

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Unequal Protection under the Safe Drinking Water Act: Exposure, Compliance Recovery, and Enforcement Responsiveness in California Community Water Systems

Permalink

<https://escholarship.org/uc/item/8pg9g6tw>

ISBN

9798197827661

Author

Bing, Qi

Publication Date

2026-07-02

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Unequal Protection under the Safe Drinking Water Act: Exposure, Compliance Recovery, and
Enforcement Responsiveness in California Community Water Systems

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Urban and Environmental Planning and Policy

by

Qi Bing

Dissertation Committee:
Professor Emeritus David L. Feldman, Chair
Professor Jae Hong Kim
Professor Jun Wu

2026

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vii
ACKNOWLEDGMENTS	ix
VITA	x
ABSTRACT OF THE DISSERTATION	xi
Chapter 1: Introduction	1
1.1 Motivation and Problem Statement	1
1.2 Objectives, Research Questions, and Contributions	4
1.3 Institutional and Regulatory Context	6
1.4 Conceptual Framework	9
1.5 Dissertation Roadmap	12
1.6 References	12
Chapter 2: Nitrate in Community Drinking Water and Birth Outcomes in California	17
2.1 Chapter Overview	17
2.2 Literature Review	18
2.3 Data	27
2.4 Methods	34
2.5 Results	42
2.6 Sensitivity Checks	67
2.7 Discussion	73
2.8 Conclusion	83

2.9 References	85
APPENDIX A: Chapter 2 Supplemental Tables and Figures	95
Chapter 3: Return to Compliance under the SDWA: Duration and Environmental Justice in California Community Water Systems	107
3.1 Chapter Overview	107
3.2 Literature Review.....	108
3.3 Data	119
3.4 Empirical Strategy	129
3.5 Results.....	134
3.6 Discussion	152
3.7 Conclusion	160
3.8 References	161
APPENDIX B: Chapter 3 Supplemental Tables and Figures	169
Chapter 4: Formal Enforcement Escalation, Compliance Dynamics, and Unequal Protection in California Community Water Systems	182
4.1 Chapter Overview	182
4.2 Literature Review.....	184
4.3 Data and Analytic Sample	193
4.4 Empirical Strategy	197
4.5 Results.....	205
4.6 Discussion	222
4.7 Conclusion	230
4.8 References.....	231

APPENDIX C: Chapter 4 Supplemental Tables and Figures	236
Chapter 5: Conclusion.....	246
5.1 Overview.....	246
5.2 Key Findings.....	246
5.3 Cross-essay Synthesis	248
5.4 Policy Implications	249
5.5 Limitations	252
5.6 Future Directions	253

LIST OF FIGURES

	Page
Figure 1.1: Integrated Conceptual Framework of Unequal Protection Under Capacity Constraints	10
Figure 2.1: Directed Acyclic Graph for Pregnancy Nitrate Exposure and Birth Outcomes	38
Figure 2.2: Adjusted Odds Ratios for Preterm Birth by Pregnancy-Average Nitrate Category ...	47
Figure 2.3: Adjusted Odds Ratios for Low Birthweight by Pregnancy-Average Nitrate Category	49
Figure 2.4: Adjusted Odds Ratios for Preterm Birth by Trimester-Specific Nitrate Exposure	52
Figure 2.5: Adjusted Odds Ratios for Low Birthweight by Trimester-Specific Nitrate Exposure... ..	55
Figure 2.6: Central Valley Heterogeneity in Adjusted Nitrate-Threshold Associations	62
Figure 2.7: Central Valley Heterogeneity in Adjusted Continuous Nitrate Associations	63
Figure 2.8: DAC Heterogeneity in Adjusted Nitrate-Threshold Associations	64
Figure 2.9: DAC Heterogeneity in Adjusted Continuous Nitrate Associations	65
Figure 2.10: SDAC Heterogeneity in Adjusted Nitrate-Threshold Associations	66
Figure 2.11: SDAC Heterogeneity in Adjusted Continuous Nitrate Associations	67
Figure 3.1: Conceptual Framework for Compliance Recovery and Return to Compliance	119
Figure 3.2: Violation Starts and Recorded Resolutions Over Time, 1995–2024	138
Figure 3.3: Recorded Resolution Share by Violation Start Cohort	139
Figure 3.4: Kaplan-Meier Recorded-Resolution Shares Within Selected Horizons, by Health- Based Status	140
Figure 3.5: Kaplan-Meier Cumulative Recorded-Resolution Curves by SDAC Status	141

Figure 3.6: Kaplan-Meier Cumulative Recorded-Resolution Curves by System Size, with SDAC Overlay	142
Figure 3.7: System-Size Contrasts Within SDAC Systems, Baseline Log-Logistic AFT Model	146
Figure 4.1: Conceptual Framework for First Formal Escalation and Compliance Progress	193
Figure 4.2: Weighted Event-Study Estimates of Quarterly Compliance Progress Around First Formal Enforcement	210
Figure 4.3: Heterogeneity by Disadvantaged Community Status in Weighted Event-Study Estimates of Quarterly Compliance Progress	216
Figure 4.4: Weighted Pre-Enforcement Coefficients from the Baseline Event-Study Specification	218
Figure 5.1: Integrated Findings Across Three Stages of Unequal Protection.....	248

LIST OF TABLES

	Page
Table 2.1: Linked and Unlinked Births in the 2012–2022 Comparison Window	32
Table 2.2: Descriptive Statistics for the Complete-Case Regression Sample	43
Table 2.3: Adjusted Odds Ratios for Preterm Birth by Pregnancy-Average Nitrate Exposure	
Specification	46
Table 2.4: Adjusted Odds Ratios for Low Birthweight by Pregnancy-Average Nitrate Exposure	
Specification	48
Table 2.5: Adjusted Odds Ratios for Preterm Birth by Trimester-Specific Nitrate Exposure	
Specification	50
Table 2.6: Adjusted Odds Ratios for Low Birthweight by Trimester-Specific Nitrate Exposure	
Specification	53
Table 2.7: Race and Ethnicity Differences in Adjusted Nitrate-Birth Outcome Associations.....	56
Table 2.8: Community-Context Heterogeneity in Adjusted Nitrate–Birth Outcome Associations	
.....	60
Table 3.1: Descriptive Characteristics of the Violation-Spell Analytic Sample by SDAC Status	
.....	136
Table 3.2: Selected Adjusted Time Ratios from the Baseline Log-Logistic AFT Model	144
Table 3.3: Rule-Family Composition Sensitivity for Core SDAC and System-Size Time Ratios	
.....	148
Table 3.4: Robustness Checks for Key SDAC-by-System-Size Time-Ratio Contrasts	150
Table 4.1: Event-Study Data Structure, Outcome Construction, and Sample Support	206

Table 4.2: Horizon-Averaged Post-Enforcement Changes in Quarterly Compliance Progress, Weighted Estimates	212
Table 4.3: Community-Context Differences in Horizon-Averaged Post-Enforcement Compliance Progress, Weighted Estimates.....	214
Table 4.4: Joint Tests of Pre-Enforcement Event-Time Coefficients, Weighted and Unweighted Specifications	219

ACKNOWLEDGMENTS

I am deeply grateful to my dissertation chair, Professor David L. Feldman, for his patience, generosity, and unwavering belief in me throughout this journey. When I most needed steadiness, his mentorship gave me the ground to continue and the confidence to finish. His kindness, sincerity, and trust were constant sources of guidance and reassurance. He taught me that true mentorship joins intellectual clarity with care. His example has shaped not only the scholar I am becoming, but also the kind of mentor I hope one day to be.

I am also sincerely thankful to my committee members, Professor Jae Kim and Professor Jun Wu, and to Professor Maura Allaire, for their thoughtful feedback and encouragement at different stages of this dissertation. I am grateful to the faculty members with whom I had the privilege to study and work, whose intellectual rigor and humane guidance shaped my development as a scholar and teacher. I also thank my graduate school friends, whose companionship made this long journey less lonely and far more meaningful. I am grateful as well to my colleagues at WaterUCI for creating a collegial and supportive research environment where I could continue to learn and grow.

I thank my family for their unconditional love, enduring patience, and steadfast support throughout graduate school. Their faith in me carried me through uncertainty and helped make this dissertation possible. I am especially grateful to my grandmothers, whose quiet strength taught me resilience, compassion, and the courage to continue. Their example will remain with me in every path that follows. I am grateful, too, to the part of myself that did not give up.

To every silver lining that shone through my clouds, thank you.

VITA

Qi Bing

2010 Bachelor of Engineering, North China University of Technology, Beijing, China

2019 Master of Urban and Regional Planning, University of California, Irvine

2026 Doctor of Philosophy in Urban and Environmental Planning and Policy,
University of California, Irvine

PROFESSIONAL APPOINTMENTS

2019–2024 Teaching Assistant, University of California, Irvine

2025–2026 Graduate Student Researcher, WaterUCI,
University of California, Irvine

FIELD OF STUDY

Urban and Environmental Planning and Policy

ABSTRACT OF THE DISSERTATION

Unequal Protection under the Safe Drinking Water Act: Exposure, Compliance Recovery, and Enforcement Responsiveness in California Community Water Systems

by

Qi Bing

Doctor of Philosophy in Urban and Environmental Planning and Policy

University of California, Irvine, 2026

Professor Emeritus David L. Feldman, Chair

This dissertation examines unequal protection in California drinking-water governance under the Safe Drinking Water Act. It asks how formally uniform regulatory standards translate into uneven protection when exposure risk, compliance recovery, and enforcement responsiveness are mediated by community water system capacity and regulatory implementation. Using statewide administrative, regulatory, water-quality, and birth-record data, the study links California births to community water system nitrate monitoring from 2011 to 2022, models time to recorded return to compliance for Safe Drinking Water Act violations, and estimates changes in quarterly compliance progress around each system's first formal enforcement action. The birth-outcome analysis finds a consistent positive association between pregnancy-average nitrate exposure and preterm birth across continuous, categorical, and threshold specifications, while low birthweight associations are weaker and more specification-dependent. The compliance-duration analysis shows that unequal protection after violations occur is not captured by a uniform demographic penalty. Instead, the clearest policy-relevant pattern appears where severely disadvantaged community status intersects with system size, indicating that recorded recovery depends on the capacity to diagnose, finance, implement, verify, and close corrective action. The enforcement analysis extends this governance perspective by treating first formal enforcement as an administrative threshold that can move

persistent noncompliance from informal compliance management into a more structured accountability pathway. Formal escalation is associated with modest and uneven changes in compliance progress, indicating that clearer obligations and heightened follow-through do not automatically translate into comparable corrective momentum across community contexts. Across these stages, the dissertation shows that unequal protection under the Safe Drinking Water Act emerges not only through uneven exposure burden, but also through the institutional processes that determine how quickly violations are resolved and whether enforcement produces timely, durable protection. The findings contribute to environmental justice, public policy, and drinking-water governance scholarship by linking exposure, recovery capacity, and regulatory responsiveness within a single implementation framework.

Chapter 1: Introduction

1.1 Motivation and Problem Statement

1.1.1 Drinking Water Governance as a Public Health Problem

Ensuring safe drinking water in the United States depends not only on the existence of health-based standards, but also on the ability of public institutions and regulated systems to detect problems, respond to violations, and sustain corrective action over time. In community water systems, compliance is produced through a governance process that connects federal and state rules to local implementation under technical, financial, and organizational constraints. When that process functions poorly, unsafe conditions may persist, remediation may be delayed, and harms may be distributed unevenly across communities. These risks are especially consequential in systems with limited technical capacity, constrained rate bases, and weak access to interim protective options.

These implementation challenges are organized through the Safe Drinking Water Act (SDWA) and carried out primarily through state primacy programs. Although the SDWA is often understood as a public health statute, it also functions as an administrative system through which monitoring results are reported, compliance status is classified, corrective actions are undertaken, and regulators decide whether and when to escalate from routine oversight to formal enforcement. The pace and form of these processes shape how quickly violations are identified, addressed, and resolved. As a result, the same violation can imply different levels of risk depending on how long it persists and whether affected households have the resources to reduce exposure during periods of noncompliance.

1.1.2 Environmental Justice, Compliance Duration, and Enforcement Responsiveness

This dissertation starts from the premise that compliance duration and enforcement responsiveness are central dimensions of regulatory performance under the SDWA. A violation is not only an indicator that a system has fallen below required standards, but also a time-bound episode during which served communities may face prolonged exposure, uncertainty, and costly coping behaviors. From an environmental justice perspective, the concern is therefore not only whether some communities experience more violations, but also whether they experience longer periods of noncompliance and weaker institutional responsiveness once problems occur.

Environmental justice is especially salient in drinking water governance because both recovery capacity and adaptive capacity are unevenly distributed. Public water systems' compliance trajectories often reflect differences in technical expertise, staffing, financing, and access to engineering support rather than simple willingness to comply. Households' ability to reduce exposure during periods of noncompliance is also unequal. When recovery is slow, the costs of purchasing bottled water, installing filtration, or coping with unreliable service can fall disproportionately on low-income households and communities with fewer alternatives, turning compliance delays into uneven protection even when regulatory standards are formally uniform.

These concerns are closely tied to how enforcement unfolds in practice. When government agencies regulate other public entities, enforcement may carry political and intergovernmental costs that make harsh punitive strategies less attractive (Wilson & Rachal, 1977; Konisky & Teodoro, 2016). Regulators often begin with routine oversight, informal actions, and technical assistance rather than stricter formal measures, because formal action against small, under-resourced community water systems can raise concerns about affordability, administrative burden, and the stability of essential service provision. Yet cooperative

approaches are most effective when backed by a credible capacity to escalate if progress stalls (Braithwaite, 2011). Formal enforcement escalation is therefore important not as an end in itself, but as a mechanism for moving persistent noncompliance into a more structured process of accountability and required corrective action, thereby reducing the risk that prolonged delay falls disproportionately on some communities (Ayres & Braithwaite, 1992).

1.1.3 California as a Policy-relevant Setting

California provides a consequential and policy-relevant setting for studying SDWA governance. The state includes a large and diverse community water system landscape, ranging from major metropolitan providers to many small systems. California is also distinctive in the way drinking-water equity has been formalized in state policy. The state recognizes a human right to safe, clean, affordable, and accessible water and has built implementation capacity around that commitment through targeted programs for small and disadvantaged systems, including the Safe and Affordable Funding for Equity and Resilience program created by SB 200 (California Water Code § 106.3; California State Water Resources Control Board [State Water Board], n.d.-a, n.d.-b; California Grants Portal, 2025). In parallel, California maintains its own drinking water standards in Title 22 and has adopted state-specific maximum contaminant levels, including contaminants without federal MCLs, illustrating an institutional commitment to drinking water protection that can exceed national baselines (State Water Board, 2024a, 2025).

California is also especially informative given the geographic concentration of persistent safe-drinking-water challenges in the San Joaquin Valley and the broader Central Valley. These regions include a large share of the state's low-income communities and have experienced chronic noncompliance in community water systems. Statewide assessments indicate that hundreds of systems have been out of compliance in recent years, that prolonged noncompliance

is a recurring problem for some systems, and that more than half of noncompliant systems are located in the San Joaquin Valley within the Central Valley (Chappelle et al., 2021). These patterns likely reflect both administrative realities and the region's distinctive hydrogeologic and agricultural context. Central Valley is characterized by high groundwater dependence and intensive agriculture, where drought increases reliance on groundwater and persistent nitrate contamination remains a central drinking water concern (State Water Board, 2013). For this reason, Central Valley provides an especially important context for understanding how nitrate contamination, baseline vulnerability, and remediation challenges intersect within California's drinking water landscape.

1.2 Objectives, Research Questions, and Contributions

1.2.1 Dissertation Objectives

This dissertation is organized as a three-essay study with the overarching objective of assessing whether drinking water governance under the SDWA in California produces patterns consistent with unequal protection under capacity constraints. It pursues that objective by examining three distinct but connected dimensions of implementation and risk. The first essay examines whether nitrate exposure in community water systems is associated with adverse birth outcomes. The second essay analyzes time to return to compliance by modeling the duration of noncompliance spells and identifying the conditions associated with faster or slower recovery. The third essay evaluates whether first formal enforcement escalation is associated with changes in compliance progress over time. Across the three essays, the dissertation also examines whether these patterns vary across disadvantaged community contexts and other relevant dimensions of capacity and vulnerability.

1.2.2 Research Questions

To operationalize these objectives, the dissertation addresses three overarching research questions.

Exposure and health. Are higher nitrate levels in community drinking water systems associated with adverse birth outcomes, and are these associations patterned by maternal demographics and place-based context, including disadvantaged communities and Central Valley geography?

Compliance duration. Which system and community characteristics are associated with longer noncompliance spells, and how do system size and disadvantaged-community context shape variation in time to return to compliance?

Enforcement responsiveness. Is a system's first formal enforcement action followed by measurable changes in compliance progress, and do those dynamics differ across disadvantaged and non-disadvantaged community contexts?

1.2.3 Contributions

This dissertation contributes to research on drinking water governance by examining three underdeveloped dimensions of SDWA implementation. Existing studies have more often emphasized violation incidence, system characteristics, or cross-sectional disparities than the public health implications of drinking water exposure, the duration of unresolved noncompliance, or compliance dynamics following formal enforcement escalation. By bringing these dimensions into a common framework, the dissertation advances an environmental justice perspective that treats protection not only as a question of whether violations occur, but also as a question of how exposure risk is distributed, how long protection remains incomplete, and whether institutional intervention produces meaningful improvement across communities.

Analytically and from a policy perspective, the dissertation centers on the community water system as the relevant unit of governance. This focus links health outcomes, compliance recovery, and enforcement dynamics to the institutional level at which compliance obligations and regulatory oversight are organized. It strengthens the policy relevance of the analysis by defining heterogeneity comparisons on a common governance-relevant basis and by highlighting performance measures more closely tied to lived protection, including noncompliance duration and post-escalation compliance progress. In doing so, the dissertation highlights the importance of oversight, technical assistance, and financing support in contexts where prolonged noncompliance and uneven protection are most likely to persist.

1.3 Institutional and Regulatory Context

1.3.1 The SDWA and Enforcement

The SDWA is the primary federal statute governing public drinking water quality in the United States. It authorizes the U.S. Environmental Protection Agency (EPA) to establish National Primary Drinking Water Regulations, including enforceable maximum contaminant levels and related requirements for monitoring, treatment, and public reporting (EPA, n.d.-a, n.d.-b; EPA, 2025d). In practice, this framework links contaminant-specific standards to routine compliance monitoring and reporting, creating the informational basis for oversight and enforcement when violations occur.

Although drinking water standards are federal, implementation is primarily carried out through state primacy programs. EPA may delegate primary enforcement responsibility to states and eligible tribes that meet statutory and regulatory requirements, after which those agencies assume the day-to-day responsibility for determining compliance and initiating enforcement actions for public water systems (EPA, 2025a; Electronic Code of Federal Regulations, n.d.).

EPA retains oversight authority and may intervene in limited circumstances, but regulatory performance depends in practice on the administrative capacity of primacy agencies as well as the organizational capacity of regulated systems (Electronic Code of Federal Regulations, n.d.).

Within this system, enforcement operates along a spectrum from informal compliance management to formal legal action. Informal actions, including warning letters, follow-up calls, meetings, and technical assistance, often function as early compliance prompts intended to encourage correction without creating independently enforceable obligations (EPA, 2025b).

When violations persist or recur, regulators may escalate to formal enforcement, which places the system on a more explicit and enforceable path toward return to compliance through instruments such as administrative orders or judicial actions (EPA, 2026b, 2025c). This sequence is consistent with EPA’s Public Water System Supervision Enforcement Response Policy, which emphasizes timely and appropriate responses to violations and recognizes a range of regulatory tools, from compliance assistance and informal communication to administrative orders and judicial actions (EPA, 2009a, 2009b).

1.3.2 Violations, Noncompliance, and Return to Compliance

Under the SDWA, compliance is defined through the NPDWRs, which establish enforceable standards and monitoring requirements for public water systems. EPA sets legal limits for regulated contaminants and specifies the testing schedules and analytical methods systems must follow (EPA, 2025d). In regulatory terms, a maximum contaminant level is “the maximum permissible level of a contaminant in water which is delivered to any user of a public water system” (40 C.F.R. § 141.2, 2025). When a system exceeds an applicable standard or fails to meet required monitoring or reporting obligations, it incurs a violation under the SDWA. Restoring compliance may then require actions such as conducting required monitoring,

submitting missing reports, implementing treatment or operational changes, and providing health-related communication when required (EPA, 2009b, 2025e).

These compliance concepts are operationalized in administrative data systems through dated records of violation status and closure. In SDWIS Federal Reporting Services, each violation is tracked using key dates and status fields that define a noncompliance timeline. A violation remains unresolved when its compliance status is recorded as Open or Known, and it is not treated as closed until it is recorded as returned to compliance (EPA, 2026). EPA's enforcement guidance reinforces this distinction by treating return to compliance, rather than partial response alone, as the operational endpoint that should be reflected in federal and state data systems (EPA, 2009a, 2009b). In this way, administrative timelines do more than document recordkeeping milestones, they make visible the observed pace at which regulated problems move from detection and documentation to correction and formal closure within the SDWA implementation process.

1.3.3 Regulatory Discretion and Capacity Constraints

Even under uniform SDWA standards, regulatory oversight and enforcement escalation can vary in practice, reflecting agency resource constraints, targeting decisions made under uncertainty, and heterogeneous local conditions. Differences in staffing, budgets, workload, and organizational structure may shape how frequently systems are monitored, how quickly agencies escalate, and how consistently follow-up occurs, particularly as regulatory demands expand (Government Accountability Office, 2009; Zhao & Lyon, 2022). Enforcement behavior may also reflect prior noncompliance, agency structure, political oversight, and the capacity of communities to draw attention to unresolved problems (Harrington, 1988; Rousseau, 2007; Konisky, 2009; Konisky & Reenock, 2013). These constraints and discretionary processes help

explain why regulatory responsiveness may differ across systems and communities even when formal standards are the same.

1.4 Conceptual Framework

1.4.1 Integrated Conceptual Model

This dissertation treats compliance not as a static status but as a time-varying process, and exposure not only as recorded regulatory exceedance but as a continuum of contaminant conditions experienced through delivered drinking water. The conceptual model links these elements through timing, capacity, and institutional response. Regulatory action shapes compliance progress and the duration of unresolved noncompliance, but those relationships depend on remediation feasibility and administrative follow-through. Exposure burden depends not only on whether contamination occurs, but also on how long unsafe or elevated conditions persist. From this perspective, environmental justice concerns arise when exposure risk, recovery capacity, and regulatory responsiveness are unevenly distributed across communities. Figure 1.1 synthesizes this framework by showing how regulatory implementation, protection pathways, and public health consequences connect the three empirical chapters to the dissertation-wide argument about unequal protection.

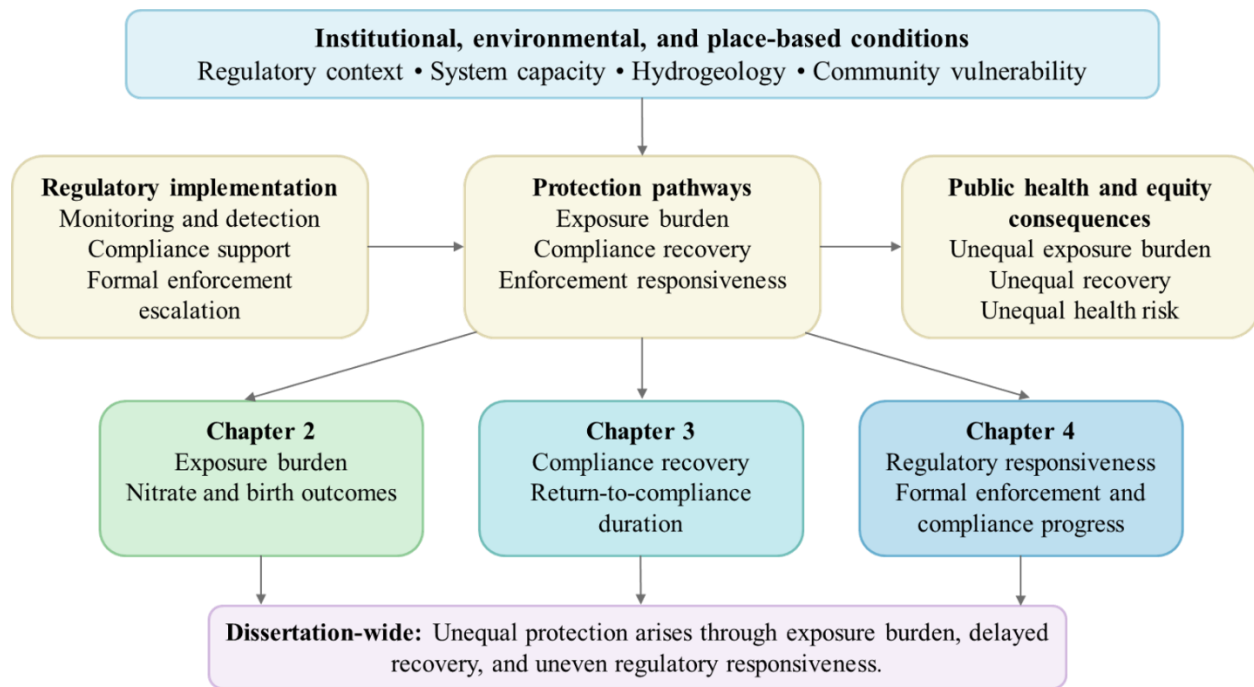


Figure 1.1: Integrated Conceptual Framework of Unequal Protection Under Capacity Constraints

Note. This figure shows how the three empirical chapters connect within the dissertation’s broader argument. Unequal protection can arise through uneven exposure burden, compliance recovery, and enforcement timing, each shaped by institutional, environmental, and place-based conditions.

1.4.2 Essay 1: Nitrate Exposure and Birth Outcomes

The first essay examines whether nitrate exposure in community drinking water systems is associated with adverse birth outcomes in California. Within the dissertation’s broader framework, this essay captures unequal protection through exposure and health burden. It treats drinking water exposure not only as an environmental condition, but also as a governance-mediated condition shaped by source-water vulnerability, monitoring, treatment capacity, and the ability of households and systems to reduce risk. In this way, the chapter links drinking water governance to early-life health risk and establishes the dissertation’s exposure-to-health dimension of unequal protection.

1.4.3 Essay 2: Compliance Duration and Return to Compliance

The second essay examines unequal protection through compliance recovery capacity. It treats return to compliance as a governance outcome, because prolonged unresolved noncompliance implies longer periods during which regulatory protection remains incomplete. In this framework, duration is meaningful not simply as elapsed time, but as an indicator of how quickly institutions and regulated systems move from violation detection to formally recognized recovery. Conceptually, the essay treats return to compliance as a feasibility-constrained recovery process rather than as a simple byproduct of enforcement intensity alone. Once a violation is identified, the pace of recorded resolution depends on whether a system can diagnose the problem, mobilize technical and financial resources, implement corrective action, and satisfy the documentation and verification steps required for formal closeout. Unequal protection therefore arises not only through where violations occur, but also through whether some communities remain in unresolved noncompliance longer because recovery is harder to achieve and formal closure is slower to complete.

1.4.4 Essay 3: Formal Enforcement Escalation and Compliance Progress

The third essay examines unequal protection through regulatory responsiveness. It treats first formal enforcement as an analytically meaningful administrative threshold because it marks the point at which a case moves from informal compliance management to a more explicit accountability pathway. Conceptually, the essay asks whether formal escalation is associated with changes in the pace of compliance progress, while recognizing that escalation timing is itself shaped by prior noncompliance, regulator discretion, and institutional constraints. In contexts where many compliance barriers are technical, financial, and organizational rather than purely motivational, formal escalation may matter less through punishment alone than through

clearer obligations, heightened accountability, and more structured follow-through. From an environmental justice perspective, the central concern is whether this shift produces comparable corrective momentum across communities or whether post-enforcement compliance progress remains weaker where capacity constraints are more binding.

1.5 Dissertation Roadmap

This dissertation proceeds in three empirical chapters. Chapter 2 examines whether nitrate exposure in community water systems is associated with adverse birth outcomes in California and whether estimated associations vary across maternal and place-based contexts. Chapter 3 analyzes time to return to compliance by modeling the duration of noncompliance spells and assessing how recovery varies across system and community conditions. Chapter 4 examines whether a system's first formal enforcement action is followed by changes in compliance progress and whether those dynamics differ across disadvantaged and non-disadvantaged community contexts. Together, these chapters examine exposure, compliance recovery, and regulatory responsiveness as complementary dimensions of SDWA implementation.

1.6 References

- Ayres, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. Oxford University Press.
- Balazs, C. L., & Ray, I. (2014). The drinking water disparities framework: On the origins and persistence of inequities in exposure. *American Journal of Public Health*, 104(4), 603–611.
- Balazs, C. L., Morello-Frosch, R., Hubbard, A., & Ray, I. (2011). Social disparities in nitrate-contaminated drinking water in California's San Joaquin Valley. *Environmental Health Perspectives*, 119(9), 1272–1278.
- Baldwin, R., & Black, J. (2008). Really responsive regulation. *The Modern Law Review*, 71(1), 59–94.

- Braithwaite, J. (2011). The essence of responsive regulation. *UBC Law Review*, 44(3), 475–520.
- California Grants Portal. (2025, September 24). *The Safe and Affordable Funding for Equity and Resilience Program (SAFER)*. <https://www.grants.ca.gov/grants/the-safe-and-affordable-funding-for-equity-and-resilience-program-safer/>
- California State Water Resources Control Board. (n.d.-a). *Human right to water*. https://www.waterboards.ca.gov/water_issues/programs/hr2w/
- California State Water Resources Control Board. (n.d.-b). *Transfer of drinking water programs*. https://www.waterboards.ca.gov/drinking_water/programs/DW_PreJuly2014.html
- California State Water Resources Control Board. (2013, February 20). *Recommendations addressing nitrate in groundwater: Report to the Legislature*. https://www.waterboards.ca.gov/water_issues/programs/nitrate_project/docs/nitrate_rpt.pdf
- California State Water Resources Control Board. (2024a, August 22). *1,2,3-Trichloropropane (1,2,3-TCP)*. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/123TCP_page.html
- California State Water Resources Control Board. (2024b, November 7). *Water quality enforcement policy*. https://www.waterboards.ca.gov/water_issues/programs/enforcement/docs/2024/2024-enforcement-policy.pdf
- California State Water Resources Control Board. (2025, April 21). *Chemicals and contaminants in drinking water*. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/Chemicalcontaminants.html
- California State Water Resources Control Board. (2026, February 11). *SAFER Drinking Water*. <https://www.waterboards.ca.gov/safer/>
- Chappelle, C., Collins, J., & Hanak, E. (2021, April 2). *Access to safe drinking water in California*. Public Policy Institute of California. <https://www.ppic.org/publication/access-to-safe-drinking-water/>
- Clemmensen, P. J., Schullehner, J., Brix, N., Sigsgaard, T., Stayner, L. T., Kolstad, H. A., & Ramlau-Hansen, C. H. (2023). Prenatal exposure to nitrate in drinking water and adverse health outcomes in the offspring: A review of current epidemiological research. *Current Environmental Health Reports*, 10(3), 250–263.
- Coffman, V. R., Jensen, A. S., Trabjerg, B. B., Pedersen, C. B., Hansen, B., Schullehner, J., Pedersen, M., & Stayner, L. T. (2022). Prenatal exposure to nitrate from drinking water and the risk of preterm birth: A Danish nationwide cohort study. *Environmental Epidemiology*, 6(5), e223.
- Currie, J., Graff Zivin, J., Meckel, K., Neidell, M., & Schlenker, W. (2013). Something in the water: Contaminated drinking water and infant health. *Canadian Journal of Economics/Revue canadienne d'économie*, 46(3), 791–810.

- DiSalvo, R. W., & Hill, E. L. (2024). Drinking water contaminant concentrations and birth outcomes. *Journal of Policy Analysis and Management*, 43(2), 368–399.
- Electronic Code of Federal Regulations. (n.d.). 40 CFR Part 142, Subpart B—Primary enforcement responsibility for public water systems. <https://www.ecfr.gov/current/title-40/chapter-I/subchapter-D/part-142/subpart-B>
- Government Accountability Office. (2009, October 15). *Clean Water Act: Longstanding issues impact EPA's and states' enforcement efforts* (GAO-10-165T). <https://www.gao.gov/products/gao-10-165t>
- Gray, W. B., & Shimshack, J. P. (2011). The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*, 5(1), 3–24.
- Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics*, 37(1), 29–53.
- Harrison, J. L. (2019). *From the inside out: The fight for environmental justice within government agencies*. MIT Press.
- Kang, K., & Silveira, B. S. (2021). Understanding disparities in punishment: Regulator preferences and expertise. *Journal of Political Economy*, 129(10), 2947–2992.
- Keohane, R. O., Macedo, S., & Moravcsik, A. (2009). Democracy-enhancing multilateralism. *International Organization*, 63(1), 1–31.
- Kirchhoff, C. J., Flagg, J. A., Zhuang, Y., & Utemuratov, B. (2019). Understanding and improving enforcement and compliance with drinking water standards. *Water Resources Management*, 33(5), 1647–1663.
- Konisky, D. M., & Reenock, C. (2013). Compliance bias and environmental (in)justice. *The Journal of Politics*, 75(2), 506–519.
- Konisky, D. M., & Teodoro, M. P. (2016). When governments regulate governments. *American Journal of Political Science*, 60(3), 559–574.
- Kubrin, C. E., Stucky, T. D., & Krohn, M. D. (2009). *Researching theories of crime and deviance*. Oxford University Press.
- Legal Information Institute. (n.d.). 40 CFR § 141.2—Definitions. <https://www.law.cornell.edu/cfr/text/40/141.2>
- Lin, L., St Clair, S., Gamble, G. D., Crowther, C. A., Dixon, L., Bloomfield, F. H., & Harding, J. E. (2023). Nitrate contamination in drinking water and adverse reproductive and birth outcomes: A systematic review and meta-analysis. *Scientific Reports*, 13(1), 563.
- Mennis, J. L. (2005). The distribution and enforcement of air polluting facilities in New Jersey. *The Professional Geographer*, 57(3), 411–422.
- Opp, S. M. (2012). Environmental justice and the Resource Conservation Recovery Act inspection and enforcement process. *International Review of Public Administration*, 17(1), 179–192.
- Ransom, K. M., Nolan, B. T., Traum, J. A., Faunt, C. C., Bell, A. M., Gronberg, J. A. M., Wheeler, D. C., Rosecrans, C. Z., Jurgens, B., Schwarz, G. E., Belitz, K., Eberts, S. M.,

- Kourakos, G., & Harter, T. (2017). A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA. *Science of the Total Environment*, 601–602, 1160–1172.
- Roberson, J. A., & Frey, M. M. (2016). An SDWA retrospective: 20 years after the 1996 amendments. *Journal AWWA*, 108(3), 22–30.
- Rousseau, S. (2007). *Economic empirical analysis of sanctions for environmental violations: A literature overview* (Working Paper No. 2007-03). Katholieke Universiteit Leuven, Centre for Economic Studies, Energy, Transport & Environmental Economics Working Paper Series.
- Royal, H., 't Mannetje, A., Hales, S., Douwes, J., Berry, M., & Chambers, T. (2024). Nitrate in drinking water and adverse pregnancy outcomes: A narrative review of epidemiological evidence and proposed biological mechanisms. *PLOS Water*, 3(1), e0000214.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328), 918–924.
- Schaider, L. A., Swetschinski, L., Campbell, C., & Rudel, R. A. (2019). Environmental justice and drinking water quality: Are there socioeconomic disparities in nitrate levels in U.S. drinking water? *Environmental Health*, 18, Article 3.
- Sherris, A. R., Baiocchi, M., Fendorf, S., Luby, S. P., Yang, W., & Shaw, G. M. (2021). Nitrate in drinking water during pregnancy and spontaneous preterm birth: A retrospective within-mother analysis in California. *Environmental Health Perspectives*, 129(5), 057001.
- Spina, F. (2015). Environmental justice and patterns of state inspections. *Social Science Quarterly*, 96(2), 417–429.
- State of California. (2024). *California Water Code § 106.3*. https://www.loginfo.ca.gov/pub/11-12/bill/asm/ab_0651-0700/ab_685_bill_20120925_chaptered.pdf
- U.S. Environmental Protection Agency. (n.d.-a). *National Primary Drinking Water Regulations*. <https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations>
- U.S. Environmental Protection Agency. (n.d.-b). *Safe Drinking Water Act (SDWA)*. <https://www.epa.gov/sdwa>
- U.S. Environmental Protection Agency. (2009a, December 8). *Enforcement Response Policy for the Public Water System Supervision (PWSS) Program under the Safe Drinking Water Act and implementation of the Enforcement Targeting Tool*. <https://www.epa.gov/enforcement/enforcement-response-policy-public-water-system-supervision-pwss-program-under-safe>
- U.S. Environmental Protection Agency. (2009b, December 8). *Proposed revision to Enforcement Response Policy for the Public Water System Supervision (PWSS) Program under the Safe Drinking Water Act and implementation of the Enforcement Targeting Tool* [Memorandum]. Office of Enforcement and Compliance Assurance. https://www.epa.gov/sites/default/files/documents/drinking_water_erp_2009.pdf

- U.S. Environmental Protection Agency. (2024, December 5). *Safe Drinking Water Act (SDWA) resources and FAQs*. <https://echo.epa.gov/help/sdwa-faqs>
- U.S. Environmental Protection Agency. (2025a, October 23). *Primacy enforcement responsibility for public water systems*. <https://www.epa.gov/dwreginfo/primacy-enforcement-responsibility-public-water-systems>
- U.S. Environmental Protection Agency. (2025b, September 30). *What is a Notice of Violation (NOV) letter?* <https://www.epa.gov/enforcement/what-notice-violation-nov-letter>
- U.S. Environmental Protection Agency. (2025c, March 3). *Criminal provisions of the Safe Drinking Water Act (SDWA)*. <https://www.epa.gov/enforcement/criminal-provisions-safe-drinking-water-act-sdwa>
- U.S. Environmental Protection Agency. (2025d, November 25). *Drinking water regulations*. <https://www.epa.gov/dwreginfo/drinking-water-regulations>
- U.S. Environmental Protection Agency. (2025e, June 11). *Public Notification Rule*. <https://www.epa.gov/dwreginfo/public-notification-rule>
- U.S. Environmental Protection Agency. (2026, January 20). *SDWIS Federal Reporting Services data dictionary*. <https://www.epa.gov/DWdata/sdwis-federal-reporting-services-data-dictionary>
- Ward, M. H., Jones, R. R., Brender, J. D., de Kok, T. M., Weyer, P. J., Nolan, B. T., Villanueva, C. M., & van Breda, S. G. J. (2018). Drinking water nitrate and human health: An updated review. *International Journal of Environmental Research and Public Health*, 15(7), 1557.
- Wilson, J. Q., & Rachal, P. (1977). Can the government regulate itself? *The Public Interest*, 46, 3–14.
- Zhao, D., & Lyon, T. P. (2022). Why does general deterrence exist? A study of drinking water quality regulation. *SSRN Electronic Journal*.

Chapter 2: Nitrate in Community Drinking Water and Birth Outcomes in California

2.1 Chapter Overview

This chapter estimates the association between nitrate in community drinking water and adverse birth outcomes in California, with particular attention to whether estimated associations differ across disadvantaged contexts. The analysis links individual birth records to community water system service areas and assigns pregnancy-average nitrate exposure using monitoring data, while adjusting for maternal and infant characteristics and time fixed effects with inference clustered at the water system level. To align results with both epidemiologic and regulatory questions, nitrate exposure is parameterized three ways, including a continuous concentration measure, a categorical dose measure, and a threshold contrast at 10 milligrams per liter as nitrogen.

Given that drinking water quality is shaped by groundwater dependence, agricultural nitrate loading, and water systems' unequal capacity for avoidance and mitigation, this chapter frames nitrate exposure as an environmental justice concern. It assesses heterogeneity along two complementary dimensions. First, it tests whether estimated associations differ by maternal race and ethnicity, with a focus on Hispanic and non-Hispanic Black mothers. Second, it tests whether associations vary across place and system contexts, including disadvantaged community status, severely disadvantaged community status, and Central Valley geography, where nitrate contamination and groundwater stress are especially salient. Heterogeneity patterns are interpreted as evidence that structural conditions, including place-based disadvantage and water system constraints, shape exposure profiles, coexposures, and baseline vulnerability rather than nitrate-specific biological susceptibility. The results are interpreted as adjusted associations

rather than causal effects, but they establish an empirical benchmark and identify where nitrate-related risks appear most concentrated, providing a foundation for future causal research and policy evaluation.

2.2 Literature Review

2.2.1 Nitrate and Health

Nitrate can influence pregnancy and fetal development through pathways linked to its reduction to nitrite, including oxygen transport disruption and endogenous nitrosation processes. Evidence from toxicology, clinical physiology, and epidemiology therefore supports examining nitrate exposure during pregnancy and assessing whether associations differ across contexts and population groups (Kardol et al., 2025; Royal et al., 2024; Clemmensen et al., 2023). One pathway involves oxygen transport impairment through nitrite and methemoglobin formation. In the body, nitrate can be reduced to nitrite, and nitrite can oxidize hemoglobin to methemoglobin, reducing oxygen carrying capacity. Although clinically overt methemoglobinemia is most clearly documented in infants, the underlying oxygen transport mechanism is relevant to pregnancy given the sensitivity of fetal growth and gestational maintenance to oxygen availability (Knobeloch et al., 2000; Ward et al., 2005; Royal et al., 2024).

A related pathway involves endogenous nitrosation and the formation of N nitroso compounds. Nitrite can participate in reactions in the body that generate these compounds, which have been studied in relation to reproductive toxicity, developmental outcomes, and cancer risk. Prior studies emphasize that nitrosation potential can vary with diet, infection or inflammation, and coexposures, and that nitrate and nitrite may therefore operate within broader exposure mixtures rather than in isolation (Manassaram et al., 2007; Ward et al., 2018; Clemmensen et al., 2023; Royal et al., 2024). Consistent with this framing, several studies consider contaminant

mixtures or correlated coexposures that may track with nitrate in drinking water and shape observed associations (Currie et al., 2013; Schaider et al., 2019; DiSalvo & Hill, 2024).

Nitrate has also been linked to thyroid-related pathways that may be relevant during pregnancy, given the role of thyroid hormones in supporting fetal neurodevelopment and growth (Ward et al., 2010; Ward et al., 2018). In addition, evidence connecting nitrate exposure to longer-run outcomes, including certain cancers, supports the broader premise that nitrate can have systemic biological effects that may be especially consequential during sensitive life stages such as pregnancy (Weyer et al., 2001; De Roos et al., 2003; Brender & Weyer, 2016; Jones et al., 2016).

2.2.2 Nitrate in Drinking Water and Birth Outcomes

Evidence linking drinking water nitrate to preterm birth has expanded in the past decade, with increasing attention to study designs that reduce confounding by stable maternal and household characteristics. The most directly relevant study for this chapter is Sherris et al. (2021), which links California births to public water system monitoring and applies within-mother comparisons, reporting elevated risk of spontaneous preterm birth at higher nitrate levels and evidence of associations below the federal standard. Existing studies in other settings generally support the relevance of tap-water nitrate exposure for pregnancy outcomes, although findings are not uniform across settings. Danish register-based analyses link household nitrate exposure to preterm birth in large administrative data, and meta-analytic synthesis across cohort studies indicates a small but measurable elevation in preterm birth risk, including specifications parameterized per-unit increases in nitrate (Coffman et al., 2022; Lin et al., 2023). Some cohorts, however, report null associations (Albouy-Llaty et al., 2016).

In addition to preterm birth, a second core outcome domain is fetal growth, where the literature underscores that drinking-water risks can be unevenly borne. Currie et al. (2013) link drinking-water testing results to New Jersey birth records and estimate effects on both low birthweight and prematurity, with larger adverse impacts among infants born to less educated mothers. A more recent study further emphasizes that health-relevant exposure variation may occur below formal regulatory triggers. DiSalvo and Hill (2024) use Pennsylvania sampling data geocoded to community water system boundaries and mother fixed effects, finding that higher contaminant concentrations among births with no violations are associated with higher risks of low birthweight and preterm birth, which reinforces the importance of modeling measured concentrations rather than relying only on violations. Additional evidence from U.S. agricultural settings examines finished-water nitrate, sometimes alongside co-occurring contaminants, in relation to birth outcomes including low birthweight and preterm birth (Stayner et al., 2016).

A related literature extends beyond these primary endpoints to other pregnancy and early-life outcomes. Register-based studies link tap-water nitrate exposure to small-for-gestational-age risk (Jensen et al., 2023). A smaller parallel literature evaluates congenital anomalies in relation to nitrate exposure or correlated agricultural contaminants and emphasizes variability across outcomes and exposure assessment approaches (Manassaram et al., 2007; Ward et al., 2018; Brender & Weyer, 2016; Brender et al., 2013). Others emphasize that estimated effects are often modest and sensitive to exposure definition, assignment, and confounding control (Clemmensen et al., 2023; Royal et al., 2024; Ward et al., 2018).

Across the literature, studies use cohort, case-control, and quasi-experimental designs, and they assign exposure using community water system monitoring data, modeled nitrate surfaces, or linked household estimates (Coffman et al., 2022; Jensen et al., 2023; Sherris et al.,

2021). Exposure definitions also vary, including system-average concentrations, residence-linked monitoring zones, and proxy indicators such as drinking water violations, which affects comparability and can introduce misclassification (Clemmensen et al., 2023; Lin et al., 2023; Royal et al., 2024; Ward et al., 2018; Young et al., 2023). Several studies use within-mother designs or finer-scale exposure assignment to reduce confounding by stable family characteristics and contextual differences (Currie et al., 2013; DiSalvo & Hill, 2024; Sherris et al., 2021). These design considerations are especially relevant for California, where nitrate concentrations, water sources, and sociodemographic composition vary sharply across space and system types (Balazs et al., 2011; Schaider et al., 2019). Even with these advances, relatively few studies combine pregnancy-period nitrate exposure linked to community water system service areas with policy-salient exposure specifications and environmental justice heterogeneity in groundwater-dependent settings where exposure and disadvantage plausibly co-occur. These themes inform the empirical strategy in this chapter, including water-system-linked exposure assignment, complementary continuous and threshold-based nitrate specifications, and heterogeneity analyses by maternal race and ethnicity and by place-based disadvantage, including Central Valley context.

2.2.3 Drinking Water Nitrate Regulation

In the United States, nitrate is regulated under the Safe Drinking Water Act (Safe Drinking Water Act, 42 U.S.C. § 300f et seq., 1974) through the National Primary Drinking Water Regulations (National Primary Drinking Water Regulations, 40 C.F.R. § 141.62, 2024). The SDWA establishes enforceable standards and monitoring mandates for public water systems, and EPA guidance clarifies how nitrate compliance is determined when laboratories report combined nitrate and nitrite measures (EPA, 2020). The current federal maximum contaminant

level for nitrate is 10 mg/L as nitrate-nitrogen (NO₃-N), a threshold historically motivated by prevention of infant methemoglobinemia rather than pregnancy or fetal-development endpoints (Knobeloch et al., 2000; Ward et al., 2005; Ward et al., 2018).

A key implication for birth-outcome research is that this compliance threshold may not align with the lowest observable risk levels for outcomes such as preterm birth or fetal growth restriction. Violation-based measures can miss health-relevant variation below the MCL and can also conflate exposure with monitoring and enforcement patterns that differ by system characteristics (DiSalvo & Hill, 2024; Pennino et al., 2017; Ward et al., 2018). These considerations suggest that pregnancy-relevant pathways may operate at concentrations below the methemoglobinemia-focused MCL, reinforcing the value of continuous exposure specifications (Clemmensen et al., 2023; Royal et al., 2024; Ward et al., 2018). These issues also shape how nitrate exposure is operationalized in epidemiologic studies, motivating both threshold-based and continuous exposure specifications. This regulatory history motivates the chapter's focus on birth outcomes, since the nitrate standard was developed around infant methemoglobinemia rather than prenatal endpoints such as preterm birth or fetal growth.

2.2.4 Drought, Groundwater Depletion, and Nitrate Risk in California

California's warming-driven hydroclimate variability increases surface water uncertainty and drought intensity, which heightens reliance on groundwater and can intensify depletion and overdraft in basins where withdrawals exceed recharge (Wada et al., 2010; Famiglietti, 2014; Swain et al., 2018). The state's drought conditions illustrate how water quantity and water quality are intertwined for groundwater-dependent households and community water systems. Sustained groundwater depletion can increase dependence on groundwater in regions where nitrate contamination is already persistent, while also forcing community water systems to alter

well depth, pumping patterns, or source mix in ways that change delivered water quality. In some cases, these operational shifts are associated with measurable changes in contaminant profiles when pumping depths change or when systems substitute among available sources (Feinstein et al., 2017; Pauloo et al., 2020).

Nitrate is especially relevant in this setting given its close connection to agricultural nitrogen inputs and its persistence in groundwater over long periods. In this sense, groundwater depletion is relevant not as a cause of nitrate contamination, but as a condition that can intensify reliance on already contaminated groundwater and complicate efforts to maintain a safe supply. Groundwater quality reflects decades of loading, and even strong source-reduction policies may take years to translate into improved drinking water quality in many aquifers (Ransom et al., 2017; Nolan et al., 1997). This persistence reinforces the plausibility that California's intensively irrigated basins, characterized by extensive fertilizer inputs, animal operations, and heavy groundwater reliance, face sustained nitrate risks that may be difficult to eliminate in the near term. These exposure dynamics make nitrate a natural environmental justice concern given the uneven distribution of both contamination burdens and the capacity to prevent and remediate them.

2.2.5 Environmental Justice, Governance Capacity, and Unequal Protection

California's nitrate problem sits at the intersection of environmental exposure, governance capacity, and historical inequities. The vulnerability is not only hydrogeologic. It also reflects structural conditions that shape which communities face higher baseline risk, which water systems can finance monitoring and treatment, and how quickly systems can respond when contaminants exceed health-based thresholds (Balazs & Ray, 2014). High-profile drinking water crises underscore that environmental health risks can be produced through institutional decision

making, resource allocation, and accountability failures, not only through natural conditions. The Flint water crisis is frequently cited as a case of regulatory failure and environmental injustice, illustrating how political marginalization can contribute to preventable exposure and delayed corrective action (Butler et al., 2016). Although Flint differs from California's nitrate context, the broader lesson is that accountability gaps and capacity constraints can translate formal equality in standards into unequal protection in practice (Balazs & Ray, 2014; Butler et al., 2016). A central implication of California's nitrate-drought-groundwater nexus is that technical remedies alone are unlikely to eliminate inequity without parallel attention to institutional capacity and accountability.

This environmental justice lens is reinforced by evidence that nitrate exposure is socially patterned across U.S. community water systems and linked to the social geography of who is served by which systems (Balazs & Ray, 2014; Schaider et al., 2019). Unequal protection arises when communities facing higher exposure burdens are also served by systems with fewer resources to prevent, detect, and remediate contamination. Empirical studies illustrate these patterns and clarify the mechanisms that connect land use, hydrogeology, and system capacity to inequitable exposure. In California's San Joaquin Valley, Balazs et al. (2011) find that, especially among smaller systems, higher shares of Latino residents and renters are associated with higher nitrate levels, consistent with disparities in drinking-water quality. National evidence reinforces this pattern and clarifies its determinants. Schaider et al. (2019) compile nitrate data for nearly 40,000 U.S. community water systems and show that agricultural land use and groundwater reliance are associated with higher nitrate, and that systems serving higher proportions of Hispanic residents have higher nitrate even after accounting for cropland and livestock production. Their results also highlight the scale of this exposure distribution, with over

5.6 million people served by systems averaging at least 5 mg/L NO₃-N during 2010 to 2014. In addition, statewide California evidence links drinking-water conditions to birth outcomes in a cumulative exposure framework. Huang et al. (2018) report associations between preterm birth and drinking-water contaminant indicators, including nitrate, alongside social stressors, supporting an environmental justice framing that connects exposure patterns to early-life health risk.

2.2.6 Central Valley as a Focal Nitrate and Environmental Justice Context

Against this background, Central Valley is a focal setting for environmental justice research on nitrate in California, given the coincidence of high exposure risk and uneven community and system capacity. The region combines drought-prone hydrologic conditions, intensive irrigated agriculture, and heavy reliance on groundwater for drinking water in ways that create sustained risk of elevated nitrate exposure, particularly for smaller groundwater-dependent community water systems. Evidence from other major basins worldwide links irrigated agriculture to groundwater pumping pressure, underscoring that this vulnerability is long-standing and structural (Siebert et al., 2010; Wada et al., 2010). As one of the most agriculturally productive regions in the United States, the Central Valley also faces high nitrogen loading from fertilizer and manure that can translate into persistent groundwater contamination (Levy et al., 2021; Ransom et al., 2017). These dynamics suggest that drought risk and groundwater depletion can compound existing inequities in groundwater-dependent regions, intensifying disparities in drinking water quality and affordability.

Evidence within the Central Valley illustrates how nitrate risks can align with social disadvantages and limited compliance capacity. In the San Joaquin Valley, Balazs et al. (2011) show that nitrate contamination is socially patterned, with low-income and minority communities

more likely to be served by systems with elevated nitrate. Related studies indicate that exposure and compliance burdens can fall disproportionately on disadvantaged communities and that system capacity shapes monitoring, reporting, and the ability to sustain compliance over time (Tariqi & Naughton, 2021; Williams et al., 2024). In this setting, exposure risk and limited system capacity often co-occur in the same places, so equal rules can yield unequal protection when remediation and administrative capacity differ across systems (Balazs & Ray, 2014; Ransom et al., 2017). This evidence supports a Central Valley focus and motivates the chapter's place-based heterogeneity analyses for nitrate exposure and birth outcomes.

