Appendix

Descriptive and Summary Statistics **2**
Table A1. Descriptive Statistics on Misconduct Behaviors 2
Table A2. Descriptive Statistics on Certification Status 2
Pre-Treatment Balance Checks **3**
Parallel Assumption Test **4**
Figure A1. Estimated difference in number of misconducts between groups 4
Figure A2. Pre-Treatment trends in misconduct by group 5
Time Series Trends in the Data **6**
Figure A3. Monthly misconduct by group 6
Figure A4. Monthly misconduct occurrence by group 7
Additional Models **8**
Table A3. Difference-in-Differences (DiD) Results, OLS Regression (individual fixed effect) 8
Table A4. Difference-in-Differences (DiD) Results, Logistic Regression (individual fixed effect) 9
Table A5. Difference-in-Differences (DiD) Results, OLS Regression (different cut-off) 10
Table A6. Difference-in-Differences (DiD) Results, Logistic Regression (different cut-off) 11
Table A7. Difference-in-Differences (DiD) Results, OLS Regression (without outliers) 12
Table A8. Difference-in-Differences (DiD) Results, Logistic Regression (without outliers) 13
Table A9. Difference-in-Differences (DiD) Results, (interaction term) 14
Middle Cohort (with last names beginning I-P) **15**
Table A10. Summary Statistics 15
Figure A7. Monthly Misconduct Occurrence 16
Information about POST Commission **17**
Background: 17
Figure A8: Conditions Requiring Mandatory Certification Revocation 18
Figure A9: Conditions Permitting Certification Suspension or Revocation 19
A note on the possibility of within agency officer reassignment, and the potential effects on our results **20**
Interview Guide **20**
Police Officers 20
Supervisors and Managers 22
Community Stakeholders and Activists 25
POST Commission Officials 25

Descriptive and Summary Statistics

Table A1. Descriptive Statistics on Misconduct Behaviors

	Observation		**Number of misconduct**			**Baseline misconduct**			**Occurrence of misconduct**	
Group	Officer	Officer-month	Mean	SD	Max	Mean	SD	Max	Mean	SD
Total	888	58608	0.034	0.352	21	0.596	1.647	17	0.018	0.134
Control	293	19338	0.029	0.284	10	0.556	1.839	17	0.017	0.131
Treated	595	39270	0.036	0.381	21	0.615	1.543	11	0.018	0.135

Table A2. Descriptive Statistics on Certification Status

Group	**Certification Status**	**Count**	**Percentage (%)**
Control	Certified*	230	**89.1**[a]
	Decertified	3	**1.2**
	Not certified**	19	7.4
	Suspended***	6	2.3
	Missing data	35	-
	Total	293	100
Treatment	Certified*	508	**88.3**[b]
	Decertified	13	**2.3**
	Not certified**	39	6.8
	Suspended***	15	2.6
	Missing data	20	-
	Total	595	100

* Includes cases coded as certified, certified/SRO, conditionally certified; ** Includes cases coded as not certified, not certified (on leave), expired; *** Could be by POST or administratively; [ab] Percentage without the missing data

Pre-Treatment Balance Checks

To ensure that our treated and control cohorts are comparable before the recertification review, we restrict the sample to all officer-months up to July 1, 2022 and conduct the following tests:

Welch two-sample *t*-test on mean monthly misconduct count. The treatment group has significantly higher average of misconduct (0.0340) than the control group (0.0239) [$t = -3.539$, df = 34,490, p = 0.00040]

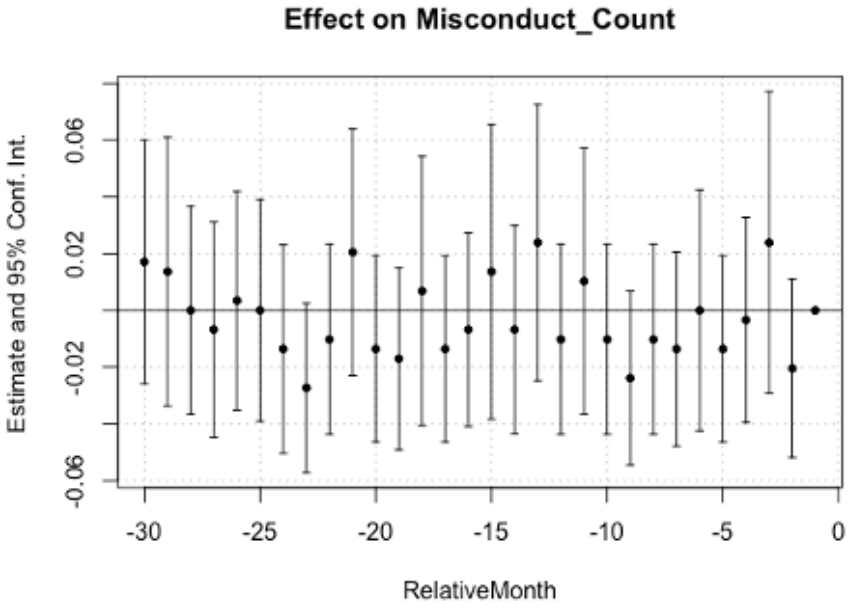
Pearson's χ^2 test (Yates' correction) on the proportion of months with any misconduct (the 0/1 indicator). There is no strong evidence of a difference in the likelihood of any misconduct between treatment and control groups (marginally non-significant at the 10% level) [$\chi^2 = 3.443$, df = 1, p = 0.0635]

