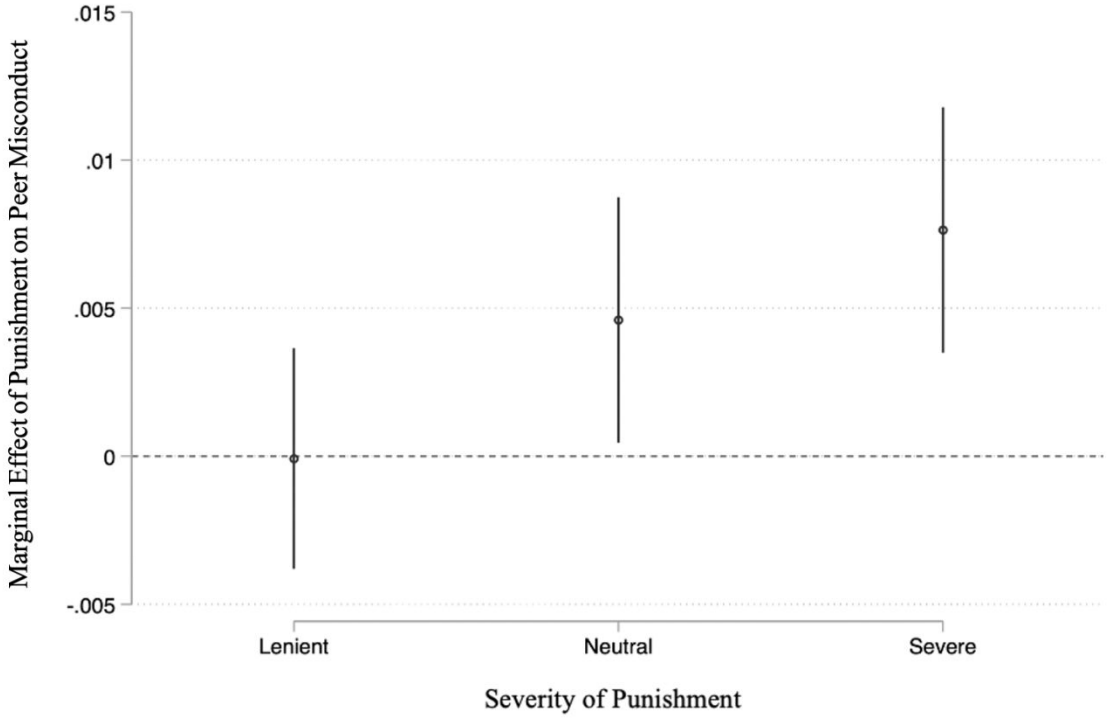


Figure 4. Robustness Specification Curve

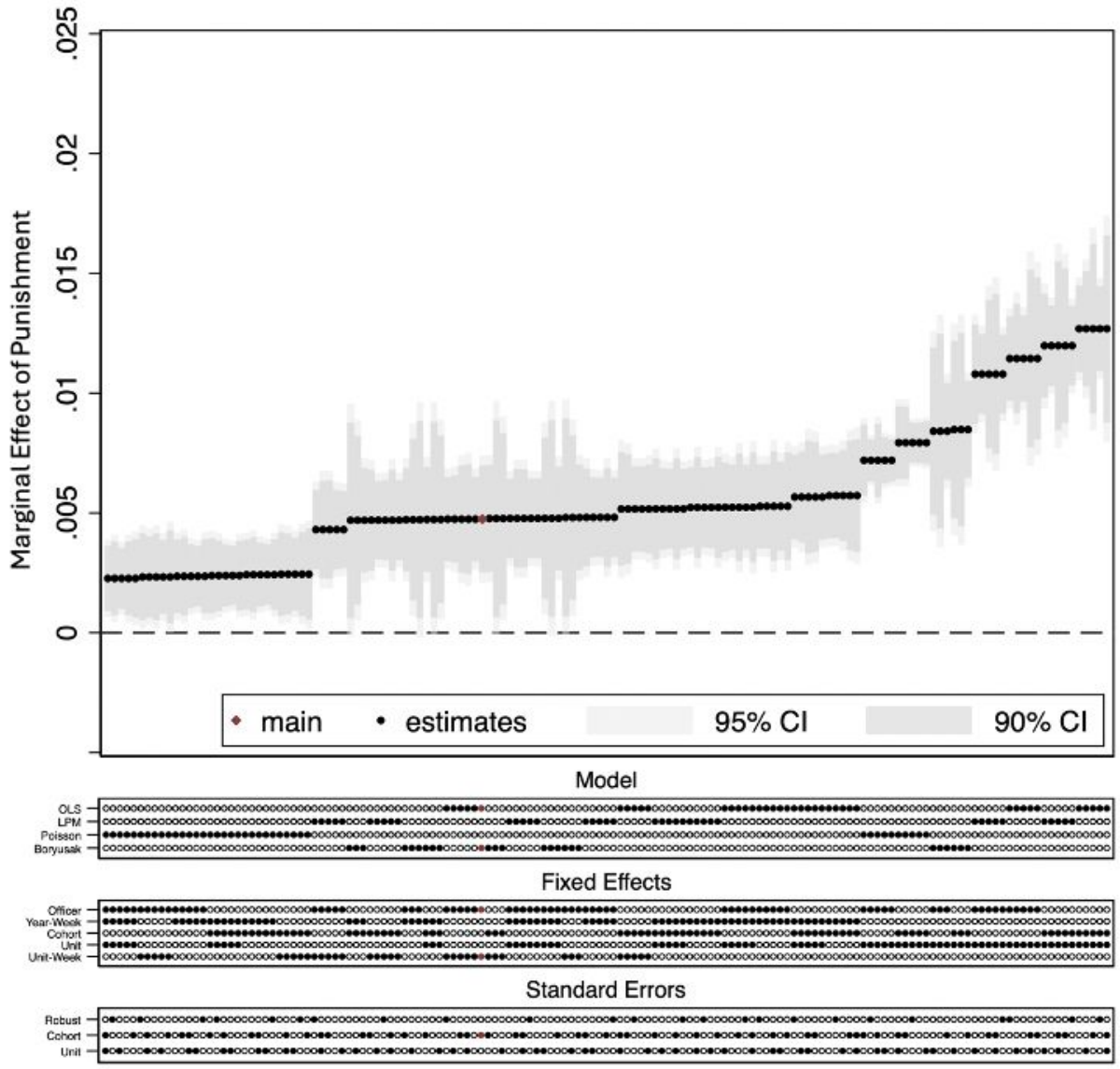


Table 1. Summary Statistics

	Observations	Mean	SD	Min	Max
Peer Misconduct	7,865,204	0.013	0.115	0	5
Punishment	7,865,204	0.00311	0.0556	0	1
Punishment Severity	7,865,204	0.000340	0.0676	-0.900	14
Lenient Punishment	7,865,204	0.00157	0.0396	0	1
Neutral Punishment	7,865,204	0.000336	0.0183	0	1
Severe Punishment	7,865,204	0.000914	0.0302	0	1

Table 2. Effect of Punishment on Peer Misconduct

	DV: Peer Misconduct					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			Borusyak et al. (2022)		
Punishment	0.00498***	0.00494***	0.00444***	0.00472***	0.00478***	0.00481***
	(0.00132)	(0.00132)	(0.00131)	(0.00131)	(0.00131)	(0.00134)
Officer FE	Y	Y	Y	Y	Y	Y
Year-Week FE	Y	Y	N	Y	Y	N
Unit FE	N	Y	N	N	Y	N
Unit × Year-Week FE	N	N	Y	N	N	Y
Mean of DV	0.013	0.013	0.013	0.013	0.013	0.013
Percent Change	39.2%	38.9%	35.0%	37.2%	37.6%	37.9%
Observations	7,865,293	7,865,293	7,865,293	5,388,449	5,388,449	5,388,449