2.2.7 Study Design Implications

The literature supports modeling drinking water nitrate as a potentially health-relevant prenatal exposure using both continuous measures and policy-salient threshold and dose range specifications (Coffman et al., 2022; Jensen et al., 2023; Lin et al., 2023; Sherris et al., 2021). It also underscores that inference depends on exposure assignment, confounding control, and selection-related biases, motivating pregnancy-average exposure measures with trimester sensitivity analyses, extensive covariate adjustment, and clustered inference at the water system level (Ward et al., 2018; DiSalvo & Hill, 2024; Clemmensen et al., 2023; Royal et al., 2024). Finally, environmental justice research highlights that exposure risk and the capacity to mitigate contamination are often aligned spatially, supporting heterogeneity analyses by maternal race and ethnicity, community disadvantage, and Central Valley location (Balazs & Ray, 2014; Schaider et al., 2019).

2.3 Data

2.3.1 Birth Certificates Data

This chapter uses individual-level birth records from the California Department of Public Health Vital Statistics Birth Statistical Master Files obtained under a research data use agreement. These files cover registered live births in California and contain information on infant outcomes, maternal characteristics, reproductive history, prenatal care, health behaviors, and payer information. The analysis focuses on births from 2011 through 2022 to align the study population with the drinking-water monitoring window used to construct nitrate exposure.

2.3.2 Drinking Water Quality Data and Nitrate Exposure Metrics

This chapter uses sample-level drinking water quality data for California public water systems from 2011 through 2022, obtained from the California State Water Resources Control Board Division of Drinking Water. The records include sampling dates, analytical results, reported units, and system identifiers and are drawn from the State Water Board's public reporting and compliance systems (State Water Board, Division of Drinking Water, 2025a).

The analysis constructs a nitrate exposure series aligned with nitrate-based drinking water standards. It retains records when the analyte corresponds to nitrate or combined nitrate-nitrite and excludes other nitrogen species subject to different regulatory treatment. Specifically, the analytic file retains observations with analyte names NITRATE, NITRATE (AS NO₃), or NITRATE-NITRITE. Numeric laboratory results are cleaned and standardized to a common nitrate metric in mg/L as nitrogen before construction of the analytic nitrate panel. The analytic nitrate panel uses usable numeric nitrate concentrations standardized to mg/L as nitrogen before exposure construction. Records that remained missing or non-quantitative in the analytic file

were excluded from continuous exposure calculations. Concentrations are right-skewed, with exceedances above the 10 mg/L as N maximum contaminant level concentrated in the upper tail.

2.3.3 CWS-Level Gestational Nitrate Concentration

This chapter links birth records to drinking water exposure by geocoding maternal residence at delivery and assigning each birth a community water system identifier corresponding to the service area expected to supply that residence. The analytic sample includes births with a valid system identifier and gestational periods that fall within the 2011 through 2022 monitoring window. This linkage supports pregnancy-average and trimester-specific exposure measures that capture temporal variation in system-level nitrate during gestation.

This chapter constructs system-level nitrate panels using monthly monitoring data. For each system-month, it computes the mean, maximum, and measurement count across all samples collected in that month. Pregnancy-average nitrate is defined as the mean of monthly system means over the gestational period, and parallel trimester-specific measures are constructed using the same logic. When a system had no nitrate monitoring in a given gestational month, pregnancy-average exposure was calculated from the available monthly values in the analytic nitrate panel; unmonitored months were not filled in during pregnancy-level exposure assignment. Temporal monitoring support varied across contributing system-years and assigned pregnancies, with the sparsest coverage concentrated in smaller systems; Appendix Table 2.A1 reports summary diagnostics for system-year monitoring cadence and pregnancy-level monitored-month support.

2.3.4 Community Water System Characteristics

This chapter assigns system characteristics and service-area sociodemographic context at the community water system level and links them to births through the public water system

identifier. It draws system characteristics from the EPA Safe Drinking Water Information System inventory for California and restricts the sample to community water systems that serve the same population year-round. When a system has multiple inventory records, this chapter retains the most recent record during 2011 through 2022.

These system characteristics are treated as time-invariant contextual measures in the birth-outcome models. System characteristics are drawn from the most recent SDWIS inventory record during 2011 through 2022, while service-area demographic variables represent 2024 CWS-level demographic measures. This approach may not capture changes in service population, source type, ownership, or community demographics over the study period. For this reason, these variables are interpreted as contextual and heterogeneity measures rather than time-varying exposures, while nitrate exposure itself is constructed from time-varying monitoring data over gestation.

This chapter measures system size using population served and categorizes systems into four groups, under 3,300, 3,300 to 9,999, 10,000 to 99,999, and 100,000 or more, with a three-level alternative under 3,300, 3,300 to 99,999, and 100,000 or more. These cut points align with EPA size-based conventions and capture differences in monitoring and treatment capacity (EPA, 2011, 2025f). It classifies primary source type as groundwater versus non-groundwater using SDWIS source codes, consistent with evidence that nitrate is most prevalent and persistent in groundwater-reliant systems (Balazs et al., 2011; Schaidt et al., 2019).

To characterize service-area sociodemographic context, this chapter uses an areal-weighted crosswalk that intersects community water system service boundaries with U.S. Census and American Community Survey geographies. For each system, it constructs areal-weighted estimates of population, race and ethnicity, housing units, and median household income and

then derives system-area percentages non-Hispanic White, non-Hispanic Black, and Hispanic, along with median household income and, where available, poverty rate.

This chapter defines disadvantaged community (DAC) status as system median household income under 80 percent of the statewide median and severely disadvantaged community (SDAC) status as under 60 percent, consistent with California Water Code §79505.5 and guidance that uses CalEnviroScreen and related criteria (Office of Environmental Health Hazard Assessment, 2023; California Environmental Protection Agency, 2022). It computes housing unit density as housing units per square mile of estimated service area to characterize urban–rural differences relevant to nitrate exposure and infrastructure constraints.

2.3.5 Central Valley

The analysis operationalizes Central Valley status using the USGS alluvial boundary for the Sacramento Valley, San Joaquin Basin, and Tulare Lake Basin developed for the Central Valley Hydrologic Model. It uses the Alluvial Boundary polygon dataset from the USGS Professional Paper on groundwater availability in the Central Valley aquifer system as the spatial reference (Faunt, 2009). It then classifies each system using a point-in-polygon assignment based on the service-area centroid, yielding a reproducible one-to-one system classification.

This analysis examines Central Valley heterogeneity in two complementary ways. First, interaction models compare associations for births served by Central Valley versus non-Central Valley systems. Second, this chapter re-estimates the primary models in a Central Valley-only subsample to assess whether statewide patterns are robust within a high-priority nitrate setting while keeping outcome definitions, covariates, and model structure identical (U.S. Geological Survey, 2009, 2021; State Water Board, 2020).

2.3.6 Study Population and Analytic Sample

The analytic sample includes births whose maternal residence at delivery is successfully geocoded and linked to a California community water system service area, yielding a non-missing system identifier. Gestational nitrate exposure is assigned using the time-varying system-quarter nitrate panel with monitoring observations from January 1, 2011, through December 31, 2022 and births occurring within the same period. Births with implausible gestational lengths, defined as under 140 or over 320 days, are excluded, and internal consistency between gestational age, birth date, and derived last menstrual period measures is checked for retained records (Wilhelm et al., 2011; von Ehrenstein et al., 2013).

After applying linkage, monitoring-window, and plausibility criteria, 3,539,019 births have pregnancy-average nitrate exposure defined. Regression analyses use a complete-case sample of 3,403,881 births after excluding observations missing outcomes or covariates, which maintains a consistent estimation sample across models. Table 2.2 summarizes this complete-case regression sample. The complete-case restriction reduces the sample by 135,138 births relative to the broader analytic exposure sample.

Linkage from birth records to community water systems was incomplete (Table 2.1). Restricting the comparison to births from 2012 through 2022, 70.2 percent of births were linked to the CWS-based analytic file and 29.8 percent were not. The 2012 to 2022 window is used because 2011 behaved as an edge year with unusually low linkage and would otherwise understate linkage in the main study period. Across 2012 to 2022, yearly linkage rates were relatively stable, generally ranging from about 67.5 to 72.2 percent. Linked and unlinked births differed modestly on observed characteristics. Compared with unlinked births, linked births were more likely to be Hispanic, less likely to be non-Hispanic White, somewhat more likely to have

public prenatal payer, and slightly less likely to have a college-educated mother. These patterns indicate that linkage was not random. However, they do not suggest that linkage disproportionately excluded the populations most central to the chapter’s environmental justice motivation, since the linked sample was modestly more disadvantaged than the unlinked sample on several observed maternal socioeconomic indicators.

At the same time, the linked sample is overwhelmingly concentrated in larger urban community water systems. Among linked births in the 2012 to 2022 comparison window, 64.59 percent were served by systems serving at least 100,000 people and another 31.46 percent were served by systems serving 10,000 to 99,999 people, while only 1.44 percent and 2.51 percent were served by systems in the two smaller size categories. In addition, 98.47 percent of linked births were linked to systems coded as urban. Although the linked CWS universe still included many small and groundwater-only systems, the chapter’s estimates therefore speak most directly to births linked to larger urban community water systems rather than to the full universe of rural births, private-well users, or other unlinked populations. The remaining non-linkage likely reflects a combination of residence outside mapped community water system service areas, private-well or other non-CWS service, incomplete service-area coverage, and geocoding or spatial-assignment limitations.

Table 2.1: Linked and Unlinked Births in the 2012–2022 Comparison Window

Characteristic	Linked	Unlinked	Difference (%)
Panel A. Linkage rate			
Total births, n	3,559,922	1,513,611	
Share of births, %	70.17	29.83	40.34
Panel B. Composition of linked and unlinked births (%)			
Singleton	97.03	96.64	0.39
Non-Hispanic White	26.56	32.42	-5.86
Non-Hispanic Black	5.42	5.72	-0.30
Hispanic	49.41	42.02	7.39
Other race/ethnicity	18.62	19.83	-1.21

Mother has college degree	30.08	34.23	-4.15
Public prenatal payer (Medi-Cal)	42.09	37.07	5.02
Uninsured prenatal payer	3.53	4.98	-1.45
Private prenatal payer	47.04	49.89	-2.85
Other prenatal payer	4.37	4.70	-0.33
Unknown prenatal payer	2.97	3.36	-0.39
Panel C. System-side profile of the linked sample (%)			
Characteristic	Linked births (birth-weighted)	Distinct linked CWSs (system-weighted)	
Size <3,300	1.44	68.27	
Size 3,300–9,999	2.51	9.49	
Size 10,000–99,999	31.46	16.84	
Size ≥100,000	64.59	4.33	
Groundwater-only	18.86	71.12	
Central Valley	20.84	28.11	
DAC	51.13	50.53	
SDAC	13.18	28.88	
Urban	98.47	52.85	

Note. Linked versus unlinked status is defined by whether a birth record matched the CWS-linked analytic file in the 2012–2022 comparison window. Percentages in Panel B are column percentages within linked and unlinked births. Panel C summarizes the system-side composition of the linked sample only. The Linked births column is birth-weighted and shows the share of linked births served by each type of CWS. The Distinct linked CWSs column is system-weighted and shows the share of unique linked systems in each category. Because unlinked births do not have assigned CWS identifiers, the current data cannot directly classify the unlinked group by water-system type. Panel C is therefore intended to clarify that the linked sample is already heavily concentrated in large, urban systems and that the remaining non-linkage should not be interpreted as a simple rural residual.

Following construction of pregnancy-average and trimester-specific nitrate exposure measures, the analysis estimates adjusted associations between nitrate exposure and the two primary outcomes, preterm birth and low birthweight, using regression models with clustered inference at the community water system level. The empirical strategy evaluates nitrate exposure using both continuous measures and policy-salient threshold specifications and examines heterogeneity by maternal race and ethnicity, disadvantaged community status, and Central Valley context. Estimates are interpreted as adjusted associations rather than causal effects, and

robustness checks assess sensitivity to alternative exposure aggregations, exclusion of extreme concentration values, and Central Valley subsample specifications.

2.4 Methods

2.4.1 Exposure Measures

The primary exposure is pregnancy-average nitrate concentration in the assigned community water system, measured in milligrams per liter as nitrogen. Pregnancy-average nitrate is computed by averaging system-month mean nitrate concentrations across gestational months with non-missing monitoring data.

The study parameterized nitrate exposure in three ways to reduce reliance on a single functional-form assumption. First, continuous pregnancy-average nitrate is modeled per 1 mg/L as nitrogen. Second, a five-category pregnancy-average nitrate variable is constructed to capture potential nonlinearity and to distinguish sub-threshold exposure ranges. Categories are under 1, 1 to under 2.5, 2.5 to under 5, 5 to under 10, and 10 or more mg/L as nitrogen, with under 1 mg/L as nitrogen as the reference. The 10 or more category corresponds to the federal and California maximum contaminant level for nitrate and provides a direct policy-salient contrast (EPA, 2025e). Cutpoints at 2.5 and 5 mg/L as nitrogen delineate exposure ranges highlighted in recent epidemiologic evidence while maintaining adequate sample size within each category given the distribution of nitrate in California systems (Coffman et al., 2022; Sherris et al., 2021; Lin et al., 2023; Royal et al., 2024; Clemmensen et al., 2023). Third, a binary indicator contrasts pregnancy-average nitrate at or above 10 mg/L as nitrogen versus below 10 mg/L as nitrogen to provide a single interpretable threshold-based estimate aligned with the maximum contaminant level.

The main categorical specification uses policy-oriented cutpoints rather than distribution-based bins to preserve substantively meaningful contrasts tied to public-health and regulatory interpretation, especially the near-threshold 5 to <10 mg/L range and the ≥ 10 mg/L as N threshold. Because these cutpoints are not distribution-based, additional sensitivity analyses re-estimate the categorical models using quintiles and quartiles of pregnancy-average nitrate. Additionally, the study further constructed secondary exposure measures including trimester-average nitrate concentrations for the first, second, and third trimesters. Trimester averages are computed by averaging system-month means across gestational months corresponding to weeks 1 through 13, 14 through 27, and 28 or more. These measures support sensitivity analyses of temporal specificity.

2.4.2 Birth outcomes

The primary outcomes are preterm birth and low birthweight. Preterm birth (PTB) is defined as delivery before 37 completed weeks of gestation (<259 days), in accordance with World Health Organization (WHO) and widely used clinical guidelines (National Center for Health Statistics, 2024; WHO, 2023; Quinn et al., 2016). Using the cleaned gestational age variable in days, the study created a binary indicator for PTB and, in secondary analyses, a very preterm birth outcome defined as delivery before 32 completed weeks (<224 days). Low birthweight (LBW) is defined as birthweight under 2,500 grams regardless of gestational age (WHO, n.d.). These outcome definitions follow standard perinatal epidemiology conventions to support comparability across studies (Lin et al., 2023; Royal et al., 2024; Sherris et al., 2021). Gestational age and birthweight values are drawn from the vital records and subjected to basic plausibility checks, while sensitivity analyses excluded extreme or implausible values.

2.4.3 Covariates

This chapter selects covariates based on perinatal epidemiology and a directed acyclic graph, Figure 2.1, that represents relationships among maternal sociodemographic factors, residential and water system context, nitrate exposure, and birth outcomes. The adjustment strategy aims to address confounding associated with maternal residential location, and therefore nitrate exposure, and baseline risk while avoiding automatic adjustment for variables that may reflect pregnancy complications or intermediates on the exposure-to-outcome pathway (Goldenberg et al., 2008; Manuck, 2017; WHO, 2023).

The primary adjustment set includes maternal age, maternal race and ethnicity, and socioeconomic measures. Models represent maternal age using four categories, under 20, 20 to 24, 25 to 34, and 35 or more years, reflecting well-established non-linear associations between maternal age and adverse outcomes (Goldenberg et al., 2008; Holzman et al., 2009). Models code maternal race and ethnicity as non-Hispanic White, non-Hispanic Black, Hispanic, and Other, consistent with literature documenting persistent disparities in preterm birth and related outcomes and supporting interpretation of environmental justice patterns (Manuck, 2017; James-Todd et al., 2014). Models capture socioeconomic position using maternal education and expected payer at delivery, which are standard proxies in birth cohort analyses (Goldenberg et al., 2008). Models also adjust for prenatal care payer, categorized as no prenatal care, public payer, private insurance, self-pay, and other or unknown payer. The analyses also adjust for reproductive history and parity to account for baseline differences in risk. This chapter includes an indicator for first-time mothers and an indicator for prior preterm birth, given their strong predictive value for subsequent preterm delivery and their plausibly exogenous status with respect to nitrate exposure in the index pregnancy (Goldenberg et al., 2008; WHO, 2023).

The extended adjustment set includes behavioral and clinical risk factors measured before or early in pregnancy. Models code maternal smoking during pregnancy as any reported smoking, given its strong association with fetal growth restriction and preterm birth and its correlation with socioeconomic and residential patterns that may also relate to drinking-water exposure (Goldenberg et al., 2008). Models include pre-pregnancy body mass index categories constructed from recorded values after excluding extreme observations, reflecting evidence that both underweight and overweight are associated with altered risks of preterm birth and abnormal fetal growth and may confound nitrate associations (Liu et al., 2019; Goldenberg et al., 2008). Some pre-pregnancy chronic condition indicators are not consistently available or reliable in the vital records fields used for this analysis period and are therefore not included in the primary models. This chapter addresses these limitations through extensive adjustment for sociodemographic and reproductive risk factors and by including birth year-quarter fixed effects that absorb statewide time-varying shocks.

It's worth noting that System-area racial and ethnic composition measures are used to describe the distribution of nitrate exposure and to support environmental justice interpretation. The main models use maternal race and ethnicity for individual-level heterogeneity tests and do not include system-level racial composition measures as default covariates, given their role as structural determinants of exposure and their close intertwining with individual-level race, ethnicity, and socioeconomic position. This choice avoids mechanically adjusting away disparities that are central to the chapter's environmental justice aims.

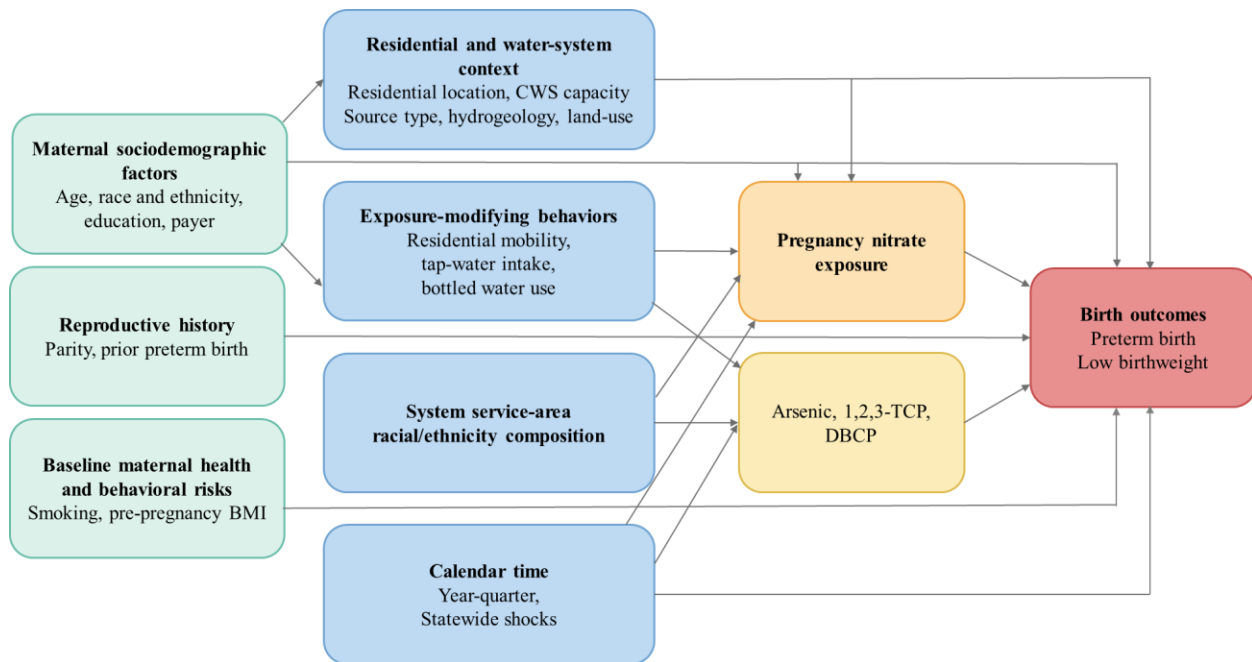


Figure 2.1: Directed Acyclic Graph for Pregnancy Nitrate Exposure and Birth Outcomes

Note. The DAG treats arsenic, 1,2,3-TCP, and DBCP as co-occurring groundwater contaminants that share upstream determinants with nitrate, including source type, hydrogeologic setting, land use, and water-system context. They are therefore shown as parallel co-exposures rather than mediators of nitrate. In the empirical analysis, nitrate is the primary exposure of interest and co-pollutant models are treated as sensitivity analyses because co-contaminant measurement is more limited.

2.4.4 Model Specification

2.4.4.1 Main specification

The study estimated adjusted associations between nitrate exposure in community water systems and adverse birth outcomes using multivariable logistic regression at the individual birth level. The outcome is a binary indicator for preterm birth (<37 completed weeks) or low birthweight (<2,500 g) for birth i served by community water system p and born in year-quarter t . Models include birth year-quarter fixed effects that are constructed from each birth's date and included to absorb statewide secular trends, seasonality, and time-varying shocks common across births. Standard errors are clustered at the CWS level to account for correlation

among births served by the same system. Results are reported as odds ratios (ORs) with 95% confidence intervals.

The main specification with continuous pregnancy-average nitrate is:

$$\text{logit}(\Pr(Y_{ipt} = 1 \mid N_{ipt}, X_{ipt}, \delta_t)) = \alpha + \beta N_{ipt} + X'_{ipt}\gamma + \delta_t \quad (2.1)$$

Here, Y_{ipt} is the binary outcome for birth i served by system p and born in year-quarter t ; N_{ipt} is pregnancy-average nitrate exposure; X_{ipt} is the covariate vector described in Section 2.4.3; and δ_t denotes birth year-quarter fixed effects. All models use standard errors clustered at the community water system level p .

To align directly with the regulatory threshold, the binary exposure model is:

$$N_{ipt}^{10} = I(N_{ipt} \geq 10)$$

$$\text{logit}(\Pr(Y_{ipt} = 1 \mid N_{ipt}^{10}, X_{ipt}, \delta_t)) = \alpha + \beta_{10} N_{ipt}^{10} + X'_{ipt}\gamma + \delta_t \quad (2.2)$$

The continuous pregnancy-average specification is the primary model, while the binary threshold and five-category specifications assess whether associations align with the regulatory threshold or vary across lower exposure ranges, including the half-MCL to MCL interval, and quartile- and quintile-based sensitivity checks assess whether the categorical pattern depends on the chosen cut points. For the five-category specification, let $D_{k,ipt}$ indicate nitrate category k and omit the reference category $< \text{mg/L}$ as N:

$$\text{logit}(\Pr(Y_{ipt} = 1 \mid \{D_{k,ipt}\}, X_{ipt}, \delta_t)) = \alpha + \sum_{k \in K} \beta_k D_{k,ipt} + X'_{ipt}\gamma + \delta_t \quad (2.3)$$

Pregnancy-average nitrate is the primary exposure specification because it captures the full prenatal exposure window and is less sensitive to sparse monitoring within shorter

gestational intervals. Trimester-specific models are retained as exploratory time-window analyses because exposure timing is relevant for assessing biological plausibility in prenatal environmental health studies. To assess this temporal specificity, the main specification is re-estimated using trimester-average nitrate exposures. Let $N_{ipt}^{(m)}$ denote trimester-average nitrate for trimester $m \in \{1,2,3\}$. Each trimester model is estimated separately:

$$\text{logit}\left(\Pr\left(Y_{ipt} = 1 \mid N_{ipt}^{(m)}, X_{ipt}, \delta_t\right)\right) = \alpha_m + \beta_m N_{ipt}^{(m)} + X'_{ipt} \gamma_m + \delta_t, m \in \{1,2,3\} \quad (2.4)$$

The primary analyses restrict to singleton births with pregnancy-average nitrate exposure defined. Regression models use a complete-case estimation sample, excluding births with missing outcomes, exposure, fixed effects identifiers, cluster identifiers, or covariates. Descriptive statistics are reported for the complete-case regression sample used in the main models, and model-specific sample sizes are reported with results when specifications differ.

2.4.4.2 Race and ethnicity heterogeneity

To examine whether the association between pregnancy-average nitrate exposure and adverse birth outcomes differs by maternal race and ethnicity, the study estimated interaction models in the singleton complete-case sample. Models include the same covariates, birth year-quarter fixed effects, and clustered standard errors as the main specification. The reference group is non-Hispanic White and Other, and interactions test whether nitrate associations differ for non-Hispanic Black and Hispanic mothers relative to the reference group.

Let $R_{i,g}$ be indicators for race and ethnicity group g , with the reference group omitted. The continuous-exposure interaction model is:

$$\text{logit}(\Pr(Y_{ipt} = 1 \mid N_{ipt}, X_{ipt}, \delta_t)) = \alpha + \beta N_{ipt} + \sum_{g \in G} \theta_g R_{ig} + \sum_{g \in G} \phi_g (N_{ipt} \times R_{ig}) + X'_{ipt} \gamma + \delta_t \quad (2.5)$$

For group g , the group-specific nitrate odds ratio is $\exp(\beta + \phi_g)$, and the ratio of odds ratios comparing group g to the reference group is $\exp(\phi_g)$. Nitrate exposure is also parameterized as the binary threshold indicator and the five-category specification, following Equations (2.2) and (2.3). When standard logistic regression exhibits convergence issues in low birthweight models under some specifications, models are fit using a generalized linear model with binomial family and logit link while retaining the same linear predictor and clustered standard errors.

The model uses maternal race and ethnicity to assess heterogeneity in adjusted associations, not to imply nitrate-specific biological susceptibility. Differences across groups are interpreted as reflecting socially patterned conditions that may shape baseline risk, co-exposures, prenatal care access, cumulative stressors, and the ability to reduce exposure when water quality concerns arise. Given that this chapter does not observe public-notification receipt, bottled-water use, filtration, or other household avoidance behaviors, these pathways are treated as plausible interpretations rather than directly tested mechanisms.

2.4.4.3 Community heterogeneity

To examine whether nitrate-birth outcome associations vary across community and system contexts, place-based heterogeneity is evaluated using interaction models that replace maternal race and ethnicity modifiers with system-level indicators relevant to environmental justice. Modifiers include disadvantaged community status, severely disadvantaged community status, and Central Valley location. Models include the same covariate set, birth year-quarter fixed effects, and clustered standard errors as the main specification.

Let M_p denote a system-level modifier. The continuous-exposure interaction model is:

$$\text{logit}(\Pr(Y_{ipt} = 1 \mid N_{ipt}, X_{ipt}, \delta_t)) = \alpha + \beta N_{ipt} + \theta M_p + \phi(N_{ipt} \times M_p) + X'_{ipt}\gamma + \delta_t \quad (2.6)$$

The nitrate odds ratio among systems with $M_p = 0$ is $\exp(\beta)$, the nitrate odds ratio among systems with $M_p = 1$ is $\exp(\beta + \phi)$, and the ratio of odds ratios comparing $M_p = 1$ to $M_p = 0$ is $\exp(\phi)$. As in the race and ethnicity analysis, nitrate exposure is parameterized as continuous, binary threshold, and five-category specifications. Low birthweight models use the generalized linear model logit specification when convergence issues arise, and all other features of the specification remain unchanged.

These community-level variables are interpreted as place-based modifiers rather than mediators of measured nitrate exposure. Because the models condition on observed pregnancy-average nitrate, the interaction terms assess whether nitrate-birth outcome associations differ across community contexts where co-exposures, public information, household avoidance options, and water-system capacity may vary.

2.5 Results

2.5.1 Summary Statistics

Table 2.2 summarizes the complete-case singleton regression sample of 3,403,881 births with defined pregnancy-average nitrate exposure and nonmissing covariates from 2011 to 2022. In this sample, 7.9% of births were preterm and 5.4% were low birthweight. Mean pregnancy-average nitrate concentration was 3.32 mg/L as N, and 0.5% of births were exposed to nitrate levels at or above 10 mg/L as N. The largest share of births, 46.5%, fell in the 2.5 to less than 5 mg/L as N category.

The sample was racially and ethnically diverse and was concentrated in large community water systems. Hispanic mothers accounted for 49.7% of births, 45.0% were non-Hispanic White or another race/ethnicity, and 5.6% were non-Hispanic Black. Most births were linked to systems serving 10,000 or more people, including 64.7% linked to systems serving at least 100,000 people. In addition, 20.6% of births were linked to Central Valley systems, 13.1% to severely disadvantaged community systems, and 51.1% to disadvantaged community systems.

Table 2.2: Descriptive Statistics for the Complete-Case Regression Sample

Characteristic	Category or measure	<i>n</i>	%	Mean (SD)	Median
Panel A. Outcomes and nitrate exposure					
Preterm birth	<37 weeks	267,959	7.9		
Low birthweight	<2,500 g	182,249	5.4		
Pregnancy-average nitrate concentration	mg/L as N	3,403,881		3.32 (1.96)	3.14
Pregnancy-average nitrate category	<1 mg/L as N	397,228	11.7		
Pregnancy-average nitrate category	1–<2.5 mg/L as N	819,497	24.1		
Pregnancy-average nitrate category	2.5–<5 mg/L as N	1,584,268	46.5		
Pregnancy-average nitrate category	5–<10 mg/L as N	586,989	17.2		
Pregnancy-average nitrate category	≥10 mg/L as N	15,899	0.5		
Panel B. Maternal and infant characteristics					
Maternal age	<20	161,911	4.8		
Maternal age	20–24	589,160	17.3		
Maternal age	25–34	1,909,937	56.1		
Maternal age	≥35	742,873	21.8		

Characteristic	Category or measure	<i>n</i>	%	Mean (SD)	Median
College degree or higher	Yes	1,019,431	29.9		
Maternal race/ethnicity	Non-Hispanic White or other race/ethnicity	1,530,540	45.0		
Maternal race/ethnicity	Non-Hispanic Black	191,476	5.6		
Maternal race/ethnicity	Hispanic	1,692,710	49.7		
Smoking during pregnancy	Yes	83,802	2.5		
First live birth	Yes	1,346,080	39.5		
Pre-pregnancy BMI	Underweight (<18.5)	140,868	4.1		
Pre-pregnancy BMI	Normal weight (18.5–24.9)	1,500,635	44.1		
Pre-pregnancy BMI	Overweight (25.0–29.9)	909,427	26.7		
Pre-pregnancy BMI	Obese (≥30)	852,951	25.1		
Previous preterm birth history	Yes	46,220	1.4		
Female infant	Yes	1,660,883	48.8		
Prenatal care payer	No prenatal care	17,744	0.5		
Prenatal care payer	Public insurance, including Medi-Cal	1,585,689	46.6		
Prenatal care payer	Private insurance	1,659,586	48.8		
Prenatal care payer	Self-pay	99,737	2.9		
Prenatal care payer	Other or unknown	41,125	1.2		
Panel C. System and community characteristics					
Central Valley system	Yes	699,919	20.6		

Characteristic	Category or measure	<i>n</i>	%	Mean (SD)	Median
Severely disadvantaged community	Yes	445,212	13.1		
Disadvantaged community	Yes	1,739,268	51.1		
Groundwater-only system	Yes	639,472	18.8		
Community water system size	<3,300 (small or very small)	49,063	1.4		
Community water system size	3,300–9,999 (medium)	84,984	2.5		
Community water system size	10,000–99,999 (large)	1,067,127	31.4		
Community water system size	≥100,000 (very large)	2,202,707	64.7		

Note. The table summarizes the complete-case singleton birth sample used in the main Chapter 2 regression models. For categorical and binary characteristics, *n* reports the number of births in the listed category and percentages are calculated within the complete-case sample. For continuous pregnancy-average nitrate concentration, the table reports the mean, standard deviation, and median. All births in this sample have defined pregnancy-average nitrate exposure, nonmissing primary outcomes and covariates, birth year-quarter fixed effects, and community water system cluster identifiers. PWS refers to public water system; DAC refers to disadvantaged community status; SDAC refers to severely disadvantaged community status.

2.5.2 Pregnancy-average Nitrate and Birth Outcomes

2.5.2.1 Preterm birth

Higher pregnancy-average nitrate exposure was consistently associated with higher adjusted odds of preterm birth across all three exposure parameterizations (Table 2.3; Figure 2.2). In the continuous model, each 1 mg/L increase in nitrate was associated with 1.9% higher odds of preterm birth (OR = 1.019, 95% CI [1.012, 1.026], $p < 0.001$). In the categorical model, odds ratios rose across nitrate categories relative to <1 mg/L as N, reaching 1.182 for ≥10 mg/L as N (95% CI [1.104, 1.265], $p = 0.017$). The threshold model also indicated higher odds of preterm birth for births exposed to nitrate concentrations at or above 10 mg/L as N (OR = 1.091,

95% CI [1.021, 1.167], $p = 0.010$). Figure 2.2 displays categorical estimates, and full model results are reported in Appendix Table 2.A3.

Table 2.3: Adjusted Odds Ratios for Preterm Birth by Pregnancy-Average Nitrate Exposure Specification

Exposure specification	Nitrate Contrast	Odds ratio	95% CI	<i>p</i> -value
Continuous	Per 1 mg/L as N	1.019	[1.012, 1.026]	<0.001***
Categorical	1–<2.5 mg/L as N vs <1 mg/L as N	1.050	[1.007, 1.094]	0.021**
Categorical	2.5–<5 mg/L as N vs <1 mg/L as N	1.104	[1.065, 1.143]	<0.001***
Categorical	5–<10 mg/L as N vs <1 mg/L as N	1.130	[1.084, 1.177]	<0.001***
Categorical	≥10 mg/L as N vs <1 mg/L as N	1.182	[1.104, 1.265]	<0.001***
Threshold	≥10 mg/L as N vs <10 mg/L as N	1.091	[1.021, 1.167]	0.010**

Note. The table reports selected nitrate exposure estimates from separate adjusted logistic regression models for preterm birth. The continuous model parameterizes pregnancy-average nitrate as a 1 mg/L increase in nitrate as nitrogen. The categorical model uses <1 mg/L as N as the reference category. The threshold model compares births with pregnancy-average nitrate ≥10 mg/L as N to births below 10 mg/L as N. All models adjust for maternal sociodemographic and clinical covariates, prenatal payer category, and birth year-quarter fixed effects, with standard errors clustered at the public water system level. OR refers to odds ratio, and CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

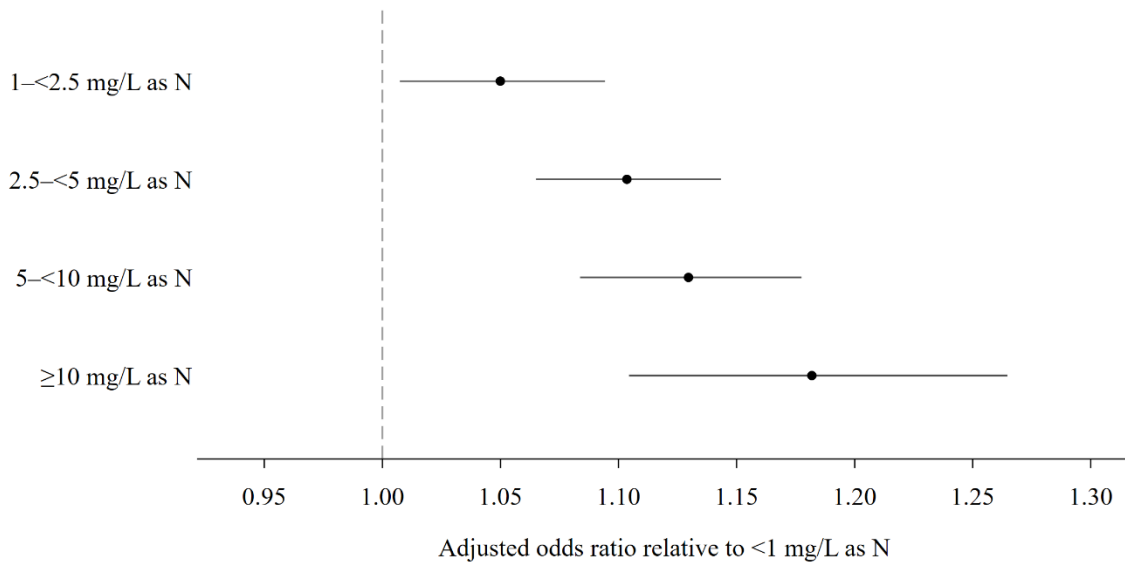


Figure 2.2: Adjusted Odds Ratios for Preterm Birth by Pregnancy-Average Nitrate Category

Note. Points show adjusted odds ratios and 95% confidence intervals from the categorical nitrate model for preterm birth. The reference group is <1 mg/L as N. The model adjusts for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

2.5.2.2 Low birthweight

Associations between pregnancy-average nitrate exposure and low birthweight were smaller and less consistent across exposure parameterizations than those for preterm birth (Table 2.4; Figure 2.3). In the continuous model, each 1 mg/L increase in nitrate was associated with 1.4% higher odds of low birthweight (OR = 1.014, 95% CI [1.006, 1.021], $p < 0.001$). In the categorical model, the lowest non-reference category was close to null (1-<2.5 mg/L as N: OR = 1.039, 95% CI [0.978, 1.104], $p = 0.212$), while the mid-range categories were modestly elevated, with odds ratios of 1.076 for 2.5-<5 mg/L as N (95% CI [1.012, 1.145], $p = 0.020$) and 1.093 for 5-<10 mg/L as N (95% CI [1.026, 1.165], $p = 0.006$). The highest category was positive but imprecise (≥ 10 mg/L as N: OR = 1.062, 95% CI [0.982, 1.148], $p = 0.132$). The

threshold specification likewise showed no evidence of an association for nitrate concentrations at or above 10 mg/L as N relative to lower concentrations (OR = 1.001, 95% CI [0.946, 1.058], $p = 0.985$). Figure 2.3 shows the categorical estimates and their confidence intervals. Full model results are reported in Appendix Tables 2.A3. In summary, the nitrate associations were more consistent for preterm birth than for low birthweight.

Table 2.4: Adjusted Odds Ratios for Low Birthweight by Pregnancy-Average Nitrate Exposure Specification

Exposure specification	Nitrate Contrast	Odds ratio	95% CI	<i>p</i> -value
Continuous	Per 1 mg/L as N	1.014	[1.006, 1.021]	<0.001***
Categorical	1–<2.5 mg/L as N vs <1 mg/L as N	1.039	[0.978, 1.104]	0.212
Categorical	2.5–<5 mg/L as N vs <1 mg/L as N	1.076	[1.012, 1.145]	0.020**
Categorical	5–<10 mg/L as N vs <1 mg/L as N	1.093	[1.026, 1.165]	0.006***
Categorical	≥10 mg/L as N vs <1 mg/L as N	1.062	[0.982, 1.148]	0.132
Threshold	≥10 mg/L as N vs <10 mg/L as N	1.001	[0.946, 1.058]	0.985

Note. The table reports selected nitrate exposure estimates from separate adjusted models for low birthweight. The continuous model parameterizes pregnancy-average nitrate as a 1 mg/L increase in nitrate as nitrogen. The categorical model uses <1 mg/L as N as the reference category. The threshold model compares births with pregnancy-average nitrate ≥10 mg/L as N to births below 10 mg/L as N. All models adjust for maternal sociodemographic and clinical covariates, prenatal payer category, community water system urban-rural context, and birth year-quarter fixed effects, with standard errors clustered at the public water system level. Low birthweight models were estimated using a generalized linear model with binomial family and logit link to improve numerical stability under high-dimensional fixed effects. OR refers to odds ratio, and CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

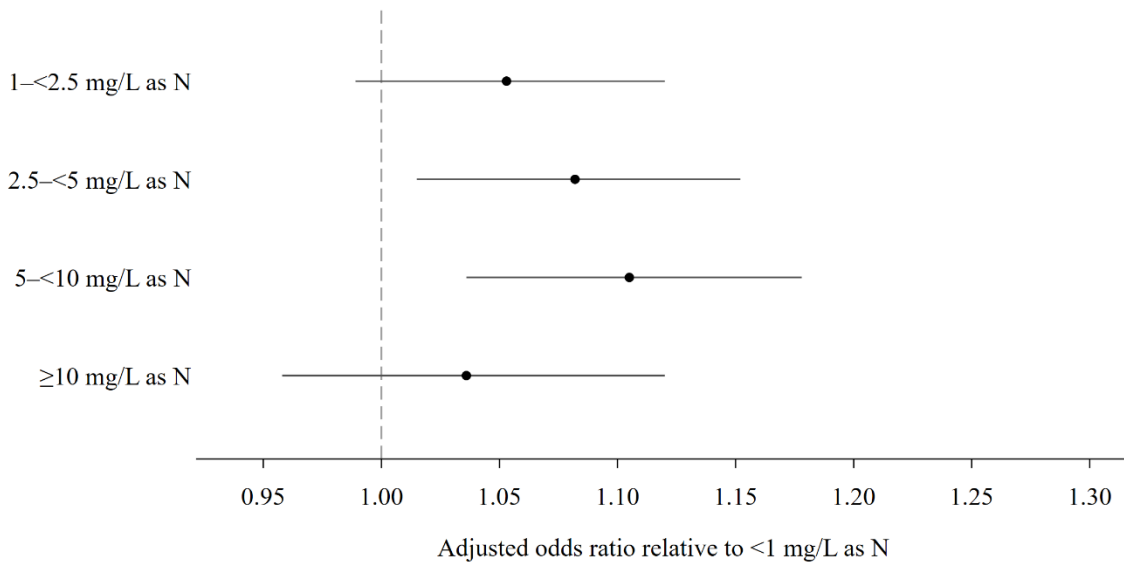


Figure 2.3: Adjusted Odds Ratios for Low Birthweight by Pregnancy-Average Nitrate Category

Note. Points show adjusted odds ratios and 95% confidence intervals from the categorical nitrate model for low birthweight. The reference group is <1 mg/L as N. The model adjusts for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

2.5.3 Trimester-specific Nitrate Exposure

Trimester-specific analyses serve as time-window tests rather than simple replications of the pooled pregnancy-average models. In prenatal exposure studies, the timing of exposure helps assess biological plausibility by showing whether associations are concentrated in particular gestational windows rather than distributed uniformly across pregnancy. This chapter estimates first-, second-, and third-trimester average nitrate exposure using continuous, threshold, and five-category parameterizations in the same regression sample and adjustment framework as the main models. Tables 2.5 and 2.6 summarize the trimester-specific estimates for preterm birth and low birthweight, and Figures 2.4 and 2.5 display the continuous trimester-specific associations.

For preterm birth, the continuous and binary threshold specifications were close to null in all three trimesters (Table 2.5; Figure 2.4). In the categorical models, however, associations varied across trimesters. In the first trimester, the 2.5–<5 mg/L as N category was associated with higher odds of preterm birth relative to <1 mg/L as N (OR = 1.123, 95% CI [1.011, 1.246], $p = 0.030$). In the second trimester, the categorical pattern was inconsistent, including an estimate below 1.0 for the ≥ 10 mg/L as N category (OR = 0.861, 95% CI [0.746, 0.994], $p = 0.041$). In the third trimester, odds ratios were above 1.0 across all categories and were largest for ≥ 10 mg/L as N (OR = 1.515, 95% CI [1.333, 1.723], $p < 0.001$), indicating a stronger positive association late in pregnancy. Given the number of contrasts across trimesters and exposure parameterizations, the non-monotonic second-trimester estimates are interpreted cautiously.

Table 2.5: Adjusted Odds Ratios for Preterm Birth by Trimester-Specific Nitrate Exposure Specification

Trimester	Exposure specification	Nitrate Contrast	Odds ratio	95% CI	<i>p</i> -value
First trimester	Continuous per 1 mg/L as N	per 1 mg/L as N	1.011	[0.997, 1.025]	0.111
	Binary ≥ 10 mg/L as N	≥ 10 vs <10 mg/L as N	0.933	[0.799, 1.089]	0.377
	Categorical nitrate	1–<2.5 mg/L as N vs <1 mg/L as N	1.029	[0.920, 1.151]	0.620
	Categorical nitrate	2.5–<5 mg/L as N vs <1 mg/L as N	1.123	[1.011, 1.246]	0.030**

	Categorical nitrate	5–<10 mg/L as N vs <1 mg/L as N	1.084	[0.966, 1.217]	0.168
	Categorical nitrate	≥10 mg/L as N vs <1 mg/L as N	0.950	[0.854, 1.057]	0.346
Second trimester	Continuous per 1 mg/L as N	per 1 mg/L as N	1.004	[0.989, 1.018]	0.625
	Binary ≥10 mg/L as N	≥10 vs <10 mg/L as N	0.971	[0.794, 1.186]	0.771
	Categorical nitrate	1–<2.5 mg/L as N vs <1 mg/L as N	0.925	[0.824, 1.040]	0.192
	Categorical nitrate	2.5–<5 mg/L as N vs <1 mg/L as N	0.996	[0.886, 1.121]	0.953
	Categorical nitrate	5–<10 mg/L as N vs <1 mg/L as N	0.993	[0.884, 1.116]	0.908
	Categorical nitrate	≥10 mg/L as N vs <1 mg/L as N	0.861	[0.746, 0.994]	0.041**
Third trimester	Continuous per 1 mg/L as N	per 1 mg/L as N	1.008	[0.995, 1.021]	0.222
	Binary ≥10 mg/L as N	≥10 vs <10 mg/L as N	1.030	[0.844, 1.257]	0.769
	Categorical nitrate	1–<2.5 mg/L as N vs <1 mg/L as N	1.115	[1.038, 1.199]	0.003***
	Categorical nitrate	2.5–<5 mg/L as N vs <1 mg/L as N	1.169	[1.061, 1.289]	0.002***

	Categorical nitrate	5–<10 mg/L as N vs <1 mg/L as N	1.142	[1.063, 1.227]	<0.001***
	Categorical nitrate	≥10 mg/L as N vs <1 mg/L as N	1.515	[1.333, 1.723]	<0.001***

Note. The table reports selected nitrate exposure estimates from separate adjusted trimester-specific models for preterm birth. Within each trimester, nitrate exposure is parameterized as a continuous measure per 1 mg/L as nitrogen, a threshold indicator comparing ≥10 mg/L as N with <10 mg/L as N, and categorical exposure groups using <1 mg/L as N as the reference category. All models adjust for maternal sociodemographic and clinical covariates, prenatal payer category, community water system urban-rural context, and birth year-quarter fixed effects, with standard errors clustered at the public water system level. Trimester-specific model samples may vary slightly because exposure must be observed for the relevant trimester. OR refers to odds ratio, and CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

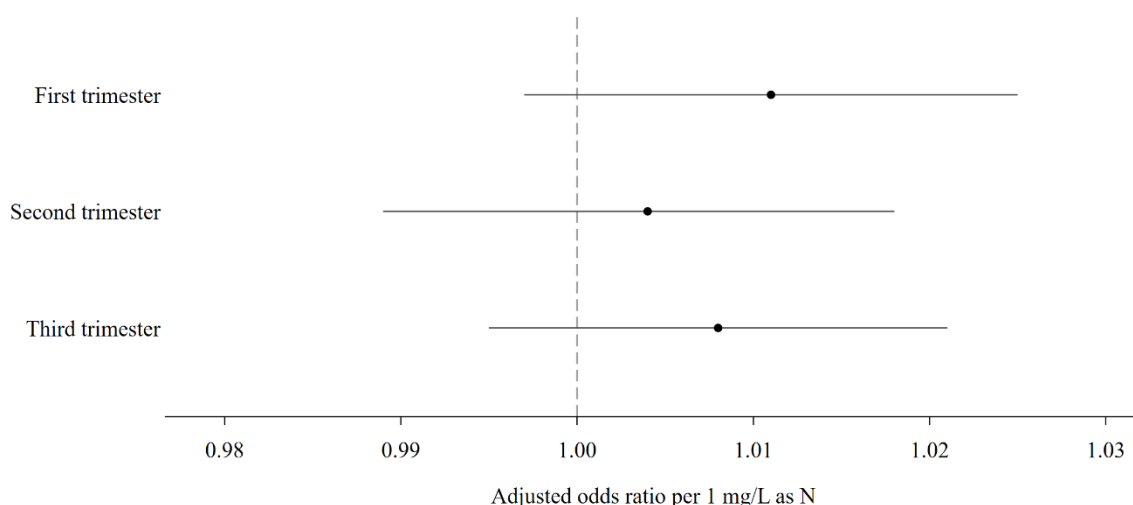


Figure 2.4: Adjusted Odds Ratios for Preterm Birth by Trimester-Specific Nitrate Exposure

Note. Points show adjusted odds ratios and 95% confidence intervals from continuous trimester-specific nitrate models for preterm birth. Exposure is parameterized as a 1 mg/L increase in nitrate as nitrogen. Models adjust for the full covariate set, birth year-quarter fixed effects, and standard errors clustered at the public water system level.

For low birthweight, the continuous and binary threshold specifications were likewise close to null across trimesters (Table 2.6; Figure 2.5). In the categorical models, the strongest

positive associations again appeared in the third trimester, where the ≥ 10 mg/L as N category was associated with higher odds of low birthweight relative to < 1 mg/L as N (OR = 1.489, 95% CI [1.302, 1.704], $p < 0.001$), and the $1 - < 2.5$ mg/L as N category was also modestly elevated (OR = 1.077, 95% CI [1.013, 1.146], $p = 0.018$). As in the preterm models, the second-trimester ≥ 10 mg/L as N estimate was below 1.0 (OR = 0.856, 95% CI [0.763, 0.960], $p = 0.008$). Overall, the trimester-specific results point most clearly to positive associations in the categorical third-trimester models, whereas the continuous and threshold specifications show little consistent evidence of trimester-specific effects. These estimates are interpreted cautiously, however, because nitrate monitoring frequency varies across systems and many system-years have limited within-year temporal support, especially among smaller systems.

Table 2.6: Adjusted Odds Ratios for Low Birthweight by Trimester-Specific Nitrate Exposure Specification

Trimester	Exposure specification	Nitrate Contrast	Odds ratio	95% CI	<i>p</i> -value
First trimester	Continuous per 1 mg/L as N	per 1 mg/L as N	1.002	[0.991, 1.014]	0.676
	Binary ≥ 10 mg/L as N	≥ 10 vs < 10 mg/L as N	0.927	[0.807, 1.066]	0.290
	Categorical nitrate	$1 - < 2.5$ mg/L as N vs < 1 mg/L as N	1.005	[0.891, 1.133]	0.940
	Categorical nitrate	$2.5 - < 5$ mg/L as N vs < 1 mg/L as N	1.073	[0.954, 1.206]	0.239
	Categorical nitrate	$5 - < 10$ mg/L as N vs < 1 mg/L as N	1.012	[0.904, 1.133]	0.833

	Categorical nitrate	≥10 mg/L as N vs <1 mg/L as N	0.943	[0.842, 1.056]	0.311
Second trimester	Continuous per 1 mg/L as N	per 1 mg/L as N	0.992	[0.980, 1.005]	0.217
	Binary ≥10 mg/L as N	≥10 vs <10 mg/L as N	0.997	[0.835, 1.192]	0.977
	Categorical nitrate	1–<2.5 mg/L as N vs <1 mg/L as N	0.913	[0.823, 1.014]	0.088*
	Categorical nitrate	2.5–<5 mg/L as N vs <1 mg/L as N	0.962	[0.864, 1.072]	0.484
	Categorical nitrate	5–<10 mg/L as N vs <1 mg/L as N	0.921	[0.835, 1.017]	0.102
	Categorical nitrate	≥10 mg/L as N vs <1 mg/L as N	0.856	[0.763, 0.960]	0.008***
Third trimester	Continuous per 1 mg/L as N	per 1 mg/L as N	0.996	[0.985, 1.007]	0.463
	Binary ≥10 mg/L as N	≥10 vs <10 mg/L as N	1.038	[0.861, 1.251]	0.698
	Categorical nitrate	1–<2.5 mg/L as N	1.077	[1.013, 1.146]	0.018**
	Categorical nitrate	2.5–<5 mg/L as N	1.103	[0.998, 1.218]	0.054*
	Categorical nitrate	5–<10 mg/L as N	1.038	[0.971, 1.109]	0.271
	Categorical nitrate	≥10 mg/L as N	1.489	[1.302, 1.704]	<0.001***

Note. The table reports selected nitrate exposure estimates from separate adjusted trimester-specific models for low birthweight. Within each trimester, nitrate exposure is parameterized as a continuous measure per 1 mg/L as nitrogen, a threshold indicator comparing ≥ 10 mg/L as N with < 10 mg/L as N, and categorical exposure groups using < 1 mg/L as N as the reference category. All models adjust for maternal sociodemographic and clinical covariates, prenatal payer category, community water system urban-rural context, and birth year-quarter fixed effects, with standard errors clustered at the public water system level. Trimester-specific model samples may vary slightly because exposure must be observed for the relevant trimester. Low birthweight models were estimated using a generalized linear model with binomial family and logit link to improve numerical stability under high-dimensional fixed effects. OR refers to odds ratio, and CI refers to confidence interval. Statistical significance: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

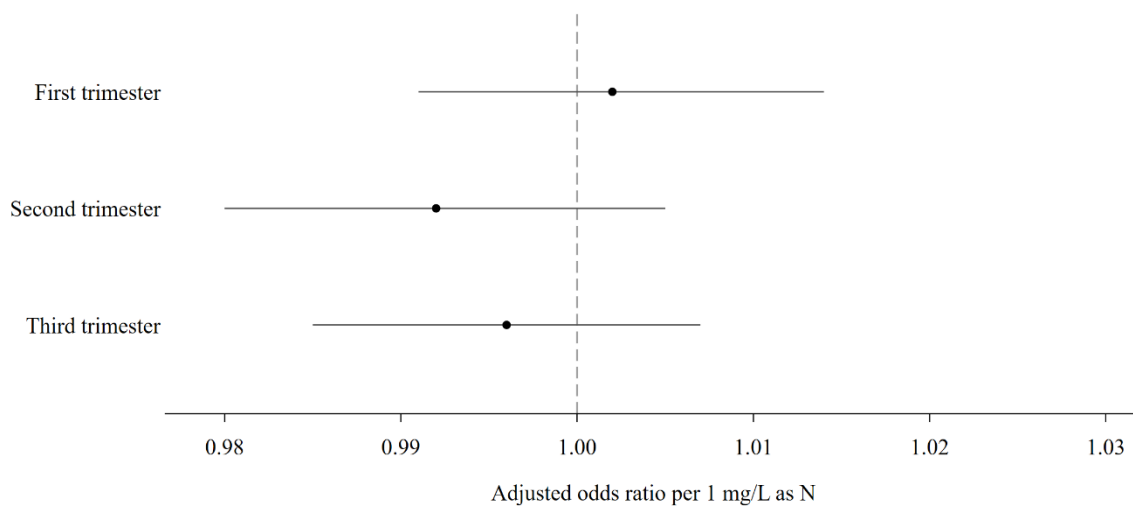


Figure 2.5: Adjusted Odds Ratios for Low Birthweight by Trimester-Specific Nitrate Exposure

Note. Points show adjusted odds ratios and 95% confidence intervals from continuous trimester-specific nitrate models for low birthweight. Exposure is parameterized as a 1 mg/L increase in nitrate as nitrogen. Models adjust for the full covariate set, birth year-quarter fixed effects, and standard errors clustered at the public water system level.