Welch two-sample *t*-test on baseline misconduct (total pre-2019 incidents), using one record per officer. There is no significant difference in officers' cumulative misconduct prior to 2019 with control and treatment groups' means being 0.5563 and 0.6151, respectively [$t = -0.471$, df $\approx$ 500, p = 0.638]

While the two cohorts do not differ significantly in the binary occurrence of misconduct or in their pre-2019 misconduct histories, the treated group has a higher monthly misconduct count before July 2022. While differences in raw means do not themselves invalidate the DiD design, they underscore the importance of our difference-in-differences specification—with cohort and time fixed effects plus baseline controls—to adjust for any remaining pre-treatment disparities.

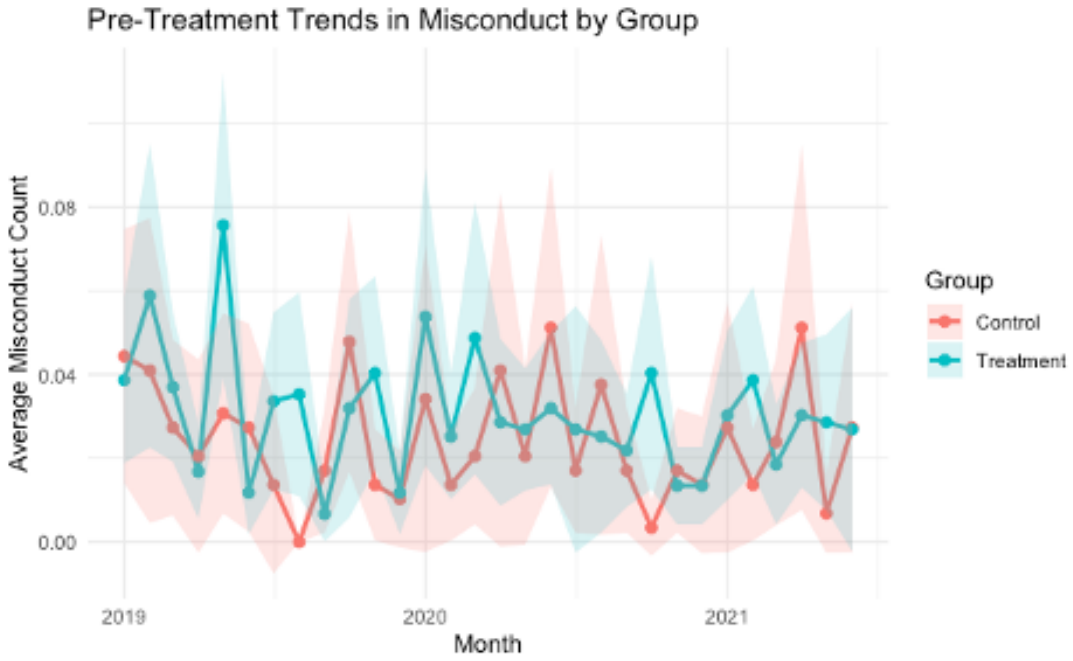
Parallel Assumption Test

Figure A1. Estimated difference in number of misconducts between groups



In this plot, each point represents the estimated difference in number of misconducts between the treated and control groups at a given RelativeMonth before the policy took effect, with vertical lines showing 95% confidence intervals. Most estimates are close to the horizontal axis (zero), with the vertical bars being quite wide and intersecting the horizontal axis. This indicates that the treatment and control groups did not differ significantly in their misconduct trends at those lead times.