*** p<0.001, ** p<0.01, * p<0.05 (two-tailed). Robust standard errors, clustered by cohort, in parentheses. The unit of analysis is the officer-week. Models 1-3 use OLS fixed effects estimators. The dependent variable equals the number of misconduct complaints the focal officer received in the week. The independent variable is a dichotomous variable equal to one in the week after an officer's cohort peer in a different unit from them was punishment for misconduct. Models 4-6 replicate these results of those in Models 1-3 using staggered diff-in-diff imputation methods from Borusyak et al. (2022). The number of observations differs because the Borusyak et al. (2022) drops observations that fall after the post-treatment horizon.

Table 3. Heterogeneous Effects of Punishment on Peer Misconduct, by Punishment Severity

	DV: Peer Misconduct					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			OLS		
Punishment Severity	0.00161*	0.00161*	0.00160*			
	(0.000701)	(-0.000699)	(0.000651)			
Lenient Punishment				-0.000643	-0.000654	-0.000827
				(0.00177)	(0.00177)	(0.00178)
Neutral Punishment				0.00452*	0.00451*	0.00385+
				(0.00209)	(0.00209)	(0.00198)
Severe Punishment				0.00689***	0.00683**	0.00644**
				(0.00205)	(0.00205)	(0.00203)
Officer FE	Y	Y	Y	Y	Y	Y
Year-Week FE	Y	Y	N	Y	Y	N
Unit FE	N	Y	N	N	Y	N
Unit × Year-Week FE	N	N	Y	N	N	Y
Mean of DV	0.0113	0.0113	0.0113	0.0113	0.0113	0.0113
Observations	7,865,293	7,865,293	7,865,293	7,865,293	7,865,293	7,865,293

*** p<0.001, ** p<0.01, * p<0.05 (two-tailed). Robust standard errors, clustered by cohort, in parentheses. The unit of analysis is the officer-week. All Models use OLS fixed effects estimators. The dependent variable equals the number of misconduct complaints the focal officer received in the week. *Punishment Severity* is a continuous variable equal to the percent difference between the number of days of suspension the peer officer reviewed and the number of suspension days the investigators recommended. *Lenient Punishment* is equal to one if the number of days of suspension the peer officer reviewed was less than the number of suspension days the investigators recommended; zero otherwise. *Neutral Punishment* is equal to one if the number of days of suspension the peer officer reviewed was equal to the number of suspension days the investigators recommended; zero otherwise. *Severe Punishment* is equal to one if the number of days of suspension the peer officer reviewed was greater than the number of suspension days the investigators recommended; zero otherwise.

Table 4. Results by Complaint Category

Dependent Variable:	Internal Complaints	External Complaints	Operational and Personal Complaints	Use of Force Complaints	Illegal Search Complaints	Lockup Procedure Complaints	Officer-Filed Complaints
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Punishment	0.00183**	0.00309**	0.00181**	0.00101+	0.00112*	0.000226	-0.000172
	-0.000684	-0.00107	-0.000658	-0.000584	-0.000499	-0.000261	-0.000264
Officer FE	Y	Y	Y	Y	Y	Y	Y
Year-Week FE	Y	Y	Y	Y	Y	Y	Y
Unit FE	Y	Y	Y	Y	Y	Y	Y
Mean of DV	0.00415	0.00796	0.00393	0.00324	0.00184	0.000673	0.000725
Observations	7,865,597	7,865,597	7,865,597	7,865,597	7,865,597	7,865,597	7,865,597

*** p<0.001, ** p<0.01, * p<0.05 (two-tailed). Robust standard errors, clustered by cohort, in parentheses. The unit of analysis is the officer-week. All regressions use OLS fixed effects estimators. Internal and external complaints are complaints that are primarily against the CPD or the public, respectively. The dependent variables in Models 3-6 use the four most prevalent complaint types. Officer-filed complaints are those levied by other officers, rather than the public. For details on complaint types, see https://directives.chicagopolice.org/forms/CPD-44.248.pdf