2.5.4 Race/Ethnicity (mother level) Heterogeneity

Table 2.7 evaluates whether the association between pregnancy-average nitrate exposure and adverse birth outcomes differs by maternal race and ethnicity using race-by-exposure

interaction models. The main text focuses on the continuous and binary exposure specifications, while the five-category interaction results are reported in Appendix Table 2.A4.

Table 2.7: Race and Ethnicity Differences in Adjusted Nitrate-Birth Outcome Associations

Outcome	Exposure specification	Estimate type and group	Estimate	95% CI	p-value
Preterm birth	Continuous (per 1 mg/L as N)	Non-Hispanic White or other race/ethnicity	1.020	[1.010, 1.030]	<0.001***
		Non-Hispanic Black	1.030	[1.010, 1.040]	<0.001***
		Non-Hispanic Black vs non-Hispanic White or other	1.010	[0.990, 1.020]	0.204
		Hispanic	1.020	[1.010, 1.030]	<0.001***
		Hispanic vs non-Hispanic White or other	1.000	[0.990, 1.010]	0.986
	Binary (≥ 10 vs <10 mg/L as N)	Non-Hispanic White or other race/ethnicity	1.040	[0.930, 1.160]	0.512
		Non-Hispanic Black	1.430	[1.230, 1.660]	<0.001***
		Non-Hispanic Black vs non-Hispanic White or other	1.380	[1.130, 1.670]	0.001***
		Hispanic	1.090	[1.000, 1.200]	0.055*

		Hispanic vs non-Hispanic White or other	1.050	[0.910, 1.220]	0.492
Low birthweight	Continuous (per 1 mg/L as N)	Non-Hispanic White or other race/ethnicity	1.010	[1.000, 1.030]	0.027**
		Non-Hispanic Black	1.020	[1.010, 1.040]	0.003***
		Non-Hispanic Black vs non-Hispanic White or other	1.010	[0.990, 1.030]	0.361
		Hispanic	1.010	[1.010, 1.020]	<0.001***
		Hispanic vs non-Hispanic White or other	1.000	[0.980, 1.010]	0.650
	Binary (≥ 10 vs < 10 mg/L as N)	Non-Hispanic White or other race/ethnicity	0.930	[0.850, 1.020]	0.129
		Non-Hispanic Black	1.170	[1.030, 1.330]	0.016**
		Non-Hispanic Black vs non-Hispanic White or other	1.250	[1.110, 1.410]	<0.001***
		Hispanic	1.020	[0.950, 1.090]	0.635
		Hispanic vs non-Hispanic White or other	1.090	[0.970, 1.230]	0.160

Note. The table reports selected estimates from separate adjusted interaction models that include race or ethnicity by nitrate exposure terms. ORs are group-specific odds ratios for the association between pregnancy-average nitrate exposure and the listed birth outcome. RORs are ratios of

odds ratios and compare the nitrate association for non-Hispanic Black or Hispanic mothers with the association for non-Hispanic White or other race/ethnicity mothers. Values above 1 indicate a stronger positive nitrate association for the listed group relative to the reference group. Models adjust for maternal age, education, smoking during pregnancy, primiparity, pre-pregnancy BMI, previous preterm birth, infant sex, prenatal payer category, community water system urban-rural context, and birth year-quarter fixed effects, with standard errors clustered at the public water system level. OR refers to odds ratio, ROR refers to ratio of odds ratios, and CI refers to confidence interval. Statistical significance: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

For preterm birth, the continuous nitrate specification showed small positive associations across maternal race and ethnicity groups, but there was no evidence that the slope differed for non-Hispanic Black or Hispanic mothers relative to the non-Hispanic White/Other reference group (Table 2.7). By contrast, the binary ≥ 10 mg/L as N specification indicated clearer heterogeneity. Elevated nitrate at or above 10 mg/L as N was not clearly associated with preterm birth among non-Hispanic White/Other births or Hispanic births, but it was associated with higher odds among non-Hispanic Black births (OR = 1.430, 95% CI [1.230, 1.660], $p < 0.001$), and the corresponding ratio of odds ratios relative to the reference group was also statistically significant (ROR = 1.380, 95% CI [1.130, 1.670], $p = 0.001$). These results indicate that heterogeneity is not evident for the continuous exposure slope but emerges more clearly at the highest exposure range.

For low birthweight, the continuous nitrate specification again showed modest positive associations across groups without evidence of slope heterogeneity (Table 2.7). In the binary ≥ 10 mg/L as N specification, elevated nitrate was not clearly associated with low birthweight among non-Hispanic White/Other births or Hispanic births, but it was associated with higher odds among non-Hispanic Black births (OR = 1.170, 95% CI [1.030, 1.330], $p = 0.016$). The relative difference compared with the non-Hispanic White/Other reference group was also statistically significant (ROR = 1.250, 95% CI [1.110, 1.410], $p < 0.001$). As with preterm birth, the clearest

race and ethnicity heterogeneity appears in the highest exposure range rather than in the continuous exposure slope.

The five-category interaction models are consistent with this interpretation. For both preterm birth and low birthweight, the largest differentials appear in the highest nitrate category, particularly for non-Hispanic Black births, whereas lower exposure categories show less consistent evidence of heterogeneity (Appendix Table 2.A4). These race and ethnicity interaction results indicate that the overall nitrate association is fairly similar across groups at lower and moderate exposure levels, while disparities become more apparent in the highest exposure range.

2.5.5 Community Water System-Level Heterogeneity

Tables 2.7 and 2.8 use the same core interaction-model framework, including year-quarter fixed effects, clustered standard errors, a common covariate set, and multiple exposure parameterizations, but they differ in the effect modifier. Table 2.7 evaluates heterogeneity by maternal race and ethnicity at the individual level, whereas Table 2.8 evaluates heterogeneity by community water system context, specifically DAC, SDAC, and Central Valley status. To aid interpretation, both tables report group-specific odds ratios and ratios of odds ratios rather than interaction coefficients alone.

Table 2.8 assesses whether the association between pregnancy-average nitrate exposure and adverse birth outcomes varies by Central Valley status, DAC status, and SDAC status. For each modifier, the table reports the reference-group odds ratio, the focal-group odds ratio, and the ratio of odds ratios summarizing heterogeneity on the odds-ratio scale. Figures 2.6 through 2.11 visualize these quantities for the binary threshold and continuous exposure specifications.

Table 2.8: Community-Context Heterogeneity in Adjusted Nitrate–Birth Outcome Associations

Modifier	Outcome	Exposure contrast	Reference-group OR [95% CI]	<i>p</i>	Focal-group OR [95% CI]	<i>p</i>	ROR [95% CI]	<i>p</i>
SDAC	Preterm	≥10 vs <10 mg/L as N	1.077 [0.992, 1.169]	0.077 *	1.150 [1.037, 1.276]	0.008 ***	1.068 [0.935, 1.220]	0.330
SDAC	Preterm	Per 5 mg/L as N	1.098 [1.057, 1.142]	<0.001 ***	1.095 [1.030, 1.163]	0.004 ***	0.997 [0.928, 1.071]	0.924
SDAC	Low birthweight	≥10 vs <10 mg/L as N	0.992 [0.934, 1.054]	0.802	0.965 [0.835, 1.115]	0.626	0.972 [0.831, 1.137]	0.724
SDAC	Low birthweight	Per 5 mg/L as N	1.047 [1.004, 1.092]	0.031 **	1.088 [1.027, 1.152]	0.004 ***	1.039 [0.969, 1.115]	0.286
DAC	Preterm	≥10 vs <10 mg/L as N	1.122 [1.036, 1.216]	0.005 ***	1.166 [1.054, 1.289]	0.003 ***	1.039 [0.914, 1.181]	0.563
DAC	Preterm	Per 5 mg/L as N	1.084 [1.045, 1.124]	<0.001 ***	1.106 [1.052, 1.163]	<0.001 ***	1.020 [0.958, 1.086]	0.528
DAC	Low birthweight	≥10 vs <10 mg/L as N	1.011 [0.949, 1.076]	0.737	1.038 [0.916, 1.176]	0.562	1.027 [0.893, 1.180]	0.712
DAC	Low birthweight	Per 5 mg/L as N	1.049 [0.998, 1.102]	0.062 *	1.055 [1.007, 1.106]	0.026 **	1.006 [0.939, 1.078]	0.860

Central Valley	Preterm	≥ 10 vs < 10 mg/L as N	1.089 [1.007, 1.176]	0.032 **	1.126 [0.966, 1.312]	0.129	1.034 [0.872, 1.227]	0.699
Central Valley	Preterm	Per 5 mg/L as N	1.097 [1.054, 1.142]	< 0.001 ***	1.108 [1.040, 1.181]	0.002 ***	1.010 [0.937, 1.089]	0.789
Central Valley	Low birthweight	≥ 10 vs < 10 mg/L as N	1.000 [0.944, 1.060]	0.991	0.853 [0.619, 1.175]	0.331	0.853 [0.615, 1.181]	0.338
Central Valley	Low birthweight	Per 5 mg/L as N	1.052 [1.009, 1.098]	0.018 **	1.069 [0.997, 1.147]	0.061 *	1.016 [0.936, 1.102]	0.702

Note. The table reports selected estimates from separate adjusted interaction models for each community-context modifier, outcome, and exposure contrast. Reference-group ORs report the nitrate association among systems not in the focal category, and focal-group ORs report the nitrate association among systems in the focal category. For SDAC rows, the focal group is severely disadvantaged community systems; for DAC rows, the focal group is disadvantaged community systems; and for Central Valley rows, the focal group is Central Valley systems. ROR refers to the ratio of odds ratios and compares the focal-group nitrate association with the corresponding reference-group association. Values above 1 indicate a stronger positive nitrate association in the focal group relative to the reference group. Binary exposure compares pregnancy-average nitrate ≥ 10 mg/L as N with < 10 mg/L as N. Continuous exposure is scaled per 5 mg/L as N. Models adjust for maternal age, education, smoking during pregnancy, primiparity, pre-pregnancy BMI, previous preterm birth, infant sex, prenatal payer category, community water system urban-rural context, and birth year-quarter fixed effects, with standard errors clustered at the public water system level. Low birthweight models were estimated using a generalized linear model with binomial family and logit link to improve numerical stability under high-dimensional fixed effects. OR refers to odds ratio, and CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

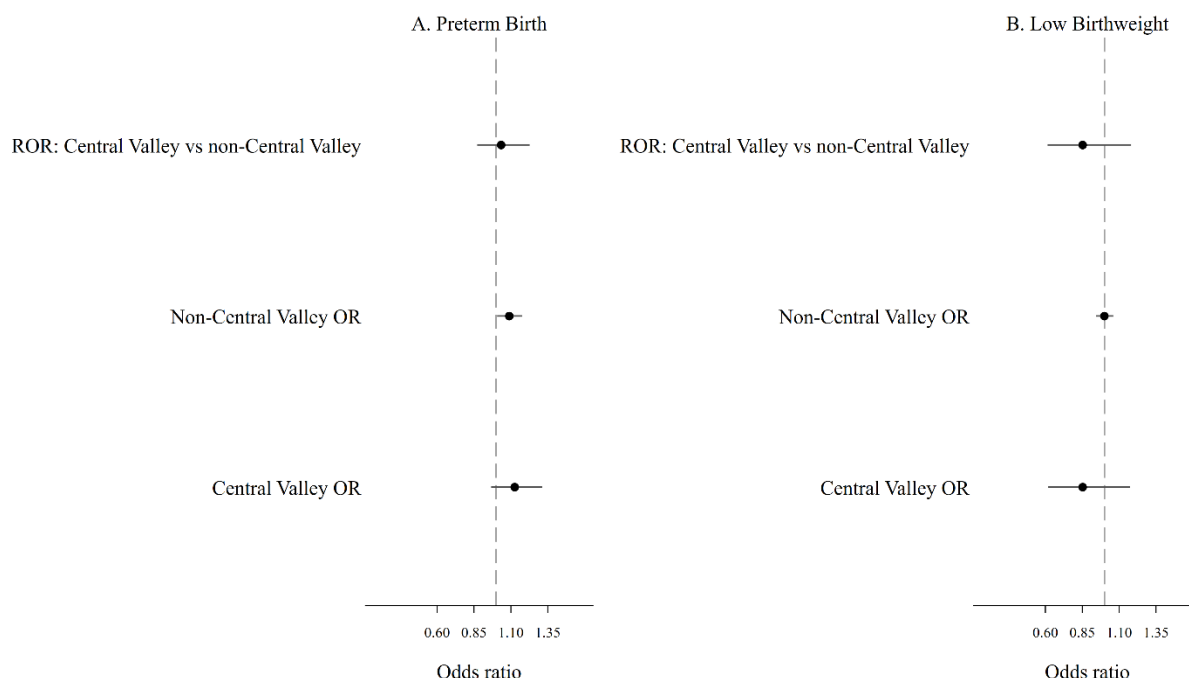


Figure 2.6: Central Valley Heterogeneity in Adjusted Nitrate-Threshold Associations

Note. Panels show group-specific odds ratios and ratios of odds ratios from Central Valley heterogeneity models. The binary exposure specification compares pregnancy-average nitrate ≥ 10 mg/L as N with < 10 mg/L as N. Points show odds ratios and horizontal lines show 95% confidence intervals. ROR refers to the ratio of odds ratios comparing the Central Valley association with the non-Central Valley association. Models adjust for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

Figure 2.6 visualizes the Central Valley threshold models, showing that the ≥ 10 versus < 10 mg/L as N association is positive for preterm birth in both Central Valley and non-Central Valley systems, while the ratio of odds ratios does not indicate statistically detectable heterogeneity. For preterm birth, estimated nitrate associations were generally positive across modifiers and exposure parameterizations, but there was little evidence that they differed significantly by SDAC, DAC, or Central Valley status. In the binary ≥ 10 mg/L as N models, elevated nitrate was associated with higher odds of preterm birth in several focal and reference groups, including DAC and SDAC systems, but the corresponding ratios of odds ratios remained close to 1.0 and were not statistically significant. The same pattern appeared in the continuous

models. Group-specific odds ratios were positive and similar across community contexts for preterm birth, while low birthweight estimates were smaller and less consistent. Across both outcomes, the heterogeneity terms showed no clear evidence of effect modification (Table 2.8; Figures 2.7, 2.9, and 2.11).

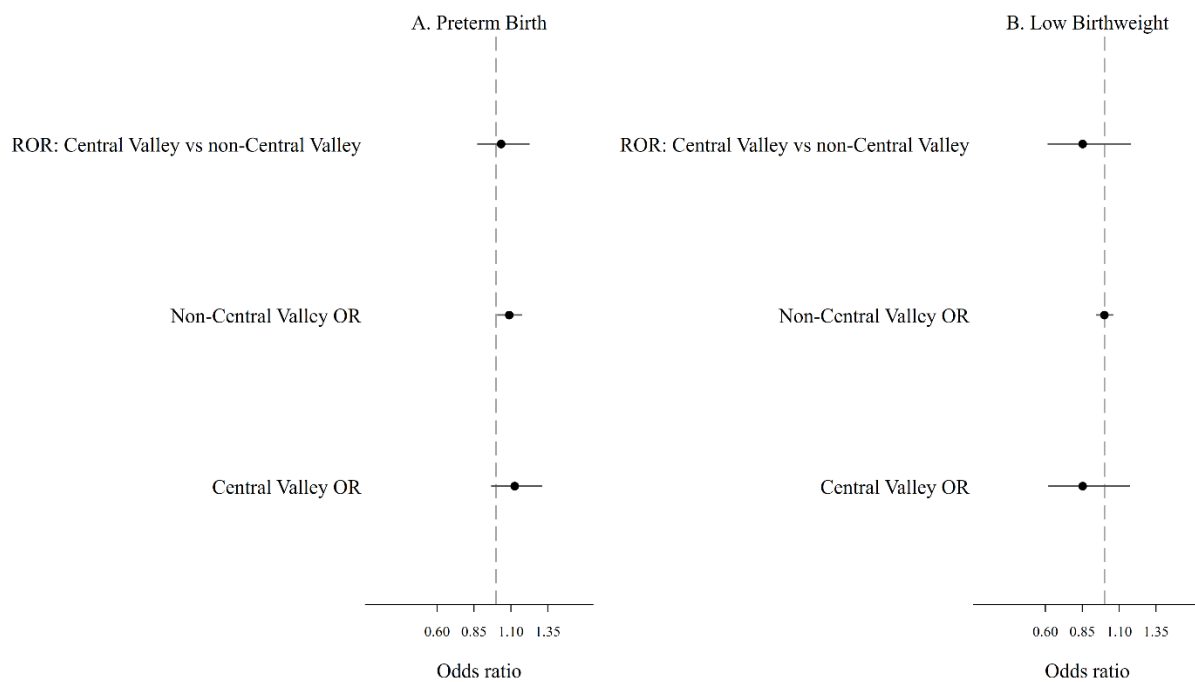


Figure 2.7: Central Valley Heterogeneity in Adjusted Continuous Nitrate Associations

Note. Panels show group-specific odds ratios and ratios of odds ratios from Central Valley heterogeneity models. The continuous exposure specification is scaled per 5 mg/L as N. Points show odds ratios and horizontal lines show 95% confidence intervals. ROR refers to the ratio of odds ratios comparing the Central Valley association with the non-Central Valley association. Models adjust for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

For low birthweight, the binary ≥ 10 mg/L as N results were generally null or imprecise, particularly for Central Valley systems. In the continuous models, several strata showed modest positive associations, but the ratios of odds ratios remained close to 1.0 and did not indicate

statistically detectable heterogeneity by Central Valley, DAC, or SDAC status. Overall, these results indicate that nitrate–birth outcome associations, especially for preterm birth, are broadly similar across community and system contexts on the odds-ratio scale (Table 2.8; Figures 2.6–2.11).

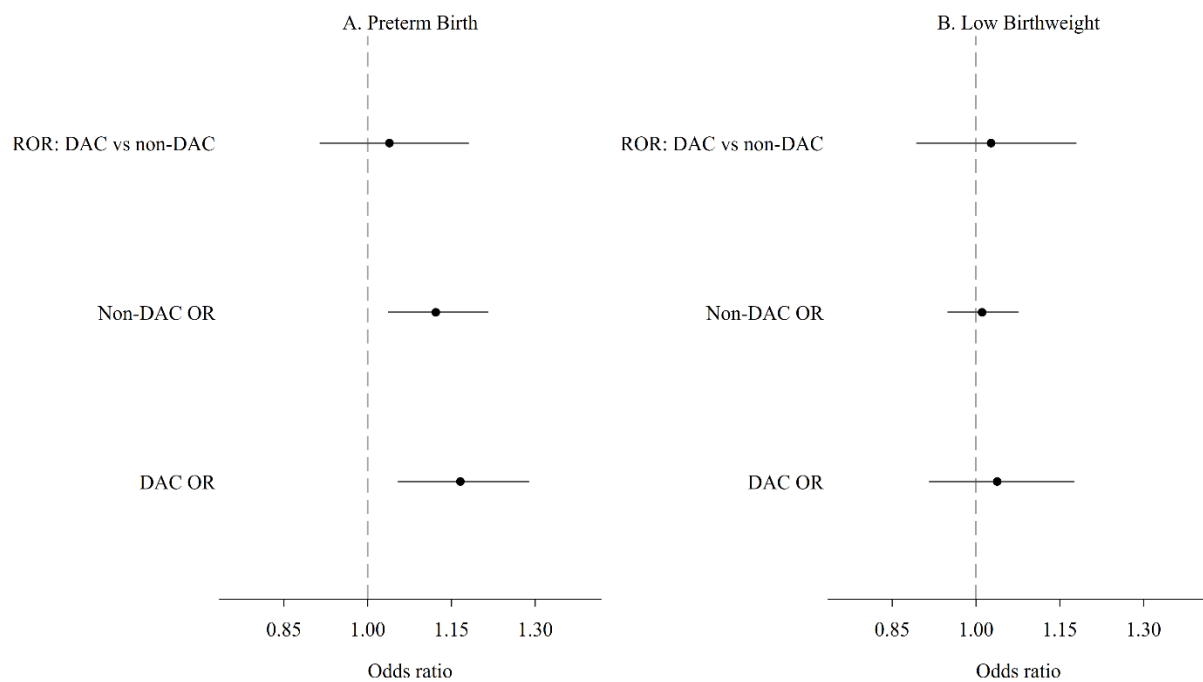


Figure 2.8: DAC Heterogeneity in Adjusted Nitrate-Threshold Associations

Note. Panels show group-specific odds ratios and ratios of odds ratios from DAC heterogeneity models. The binary exposure specification compares pregnancy-average nitrate ≥ 10 mg/L as N with < 10 mg/L as N. Points show odds ratios and horizontal lines show 95% confidence intervals. ROR refers to the ratio of odds ratios comparing the DAC association with the non-DAC association. Models adjust for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

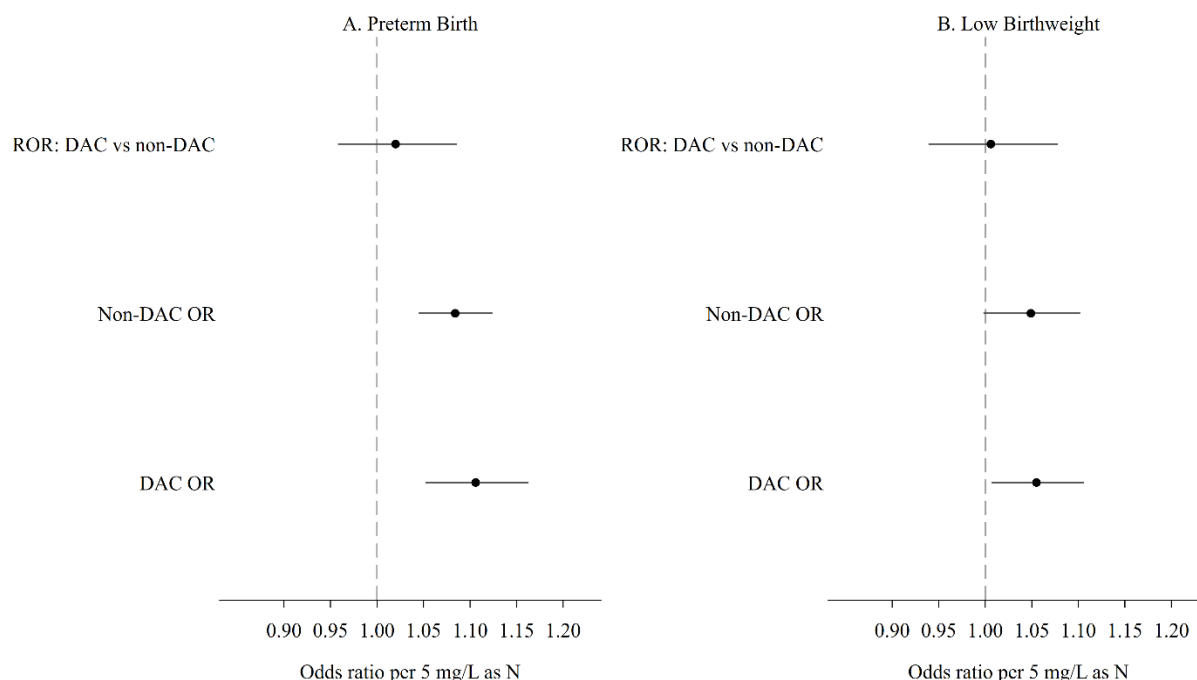


Figure 2.9: DAC Heterogeneity in Adjusted Continuous Nitrate Associations

Note. Panels show group-specific odds ratios and ratios of odds ratios from DAC heterogeneity models. The continuous exposure specification is scaled per 5 mg/L as N. Points show odds ratios and horizontal lines show 95% confidence intervals. ROR refers to the ratio of odds ratios comparing the DAC association with the non-DAC association. Models adjust for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

The categorical interaction models are consistent with this interpretation. Across SDAC, DAC, and Central Valley comparisons, some focal-group point estimates are somewhat larger in higher nitrate categories, but the categorical ratios of odds ratios do not show a clear or systematic pattern of community-level effect modification.

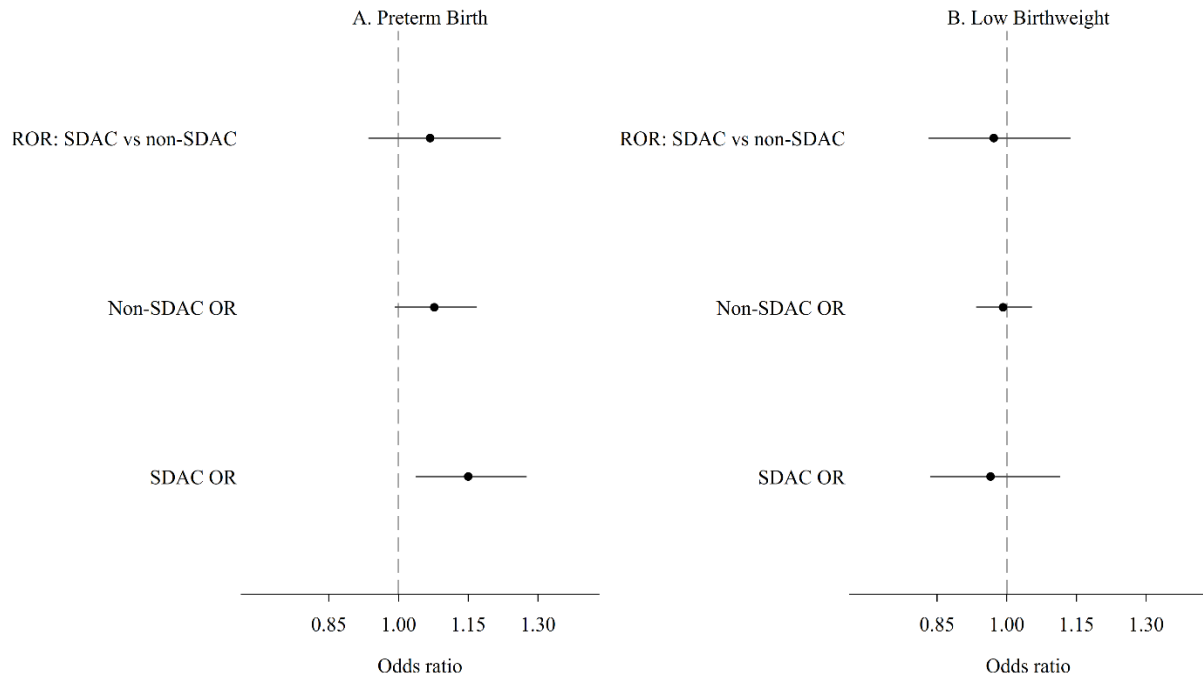


Figure 2.10: SDAC Heterogeneity in Adjusted Nitrate-Threshold Associations

Note. Panels show group-specific odds ratios and ratios of odds ratios from SDAC heterogeneity models. The binary exposure specification compares pregnancy-average nitrate ≥ 10 mg/L as N with < 10 mg/L as N. Points show odds ratios and horizontal lines show 95% confidence intervals. ROR refers to the ratio of odds ratios comparing the SDAC association with the non-SDAC association. Models adjust for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

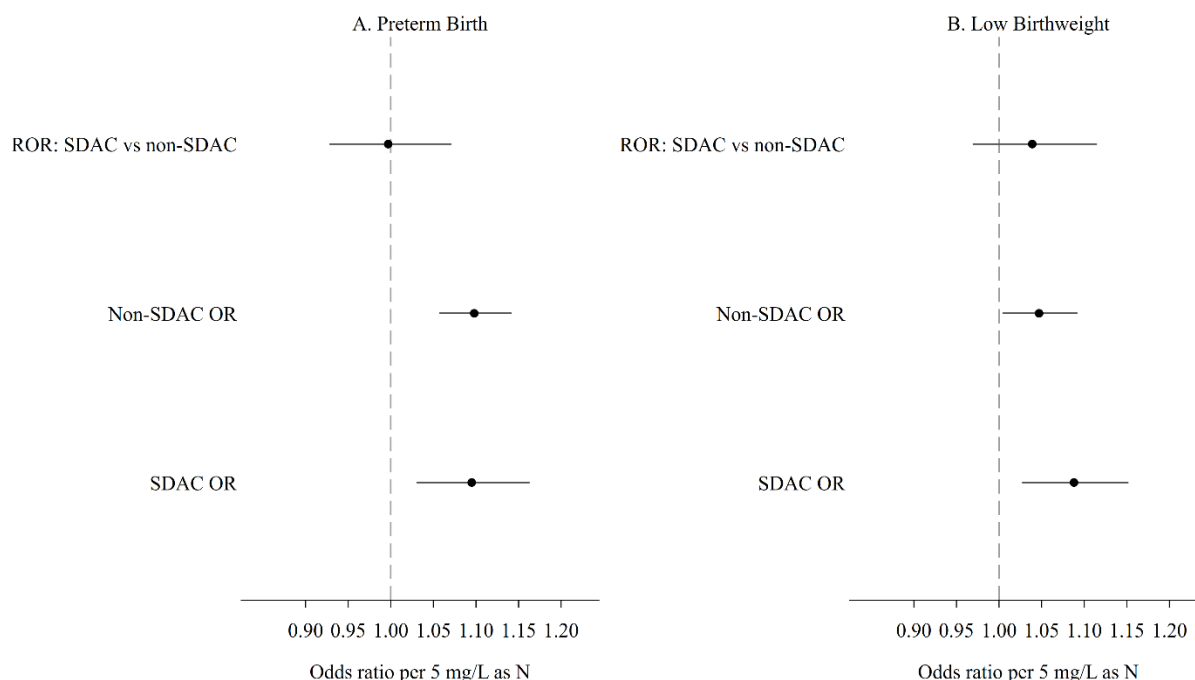


Figure 2.11: SDAC Heterogeneity in Adjusted Continuous Nitrate Associations

Note. Panels show group-specific odds ratios and ratios of odds ratios from SDAC heterogeneity models. The continuous exposure specification is scaled per 5 mg/L as N. Points show odds ratios and horizontal lines show 95% confidence intervals. ROR refers to the ratio of odds ratios comparing the SDAC association with the non-SDAC association. Models adjust for the full covariate set and birth year-quarter fixed effects, with standard errors clustered at the public water system level.

2.6 Sensitivity Checks

Sensitivity analyses assessed whether the main conclusions were robust to alternative outcomes and contextual specifications. This chapter examined very preterm birth as a more severe outcome, tested heterogeneity by groundwater-only system status, and re-estimated the primary models within a Central Valley-only subsample. These checks evaluated whether nitrate–birth outcome associations changed in settings characterized by greater aquifer vulnerability, distinct environmental justice profiles, or a rarer and more clinically severe

endpoint (Sherris et al., 2021; Royal et al., 2024). All sensitivity models retained the same exposure construction, covariate set, fixed effects, and clustering strategy as the main analyses.

2.6.1 Very Preterm Birth (VPTB)

Very preterm birth, defined here as delivery before 32 completed weeks of gestation, is examined as a sensitivity outcome because it is a rarer and more clinically severe endpoint than preterm birth defined at fewer than 37 weeks. It is also less likely to be driven by borderline gestational-age classification. Appendix Table 2.A2 reports descriptive statistics for supplemental outcome measures, including very preterm birth. Because very preterm birth is uncommon, this sensitivity check focuses on pooled pregnancy-average nitrate models rather than extending the full trimester-specific, subgroup, and heterogeneity analyses to a lower-prevalence outcome where estimates are more vulnerable to sparse-cell instability.

The pooled models were broadly consistent with the main preterm birth results. In the continuous specification, higher pregnancy-average nitrate was associated with higher odds of very preterm birth, with an odds ratio of 1.029 per 1 mg/L as N increase (95% CI [1.019, 1.039], $p < 0.001$; Appendix Table 2.A5). The categorical specification likewise showed a positive gradient across exposure categories relative to <1 mg/L as N, with the largest estimate in the highest exposure category. By contrast, the binary ≥ 10 versus <10 mg/L as N specification was not statistically distinguishable from the null. This pattern suggests that the positive nitrate association observed for preterm birth is also visible for the more severe very preterm endpoint, while reinforcing that dichotomizing exposure at the regulatory threshold is less informative for a rarer outcome.

Overall, the very preterm birth sensitivity analysis supports the main chapter interpretation. Continuous and multi-category nitrate specifications show positive associations,

whereas the binary threshold model is less precise and less informative. The analysis reinforces the primary preterm birth findings without extending the chapter's timing or heterogeneity claims to a rarer outcome.

2.6.2 Groundwater-only Systems

This sensitivity analysis examines whether nitrate birth-outcome associations differ between births served by groundwater-only community water systems and those served by non-groundwater systems. The motivation is that nitrate exposure in groundwater-only systems may reflect different source-water conditions, treatment constraints, or correlated contaminant profiles than exposure in surface-water or mixed-source systems. The models estimate nitrate-by-groundwater-only interactions while retaining the same exposure definitions, covariate set, fixed effects, and clustering strategy used in the main analyses.

The results show little evidence that nitrate associations differ by groundwater-only system status. For preterm birth, both the binary and continuous specifications show positive nitrate associations in non-groundwater and groundwater-only systems, but the ratios of odds ratios are close to 1.0 and not statistically significant (Appendix Table 2.A6). For low birthweight, the pattern is similar. Continuous nitrate is positively associated with low birthweight in both system types, while the binary threshold specification is null or imprecise. The interaction estimates do not indicate meaningful heterogeneity by groundwater-only status.

Overall, the groundwater-only sensitivity analysis supports the main chapter interpretation. Nitrate birth-outcome associations are broadly similar across groundwater-only and non-groundwater systems, and the results do not indicate that groundwater-only systems drive the primary findings.

2.6.3 Central Valley Subsample Sensitivity

To assess whether the main results hold within California's Central Valley, this chapter re-estimates the nitrate birth-outcome models after restricting the sample to births served by Central Valley systems. This sensitivity analysis is substantively important because Central Valley combines intensive agriculture, groundwater dependence, and persistent nitrate concerns. The models retain the same exposure definitions, covariate set, fixed effects, and clustering strategy used in the main analyses.

Within the Central Valley subsample, the continuous nitrate specification remains positively associated with both outcomes. Each 1 mg/L increase in pregnancy-average nitrate is associated with higher odds of preterm birth (OR = 1.021, 95% CI [1.008, 1.035], $p = 0.002$) and low birthweight (OR = 1.019, 95% CI [1.007, 1.031], $p = 0.001$) (Appendix Table 2.A7). By contrast, the binary ≥ 10 mg/L as N indicator is not statistically distinguishable from the null for either outcome. These results show that the positive continuous nitrate association remains present within Central Valley systems, while the threshold specification provides less stable information in the restricted subsample.

Overall, the Central Valley-only sensitivity analysis supports the chapter's main interpretation. Continuous pregnancy-average nitrate remains positively associated with adverse birth outcomes in this focal nitrate-vulnerable region, whereas the binary threshold contrast is imprecise and not statistically significant. This pattern reinforces the value of modeling nitrate as a continuous exposure rather than relying only on the regulatory threshold.

2.6.4 Alternative Categorical Cutpoints

The main categorical nitrate specification uses policy-oriented cutpoints to preserve substantively meaningful contrasts tied to regulatory and public-health interpretation rather than

equal sample size across exposure groups. To assess whether the categorical results depend on that choice, this chapter re-estimates the models using distribution-based quintiles and quartiles of pregnancy-average nitrate. The quintile cutpoints occur at 1.643, 2.668, 3.456, and 4.791 mg/L as N, while the quartile cutpoints occur at 2.010, 3.143, and 4.340 mg/L as N.

The alternative categorizations produce a similar positive pattern for both preterm birth and low birthweight. Relative to the lowest exposure group, higher quintiles and quartiles are associated with higher odds of both outcomes, with the clearest elevations appearing in the upper exposure groups. These results show that the chapter's categorical findings are not driven by the policy-based cutpoints used in the main specification and remain consistent under distribution-based categorizations of pregnancy-average nitrate. Appendix Table 2.A8 reports the quintile and quartile cutpoints and the corresponding sensitivity models.

2.6.5 Within-System Spatial Heterogeneity in Nitrate Exposure

To assess whether system-level averaging masks higher nitrate concentrations at particular sampling locations within community water systems, this chapter re-estimates the main models using a maximum sampling-location exposure metric. This sensitivity analysis constructs a monthly nitrate series from the highest sampling-location mean observed within each system-month and links that series to births using the same gestational averaging approach as the main exposure measure.

Within-system spatial variation is substantively important in the monitoring data. Among systems with identifiable sampling-point names, 64.6 percent had more than one distinct sampling location, with a median of two locations and an interquartile range from one to four. In addition, 40.1 percent of nitrate system-months included more than one sampled location. Among multi-location system-months, 23.3 percent of months below 10 mg/L as N using the

system mean crossed the 10 mg/L threshold using the maximum-location series. These patterns indicate that system-level averaging can mask higher localized nitrate conditions within some systems.

The maximum-location sensitivity models preserve the main pattern. For preterm birth, pregnancy-average maximum-location nitrate remains positively associated with the outcome in the continuous, higher categorical, and threshold specifications. For low birthweight, the associations are smaller than those for preterm birth but remain positive in the continuous, mid-range categorical, highest-category, and threshold specifications. These results show that the main findings are not driven solely by dilution from system-level averaging and remain visible under an alternative exposure construction designed to capture higher localized nitrate concentrations within systems (Appendix Table 2.A9). This supports the chapter's broader interpretation that the nitrate associations are robust to a more spatially conservative exposure metric.

2.6.6 Co-Occurring Contaminants and Matched-Sample Co-Pollutant Adjustment

A related concern is that nitrate may proxy for a broader mixture of groundwater contaminants in agricultural regions. In the pregnancy-level data, nitrate was only weakly correlated with arsenic, 1,2,3-TCP, and DBCP, with Pearson correlations of 0.0666, 0.0822, and 0.1571, respectively. Matched-sample co-pollutant models also showed little change in the nitrate coefficient after adjustment for each contaminant. The strongest comparison comes from the arsenic overlap sample, which is substantially larger than the 1,2,3-TCP and DBCP overlap samples. In that sample, the preterm nitrate coefficient changed from 0.0117 to 0.0111, and the low birthweight nitrate coefficient changed from 0.0068 to 0.0064 after arsenic adjustment.

The 1,2,3-TCP and DBCP checks provide more limited supplementary evidence because they rely on smaller overlap samples. In those restricted samples, the preterm nitrate coefficient is already negative in the nitrate-only model, indicating that the sign difference reflects the composition of the overlap sample rather than the effect of adding the co-pollutant. Overall, these checks do not remove the broader conceptual concern that nitrate can be part of a larger contamination mixture, but they show that the main nitrate associations are not explained by strong collinearity with these three measured co-contaminants. Appendix Table 2.A10 reports the matched-sample estimates before and after co-pollutant adjustment.

2.7 Discussion

2.7.1 Interpretation and Implications

These findings add to a growing body of evidence that nitrate in drinking water is relevant to perinatal health within the exposure range observed in regulated community water systems. They also show that the apparent strength of the association depends on how exposure is parameterized. For preterm birth, the association appears across continuous, categorical, and binary threshold specifications, with a graded pattern across exposure categories. This pattern indicates that a framework focused only on exceedance of the 10 mg/L standard may overlook meaningful variation in risk across the broader exposure distribution. In policy terms, a compliance threshold and a health-relevant exposure gradient are not necessarily the same thing. Efforts that reduce nitrate from moderate to lower levels may therefore matter for maternal and infant health, even when systems do not cross the regulatory threshold.

The low birthweight results point in a related but more qualified direction. Estimated associations are smaller and more sensitive to exposure specification. Continuous nitrate and mid-range exposure categories show modest positive associations, whereas the binary threshold

comparison is close to null. This pattern suggests that low birthweight is not well captured here as a simple high-versus-low exposure contrast. A defensible interpretation is that the signal is either distributed across a wider range of nitrate concentrations or diluted because low birthweight reflects multiple etiologic pathways and may be less responsive to a single community-level environmental exposure measure. For water agencies and practitioners, this again points to the value of continuous monitoring and incremental exposure reduction when assessing maternal and child health relevance.

The trimester-specific results add an epidemiologic timing perspective. These models are useful because exposure timing can inform biologic plausibility even when the evidence does not isolate a single definitive critical window. Potential nitrate-related pathways, including oxygen transport disruption through methemoglobinemia, endogenous nitrosation, and thyroid-related mechanisms, may not be equally relevant across gestation. In that sense, the stronger positive estimates observed in some third-trimester categorical models are plausibly consistent with greater vulnerability in late pregnancy, when fetal growth and birthweight accumulation are especially salient. At the same time, the lack of consistency across continuous, threshold, and categorical parameterizations requires caution. This chapter therefore treats the trimester models as biologically motivated time-window tests that provide suggestive evidence of temporal specificity, while treating the pregnancy-average models as the more stable summary of the overall nitrate-birth outcome association.

From an environmental justice perspective, two aspects of the results are especially important. The first is the persistence of positive continuous associations within the Central Valley. In the Central Valley-only subsample, continuous nitrate remains positively associated with both preterm birth and low birthweight. This does not mean that Central Valley systems

show a statistically distinct nitrate-birth outcome association, since the place-based interaction tests are generally limited. It does, however, show that the statewide association remains visible in the region where nitrate management is most policy-salient. That finding supports continued attention to Central Valley systems in efforts related to source protection, treatment investment, consolidation, and long-term groundwater quality management.

The second environmental justice implication concerns heterogeneity by maternal race and ethnicity. The clearest difference appears in the highest exposure contrast: in the binary models comparing concentrations at or above 10 mg/L with lower concentrations, births to non-Hispanic Black mothers show larger estimated associations for both preterm birth and low birthweight than births in the non-Hispanic White or Other reference group. By contrast, continuous slope heterogeneity is not statistically detectable. This pattern suggests that heterogeneity is concentrated at the upper end of the observed exposure distribution rather than distributed evenly across the nitrate range. It is most consistent with structural inequality in exposure and vulnerability, including unequal co-exposures, unequal protection, differences in prenatal care access, and other background conditions correlated with both place and exposure. The results therefore point to the importance of targeting protection and risk reduction in communities that are more likely to experience both high-end nitrate exposure and broader forms of vulnerability.

The findings also have implications for how regulators, water systems, and public-health agencies think about compliance, prioritization, and risk communication. A single maximum contaminant level remains administratively necessary, but the empirical pattern observed here indicates that a threshold-centered compliance framework does not fully align with how health-relevant exposure gradients appear in observational data, especially for preterm birth. One

practical implication is that systems with persistently moderate nitrate levels may still warrant preventive attention, technical assistance, or closer surveillance even when they remain below the regulatory limit. In Central Valley communities shaped by agricultural loading and legacy groundwater contamination, prevention should therefore combine long-term nitrate reduction with near-term exposure interruption. Recent Central Valley nitrate programs illustrate the practical components of this approach, including targeted testing, direct outreach, replacement drinking water, bottled-water deliveries, filling stations, and assistance when nitrate exceeds the drinking-water standard (State Water Board, 2025a).

For community water systems, the parallel implication is to connect compliance management with nitrate-specific monitoring, clear risk communication for pregnant mothers and households preparing infant formula, interim safe-water provision when elevated nitrate is detected or strongly suspected, and treatment or consolidation strategies where durable system-level correction is needed. Point-of-use or point-of-entry treatment may also be relevant in some small-system or household contexts, but only as a managed protection strategy rather than as an informal household responsibility. State and federal guidance indicates that treatment devices must be appropriate for nitrate, certified or registered where applicable, properly installed, regularly maintained, and verified through treated-water testing; when used as part of a public-water-system compliance strategy, they also require system responsibility, monitoring, customer access, and sustained public education (State Water Board, 2025b; EPA, 2006). These near-term measures are not substitutes for source reduction, aquifer protection, treatment investment, and long-run planning. They are necessary complements because nitrate stored in soils and groundwater can continue moving through aquifer systems for decades, so improvements in land

and nutrient management may take many years to translate into lower delivered-water concentrations (Harter et al., 2012; Ransom et al., 2017).

This operational perspective also connects to broader planning questions in vulnerable rural and agricultural communities. Recent justice-oriented studies emphasize that durable responses to environmental burdens in low-income agricultural regions require integrating water, land-use, and socioeconomic objectives rather than treating each problem in isolation (Fernandez-Bou et al., 2021). That perspective is especially relevant for nitrate because reducing exposure requires both drinking-water interventions and upstream changes in land and nutrient management. Related research also points to the importance of place-based planning that accounts for cumulative vulnerability and community buffering capacity when designing adaptation and environmental health interventions (Tariqi & Naughton, 2021). In this sense, the Central Valley-focused and place-based analyses in this chapter are relevant not only because the region faces elevated nitrate vulnerability, but also because nitrate exposure is embedded within a broader landscape of agricultural dependence, groundwater stress, infrastructure constraints, and uneven social protection.

The sensitivity analyses clarify what should be emphasized and what should not be overread. The very preterm birth models reinforce the main pattern in the continuous and multi-category specifications, which is notable given the clinical severity of very preterm birth and its lower vulnerability to borderline gestational-age classification. By contrast, threshold-based and interaction-heavy models are less stable for very preterm birth and in restricted subsamples, as expected when exposure is dichotomized or when rarer outcomes are analyzed in smaller populations. These less stable results do not overturn the broader pattern. The strongest signal in this chapter lies in the repeated appearance of positive continuous and multi-category

associations, especially for preterm birth, alongside more selective evidence of heterogeneity at the highest exposure range.

Taken together, the chapter supports a cautious but clear interpretation. Nitrate-birth outcome associations are detectable in population-scale linked administrative data, but their visibility depends on outcome choice, exposure parameterization, and the limits of community-level exposure assignment. The most consistent evidence appears for preterm birth and for continuous or multi-category exposure specifications. The clearest heterogeneity signals emerge at the upper end of the exposure distribution, particularly for births to non-Hispanic Black mothers, while place-based modifiers show less evidence of differential associations on the odds-ratio scale. These findings support a view of nitrate as a public health concern that is not confined to rare regulatory exceedances and is best understood within a broader framework of cumulative vulnerability, place-based inequality, and long-run drinking water governance.

2.7.2 Limitations

A central limitation of this chapter is that pregnancy exposure is assigned using maternal residence at delivery. This approach can misclassify nitrate exposure when mothers move during pregnancy, particularly if a move changes the community water system serving the residence. The limitation is more consequential here because the California birth records available for analysis do not include a unique maternal identifier, which prevents direct measurement of within-pregnancy moves or linkage across repeated births. As a result, this chapter cannot estimate key diagnostic quantities, such as the proportion of mothers who changed community water systems during pregnancy and cannot implement within-mother designs that would help account for stable maternal characteristics. A related limitation concerns linkage and external validity. Although linked births were modestly more disadvantaged than unlinked births on

several observed maternal socioeconomic indicators, the linked analytic sample was overwhelmingly urban and heavily concentrated in larger community water systems. The chapter's estimates therefore speak most directly to births linked to larger urban CWSs rather than to the full universe of rural births, private-well users, or other unlinked populations.

A related concern is avoidance behavior. If mothers selectively relocate or change water-use practices in response to perceived water-quality risk, estimated associations could be biased in ways that are difficult to predict. At the same time, existing evidence suggests that contamination-driven relocation may be limited. For example, Currie et al. (2013) find little evidence that mothers most affected by contamination were more likely to move away, which reduces concern that systematic mobility is the primary driver of estimated associations. More broadly, reviews in pregnancy environmental epidemiology identify residential mobility as a common source of exposure error and recommend sensitivity analyses where feasible (Clemmensen et al., 2023; Royal et al., 2024). In this chapter, residential mobility is therefore treated primarily as a source of exposure misclassification that may weaken estimated associations, while acknowledging that more complex mobility patterns cannot be ruled out with the available data. Related study using mother fixed effects and contaminant sampling data illustrates one possible strategy for reducing bias from stable maternal traits, but it also underscores the importance of high-quality linkage between maternal residence and environmental measures (DiSalvo & Hill, 2024).

A related limitation concerns monitoring cadence and, therefore, exposure precision. The system-year summaries show that nitrate monitoring was often sparse even among the system-years that contributed births to the analytic sample, with a median of 2 sampled months per contributing system-year and 58.1 percent of system-years having 2 or fewer sampled months

(Appendix Table 2.A1). The higher pregnancy-level median of 8 monitored gestational months reflects where most linked births occur rather than a contradiction in monitoring cadence. Because the pregnancy-level summary is effectively birth-weighted, it is dominated by larger, more frequently monitored systems, whereas the unweighted system-year summary reflects the typical contributing system-year, including many smaller systems with sparse monitoring. This matters because pregnancy-average nitrate is assigned from observed system-level monitoring rather than continuously observed household exposure. When temporal monitoring is limited, assigned exposure is supported by less direct information and is more vulnerable to measurement error. To the extent that this error is nondifferential with respect to birth outcomes, it would be expected to attenuate estimated associations toward the null rather than create spurious positive associations.

A second limitation concerns the gap between system-level regulatory monitoring and individual ingested dose. Assigning nitrate at the community water system level is appropriate for this chapter because the CWS is the regulatory and distribution-system unit, and compliance samples provide the best available population-scale measure of delivered drinking-water quality. However, these data cannot fully capture bottled-water substitution, filtration, household water consumption, short-term concentration changes, or within-system spatial variation. These factors may create exposure measurement error between measured system nitrate and individual dose, even when the system-level assignment is policy-relevant.

Residual confounding also remains plausible. Nitrate is correlated with agricultural intensity and may co-vary with other contaminants, including pesticides, as well as with broader contextual exposures such as air pollution, heat, and other place-based risk factors that are not fully measured in vital records. Although the models adjust for a substantial set of maternal and

pregnancy covariates and include time fixed effects, these controls cannot eliminate confounding from unobserved co-exposures or from structural conditions that vary with nitrate over time and space.

Finally, outcome measurements in vital records is generally strong but not error-free. Gestational age and birthweight may contain some recording error, and broad outcomes such as preterm birth combine multiple etiologic pathways that may not respond uniformly to nitrate exposure. The estimates reported in this chapter should therefore be interpreted as adjusted population-level associations that are consistent with, but do not by themselves identify, specific biological mechanisms.

2.7.3 Future Directions

Several directions would strengthen both inference and policy relevance in this research area. A first priority is improving exposure assignment, while recognizing that direct tracking of maternal residence throughout pregnancy is difficult to implement at statewide scale. Future work could make progress through stable-residence sensitivity analyses, smaller linked health-system or insurance cohorts, or address-history data where available. These approaches would help assess exposure misclassification from residential mobility and could support within-mother designs that better account for stable maternal characteristics (Clemmensen et al., 2023; DiSalvo & Hill, 2024; Royal et al., 2024).

A second priority is complementing community water system-level exposure assignment with more refined measures of individual exposure opportunity. More precise linkage to service areas, distribution zones, or household-level monitoring data could reduce error from within-system heterogeneity and temporal averaging. An important direction for future research is improved information on bottled-water use, filtration, and water consumption, since these

behaviors shape the relationship between measured system nitrate and actual ingested dose (Coffman et al., 2022; Jensen et al., 2023; Ward et al., 2018).

A third direction is to engage more directly with co-exposures and mixtures. In agricultural regions, nitrate may move with pesticides, heat, air pollution, and other environmental stressors. As a result, some subgroup differences observed in this chapter may reflect broader exposure bundles and unequal background conditions rather than nitrate alone. Future studies should therefore test whether the associations reported here persist when better measures of correlated environmental exposures are incorporated (Royal et al., 2024; Tariqi & Naughton, 2021).