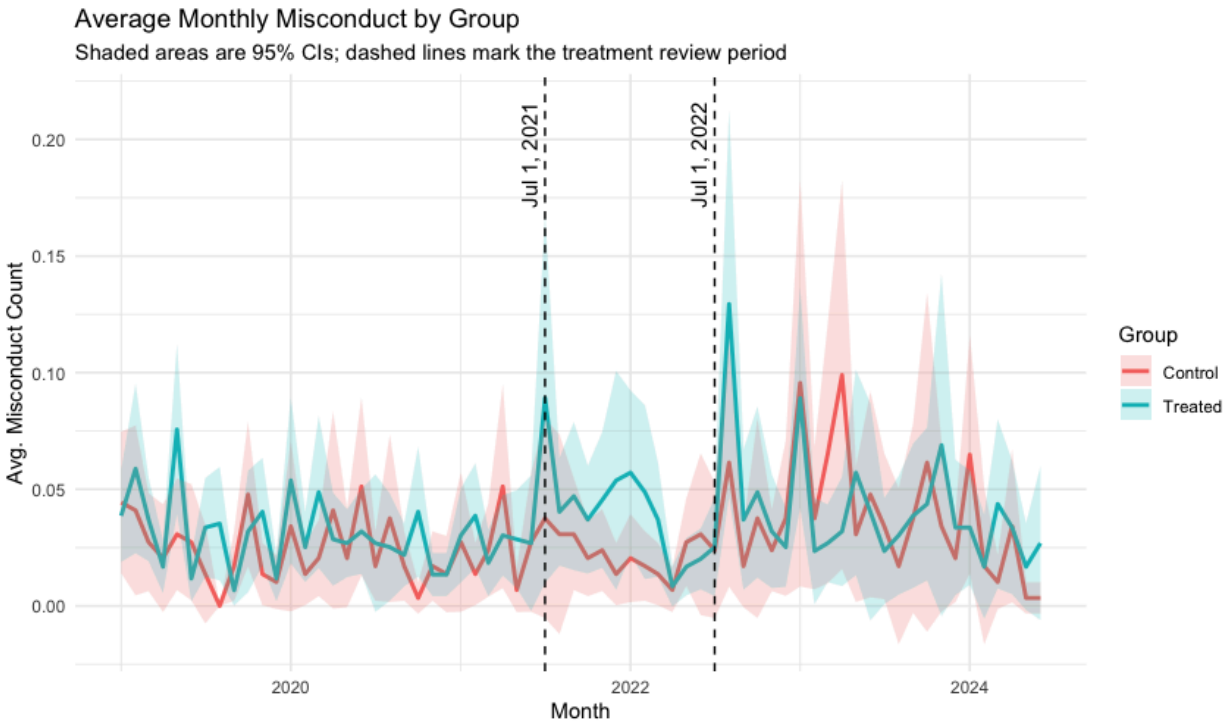
Figure A2. Pre-Treatment trends in misconduct by group



We plot the average misconduct trends for both the treatment and control groups over the pre-treatment period. With the exception of a few isolated points, the trends run in parallel, and the confidence intervals around the average estimates overlap. This visual evidence supports the parallel trends assumption essential to our difference-in-differences analysis

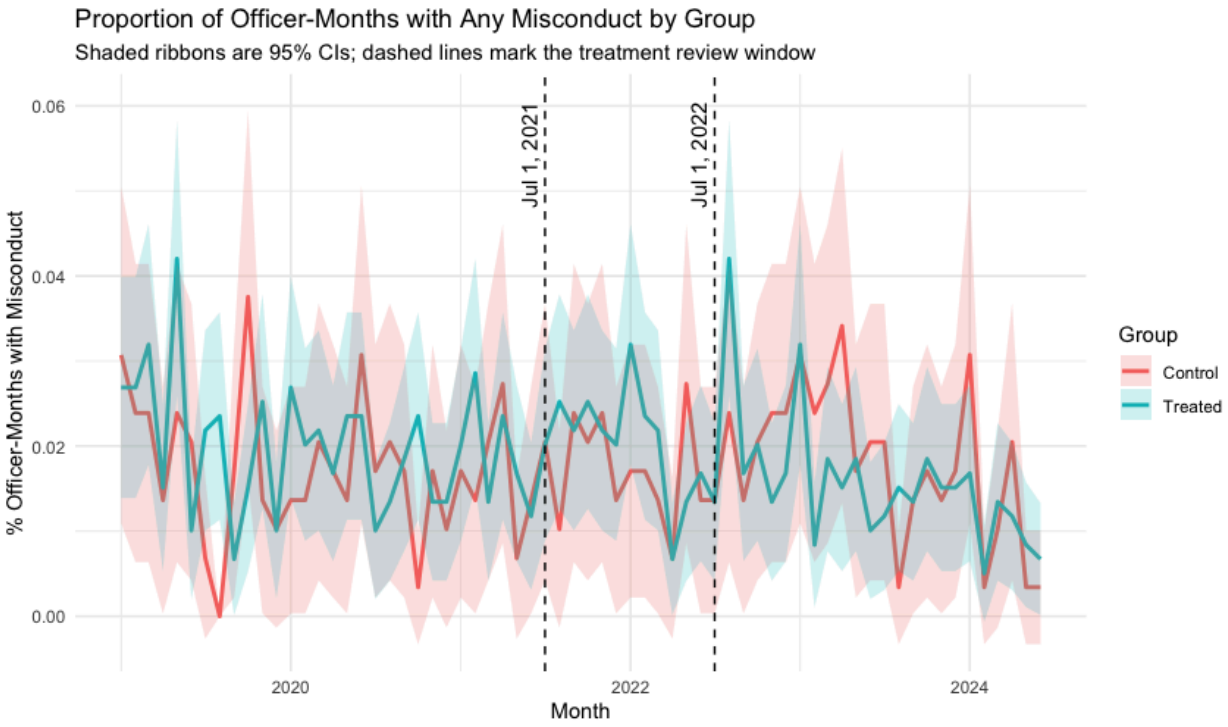
Time Series Trends in the Data

Figure A3. Monthly misconduct by group



This plot shows the number of misconduct by group (treatment vs. control) over time.

Figure A4. Monthly misconduct occurrence by group



This plot displays the monthly proportion of officers committing at least one misconduct over time in both the treatment and control groups.

Additional Models

Table A3. Difference-in-Differences (DiD) Results, OLS Regression (*individual fixed effect*)

	Dependent Variable: *Count of Misconducts*	
	(1)	(2)
DiD	-0.008 (0.007)	-0.004 (0.007)
Individual (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Adjusted R^2	-0.004	-0.003
RMSE	0.350	0.341
Observation	58,608	57,024

**p<.10, ** p<.05, *** p<.01*

In our fixed-effects logistic specification, two automatic adjustments occur. First, any officer whose outcome is always 0 or always 1 across all months is dropped (264 observations in Model 1; 894 in Model 2), since the logit likelihood requires within-officer variation. Second, like in the OLS fixed-effects model, Baseline_Misconduct is constant for each officer and is perfectly collinear with the individual dummies, so it is removed by the estimator.

Table A4. Difference-in-Differences (DiD) Results, Logistic Regression (*individual fixed effect*)

	Dependent Variable: *Misconduct Occurrence*	
	(1)	(2)
DiD	-0.293** (0.147)	-0.258* (0.151)
Individual (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Log-Likelihood	-5,166.4	-5,032.7
BIC	20,758.3	20,290.1
Observation	58,344	56,130

**p<.10, ** p<.05, *** p<.01*

In our fixed-effects logistic specification, there were two automatic adjustments. First, those with "all-0" or "all-1" fixed-effect groups are dropped (264 and 894 observations for Model 1 and Model 2, respectively). That means any officer for which the outcome is always 0 or always 1 across all months provides no within-officer variation for the logit likelihood. Because the logit estimator cannot separate a constant outcome from the fixed-effect intercept, those officers and their observations are excluded entirely. Also, just as in the OLS fixed-effects model, Baseline_Misconduct is a constant for each officer. Once we include individual dummies, there is no remaining within-officer variation in Baseline_Misconduct to estimate. As a result, it is perfectly collinear with the officer fixed effects and is removed by the estimation routine.

Table A5. Difference-in-Differences (DiD) Results, OLS Regression (*different cut-off*)

	Dependent Variable: *Count of Misconducts*	
	(1)	(2)
DiD	-0.002 (0.007)	-0.004 (0.006)
Baseline Misconduct	0.003** (0.001)	0.002** (0.001)
Group (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Clustered Standard Errors (Officer)	✓	✓
Adjusted R^2	0.002	0.002
RMSE	0.342	0.304
Observation	47,952	46,368

**p<.10, **p<.05, ***p<.01*

Table A5 re-estimates our primary OLS difference-in-differences models using July 1, 2021 as the "treatment" cutoff (i.e., the start of the A–H review window). By shifting the post-indicator earlier—when no A–H officers had yet completed recertification—the DiD coefficients are insignificant, confirming that there is no anticipatory decline in misconduct prior to actual recertification.

Table A6. Difference-in-Differences (DiD) Results, Logistic Regression (*different cut-off*)

	Dependent Variable: *Misconduct Occurrence*	
	(1)	(2)
DiD	-0.162 (0.150)	-0.231 (0.153)
Baseline Misconduct	0.050*** (0.011)	0.044*** (0.012)
Group (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Clustered Standard Errors (Officer)	✓	✓
Log-Likelihood	-4,530.3	-4,312.0
BIC	9,675.0	9,236.4
Observation	47,952	46,368

**p<.10, **p<.05, ***p<.01*

Table A6 re-estimates our primary logit difference-in-differences models using July 1, 2021 as the "treatment" cutoff. The DiD coefficients become statistically insignificant, confirming that there is no anticipatory decline in misconduct occurrence rate prior to actual recertification.

Table A7. Difference-in-Differences (DiD) Results, OLS Regression (*without outliers*)

	Dependent Variable: *Count of Misconducts*	
	(1)	(2)
DiD	-0.008 (0.006)	-0.007 (0.007)
Baseline Misconduct	0.000 (0.000)	0.000** (0.001)
Group (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Clustered Standard Errors (Officer)	✓	✓
Adjusted R^2	0.001	0.001
RMSE	0.315	0.314
Observation	58,146	56,658

**p<.10, **p<.05, ***p<.01*

To check whether a few officers with extremely high misconduct counts unduly influence our DiD estimates, we identify the top 1 percent of officers by total misconduct over the full panel and re-estimate our main OLS and logit models (results in Tables A7 and A8, respectively). The key DiD coefficients remain essentially unchanged, indicating that extreme "high-misconduct" officers are not driving the results. However, in Table A7 (the OLS specification), the coefficient on baseline misconduct loses statistical significance—possibly because those highest-misconduct officers also had the largest pre-2019 misconduct histories.