A fourth priority is stronger causal identification. This chapter estimates adjusted associations, not causal effects. Future work should build on these results using designs that exploit quasi-experimental variation in drinking water quality, such as drought-related source shifts, treatment upgrades, well shutdowns, consolidation events, or other changes in supply and infrastructure. California may be especially well suited to this type of analysis given the meaningful regulatory, hydrologic, and operational changes that water systems experience over time.

A fifth direction is to deepen the environmental justice dimension of the analysis. The limited place-based heterogeneity in slope estimates does not imply that inequity is unimportant, since unequal exposure can persist even when effect-size differences are modest. Future research should continue to develop place-based frameworks that connect water quality, land use, and rural vulnerability. Justice-oriented studies emphasize that durable responses in nitrate-vulnerable agricultural regions require integrating water, land, and socioeconomic objectives rather than treating contamination as an isolated technical problem (Fernandez-Bou et al., 2021).

In the Central Valley especially, this suggests value in research that examines whether source reduction, groundwater protection, cropland transition, or regional governance changes produce joint benefits for drinking water quality and maternal and infant health.

Finally, future work should broaden the outcome side of the analysis. Preterm birth and low birthweight are important summary endpoints, but they combine multiple etiologic pathways. More specific outcomes, such as spontaneous versus medically indicated preterm birth, fetal growth restriction, or longer-run child health measures, could help clarify which mechanisms are most consistent with nitrate-related associations.

These directions suggest that the next stage of research should focus on better exposure measurement, stronger causal identification, and deeper integration of environmental justice and water-governance perspectives. This chapter provides a foundation by showing that nitrate-related birth outcome associations are visible in statewide linked administrative data. Future work can build on that foundation by tracing those associations more precisely, identifying when they change, and clarifying how they translate into unequal health risk across places and populations.

2.8 Conclusion

Using a statewide linkage between California birth records and community water systems, this chapter finds a consistent positive association between pregnancy-average nitrate exposure and preterm birth. In the continuous specification, each 1 mg/L increase in pregnancy-average nitrate was associated with higher adjusted odds of preterm birth (OR = 1.019, 95% CI [1.012, 1.026]). The categorical and binary threshold specifications showed the same positive direction, indicating that the preterm association is not confined to a single functional form and remains visible both below and above the 10 mg/L as N regulatory threshold.

For low birthweight, the overall pattern was weaker and more sensitive to exposure parameterization. Continuous nitrate exposure and mid-range categorical contrasts were positively associated with low birthweight, while the binary ≥ 10 versus < 10 mg/L as N comparison was close to null. This pattern shows that low birthweight associations are more evident when nitrate is modeled continuously or across multiple exposure categories than when exposure is reduced to a single regulatory-threshold contrast.

The heterogeneity results were selective rather than uniform. Most interaction tests were small on the odds-ratio scale, especially for continuous exposure slopes and for place-based modifiers such as DAC, SDAC, and Central Valley status. The clearest heterogeneity signal appeared in the binary ≥ 10 mg/L as N specification for maternal race and ethnicity. For both preterm birth and low birthweight, the threshold association was larger among births to non-Hispanic Black mothers than among births to the non-Hispanic White/Other reference group. In contrast, continuous-exposure slope heterogeneity by maternal race and ethnicity was not statistically detectable, and community-level modifiers showed broadly similar associations across focal and reference groups.

Sensitivity analyses reinforced the same core interpretation. Very preterm birth, Central Valley-only models, alternative categorical cutpoints, maximum-location exposure models, groundwater-only comparisons, and co-pollutant adjustment checks all supported the broader pattern that positive nitrate associations are most consistent when exposure is modeled continuously or across multiple categories. Threshold and interaction specifications were less stable, especially for rarer outcomes and restricted subsamples.

Overall, this chapter shows that higher nitrate concentrations in community water systems are associated with higher odds of preterm birth and, more weakly, low birthweight,

with the most consistent evidence appearing for preterm birth. The findings support the dissertation's broader unequal-protection framework by showing that drinking-water exposure is not only a compliance issue, but also a population-health concern. The results also show that exposure parameterization matters. Associations are clearer when nitrate is modeled as a continuous or multi-category exposure than when it is reduced to a single regulatory threshold.

2.9 References

- Acquah, S., & Allaire, M. (2023). Disparities in drinking water quality: evidence from California. *Water Policy*, 25(2), 69–86.
- Albouy-Llaty, M., Limousi, F., Carles, C., Dupuis, A., Rabouan, S., & Migeot, V. (2016). Association between exposure to endocrine disruptors in drinking water and preterm birth, taking neighborhood deprivation into account: a historic cohort study. *International Journal of Environmental Research and Public Health*, 13(8), 796.
- Balazs, C. L., Morello-Frosch, R., Hubbard, A., & Ray, I. (2011). Social disparities in nitrate-contaminated drinking water in California's San Joaquin Valley. *Environmental Health Perspectives*, 119(9), 1272–1278.
- Balazs, C. L., & Ray, I. (2014). The drinking water disparities framework: on the origins and persistence of inequities in exposure. *American Journal of Public Health*, 104(4), 603–611.
- Blaisdell, J., Turyk, M. E., Almberg, K. S., Jones, R. M., & Stayner, L. T. (2019). Prenatal exposure to nitrate in drinking water and the risk of congenital anomalies. *Environmental Research*, 176, 108553.
- Brender, J. D., Weyer, P. J., Romitti, P. A., Mohanty, B. P., Shinde, M. U., Vuong, A. M., Sharkey, J. R., Dwivedi, D., Horel, S. A., Kantamneni, J., Huber, J. C., Jr., Zheng, Q., Werler, M. M., Kelley, K. E., Griesenbeck, J. S., Zhan, F. B., Langlois, P. H., Suarez, L., Canfield, M. A., & National Birth Defects Prevention Study. (2013). Prenatal nitrate intake from drinking water and selected birth defects in offspring of participants in the National Birth Defects Prevention Study. *Environmental Health Perspectives*, 121(9), 1083–1089.
- Brender, J. D., & Weyer, P. J. (2016). Agricultural compounds in water and the risk of birth defects. *Current Environmental Health Reports*, 3(2), 144–152.
- Brooks, B. W., et al. (2019). Environmental health practice challenges and research needs for safe drinking water. *Environmental Health Perspectives*, 127(8), 085001.
- Butler, L. J., Scammell, M. K., & Benson, E. B. (2016). The Flint, Michigan, water crisis: A case study in regulatory failure and environmental injustice. *Environmental Justice*, 9(4), 93–97.

- California Department of Public Health. (2008, August 7). *Arsenic MCL revision regulations: Regulation text (DPH-04-017; California Code of Regulations, Title 22, Chapter 15)*. California Department of Public Health.
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/documents/arsenic_mcl/DPH-17-04FinalRegText2008-27-08.pdf
- California Environmental Protection Agency. (2022, May). *Designation of disadvantaged communities using CalEnviroScreen 4.0*. Sacramento, CA: CalEPA.
<https://oehha.ca.gov/calenviroscreen/sb535>
- California Natural Resources Agency. (n.d.). *Alluvial boundary of California's Central Valley*.
<https://data.cnra.ca.gov/dataset/alluvial-boundary-of-californias-central-valley>
- California State Water Resources Control Board. (2020, December 2). *Nitrate Project*.
https://www.waterboards.ca.gov/water_issues/programs/nitrate_project/
- California State Water Resources Control Board. (2023, September). Chapter 10: Economically disadvantaged communities. In *Draft staff report: Sacramento/Delta update to the Bay-Delta Plan*. https://www.waterboards.ca.gov/bay_delta/docs/2023/staff-report/ch10-dacs.pdf
- California State Water Resources Control Board. (2025a, June 27). *Residential water treatment devices*.
https://www.waterboards.ca.gov/drinking_water/certlic/device/watertreatmentdevices.html
- California State Water Resources Control Board. (2025b, February 27). *State expands program that has provided free replacement water for over 1,800 nitrate-impacted wells*.
https://www.waterboards.ca.gov/press_room/press_releases/2025/pr20250227-cv-salts-expansion.html
- California State Water Resources Control Board, Division of Drinking Water. (2017, December 29). *California regulations related to drinking water: Titles 17 and 22 California Code of Regulations (Last updated December 29, 2017)*
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/documents/lawbook/dwregulations-2017-12-29.pdf
- California State Water Resources Control Board, Division of Drinking Water. (2025a, April 22). *Nitrates and nitrites in drinking water*.
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/Nitrate.html
- California State Water Resources Control Board, Division of Drinking Water. (2025b, July). *Drinking water program annual compliance report: Calendar year 2024*.
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/Publications.html
- California Water Code § 79505.5; Northern California Water Association. (2023). *Disadvantaged communities in Northern California*.
<https://law.justia.com/codes/california/code-wat/division-26-5/chapter-1/section-79505-5/>
- Centers for Disease Control and Prevention / ATSDR. (2015). *Nitrate/nitrite toxicity: What are U.S. standards and regulations for nitrate exposure?*
https://archive.cdc.gov/www_atsdr_cdc_gov/csem/nitrate-nitrite/standards.html

- Clemmensen, P. J., Schullehner, J., Brix, N., Sigsgaard, T., Stayner, L. T., Kolstad, H. A., & Ramlau-Hansen, C. H. (2023). Prenatal exposure to nitrate in drinking water and adverse health outcomes in the offspring: A review of current epidemiological research. *Current Environmental Health Reports*, 10(3), 250–263.
- Chen, W., Qian, L., Shi, J., & Franklin, M. (2018). Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Medical Research Methodology*, 18, 63.
- Coffman, V. R., Jensen, A. S., Trabjerg, B. B., Pedersen, C. B., Hansen, B., Schullehner, J., Pedersen, M., & Stayner, L. T. (2022). Prenatal exposure to nitrate from drinking water and the risk of preterm birth: A Danish nationwide cohort study. *Environmental Epidemiology*, 6(5), e223.
- Croen, L. A., Todoroff, K., & Shaw, G. M. (2001). Maternal exposure to nitrate from drinking water and diet and risk for neural tube defects. *American Journal of Epidemiology*, 153(4), 325–331.
- Currie, J., Graff Zivin, J., Meckel, K., Neidell, M., & Schlenker, W. (2013). Something in the water: Contaminated drinking water and infant health. *Canadian Journal of Economics/Revue canadienne d'économique*, 46, 791–810.
- De Roos, A. J., Ward, M. H., Lynch, C. F., & Cantor, K. P. (2003). Nitrate in public water supplies and the risk of colon and rectum cancers. *Epidemiology*, 14(6), 640–649.
- DiSalvo, R. W., & Hill, E. L. (2024). Drinking water contaminant concentrations and birth outcomes. *Journal of Policy Analysis and Management*, 43(2), 368–399.
- Ebdrup, N. H., Schullehner, J., Knudsen, U. B., Liew, Z., Thomsen, A. M. L., Lyngsø, J., Bay, B., Arendt, L. H., Clemmensen, P. J., Sigsgaard, T., Hansen, B., & Ramlau-Hansen, C. H. (2022). Drinking water nitrate and risk of pregnancy loss: A nationwide cohort study. *Environmental Health*, 21(1), Article 87.
- Famiglietti, J. S. (2014). The global groundwater crisis. *Nature Climate Change*, 4(11), 945–948.
- Faunt, C. C. (Ed.). (2009). Groundwater availability of the Central Valley aquifer, California (U.S. Geological Survey Professional Paper 1766). U.S. Geological Survey.
- Feinstein, L., Phurisamban, R., Ford, A., Tyler, C., & Crawford, A. (2017). *Drought and equity in California*. Pacific Institute, 80.
- Fernandez-Bou, A. S., Ortiz-Partida, J. P., Dobbin, K. B., Flores-Landeros, H., Bernacchi, L. A., & Medellín-Azuara, J. (2021). Underrepresented, understudied, underserved: Gaps and opportunities for advancing justice in disadvantaged communities. *Environmental Science & Policy*, 122, 92–100.
- Fox, M. P., Lash, T. L., & Greenland, S. (2005). A method to automate probabilistic sensitivity analyses of misclassified binary variables. *International Journal of Epidemiology*, 34(6), 1370–1376.
- Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *The Lancet*, 371(9606), 75–84.

- Harter, T., Lund, J. R., Darby, J., Fogg, G. E., Howitt, R., Jessoe, K. K., Pettygrove, G. S., Quinn, J. F., Viers, J. H., Boyle, D. B., Canada, H. E., DeLaMora, N., Dzurella, K. N., Fryjoff-Hung, A., Hollander, A. D., Honeycutt, K. L., Jenkins, M. W., Jensen, V. B., King, A. M., Kourakos, G., Lopez, E. M., Mayzelle, M. M., McNally, A., Medellín-Azuara, J., & Rosenstock, T. S. (2012). *Addressing nitrate in California's drinking water with a focus on Tulare Lake Basin and Salinas Valley groundwater: Report for the State Water Resources Control Board Report to the Legislature*. Center for Watershed Sciences, University of California, Davis.
- Harville, E. W., Viikari, J., Raitakari, O. T., & Juonala, M. (2018). Pre-pregnancy cardiovascular risk factors and racial disparities in preterm birth. *Paediatric and Perinatal Epidemiology*, 32(4), 280–289.
- Holtby, C. E., Guernsey, J. R., Allen, A. C., VanLeeuwen, J. A., Allen, V. M., & Gordon, R. J. (2014). A population-based case-control study of drinking-water nitrate and congenital anomalies using geographic information systems (GIS) to develop individual-level exposure estimates. *International Journal of Environmental Research and Public Health*, 11(2), 1803–1823.
- Holzman, C., Eyster, J., Kleyn, M., Messer, L. C., Kaufman, J. S., Laraia, B. A., & Elo, I. T. (2009). Maternal weathering and risk of preterm delivery. *American Journal of Public Health*, 99(10), 1864–1871.
- Huang, H., Woodruff, T. J., Baer, R. J., Bangia, K., August, L. M., Jellife-Palowski, L. L., Padula, A. M., & Sirota, M. (2018). Investigation of association between environmental and socioeconomic factors and preterm birth in California. *Environment International*, 121, 1066–1078.
- James-Todd, T., Karumanchi, S. A., & Hibert, E. N. (2014). Race/Ethnicity, educational attainment, and pregnancy complications in the United States. *Paediatric and Perinatal Epidemiology*, 28(4), 346–355.
- Jauniaux, E., Jeremiah, L., Richardson, B., & Rogozińska, E. (2024). Exposure to drinking water pollutants and non-syndromic birth defects: a systematic review and meta-analysis synthesis. *BMJ Open*, 14(11), e084122.
- Jensen, A. S., Coffman, V. R., Schullehner, J., Trabjerg, B. B., Pedersen, C. B., Hansen, B., Olsén, J., Pedersen, M., Stayner, L. T., & Sigsgaard, T. (2023). Prenatal exposure to tap water containing nitrate and the risk of small-for-gestational-age: A nationwide register-based study of Danish births, 1991–2015. *Environment International*, 174, 107883.
- Jones, R. R., Weyer, P. J., DellaValle, C. T., Inoue-Choi, M., Anderson, K. E., Cantor, K. P., Krasner, S., Robien, K., Beane Freeman, L. E., Silverman, D. T., & Ward, M. H. (2016). Nitrate from drinking water and diet and bladder cancer among postmenopausal women in Iowa. *Environmental Health Perspectives*, 124(11), 1751–1758.
- Kardol, L. R., Rhaman, A., Kelty, E., Morgan, S., Russell, D. J., Gebremedhin, A., Bourke, S., & Wyrwoll, C. S. (2026). Prenatal nitrate and nitrite intake in mammals: A scoping review of effects and associations with pregnancy and maternal and child health. *International Journal of Hygiene and Environmental Health*, 271, 114712.

- Knobeloch, L., Salna, B., Hogan, A., Postle, J., & Anderson, H. (2000). Blue babies and nitrate-contaminated well water. *Environmental Health Perspectives*, 108(7), 675.
- Lash, T. L., Fox, M. P., MacLehose, R. F., Maldonado, G., McCandless, L. C., & Greenland, S. (2014). Good practices for quantitative bias analysis. *International Journal of Epidemiology*, 43(6), 1969–1985.
- Levy, Z. F., Jurgens, B. C., Burow, K. R., Voss, S. A., Faulkner, K. E., Arroyo-Lopez, J. A., & Fram, M. S. (2021). Critical aquifer overdraft accelerates degradation of groundwater quality in California's Central Valley during drought. *Geophysical Research Letters*, 48(17), e2021GL094398.
- Lin, L., St Clair, S., Gamble, G. D., Crowther, C. A., Dixon, L., Bloomfield, F. H., & Harding, J. E. (2023). Nitrate contamination in drinking water and adverse reproductive and birth outcomes: a systematic review and meta-analysis. *Scientific Reports*, 13(1), 563.
- Liu, B., Xu, G., Sun, Y., Du, Y., Gao, R., Snetselaar, L. G., Santillan, M. K., & Bao, W. (2019). Association between maternal pre-pregnancy obesity and preterm birth according to maternal age and race or ethnicity: A population-based study. *The Lancet Diabetes & Endocrinology*, 7(9), 707–714.
- Manassaram, D. M., Backer, L. C., & Moll, D. M. (2007). A review of nitrates in drinking water: maternal exposure and adverse reproductive and developmental outcomes. *Ciência & Saúde Coletiva*, 12, 153–163.
- Manuck, T. A. (2017). Racial and ethnic differences in preterm birth: A complex, multifactorial problem. *Seminars in Perinatology*, 41(8), 511–518.
- Mathewson, P. D., Evans, S., Byrnes, T., Joos, A., & Naidenko, O. V. (2020). Health and economic impact of nitrate pollution in drinking water: a Wisconsin case study. *Environmental Monitoring and Assessment*, 192(11), 724.
- Migeot, V., Albouy-Llaty, M., Carles, C., Limousi, F., Strezlec, S., Dupuis, A., & Rabouan, S. (2013). Drinking-water exposure to a mixture of nitrate and low-dose atrazine metabolites and small-for-gestational age (SGA) babies: a historic cohort study. *Environmental Research*, 122, 58–64.
- Nolan, B. T., Ruddy, B. C., Hitt, K. J., & Helsel, D. R. (1997). Risk of Nitrate in groundwaters of the United States a national perspective. *Environmental Science & Technology*, 31(8), 2229–2236.
- Montoya, E. S. (2025). Birth outcome effects of nitrate contamination in drinking water. *Journal of Environmental Economics and Management*, 133, 103197.
- National Center for Health Statistics. (2024). *User guide to the 2022 natality public use file*. Centers for Disease Control and Prevention.
https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/natality/UserGuide2022.pdf
- National Primary Drinking Water Regulations, 40 C.F.R. § 141.62 (2024).
- Pauloo, R. A., Escrivá-Bou, A., Dahlke, H., Fencel, A., Guillon, H., & Fogg, G. E. (2020). Domestic well vulnerability to drought duration and unsustainable groundwater

- management in California's Central Valley. *Environmental Research Letters*, 15(4), 044010.
- Pennino, M. J., Compton, J. E., & Leibowitz, S. G. (2017). Trends in drinking water nitrate violations across the United States. *Environmental Science & Technology*, 51(22), 13450–13460.
- Preda, A., Iliescu, D.-G., Comănescu, A., Zorilă, G.-L., Vladu, I. M., Forțofoiu, M.-C., Țenea-Cojan, T. S., Preda, S.-D., Diaconu, I.-D., Moța, E., Gheorghe, I.-O., & Moța, M. (2023). Gestational diabetes and preterm birth: What do we know? Our experience and mini-review of the literature. *Journal of Clinical Medicine*, 12(14), Article 4572.
- Office of Environmental Health Hazard Assessment. (2018, May). *Public health goals for nitrate and nitrite in drinking water*. <https://oehha.ca.gov/sites/default/files/media/downloads/water/chemicals/phg/nitratephg051118.pdf>
- Office of Environmental Health Hazard Assessment. (2023). *CalEnviroScreen 4.0 report update*. Sacramento, CA: OEHHHA. <https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-40>
- Quinn, J.-A., Munoz, F. M., Gonik, B., Frau, L., Cutland, C., Mallett-Moore, T., Kissou, A., Wittke, F., Das, M., Nunes, T., Pye, S., Watson, W., Alguacil Ramos, A.-M., Cordero, J. F., Huang, W.-T., Kochhar, S., Buttery, J., & Brighton Collaboration Preterm Birth Working Group. (2016). Preterm birth: Case definition & guidelines for data collection, analysis, and presentation of immunisation safety data. *Vaccine*, 34(49), 6047–6056.
- Ransom, K. M., Nolan, B. T., Traum, J. A., Faunt, C. C., Bell, A. M., Gronberg, J. A. M., Wheeler, D. C., Rosecrans, C. Z., Jurgens, B. C., Schwarz, G. E., Belitz, K., Eberts, S. M., Kourakos, G., & Harter, T. (2017). A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA. *Science of the Total Environment*, 601–602, 1160–1172.
- Royal, H., 't Mannetje, A., Hales, S., Douwes, J., Berry, M., & Chambers, T. (2024). Nitrate in drinking water and adverse pregnancy outcomes: A narrative review of epidemiological evidence and proposed biological mechanisms. *PLOS Water*, 3(1), e0000214.
- Safe Drinking Water Act, 42 U.S.C. § 300f et seq. (1974).
- Schaider, L. A., Swetschinski, L., Campbell, C., & Rudel, R. A. (2019). Environmental justice and drinking water quality: are there socioeconomic disparities in nitrate levels in U.S. drinking water? *Environmental Health*, 18(1), 3.
- Semprini, J. (2025). Early prenatal nitrate exposure and birth outcomes: A study of Iowa's public drinking water (1970–1988). *PLOS Water*, 4(6), e0000329.
- Sherris, A. R., Baiocchi, M., Fendorf, S., Luby, S. P., Yang, W., & Shaw, G. M. (2021). Nitrate in drinking water during pregnancy and spontaneous preterm birth: a retrospective within-mother analysis in California. *Environmental Health Perspectives*, 129(5), 057001.

- Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation—a global inventory. *Hydrology and Earth System Sciences*, 14(10), 1863–1880.
- Snowden, J. M., Rose, S., & Mortimer, K. M. (2011). Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *American Journal of Epidemiology*, 173(7), 731–738.
- State Water Resources Control Board. (2023, September). *Chapter 10: Economically disadvantaged communities (Draft staff report: Sacramento/Delta update to the Bay-Delta Plan)*. https://www.waterboards.ca.gov/bay_delta/docs/2023/staff-report/ch10-dacs.pdf
- Stayner, L., et al. (2016). Atrazine and nitrate in drinking water and the risk of preterm delivery and low birth weight in four midwestern states. *Environmental Research*, 152, 294–303.
- Sum, S. (2024). Impact of Droughts on Served Drinking Water Disparities in California, 2007–2020. *American Journal of Public Health*, 114(9), 935–945.
- Swain, D. L., Langenbrunner, B., Neelin, J. D., & Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. *Nature Climate Change*, 8(5), 427–433.
- Tariqi, A. Q., & Naughton, C. C. (2021). Water, health, and environmental justice in California: geospatial analysis of nitrate contamination and thyroid cancer. *Environmental Engineering Science*, 38(5), 377–388.
- Temkin, A., Evans, S., Manidis, T., Campbell, C., & Naidenko, O. V. (2019). Exposure-based assessment and economic valuation of adverse birth outcomes and cancer risk due to nitrate in United States drinking water. *Environmental Research*, 176, 108442.
- U.S. Environmental Protection Agency. (n.d.). *SDWIS metadata: PRIMARY_SOURCE_CODE*. https://enviro.epa.gov/enviro/EF_METADATA_HTML.sdwis_page?p_column_name=PRIMARY_SOURCE_CODE
- U.S. Environmental Protection Agency. (1997). *EPA Method 300.1 (Revision 1.0): Determination of inorganic anions in drinking water by ion chromatography*. <https://www.epa.gov/sites/default/files/2015-06/documents/epa-300.1.pdf>
- U.S. Environmental Protection Agency. (2006, April). *Point-of-use or point-of-entry treatment options for small drinking water systems* (EPA 815-R-06-010). https://www.epa.gov/sites/default/files/2015-09/documents/guide_smallsystems_pou-poe_june6-2006.pdf
- U.S. Environmental Protection Agency. (2008, January). *Drinking water health advisory for 2,4-dinitrotoluene and 2,6-dinitrotoluene* (EPA 822-R-08-010). https://www.epa.gov/sites/default/files/2014-09/documents/drinking_water_health_advisory_for_24_and_26_dinitrotoluene.pdf
- U.S. Environmental Protection Agency. (2011, July). *National characteristics of drinking water systems serving 10,000 or fewer people*. Washington, DC: U.S. EPA. EPA 816-R-10-022. <https://www.epa.gov/sites/default/files/2015-12/documents/epa816r10022.pdf>
- U.S. Environmental Protection Agency. (2020, November 30). *Use of total nitrate and nitrite analysis for compliance determinations with the nitrate maximum contaminant level – 40*

- CFR §141.23 (Memorandum)*. https://www.epa.gov/sites/default/files/2021-01/documents/wsg_213_nitrate_wsg_11-30-2020_signed_508-compliantfinal.pdf
- U.S. Environmental Protection Agency. (2022, October). *DWSRF disadvantaged community definitions report*. https://www.epa.gov/system/files/documents/2022-10/DWSRF%20DAC%20Definitions%20Report_October%202022%20Updates_FINAL_508.pdf
- U.S. Environmental Protection Agency. (2025a, April 25). *SDWIS Federal Reporting Services data dictionary (Primary Source Code definitions)*. <https://www.epa.gov/DWdata/sdwis-federal-reporting-services-data-dictionary>
- U.S. Environmental Protection Agency. (2025b, June 24). *Learn about the unregulated contaminant monitoring rule*. <https://www.epa.gov/dwucmr/learn-about-unregulated-contaminant-monitoring-rule>
- U.S. Environmental Protection Agency. (2025c, July 15). *Regulatory Determination 2 support documents for 2,4-dinitrotoluene and 2,6-dinitrotoluene*. <https://www.epa.gov/ccl/regulatory-determination-2-support-documents-24-dinitrotoluene-and-26-dinitrotoluene>
- U.S. Environmental Protection Agency. (2025d, September 10). *Second unregulated contaminant monitoring rule*. <https://www.epa.gov/dwucmr/second-unregulated-contaminant-monitoring-rule>
- U.S. Environmental Protection Agency. (2025e, December 1). *National primary drinking water regulations*. <https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations>
- U.S. Environmental Protection Agency. (2026, March 16). *Information about public water systems*. <https://www.epa.gov/dwreginfo/information-about-public-water-systems>
- U.S. Geological Survey. (n.d.). *Central Valley Hydrologic Model (CVHM): Numerical model and supporting datasets*. California Water Science Center. <https://ca.water.usgs.gov/projects/central-valley/central-valley-hydrologic-model.html>
- U.S. Geological Survey. (2009). *California's Central Valley groundwater study: A powerful new tool to assess water resources in California's Central Valley (Fact Sheet 2009–3057)*. <https://pubs.usgs.gov/fs/2009/3057>
- U.S. Geological Survey. (2021, September 1). *Increased pumping in California's Central Valley during drought worsens groundwater quality*. <https://www.usgs.gov/news/state-news-release/increased-pumping-californias-central-valley-during-drought-worsens>
- U.S. Geological Survey. (2021). *Community water system service area boundaries and groundwater quality in California's Central Valley*. Reston, VA: USGS. <https://www.usgs.gov/centers/california-water-science-center/central-valley-programs>
- VanDerslice, J. (2011). Drinking water infrastructure and environmental disparities: Evidence and methodological considerations. *American Journal of Public Health*, 101(Suppl 1), S109–S114.

- von Ehrenstein, O. S., Wilhelm, M., & Ritz, B. (2013). Maternal occupation and term low birth weight in a predominantly Latina population in Los Angeles, California. *Journal of Occupational and Environmental Medicine*, 55(9), 1046–1051.
- Vuong, A. M., Shinde, M. U., Brender, J. D., Shipp, E. M., Huber, J. C., Jr., Sharkey, J. R., McDonald, T. J., Werler, M. M., Kelley, K. E., Griesenbeck, J. S., Langlois, P. H., Canfield, M. A., & National Birth Defects Prevention Study Investigators. (2016). Prenatal exposure to nitrosatable drugs, dietary intake of nitrites, and preterm birth. *American Journal of Epidemiology*, 183(7), 634–642.
- Wada, Y., Van Beek, L. P., Van Kempen, C. M., Reckman, J. W., Vasak, S., & Bierkens, M. F. (2010). Global depletion of groundwater resources. *Geophysical Research Letters*, 37(20).
- Ward, M. H., Theo, M. D., Levallois, P., Brender, J., Gulis, G., Nolan, B. T., & VanDerslice, J. (2005). Workgroup report: drinking-water nitrate and health—recent findings and research needs. *Environmental Health Perspectives*, 113(11), 1607.
- Ward, M. H., Kilfoy, B. A., Weyer, P. J., Anderson, K. E., Folsom, A. R., & Cerhan, J. R. (2010). Nitrate intake and the risk of thyroid cancer and thyroid disease. *Epidemiology*, 21(3), 389–395.
- Ward, M. H., Jones, R. R., Brender, J. D., de Kok, T. M., Weyer, P. J., Nolan, B. T., Villanueva, C. M., & van Breda, S. G. J. (2018). Drinking water nitrate and human health: An updated review. *International Journal of Environmental Research and Public Health*, 15(7), Article 1557.
- Weyer, P. J., Cerhan, J. R., Kross, B. C., Hallberg, G. R., Kantamneni, J., Breuer, G., Jones, M. P., Zheng, W., & Lynch, C. F. (2001). Municipal drinking water nitrate level and cancer risk in older women: The Iowa Women’s Health Study. *Epidemiology*, 12(3), 327–338.
- Wilhelm, M., Ghosh, J. K., Su, J., Cockburn, M., Jerrett, M., & Ritz, B. (2011). Traffic-related air toxics and preterm birth: A population-based case-control study in Los Angeles County, California. *Environmental Health*, 10, 89.
- Williams, S. A., Megdal, S. B., Zuniga-Teran, A. A., Quanrud, D. M., & Christopherson, G. (2024). Equity assessment of groundwater vulnerability and risk in drinking water supplies in arid regions. *Water*, 16(23), 3520.
- World Health Organization. (n.d.). *Low birthweight (% newborns who weigh <2.5 kg)*. Global Health Observatory Indicator Metadata Registry. <https://www.who.int/data/gho/indicator-metadata-registry/imr-details/low-birthweight-%28-newborns-who-weigh-2.5kg%29>
- World Health Organization. (2014). *Global nutrition targets 2025: Low birth weight policy brief*. World Health Organization.
- World Health Organization. (2016). *Nitrate and nitrite in drinking-water*. World Health Organization. <https://www.who.int/docs/default-source/wash-documents/wash-chemicals/nitrate-nitrite-background-document.pdf>
- World Health Organization. (2023, May 10). *Preterm birth*. <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>

- World Health Organization, United Nations Children's Fund, & United Nations Population Fund. (2023, May 9). *Born too soon: Decade of action on preterm birth*. World Health Organization. <https://www.who.int/publications/i/item/9789240073890>
- Yelland, L. N., Salter, A. B., & Ryan, P. (2011). Performance of the modified Poisson regression approach for estimating relative risks from clustered prospective data. *American Journal of Epidemiology*, 174(8), 984–992.
- Young, H. A., Kolivras, K. N., Krometis, L. A. H., Marcillo, C. E., & Gohlke, J. M. (2023). Examining the association between safe drinking water act violations and adverse birth outcomes in Virginia. *Environmental Research*, 218, 114977.
- Zou, G. (2004). A modified Poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*, 159(7), 702–706.

APPENDIX A: Chapter 2 Supplemental Tables and Figures

Appendix Table 2.A1: System-Year Nitrate Monitoring Cadence Among System-Years Contributing to the Analytic Birth Sample

Panel A. Overall and subgroup monitoring cadence

Group	System-years	Sampled months			≤ 2 sampled months (%)	Sampled quarters			≤ 2 sampled quarters (%)
		p25	P50	p75		p25	P50	p75	
Overall	12,178	1	2	5	58.1	1	2	4	61.1
Groundwater r-only	8,096	1	2	4	63.3	1	1	4	66.2
Not groundwater	4,046	1	3	9	47.4	1	2	4	50.8
Central Valley	3,786	1	2	4	57.4	1	2	4	60.5
Not Central Valley	8,392	1	2	5	58.4	1	2	4	61.4
DAC	6,555	1	2	4	59.7	1	2	4	62.8
Non-DAC	5,623	1	2	5	56.2	1	2	4	59.2
SDAC	3,728	1	2	4	61.7	1	1	4	65.2
Non-SDAC	8,450	1	2	5	56.5	1	2	4	59.3

Panel B. Monitoring cadence by system size

Group	System -years	Sampled months			≤2 sampled months (%)	Sampled quarters			≤2 sampled quarters (%)
		p25	p50	p75		p25	p50	p75	
<3,300 served	6,666	1	1	3	73.0	1	1	3	74.9
3,300–9,999 served	1,541	1	2	5	55.2	1	2	4	59.8
10,000–99,999 served	3,091	1	4	8	37.9	1	3	4	42.4
≥100,000 served	822	4	11	12	15.7	3	4	4	20.2

Note. The table reports nitrate monitoring cadence for system-years with at least one birth in the analytic sample. Sampled months count the number of distinct calendar months in a system-year with at least one nitrate sample, and sampled quarters count the number of distinct calendar quarters in a system-year with at least one nitrate sample. Percentages with two or fewer sampled months or quarters summarize the share of system-years with relatively sparse temporal monitoring. Panel A reports overall and contextual subgroup patterns, and Panel B reports patterns by system size. Subgroup totals may differ from the overall total where source type or system-size classification is missing. DAC refers to disadvantaged community status; SDAC refers to severely disadvantaged community status.

Appendix Table 2.A2: Supplemental Birth-Outcome Descriptive Statistics in the Analytic Exposure Sample

Variable	Definition or unit	N	%	Mean (SD)	Median
Very preterm birth	<32 weeks	38,111	1.1		
Very low birthweight	<1,500 g	63,309	1.8		
Birthweight	grams	3,538,347	—	3315.02 (536.48)	3330
Gestational age	days	3,539,019	—	274.2 (14.0)	275

Note. The table reports supplemental descriptive statistics for secondary birth outcomes and continuous outcome measures in the broader analytic exposure sample. Very preterm birth is defined as birth before 32 completed weeks of gestation, and very low birthweight is defined as birthweight below 1,500 grams. Percentages are calculated relative to the analytic exposure

sample where the corresponding outcome is defined. Birthweight and gestational age use cleaned outcome variables.

Appendix Table 2.A3: Full Adjusted Birth-Outcome Models by Pregnancy-Average Nitrate Exposure Specification

Panel A. Preterm birth

Variable	Continuous nitrate	Categorical nitrate	Threshold nitrate
Nitrate exposure			
Pregnancy-average nitrate, per 1 mg/L as N	1.019 [1.012, 1.026]***	—	—
1-<2.5 mg/L as N vs <1	—	1.050 [1.007, 1.094]**	—
2.5-<5 mg/L as N vs <1	—	1.104 [1.065, 1.143]***	—
5-<10 mg/L as N vs <1	—	1.130 [1.084, 1.177]***	—
>=10 mg/L as N vs <1	—	1.182 [1.104, 1.265]***	—
>=10 vs <10 mg/L as N	—	—	1.091 [1.021, 1.167]**
Maternal and infant covariates			
Maternal age: 20-24	0.864 [0.847, 0.881]***	0.863 [0.846, 0.881]***	0.863 [0.846, 0.880]***
Maternal age: 25-34	0.974 [0.954, 0.996]**	0.974 [0.953, 0.995]**	0.973 [0.952, 0.995]**
Maternal age: >=35	1.385 [1.343, 1.428]***	1.384 [1.341, 1.429]***	1.384 [1.342, 1.427]***
College degree or higher	0.781 [0.769, 0.793]***	0.781 [0.769, 0.793]***	0.782 [0.769, 0.794]***
Non-Hispanic Black	1.411 [1.361, 1.464]***	1.409 [1.360, 1.459]***	1.407 [1.354, 1.461]***
Hispanic	1.128 [1.087, 1.171]***	1.127 [1.087, 1.168]***	1.135 [1.095, 1.176]***
Smoking during pregnancy	1.319 [1.286, 1.353]***	1.318 [1.285, 1.351]***	1.314 [1.280, 1.349]***
First live birth	1.119 [1.093, 1.146]***	1.119 [1.092, 1.146]***	1.119 [1.093, 1.145]***
BMI: Normal (18.5-24.9)	0.836 [0.815, 0.857]***	0.836 [0.816, 0.857]***	0.834 [0.813, 0.855]***

Variable	Continuous nitrate	Categorical nitrate	Threshold nitrate
BMI: Overweight (25.0-29.9)	0.875 [0.848, 0.903]***	0.875 [0.848, 0.903]***	0.873 [0.846, 0.900]***
BMI: Obese (≥ 30)	0.993 [0.959, 1.028]	0.994 [0.960, 1.029]	0.991 [0.957, 1.025]
History of previous preterm birth	4.154 [3.958, 4.359]***	4.156 [3.960, 4.362]***	4.136 [3.938, 4.344]***
Female infant	0.839 [0.832, 0.846]***	0.839 [0.832, 0.846]***	0.839 [0.832, 0.846]***
Prenatal payer: No prenatal care	5.903 [5.223, 6.672]***	5.902 [5.221, 6.673]***	5.832 [5.175, 6.572]***
Prenatal payer: Public insurance	1.660 [1.512, 1.822]***	1.656 [1.508, 1.818]***	1.645 [1.505, 1.798]***
Prenatal payer: Private insurance	1.421 [1.297, 1.557]***	1.420 [1.295, 1.556]***	1.409 [1.291, 1.538]***
Prenatal payer: Other or unknown	1.674 [1.502, 1.865]***	1.670 [1.500, 1.860]***	1.649 [1.481, 1.837]***
Urban/higher-density PWS	1.093 [1.036, 1.154]***	1.081 [1.027, 1.138]***	1.111 [1.051, 1.174]***

Panel B. Low birthweight

Variable	Continuous nitrate	Categorical nitrate	Threshold nitrate
Nitrate exposure			
Pregnancy-average nitrate, per 1 mg/L as N	1.014 [1.006, 1.021]***	—	—
1-<2.5 mg/L as N vs <1	—	1.039 [0.978, 1.104]	—
2.5-<5 mg/L as N vs <1	—	1.076 [1.012, 1.145]**	—
5-<10 mg/L as N vs <1	—	1.093 [1.026, 1.165]***	—
≥ 10 mg/L as N vs <1	—	1.062 [0.982, 1.148]	—
≥ 10 vs <10 mg/L as N	—	—	1.001 [0.946, 1.058]
Maternal and infant covariates			
Maternal age: 20-24	0.977 [0.945, 1.010]	0.977 [0.945, 1.009]	0.976 [0.944, 1.009]
Maternal age: 25-34	1.129 [1.091, 1.168]***	1.128 [1.091, 1.167]***	1.128 [1.090, 1.167]***
Maternal age: ≥ 35	1.577 [1.523, 1.633]***	1.576 [1.522, 1.633]***	1.576 [1.522, 1.632]***

Variable	Continuous nitrate	Categorical nitrate	Threshold nitrate
College degree or higher	0.793 [0.773, 0.813]***	0.793 [0.773, 0.813]***	0.793 [0.774, 0.813]***
Non-Hispanic Black	1.815 [1.747, 1.887]***	1.813 [1.745, 1.883]***	1.811 [1.742, 1.882]***
Hispanic	1.062 [1.028, 1.097]***	1.061 [1.027, 1.096]***	1.067 [1.033, 1.102]***
Smoking during pregnancy	1.433 [1.358, 1.512]***	1.432 [1.355, 1.512]***	1.429 [1.355, 1.507]***
First live birth	1.642 [1.611, 1.672]***	1.641 [1.610, 1.673]***	1.641 [1.611, 1.672]***
BMI: Normal (18.5-24.9)	0.675 [0.659, 0.691]***	0.675 [0.659, 0.691]***	0.673 [0.658, 0.690]***
BMI: Overweight (25.0-29.9)	0.636 [0.616, 0.657]***	0.636 [0.617, 0.657]***	0.635 [0.615, 0.655]***
BMI: Obese (≥ 30)	0.664 [0.644, 0.685]***	0.664 [0.645, 0.685]***	0.663 [0.643, 0.684]***
History of previous preterm birth	4.136 [3.931, 4.352]***	4.138 [3.932, 4.354]***	4.124 [3.917, 4.341]***
Female infant	1.110 [1.098, 1.122]***	1.110 [1.098, 1.122]***	1.110 [1.098, 1.122]***
Prenatal payer: No prenatal care	5.002 [4.438, 5.637]***	4.999 [4.435, 5.635]***	4.954 [4.401, 5.576]***
Prenatal payer: Public insurance	1.733 [1.572, 1.910]***	1.730 [1.568, 1.907]***	1.720 [1.564, 1.892]***
Prenatal payer: Private insurance	1.505 [1.369, 1.655]***	1.504 [1.367, 1.654]***	1.494 [1.362, 1.639]***
Prenatal payer: Other or unknown	1.795 [1.592, 2.024]***	1.792 [1.586, 2.024]***	1.773 [1.579, 1.991]***
Urban/higher-density PWS	1.145 [1.067, 1.229]***	1.135 [1.057, 1.219]***	1.158 [1.076, 1.246]***

Note. Values are odds ratios with 95% confidence intervals. Panel A reports adjusted logistic regression models for preterm birth. Panel B reports adjusted generalized linear models with binomial family and logit link for low birthweight. Within each panel, the three model columns differ only by pregnancy-average nitrate exposure specification; all other covariates are the same. Categorical nitrate models use <1 mg/L as N as the reference category, and threshold models compare ≥ 10 mg/L as N with <10 mg/L as N. All models adjust for maternal age, education, race and ethnicity, smoking during pregnancy, first live birth, pre-pregnancy body mass index, previous preterm birth, infant sex, prenatal payer, public water system urban-rural context, and birth year-quarter fixed effects. Birth year-quarter fixed effects are included but not

displayed. Standard errors are clustered at the public water system level. Omitted reference categories are maternal age <20, no college degree, non-Hispanic White or other race and ethnicity, no smoking during pregnancy, not first live birth, underweight BMI, no previous preterm birth, male infant, self-pay prenatal care payer, and rural or low-density public water system. OR refers to odds ratio; CI refers to confidence interval; PWS refers to public water system. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 2.A4: Categorical Nitrate Interactions by Maternal Race and Ethnicity

Outcome	Nitrate category	Estimate type and group	OR/ROR [95% CI]	<i>p</i> -value
Preterm birth	1-<2.5 mg/L as N	OR: NH White/Other	1.040 [0.994, 1.087]	0.092*
Preterm birth	1-<2.5 mg/L as N	ROR: NH Black	0.948 [0.869, 1.034]	0.226
Preterm birth	1-<2.5 mg/L as N	ROR: Hispanic	1.038 [0.978, 1.101]	0.218
Preterm birth	2.5-<5 mg/L as N	OR: NH White/Other	1.091 [1.031, 1.156]	0.003***
Preterm birth	2.5-<5 mg/L as N	ROR: NH Black	0.953 [0.869, 1.044]	0.298
Preterm birth	2.5-<5 mg/L as N	ROR: Hispanic	1.039 [0.949, 1.137]	0.408
Preterm birth	5-<10 mg/L as N	OR: NH White/Other	1.115 [1.048, 1.185]	<0.001***
Preterm birth	5-<10 mg/L as N	ROR: NH Black	1.010 [0.932, 1.094]	0.808
Preterm birth	5-<10 mg/L as N	ROR: Hispanic	1.037 [0.964, 1.115]	0.331
Preterm birth	≥10 mg/L as N	OR: NH White/Other	1.109 [0.987, 1.246]	0.083*
Preterm birth	≥10 mg/L as N	ROR: NH Black	1.327 [1.089, 1.617]	0.005***
Preterm birth	≥10 mg/L as N	ROR: Hispanic	1.097 [0.944, 1.276]	0.228
Low birthweight	1-<2.5 mg/L as N	OR: NH White/Other	1.020 [0.923, 1.128]	0.693
Low birthweight	1-<2.5 mg/L as N	ROR: NH Black	0.974 [0.846, 1.121]	0.711
Low birthweight	1-<2.5 mg/L as N	ROR: Hispanic	1.061 [0.959, 1.173]	0.254
Low birthweight	2.5-<5 mg/L as N	OR: NH White/Other	1.065 [0.947, 1.197]	0.296
Low birthweight	2.5-<5 mg/L as N	ROR: NH Black	0.970 [0.837, 1.125]	0.686

Outcome	Nitrate category	Estimate type and group	OR/ROR [95% CI]	p-value
Low birthweight	2.5-<5 mg/L as N	ROR: Hispanic	1.044 [0.920, 1.185]	0.505
Low birthweight	5-<10 mg/L as N	OR: NH White/Other	1.088 [0.970, 1.220]	0.150
Low birthweight	5-<10 mg/L as N	ROR: NH Black	1.026 [0.887, 1.186]	0.731
Low birthweight	5-<10 mg/L as N	ROR: Hispanic	1.024 [0.915, 1.146]	0.678
Low birthweight	≥10 mg/L as N	OR: NH White/Other	0.978 [0.852, 1.123]	0.752
Low birthweight	≥10 mg/L as N	ROR: NH Black	1.228 [1.036, 1.455]	0.018**
Low birthweight	≥10 mg/L as N	ROR: Hispanic	1.139 [0.976, 1.330]	0.098*

Note. OR refers to odds ratio, ROR refers to ratio of odds ratios, CI refers to confidence interval, NH refers to non-Hispanic, and PWS refers to public water system. Rows labeled OR show category-specific associations for non-Hispanic White/Other mothers relative to the <1 mg/L as N nitrate reference category. Rows labeled ROR compare the category-specific association for non-Hispanic Black or Hispanic mothers with the corresponding association for non-Hispanic White/Other mothers within the same nitrate category. Models include the same adjustment set, birth year-quarter fixed effects, and PWS-clustered standard errors as the main race and ethnicity heterogeneity models. Low birthweight models were estimated using a generalized linear model with binomial family and logit link for numerical stability under high-dimensional fixed effects. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 2.A5: Sensitivity Models for Very Preterm Birth Using Pregnancy-Average Nitrate Exposure

Specification	Contrast	OR	95% CI	p-value
Continuous	per 1 mg/L as N	1.029	[1.019, 1.039]	<0.001***
Binary	≥10 vs <10 mg/L as N	1.102	[0.922, 1.317]	0.285
Categorical	1-<2.5 vs <1 mg/L as N	1.073	[1.000, 1.151]	0.048**
Categorical	2.5-<5 vs <1 mg/L as N	1.147	[1.086, 1.212]	<0.001***
Categorical	5-<10 vs <1 mg/L as N	1.224	[1.146, 1.308]	<0.001***
Categorical	≥10 vs <1 mg/L as N	1.240	[1.035, 1.485]	0.020**

Note. Odds ratios and 95% confidence intervals are from adjusted models of very preterm birth, defined as birth before 32 completed weeks of gestation. Continuous exposure is parameterized per 1 mg/L as N. Categorical models use <1 mg/L as N as the reference category. Binary models compare nitrate concentrations ≥10 mg/L as N with <10 mg/L as N. Models use the same adjustment set, birth year-quarter fixed effects, and public water system-clustered standard errors

as the main pregnancy-average nitrate models. OR refers to odds ratio; CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 2.A6: Groundwater-Only Heterogeneity in Adjusted Nitrate and Birth Outcome Associations

Outcome	Exposure	Non-groundwater OR [95% CI]	Groundwater-only OR [95% CI]	ROR [95% CI]	<i>p</i> -value
Preterm birth	≥ 10 vs < 10 mg/L as N	1.091 [1.009, 1.180]	1.108 [0.961, 1.278]	1.016 [0.864, 1.194]	0.852
Preterm birth	Per 5 mg/L as N	1.099 [1.053, 1.148]	1.105 [1.058, 1.155]	1.005 [0.946, 1.069]	0.863
Low birthweight	≥ 10 vs < 10 mg/L as N	0.982 [0.927, 1.040]	1.021 [0.798, 1.305]	1.040 [0.808, 1.338]	0.763
Low birthweight	Per 5 mg/L as N	1.054 [1.008, 1.102]	1.077 [1.027, 1.130]	1.022 [0.958, 1.091]	0.509

Note. Odds ratios and 95% confidence intervals are from interaction models comparing births served by groundwater-only community water systems with births served by non-groundwater systems. Binary exposure compares nitrate concentrations ≥ 10 mg/L as N with < 10 mg/L as N. Continuous exposure is scaled per 5 mg/L as N. ROR denotes the ratio of odds ratios comparing the groundwater-only association with the non-groundwater association, and the *p*-value column corresponds to the ROR. Models adjust for the same covariates, birth year-quarter fixed effects, and public water system-clustered standard errors as the main models. OR refers to odds ratio; CI refers to confidence interval; ROR refers to ratio of odds ratios.

Appendix Table 2.A7: Central Valley-Only Sensitivity Models for Pregnancy-Average Nitrate Exposure

Outcome	Exposure specification	Contrast	OR [95% CI]	<i>p</i> -value
Preterm birth	Binary	≥ 10 vs < 10 mg/L as N	1.114 [0.951, 1.303]	0.180
Preterm birth	Continuous	Per 1 mg/L as N	1.021 [1.008, 1.035]	0.002***
Low birthweight	Binary	≥ 10 vs < 10 mg/L as N	0.866 [0.630, 1.191]	0.377
Low birthweight	Continuous	Per 1 mg/L as N	1.019 [1.007, 1.031]	0.001***

Note. Odds ratios and 95% confidence intervals are from adjusted models estimated within the Central Valley subsample. Continuous exposure is parameterized per 1 mg/L as N. Binary exposure compares nitrate concentrations ≥ 10 mg/L as N with < 10 mg/L as N. Models use the same covariate set, birth year-quarter fixed effects, and public water system-clustered standard errors as the main pregnancy-average nitrate models. OR refers to odds ratio; CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 2.A8: Quintile and Quartile Sensitivity Models for Nitrate Exposure

Panel A. Quintile cutpoints and within-quintile exposure distribution

Quintile	N	Minimum	p25	Median	p75	Maximum
Q1 (lowest)	707,964	0.000	0.516	0.883	1.256	1.643
Q2	707,869	1.643	2.010	2.285	2.471	2.668
Q3	708,123	2.668	2.904	3.143	3.310	3.456
Q4	707,263	3.456	3.724	3.974	4.340	4.791
Q5 (highest)	707,800	4.792	5.309	6.069	6.944	67.591

Panel B. Quintile sensitivity models

Outcome	Exposure contrast	OR [95% CI]	p-value
Preterm birth	Q2 vs Q1	1.038 [0.995, 1.082]	0.086*
Preterm birth	Q3 vs Q1	1.080 [1.034, 1.127]	<0.001***
Preterm birth	Q4 vs Q1	1.136 [1.065, 1.213]	<0.001***
Preterm birth	Q5 vs Q1	1.115 [1.072, 1.160]	<0.001***
Low birthweight	Q2 vs Q1	1.023 [0.974, 1.075]	0.365
Low birthweight	Q3 vs Q1	1.057 [1.004, 1.111]	0.033**
Low birthweight	Q4 vs Q1	1.101 [1.035, 1.170]	0.002***
Low birthweight	Q5 vs Q1	1.078 [1.026, 1.132]	0.003***

Panel C. Quartile cutpoints and within-quartile exposure distribution

Quartile	N	Minimum	p25	Median	p75	Maximum
Q1 (lowest)	894,951	0.000	0.573	1.057	1.566	2.010
Q2	875,775	2.010	2.340	2.590	2.868	3.143
Q3	883,625	3.143	3.344	3.591	3.924	4.340
Q4 (highest)	884,668	4.340	4.910	5.709	6.683	67.591

Panel D. Quartile sensitivity models

Outcome	Exposure contrast	OR [95% CI]	p-value
Preterm birth	Q2 vs Q1	1.057 [1.023, 1.092]	0.001***
Preterm birth	Q3 vs Q1	1.114 [1.055, 1.176]	<0.001***
Preterm birth	Q4 vs Q1	1.111 [1.072, 1.152]	<0.001***
Low birthweight	Q2 vs Q1	1.034 [0.992, 1.077]	0.111
Low birthweight	Q3 vs Q1	1.093 [1.043, 1.146]	<0.001***
Low birthweight	Q4 vs Q1	1.074 [1.029, 1.121]	0.001***

Note. Panels A and C report pregnancy-average nitrate distributions in mg/L as N in the singleton analytic exposure sample. Quintile cutpoints are 1.643, 2.668, 3.456, and 4.791 mg/L as N; quartile cutpoints are 2.010, 3.143, and 4.340 mg/L as N. Panels B and D report adjusted ORs from complete-case singleton regression models using the same covariates, birth year-quarter fixed effects, and public water system-clustered standard errors as the main Chapter 2 models. Q1 is the reference category. OR refers to odds ratio; CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 2.A9: Sensitivity Models Using Pregnancy-Average Maximum Sampling-Location Nitrate Exposure

Outcome	Specification	Contrast	OR [95% CI]	p-value
Preterm birth	Continuous	Per 1 mg/L as N	1.010 [1.006, 1.015]	<0.001***
Preterm birth	Categorical	1-<2.5 mg/L as N vs <1 mg/L as N	1.021 [0.986, 1.057]	0.245
Preterm birth	Categorical	2.5-<5 mg/L as N vs <1 mg/L as N	1.049 [1.013, 1.086]	0.007***
Preterm birth	Categorical	5-<10 mg/L as N vs <1 mg/L as N	1.124 [1.085, 1.164]	<0.001***
Preterm birth	Categorical	≥10 mg/L as N vs <1 mg/L as N	1.151 [1.084, 1.223]	<0.001***
Preterm birth	Threshold	≥10 mg/L as N vs <10 mg/L as N	1.071 [1.013, 1.132]	0.016**
Low birthweight	Continuous	Per 1 mg/L as N	1.006 [1.003, 1.010]	<0.001***
Low birthweight	Categorical	1-<2.5 mg/L as N vs <1 mg/L as N	1.044 [0.999, 1.091]	0.055*
Low birthweight	Categorical	2.5-<5 mg/L as N vs <1 mg/L as N	1.063 [1.020, 1.109]	0.004***

Outcome	Specification	Contrast	OR [95% CI]	p-value
Low birthweight	Categorical	5-<10 mg/L as N vs <1 mg/L as N	1.123 [1.082, 1.165]	<0.001***
Low birthweight	Categorical	≥10 mg/L as N vs <1 mg/L as N	1.116 [1.071, 1.163]	<0.001***
Low birthweight	Threshold	≥10 mg/L as N vs <10 mg/L as N	1.031 [1.005, 1.059]	0.019**

Note. Odds ratios and 95% confidence intervals are from adjusted models using pregnancy-average maximum sampling-location nitrate exposure. For each system-month, this measure uses the highest sampling-location mean observed within the system-month and then applies the same gestational averaging approach as the main exposure measure. Continuous exposure is per 1 mg/L as N; categorical exposure uses <1 mg/L as N as the reference group; threshold exposure compares ≥10 with <10 mg/L as N. Models use the same covariates, birth year-quarter fixed effects, and public water system-clustered standard errors as the main models. Low-birthweight models use a generalized linear model with binomial family and logit link. This measure is an upper-bound sensitivity check rather than a direct household-level exposure measure. OR refers to odds ratio; CI refers to confidence interval. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 2.A10: Matched-Sample Nitrate Estimates Before and After Co-Pollutant Adjustment

Outcome	Co-pollutant	Matched-sample N	Nitrate coefficient before	Nitrate coefficient after	OR before	OR after	Change in coefficient
Preterm birth	Arsenic	2,362,690	0.0117	0.0111	1.0118	1.0112	-0.0006
Preterm birth	1,2,3-TCP	870,489	-0.0024	-0.0027	0.9976	0.9973	-0.0003
Preterm birth	DBCP	394,212	-0.0046	-0.0057	0.9954	0.9943	-0.0011
Low birthweight	Arsenic	2,362,714	0.0068	0.0064	1.0068	1.0064	-0.0004
Low birthweight	1,2,3-TCP	870,500	0.0011	0.0008	1.0011	1.0008	-0.0003
Low birthweight	DBCP	394,212	0.0019	0.0027	1.0019	1.0027	+0.0008

Note. The table reports matched-sample co-pollutant sensitivity models for preterm birth and low birthweight. Each row compares the nitrate estimate from a nitrate-only model with the nitrate estimate from a model that adds the listed co-pollutant as a continuous covariate, estimated on

the same overlap sample. The nitrate coefficient is scaled per 1 mg/L as N, and odds ratios are calculated by exponentiating the corresponding coefficient. Arsenic models provide the strongest comparison because they use the largest overlap sample. The 1,2,3-TCP and DBCP models use smaller subsets and are interpreted as supplementary checks. For preterm birth, the nitrate estimate is already negative in the 1,2,3-TCP and DBCP overlap samples before co-pollutant adjustment, indicating that those differences reflect the restricted overlap samples rather than the co-pollutant adjustment itself. OR refers to odds ratio.