Table A8. Difference-in-Differences (DiD) Results, Logistic Regression (*without outliers*)

	Dependent Variable: *Misconduct Occurrence*	
	(1)	(2)
DiD	-0.289** (0.148)	-0.278* (0.149)
Baseline Misconduct	0.023*** (0.009)	0.021** (0.009)
Group (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Clustered Standard Errors (Officer)	✓	✓
Log-Likelihood	-5,171.0	-5,062.4
BIC	11,099.0	10,879.9
Observation	58,146	56,658

**p<.10, ** p<.05, *** p<.01*

Table A9. Difference-in-Differences (DiD) Results, (*interaction term*)

	Dependent Variable: *Count of Misconducts*	Dependent Variable: *Misconduct Occurrence*
	(1)	(2)
DiD	0.004 (0.007)	-0.255* (0.150)
Baseline Misconduct	0.003** (0.001)	0.044*** (0.014)
DiD*Baseline Misconduct	-0.006***	-0.057 (0.050)
Group (fixed effect)	✓	✓
Year-month (fixed effect)	✓	✓
Clustered Standard Errors (Officer)	✓	✓
Adjusted R^2	0.002	-
RMSE	0.352	-
Log-Likelihood	-	-5,247.4
BIC	-	11,263.3
Observation	58,608	58,608

**p<.10, ** p<.05, *** p<.01*

Middle Cohort (with last names beginning I-P)

Table A10. Summary Statistics

Period	Average Count	SE (Count)	Occurrence	SE (Occurrence)	Number of months
Pre	0.0356	0.0044	0.0192	0.0017	6256
Post	0.0163	0.0027	0.0077	0.0010	6624

We run a Welch two-sample t-test comparing mean monthly misconduct counts before and after the I–P cohort's recertification. In the post-treatment period, the cohort's average monthly count (0.0163) is significantly lower than in the pre-treatment period (0.0356) [$t = -3.733$, df = 10,425, $p < 0.001$]. A Pearson χ^2 test (with Yates' correction) on the binary indicator for any misconduct also shows a significant decline in the share of officer-months with at least one incident after recertification [$\chi^2 = 31.51$, df = 1, $p < 0.001$]. These simple pre/post comparisons do not control for month–year effects or other time-varying factors, but they still serve as a useful robustness check on our main findings

Figure A6. Monthly Average Misconduct

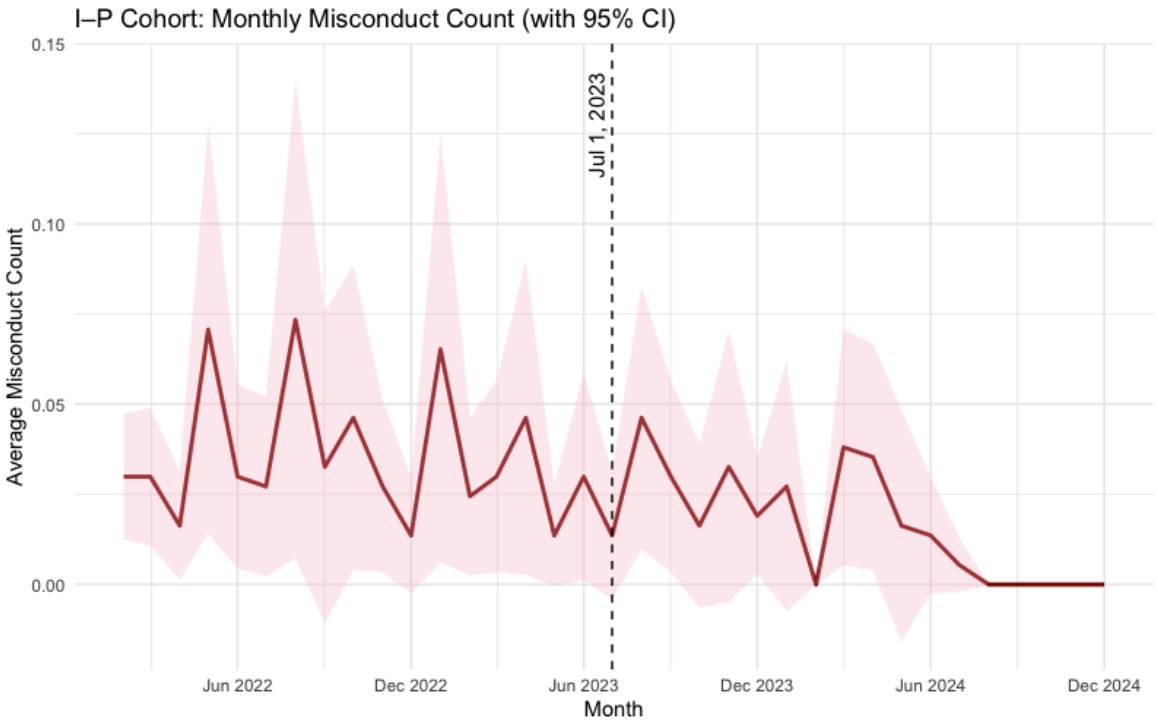
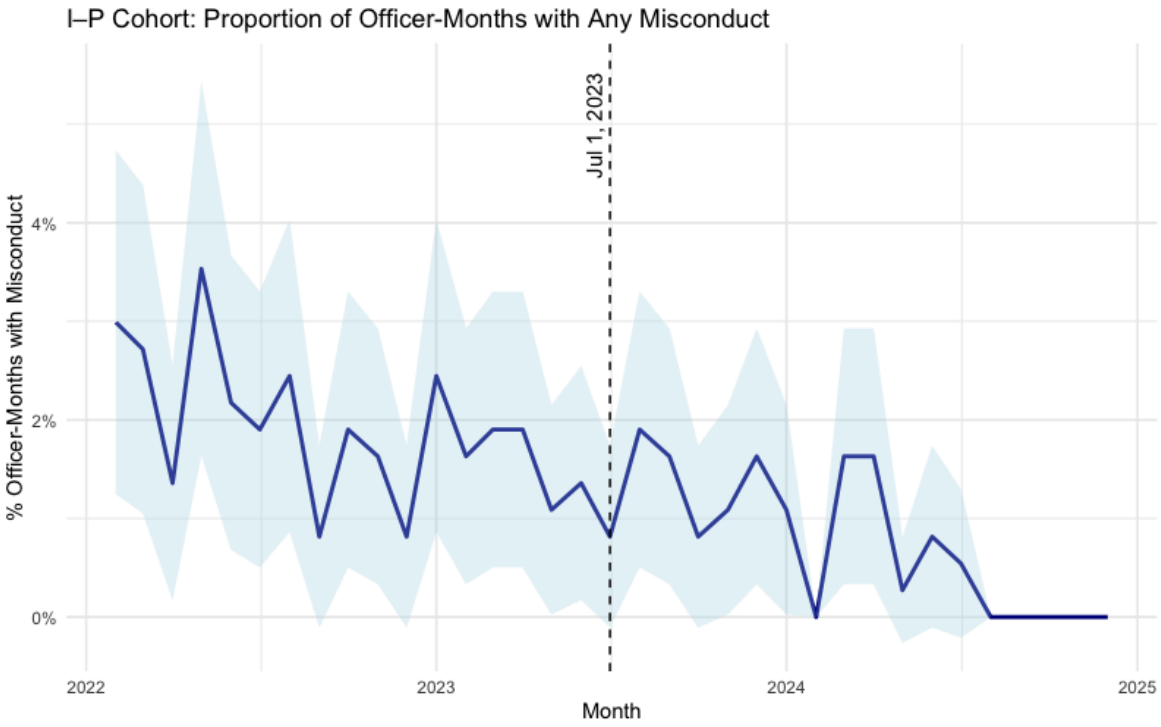


Figure A7. Monthly Misconduct Occurrence



Information about POST Commission

Background:

Beginning in 2021, all active-duty officers in Massachusetts were placed on a phased recertification schedule based on the alphabetical order of their last names. Once certified, officers must renew their certification every three years through a dual-review process. The process comprises two key phases: an internal departmental review and an external review by the POST Commission, which functions as the final and legally binding arbiter or an officer's fitness to serve. The internal review comprises three distinct phases. First, officers complete a detailed questionnaire addressing any history of misconduct both on and off duty. This includes specific questions related to use of force, bias, civil allegations of violence or abuse, issues concerning their license to carry, social media presence, and other indicators of fitness to serve (Recertification – MA POST Commission n.d.). In the second phase, the completed questionnaire is submitted to a designated evaluator—an officer of higher rank appointed by the agency head—who conducts an interview with the officer to further assess their suitability for continued service. In the third phase, the evaluator's findings, along with the completed questionnaire, the officer's misconduct history, training records, and any other relevant documentation, are forwarded to the agency head for a comprehensive review. As part of this final assessment, the agency head may consult supervisors, colleagues, internal affairs personnel, training and performance evaluations, external disciplinary records, and city officials involved in misconduct litigation and labor agreements. If the agency head determines the officer remains fit to serve, they submit an attestation to the POST Commission on the officer's behalf (Recertification-MA POST Commission). Following internal review, officers undergo an external assessment by the POST Commission. Here, for the first time in Massachusetts history, law enforcement officers are subject to certification by an independent, civilian-majority, centralized state body with subpoena power and a legal mandate to determine whether they may continue serving in any law enforcement capacity. This represents a significant increase in officers' exposure to external scrutiny and credible sanctions.. Empowered to revoke, suspend, or condition certification for a range of violations—including on- or off-duty misconduct, certification fraud, training deficiencies, and criminal convictions)—the Commission reviews the attestation materials alongside the officer's training and disciplinary records. Based on this assessment, the commission may certify, recertify, conditionally certify, or revoke an officers certification (Recertification-MA POST Commission). Officers facing decertification are given the opportunity to appear at a public deliberation regarding their certification status. After a decision is taken, their certification status is then communicated to their agency head and those who are 8 decertified are placed on the National Decertification Index, which is currently accessible to law enforcement agencies in the United States (Kaste 2025). Officers who fail to obtain certification are legally barred from working in any law enforcement capacity in Massachusetts (General Law-Chapter 6E.). Importantly, this means that when agencies terminate or suspend officers for egregious or unlawful misconduct, the resulting decertification falls outside the scope of collective bargaining protections. As a result, disciplinary actions resulting in the termination of employment that were previously vulnerable to reversal through arbitration are no longer subject to the same constraints, reducing the likelihood that