Chapter 3: Return to Compliance under the SDWA: Duration and Environmental Justice in California Community Water Systems

3.1 Chapter Overview

Chapter 2 examined unequal protection through differences in nitrate exposure and associated birth outcomes. This chapter shifts attention from exposure burden to compliance recovery by asking whether some community water systems remain out of compliance longer than others once violations occur. This shift moves the analysis from the distribution of health-relevant risk to the governance processes that shape how quickly regulatory protection is restored.

Return to compliance is a central measure of drinking-water protection because violations remain consequential until they are corrected, documented, and formally closed. Some violations reflect direct contaminant exceedances, while others involve monitoring, reporting, treatment technique, or public-notification requirements. In each case, longer unresolved spells indicate a longer period in which the system has not yet returned to full regulatory compliance. Although prior research shows that enforcement and oversight can improve compliance under environmental laws, less is known about what shapes the duration of noncompliance under the SDWA, especially among community water systems facing uneven technical, managerial, financial, and administrative capacity.

This chapter examines whether return-to-compliance duration varies across community disadvantages and system capacity. The central expectation is that unequal recovery is most likely where disadvantaged community context intersects with limited technical, managerial, and financial capacity. The outcome of interest is return-to-compliance duration, measured as the

number of days from the start of a violation to its first recorded resolution. The analysis therefore focuses on whether systems serving disadvantaged communities experience longer unresolved spells, and whether those patterns differ by system size. It also accounts for compliance history and violation type, recognizing that repeated violations, concurrent unresolved violations, and more complex rule families may lengthen the path to formal closure. By modeling violation duration directly, this chapter complements prior studies focused on violation incidence and asks whether unequal protection also appears after violations occur, through systematically slower compliance recovery.

3.2 Literature Review

3.2.1 Return to Compliance as Restored Protection

Return to compliance is a central objective of drinking water governance because it marks whether regulatory protection has been restored after standards are violated. A duration perspective is especially important given that the consequences of noncompliance are time-dependent. The longer a violation remains unresolved, the longer exposure risks may persist and the longer the restoration of protection is delayed. Return-to-compliance duration therefore provides a governance-relevant performance lens that is distinct from violation incidence alone.

This chapter interprets return to compliance as a feasibility-constrained recovery process rather than as a simple byproduct of enforcement intensity. Once a violation is identified, the timing of resolution depends on whether a system can diagnose the problem, mobilize resources, implement corrective action, and satisfy the procedural requirements necessary for formal closure. Capacity-oriented perspectives help explain variation in compliance recovery by emphasizing that organizations differ in the analytical, operational, and coordination capabilities they can mobilize under constrained resources (Andrews & Boyne, 2010; Wu et al., 2015).

Observed recovery may also be slowed by procedural demands, including documentation, reporting, approvals, and coordination, even when substantive remediation is underway (Herd & Moynihan, 2018).

This feasibility-based interpretation is particularly appropriate in the context of community water systems. In public service settings characterized by intergovernmental oversight, compliance recovery often depends not only on technical correction but also on planning, coordination, and verification processes that connect remediation to formally recognized resolution (Konisky & Teodoro, 2016). In drinking water governance, noncompliance persistence reflects both what systems can practically do given their technical, managerial, and financial capacity and the institutional steps required to document and confirm return to compliance (Blanchard & Eberle, 2013; Allaire & Acquah, 2022). Observed return-to-compliance duration should therefore be understood as a function of both remediation feasibility and the administrative processes that translate corrective action into recorded compliance.

3.2.2 Remediation Feasibility: Technical, Managerial, and Financial Capacity

A central determinant of return-to-compliance duration is remediation feasibility, which is commonly understood through technical, managerial, and financial capacity. Federal statute and EPA guidance frame capacity development as both an entry condition for new systems and a continuing strategy for helping existing systems acquire and maintain the capacity needed to comply with national drinking water regulations (42 U.S.C. § 300g-9; EPA, 1999, 2025a). California's primacy agency likewise treats technical, managerial, and financial capacity as an explicit organizing framework for improving system performance and long-term sustainability, which reinforces its policy relevance for explaining variation in return-to-compliance duration in California community water systems (State Water Board, 2022, 2025a). Under this framework,

compliance is not only a technical endpoint, but also an organizational capacity to deliver safe service consistently over time.

EPA's technical, managerial, and financial framework emphasizes that compliance depends not only on infrastructure installation but also on the sustained capacity to operate systems effectively, monitor performance, and manage compliance over time (EPA, 2025c, 2025g). In this sense, system size provides a useful proxy for remediation capacity, as economies of scale shape the resources, expertise, and operational flexibility available for corrective action. Consistent with this view, drinking water research often treats system size as an indicator of technical, managerial, and financial capacity, with smaller systems typically facing greater compliance vulnerability and more limited remediation feasibility than larger systems (Statman-Weil et al., 2020).

Managerial capacity is also important for return-to-compliance duration, with the identification of a technical remedy alone insufficient to ensure timely corrective action. Rather, remedies must be translated into coordinated and sustained implementation. In this sense, managerial capacity includes the ability to coordinate actors, sequence decisions, manage implementation routines, and sustain corrective efforts across organizational handoffs, capacities that public management and implementation research treats as central to organizational performance and policy delivery (Andrews & Boyne, 2010; O'Toole & Montjoy, 1984; Hill & Hupe, 2021). Remediation in water systems commonly depends on interaction among utilities, engineers, contractors, financing agencies, and regulators. When these coordination processes are weak, delays can emerge at multiple points, including procurement, contracting, approvals, and project execution. From a duration perspective, managerial capacity therefore shapes how quickly systems can convert a recognized compliance problem into completed corrective action.

Financial capacity constitutes a third core dimension of remediation feasibility. Financial stress can slow compliance recovery by constraining the staffing, planning, and administrative capacity needed to prepare financing applications, manage procurement, coordinate across governments, and sustain corrective projects over time (Hendrick, 2004; Maher & Deller, 2007; Warner et al., 2021). Compliance recovery often requires upfront capital investment as well as sustained operation and maintenance expenditures. Financial capacity therefore remains a relevant broader capacity signal even for public or quasi-public providers operating under distinct utility governance arrangements, insofar as it can shape the managerial talent, intergovernmental coordination, and political support needed to keep corrective efforts on schedule (Terman & Feiock, 2015).

Two related theoretical perspectives extend this account by showing why remediation feasibility can vary even among systems facing similar technical problems. One centers on resource mobilization. Technical, managerial, and financial capacity matters not only for internal operations, but also for a system's ability to secure and manage the external inputs needed for correction. Because organizational performance depends in part on access to critical resources controlled by outside actors (Pfeffer & Salancik, 1978), remediation timing may depend on whether a system can obtain affordable financing, contract with qualified technical providers, and coordinate effectively across agencies and partners. Systems with stronger managerial and financial capability are therefore generally better positioned to assemble and manage the external resources required for durable correction. A related perspective comes from organizational slack. Organizational theory defines slack as resources beyond minimum operating requirements that buffer organizations against shocks and support adaptation, which makes limited slack an important source of backlog, delayed sequencing, and slower compliance recovery when new

demands arise (Cyert & March, 2020; Bourgeois, 1981). In the drinking water context, slack may take practical forms such as redundant staffing, contracted support, managerial bandwidth, financial reserves, and operational redundancy. These resources can help systems absorb unexpected costs, maintain oversight, and sustain corrective action through verified closeout. From this perspective, observed variation in return-to-compliance timing across community water systems reflects systematic differences in the resources and organizational capabilities needed to mobilize, implement, and sustain durable correction.

3.2.3 Persistent Noncompliance and Cumulative Compliance Burden

Empirical evidence of repeat violations and spatial clustering suggests that many return-to-compliance spells reflect ongoing constraints rather than isolated shocks, which makes duration a substantively meaningful outcome. Prior study documents persistent noncompliance in drinking water regulation, including repeated violations within systems and geographic clustering across communities (Allaire et al., 2018). These patterns strengthen the case for a duration-based perspective by indicating that public-health protection depends not only on whether violations occur, but also on how long systems remain out of compliance once problems are detected. In this chapter, cumulative compliance burden refers to the way unresolved remediation tasks, competing organizational demands, and compliance-related administrative work can accumulate over time and slow return to compliance (Pierson, 2000; March & Simon, 1958; Moynihan et al., 2015). This concept is analytically important, as it identifies mechanisms through which compliance recovery may be delayed even when systems are actively working toward resolution.

Cumulative compliance burden operates through four related mechanisms. State dependence highlights the temporal persistence created by unresolved prior problems (Pierson,

2000). Bounded attention explains how overlapping violations create internal organizational bottlenecks as systems triage among competing compliance tasks (March & Simon, 1958). Administrative workload adds a further layer of delay by requiring monitoring, documentation, communication, and other follow-up work in addition to remediation itself (Moynihan et al., 2015). Resource dependence extends the argument beyond the organization by showing that recovery can also stall when systems rely on outside actors for financing, technical expertise, labor, and approvals (Pfeffer & Salancik, 1978). In this formulation, cumulative burden captures how compliance tasks stack and lengthen return-to-compliance duration within and across spells, apart from a system's baseline technical, managerial, and financial capacity to implement a remedy. This framework clarifies why compliance history and concurrency are theoretically relevant predictors of return-to-compliance duration.

State dependence helps explain why compliance recovery can become harder over time. The practical difficulty of resolving a new violation depends in part on the stock of unresolved problems and on the organizational capacity already committed to ongoing corrective work. In Pierson's account of path dependence, early constraints and provisional choices can become self-reinforcing by narrowing feasible alternatives and raising the costs of later adjustment (Pierson, 2000). This logic is most plausible when systems rely on interim fixes or defer more durable capital solutions, as shifting to a long-term remedy may then require larger projects, longer financing and procurement timelines, and additional coordination before resolution can be achieved. This implies that return-to-compliance duration may vary systematically with compliance history, with systems entering a spell with larger backlogs or more persistent unresolved problems facing higher switching and coordination costs.

A related mechanism is bounded attention. Classic organizational theory emphasizes that organizations operate under limits in attention and information processing, thereby requiring them to prioritize and sequence tasks under competing demands (March & Simon, 1958). When multiple violations overlap, community water systems must triage across interdependent corrective actions and manage scheduling bottlenecks, so delay can arise from sequencing and queueing constraints rather than from low underlying capacity alone. Concurrency should therefore be associated with longer return-to-compliance duration, as stacked compliance demands can overwhelm managerial attention and slow implementation.

Administrative workload provides a third burden-based mechanism. Return to compliance requires not only substantive remediation but also follow-up self-monitoring, documentation, regulator communication, and other compliance-related tasks that consume time and organizational effort. Administrative burden scholarship clarifies that learning and compliance costs embedded in administrative systems can slow progress when demands accumulate (Moynihan et al., 2015). The relevant point is not to treat administrative follow-up as separate from remediation, but to recognize that overlapping compliance episodes increase the workload required to move problems toward closure. This dynamic helps explain why return-to-compliance duration can lengthen as unresolved issues accumulate.

Cumulative compliance burden may also intensify when overlapping demands draw on the same external support. Resource dependence theory emphasizes that organizational performance is shaped by reliance on outside actors for critical inputs (Pfeffer & Salancik, 1978). In the community water system context, financing, engineering expertise, specialized labor, and intergovernmental approvals can all become pacing factors in remediation. As multiple compliance demands overlap, systems may face queueing not only internally but also in access

to these external supports, reinforcing the expectation that compliance history and concurrency predict longer return-to-compliance duration. This perspective also clarifies the distinction between cumulative burden and baseline technical, managerial, and financial capacity.

Cumulative burden concerns how overlapping demands accumulate and slow progress over time, whereas baseline capacity shapes the remedies a system can realistically mobilize and sustain (EPA, 1999, 2025a).

3.2.4 Administrative Verification, Closeout, and Recording Cadence

Even when remediation feasibility and cumulative burden shape the pace of recovery, the timing observed in administrative data also reflects the institutional processes through which return to compliance is verified and recorded. Observed return-to-compliance duration therefore captures not only the time required to remedy a violation but also the institutional work needed to document, confirm, and close cases in program records. Accountability systems rely on routinized practices of verification and validation, administrative closure can lag substantive improvement (Power, 1997). In SDWA implementation, follow-up procedures, evidentiary standards, and case-closure practices further indicate that recorded duration includes a meaningful administrative closeout component (EPA, 2009). These requirements also create compliance and learning costs tied to documentation, monitoring schedules, and regulator follow-up, particularly when cases are complex or overlapping (Moynihan et al., 2015).

Administrative records do not capture compliance transitions in a purely real-time or mechanical way. The timing of recorded resolution also reflects how agencies process, verify, and code cases, not only the underlying pace of remediation (Bovens & Zouridis, 2002). In SDWIS, resolution dates therefore reflect institutional determinations and coding practices as well as substantive progress toward correction (EPA ECHO documentation). Reporting cadence

and data-quality constraints can introduce additional timing variation, since administrative datasets capture verification and reporting cycles rather than real-time compliance transitions. Public reporting systems may incorporate quarter-based verification windows and reporting lags, while oversight assessments have identified uneven completeness and reliability in state-reported data, limitations that matter when duration is measured from administrative records (U.S. GAO, 2011). These institutional features suggest that return-to-compliance timing is best interpreted as restored protection as administratively verified, a perspective that informs both empirical design choices and the limits of causal claims based on administrative timing measures (Moynihan, 2008; Bevan & Hood, 2006).

3.2.5 Unequal Protection through Unequal Recovery

In the drinking water context, unequal protection may operate in part through unequal noncompliance recovery capacity, as community conditions can shape whether, and how quickly, water systems mobilize the resources required for durable correction and verified return to compliance. The Drinking Water Disparities Framework makes this point directly by tracing how sociopolitical and economic conditions operate through multilevel actors to constrain access not only to safe water, but also to the financial and organizational resources needed to sustain protection over time (Balazs & Ray, 2014). Recovery capacity is therefore shaped not only by internal system characteristics, but also by the communities that systems serve.

One pathway runs through finance. Water systems serving disadvantaged communities may face tighter revenue and affordability constraints, which can weaken reserves, limit borrowing capacity, and reduce the political feasibility of rapid rate adjustments needed to finance corrective action (Patterson & Doyle, 2021). When financing cannot be assembled quickly and reliably, remediation is more likely to proceed incrementally, procurement may be

delayed, and longer intervals may pass before corrective performance can be demonstrated and formally closed out.

A second pathway runs through managerial and technical capacity. Disadvantaged and small-system contexts are more likely to face thin administrative capacity for grant and loan navigation, contracting, compliance documentation, and project management, all of which can shape the pace of implementation even when a remedy is technically identifiable (Dobbin & Fencl, 2021; Balazs & Ray, 2014). Limited access to certified operators, specialized engineers, and contractors can also produce start-stop implementation and slower stabilization of corrective actions, which in turn delays verified closeout (Dobbin & Fencl, 2021; Switzer & Teodoro, 2017, 2018). If unequal protection operates through unequal recovery capacity, disadvantaged-community indicators should be associated with longer return-to-compliance duration, while financing, technical assistance, and consolidation can be interpreted as corrective pathways that relax these constraints and shorten exposure burdens.

3.2.6 Corrective Pathways in Compliance Recovery

Corrective pathways help explain variation in return-to-compliance duration by serving as the institutional mechanisms through which binding constraints on capital access, implementation capacity, and specialized expertise can be relaxed. At the federal level, the Drinking Water State Revolving Fund supports not only infrastructure finance but also set-aside resources, including a dedicated small-system technical assistance set-aside and related program flexibilities, while EPA's Training and Technical Assistance grant programs fund organizations that work directly with small systems to build the technical, managerial, and financial capacity needed to achieve and maintain SDWA compliance (EPA, 2025a, 2026g).

California makes these compliance-recovery pathways especially salient by treating financing, technical assistance, and institutional restructuring as explicit strategies for achieving safe and sustainable service, particularly for small and disadvantaged systems. The State Water Board's SAFER-aligned approach emphasizes support for water partnerships and provides guidance on funding and incentives for consolidation and regionalization, while also maintaining authority to order mandatory consolidation or extension of service for systems that fail, or are at risk of failing, to provide safe drinking water consistently (State Water Board, 2026a, 2025a, 2026b, 2026d). These features underscore that return-to-compliance duration depends not only on the underlying technical problem, but also on whether systems can access and navigate supports that relax recovery constraints.

These corrective pathways reinforce the chapter's feasibility-based account of compliance recovery. Financing, technical assistance, and consolidation are not peripheral policy tools. They are part of the institutional environment that shapes the practicability, pace, and durability of return to compliance, especially where baseline capacity is limited and recovery constraints are most binding. Figure 3.1 summarizes this chapter's conceptual framework by linking violation burden, system and community capacity, corrective pathways, and administrative verification to recorded return-to-compliance duration.

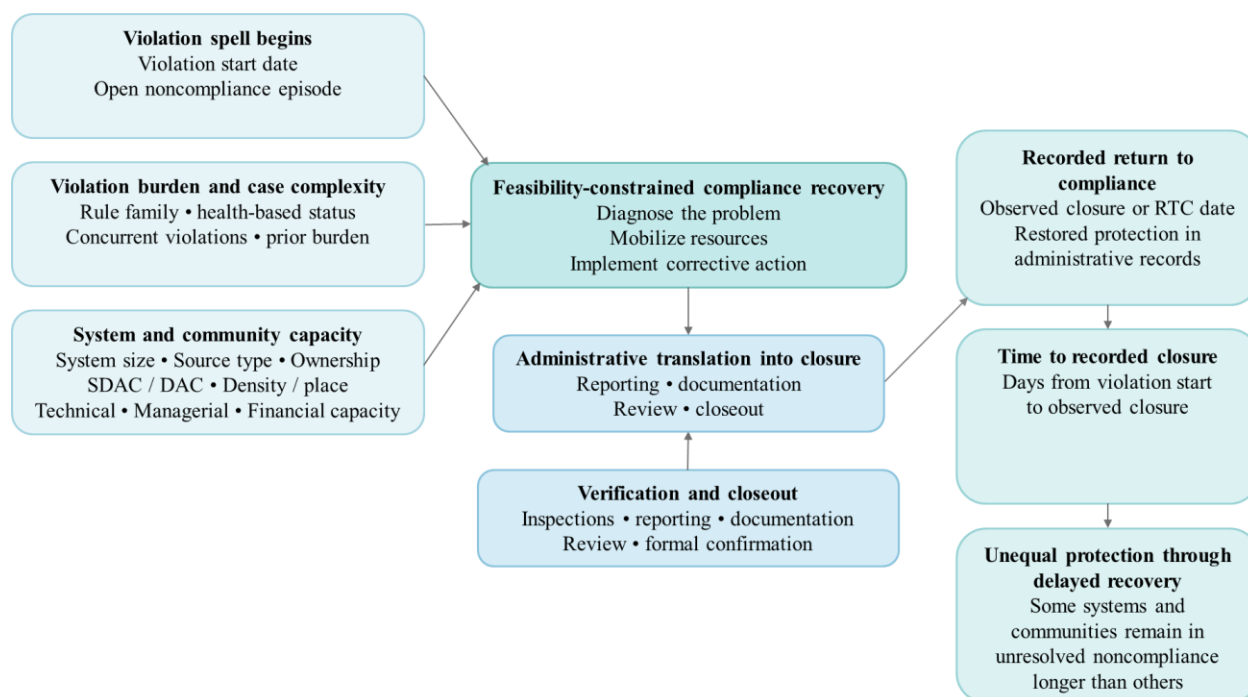


Figure 3.1: Conceptual Framework for Compliance Recovery and Return to Compliance

Note. This figure frames return to compliance as a recovery process shaped by violation burden, system and community capacity, and administrative verification and closeout. In this chapter, time to recorded resolution reflects both remediation feasibility and the administrative steps required for documented return to compliance in the administrative record.

3.3 Data

3.3.1 Data and Analytic Sample

This chapter draws on the EPA’s Violations and Enforcement file obtained through the Enforcement and Compliance History Online (ECHO) system. For each California community water system identified by PWSID, the file records violation-level information, including violation identifiers, begin and end dates, return-to-compliance dates, violation status, violation category code, violation family code, health-based and major-violation indicators, and public notification tier. It also links each violation to associated enforcement records, including

enforcement identifiers, enforcement dates, enforcement action types, and enforcement action categories.

Most violation and enforcement activity enters the data on a quarterly basis. This chapter defines the study period from the first quarter of 1995 through the fourth quarter of 2024. The analysis begins in 1995, as that year follows the initial period in which community water systems began reporting violations to EPA and offers the most complete coverage of violations and enforcement activity. This start date also has substantive relevance, as the number of contaminants regulated under the SDWA increased from 31 in 1990 to 83 by 1995, expanding the scope of regulatory compliance during the study period (Benneer & Olmstead, 2008). The raw California file contains 59,062 observations of violations and associated enforcement actions from 1995 to 2024.

The source data may report multiple enforcement actions and compliance milestones for the same violation identifier, but these entries represent the evolving regulatory response to a single violation rather than separate outcome events. The analysis therefore treats each unique violation as one survival spell and uses the violation begin date as the spell origin. This procedure produces one record per violation identifier for all descriptive statistics and survival models. The chapter defines return to compliance as the survival outcome for each spell, with event coding, censoring, and duration measurement described in the following section. The final analytic sample contains 10,714 violation spells with observed begin dates between January 1, 1995, and December 31, 2024. Given that no included spell begins before January 1, 1995, the analysis does not involve delayed entry for pre-1995 carryover spells.

3.3.2 Outcome Construction: Return-to-Compliance Duration

This chapter defines the survival event as recorded return to compliance in the administrative record. In the main analysis, a violation is treated as an event when its status is coded as Resolved¹ on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as return-to-compliance events. Instead, they remain in the risk set and are right-censored if no Resolved status is observed by the end of follow-up. This coding treats Addressed as evidence of regulatory action or attempted correction rather than confirmed return to compliance. The chapter constructs survival time using the return-to-compliance date when available and supplements it with other administrative status-date fields only when needed to identify the timing of recorded resolution.

This chapter measures survival time as the number of days from the violation begin date to the recorded resolution date. For violations that do not reach Resolved status during the study window, the analysis measures duration from the violation begin date to the administrative censor date of December 31, 2024. The main specification therefore treats the outcome as time to recorded return to compliance in the administrative data.

The analytic sample is organized at the violation-spell level rather than the system level, since individual community water systems may contribute multiple spells over the study period. A violation spell is defined by the administrative violation record rather than by manually combining violations that occur close together in time. If a system returns to compliance and later incurs the same type of violation in a subsequent quarter or compliance period, the later

¹ In ECHO/SDWIS, “Resolved” indicates that a violation has at least one resolving enforcement action; “Addressed” indicates that one or more formal enforcement actions have occurred but the violation is not resolved or archived; “Unaddressed” indicates no formal enforcement action; and “Archived” indicates a violation that is unresolved but no longer contributes to the system’s current compliance status. In the main analysis, only Resolved is treated as the event.

record is treated as a new spell rather than as a continuation of the earlier one. In the sample, 1,567 systems contribute 10,714 violation spells, and about 70 percent of systems contribute more than one spell. The median system contributes three spells, with an interquartile range from one to six spells. This means that recurrent violators are common in the analytic sample rather than excluded from it. At the same time, the chapter does not estimate a full recurrent-event survival model. Instead, it treats each observed violation spell as a separate duration outcome and addresses repeated spells within systems by clustering standard errors at the community water system level and by including prior violation burden and recent health-based noncompliance measures to capture cumulative compliance burden.

3.3.3 Violation and Enforcement Characteristics

3.3.3.1 Violation characteristics

This chapter characterizes violations using category, rule family, and public notification tier. Violation category code distinguishes Maximum Contaminant Level (MCL), Maximum Residual Disinfectant Level (MRDL), Treatment Technique (TT), Monitoring and Reporting (MR), Monitoring (MON), Reporting (RPT), and other violations. EPA identifies MCL, MRDL, and TT violations as health-based, whereas monitoring, reporting, and related procedural violations are coded separately and do not necessarily imply the same level of immediate health risk. The analysis therefore includes a binary indicator for whether a violation is health-based as a measure of substantive severity (EPA, 2025e, 2026f).

The chapter also regroups violation family codes into broader rule categories to capture meaningful differences in health relevance and compliance complexity. The resulting categories are nitrate, arsenic, total coliform, disinfection byproducts, public notification and Consumer Confidence Rule violations, lead and copper, and a residual category that includes other

contaminants and rule families regulated under the SDWA, including volatile and synthetic organic chemicals, inorganic chemicals, radionuclides, and contaminants covered by the Surface Water Treatment Rules and Ground Water Rule (EPA, 2026b). This regrouping reduces the number of rule categories while preserving substantive variation in the technical and public-health context of violations. Total coliform violations indicate microbial contamination risk and are commonly addressed through disinfection, filtration, and corrective actions required under the Ground Water Rule and the Revised Total Coliform Rule (EPA, 2026d). Nitrate violations arise when concentrations exceed safe levels and pose acute risks such as methemoglobinemia in infants, often in settings affected by agricultural runoff, septic systems, or wastewater discharge (Ward et al., 2005; EPA, 2007b). Arsenic violations reflect exceedances of the Maximum Contaminant Level for a naturally occurring toxicant associated with chronic health effects, including cancer and cardiovascular disease (Smith et al., 2000; Steinmaus et al., 2003). Disinfection byproduct violations capture contaminants formed when disinfectants react with organic matter during treatment and are monitored due to their potential carcinogenic and reproductive health effects (Villanueva et al., 2004; Richardson et al., 2007; EPA, 2025f). Lead and copper violations indicate contamination linked to plumbing and distribution systems, with especially serious neurotoxic implications for children (EPA, 2026c; Hanna-Attisha et al., 2016). Public Notification and Consumer Confidence Rule violations reflect failures to provide required information to consumers, making them relevant indicators of regulatory accountability and risk communication (EPA, 2025d, 2026e).

The analysis also controls for public notification tier to capture differences in regulatory urgency. Public notification tier is related to violation type but is not fully determined by rule family in the analytic sample. Under the SDWA Public Notification Rule, Tier 1 applies to

violations or situations with immediate potential to have serious adverse effects on human health, Tier 2 applies to other serious violations or situations with potential to have serious but not immediate adverse effects, and Tier 3 applies to other violations and situations not requiring Tier 1 or Tier 2 notice. In this sample, public-notification tier varies within several rule families, including Total Coliform Rule and nitrate violations, but it is also closely tied to violation type. For this reason, the baseline specification uses rule-family indicators to capture differences in violation type and regulatory context, while Tier 1 public-notification measures are used to capture concurrent compliance burden rather than as a separate focal classification of the main violation (Benneer & Olmstead, 2008; Grooms, 2016; Baker et al., 2023).

3.3.3.2 Compliance history and concurrent violations

The analysis controls for prior violation history and concurrent violations. Repeated noncompliance can shape both compliance recovery and regulatory response. Systems with a history of repeated violations may face persistent operational, managerial, and resource constraints that limit their ability to comply with SDWA requirements, particularly in disadvantaged communities (Allaire et al., 2018; Balazs & Ray, 2014). Regulatory agencies may also adjust monitoring and enforcement in light of observed compliance histories, hence prior violations can influence the likelihood, intensity, and timing of subsequent regulatory action (Carpenter, 2001; Dal Bó, 2006). Including indicators of violation history therefore helps distinguish the association of system and community characteristics with return-to-compliance duration from the influence of accumulated compliance problems. Prior studies specifically suggest accounting for concurrent violations, including monitoring and reporting violations, as overlapping noncompliance can complicate correction and regulatory follow-up (Kang & Silveira, 2021; Zhao & Lyon, 2022).

To operationalize violation history, the chapter constructs a count of violations occurring within the five years preceding each reported violation. The analysis also constructs alternative counts using three-year and ten-year windows, but the five-year measure serves as the primary indicator, as the three-year window yields relatively few prior violations for most systems, whereas the ten-year window captures a broader and less proximate compliance history. To capture more immediate overlap in health-relevant noncompliance, the chapter also counts health-based violations occurring within one quarter of a given reported violation. This choice aligns with the quarterly reporting structure of the data and captures recent compliance conditions that may plausibly influence recovery dynamics. Additionally, the analysis includes binary indicators for concurrent violations, defined as whether a given violation overlaps with other violations in the prior quarter or during its compliance duration, and whether health-based violations occur within one quarter before or after the reported violation. These measures capture the extent to which a violation arises in the context of stacked compliance demands rather than as an isolated episode.

3.3.4 Community Water System Characteristics

3.3.4.1 System size, ownership, and source type

Community water systems provide year-round service to at least 25 people or 15 service connections. This chapter draws system-level characteristics from SDWIS, including system size, primary water source, and ownership type. Following EPA guidance, the analysis classifies system size by population served (EPA, 2013). Large systems serve more than 50,000 people, medium systems serve 3,301 to 50,000 people, and small systems serve fewer than 3,300 people.

The analysis classifies primary water source into three categories: surface water, groundwater, and purchased water. Surface-water systems rely on surface-water sources.

Groundwater systems rely on groundwater or groundwater under the direct influence of surface water. Purchased-water systems obtain water through wholesale or other purchased-water arrangements. Ownership is coded as public or private. Public systems include federal, local governmental, and tribal ownership, while private systems include privately owned and public-private systems.

These characteristics capture important dimensions of remediation feasibility and regulatory context. System size serves as a well-established proxy for technical, managerial, and financial capacity, with smaller systems often facing greater constraints in maintaining compliance and implementing corrective action (Acquah & Allaire, 2023). Ownership may also shape regulatory interactions, given that public and private providers operate under different institutional incentives and governance arrangements. Source type is relevant, given variation across water sources in treatment complexity, compliance vulnerability, and operational risk. For example, purchased-water systems may face different compliance dynamics than systems that treat source water directly (Wallsten & Kosec, 2008).

3.3.4.2 Housing density and urban-rural context

To capture spatial context, the analysis constructs a continuous measure of the natural log of housing units per square mile within each water system's service area. Housing density equals the total number of housing units divided by service area in square miles. This measure serves as an indicator of settlement intensity and, more indirectly, of the socioeconomic and infrastructural context in which a community water system operates. Its use is also consistent with the U.S. Census Bureau's urban-rural classification framework, which incorporates housing-unit density in defining urban areas (U.S. Census Bureau, 2024; U.S. Department of Agriculture, Economic Research Service, 2025). For descriptive purposes, the chapter additionally classifies systems as

urban or rural using a housing-density threshold. Systems with housing density at or above 0.0005 units per square meter, equivalent to approximately 500 units per square kilometer or 1,295 units per square mile, are classified as urban. Systems below this threshold are classified as rural.

3.3.5 Community Demographics

This chapter links California community water system service areas to Census-based demographic data in order to construct water-system-level measures of community context. The analysis uses GIS shapefiles of service areas obtained from the State Water Resources Control Board. It draws demographic data at the Zip Code Tabulation Area level from the U.S. Census for 2000 and 2010 and from the five-year American Community Survey for 2014 to 2019. For intercensal years, the analysis estimates values using monotone piecewise cubic interpolation. It then standardizes all data to 2010 ZCTA boundaries. To assign demographic characteristics to each service area, the chapter uses areal interpolation to calculate population-weighted averages of demographic variables within each water system. The weights reflect both the share of each ZCTA's land area that falls within the service area boundary and the total population of the ZCTA. These weights therefore approximate the share of the service population attributable to each intersecting ZCTA (Allaire & Acquah, 2022).

The chapter constructs indicators of economic disadvantage and racial/ethnic service context using statewide median household income and population-composition measures, consistent with prior study on drinking water inequality (Allaire & Acquah, 2022). The State Water Resources Control Board defines disadvantaged communities as communities with median household income below 80% of the statewide median and severely disadvantaged communities as communities below 60% of the statewide median (State Water Board, 2015). To

operationalize these definitions over the study period, the analysis compiles California annual median household income estimates from decennial Census data and American Community Survey one-year estimates for 1995 to 2024, converts all values to 2017 dollars, and applies the corresponding 80% and 60% thresholds to the statewide average median household income over the full study period. Under this approach, the DAC threshold corresponds to approximately \$56,800 in 2017 dollars and the SDAC threshold to approximately \$42,600. The chapter also constructs indicators for African American-serving and Hispanic-serving systems using statewide population-share benchmarks averaged over the study period (Allaire & Acquah, 2022). A system is coded as African American-serving if its average African American population share exceeds the statewide average of 6.3% and as Hispanic-serving if its average Hispanic population share exceeds the statewide average of 36.6%. In addition, the analysis constructs an indicator for systems that are both DAC and Hispanic-serving in order to capture heterogeneity in income and system characteristics across Hispanic-serving systems (Allaire & Acquah, 2022).

3.3.6 Policy Periods and Additional Temporal Controls

The baseline models include fixed effects for the calendar year in which each violation spell begins. These start-year fixed effects provide the main temporal adjustment in the chapter and flexibly account for statewide changes in SDWA implementation, economic conditions, drought context, regulatory priorities, and administrative practice across violation-start cohorts. Right-censoring is handled directly through the survival framework, while start-year fixed effects adjust for calendar-year differences among violation cohorts.

As a sensitivity check, this chapter also classifies violations into broader policy periods to assess whether the main findings depend on using fine-grained year fixed effects. These policy

periods summarize major contextual shifts in the compliance environment, including the period following California’s Human Right to Water policy shift and the COVID-19 operational disruption. This broader epoch specification is not the preferred baseline specification. Instead, it provides a parsimonious alternative that tests whether the main SDAC and system-size patterns remain similar when calendar time is controlled through policy-anchored periods rather than year-by-year indicators.

3.4 Empirical Strategy

Building on the preceding framework, this chapter treats return-to-compliance duration as a feasibility-based recovery outcome rather than as a simple reflection of enforcement intensity alone. The empirical analysis therefore centers on three sets of factors. Compliance history and concurrent violations capture cumulative burden and queueing dynamics. System characteristics and community context capture variation in technical, managerial, and financial capacity, including the ability to mobilize financing, expertise, and administrative effort for durable correction. Institutional and administrative processes also shape the observed timing measure, since recorded resolution reflects verification, documentation, and formal closeout as well as substantive remediation.

Under this framework, the estimates are interpreted as associations consistent with differences in compliance recovery capacity and burden rather than as direct estimates of deterrence effects. This distinction is especially important for the chapter’s environmental justice analysis. Disadvantaged-community indicators are used to assess whether return-to-compliance duration differs systematically across community contexts in ways consistent with unequal recovery capacity and, in turn, unequal protection in practice.

3.4.1 Outcome, Censoring, and Survival-time Construction

The analysis models the duration from the beginning of a violation spell to recorded return to compliance. The unit of analysis is the violation spell. The main outcome is therefore time to recorded resolution. A spell is treated as an event when the violation is coded as Resolved on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as events. They remain under observation until they either reach Resolved status or are right-censored on December 31, 2024. This construction reflects the administrative logic of the data, in which recorded resolution may capture substantive remediation, verification, documentation, and formal closeout. The timing measure should therefore be interpreted as duration to recorded return to compliance in the administrative record rather than as a purely technical measure of remediation timing.

Let T_{iv} denote the duration in days for violation spell v in system i , and let δ_{iv} indicate whether the spell reaches observed resolution during follow-up. Survival time is defined as

$$T_{iv} = d_{iv}^{close} - d_{iv}^{begin} \text{ if } \delta_{iv} = 1, \quad T_{iv} = d^{censor} - d_{iv}^{begin} \text{ if } \delta_{iv} = 0 \quad (3.1)$$

where d^{censor} is December 31, 2024. The event indicator equals one when a violation is coded as Resolved on or before the censor date and equals zero otherwise. Violations coded as Addressed, Archived, or Unaddressed are treated as right-censored if no Resolved status is observed by December 31, 2024.

3.4.2 Modeling Approach and Distribution Choice

This chapter uses a survival framework to examine how long violation spells remain unresolved while accounting for spells that remain open at the end of the observation period. The approach preserves the temporal dimension of compliance recovery and accommodates right-

censored outcomes. A Cox proportional hazards model was initially considered as a standard semiparametric approach to duration analysis. However, diagnostic checks did not support the proportional hazards assumption for key covariates. That limitation is relevant for the chapter's substantive aims. The analysis is concerned not simply with whether resolution becomes more or less likely at a given moment, but with whether some violations persist for systematically longer periods after accounting for differences in community context, system characteristics, compliance history, and type of violation.

An accelerated failure time (AFT) framework is therefore more appropriate because rather than requiring proportional hazards, it models covariate associations directly on the time scale (Wei, 1992). The AFT specification also offers a more intuitive interpretation for the present study. Coefficients are reported as time ratios, which indicate whether a covariate is associated with longer or shorter return-to-compliance duration. A time ratio above one indicates longer time to recorded return to compliance, whereas a time ratio below one indicates faster resolution. This interpretation aligns with the chapter's focus on delay, burden, and unequal recovery capacity.

For this chapter, among standard AFT families, the log-logistic specification provided the best overall fit in model-comparison exercises based on information criteria, with Weibull performing second best. The main results therefore use the log-logistic AFT model, while Weibull models are retained as a sensitivity check.

3.4.3 Baseline Log-Logistic AFT Specification

The baseline specification estimates how community demographics, system characteristics, compliance history, and violation characteristics are associated with return-to-compliance duration. The model is estimated on the violation-spell analytic sample using a log-

logistic AFT specification with fixed effects for violation start year. Standard errors are clustered at the community water system level to account for repeated violation spells within systems. The chapter therefore captures recurrent violators in the sample, but it does not estimate a full recurrent-event survival model that explicitly models event order or within-system recurrence processes.

The baseline estimating equation is

$$\ln(T_{iv}) = \alpha_{y(iv)} + X'_{iv}\beta + \phi'(Z_i \times S_i) + \varepsilon_{iv} \quad (3.2)$$

where $\alpha_{y(iv)}$ denotes fixed effects for the calendar year in which spell v begins, X'_{iv} includes the remaining covariates, Z_i denotes system-size category, S_i denotes SDAC status, and ε_{iv} follows the standard logistic distribution. Under this specification, T_{iv} follows a log-logistic distribution. The interaction between system size and SDAC status is included because the conceptual framework and descriptive evidence both suggest that community disadvantage and organizational scale may jointly shape the feasibility of durable correction.

The covariates capture several dimensions of compliance recovery. Community-context measures include DAC status and race and ethnicity indicators. System-level covariates include ownership type, primary water source, and housing density. Compliance-history measures capture prior violation burden and recent health-based noncompliance. Violation-level controls include grouped rule-family indicators and an indicator for health-based status. These variables adjust for observable differences in community context, system capacity, accumulated burden, and violation type that may shape the pace of recorded resolution.

The baseline model also allows the shape of the duration distribution to vary by system size. The ancillary equation is

$$\ln(\sigma_{iv}) = Z_i' \Upsilon$$

where Z_i' again denotes system-size category. Allowing the ancillary parameter to vary by size accommodates the possibility that long-duration spells are distributed differently across organizational scales, especially in the upper tail of unresolved violations.

The main interpretation of the baseline model centers on post-estimation contrasts rather than on individual coefficients alone. In particular, the chapter reports the SDAC contrast within each size category and the size contrasts within SDAC and non-SDAC systems. This approach keeps the environmental justice question at the center of interpretation and expresses the results in substantively meaningful comparative terms.

3.4.4 Rule-Family Composition Sensitivity

Return-to-compliance duration may differ across rule families given variation across violation types in technical complexity, corrective pathways, and administrative closeout requirements. As a result, baseline differences across SDAC and system-size groups could partly reflect compositional differences in the mix of violations rather than differences in recovery capacity alone.

To assess this possibility, the chapter estimates a pooled heterogeneity model that allows the SDAC and system-size gradients to vary across major rule families. From this model, it reports SDAC-by-size contrasts under a common rule-family distribution. These standardized contrasts test whether the baseline environmental justice patterns remain after reducing the influence of rule-composition differences across groups. The results of this specification are discussed in the sensitivity analyses section as a check on whether the baseline findings are robust to heterogeneity in rule-family composition.

3.4.5 Temporal Sensitivity Specifications

The baseline specification uses fixed effects for violation start year to absorb calendar-time variation in recorded resolution. This is the preferred temporal-control strategy because it flexibly accounts for statewide regulatory, economic, drought-related, and administrative changes across the study period, while the survival framework accounts for right-censoring of spells that remain unresolved by the end of follow-up. As a sensitivity check, the chapter re-estimates the baseline model using broader policy-epoch indicators in place of start-year fixed effects. These epochs are defined from the violation start year and are intended to capture major regulatory and administrative regimes, including the post-Human Right to Water period and the COVID-era disruption. The epoch specification therefore provides a parsimonious robustness check rather than the main temporal adjustment.

An additional sensitivity specification also allows the ancillary parameter to vary across broader temporal regimes. This extension relaxes the assumption that the shape of the duration distribution is constant across policy periods. These temporal sensitivity checks assess whether the main SDAC and system-size findings depend on the choice between fine-grained year fixed effects and broader policy-epoch controls. The results of these temporal sensitivity models are discussed in the sensitivity analyses section as a check on whether the main patterns are robust to broader policy-period controls.

3.5 Results

3.5.1 Descriptive resolution patterns

Table 3.1 summarizes the analytic sample. The study includes 10,714 violation spells, of which 78.5 percent reached recorded resolution by December 31, 2024. The median time to recorded resolution is 91 days overall, but the interquartile range extends from 30 to 731 days,

indicating substantial heterogeneity in how long violations remain unresolved in the administrative record. At the descriptive level, SDAC spells are slightly less likely than non-SDAC spells to reach resolution by the end of follow-up, 77.2 percent versus 79.1 percent, while median time to recorded resolution is similar across groups. This pattern shows that simple central-tendency comparisons do not fully capture how resolution timing differs across community contexts. Appendix Table 3.A1 reports the distribution of violation rule families by SDAC status. The violation mix differs across SDAC and non-SDAC spells, with SDAC spells showing somewhat higher shares of disinfection byproduct and arsenic violations and somewhat lower shares of nitrate and other-rule violations. This descriptive difference supports the inclusion of rule-family indicators in the baseline AFT models and the additional rule-family sensitivity specification.

The composition of SDAC and non-SDAC spells also differs in ways that matter for later interpretation. SDAC spells are less concentrated in the zero-prior-violation category and more concentrated in higher-burden categories, indicating that they more often arise in settings with accumulated compliance problems. They are also somewhat more likely to have a recent prior health-based violation. These differences matter because any observed SDAC disparity in resolution timing may reflect not only community disadvantage itself, but also the broader compliance environments in which SDAC violations are embedded.

Organizational and community characteristics show additional contrasts across the two groups. SDAC spells are more concentrated in medium-sized systems, publicly owned systems, surface-water systems, and Hispanic-serving systems. Non-SDAC spells are more concentrated in the smallest systems, private systems, groundwater systems, and, to a lesser extent, African American-serving systems. Mean log housing density differs little across groups. These

descriptive differences show that SDAC and non-SDAC spells are not directly comparable in raw form. They arise in different kinds of systems and compliance contexts, which underscores the need for adjusted models.

Table 3.1: Descriptive Characteristics of the Violation-Spell Analytic Sample by SDAC Status

Characteristic	Overall	Non-SDAC	SDAC
N	10,714	7,423	3,291
Time to recorded resolution, days	91.0 [30.0, 731.0]	91.0 [30.0, 731.0]	91.0 [30.0, 682.0]
Recorded resolution by 12/31/2024			
No	2,302 (21.5)	1,552 (20.9)	750 (22.8)
Yes	8,412 (78.5)	5,871 (79.1)	2,541 (77.2)
Health-based violation			
Yes	5,311 (49.6)	3,873 (52.3)	1,438 (43.7)
No	5,391 (50.4)	3,538 (47.7)	1,853 (56.3)
Prior violation burden			
0	1,994 (18.6)	1,584 (21.4)	410 (12.5)
1–5	3,297 (30.8)	2,201 (29.7)	1,096 (33.3)
6–10	1,320 (12.3)	866 (11.7)	454 (13.8)
11+	4,091 (38.2)	2,760 (37.2)	1,331 (40.4)
Recent prior health-based violation			
No	9,246 (87.0)	6,427 (87.5)	2,819 (85.9)
Yes	1,380 (13.0)	916 (12.5)	464 (14.1)
System size			
≤3,300 served	8,230 (76.8)	5,909 (79.6)	2,321 (70.5)
3,301–50,000 served	2,028 (18.9)	1,153 (15.5)	875 (26.6)
>50,000 served	456 (4.3)	361 (4.9)	95 (2.9)

Characteristic	Overall	Non-SDAC	SDAC
Ownership			
Public	4,275 (39.9)	2,621 (35.4)	1,654 (50.3)
Private	6,427 (60.1)	4,790 (64.6)	1,637 (49.7)
Primary source			
Groundwater	7,588 (70.9)	5,370 (72.5)	2,218 (67.4)
Surface water	1,917 (17.9)	1,159 (15.6)	758 (23.0)
Purchased water	1,197 (11.2)	882 (11.9)	315 (9.6)
Log housing density	3.9 (2.3)	3.9 (2.2)	3.8 (2.5)
African American-serving system			
No	9,725 (91.9)	6,656 (91.3)	3,069 (93.3)
Yes	860 (8.1)	638 (8.7)	222 (6.7)
Hispanic-serving system			
No	7,016 (66.3)	5,457 (74.8)	1,559 (47.4)
Yes	3,569 (33.7)	1,837 (25.2)	1,732 (52.6)

Note. Values are *n* (column %) for categorical variables, mean (SD) for log housing density, and median [IQR] for time to recorded resolution. The unit of analysis is the violation spell.

Recorded resolution refers to violations coded as addressed or resolved by December 31, 2024; unresolved spells are right-censored at that date. SDAC = severely disadvantaged community.

Counts may vary across rows because of item-specific missingness. The table is descriptive and does not adjust for rule-family composition, system characteristics, or other covariates.

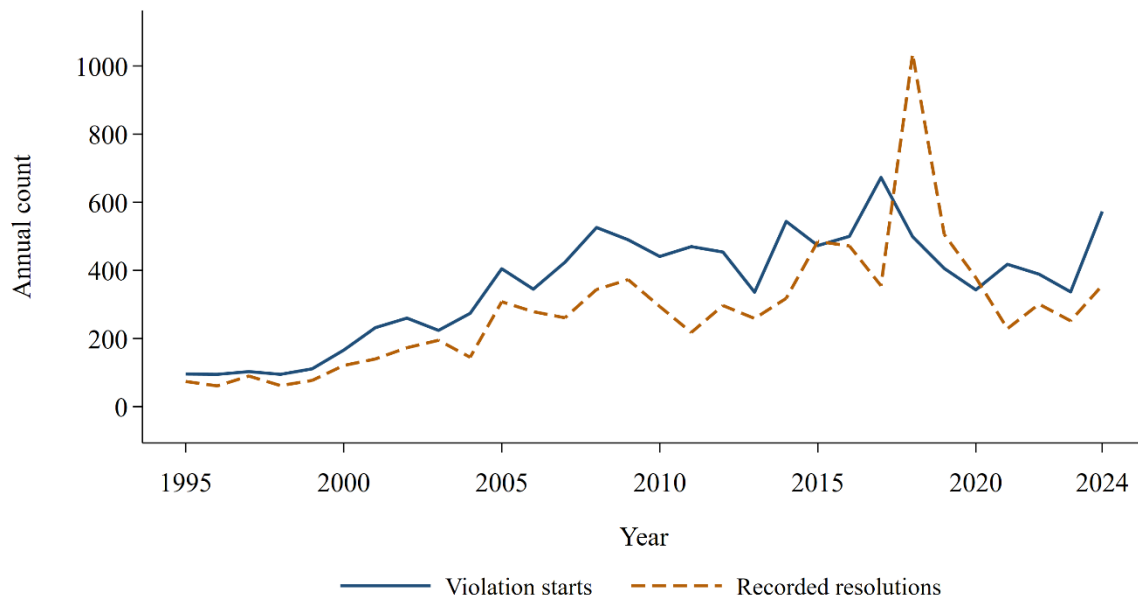


Figure 3.2: Violation Starts and Recorded Resolutions Over Time, 1995–2024

Note. The figure plots annual counts of violation starts and annual counts of recorded resolutions in the violation-spell analytic sample. Recorded resolutions reflect the chapter’s main outcome definition based on Resolved status in the administrative record. Trends are descriptive and do not adjust for changing cohort composition or follow-up length.

The temporal distribution of violations further clarifies the structure of the sample.

Annual starts and recorded resolutions do not move in perfect synchrony over the study period.

Violation starts rise into the 2000s, while recorded resolutions accelerate later, indicating that the growth in noncompliance inflow and the pace of administrative closeout followed related but distinct trajectories. Figure 3.2 provides this broad temporal orientation. The main implication is that resolution timing should be interpreted within a changing administrative environment rather than as a static process over the full 1995 to 2024 period.