sanctions against officers will be overturned. This dynamic is especially salient given that, in 2023, one-third of officer terminations in the city of Boston were overturned through arbitration, even in cases involving extreme and egregious misconduct such as documented patterns of sexual assault against minors (Katcher 2024). Importantly, the POST Commission maintains both mandatory and discretionary authority. Certification can be denied if the officer fails to meet professional standards, has a felony conviction, or appears in the National Decertification Index (NDI) (555 CMR 7.00). The Commission is required to revoke certification under specific conditions (Figure A8) and permitted to suspend or revoke certification under broader circumstances (Figure A9).

Figure A8: Conditions Requiring Mandatory Certification Revocation

After a hearing, the POST Commission is *required* to revoke an officer's certification if they find clear and convincing evidence that:

- the officer has been convicted of a felony;
- initial certification was obtained through falsification or fraud, or was revoked by any other jurisdiction;
- the officer was terminated by their appointing agency for making a false arrest, falsifying or tampering with evidence, providing false testimony, engaging in conduct that would constitute a hate crime, or taking a bribe while on duty;
- the officer has been convicted of submitting false timesheets
- the officer knowingly files a written police report containing a false statement or commits perjury
- the officer tampers with a record for use in an official proceeding
- used excessive force resulting in bodily injury or death;
- the officer used a chokehold or excessive force, in violation of section 14, which restricts the use of force to necessary and proportionate circumstances and mandates that officers engage in de-escalation tactics where possible
- The officer engaged in witness intimidation;
- The officer failed to intervene or attempt to intervene to prevent another officer from engaging in prohibited behavior, including but not limited to the use of excessive force
- the commission determines the officer is dangerous to the public

Figure A9: Conditions Permitting Certification Suspension or Revocation

Additionally, the commission, after a hearing, *may* revoke or suspend an officer's certification if they find evidence that

- the officer has committed a misdemeanor
- was biased on the basis of race, ethnicity, sex, gender identity, sexual orientation, religion, mental or physical disability, immigration status or socioeconomic or professional level in their conduct;
- failed to comply with this chapter or commission regulations, reporting requirements or training requirements;
- was biased on the basis of race, ethnicity, sex, gender identity, sexual orientation, religion, mental or physical disability, immigration status or socioeconomic or professional level in their conduct;
- used excessive force;
- failed to respond an incident according to established procedure;
- has a pattern of unprofessional police conduct;
- was untruthful, except for a statement or action that mandates revocation pursuant to subsection (a);was suspended or terminated by their appointing agency for disciplinary reasons, except those reasons which mandate revocation pursuant to said subsection (a);
- fails to intervene to prevent another officer from engaging in prohibited conduct or behavior, except a failure to intervene in conduct that mandates revocation pursuant to said subsection (a); or
- would benefit in their job performance if retrained.

A note on the possibility of within agency officer reassignment, and the potential effects on our results

It would be reasonable to expect that agency heads may reassign more problematic officers to lower-contact policing roles in the wake of recertification as a way to reduce their involvement in situations where they might reoffend. Although interviews with superior officers indicate no changes to agency organization and officer assignment as a result of recertification processes, POST data does not record changes in the specific within-agency assignment of officers unless it is part of disciplinary proceedings (i.e. demotion), and even these cases are not uniformly coded in POST data in a way that allows for robust analysis. Arguably, these administrative gymnastics do present the possibility of variation in misconduct outcomes for our treatment group simply because these officers have largely been removed from situations where it is possible to commit certain types of misconduct. At the same time, departments have little incentive to make these reassignments unless required, given personnel shortages and collective bargaining agreements that make officer reassignment or informal demotion costly to agencies. While it would be heartening to find that reassignment of problematic officers away from situations where they might be a danger to the general public were occurring, our results show that many officers with prior misconduct in fact remain in the sample of officers with sustained allegations of misconduct to continue reoffending, even in the post-treatment period.

Interview Guide

The following questions formed the basis for semi-structured interviews conducted during this project. Some may have been admitted or expanded upon during interviews.

Police Officers

General Background

- What is your name, age, and current rank?
- Can you describe your typical shift duties?
- Are you part of a specialized unit?
- Does your unit have opportunities to participate in external training programs?

Experience with Recertification

- Can you walk me through your experience with the recertification interview and questionnaire?
- How did you prepare for the interview and questionnaire?
- Did you find the questions in the recertification questionnaire fair and relevant to your duties?
- How did this evaluation process compare to other routine evaluations you've experienced?
- How do you view the role of internal affairs in this process?

Perceived Scrutiny and Oversight

- Did this process change your relationship with commanders or managers?
- Have you noticed any changes in how closely your work is being monitored?

Supervisory Interviews and Assessments

- How did your interview with your supervisor proceed?
- Did your supervisor ask any questions beyond the standard questionnaire?
- Did you feel that your responses were assessed fairly?
- Did you receive any feedback following your interview?

Impact on Behavior

- Have you changed how you approach your work or decision-making since going through the recertification process?
- Did the interview or questionnaire make you more cautious while on duty?

Fear of Sanctions and Past Misconduct

- Are you more concerned about potential penalties for misconduct under the current system?
- In your view, how does having a prior history of misconduct affect the recertification process?

- Research suggests that officers with prior misconduct are less likely to change behavior in response to reforms. Why do you think this is the case?

Peer Influence and Discussion

- Did you and your colleagues discuss the recertification interview and questionnaire among yourselves?

Duty to Intervene and Peer Accountability

- How do you feel about the reform requiring officers to intervene against excessive force?
- Do you think this requirement has changed how officers perceive misconduct within the ranks?
- Do you believe this policy will lead to greater accountability?

Civilian Oversight and Public Trust

- What is your opinion on civilian involvement in misconduct investigations?
- Do you believe civilians understand policing well enough to evaluate disciplinary cases?

Supervisors and Managers

Evaluating Officers

- How do you evaluate officers during the recertification process, beyond the questionnaire and interview?
- Have you encountered instances where officers tried to downplay or omit incidents?

Effectiveness of Reform

- In your opinion, is the recertification process an effective mechanism for ensuring officer accountability?
- Have you observed any patterns in how officers with and without prior misconduct respond to recertification?

- Why might officers with a history of misconduct be less responsive to reform?
- Have you seen behavior change among officers since decertification risks increased?
- Do you observe differences in misconduct between experienced officers and new recruits?
- Have you noticed any change in the types of misconduct (e.g., use of force, bias)?
- Have any officers or supervisors shown resistance to the decertification process?
- What role does the police union play in shaping officer responses to oversight?

Department Culture

- Have you noticed changes in departmental culture since the introduction of recertification?

Agency-level Adjustments

- How has your agency adapted to the potential for officer decertification under POST?
- Have internal policies or disciplinary procedures changed as a result?
- What actions has your agency taken to maintain compliance with certification standards?
- Have recruitment or screening procedures changed due to stricter decertification criteria?

Internal Culture and Morale

- Have you observed changes in how officers perceive job security or morale?
- How have officers reacted when colleagues are decertified?

Misconduct Reporting and Oversight

- Has there been any change in internal reporting since the decertification policy began?
- Do officers and supervisors now take complaints or use-of-force reviews more seriously?
- Are officers more reluctant to report peer misconduct?
- Has the frequency of internal investigations changed?

Reputation and Public Trust

- Has the decertification process affected the agency's community reputation?
- Have public complaints increased or decreased since its introduction?
- Since decertification is rare and often publicized, have you noticed changes in community perception following such events?

Broader Structural Effects

- Has the threat of decertification influenced officer retention?
- Have any officers resigned to avoid recertification scrutiny?
- How does your agency compare with others in decertification rates, and why?

City Officials / Solicitors

- How have officers, supervisors, unions, city solicitors, and residents responded to the reform?
- What role have police unions played in shaping this process?
- Have you observed changes in officer operations or daily practices?
- How has your city handled misconduct, especially in high-profile or criminal cases?
- Have termination procedures shifted?
- How has the reform affected the everyday work of officers and supervisors?
- How does this reform compare to others your city has implemented?
- How does it align with existing city oversight frameworks?
- Has the reform increased the administrative burden on the city or department?
- How are suspensions handled, especially for uncertified officers?

- What happens to officers with entrenched misconduct patterns?
- Do officers vary in their responsiveness to reform?

Community Stakeholders and Activists

- How has the community received the reforms?
- How do these reforms compare with earlier initiatives or existing oversight institutions?
- What opportunities exist for the public to hold officers accountable?
- How effective are the current accountability mechanisms?
- One focus of this study is constituent constraint on officers. How do these reforms relate to that?

POST Commission Officials

1. Are there discrepancies in data between the pre-POST period and the present?
2. Can you explain the complaint reporting process—both direct and agency-reported?
3. How does POST ensure that departments comply with misconduct reporting?
4. What mechanisms ensure quality and consistency in certification reporting?
5. Do agency heads or internal affairs officials receive training on POST-related responsibilities?
6. Is there a way to ensure reporting is standardized across departments?
7. How can officers contest POST rulings or sanctions?
8. What role do police unions play in these processes?
9. Can you describe the hearing process? Are they open to public participation?
10. How does POST handle officers who resign or are terminated before formal decertification?
11. Have you encountered resistance from agencies during implementation?

12. Have you observed any patterns or changes in misconduct since POST began?