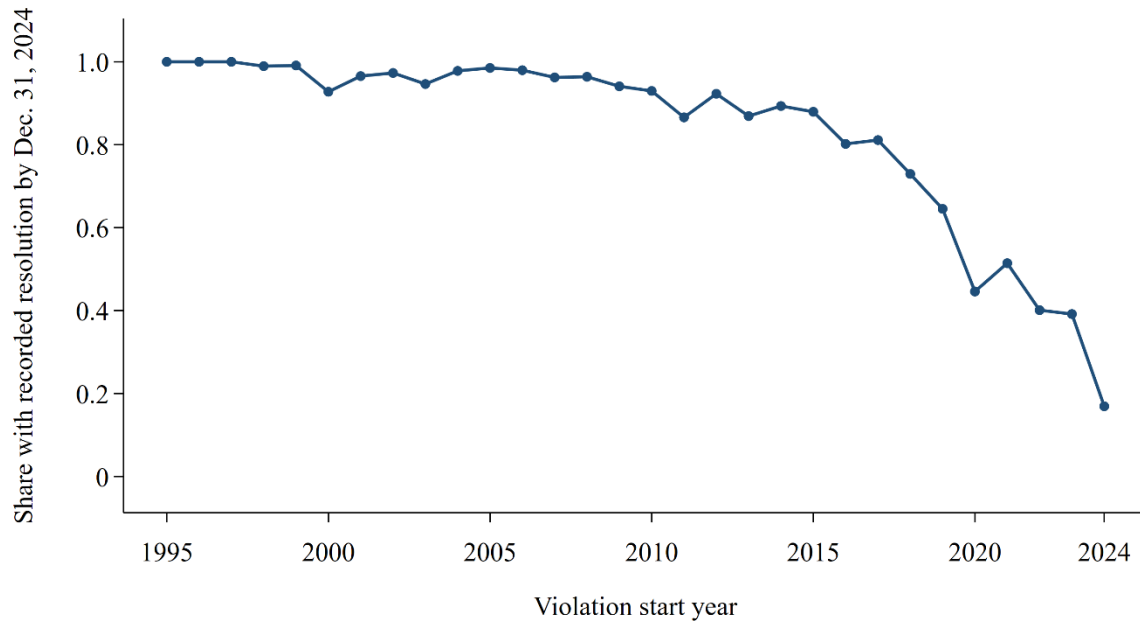


Figure 3.3: Recorded Resolution Share by Violation Start Cohort

Note. The figure shows the share of each violation start-year cohort that reaches recorded resolution by December 31, 2024. Lower resolution shares for more recent cohorts reflect shorter follow-up windows under right-censoring rather than necessarily lower long-run return-to-compliance probabilities.

Administrative censoring is a central feature of the sample and shapes how cohort-level resolution patterns should be interpreted. Earlier start-year cohorts approach near-complete recorded resolution by the end of follow-up, whereas more recent cohorts show lower observed resolution shares. This decline reflects the shorter time available for recent cohorts to reach resolution before the December 31, 2024 censoring date rather than weaker long-run resolution. Figure 3.3 illustrates this cohort pattern directly, while Appendix Figure 3.A3 shows the corresponding rise in unresolved shares as follow-up windows shorten. Appendix Figures 3.A1 and 3.A2 provide additional descriptive detail by health-based status, showing annual starts and recorded resolutions, as well as cohort-level recorded-resolution patterns.

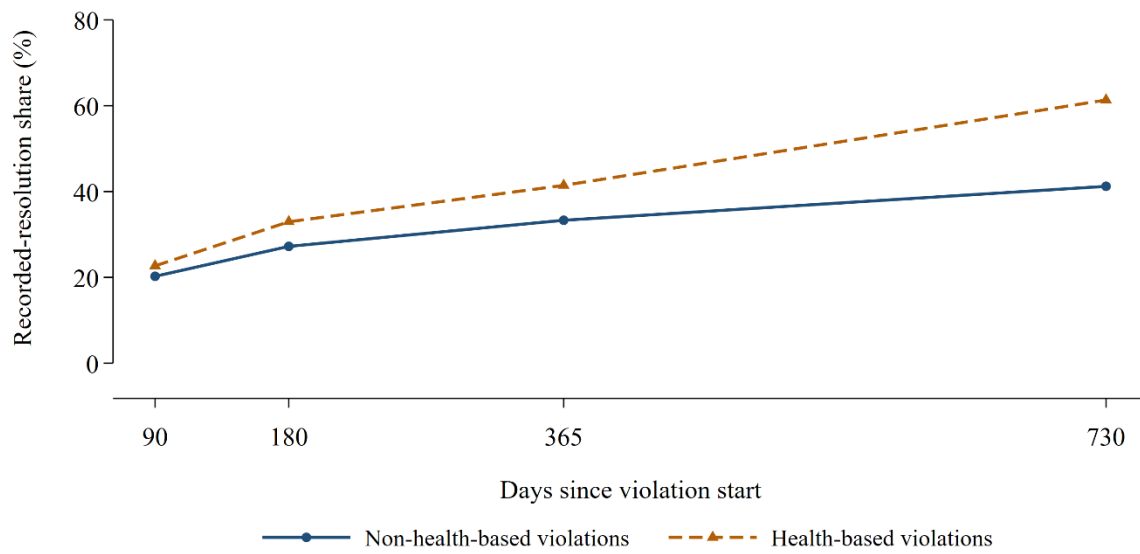


Figure 3.4: Kaplan-Meier Recorded-Resolution Shares Within Selected Horizons, by Health-Based Status

Note. The figure reports Kaplan-Meier cumulative recorded-resolution shares at selected horizons for health-based and non-health-based violations. Higher values indicate a larger share of violations reaching recorded resolution by the specified horizon. Results are descriptive and unadjusted, based on the violation-spell analytic sample.

Violation type also corresponds to meaningfully different timing patterns before covariate adjustment. Health-based violations reach recorded resolution more quickly than non-health-based violations, especially at shorter horizons, indicating that the pace of administrative closure differs across substantively distinct categories of noncompliance. Figure 3.4 shows this pattern using Kaplan-Meier recorded-resolution shares at selected time horizons. The appendix figures reinforce the same point from complementary angles. Appendix Figures 3.A1 and 3.A2 disaggregate annual starts, annual recorded resolutions, and cohort-level recorded-resolution shares by health-based status, while Appendix Figure 3.A4 extends the Kaplan-Meier comparison across broader policy periods. These descriptive results suggest that resolution

timing varies not only across cohorts, but also across violation types in ways that warrant adjustment in the multivariable analysis.

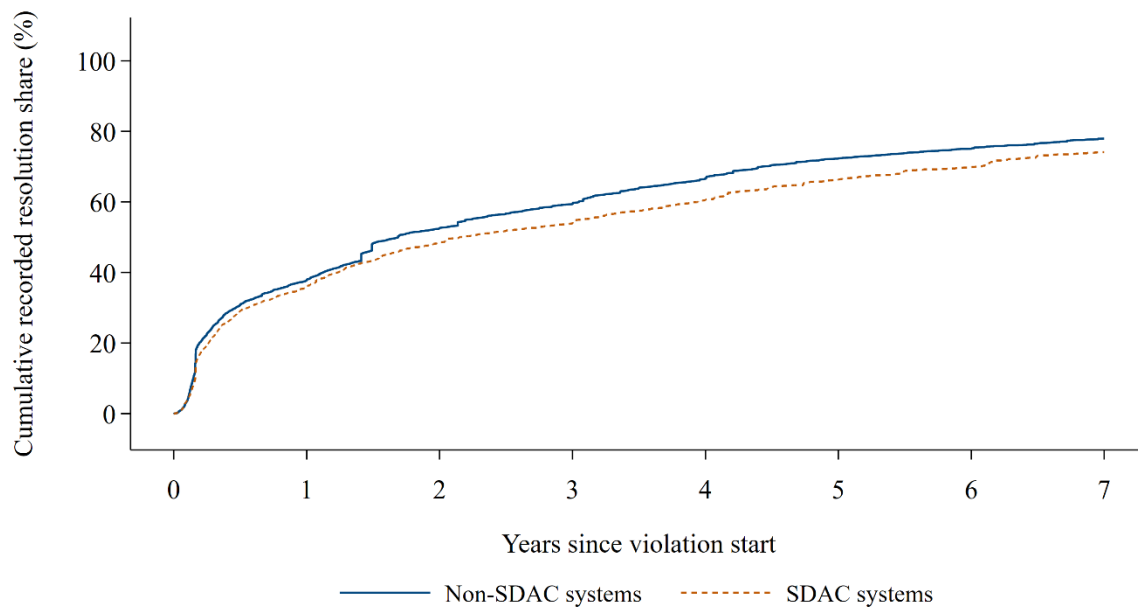


Figure 3.5: Kaplan-Meier Cumulative Recorded-Resolution Curves by SDAC Status

Note. Failure denotes the recorded-resolution event, so the curves plot the cumulative probability of recorded resolution over time as $1 - S(t)$. Results are descriptive and unadjusted. The figure visualizes overall differences in recorded-resolution timing between SDAC and non-SDAC systems before multivariable adjustment.

At the aggregate level, the difference is present but modest. SDAC systems show somewhat slower early closure, with narrower gaps at longer horizons. Figure 3.5 visualizes this overall pattern using Kaplan-Meier failure curves, where failure denotes the recorded-resolution event. This figure suggests some disparity in resolution timing, but not one that appears especially large or uniform across the duration distribution.

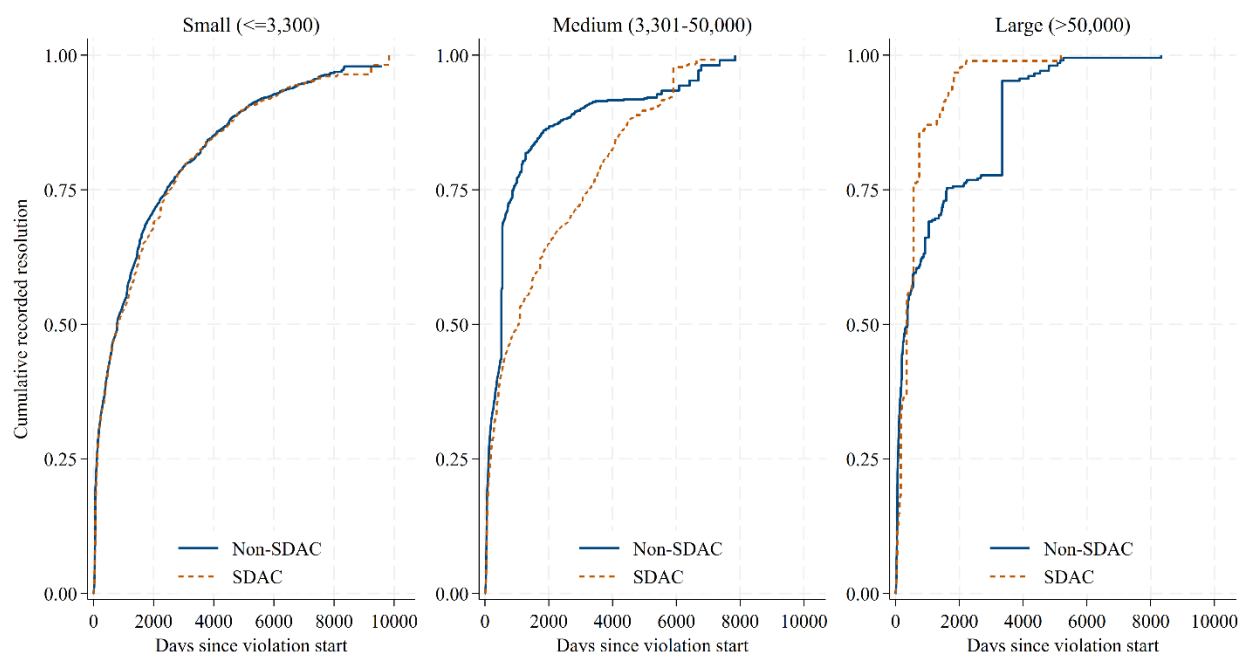


Figure 3.6: Kaplan-Meier Cumulative Recorded-Resolution Curves by System Size, with SDAC Overlay

Note. Panels show cumulative recorded-resolution curves separately by system-size category, with SDAC and non-SDAC systems overlaid within each panel. Failure denotes the recorded-resolution event, so the curves plot the cumulative probability of recorded resolution over time as $1 - S(t)$. Results are descriptive and unadjusted. The sample uses single violation spells and administrative follow-up through December 31, 2024.

A more revealing pattern appears once organizational scale is taken into account. The descriptive separation between SDAC and non-SDAC systems is not constant across system sizes. Instead, the clearest divergence appears in the medium-size category, while the SDAC and non-SDAC curves are more similar in other strata. Figure 3.6 shows this size-conditioned pattern directly, indicating that disadvantaged-community differences in return-to-compliance duration are shaped by organizational scale rather than captured by a single average SDAC effect. Appendix Table 3.A6 reports the corresponding Kaplan-Meier recorded-resolution shares at 90, 180, and 365 days, providing numeric support for the descriptive size-conditioned pattern.

These descriptive results motivate the multivariable strategy adopted in the next sections. Return-to-compliance duration is highly dispersed, observed resolution patterns are shaped by censoring and cohort structure, and SDAC and non-SDAC spells differ systematically in compliance history, system composition, and community context. Most importantly, the chapter's central environmental justice pattern lies in the interaction between SDAC status and system size. The baseline AFT models therefore assess whether this pattern persists after adjustment for community characteristics, system attributes, compliance history, violation type, and calendar time, while subsequent sensitivity analyses examine whether it remains after accounting for rule-family composition and broader temporal regimes.

3.5.2 Baseline Model

Table 3.2 presents results from the baseline log-logistic AFT model. The central finding is that return-to-compliance duration differs by disadvantaged-community status primarily through its interaction with system size rather than through a uniform average SDAC effect. In the large-system reference category, SDAC status is associated with faster recorded resolution relative to non-SDAC systems. This pattern changes across size strata. The positive and statistically significant SDAC interactions for small and medium systems indicate that the large-system SDAC advantage diminishes and reverses when SDAC status intersects with smaller organizational scale. To make this interaction more interpretable, Figure 3.7 presents post-estimation contrasts comparing small and medium SDAC systems with large SDAC systems. These contrasts show that, within SDAC systems, both small and medium systems have substantially longer expected recorded-resolution times than large systems, with the larger contrast appearing among medium-sized SDAC systems.

Table 3.2: Selected Adjusted Time Ratios from the Baseline Log-Logistic AFT Model

Characteristic	TR [95% CI]	<i>p</i> -value
Community context and size heterogeneity		
DAC system	1.042 [0.857, 1.268]	0.681
SDAC system, among large systems	0.296 [0.114, 0.770]	0.013**
Hispanic-serving system	0.979 [0.720, 1.331]	0.893
DAC × Hispanic-serving system	1.051 [0.744, 1.486]	0.778
African American-serving system	0.852 [0.710, 1.023]	0.087*
Small vs large, non-SDAC	0.877 [0.442, 1.741]	0.708
Medium vs large, non-SDAC	0.769 [0.375, 1.577]	0.473
Small system × SDAC interaction	3.141 [1.191, 8.278]	0.021**
Medium system × SDAC interaction	4.386 [1.467, 13.115]	0.008***
System and compliance context		
Log housing density, +1 unit	1.012 [0.969, 1.056]	0.594
Surface water source	1.190 [0.901, 1.571]	0.220
Purchased water source	0.867 [0.693, 1.085]	0.214
Private ownership	0.878 [0.733, 1.051]	0.155
Prior violations: 6-10	1.191 [0.968, 1.466]	0.097*
Prior violations: 11 or more	2.305 [1.752, 3.031]	<0.001***
Concurrent Tier 1 public-notification violation	2.021 [1.173, 3.484]	0.011**
Concurrent health-based violation	1.130 [0.982, 1.301]	0.089*
Health-based violation	0.948 [0.816, 1.100]	0.481

Note. Entries are selected adjusted time ratios from the baseline log-logistic accelerated failure time model. The unit of analysis is the violation spell, and the outcome is days from violation start to recorded resolution. A spell is treated as resolved only when the violation status is coded as Resolved on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as return-to-compliance events and are right-censored if no Resolved status is observed by the end of follow-up. Time ratios greater than 1 indicate longer time to recorded resolution, and time ratios below 1 indicate shorter time to recorded resolution. Large non-SDAC systems are the reference category for the SDAC-by-system-size terms. African American-serving and Hispanic-serving systems are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. The model includes rule-family indicators and violation start-year fixed effects, with standard errors clustered at the community water system level. The ancillary parameter is allowed to vary by system size. Appendix Table 3.A2 reports the full coefficient table. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

To make the AFT estimates more concrete, this chapter translated the baseline log-logistic model into adjusted predicted median times to recorded resolution. Under the baseline specification, the predicted median time to recorded resolution was about 539 days for large non-SDAC systems, 160 days for large SDAC systems, 439 days for small SDAC systems, and 538 days for medium SDAC systems. These adjusted durations clarify the substantive pattern. The results do not indicate a uniform SDAC penalty across all systems. The clearest disparity appears within SDAC systems by organizational scale. Relative to large SDAC systems, predicted time to recorded resolution was about 280 days longer for small SDAC systems and about 378 days longer for medium SDAC systems. Large SDAC systems also showed substantially shorter predicted time to recorded resolution than large non-SDAC systems, although that contrast was only marginally significant in the day-scale translation. This large-system contrast should therefore be interpreted cautiously. Because the translated day-scale difference is only marginally significant, the more defensible conclusion is not that large SDAC systems have a clear resolution advantage, but that the chapter does not show a robust large-system SDAC penalty once the comparison is expressed in predicted days.

Other demographic and system-level covariates show comparatively limited independent associations in the baseline model. DAC status, the Hispanic-serving indicator, and their interaction are not statistically distinguishable from the reference pattern once the full covariate set is included. The main effects for small and medium systems are also not significant after the SDAC interaction is taken into account, and housing density, water source, and private ownership do not show clear associations with resolution timing. These results show that the model's most important source of heterogeneity lies in the interaction between SDAC status and system size rather than in average demographic or system differences considered separately.

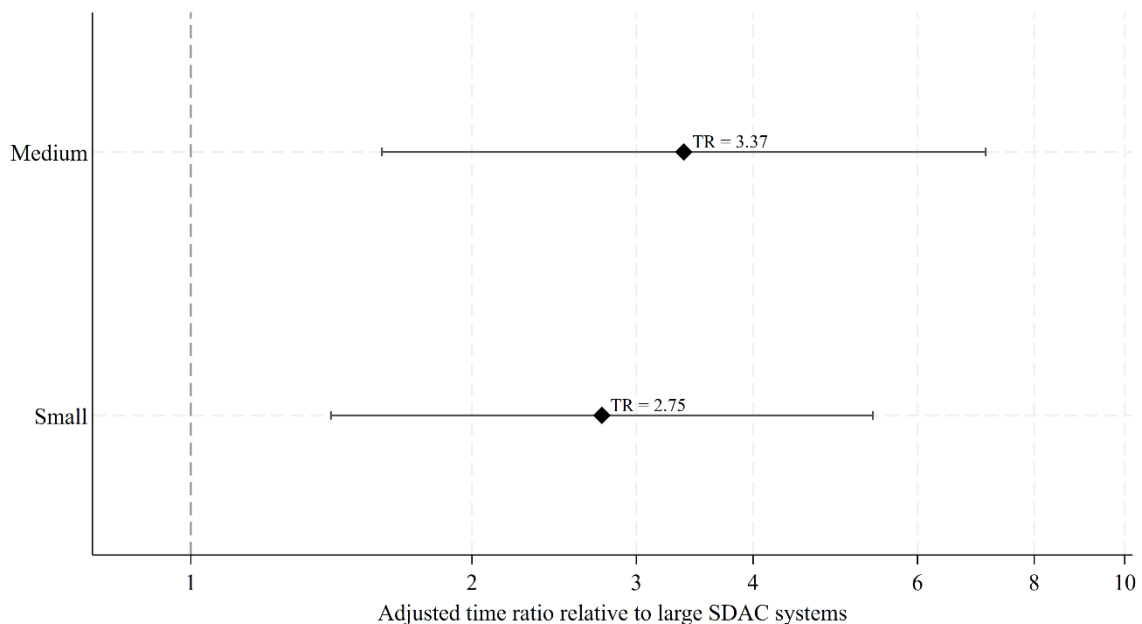


Figure 3.7: System-Size Contrasts Within SDAC Systems, Baseline Log-Logistic AFT Model

Note. Points show post-estimation adjusted time ratios and horizontal bars show 95% confidence intervals from the baseline log-logistic accelerated failure time model. Small and medium SDAC systems are compared with large SDAC systems. Time ratios greater than 1 indicate longer time to recorded resolution. The vertical dashed line marks no difference, TR = 1.

By contrast, the rest of the baseline model points to a different source of variation.

Measures of accumulated compliance burden are strongly associated with slower closure.

Systems with the heaviest prior violation histories experience substantially longer times to recorded resolution, and concurrent compliance challenges, especially concurrent Tier 1 public-notification violations, are also associated with delay. The focal health-based indicator itself is not significant once rule-family and concurrent-violation controls are included. At the same time, several rule families exhibit much faster resolution than the omitted reference category, indicating that violation composition is an important source of timing variation. Appendix Table 3.A2 reports the full baseline coefficient table. These pronounced rule-family differences motivate the next section, which examines whether the SDAC-by-size pattern persists after accounting more flexibly for rule-family composition.

3.5.3 Sensitivity to Rule-Family Composition

Descriptively, the distribution of major rule families differs across SDAC status and system-size categories, reinforcing the concern that raw contrasts may partly reflect differences in violation categories rather than recovery capacity alone. Appendix Figure 3.A5 summarizes these compositional patterns. Table 3.3 therefore tests whether the baseline SDAC-by-system-size pattern is sensitive to differences in rule-family composition. This specification extends the baseline log-logistic AFT model by allowing the SDAC and system-size gradients to vary across major rule families.

The central result remains the same. In the large-system reference category, SDAC status is associated with faster recorded resolution, but that pattern changes sharply across size strata. The interaction terms indicate that the SDAC association is substantially less favorable in smaller systems, with the clearest evidence of slower resolution again appearing among medium-sized SDAC systems. The core size-conditioned SDAC pattern therefore persists even after the model allows more flexible heterogeneity across violation types. Appendix Figure 3.A6 further

reports rule-standardized SDAC and SDAC-by-system-size contrasts under alternative weighting schemes, showing that the main pattern is not driven by a single rule-family weighting choice.

Table 3.3: Rule-Family Composition Sensitivity for Core SDAC and System-Size Time Ratios

Characteristic	TR [95% CI]	<i>p</i> -value
Core SDAC and system-size heterogeneity		
DAC system	1.039 [0.852, 1.267]	0.705
SDAC system, among large systems	0.107 [0.018, 0.627]	0.013**
Hispanic-serving system	0.898 [0.641, 1.258]	0.532
DAC × Hispanic-serving system	1.102 [0.750, 1.620]	0.621
African American-serving system	0.867 [0.734, 1.025]	0.095*
Small vs large, non-SDAC	0.459 [0.082, 2.572]	0.376
Medium vs large, non-SDAC	0.348 [0.054, 2.223]	0.265
Small system × SDAC interaction	8.504 [1.324, 54.614]	0.024**
Medium system × SDAC interaction	19.534 [2.716, 140.472]	0.003***
Compliance burden and concurrent violations		
Prior violations: 11 or more	2.193 [1.638, 2.935]	<0.001***
Concurrent Tier 1 public-notification violation	1.137 [0.995, 1.300]	0.059*
Concurrent health-based violation	1.719 [1.047, 2.822]	0.032**

Note. Entries are selected adjusted time ratios from a pooled log-logistic accelerated failure time model that allows SDAC and system-size associations to vary across major rule families. The unit of analysis is the violation spell, and the outcome is days from violation start to recorded resolution. A spell is treated as resolved only when the violation status is coded as Resolved on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as return-to-compliance events and are right-censored if no Resolved status is observed by the end of follow-up. Time ratios greater than 1 indicate longer time to recorded resolution, and time ratios below 1 indicate shorter time to recorded resolution. Large non-SDAC systems are the reference category for the SDAC-by-system-size terms. African American-serving and Hispanic-serving systems are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. The model retains the baseline covariate set, includes start-year fixed effects, clusters standard errors at the community water system level, and allows the ancillary parameter to vary by system size. Appendix Table 3.A3 reports the expanded coefficient table, including rule-family terms and rule-family-by-SDAC interactions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Most other demographic and system-level results remain secondary in this specification.

DAC status, the Hispanic-serving indicator, and their interaction show little evidence of an

independent association with resolution timing once the full covariate set is included. Housing density, water source, and private ownership likewise remain weak predictors relative to the SDAC-by-size interaction. This stability implies that the chapter's main environmental justice finding is not an artifact of simpler model specification. Instead, it continues to emerge even when the model allows rule-family-specific differences in how SDAC status and system scale relate to resolution timing.

At the same time, the rule-family interactions confirm that resolution timing differs meaningfully across violation types. Several rule-family terms and rule-by-SDAC interaction terms are large, indicating that the relationship between community disadvantage and recorded resolution is not constant across all categories of violations. Measures of accumulated compliance burden also remain important, with chronic prior violation history and concurrent compliance challenges continuing to predict slower closure. Appendix Table 3.A3 reports the expanded coefficient set for this pooled rule-family specification.

These results indicate that rule-family composition affects the magnitude of particular contrasts but does not overturn the chapter's central conclusion that SDAC-related differences in return-to-compliance duration are conditioned by system size. The final subsection broadens the robustness analysis by examining whether that pattern remains stable under alternative distributional assumptions, follow-up windows, outcome definitions, and temporal controls.

3.5.4 Additional Robustness and Temporal Sensitivity Checks

Table 3.4 shows that the chapter's central finding remains stable across the main robustness exercises. Across checks, the SDAC contrast is not uniform across system sizes. The clearest evidence of slower resolution continues to appear among small and medium SDAC systems relative to large SDAC systems. The magnitude of individual contrasts changes across

specifications, but the overall pattern remains consistent with the baseline interpretation that disadvantaged-community differences in return-to-compliance duration are conditioned by organizational scale rather than captured by a single average SDAC effect.

Table 3.4: Robustness Checks for Key SDAC-by-System-Size Time-Ratio Contrasts

Contrast	Baseline log-logistic	Weibull AFT	Pooled rules +winsorized density	Start-epoch fixed effects	180-day resolved-only event
Large SDAC vs large non-SDAC systems	0.296**	0.263**	0.297**	0.293**	0.471***
Small vs large systems within SDAC	2.755***	3.189**	3.131**	3.216**	2.072***
Medium vs large systems within SDAC	3.372***	3.936**	4.366***	4.500**	2.373***

Note. Entries report focal post-estimation time-ratio contrasts from the baseline and sensitivity specifications. Values greater than 1 indicate longer time to recorded resolution relative to the stated reference group, and values below 1 indicate shorter time to recorded resolution. The pooled-rule-family specification combines rare rule families and replaces log housing density with a winsorized measure. The start-epoch specification replaces violation start-year fixed effects with broader start-period fixed effects. The 180-day specification treats only Resolved status observed within 180 days as the event and censors all spells at day 180. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.5.4.1 Distributional Sensitivity

The first check evaluates whether the main result depends on the choice of AFT distribution. Appendix Table 3.A4 reports the full Weibull specification, which preserves the baseline covariate structure while replacing the log-logistic duration family. The key contrasts remain substantively similar under this alternative distribution. This result shows that the baseline SDAC-by-system-size pattern is not an artifact of the log-logistic functional form. The

central heterogeneity identified in the main model reflects a stable feature of the data rather than a distribution-specific modeling choice.

3.5.4.2 Sensitivity to follow-up window and outcome definition

This set of checks examines whether the findings depend on how the observation window or resolution outcome is defined. Appendix Table 3.A5 re-estimates system-size contrasts under an earlier administrative censoring date, while Appendix Table 3.A7 considers an alternative specification that defines the event within a shorter 180-day horizon. These checks address two related concerns, whether the main results are driven by the longest right-tail spells that remain unresolved near the end of follow-up, and whether the observed SDAC pattern depends on the specific resolution horizon embedded in the main event definition.

The results remain broadly consistent with the baseline interpretation. Although some estimates become less precise or change in magnitude, the central pattern of size-conditioned SDAC heterogeneity remains. This stability shows that the chapter's main finding is not a byproduct of the December 2024 administrative cap or of very long-duration resolution spells. Appendix Table 3.A6 complements these model-based checks by reporting Kaplan-Meier recorded-resolution shares at selected horizons by SDAC status and system size, providing a nonparametric descriptive view of the same timing pattern.

3.5.4.3 Sensitivity to temporal controls and related specification choices

The final set of checks examines whether the baseline findings are sensitive to broader specification choices, particularly the treatment of calendar time and the handling of less common rule categories or influential continuous covariates. In place of start-year fixed effects, one alternative specification uses broader policy-epoch indicators, testing whether the baseline results depend on fine-grained year controls rather than on more general institutional periods.

Additional checks pool rare rule families and use a winsorized version of housing density to assess whether the baseline findings are driven by sparse rule categories or extreme covariate values.

These alternative specifications do not alter the chapter's central conclusion. The SDAC-by-system-size pattern remains visible after replacing year fixed effects with epoch controls and after making modest adjustments to rule-family and density specification. The robustness tests suggest that the main result is not narrowly tied to a single AFT distribution, a particular follow-up window, or one temporal-control strategy. Instead, the evidence consistently points to the same substantive conclusion that disadvantaged-community differences in return-to-compliance duration are most pronounced where SDAC status intersects with smaller organizational scale.

3.6 Discussion

3.6.1 Interpretation and Implications

One implication of this chapter is conceptual. Because the outcome captures recorded resolution, return-to-compliance duration should be understood as an institutional recovery measure. Protection is restored only when unsafe or noncompliant conditions are corrected and formally resolved in the regulatory record. From this perspective, duration is not a secondary administrative detail. It is part of how regulatory protection is delivered.

A second implication is that the chapter's main environmental justice pattern is conditional rather than uniform. The baseline and sensitivity analyses do not show a single average SDAC penalty across all systems. Nor do they show a clear additional SDAC penalty among the smallest systems, where capacity constraints are likely common across SDAC and non-SDAC systems alike. Instead, the clearest adjusted pattern appears within SDAC systems by organizational scale: large SDAC systems have shorter predicted times to recorded resolution

than small and medium SDAC systems. This pattern is more consistent with a capacity-conditioned account of recovery than with a simple demographic-disadvantage effect.

Disadvantage-related differences in recorded resolution are most visible when community disadvantage intersects with limited organizational scale, while the smallest systems may face substantial technical, managerial, and financial constraints regardless of SDAC status.

The day-scale translation of the baseline model shows that the estimated disparities are large in practical terms, but the magnitudes require careful interpretation. These are adjusted predicted durations for selected system-by-SDAC profiles, not raw medians for all spells, and they measure time to recorded resolution rather than engineering remediation alone. In the adjusted predictions, time to recorded resolution for small and medium SDAC systems is roughly 280 to 378 days longer than for large SDAC systems. This comparison should be interpreted as a size contrast within SDAC systems, not as evidence that small SDAC systems are clearly slower than small non-SDAC systems. Interpreted through the chapter's restored-protection framework, differences of that magnitude indicate substantial variation in recovery capacity across organizational scale within disadvantaged-community contexts.

These findings have direct policy implications for California's drinking-water governance. California's human right to water framework recognizes a right to safe, clean, affordable, and accessible water, and the SAFER program reflects an institutional recognition that restoring safe drinking water often requires both immediate protection and long-term system sustainability (State Water Board, 2026c, 2026e). As summarized by Petek (2020), SAFER and the Safe Drinking Water Fund were designed to help systems provide an adequate and affordable supply of safe drinking water in both the near and long term, with support available for long-term solutions, consolidations, operations and maintenance, emergency water supplies, appointed

administrators, and administrative costs. Read alongside those policy commitments, the results here suggest that prolonged return-to-compliance duration should be treated as an implementation signal for where assistance and oversight need to be concentrated most urgently, especially in smaller SDAC systems facing repeated or overlapping compliance burdens.

The chapter also speaks to enforcement design. EPA's Public Water System Supervision Enforcement Response Policy emphasizes timely and appropriate responses aimed at returning systems to compliance, while California's enforcement framework describes a progressive approach that begins with less formal measures and escalates when necessary (EPA, 2009; State Water Board, 2024a, 2024b). Read against those frameworks, the results suggest that duration to recorded resolution should be treated as a core performance concern, not only violation occurrence or enforcement initiation. In practical terms, agencies should monitor which systems violate, which systems remain unresolved the longest, and whether long-duration cases are concentrated in the disadvantaged and capacity-constrained settings identified in this chapter. A duration-focused view of enforcement would therefore make follow-through, verification, and closure timing part of equity-oriented regulatory performance.

A related implication is that enforcement alone is unlikely to be sufficient where prolonged noncompliance reflects deeper recovery constraints. The chapter's burden variables and concurrent-violation patterns suggest that long time to recorded resolution often arises from cumulative strain rather than from the focal violation in isolation. This pattern supports coupling enforcement escalation with capacity-building responses, while also recognizing that durable recovery may require changes to the system's organizational structure. For small or medium SDAC systems facing repeated or overlapping compliance burdens, relevant structural solutions include physical consolidation, extension of service from a more capable neighboring system,

regional partnerships, shared management, local resource sharing, and appointed administrative support. California's drinking-water partnership framework identifies related tools, including resource sharing, shared operators and management services, physical consolidation, regionalization, and managerial consolidation (State Water Board, 2026a). These tools can expand the technical, managerial, and financial capacity needed to diagnose problems, implement corrective action, and complete formal closeout. They also align with California's current policy framework, which links funding incentives, SAFER support, and mandatory consolidation or extension-of-service pathways for systems that fail, or are at risk of failing, to provide safe drinking water (State Water Board, 2026b, 2026d, 2026e). The policy implication is not that assistance should replace enforcement. Rather, progressive enforcement is most equitable when paired with mechanisms that address the feasibility of durable correction in the systems most likely to remain unresolved.

A further implication is that external support is not automatically sufficient when internal implementation capacity is weak. Even where financing, technical assistance, emergency support, or consolidation planning is available, systems may lack the trained personnel, managerial continuity, and administrative capacity needed to use those resources effectively. Recovery can therefore be slowed not only by the absence of external resources, but also by limited absorptive capacity within the system itself. This helps explain why prolonged noncompliance may persist even in a policy environment with substantial assistance, and why capacity-building must include the organizational capability to identify appropriate remedies, coordinate implementation, maintain documentation, and convert outside support into durable correction.

The findings also point to the importance of administrative cadence. Because the chapter's outcome reflects recorded resolution, delays in verification, documentation, or formal closeout can themselves become part of unequal protection in practice. This does not mean that every long-duration spell is an administrative artifact. It means that agencies concerned with equity should pay attention to case management and closeout timing, especially where disadvantaged communities face lengthy unresolved spells. More broadly, the chapter shows that environmental justice in drinking-water regulation is shaped not only by the distribution of violations and sanctions, but also by the speed with which institutions convert noncompliance episodes into verified resolution.

3.6.2 Limitations

This chapter should be interpreted in light of several limitations arising from the administrative data, the construction of the timing outcome, and the observational design. The estimates identify adjusted associations in return-to-compliance duration rather than causal effects of disadvantaged-community status, system size, or other covariates. The results therefore support an interpretation centered on unequal recovery patterns and differences in compliance burden or recovery capacity, but they do not establish that any single observed factor directly causes slower or faster return to compliance.

A first limitation concerns outcome measurement. Observed resolution reflects substantive correction of the underlying violation, but it also depends on verification, documentation, reporting cadence, and formal closeout. Given that all systems are observed within California's SDWA implementation framework, this chapter does not assume that systems face entirely different external regulatory rules or arbitrary administrative processes. Rather, the limitation is that the administrative record cannot fully separate the time required to

correct the underlying problem from the time required to document, verify, and record return to compliance. These within-system steps may still vary in practice because systems differ in technical, managerial, staffing, and administrative capacity.

In addition, regulatory oversight geography may also shape return-to-compliance timing. In California, drinking-water oversight is administered through the State Water Board's Division of Drinking Water (DDW) district structure, and some county Local Primacy Agencies (LPA) regulate small public water systems with 200 or fewer service connections. Because the analytic file does not include a consistently validated DDW district or LPA oversight identifier, this chapter does not include regulatory-office fixed effects. Future work could examine whether recovery duration varies across DDW districts or LPA jurisdictions.

A second limitation follows from censoring and the finite observation window. Violations unresolved by December 31, 2024 are right-censored, and more recent cohorts necessarily have shorter follow-up than earlier cohorts. Although the survival framework is designed to accommodate censoring, administrative censoring still constrains what can be learned from later cohorts, especially for long-duration spells. The results should therefore not be interpreted as if all cohorts had equal opportunity to reach closure.

A third limitation concerns unobserved mechanisms. The chapter interprets differences in duration as consistent with variation in compliance recovery capacity, cumulative burden, and administrative follow-through, but the data do not directly observe several processes that likely matter, including financing constraints, technical assistance intensity, staffing capacity, treatment complexity, or the timing and intensity of regulator engagement. These processes are only partially proxied through system characteristics, compliance history, ownership, rule family, and

concurrent violations. As a result, the chapter identifies where unequal recovery is concentrated more clearly than it identifies the precise mechanisms that produce it.

A fourth limitation concerns measurement of community context. The chapter's demographic indicators, including SDAC and related community measures, are system-level and therefore necessarily coarse. They capture broad contextual inequality but may mask substantial heterogeneity within service areas. They also do not fully represent all dimensions of disadvantage relevant to compliance recovery, such as language isolation, local political capacity, housing precarity, or cumulative infrastructure stress.

A related limitation is that the analytic design compresses a complex compliance process into single terminal violation spells. This choice is appropriate for the chapter's central question because it avoids repeated counting of administrative updates within the same unresolved episode and produces a tractable survival framework. Nevertheless, it abstracts from intermediate stages such as partial correction, repeated regulator-system exchanges, and delay between substantive correction and formal closeout. The chapter therefore provides evidence on total recovery time rather than a full decomposition of the compliance process.

Finally, external validity is bounded by the chapter's institutional and geographic setting. The analysis focuses on California community water systems over a specific regulatory period, and both the administrative process and the meaning of disadvantaged-community designations are shaped by California's legal, fiscal, and governance context. The results are therefore most directly informative about unequal compliance recovery within this setting.

These limitations qualify the chapter's interpretation but do not diminish its central contribution. Even with observational and measurement constraints, the results identify a consistent pattern of unequal recovery in which return-to-compliance duration varies

systematically across systems and communities. The discussion therefore turns to what those patterns imply for environmental justice, drinking water governance, and policy design in California.

3.6.3 Future Directions

Several next steps follow from these findings. A first priority is to identify the mechanisms that link disadvantaged-community context and system size to longer recovery times. The current analysis shows where unequal recovery is concentrated, but it cannot determine whether the main bottlenecks arise from financing constraints, engineering complexity, limited managerial capacity, regulator-side closeout delays, or some combination of these factors. Future work should therefore connect administrative duration records to richer information on technical assistance, funding awards, consolidation efforts, infrastructure investments, and case-management timelines. These linked data would make it possible to evaluate whether technical and financial assistance programs shorten time to recorded resolution, and whether those effects differ across small, disadvantaged, or chronically noncompliant systems.

A second direction is to model the compliance process more explicitly over time. This chapter focuses on observed violation-spell duration to capture total time to recorded resolution, which is the most policy-relevant summary measure for the present purpose. Future research could build on that design by separating intermediate stages such as violation onset, enforcement escalation, corrective action, verification, and administrative closeout.

A related extension would be to model recurrent noncompliance more explicitly at the system level. This chapter is organized around observed violation spells and asks which spell, system, and community characteristics are associated with longer time to recorded resolution. A

recurrent-event survival design would support a different but complementary set of questions. For example, it could examine whether second and later spells within the same system tend to persist longer than earlier spells, whether recurrence patterns differ across system size or disadvantaged-community status, and whether some systems exhibit persistent unobserved vulnerability across successive spells. Such a design could also distinguish more clearly between cumulative compliance burden as a measured covariate and recurrence as a dynamic within-system process. That extension would be especially useful for studying whether chronic noncompliance reflects event-specific difficulty, durable system-level constraints, or both.

A third direction is to connect recovery duration more directly to community consequences. Prolonged noncompliance is important, but its significance ultimately depends on what it means for exposure burden, trust in public institutions, affordability pressures, and service reliability. Future work could therefore link duration measures to drinking water quality data, emergency notifications, household coping costs, or health-relevant exposure indicators.

Finally, the chapter points toward a broader research agenda on unequal protection under environmental regulation. This chapter suggests that environmental inequality in drinking water systems is not only about where risks are located, but also about how institutions respond over time and how quickly protection is restored once failures occur. It shows that inequality can emerge through differences in recovery time once noncompliance has begun. That insight complements the dissertation's other chapters, which examine unequal protection through exposure and health burden and through the timing of formal enforcement escalation.

3.7 Conclusion

This chapter examined return-to-compliance duration as a measure of how regulatory protection is restored in practice once drinking water violations occur. Using a survival

framework applied to California community water system violations, the analysis showed that resolution timing is uneven across systems and settings. The clearest pattern is not a uniform disadvantaged-community penalty across all cases. Instead, the results consistently indicate that differences in return-to-compliance duration are conditioned by organizational scale, with the strongest evidence of slower recorded resolution appearing where disadvantaged-community status intersects with smaller system size.

The chapter also showed that prolonged duration is tied to broader compliance context rather than to focal violation alone. Prior violation burden, concurrent compliance challenges, and rule-family composition are all closely associated with how long violations remain unresolved. These findings reinforce the chapter's central interpretation that return-to-compliance duration reflects a feasibility-based recovery process shaped by cumulative burden, institutional capacity, and administrative follow-through.

The chapter's main contribution is therefore both empirical and conceptual. Empirically, it identifies a recurring pattern of unequal recovery in which some communities remain in unresolved noncompliance for longer periods than others. Conceptually, it argues that duration itself is an important dimension of unequal protection. Safe drinking water governance is not only about whether standards are violated or whether agencies eventually act. It is also about whether communities receive timely restoration of compliance once a violation has been identified.

3.8 References

- Acquah, S., & Allaire, M. (2023). Disparities in drinking water quality: evidence from California. *Water Policy*, 25(2), 69–86.
- Allaire, M., & Acquah, S. (2022). Disparities in drinking water compliance: Implications for incorporating equity into regulatory practices. *AWWA Water Science*, 4(2), e1274.

- Allaire, M., Wu, H., & Lall, U. (2018). National trends in drinking water quality violations. *Proceedings of the National Academy of Sciences*, 115(9), 2078–2083.
- Allison, P. D. (2010). *Survival analysis using SAS: A practical guide (2nd ed.)*. SAS Institute.
- Andrews, R., & Boyne, G. A. (2010). Capacity, leadership, and organizational performance: Testing the black box model of public management. *Public Administration Review*, 70(3), 443–454.
- Association of State Drinking Water Administrators. (n.d.). *The public water system supervision (PWSS) program*. <https://www.asdwa.org/pwss/>
- Baker, J., Benneer, L., & Olmstead, S. (2023). Does information disclosure reduce drinking water violations in the United States? *Journal of the Association of Environmental and Resource Economists*, 10(3), 787–818.
- Balazs, C. L., & Ray, I. (2014). The drinking water disparities framework: On the origins and persistence of inequities in exposure. *American Journal of Public Health*, 104(4), 603–611.
- Benneer, L. S., & Olmstead, S. M. (2008). The impacts of the "right to know": Information disclosure and the violation of drinking water standards. *Journal of Environmental Economics and Management*, 56(2), 117–130.
- Benneer, L. S., Jessoe, K. K., & Olmstead, S. M. (2009). Sampling out: Regulatory avoidance and the total coliform rule. *Environmental Science & Technology*, 43(14), 5176–5182.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538.
- Blanchard, C. S., & Eberle, W. D. (2013). Technical, managerial, and financial capacity among small water systems. *Journal-American Water Works Association*, 105(5), E229-E235.
- Bourgeois, L. J., III. (1981). On the measurement of organizational slack. *Academy of Management Review*, 6(1), 29–39.
- Bovens, M., & Zouridis, S. (2002). From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control. *Public Administration Review*, 62(2), 174–184.
- Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling: A guide for social scientists*. Cambridge University Press.
- California Department of Water Resources. (2015, February 20). *Californians without safe water and sanitation: California Water Plan Update 2013*. https://www.waterboards.ca.gov/rwqcb3/water_issues/programs/enviro_justice/ej_docs/20_15ca_wo_safewater_san.pdf
- California State Water Resources Control Board. (2022). *California capacity development strategy for public water systems: 2022 revised draft strategy outline*. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/docs/2022-CapDev-Revised-Draft-Strategy.pdf

- California State Water Resources Control Board. (2024a). *2024 Water Code Section 13385(o) enforcement report*.
https://www.waterboards.ca.gov/publications_forms/publications/2024_13385report.html
- California State Water Resources Control Board. (2024b). *Water quality enforcement policy*. California Environmental Protection Agency.
https://www.waterboards.ca.gov/water_issues/programs/enforcement/docs/2024/2024-enforcement-policy.pdf
- California State Water Resources Control Board. (2025a, November 10). *Safe and Affordable Funding for Equity and Resilience (SAFER)*.
https://www.waterboards.ca.gov/water_issues/programs/grants_loans/sustainable_water_solutions/safer.html
- California State Water Resources Control Board. (2025b, July 23). *TMF – Capacity development*.
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/TMF.html
- California State Water Resources Control Board. (2026a, April 21). *Drinking water partnerships and consolidation*.
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/waterpartnership.html
- California State Water Resources Control Board. (2026b, January 7). *Funding and incentives for consolidation and regionalization projects*.
https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/fundingincentives.html
- California State Water Resources Control Board. (2026c, February 10). *Human right to water*.
https://www.waterboards.ca.gov/water_issues/programs/hr2w/
- California State Water Resources Control Board. (2026d, April 7). *Mandatory consolidation: Community and program info*.
https://www.waterboards.ca.gov/drinking_water/programs/compliance/
- California State Water Resources Control Board. (2026e, February 11). *SAFER drinking water*.
<https://www.waterboards.ca.gov/safer/>
- Carpenter, D. P. (2001). *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862–1928*. Princeton University Press.
- Chappelle, C., Hanak, E., & colleagues. (2021). *Water affordability in California*. PPIC Water Policy Center.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187–220.
- Cyert, R. M., & March, J. G. (2020). *A behavioral theory of the firm*. Barakaldo Books.
- Dal Bó, E. (2006). Regulatory capture: A review. *Oxford Review of Economic Policy*, 22(2), 203–225.
- Dobbin, K. B., & Fencl, A. L. (2021). Institutional diversity and safe drinking water provision in the United States. *Utilities Policy*, 73, 101306.

- Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57–74.
- Friesen, L. (2003). Targeting enforcement to improve compliance with environmental regulations. *Journal of Environmental Economics and Management*, 46(1), 72–85.
- Gray, W. B., & Shimshack, J. P. (2011). The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*, 5(1), 3–24.
- Grooms, K. K. (2016). Does water quality improve when a Safe Drinking Water Act violation is issued? A study of the effectiveness of the SDWA in California. *The BE Journal of Economic Analysis & Policy*, 16(1), 1–23.
- Gunningham, N., Kagan, R. A., & Thornton, D. (2003). *Shades of green: business, regulation, and environment*. Stanford University Press.
- Hanna-Attisha, M., LaChance, J., Sadler, R. C., & Champney Schnepf, A. (2016). Elevated blood lead levels in children associated with the Flint drinking water crisis: A spatial analysis of risk and public health response. *American Journal of Public Health*, 106(2), 283–290.
- Hendrick, R. (2004). Assessing and measuring the fiscal health of local governments: Focus on Chicago suburban municipalities. *Urban Affairs Review*, 40(1), 78–114.
- Herd, P., & Moynihan, D. P. (2018). *Administrative burden: Policymaking by other means*. Russell Sage Foundation.
- Hill, M., & Hupe, P. (2021). *Implementing public policy: An introduction to the study of operational governance (4th ed.)*. SAGE Publications.
- Kang, K., & Silveira, B. S. (2021). Understanding disparities in punishment: Regulator preferences and expertise. *Journal of Political Economy*, 129(10), 2947–2992.
- Kleinbaum, D. G., & Klein, M. (2012). *Survival analysis: A self-learning text (3rd ed.)*. Springer.
- Konisky, D. M., & Reenock, C. (2013). Compliance bias and environmental (in)justice. *The Journal of Politics*, 75(2), 506–519.
- Konisky, D. M., & Teodoro, M. P. (2016). When governments regulate governments. *American Journal of Political Science*, 60(3), 559–574.
- Legislative Analyst's Office. (2020, November). *Expanding access to safe and affordable drinking water in California: A status update*.
<https://lao.ca.gov/reports/2020/4294/drinking-water-access-111020.pdf>
- Maher, C. S., & Deller, S. C. (2007). Municipal responses to fiscal stress. *International Journal of Public Administration*, 30(12–14), 1549–1572.
- March, J. G., & Simon, H. A. (1958). *Organizations*. John Wiley & Sons.
- McDonald, Y. J., & Jones, N. E. (2018). Drinking water violations and environmental justice in the United States, 2011–2015. *American Journal of Public Health*, 108(10), 1401–1407.
- Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Georgetown University Press.

- Moynihan, D., Herd, P., & Harvey, H. (2015). Administrative burden: Learning, psychological, and compliance costs in citizen–state interactions. *Journal of Public Administration Research and Theory*, 25(1), 43–69.
- Nohria, N., & Gulati, R. (1996). Is slack good or bad for innovation? *Academy of Management Journal*, 39(5), 1245–1264.
- O’Toole, L. J., Jr., & Montjoy, R. S. (1984). Interorganizational policy implementation: A theoretical perspective. *Public Administration Review*, 44(6), 491–503.
- Patterson, L. A., & Doyle, M. W. (2021). Measuring water affordability and the financial capability of utilities. *AWWA Water Science*, 3(6), e1260.
- Petek, G. (2020, November). *Expanding access to safe and affordable drinking water in California: A status update* (Report No. 4294). Legislative Analyst’s Office. <https://lao.ca.gov/reports/2020/4294/drinking-water-access-111020.pdf>
- Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. Harper & Row.
- Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251–267.
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford University Press.
- Richardson, S. D., Plewa, M. J., Wagner, E. D., Schoeny, R., & DeMarini, D. M. (2007). Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. *Mutation Research/Reviews in Mutation Research*, 636(1-3), 178–242.
- Safe Drinking Water Act, 42 U.S.C. § 300g–9.
- Shimshack, J. P., & Ward, M. B. (2005). Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*, 50(3), 519–540.
- Shimshack, J. P. (2014). The economics of environmental monitoring and enforcement. *Annual Review of Resource Economics*, 6(1), 339–360.
- Smith, A. H., Lingas, E. O., & Rahman, M. (2000). Contamination of drinking-water by arsenic in Bangladesh: A public health emergency. *Bulletin of the World Health Organization*, 78(9), 1093–1103.
- Statman-Weil, Z., Nanus, L., & Wilkinson, N. (2020). Disparities in community water system compliance with the Safe Drinking Water Act. *Applied Geography*, 121, 102264.
- Steinmaus, C., Yuan, Y., Bates, M. N., & Smith, A. H. (2003). Case-control study of bladder cancer and drinking water arsenic in the western United States. *American Journal of Epidemiology*, 158(12), 1193–1201.
- Switzer, D., & Teodoro, M. P. (2017). The color of drinking water: Class, race, ethnicity, and Safe Drinking Water Act compliance. *Journal AWWA*, 109(9), 40–45.
- Switzer, D., & Teodoro, M. P. (2018). Class, race, ethnicity, and justice in Safe Drinking Water Act compliance. *Social Science Quarterly*, 99(2), 524–535.

- Terman, J. N., & Feiock, R. C. (2015). Improving outcomes in fiscal federalism: Local political leadership and administrative capacity. *Journal of Public Administration Research and Theory*, 25(4), 1059–1080.
- U.S. Census Bureau. (2024, December 16). *Urban and rural*. U.S. Department of Commerce. <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html>
- U.S. Department of Agriculture, Economic Research Service. (2025, January 7). *Rural-urban continuum codes: Documentation*. <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation>
- U.S. Environmental Protection Agency. (n.d.). *SDWA resources and FAQs*. ECHO: Enforcement and Compliance History Online. <https://echo.epa.gov/help/sdwa-faqs>
- U.S. Environmental Protection Agency. (1998). *Guidance on implementing the capacity development provisions of the Safe Drinking Water Act Amendments of 1996* (EPA 816-R-98-006). <https://19january2021snapshot.epa.gov/sites/static/files/2015-04/documents/epa816r98006.pdf>
- U.S. Environmental Protection Agency. (1999, July). *Handbook for capacity development: Developing water system capacity under the Safe Drinking Water Act as amended in 1996* (EPA 816-R-99-012). <https://www.epa.gov/sites/default/files/2015-04/documents/epa816r99012.pdf>
- U.S. Environmental Protection Agency. (2007a, December). *Analysis on the use of DWSRF set-asides: Promoting capacity development*. https://www.epa.gov/sites/default/files/2017-11/documents/2007_set-asides_analysis.pdf
- U.S. Environmental Protection Agency. (2007b). *Nitrates and nitrites: TEACH chemical summary*. https://archive.epa.gov/teach/web/pdf/nitrates_summary.pdf
- U.S. Environmental Protection Agency. (2009, December 8). *Enforcement Response Policy for the Public Water System Supervision (PWSS) Program under the Safe Drinking Water Act and implementation of the Enforcement Targeting Tool*. <https://www.epa.gov/sites/default/files/2015-09/documents/drinking-water-erp-2009.pdf>
- U.S. Environmental Protection Agency. (2013). *Providing safe drinking water in America: 2013 national public water systems compliance report* (EPA 305-R-15-001).
- U.S. Environmental Protection Agency. (2017, June 6). *Implementation of Drinking Water State Revolving Fund (DWSRF)-related Safe Drinking Water Act amendments in the Water Infrastructure Improvements for the Nation (WIIN) Act* (Memorandum). https://www.epa.gov/sites/default/files/2019-07/documents/implementation_of_dwsrf-related_sdwa_amendments_in_the_wiin_act.pdf
- U.S. Environmental Protection Agency. (2025a, December 23). *About the Drinking Water State Revolving Fund (DWSRF) set-asides*. <https://www.epa.gov/dwcapacity/about-drinking-water-state-revolving-fund-dwsrf-set-asides>
- U.S. Environmental Protection Agency. (2025b, August 20). *How the Drinking Water State Revolving Fund works*. <https://www.epa.gov/dwsrf/how-drinking-water-state-revolving-fund-works>

- U.S. Environmental Protection Agency. (2025c, September 9). *Learn about capacity development*. <https://www.epa.gov/dwcapacity/learn-about-capacity-development>
- U.S. Environmental Protection Agency. (2025d, June 11). *Public Notification Rule*. <https://www.epa.gov/dwreginfo/public-notification-rule>
- U.S. Environmental Protection Agency. (2025e, December 9). *SDWA Data Download Summary and Data Element Dictionary*. ECHO: Enforcement and Compliance History Online. <https://echo.epa.gov/tools/data-downloads/sdwa-download-summary#dataelement>
- U.S. Environmental Protection Agency. (2025f). *Stage 1 and Stage 2 Disinfectants and Disinfection Byproducts Rules*. <https://www.epa.gov/dwreginfo/stage-1-and-stage-2-disinfectants-and-disinfection-byproducts-rules>
- U.S. Environmental Protection Agency. (2025g, July 25). *Technical, managerial and financial (TMF) capacity resources for small drinking water systems*. <https://www.epa.gov/dwcapacity/technical-managerial-and-financial-tmf-capacity-resources-small-drinking-water-systems>
- U.S. Environmental Protection Agency. (2026a, March 11). *Building the capacity of drinking water systems*. <https://www.epa.gov/dwcapacity>
- U.S. Environmental Protection Agency. (2026b, March 16). *Chemical contaminant rules*. <https://www.epa.gov/dwreginfo/chemical-contaminant-rules>
- U.S. Environmental Protection Agency. (2026c, January 2). *Lead and Copper Rule*. <https://www.epa.gov/dwreginfo/lead-and-copper-rule>
- U.S. Environmental Protection Agency. (2026d, February 12). *Revised Total Coliform Rule and Total Coliform Rule*. <https://www.epa.gov/dwreginfo/revised-total-coliform-rule-and-total-coliform-rule>
- U.S. Environmental Protection Agency. (2026e, February 24). *Safe Drinking Water Act: Consumer Confidence Reports (CCR)*. <https://www.epa.gov/ccr>
- U.S. Environmental Protection Agency. (2026f, February 27). *SDWIS Federal Reporting Services data dictionary*. <https://www.epa.gov/DWdata/sdwis-federal-reporting-services-data-dictionary>
- U.S. Environmental Protection Agency. (2026g, April 20). *Training and technical assistance to improve water quality and enable small public water systems to provide safe drinking water*. <https://www.epa.gov/dwcapacity/training-and-technical-assistance-improve-water-quality-and-enable-small-public-water-0>
- U.S. Government Accountability Office. (2011, June 17). *Drinking water: Unreliable state data limit EPA's ability to target enforcement priorities and communicate water systems' performance* (GAO-11-381). <https://www.gao.gov/products/gao-11-381>
- Villanueva, C. M., Cantor, K. P., Cordier, S., Jaakkola, J. J. K., King, W. D., Lynch, C. F., Porru, S., & Kogevinas, M. (2004). Disinfection byproducts and bladder cancer: A pooled analysis. *Epidemiology*, 15(3), 357–367.

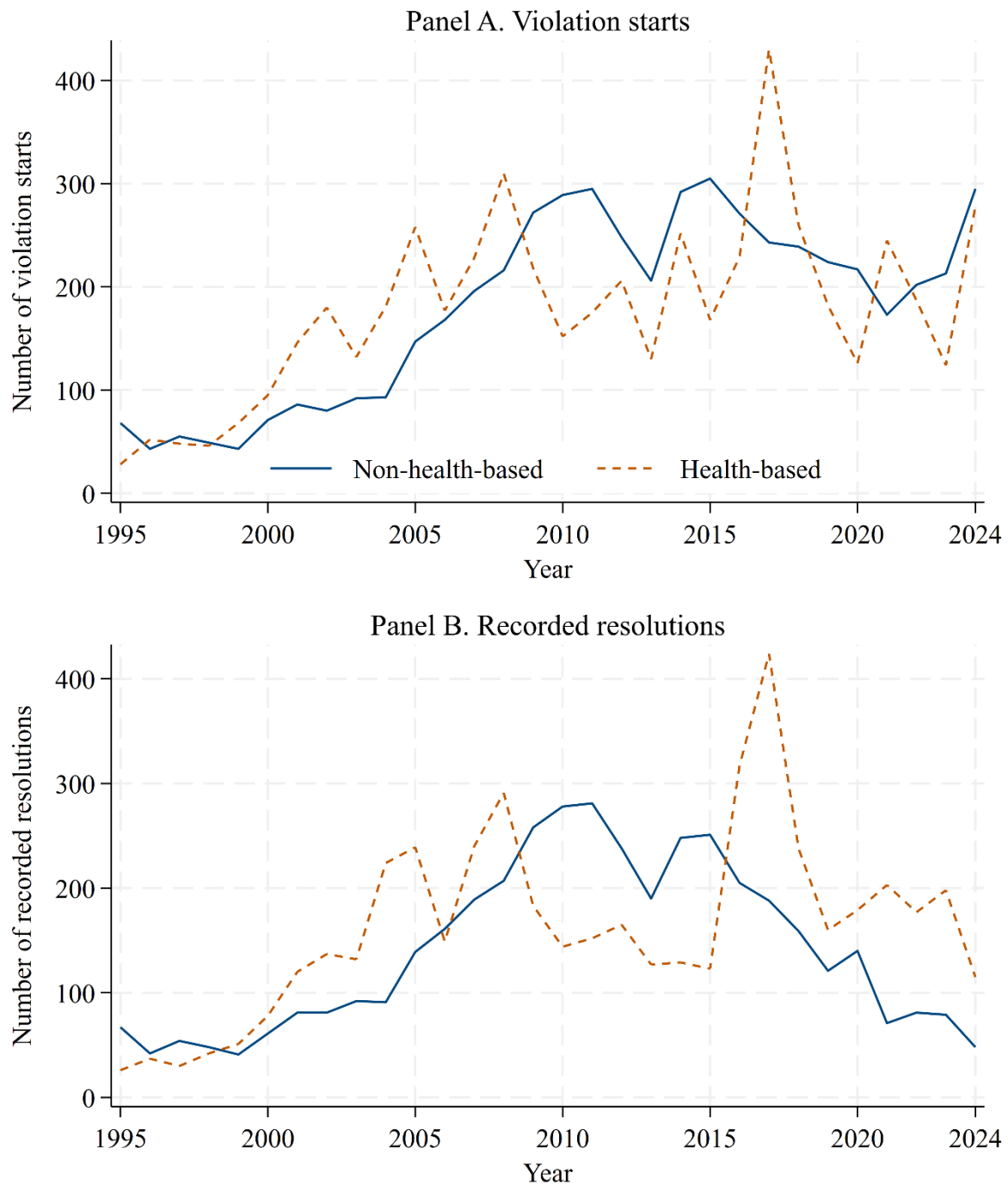
- Ward, M. H., Theo, M. D., Levallois, P., Brender, J., Gulis, G., Nolan, B. T., & VanDerslice, J. (2005). Workgroup report: drinking-water nitrate and health—recent findings and research needs. *Environmental health perspectives*, 113(11), 1607.
- Wallsten, S., & Kosec, K. (2008). The effects of ownership and benchmark competition: An empirical analysis of US water systems. *International Journal of Industrial Organization*, 26(1), 186–205.
- Warner, M. E., Aldag, A., & Kim, Y. (2021). Pragmatic municipalism: U.S. *local government responses to fiscal stress*. *Public Administration Review*, 81(3), 389–398.
- Wei, L. J. (1992). The accelerated failure time model: A useful alternative to the Cox regression model in survival analysis. *Statistics in Medicine*, 11(14–15), 1871–1879.
- Wu, X., Ramesh, M., & Howlett, M. (2015). Policy capacity: A conceptual framework for understanding policy competences and capabilities. *Policy and Society*, 34(3–4), 165–171.
- Zhao, D., & Lyon, T. P. (2022). *Why does general deterrence exist? A study of drinking water quality regulation*. *SSRN Electronic Journal*.

APPENDIX B: Chapter 3 Supplemental Tables and Figures

Appendix Table 3.A1: Violation Rule-Family Distribution by SDAC Status

Rule family	Non-SDAC		SDAC		Total	
	n	%	n	%	n	%
Total Coliform Rules	1,923	27.5	805	25.5	2,728	26.9
Nitrates	824	11.8	292	9.3	1,116	11.0
Arsenic	993	14.2	508	16.1	1,501	14.8
Disinfection Byproducts	1,033	14.8	658	20.8	1,691	16.7
Lead and Copper Rule	482	6.9	239	7.6	721	7.1
Public Notice and Consumer Confidence Rule	169	2.4	72	2.3	241	2.4
Others	1,564	22.4	582	18.4	2,146	21.2
Total	6,988	100.0	3,156	100.0	10,144	100.0

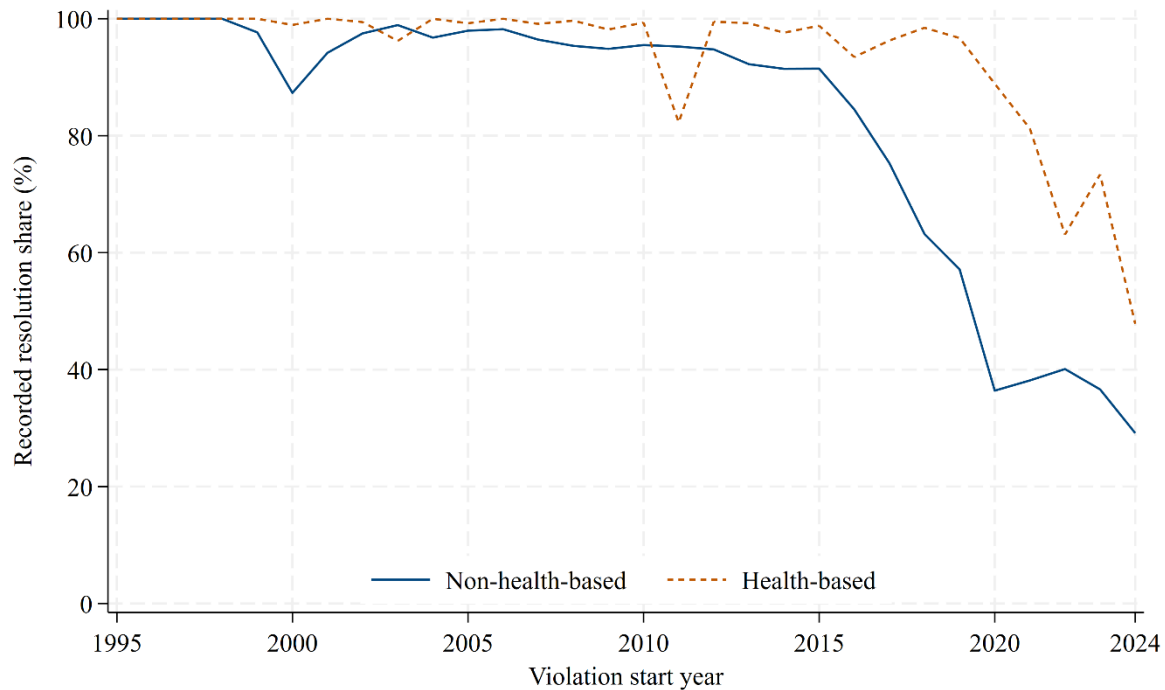
Note. Values are *n* and column percentages of violation spells within SDAC status. The table excludes 570 spells with missing rule-family code or SDAC status, leaving 10,144 spells for this descriptive comparison. A Pearson chi-square test indicates that rule-family composition differs by SDAC status, $\chi^2(6) = 86.77$, $p < 0.001$. SDAC = severely disadvantaged community.



Appendix Figure 3.A1: Violation Starts and Recorded Resolutions by Year, Health-Based Versus Non-Health-Based Violations

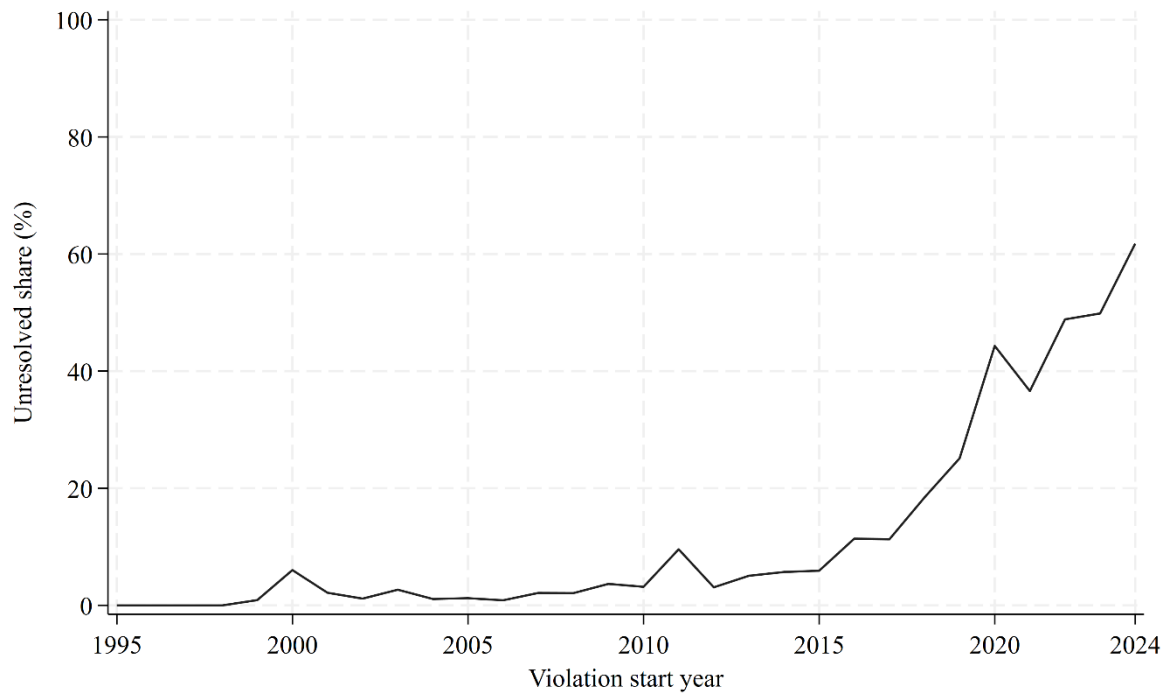
Note. Panel A plots annual counts of violation start with the violation-spell analytic sample. Panel B plots annual counts of recorded resolutions among violations that reached recorded resolution. The blue solid line represents non-health-based violations, and the orange dashed line

represents health-based violations. Health-based status is based on the violation-level health-based indicator. Administrative follow-up is censored at December 31, 2024.



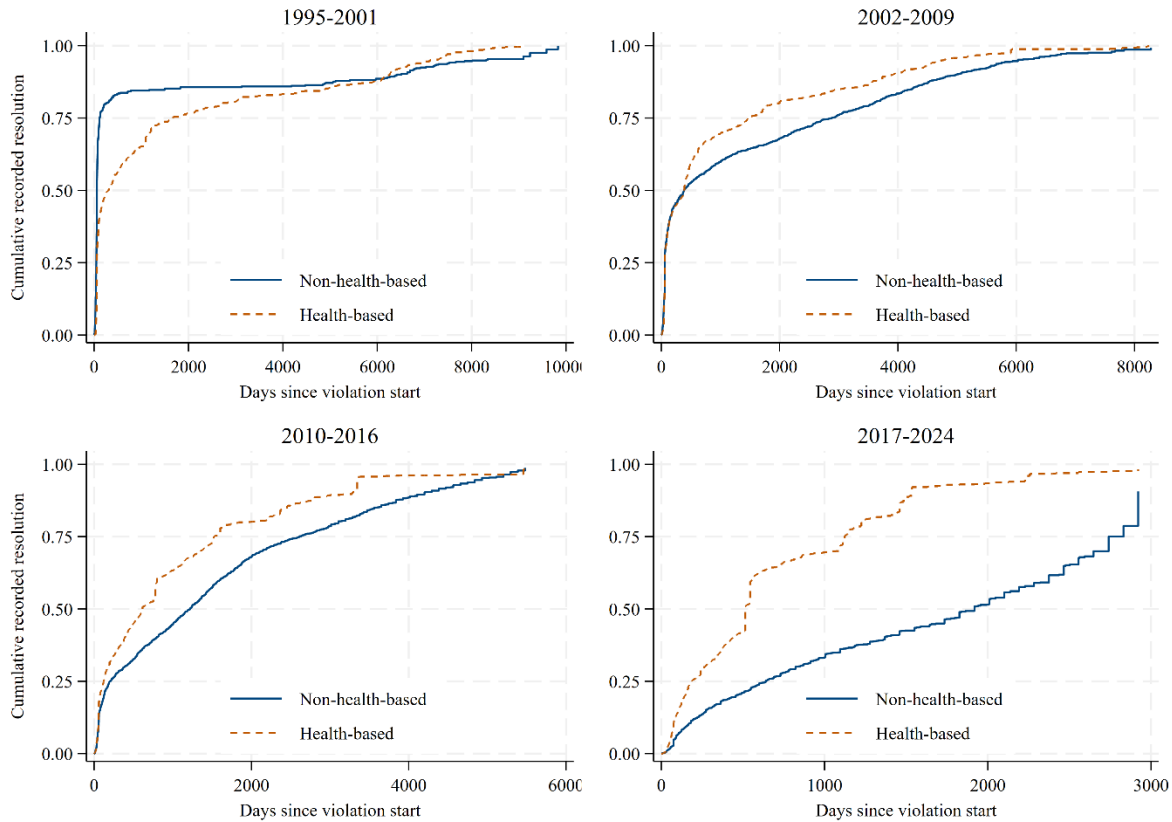
Appendix Figure 3.A2: Recorded Resolution Share by Violation Start Cohort, Health-Based Versus Non-Health-Based Violations

Note. The figure shows the share of each violation start-year cohort that reaches recorded resolution by December 31, 2024, separately for health-based and non-health-based violations. Lower recorded-resolution shares for more recent cohorts reflect shorter follow-up windows under right-censoring rather than necessarily lower long-run return-to-compliance probabilities.



Appendix Figure 3.A3: Unresolved Share by Violation Start Cohort

Note. The figure shows the share of each violation start-year cohort that remains unresolved by the December 31, 2024 administrative censoring date. Higher unresolved shares among recent cohorts primarily reflect shorter available follow-up time before censoring.



Appendix Figure 3.A4: Kaplan-Meier Cumulative Recorded-Resolution Curves by Policy Period and Health-Based Status

Note. Failure denotes the recorded-resolution event, so the curves plot the cumulative probability of recorded resolution over time as $1 - S(t)$. Panels group violations by violation start year. The blue solid line represents non-health-based violations, and the orange dashed line represents health-based violations. Results are descriptive and unadjusted, based on the violation-spell analytic sample with administrative follow-up censored at December 31, 2024.

Appendix Table 3.A2: Full Baseline Log-Logistic AFT Time-Ratio Table

Covariate	TR [95% CI]	<i>p</i> -value
Community context and size heterogeneity		
DAC system	1.042 [0.857, 1.268]	0.681
SDAC system, among large systems	0.296 [0.114, 0.770]	0.013**
Hispanic-serving system	0.979 [0.720, 1.331]	0.893
DAC × Hispanic-serving system	1.051 [0.744, 1.486]	0.778
African American-serving system	0.852 [0.710, 1.023]	0.087*

Covariate	TR [95% CI]	p-value
Small system, among non-SDAC systems	0.877 [0.442, 1.741]	0.708
Medium system, among non-SDAC systems	0.769 [0.375, 1.577]	0.473
Small system × SDAC interaction	3.141 [1.191, 8.278]	0.021**
Medium system × SDAC interaction	4.386 [1.467, 13.115]	0.008***
System and compliance context		
Log housing density, +1 unit	1.012 [0.969, 1.056]	0.594
Surface water source	1.190 [0.901, 1.571]	0.220
Purchased water source	0.867 [0.693, 1.085]	0.214
Private ownership	0.878 [0.733, 1.051]	0.155
Prior violations: 1-5	0.988 [0.844, 1.156]	0.877
Prior violations: 6-10	1.191 [0.968, 1.466]	0.097*
Prior violations: 11 or more	2.305 [1.752, 3.031]	<0.001***
Concurrent Tier 1 public-notification violation	2.021 [1.173, 3.484]	0.011**
Concurrent health-based violation	1.130 [0.982, 1.301]	0.089*
Health-based violation	0.948 [0.816, 1.100]	0.481
Rule-family indicators		
Rule: Total Coliform Rules	0.090 [0.071, 0.114]	<0.001***
Rule: Nitrates	0.434 [0.320, 0.587]	<0.001***
Rule: Arsenic	0.573 [0.386, 0.851]	<0.001***
Rule: Disinfection Byproducts	0.342 [0.249, 0.470]	<0.001***
Rule: Lead and Copper Rule	0.252 [0.185, 0.342]	<0.001***
Rule: Public Notice and Consumer Confidence Rule	0.149 [0.110, 0.202]	<0.001***

Note. Adjusted time ratios are from the full baseline log-logistic accelerated failure time model. The unit of analysis is the violation spell, and the outcome is days from violation start to recorded resolution. A spell is treated as resolved only when the violation status is coded as Resolved on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as return-to-compliance events and are right-censored if no Resolved status is observed by the end of follow-up. Time ratios greater than 1 indicate longer time to recorded resolution, and time ratios below 1 indicate shorter time to recorded resolution. Large non-SDAC systems are the reference category for the SDAC-by-system-size terms. African American-serving and Hispanic-serving systems are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. The model includes violation start-year fixed effects, grouped rule-family indicators, and a size-specific

ancillary parameter, with standard errors clustered at the community water system level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Appendix Table 3.A3: Expanded Pooled Rule-Family Log-Logistic AFT Time-Ratio Table

Covariate	TR [95% CI]	<i>p</i> -value
Community context and system-size heterogeneity		
DAC system	1.039 [0.852, 1.267]	0.705
SDAC system, among large systems	0.107 [0.018, 0.627]	0.013**
Hispanic-serving system	0.898 [0.641, 1.258]	0.532
DAC × Hispanic-serving system	1.102 [0.750, 1.620]	0.621
African American-serving system	0.867 [0.734, 1.025]	0.095*
Small vs large, non-SDAC	0.459 [0.082, 2.572]	0.376
Medium vs large, non-SDAC	0.348 [0.054, 2.223]	0.265
Small system × SDAC interaction	8.504 [1.324, 54.614]	0.024**
Medium system × SDAC interaction	19.534 [2.716, 140.472]	0.003***
System and compliance context		
Log housing density, +1 unit	1.016 [0.974, 1.059]	0.463
Surface water source	1.141 [0.883, 1.472]	0.313
Purchased water source	0.837 [0.643, 1.089]	0.185
Private ownership	0.852 [0.702, 1.034]	0.105
Prior violations: 1-5	0.907 [0.763, 1.079]	0.272
Prior violations: 6-10	1.123 [0.889, 1.418]	0.332
Prior violations: 11 or more	2.193 [1.638, 2.935]	<0.001***
Concurrent Tier 1 public-notification violation	1.137 [0.995, 1.300]	0.059*
Concurrent health-based violation	1.719 [1.047, 2.822]	0.032**
Rule-family terms		
Rule: Total Coliform Rules	0.051 [0.009, 0.277]	0.001***
Rule: Nitrates	0.186 [0.027, 1.278]	0.087*
Rule: Arsenic	0.208 [0.029, 1.476]	0.116
Rule: Disinfection Byproducts	0.166 [0.026, 1.074]	0.059*
Rule: Lead and Copper Rule	0.188 [0.031, 1.136]	0.069*
Health-based violation	0.968 [0.833, 1.124]	0.669
Rule-family interactions		
Rule: Total Coliform Rules × SDAC	12.396 [2.026, 75.851]	0.006***
Rule: Total Coliform Rules × SDAC × Small system	0.094 [0.014, 0.638]	0.016**

Covariate	TR [95% CI]	p-value
Rule: Total Coliform Rules × SDAC × Medium system	0.036 [0.005, 0.287]	0.002***
Rule: Total Coliform Rules × Small system	2.180 [0.383, 12.418]	0.380
Rule: Total Coliform Rules × Medium system	2.859 [0.447, 18.277]	0.267
Rule: Nitrates × SDAC	13.458 [1.031, 175.646]	0.047**
Rule: Nitrates × SDAC × Small system	0.060 [0.004, 0.950]	0.046**
Rule: Nitrates × SDAC × Medium system	0.031 [0.002, 0.515]	0.015**
Rule: Nitrates × Small system	3.641 [0.502, 26.425]	0.201
Rule: Nitrates × Medium system	2.321 [0.278, 19.368]	0.437
Rule: Arsenic × SDAC	4.379 [0.491, 39.073]	0.186
Rule: Arsenic × SDAC × Small system	0.155 [0.014, 1.733]	0.130
Rule: Arsenic × SDAC × Medium system	0.051 [0.004, 0.635]	0.021**
Rule: Arsenic × Small system	4.005 [0.516, 31.094]	0.184
Rule: Arsenic × Medium system	6.170 [0.707, 53.840]	0.100*
Rule: Disinfection Byproducts × SDAC	8.090 [1.175, 55.680]	0.034**
Rule: Disinfection Byproducts × SDAC × Small system	0.166 [0.021, 1.292]	0.086*
Rule: Disinfection Byproducts × SDAC × Medium system	0.171 [0.017, 1.753]	0.137
Rule: Disinfection Byproducts × Small system	2.084 [0.313, 13.852]	0.447
Rule: Disinfection Byproducts × Medium system	2.318 [0.289, 18.621]	0.429
Rule: Lead and Copper Rule × SDAC	81.423 [12.670, 523.248]	<0.001***
Rule: Lead and Copper Rule × SDAC × Small system	0.012 [0.002, 0.085]	<0.001***
Rule: Lead and Copper Rule × SDAC × Medium system	0.008 [0.001, 0.071]	<0.001***
Rule: Lead and Copper Rule × Small system	1.754 [0.282, 10.914]	0.547
Rule: Lead and Copper Rule × Medium system	1.526 [0.210, 11.075]	0.676

Note. Adjusted time ratios are from the expanded pooled log-logistic accelerated failure time model that interacts major rule families with SDAC status and system size. The unit of analysis is the violation spell, and the outcome is days from violation start to recorded resolution. A spell is treated as resolved only when the violation status is coded as Resolved on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as return-to-compliance events and are right-censored if no Resolved status is observed by the end of follow-up. Time ratios greater than 1 indicate longer time to recorded resolution, and time ratios below 1 indicate shorter time to recorded resolution. Large non-SDAC systems are the reference category for the SDAC-by-system-size terms. African American-serving and Hispanic-serving systems

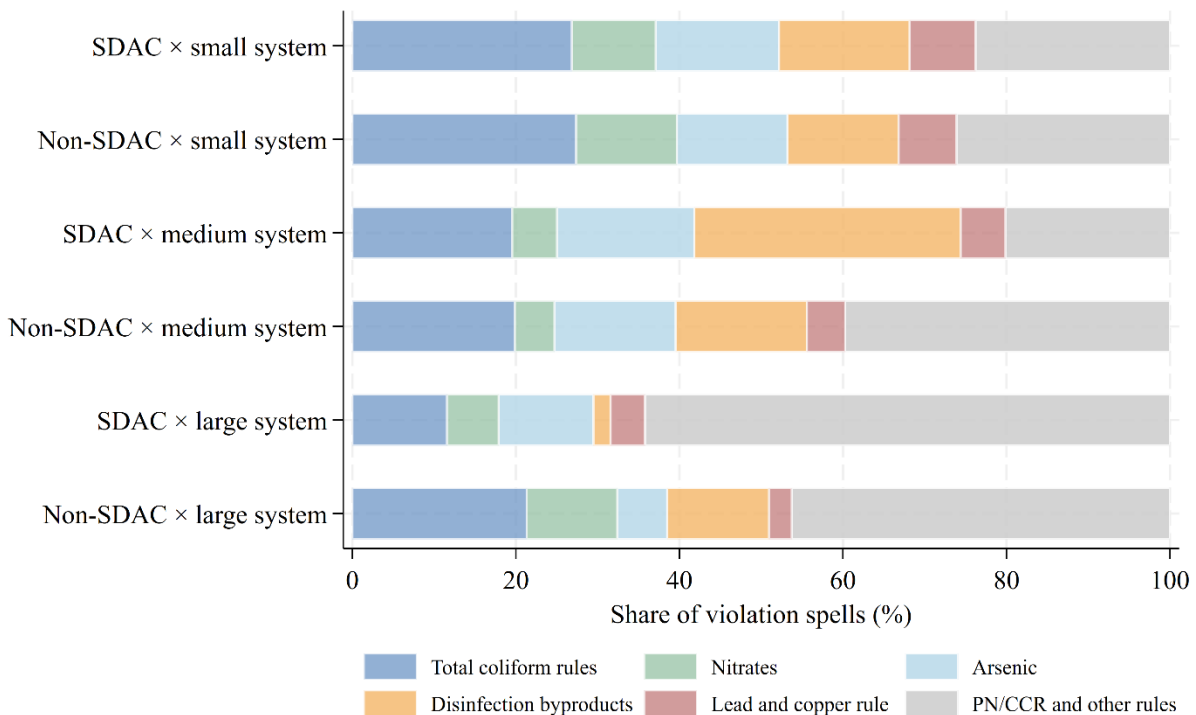
are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. The model retains the baseline covariate set, includes violation start-year fixed effects, and allows the ancillary parameter to vary by system size, with standard errors clustered at the community water system level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 3.A4: Full Weibull AFT Time-Ratio Table

Covariate	TR [95% CI]	p-value
Community context and size heterogeneity		
DAC system	1.041 [0.762, 1.420]	0.802
SDAC system, among large systems	0.263 [0.091, 0.758]	0.013**
Hispanic-serving system	0.995 [0.692, 1.430]	0.978
DAC × Hispanic-serving system	1.061 [0.683, 1.649]	0.792
African American-serving system	0.808 [0.602, 1.086]	0.157
Small system, among non-SDAC systems	0.569 [0.257, 1.259]	0.164
Medium system, among non-SDAC systems	0.556 [0.253, 1.222]	0.144
Small system × SDAC interaction	3.189 [1.098, 9.267]	0.033**
Medium system × SDAC interaction	3.936 [1.188, 13.035]	0.025**
System and compliance context		
Log housing density, +1 unit	0.974 [0.915, 1.037]	0.413
Surface water source	1.152 [0.812, 1.635]	0.428
Purchased water source	0.781 [0.573, 1.066]	0.119
Private ownership	0.960 [0.757, 1.216]	0.733
Prior violations: 1-5	0.959 [0.762, 1.208]	0.723
Prior violations: 6-10	1.317 [0.992, 1.748]	0.057*
Prior violations: 11 or more	2.442 [1.692, 3.526]	<0.001***
Concurrent Tier 1 public-notification violation	1.821 [0.967, 3.429]	0.063*
Concurrent health-based violation	1.345 [1.054, 1.717]	0.017**
Health-based violation	1.137 [0.903, 1.432]	0.275
Rule-family indicators		
Rule: Total Coliform Rules	0.119 [0.087, 0.162]	<0.001***
Rule: Nitrates	0.542 [0.386, 0.763]	<0.001***
Rule: Arsenic	0.635 [0.382, 1.057]	0.081*
Rule: Disinfection Byproducts	0.386 [0.273, 0.547]	<0.001***

Covariate	TR [95% CI]	p-value
Rule: Lead and Copper Rule	0.393 [0.286, 0.539]	<0.001***
Rule: Public Notice and Consumer Confidence Rule	0.171 [0.119, 0.247]	<0.001***

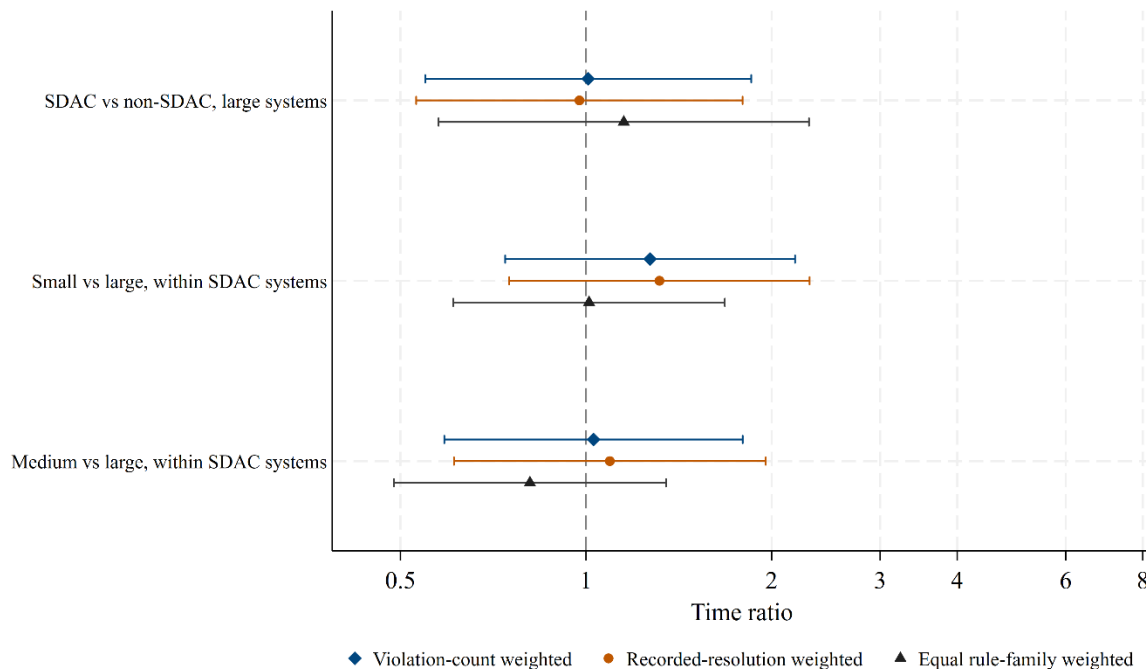
Note. Entries are adjusted time ratios from the Weibull accelerated failure time model estimated on the violation-spell analytic sample. The unit of analysis is the violation spell, and the outcome is days from violation start to recorded resolution. A spell is treated as resolved only when the violation status is coded as Resolved on or before December 31, 2024. Violations coded as Addressed, Archived, or Unaddressed are not treated as return-to-compliance events and are right-censored if no Resolved status is observed by the end of follow-up. Time ratios greater than 1 indicate longer time to recorded resolution, and values below 1 indicate shorter time to recorded resolution. Large non-SDAC systems are the reference category for the SDAC-by-system-size terms. African American-serving and Hispanic-serving systems are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. The specification mirrors the baseline covariate set, includes violation start-year fixed effects and grouped rule-family indicators, allows the ancillary parameter to vary by system size, and clusters standard errors at the community water system level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



Appendix Figure 3.A5: Rule-Family Mix by SDAC Status and System Size, 1995–2024

Note. Bars show the percentage distribution of violation spells across major rule families within each SDAC-by-system-size group. Percentages sum to 100 within each group. SDAC = severely

disadvantaged community. PN/CCR = Public Notice and Consumer Confidence Rule. System-size categories follow the chapter’s analytic classification.



Appendix Figure 3.A6: Rule-Standardized SDAC-by-System-Size Contrasts from the Log-Logistic AFT Model

Note. Points show estimated time ratios and horizontal bars show 95% confidence intervals. The vertical dashed line marks no difference, TR = 1. Weighting schemes standardize rule-family composition using violation-count weights, recorded-resolution weights, or equal weights across major rule families. Time ratios greater than 1 indicate longer time to recorded resolution. SDAC = severely disadvantaged community.

Appendix Table 3.A5: Earlier Administrative-Censoring-Date Sensitivity for System-Size Time-Ratio Contrasts

Contrast	TR, 2024 censoring date	TR, 2023 censoring date	% change	<i>p</i> -value, 2024	<i>p</i> -value, 2023
Small vs large systems within non-SDAC	0.877	0.856	-2.41	0.708	0.748

Medium vs large systems within non-SDAC	0.769	0.752	-2.24	0.473	0.567
Small vs large systems within SDAC	2.755	2.957	7.34	0.003***	0.002***
Medium vs large systems within SDAC	3.372	3.549	5.25	0.001***	0.001***
Resolved spells	9,143	8,726	—	—	—
Censored spells	1,571	1,988	—	—	—

Note. Entries compare focal post-estimation time-ratio contrasts under the baseline December 31, 2024 administrative censoring date and an earlier December 31, 2023 administrative censoring date. Values greater than 1 indicate longer time to recorded resolution relative to the stated reference group, and values below 1 indicate shorter time to recorded resolution. The specification retains the baseline log-logistic AFT structure and pools rare rule families. Positive percentage changes indicate larger time-ratio contrasts under the earlier censoring date. Resolved and censored spell counts are shown for each censoring date. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 3.A6: Kaplan-Meier Recorded-Resolution Shares by SDAC Status and System Size

SDAC status	System size	90 days	180 days	365 days
Non-SDAC	Small system	32.04%	42.31%	50.42%
Non-SDAC	Medium system	39.27%	48.72%	55.69%
Non-SDAC	Large system	38.94%	49.40%	59.92%
SDAC	Small system	32.22%	39.67%	49.64%
SDAC	Medium system	40.73%	52.56%	58.98%

SDAC	Large system	49.08%	59.38%	88.10%
------	--------------	--------	--------	--------

Note. Entries report Kaplan-Meier cumulative recorded-resolution shares at each horizon. The recorded-resolution event is defined by Resolved status in the administrative record. Values are descriptive and unadjusted, based on the violation-spell sample under the December 31, 2024 administrative censoring date. The table provides numeric support for the descriptive system-size patterns shown in Figure 3.6.

Appendix Table 3.A7: 180-Day Resolved-Only Event Sensitivity for Selected AFT Time-Ratio Contrasts

Contrast	TR	95% CI	<i>p</i> -value
Large SDAC vs large non-SDAC systems	0.471	[0.295, 0.751]	0.002***
Small vs large systems within SDAC	2.072	[1.287, 3.336]	0.003***
Medium vs large systems within SDAC	2.373	[1.416, 3.976]	0.001***

Note. The event is redefined as recorded resolution within 180 days, and spells without observed Resolved status by day 180 are censored at day 180. Entries report focal post-estimation time-ratio contrasts from the otherwise baseline log-logistic AFT specification. Time ratios greater than 1 indicate longer time to recorded resolution relative to the stated reference group, and values below 1 indicate shorter time to recorded resolution. This sensitivity check assesses whether the main findings are driven by very long right-tail resolution times. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 4: Formal Enforcement Escalation, Compliance Dynamics, and Unequal Protection in California Community Water Systems

4.1 Chapter Overview

This chapter shifts attention from recovery duration to regulatory response by asking whether a system's first formal enforcement action is followed by measurable improvement in compliance progress. The chapter uses the term formal enforcement in a specific administrative sense. Informal enforcement refers to regulatory communication and compliance management actions, such as warning letters, notices of violation, phone calls, meetings, and other prompts that notify a system of noncompliance or encourage correction but do not create the same enforceable compliance pathway. Formal enforcement refers to administrative actions that move a case into a more explicit accountability process by specifying required corrective steps, deadlines, and potential consequences for failing to return to compliance. First formal enforcement is therefore not treated simply as a punitive sanction. It is treated as the point at which unresolved noncompliance is formalized into a more structured pathway toward correction.

The chapter examines the governance process through which regulators seek to move persistent noncompliance toward restored protection. This question is analytically important because drinking-water noncompliance often reflects not only deterrence incentives, but also technical, managerial, and financial capacity constraints. Much of the environmental enforcement literature focuses on regulated entities that weigh compliance costs against the expected costs of detection and sanctions. Community water systems differ from that classic setting because they provide an essential public service and often face barriers related to

infrastructure, financing, staffing, engineering capacity, and administrative follow-through. SDWA enforcement therefore provides a useful setting for assessing whether escalation matters in a context where regulatory action may operate through structured obligations, case management, consultation, and accountability rather than punishment alone.

The question also has practical significance. Formal enforcement is one of the main tools regulators use when informal compliance management has not produced sufficient progress. Formal actions can require corrective steps, impose deadlines, increase documentation requirements, and focus regulatory attention on unresolved violations. Observed gains after escalation would suggest that this administrative threshold helps move systems toward compliance. Limited, delayed, or uneven gains would instead suggest that formal enforcement alone may be insufficient without additional technical, financial, or managerial support. This concern is especially important for disadvantaged communities, where the same enforcement action may be associated with weaker corrective momentum if underlying capacity constraints are more severe.

This empirical analysis therefore examines California community water systems' compliance trajectories around each system's first formal enforcement action. The outcome is quarterly compliance progress, measured through resolution activity among open violations in a system-quarter panel. The analysis first estimates whether compliance progress changes after formal escalation and then evaluates whether those changes differ across disadvantaged and non-disadvantaged community contexts. Through this design, the chapter complements the preceding analysis of noncompliance duration and evaluates whether formal enforcement is associated with comparable regulatory responsiveness across community systems.

4.2 Literature Review

4.2.1 Regulator-Regulatee Interaction and the Behavioral Logic of Enforcement

Environmental enforcement research begins from the premise that compliance is jointly produced by regulated entities' behavioral responses and regulators' monitoring and intervention decisions. In the enforcement economics tradition, compliance incentives depend on expected consequences, especially perceived detection risk and the severity of sanctions conditional on detection (Becker, 1968; Polinsky & Shavell, 2000; Rousseau, 2007). Regulators increase the expected cost of noncompliance through monitoring tools such as inspections and reporting requirements and through sanctions such as penalties and orders (Gray & Shimshack, 2011; Earnhart & Segerson, 2012). This relationship is dynamic. Regulators adjust monitoring and sanction intensity in response to observed performance, while regulated entities update expectations about regulatory attention and likely consequences based on prior enforcement activity and adjust their compliance behavior accordingly (Cory & Rahman, 2009; Arguedas & Rousseau, 2012). This interactional framework provides a behavioral foundation for why enforcement timing is important.

Regulators choose monitoring and penalty intensity under resource constraints and political and administrative costs, weighing expected compliance benefits against enforcement burdens and potential collateral harms (Rousseau, 2007; Gray & Shimshack, 2011; Breyer, 2021; Konisky & Reenock, 2015; Salzman & Mackay Peltzer, 2025). These tradeoffs imply selection in both the type and timing of enforcement actions. In addition to compliance history, regulators may prioritize cases based on violation and contaminant characteristics such as severity, category, persistence, and technical complexity, as these features shape public health risk, the feasibility of corrective action, and the expected returns to regulatory effort. Selective-

enforcement theory likewise suggests that regulators may avoid aggressive action where sanctions are politically costly or practically counterproductive, such as at facilities that are financially fragile or central to local service provision (Deily & Gray, 1991). In the drinking-water context, this logic extends to cases in which state primacy agencies must weigh affordability concerns and service continuity alongside formal compliance objectives. Oversight therefore unfolds through a sequence of decisions rather than a single intervention, and the implications of that sequence depend on the institutional setting in which compliance is pursued.

4.2.2 SDWA Implementation, Primacy Capacity, and Return to Compliance as the Administrative Objective

Empirical evidence of persistent and geographically clustered drinking water noncompliance suggests that SDWA enforcement operates as an ongoing governance process rather than as a simple technical pipeline (Allaire et al., 2018; Association of State Drinking Water Administrators [ASDWA], 2013). Although the compliance system is often described procedurally, with monitoring identifying violations that trigger corrective action, persistent noncompliance indicates that oversight also depends on discretionary decisions about follow-up, prioritization, and escalation under administrative constraints. SDWA oversight is organized around return to compliance, making the timing and progression of corrective action central to regulatory practice. Because primary implementation responsibility rests with state primacy agencies, regulatory success is often framed less in terms of punishment than in terms of whether systems move toward, and are documented as achieving, return to compliance (EPA, 2009c). This administrative orientation makes time substantively important by shaping whether regulators respond, how quickly systems move toward correction, and how oversight tools are sequenced as noncompliance persists.

Primacy capacity is therefore central to how enforcement unfolds in practice. State drinking water programs must allocate limited staffing, technical support, and legal resources across large and heterogeneous portfolios of systems, violations, and corrective needs. These constraints shape what follow-up is feasible, how aggressively agencies respond, and how quickly cases move through the compliance process. SDWA implementation is therefore not the mechanical application of fixed rules, but an administratively mediated process shaped by organizational capacity, competing demands, and case prioritization. In that context, the distinction between informal engagement and formal enforcement becomes especially important, because formalization marks a shift in how the compliance pathway is structured, monitored, and enforced.

4.2.3 First Formal Enforcement as a Structured Accountability Threshold

SDWA oversight operates through a sequenced toolkit rather than a single intervention. Enforcement scholarship often describes this progression as an enforcement ladder, in which initial violations may prompt informal action while repeated or persistent noncompliance can trigger more formal responses (Earnhart & Jacobson, 2025). In the SDWA context, EPA's enforcement framework distinguishes informal engagement from formal enforcement by defining formal actions as independently enforceable instruments that specify what is required for return to compliance, establish an enforceable compliance schedule, and carry consequences if deadlines are not met (EPA, 2009c). Empirical drinking water research likewise suggests that systems respond differently to informal and formal interventions, with the two associated with distinct violation-duration patterns (Kirchhoff et al., 2019).

First formal enforcement is analytically meaningful because it changes the administrative status of unresolved noncompliance. Before this point, regulators may rely on warnings,

communication, meetings, and other informal efforts to encourage correction. Once a formal action is issued, the case is tied to more explicit requirements, deadlines, documentation expectations, and follow-up authority (EPA, 2009c). The mechanism is therefore not simply that the regulator signals concern, but that the system's obligations become more specified and more readily trackable in the administrative process. (Baldwin & Black, 2008).

This shift remains significant even when monetary penalties are limited or when regulators combine formal orders with consultation and compliance support. Its importance lies not only in punishment, but also in how formalization makes noncompliance more salient for system decision-makers, regulators, and affected communities. In many SDWA cases, return to compliance requires investments and operational changes that are costly, technically complex, and not immediately feasible. Under these conditions, formal enforcement may shape compliance progress by clarifying obligations, imposing timelines, focusing regulatory attention, and making continued delay more visible, rather than relying primarily on financial sanctions. Deterrence theory similarly emphasizes that compliance responses depend not only on penalty severity, but also on the perceived certainty and timing of enforcement, and that enforcement actions can signal increased regulatory attention (Glicksman & Earnhart, 2007). These features make first formal enforcement a substantively meaningful threshold for examining changes in compliance trajectories.

4.2.4 Institutional Constraints and the Feasibility of Escalation

A core complication in SDWA enforcement is that many regulated entities are themselves governments. Because most community water systems are publicly owned, oversight often takes the form of governments regulating other governments. This institutional setting has important implications for enforcement. Public facilities may be more likely to violate yet face

fewer or less severe punitive actions, reflecting the political and administrative costs of sanctioning peer public entities (Wilson & Rachal, 1977; Scheberle, 2004; Konisky & Teodoro, 2016). Unlike private firms, public systems rarely face exit and often provide essential local services, which can raise the costs and reduce the expected effectiveness of aggressive sanctions.

These features align with broader selective-enforcement logic. When enforcement is costly and the regulated entity is institutionally embedded or capacity constrained, regulators may prioritize negotiation, sequenced requirements, and problem-solving support over rapid movement toward punitive measures (McKean, 1980). In the SDWA context, escalation timing may therefore reflect not only compliance history and public health risk, but also judgments about feasibility, political cost, affordability, service continuity, and the likely effectiveness of available tools. A delayed or less punitive response need not indicate administrative indifference. It may instead reflect the practical limits of enforcing against community water systems whose failures also have consequences for the communities they serve.

This institutional perspective helps explain why first formal enforcement should not be understood as an automatic response to a fixed threshold of noncompliance. Escalation instead emerges through decisions made under bounded capacity, competing demands, and uncertainty about which intervention is most likely to move a system toward compliance. The feasibility of escalation is therefore part of the enforcement process itself. Understanding when formal action occurs requires attention not only to system behavior, but also to the institutional constraints that shape what regulators can credibly and practically do.

4.2.5 Responsive Regulation and the Interpretation of Enforcement Timing

Responsive-regulation theory is directly relevant to enforcement timing by conceptualizing oversight as a staged escalation process in which regulators begin with dialogue,

persuasion, and support and move to more coercive responses when lower-level interventions fail (Ayres & Braithwaite, 1992; Braithwaite, 2011). In this framework, the transition to formal enforcement is not simply a stronger sanction. It is a timed shift within an enforcement sequence that reflects prior regulatory engagement and a judgment that earlier efforts have not produced sufficient compliance progress. This perspective fits the SDWA context closely, where regulators often prioritize corrective action and return to compliance over punishment, but still require a credible escalation pathway to address persistent noncompliance. The timing of first formal enforcement therefore matters because it marks the point at which regulators determine that a more structured intervention has become necessary.

Responsive regulation also underscores the importance of legitimacy and procedural fairness. Responsive regulation emphasizes that regulators can strengthen voluntary cooperation by beginning with lower-level engagement and escalating only after those efforts prove insufficient (Braithwaite, 2011). From this perspective, formal escalation matters not simply as a harsher intervention, but as a signal that the regulator-system relationship has shifted into a more formalized mode of accountability. That shift can increase the salience and credibility of compliance obligations even when monetary penalties are limited. Formalization may therefore influence behavior through multiple channels, including deterrence, heightened scrutiny, and the clearer structuring of corrective expectations.

Really responsive regulation extends this argument by clarifying why escalation timing should be expected to be discretionary and institutionally conditioned rather than random.

Baldwin and Black (2008) argue that regulation must respond not only to regulated actors' behavior, but also to the institutional environment, the different logics of regulatory tools, and the practical demands of implementation and evaluation. This perspective shifts attention away

from a mechanical enforcement ladder and toward the institutional conditions under which escalation decisions are made and implemented. A central implication is that the timing of first formal enforcement often reflects accumulated signals of persistence and risk, information limits, administrative feasibility, and organizational frictions rather than an exogenous intervention point (Baldwin & Black, 2008).

These theoretical perspectives suggest that enforcement timing should not be understood as exogenous or purely rule-driven. Given the sequential and discretionary nature of escalation, compliance behavior may begin changing before formalization through earlier monitoring, informal pressure, or anticipatory adjustment. The recorded timing of first formal enforcement may therefore not align perfectly with when behavioral change begins. This interpretive caution matters because formal escalation is best understood as an institutional threshold embedded within a broader enforcement sequence rather than as a discrete intervention detached from prior regulatory activity.

4.2.6 Environmental Justice, Unequal Protection, and the Distribution of Regulatory Response

Environmental justice scholarship reframes enforcement as a question of unequal protection in implementation rather than simply a question of whether statutory standards are formally uniform (Konisky, 2015). In this view, regulatory protection depends on how agencies allocate monitoring attention, decide when to escalate, and sustain follow-up over time. As these processes are shaped by discretion, capacity, and institutional priorities, enforcement need not operate evenly across communities even when legal standards are uniform. This implementation-centered perspective is especially important in drinking water governance, where protection depends not only on the existence of standards, but also on whether violations are detected,

whether oversight intensifies when problems persist, and whether corrective action is ultimately secured.

Executive Order 12898 and subsequent federal reforms framed environmental justice as an implementation problem by directing agencies to address disproportionate harm and by recognizing enforcement as a mechanism for reducing environmental inequities (Konisky, 2015; Konisky & Reenock, 2015). California's 2024 Water Quality Enforcement Policy reflects a similar equal-protection logic by defining environmental justice as fair treatment and meaningful involvement in the implementation and enforcement of environmental laws and policies and by linking enforcement to the reduction of disproportionate pollution burdens (State Water Board, 2023). These federal and state frameworks position enforcement not only as a compliance mechanism, but also as a distributional governance process whose timing and intensity may vary across communities under unequal institutional conditions.

One mechanism through which unequal protection can emerge is compliance bias. Observed noncompliance reflects not only regulated entities' behavior, but also regulators' detection, monitoring, and response. If detection is systematically weaker in marginalized communities due to uneven monitoring, lower complaint capacity, or reduced political pressure, then observed violations and enforcement activity may understate underlying risk in precisely where protection is most needed (Konisky & Reenock, 2013). Related studies on environmental riskscape similarly shows that enforcement allocation can reflect both risk mitigation and political incentives under resource constraints, and that community mobilization can shape how much regulatory attention particular places receive (Konisky & Reenock, 2018).

A related mechanism concerns the unequal availability of third-party amplifiers that strengthen or weaken the practical force of enforcement. Smart-regulation perspectives

emphasize that implementation depends not only on the regulator-regulated entity relationship, but also on the wider network of actors that can add visibility, pressure, and support (Gunningham & Sinclair, 2017). In the drinking water context, formal escalation may become more consequential when it activates broader channels of accountability, including public disclosure, community pressure, media attention, political scrutiny, and pathways to technical or financial assistance. Disadvantaged or resource-constrained communities, however, may have fewer supportive intermediaries and face greater barriers to mobilization, potentially weakening both external pressure and supportive pathways that make escalation effective in practice.

These mechanisms are especially relevant in drinking water governance because compliance burdens and exposure risks are not evenly distributed and often intersect with both system capacity and community disadvantage. State drinking water programs also operate under documented staffing and funding shortfalls that can slow follow-up and limit the consistency of enforcement attention (ASDWA, 2013). These literatures motivate the chapter's focus on whether formal escalation is associated with stronger compliance progress and whether those dynamics differ across disadvantaged-community contexts. More broadly, they frame enforcement timing as part of the unequal distribution of regulatory protection in practice. Figure 4.1 summarizes this chapter's conceptual framework by showing how first formal escalation may shift compliance management into a more structured accountability pathway while post-escalation progress remains shaped by capacity constraints, prior informal pressure, and administrative timing.

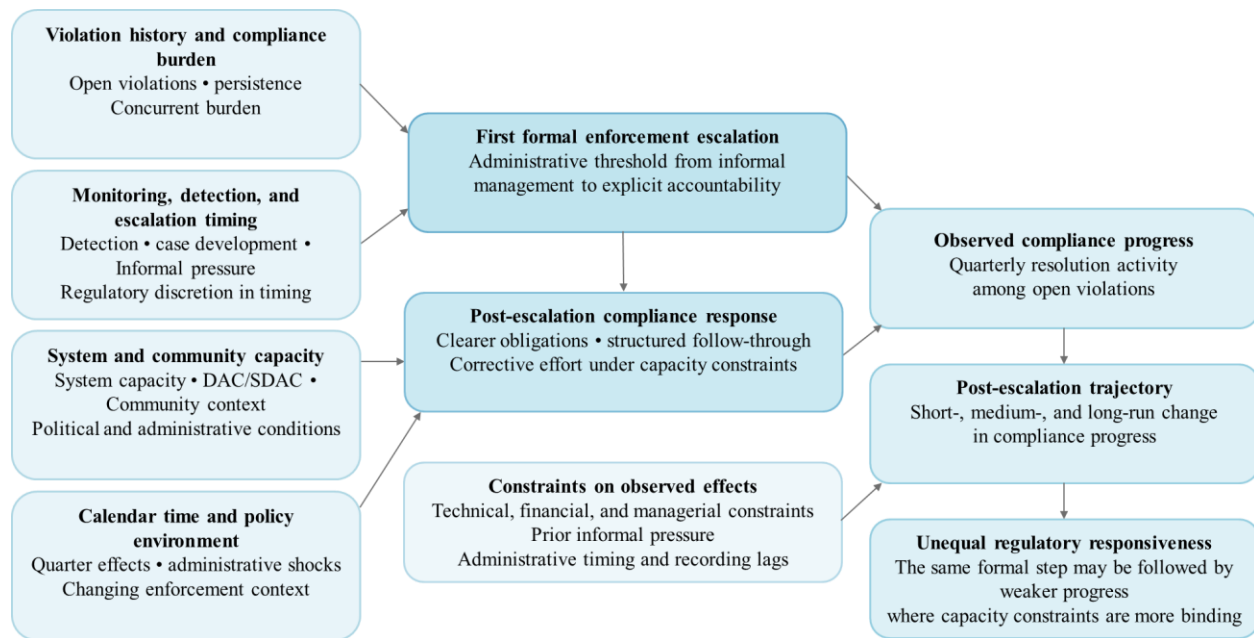


Figure 4.1: Conceptual Framework for First Formal Escalation and Compliance Progress

Note. This figure frames first formal escalation as an administrative threshold that may alter compliance management and subsequent compliance progress. It also highlights why post-escalation gains may be limited or uneven across contexts, including capacity constraints, prior informal pressure, and administrative timing or recording lags.

4.3 Data and Analytic Sample

This chapter draws on the same underlying administrative and contextual data described in Chapter 3, including violation and enforcement records for California community water systems and system-level community demographic measures linked through service-area boundaries. This chapter nevertheless reorganizes those data into a system-quarter panel centered on each system’s first formal enforcement action, reflecting the different empirical design. This section briefly summarizes the shared data foundation and then describes the chapter-specific analytic sample.

4.3.1 Shared Administrative and Contextual Data

The core administrative data come from the EPA's Violations and Enforcement records obtained through the Enforcement and Compliance History Online system for California community water systems identified by PWSID. As described in Chapter 3, these records include violation-level information such as violation identifiers, violation begin and end dates, return-to-compliance dates, violation status, violation category and family codes, health-based indicators, and public notification tier. The same files also link violations to associated enforcement records, including enforcement identifiers, enforcement dates, enforcement action types, and enforcement action categories. The study period extends from the first quarter of 1995 through the fourth quarter of 2024.

This chapter also uses the same system-level community demographic measures developed for Chapter 3. California community water system service-area boundaries were linked to Census and American Community Survey data to construct demographic characteristics for the populations served by each system. These measures include median household income and racial and ethnic composition and are used to characterize the community contexts in which enforcement unfolds. The data source and construction procedures are not repeated here in full, as they are identical to those described in Chapter 3.

4.3.2 Enforcement Classification

The U.S. EPA distinguishes among three broad enforcement categories: administrative, civil, and criminal actions. Administrative enforcement is carried out through agency processes, such as orders, notices, and consent agreements, without requiring court litigation. Civil enforcement involves judicial action seeking court-ordered remedies, penalties, or injunctive relief, whereas criminal enforcement is reserved for particularly serious or willful violations and

may involve prosecution (EPA, n.d., 2026). All enforcement activities examined in this chapter are administrative rather than civil or criminal. This focus reflects the institutional realities of SDWA implementation, where regulators commonly rely on administrative tools to restore compliance without resorting to litigation. For the purposes of this chapter, the key distinction is between informal and formal enforcement. Informal actions generally notify systems of noncompliance or encourage corrective action without imposing the same enforceable compliance schedule. Formal actions, by contrast, place the system on an enforceable path toward return to compliance by specifying required corrective steps and deadlines.

The chapter relies on the enforcement action category variable, which groups detailed enforcement action types into broader classifications. Formal enforcement actions include instruments such as formal administrative orders, formal notices of noncompliance, and administrative consent orders issued by state authorities or administrative law judges associated with state and regional offices (Shimshack, 2014). Informal enforcement actions include less punitive forms of regulatory communication, such as warning letters, telephone calls, and notices of violation, that notify the regulated entity of an infringement without imposing the same enforceable compliance schedule (EPA, 2009c).

The chapter uses the enforcement action category variable, which summarizes detailed enforcement action types into broader categories of formal and informal actions. The data also identify resolving actions, which indicate that the system has returned to compliance, that no further action is necessary, or that the violated rule no longer applies (EPA, 2025). Given resolving actions mark compliance resolution rather than an antecedent enforcement response, the chapter does not treat them as explanatory enforcement actions in the main analysis.

4.3.3 Community Water System Demographic Context

To examine whether enforcement dynamics differ across community contexts relevant to environmental justice, this chapter uses the same system-level demographic indicators constructed for Chapter 3. These measures are based on California community water system service areas linked to Census and American Community Survey data and include median household income and racial and ethnic composition. Following prior drinking water environmental justice research and California policy practice, the chapter uses these data to identify disadvantaged and severely disadvantaged community systems and to construct indicators for systems serving relatively high shares of Hispanic and African American residents.

These measures are treated here as contextual indicators of the communities served by each system rather than as individual-level characteristics. In Chapter 4, their primary function is to support heterogeneity analyses that assess whether compliance progress following formal enforcement differs across disadvantaged and non-disadvantaged settings and across related dimensions of community disadvantage.

4.3.4 Analytic Sample and Unit of Analysis

Although the underlying administrative data are the same as those used in Chapter 3, the unit of analysis and sample construction differ because this chapter focuses on dynamic changes in compliance progress around first formal enforcement. The unit of analysis is the system-quarter. Calendar time is indexed by year-quarter, and each treated system is centered on the first quarter in which a formal enforcement action is recorded. For each system, the data are processed so that at most one first-formal-enforcement quarter is retained.

The full panel includes all observed California community water system-quarters in the Chapter 4 study period. The event-study estimation sample is narrower. It includes systems that

eventually receive a first formal enforcement action and contribute nonmissing event time, outcome, and weight variables within the analytic event window. This construction yields a final event-study estimation sample of 6,961 system-quarters across 804 systems. The resulting estimates therefore describe compliance dynamics among systems that eventually cross the formal-enforcement threshold, rather than all community water systems in the state.

The outcome is not restricted to a single contaminant, rule family, or public-notification tier. Instead, it summarizes quarterly resolution activity across violations open at the start of the quarter for each system. This design captures system-level compliance progress around first formal enforcement rather than contaminant-specific or tier-specific return-to-compliance dynamics. The next section defines formal-enforcement timing, benchmark-adjusted compliance progress, and the event-study model.

4.4 Empirical Strategy

This chapter examines how compliance progress changes around each system's first formal enforcement action using an event-study design in a staggered-timing system-quarter panel (Borusyak et al., 2024; Rambachan & Roth, 2023). The design asks whether a system's pace of violation resolution changes before and after formal escalation, after adjusting quarterly compliance progress by a same-calendar-quarter benchmark. This benchmark places treated system-quarters in relation to broader statewide compliance conditions in the same calendar quarter, while the event-study comparison itself is defined relative to the omitted pre-enforcement quarter, $k = -2$. This comparison is useful, as it places treated and comparison systems under the same statewide conditions at the same time, including broader policy, administrative, and environmental conditions that may affect compliance progress.

4.4.1 Formal-Enforcement Timing and Outcome Construction

The key event in this chapter is a system's first formal enforcement action. For each community water system i , let g_i denote the calendar quarter in which its first formal enforcement action is first recorded. Event time is then defined relative to that quarter, so that each treated system is re-centered around its own first formal enforcement date. This construction allows the analysis to compare compliance progress in quarters before and after first formal escalation even though systems enter formal enforcement in different calendar quarters.

Formally, event time is defined as

$$k_{it} = t - g_i$$

where k_{it} is event time for system i in calendar quarter t , t indexes calendar quarter, and g_i denotes the quarter in which system i first receives a formal enforcement action.

To reduce contamination around the enforcement date, the analysis uses a donut specification that excludes the quarter immediately before formal-enforcement ($k = -1$) and the formal-enforcement quarter itself ($k = 0$) (Hausman & Rapson, 2018; Barreca et al., 2016; Dahl et al., 2014). The quarter immediately before formal enforcement may already reflect regulatory pressure if systems respond to warnings, negotiations, or other signs that escalation is imminent. The formal-enforcement quarter itself may also combine pre- and post-enforcement days because formal actions can occur at any point within the quarter. To avoid mixing those transitional periods into the reference category, the chapter omits both quarters and uses the second quarter before formal-enforcement ($k = -2$) as the comparison point for the event-study coefficients (Miller, 2023; Schmidheiny & Siegloch, 2023; Clarke & Tapia-Schythe, 2021). Thus, $k = -2$ is the event-study reference period. The broader same-calendar-quarter benchmark described below

is used to construct the outcome, not to define the event-study reference period. The main event window includes six quarters before and twelve quarters after first formal enforcement. The six-quarter pre-period and twelve-quarter post-period are empirical design choices rather than formal regulatory thresholds. The window provides enough pre-enforcement quarters to assess baseline dynamics while keeping the post-enforcement horizon within a range where event-time support remains adequate and estimates remain interpretable. This choice is consistent with the quarterly structure of SDWA enforcement tracking and EPA’s expectation that serious, multiple, or continuing violations either return to compliance or be addressed by formal enforcement within six months (EPA, 2009a, 2009b; EPA, 2024).

The outcome is a system-quarter measure of compliance progress. For each system-quarter, the chapter calculates the share of violations that were open at the start of the quarter and were resolved during that same quarter. This quarterly resolution rate captures how quickly the system moves existing noncompliance toward resolution within a given quarter.

Let the quarterly resolution rate be defined as

$$rate_{it} = \frac{resolved_{it}}{openstart_{it}}$$

where $rate_{it}$ is the quarterly resolution rate for system i in quarter t , $resolved_{it}$ is the number of violations resolved during quarter t , and $openstart_{it}$ is the number of violations open at the start of that quarter.

To make this outcome comparable across calendar quarters, the chapter expresses each treated system’s quarterly resolution rate relative to a same-calendar-quarter benchmark. In the current specification, this benchmark is operationalized as the average quarterly resolution rate among systems that do not receive first formal enforcement during the panel. The benchmark-

adjusted outcome is defined as the treated system's own quarterly resolution rate minus this same-calendar-quarter benchmark and is reported in percentage points. The dependent variable is therefore

$$Y_{it} = rate_{it} - \overline{rate}_t^{NYT}$$

Y_{it} measures whether a treated system resolves a larger or smaller share of its open violations than the contemporaneous same-calendar-quarter benchmark facing the same statewide conditions in the same quarter. $rate_{it}$ is the quarterly resolution rate for system i in quarter t , and $\overline{rate}_t^{NYT}$ is the average quarterly resolution rate among systems that did not receive a first formal enforcement action by quarter t . Results are reported in percentage points.

This construction has a two-step difference structure. The first difference compares the treated system's own quarterly resolution rate with the same-calendar-quarter benchmark. This step constructs the dependent variable. The second difference is the event-study comparison, which compares the constructed benchmark-adjusted outcome at event time k with the constructed outcome in the omitted reference quarter, $k = -2$. The same-calendar-quarter benchmark is therefore used in outcome construction, while $k = -2$ defines the event-study comparison. This benchmark is used only to construct the dependent variable. The event-study comparison itself is made across event time within the constructed benchmark-adjusted outcome, relative to the omitted reference quarter, $k = -2$.

A quarterly resolution rate of zero is retained as an observed outcome when a system has violations open at the start of the quarter but records no resolutions during that quarter. This coding is important for the event-study design because zero resolution reflects observed lack of quarterly progress rather than missing information. The resulting outcome differs from the

violation-level duration measure used in Chapter 3. Chapter 3 examines how long individual violations remain unresolved, whereas Chapter 4 examines whether formal escalation is followed by a change in the quarterly pace of resolution at the system level.

4.4.2 Baseline Event-Study Specification

The baseline model estimates how benchmark-adjusted quarterly compliance progress changes over event time around first formal enforcement. The specification includes event-time indicators, system fixed effects, and calendar-quarter fixed effects. System fixed effects account for stable differences across community water systems, including long-standing capacity, governance, and administrative conditions. Calendar-quarter fixed effects account for statewide influences that affect all systems in a given quarter, including policy changes, staffing conditions, drought, COVID-related disruption, and other common shocks. Standard errors are clustered at the system level.

The baseline estimating equation is

$$Y_{it} = \sum_{(k \in K, k \neq -2, -1, 0)} \beta_k 1(k_{it} = k) + \alpha_i + \lambda_t + \varepsilon_{it} \quad (4.1)$$

where Y_{it} is the benchmark-adjusted quarterly resolution outcome for system i in quarter t , defined as the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark. $1(k_{it} = k)$ is an indicator for event time k , α_i denotes system fixed effects, and λ_t denotes calendar-quarter fixed effects. The omitted reference period is $k = -2$, while the donut quarters $k = -1$ and $k = 0$ are excluded from estimation. The coefficients β_k therefore trace changes in benchmark-adjusted quarterly compliance progress at each event time relative to the second quarter before first formal enforcement.

Substantively, positive coefficients indicate that treated systems' benchmark-adjusted quarterly compliance progress is higher at event time k than in the omitted reference quarter, $k = -2$. Negative coefficients indicate lower benchmark-adjusted progress relative to that reference quarter. The coefficients should not be read as a simple comparison between the broad post-enforcement window and the broad pre-enforcement window, nor as a direct comparison between treated systems and the same-calendar-quarter benchmark systems. The same-calendar-quarter benchmark is used to construct the dependent variable. The event-study comparison is between the constructed outcome at each event time k and the constructed outcome at $k = -2$.

This distinction matters because the chapter uses a two-step difference structure. The first difference constructs the dependent variable by comparing the treated system's own quarterly resolution rate with the same-calendar-quarter benchmark. The second difference is the event-study comparison, which compares the constructed benchmark-adjusted outcome at event time k with the constructed outcome in the omitted reference quarter, $k = -2$. Thus, pre-enforcement information enters the model through the omitted $k = -2$ reference quarter and through the estimated pre-event lead coefficients for $k = -6$ through $k = -3$.

The estimates are interpreted as changes in compliance progress around first formal enforcement rather than as fully causal effects. This caution is important because first formal enforcement is not randomly assigned. Regulators may escalate after repeated violations, prolonged noncompliance, insufficient progress, public health concern, or case-specific information developed through informal oversight. Systems may also respond to warnings, negotiations, technical assistance, case development, or anticipated escalation before the formal action appears in the administrative record. The design therefore characterizes how benchmark-

adjusted quarterly compliance progress changes around a meaningful institutional threshold while accounting for stable system differences and common quarter-specific shocks.

4.4.3 Horizon-Averaged Effects and Weighting

Although the event-study coefficients show quarter-by-quarter dynamics, the chapter also reports horizon-averaged effects to summarize post-enforcement patterns more clearly. A small set of grouped post-enforcement estimates is often easier to interpret than a long sequence of quarterly coefficients. The chapter therefore averages post-enforcement changes over three pre-specified horizons: a short-run horizon covering quarters 1 through 3 after first formal enforcement, a mid-run horizon covering quarters 4 through 8, and a long-run horizon covering quarters 9 through 12.

Let H denote a post-enforcement horizon. The equal-weighted horizon average is

$$\bar{\beta}_H = \left(\frac{1}{|H|} \right) \sum_{(k \in H)} \beta_k$$

where $\bar{\beta}_H$ is the equal-weighted average effect for post-enforcement horizon H , $|H|$ is the number of event-time quarters included in that horizon, and β_k is the event-study coefficient for quarter k . This quantity summarizes the average post-enforcement change in compliance progress over the specified interval.

The chapter also reports exponentially weighted horizon averages that place somewhat greater emphasis on earlier quarters within each horizon. A general expression is

$$\bar{\beta}_H^\omega = \sum_{k \in H} w_k \beta_k, \quad \sum_{k \in H} w_k = 1$$

with weights w_k declining across quarters within the same horizon. This weighted version is useful because any effect of formal enforcement may be strongest soon after escalation and may attenuate over time. Reporting both equal-weighted and exponentially weighted summaries makes the conclusions less dependent on a single averaging rule.

In addition to unweighted models that treat each system-quarter equally, the chapter also estimates weighted event-study specifications using analytic weights proportional to the number of violations open at the start of the quarter. Quarters with more open violations receive greater weight, as the quarterly resolution rate is more informative when more violations are at risk of resolution. In notation,

$$\omega_{it} \propto \text{openstart}_{it}, \quad \bar{\omega} = \text{among quarters with } \text{openstart}_{it} > 0$$

so that the weights are normalized to have mean one among system-quarters with at least one open violation at the start. This weighting strategy places greater emphasis on quarters in which the outcome is measured from a larger underlying pool of unresolved violations.

These alternative summaries and weighting schemes serve two purposes. First, they make dynamic results easier to interpret at the short-, mid-, and long-run horizons. Second, they help assess whether the main conclusions depend on how post-enforcement dynamics are summarized or on whether greater weight is given to quarters with larger numbers of unresolved violations.

4.4.4 Heterogeneity by community context

The chapter next examines whether changes in compliance progress around first formal enforcement differ across community contexts relevant to environmental justice. To do so, it estimates interacted event-study models using the system-level demographic indicators described

in Section 4.3, including disadvantaged community status, severely disadvantaged community status, and related racial and ethnic service-area measures.

A generic interacted specification can be written as

$$Y_{it} = \sum_{k \in K, k \neq -2, -1, 0} [\beta_k 1(k_{it} = k) + \theta_k (1(k_{it} = k) \times Z_i)] + \alpha_i + \lambda_t + \varepsilon_{it} \quad (4.2)$$

where Z_i is a time-invariant system-level community indicator, such as disadvantaged community status. Under this specification, the β_k terms describe the event-time pattern for the reference group, and the θ_k terms capture how that pattern differs for systems with the focal community characteristic.

These heterogeneity analyses ask whether formal escalation is followed by similar patterns of compliance progress across systems serving different communities, or whether the apparent effects of formal enforcement are weaker, delayed, or less durable in more disadvantaged settings. This question is central to the chapter's broader unequal-protection framework. If the same formal enforcement step is associated with stronger compliance progress in some contexts than in others, then enforcement may operate unevenly in practice even when the formal legal instrument is the same.

4.5 Results

4.5.1 Analytic Sample and Descriptive Summaries

Table 4.1 summarizes the event-study data structure, outcome construction, and estimation-sample support. Panel A shows that the main event-study window spans six quarters before through twelve quarters after first formal enforcement, excludes the donut quarters $k = -1$

and $k = 0$, and uses $k = -2$ as the omitted reference quarter. The main regression sample includes 6,959 treated system-quarters from 804 treated systems with valid benchmark-adjusted outcomes.

Panel B clarifies how the dependent variable is constructed. Treated systems' own quarterly resolution rate averaged 7.44 percentage points, while the same-calendar-quarter benchmark assigned to treated observations averaged 9.01 percentage points. The resulting benchmark-adjusted quarterly resolution outcome averaged -1.59 percentage points. These values should not be read as two regression groups. Rather, the same-calendar-quarter benchmark is used to construct the dependent variable for treated system-quarters.

Table 4.1: Event-Study Data Structure, Outcome Construction, and Sample Support

Panel A. Estimation Sample and Event-Time Structure

Measure	Value
Full CWS-quarter panel	286,770 system-quarters
Distinct CWSs in full panel	2,370 systems
Systems receiving first formal enforcement during the panel	1,398 systems
Systems without first formal enforcement during the panel	972 systems
Main event-study window	$k = -6$ to $k = +12$
Donut quarters excluded from estimation	$k = -1$ and $k = 0$
Reference quarter	$k = -2$
Treated system-quarters in estimation window	6,961
Treated system-quarters with benchmark-adjusted outcome	6,959
Distinct treated systems with benchmark-adjusted outcome	804 systems
Missing constructed-outcome cases	2
Benchmark-source system-quarters with defined resolution rates	6,607
Distinct benchmark-source systems	455 systems

Note. The main regression sample contains 6,959 treated system-quarters from 804 treated systems. The 2 missing constructed-outcome cases are missing the same-calendar-quarter benchmark; the treated-system resolution rate is not missing. Benchmark-source observations come from systems that do not receive first formal enforcement during the panel and have defined quarterly resolution rates in the relevant calendar quarter.

Panel B. Outcome Construction and Assigned Benchmark Values

Measure	N	Mean	SD	Median	Min	Max
Same-calendar-	6,959	9.01	9.70	5.21	0.00	66.67

quarter benchmark resolution rate assigned to treated observations, pp						
Treated system's own quarterly resolution rate, pp	6,961	7.44	25.21	0.00	0.00	100.0 0
Benchmark- adjusted quarterly resolution outcome, pp	6,959	-1.59	25.37	-4.55	-66.67	100.0 0

Note. The benchmark-adjusted quarterly resolution outcome equals the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark. The benchmark is used to construct the dependent variable for treated system-quarters; the table does not compare treated observations and benchmark-source observations as two regression groups.

Panel C. Treated Systems' Pre- and Post-Enforcement Information

Event-time window	Treated- rate N	Constructed- outcome N	Mean treated resolution rate, pp	Mean benchmark, pp	Mean adjusted outcome, pp
Pre-enforcement window, $k = -6$ to -3	1,312	1,312	4.63	9.06	-4.43
Reference quarter, $k = -2$	422	422	2.37	9.90	-7.53
Excluded donut quarters, $k = -1$ and $k = 0$	1,132	Not used	16.05	10.60	5.45
Post-enforcement window, $k = +1$ to $+12$	5,227	5,225	8.56	8.92	-0.40

Note. Rows are mutually exclusive. The event-study estimation sample excludes the donut quarters, so the treated-rate N used for estimation is 6,961. The constructed outcome is available for 6,959 of these observations because two post-enforcement observations are missing the same-calendar-quarter benchmark. The event-study model does not compare the broad post-enforcement row with the broad pre-enforcement row. Each event-time coefficient compares the constructed outcome at event time k with the constructed outcome at the reference quarter, $k = -2$. Zero pre-enforcement resolution rates are observed zero progress when open violations exist at quarter start, not missing pre-enforcement information.

Panel D. System-Context Composition of the Main Estimation Sample

System characteristic	Treated system-quarter N	Estimation sample, %
DAC system	3,545	50.91
SDAC system	2,233	32.09
African American-serving system	475	6.83
Hispanic-serving system	2,097	30.13
DAC × Hispanic-serving system	1,392	20.00
System-quarter observations, N	6,959	100

Note. Counts are treated system-quarters in the main estimation sample with valid benchmark-adjusted outcomes. Percentages are calculated over the 6,959 treated system-quarter observations. African American-serving and Hispanic-serving systems are defined as systems whose served-population shares exceed the statewide average benchmark over the study period. Percentages are calculated over treated system-quarters in the main estimation sample. African American-serving and Hispanic-serving systems are defined as systems whose served-population shares exceed the statewide average benchmark over the study period.

Panel C summarizes treated systems' own pre- and post-enforcement information. The reference quarter, $k = -2$, had an average treated resolution rate of 2.37 percentage points and an average benchmark-adjusted outcome of -7.53 percentage points. The event-study coefficients compare each estimated event-time quarter with this constructed outcome in the reference quarter. The broad pre- and post-enforcement rows are descriptive summaries only; the model estimates event-time coefficients relative to $k = -2$. Panel D reports the community-context composition and effective subgroup support in the main estimation sample. Among the 6,959 treated system-quarter observations, 3,545 involve DAC systems, 2,233 involve SDAC systems, 2,097 involve Hispanic-serving systems, 1,392 involve systems that are both DAC and Hispanic-serving, and 475 involve African American-serving systems. These subgroup counts and percentages motivate the heterogeneity analyses reported later in the chapter and help contextualize differences in precision across moderator contrasts.

This compositional pattern is important for interpreting the chapter's scope. The event-study estimates describe enforcement dynamics within the population of systems that eventually

receive a first formal enforcement action, not within all California community water systems. The overrepresentation of DAC, SDAC, and high-Hispanic-share systems in the estimation sample therefore reinforces the chapter's unequal-protection framing and motivates the heterogeneity analyses presented later in Section 4.5.4. Appendix Tables 4.A1 through 4.A3 provide supporting information on variable definitions, missingness, and event-time support. With the sample structure established, the next subsection turns to the baseline event-study estimates shown in Figure 4.2 and the post-enforcement summaries reported in Table 4.2.

4.5.2 Baseline Event-Study Results

Figure 4.2 presents the baseline weighted, tail-trimmed event-study estimates of quarterly compliance progress around first formal enforcement. Overall, the pattern indicates a modest and temporary improvement following formal escalation. The estimated post-enforcement gains do not appear immediately. Instead, the coefficients remain close to zero in the first two post-enforcement quarters, rise more clearly around quarters 3 and 4, and then gradually attenuate over the remainder of the event window. This timing is consistent with the expectation that formal enforcement may require time to translate into observable resolution activity, particularly when returning to compliance depends on administrative follow-up, technical adjustments, or operational changes that cannot be completed within a single quarter.

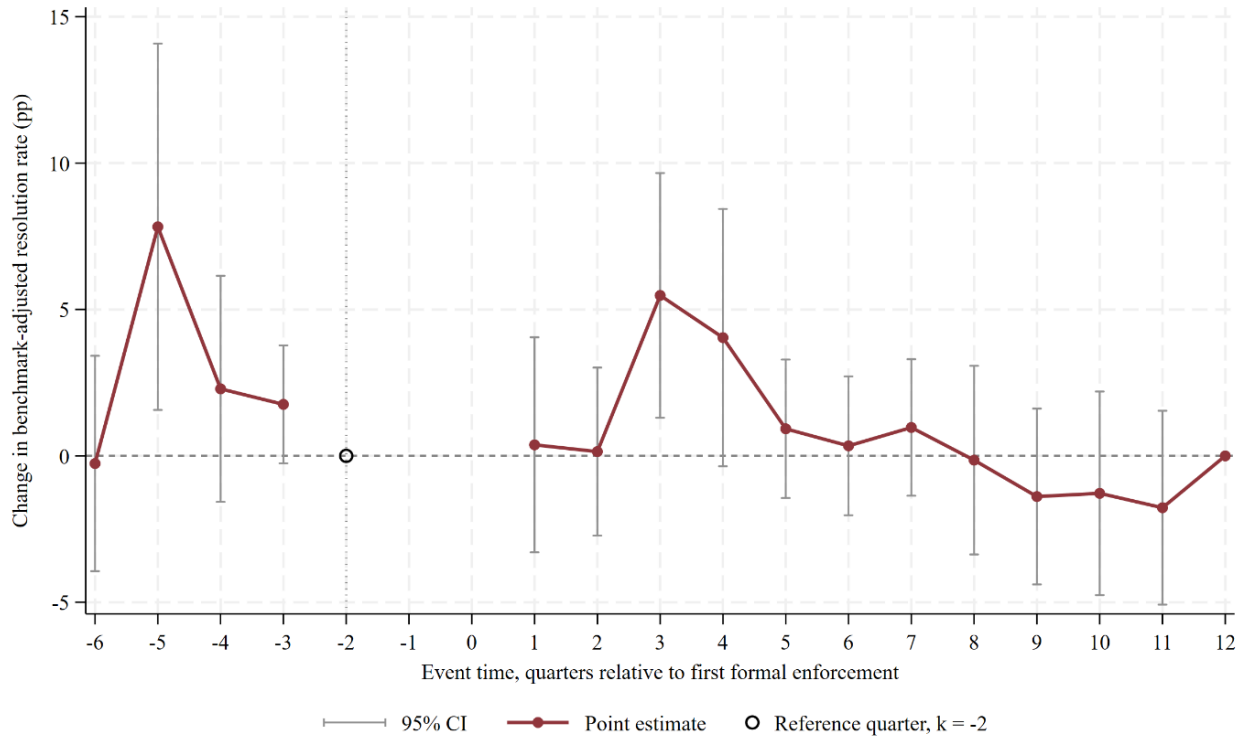


Figure 4.2: Weighted Event-Study Estimates of Quarterly Compliance Progress Around First Formal Enforcement

Note. Points show event-time estimates from the weighted event-study specification, and vertical bars show 95% confidence intervals. The outcome is the benchmark-adjusted quarterly resolution rate, measured in percentage points relative to the same-calendar-quarter benchmark. The omitted reference quarter is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation and appear as empty positions on the x-axis. Estimates are weighted by normalized open violations at the start of the quarter. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Results are descriptive and should not be interpreted as causal effects of enforcement timing.

The magnitude of the estimated improvement is limited. As shown in Figure 4.2, the largest positive point estimates occur in the early-to-middle post-enforcement period, but the confidence intervals remain wide enough that the quarter-specific estimates should be interpreted cautiously. By the later post-enforcement quarters, the coefficients return toward zero and become slightly negative. The baseline figure therefore points to a delayed and temporary increase in compliance progress rather than a large or sustained post-enforcement shift.

The lead coefficients in Figure 4.2 also require caution because the pre-enforcement path is not flat. This pattern is consistent with the chapter’s methodological discussion that first formal enforcement is an endogenous institutional threshold rather than an exogenous intervention. Formal escalation may be preceded by informal pressure, case development, or anticipatory system response, and the recorded enforcement quarter may not align exactly with when behavior begins to change. The event study is therefore interpreted as characterizing changes in compliance progress around first formal enforcement rather than identifying a causal effect of enforcement timing. More formal pre-enforcement diagnostics are presented in Section 4.5.5. Appendix Figure 4.A1 reports the corresponding unweighted baseline event-study estimates, providing a sensitivity check on whether the dynamic pattern in Figure 4.2 depends on weighting by open violations at the start of the quarter.

In sum, Figure 4.2 suggests that formal enforcement is followed by a modest and temporary improvement in quarterly compliance progress. Because quarter-specific estimates are somewhat noisy, Table 4.2 provides a more compact summary of this pattern using horizon-averaged post-enforcement estimates.

4.5.3 Horizon-Averaged Post-Enforcement Effects

Table 4.2 summarizes the weighted event-study results by averaging post-enforcement coefficients over short-run, mid-run, and long-run horizons. This grouped presentation clarifies the baseline pattern in Figure 4.2 while reducing reliance on individual quarter-specific estimates. Under both equal weighting and exponential weighting, the short-run and mid-run averages are positive, whereas the long-run averages are near zero and slightly negative. However, none of the horizon-averaged estimates is statistically distinguishable from zero at conventional levels, and all confidence intervals include zero. The table therefore provides

descriptive evidence of a modest early post-enforcement increase in compliance progress, but it does not support a statistically significant average post-enforcement change across the grouped horizons.

Table 4.2: Horizon-Averaged Post-Enforcement Changes in Quarterly Compliance Progress, Weighted Estimates

Horizon	Averaging scheme	Effect (pp)	95% CI	<i>p</i> -value
Short run, quarters 1–3	Equal-weighted	2.00	[-0.44, 4.44]	0.108
Short run, quarters 1–3	Exponential	1.60	[-0.92, 4.12]	0.213
Mid run, quarters 4–8	Equal-weighted	1.23	[-1.22, 3.67]	0.326
Mid run, quarters 4–8	Exponential	1.67	[-0.91, 4.24]	0.205
Long run, quarters 9–12	Equal-weighted	-1.11	[-3.35, 1.13]	0.331
Long run, quarters 9–12	Exponential	-1.21	[-3.67, 1.25]	0.335

Note. Estimates are from the weighted, tail-trimmed event-study specification. The outcome is quarterly compliance progress, measured in percentage points as the treated system’s own quarterly resolution rate minus the same-calendar-quarter benchmark. In the current specification, the benchmark is operationalized as the average resolution rate among systems that do not receive first formal enforcement during the panel. Short-run, mid-run, and long-run horizons correspond to post-enforcement quarters 1–3, 4–8, and 9–12, respectively. Equal-weighted averages assign the same weight to each available post-event quarter within a horizon. Exponentially weighted averages place slightly greater weight on earlier quarters within the same horizon. All models include system and calendar-quarter fixed effects. Standard errors are clustered at the system level. Analytic weights are based on normalized open violations at quarter start. Donut quarters $k = -1$ and $k = 0$ are excluded, and the omitted reference period is $k = -2$. Estimates should be interpreted as horizon-average differences in benchmark-adjusted quarterly compliance progress relative to $k = -2$. pp = percentage points. Ninety-five percent confidence intervals are shown in brackets.

The short-run estimates indicate a limited positive shift in compliance progress after formal escalation. The mid-run estimates for quarters four through eight are similar in

magnitude. All four estimates are positive, but none is statistically distinguishable from zero at conventional thresholds. By contrast, the long-run estimates for quarters nine through twelve are slightly negative under both averaging schemes. These estimates are likewise imprecise and remain statistically indistinguishable from zero. Substantively, they suggest that any early post-enforcement gains do not persist into the later post-enforcement period.

The comparison between equal-weighted and exponentially weighted summaries provides limited evidence that earlier quarters within each horizon carry somewhat larger gains, because the exponentially weighted estimates are slightly larger than the equal-weighted estimates in the short and mid run. At the same time, the differences across weighting schemes are small and do not alter the substantive interpretation. Figure 4.2 and Table 4.2 indicate that first formal enforcement is associated with, at most, a modest increase in quarterly compliance progress that emerges with some delay and then fades. The next subsection examines whether this average pattern conceals larger differences across community contexts relevant to environmental justice.

4.5.4 Heterogeneity by Community Context

Table 4.3 examines whether post-enforcement compliance progress differs across community contexts using the same system-quarter event-study framework as the baseline analysis. These estimates are not based on separate subgroup-only datasets. Instead, each row compares the horizon-averaged event-study pattern for a focal group with the corresponding pattern for its reference group. The row labels state the comparison directly. For example, the DAC row reports DAC systems minus non-DAC systems, and the SDAC row reports SDAC systems minus non-SDAC systems. Negative estimates indicate weaker post-enforcement compliance progress for the focal group relative to the reference group, measured in percentage

points. These estimates should therefore be read as differences in post-enforcement responsiveness across community contexts, not as absolute resolution rates within each subgroup.

Table 4.3 and Figure 4.3 show that the clearest heterogeneity appears by disadvantaged community status. DAC systems exhibit weaker post-enforcement compliance progress than non-DAC systems across all three horizons. The estimated DAC difference is –6.47 percentage points in the short run, –4.93 percentage points in the mid run, and –5.18 percentage points in the long run. Figure 4.3 shows the same pattern dynamically, with DAC systems displaying weaker post-enforcement improvement across much of the post-enforcement period. These results indicate that first formal enforcement is associated with smaller gains in quarterly compliance progress among DAC systems than among non-DAC systems within the event-study framework. Appendix Figures 4.A2 through 4.A5 provide parallel weighted event-study displays for the remaining community-context moderators, including SDAC systems, African American-serving systems, Hispanic-serving systems, and systems that are both DAC and Hispanic-serving. These appendix figures show the quarter-specific subgroup patterns underlying the horizon-averaged comparisons in Table 4.3, while the main text focuses on DAC status because it is the clearest and most stable heterogeneity pattern.

Table 4.3: Community-Context Differences in Horizon-Averaged Post-Enforcement Compliance Progress, Weighted Estimates

Comparison: focal group minus reference group	Short run, quarters 1–3 Effect [95% CI] (pp)	Mid run, quarters 4–8 Effect [95% CI] (pp)	Long run, quarters 9–12 Effect [95% CI] (pp)
DAC systems minus non-DAC systems	-6.47 [-11.31, -1.62]***	-4.93 [-9.04, -0.82]**	-5.18 [-9.01, -1.35]***
SDAC systems minus non-SDAC systems	-2.11 [-7.08, 2.86]	0.55 [-3.14, 4.24]	0.01 [-3.61, 3.63]

Hispanic-serving systems minus other systems	-5.24 [-10.62, 0.15]*	-1.28 [-6.41, 3.84]	-3.43 [-7.84, 0.98]
DAC and Hispanic-serving systems minus all other systems	-6.44 [-12.87, 0.00]*	-1.40 [-7.36, 4.56]	-3.91 [-9.08, 1.26]
African American-serving systems minus other systems	0.65 [-4.55, 5.84]	-1.85 [-7.16, 3.46]	-2.19 [-7.32, 2.94]

Note. Entries report equal-weighted focal-minus-reference differences in horizon-averaged post-enforcement compliance progress, expressed in percentage points. Each row is estimated using the same system-quarter event-study framework as the baseline analysis, not a separate subgroup-only dataset. The focal group is named first in each comparison, and the reference group is named second. Negative estimates indicate weaker post-enforcement compliance progress for the focal group relative to the reference group. The outcome is quarterly compliance progress, measured in percentage points as the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark. Short-run, mid-run, and long-run horizons correspond to post-enforcement quarters 1–3, 4–8, and 9–12, respectively. Hispanic-serving and African American-serving systems are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. Models include system and calendar-quarter fixed effects. Donut quarters $k = -1$ and $k = 0$ are excluded, and the omitted reference period is $k = -2$. Standard errors are clustered at the system level. Analytic weights are based on normalized open violations at quarter start. Ninety-five percent confidence intervals are shown in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The substantive magnitude of the DAC gap is meaningful. A difference of roughly 5 to 6 percentage points implies that, for every 100 violations open at the start of the quarter, DAC systems experience about 5 to 6 fewer additional resolutions in the relevant post-enforcement horizon than non-DAC systems. The average post-enforcement estimates in the baseline analysis are modest and statistically imprecise, but the heterogeneity results indicate that observed post-enforcement compliance progress is not evenly distributed across community contexts.

The remaining community-context moderators show weaker and less consistent patterns. In Table 4.3, the SDAC estimates are close to zero in the mid- and long-run horizons, while the Hispanic-serving and DAC $\times$ Hispanic-serving contrasts are negative in the short run but do not clearly persist across later horizons. The African American-serving-system comparison is especially difficult to interpret because the estimates are small, statistically indistinguishable from zero, and based on the smallest focal subgroup. Effective subgroup support helps

contextualize these differences in precision. As shown in Table 4.1, the focal subgroup counts range from 3,545 treated system-quarters for DAC systems to 2,233 for SDAC systems, 2,097 for Hispanic-serving systems, 1,392 for DAC $\times$ Hispanic-serving systems, and 475 for African American-serving systems. Because the average post-enforcement change is modest, subgroup contrasts require substantial precision. The smaller subgroup contrasts therefore require greater caution, especially the African American-serving-system comparison.

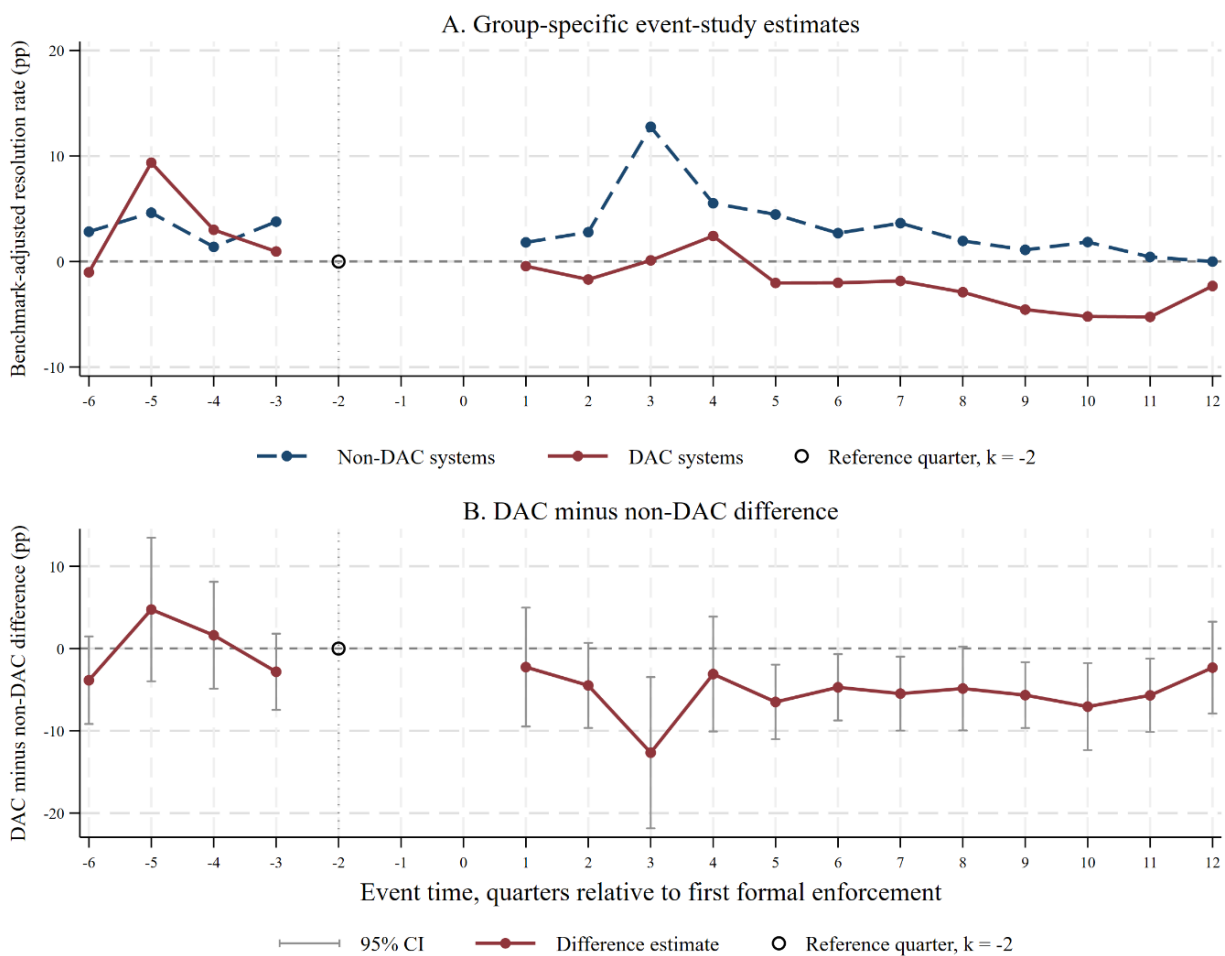


Figure 4.3: Heterogeneity by Disadvantaged Community Status in Weighted Event-Study Estimates of Quarterly Compliance Progress

Note. Panel A shows weighted event-study estimates separately for non-DAC and DAC systems. Panel B shows the DAC minus non-DAC difference, with vertical bars indicating 95%

confidence intervals. The outcome is the benchmark-adjusted quarterly resolution rate, measured in percentage points relative to the same-calendar-quarter benchmark. The omitted reference quarter is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation and appear as empty positions on the x-axis. Estimates are weighted by normalized open violations at the start of the quarter. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Results are descriptive and should not be interpreted as causal effects of enforcement timing.

Appendix Figures 4.A2 through 4.A5 provide parallel weighted event-study displays for the remaining community-context moderators, including SDAC systems, African American-serving systems, Hispanic-serving systems, and systems that are both DAC and Hispanic-serving. Appendix Table 4.A3 reports the corresponding unweighted horizon-averaged comparisons. The DAC pattern remains negative under the unweighted specification, although with weaker precision in the mid- and long-run horizons. The remaining moderators are less stable. The African American-serving-system comparison becomes positive in the unweighted specification, but this pattern is not visible in the weighted estimates and is based on the smallest focal subgroup. It is therefore treated as sensitivity to weighting rather than as a robust heterogeneity result.

4.5.5 Diagnostics and Selected Sensitivity Checks

Given that this chapter interprets the event-study estimates as descriptive evidence on compliance dynamics rather than as fully causal effects, the diagnostic and sensitivity checks focus on the assumptions and design choices most important for interpretation. The checks examine whether pre-enforcement dynamics are present, whether weighting affects the stability of the pre-event path, and whether the horizon-averaged post-enforcement summaries depend on the exact post-event window definitions or support thresholds.

4.5.5.1 Pre-Enforcement Dynamics

Table 4.4 and Figure 4.4 assess pre-enforcement dynamics in the weighted and unweighted event-study specifications. For the full sample, the weighted joint test of the lead coefficients yields a p -value of 0.056, compared with 0.008 in the unweighted specification. This comparison indicates that weighting improves the stability of the pre-enforcement path, although it does not eliminate concern about non-flat leads altogether. This pattern matches the chapter's interpretation of first formal enforcement as an endogenous institutional threshold. Pre-enforcement quarters may already be institutionally active periods, reflecting informal oversight, case development, partial correction, or anticipatory system responses before the formal action is recorded.

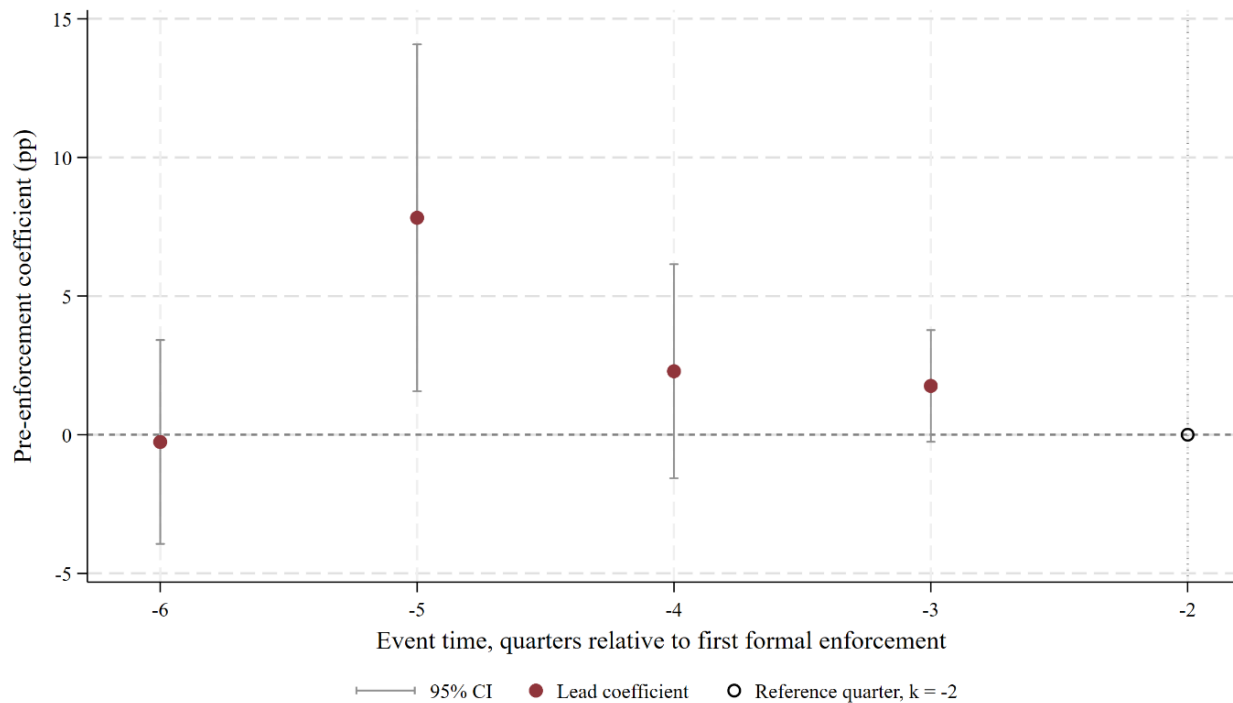


Figure 4.4: Weighted Pre-Enforcement Coefficients from the Baseline Event-Study Specification

Note. Points show pre-enforcement lead coefficients from the weighted baseline event-study specification, and vertical bars show 95% confidence intervals. The outcome is the benchmark-adjusted quarterly resolution rate, measured in percentage points relative to the same-calendar-quarter benchmark. Lead coefficients are measured relative to the omitted reference quarter, $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation. Estimates are weighted by normalized open violations at the start of the quarter. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Results are descriptive and should be interpreted as a pre-enforcement diagnostic rather than as a causal test.

Figure 4.4 conveys the same message visually. In the weighted specification, the lead coefficients are less erratic than in the unweighted model, but they are not uniformly close to zero. The pre-enforcement path is therefore more stable under weighting, yet still not flat. This evidence supports a cautious interpretation of the baseline event-study results. The estimates are best understood as describing changes in quarterly compliance progress around formal escalation rather than identifying a fully exogenous causal effect.

Table 4.4: Joint Tests of Pre-Enforcement Event-Time Coefficients, Weighted and Unweighted Specifications

Group	Weighted <i>p</i> -value	Unweighted <i>p</i> -value
Overall sample	0.056*	0.008***
DAC systems minus non-DAC systems	0.237	0.913
SDAC systems minus non-SDAC systems	0.886	0.556
African American-serving systems minus other systems	0.027**	0.143
Hispanic-serving systems minus other systems	0.452	0.842
DAC and Hispanic-serving systems minus all other systems	0.449	0.808

Note. Entries report *p*-values from joint tests of the pre-enforcement event-time coefficients for quarters $k = -6, -5, -4$, and -3 . For the overall sample, the test evaluates whether the overall pre-enforcement coefficients are jointly equal to zero. For community-context rows, the test

evaluates whether the focal-group-minus-reference-group pre-enforcement difference coefficients are jointly equal to zero. The omitted reference period is $k = -2$, and donut quarters $k = -1$ and $k = 0$ are excluded. The weighted specification uses analytic weights based on normalized open violations at quarter start; the unweighted specification treats each system-quarter equally. Smaller p -values indicate stronger evidence of pre-enforcement dynamics before formal enforcement. These tests are diagnostic and should be interpreted as evidence about pre-event patterns rather than post-enforcement effects. Statistical significance: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

The subgroup tests in Table 4.4 also show that pre-enforcement dynamics vary across moderators. For DAC and SDAC, the weighted and unweighted p -values do not indicate strong evidence of differential pre-enforcement movement. This strengthens the interpretation of the DAC heterogeneity pattern reported in Section 4.5.4, because the main disadvantaged-community split is not accompanied by pronounced lead-period instability. By contrast, the African American-serving-system split yields a weighted p -value of 0.027, while the unweighted p -value is 0.143. Together with the weak substantive pattern reported earlier, this inconsistency indicates that the African American subgroup results require caution. The Hispanic and DAC $\times$ Hispanic-serving-system interactions do not show comparable evidence of problematic lead dynamics.

Overall, the diagnostic evidence supports a restrained conclusion. Weighting improves the behavior of the pre-enforcement path, and the main DAC heterogeneity result is not accompanied by strong evidence of differential lead movement. At the same time, the diagnostics do not support a strict exogeneity interpretation of first formal enforcement. The chapter therefore interprets the event-study estimates as evidence on compliance dynamics around formal escalation while reserving fully causal claims.

4.5.5.2 Horizon-Window and Support Sensitivity

Appendix Table 4.A4 assesses whether the weighted post-enforcement summaries depend on the exact upper bound used to define the mid-run and long-run horizons. For the mid run, the average effect is recomputed over windows 4–6, 4–7, and 4–8. For the long run, the average effect is recomputed over windows 9–10, 9–11, and 9–12. Across these alternatives, the substantive pattern remains stable. Mid-run averages remain modestly positive under both equal and exponential weighting, while long-run averages remain near zero and slightly negative. Although the exact magnitudes vary somewhat with the window definition, the interpretation does not change. In all cases, the confidence intervals include zero. These results show that the chapter’s substantive conclusions do not depend on a particular choice of post-enforcement window.

Support-based trimming also does not materially affect the horizon summaries. Tightening the minimum support threshold for post-enforcement bins from at least two to at least four distinct calendar quarters leaves the weighted horizon averages unchanged. This exercise is not shown as a separate table, as it does not alter the estimates. Together, these checks indicate that the baseline interpretation is not an artifact of how late post-enforcement bins are handled or how the mid-run and long-run horizons are capped.

This chapter also reports a restricted-sample sensitivity check focused on support for the omitted reference quarter, $k = -2$. Because the event-study coefficients are interpreted relative to this reference period, the check restricts the model to systems with valid constructed-outcome support at $k = -2$. After singleton drops, the restricted model includes 4,761 observations from 422 systems. The horizon-averaged estimates differ somewhat in magnitude from the full-sample estimates but lead to the same broad interpretation. In the full model, the short-, mid-, and long-

run pooled estimates are approximately +2.00, +1.23, and -1.11 percentage points, respectively. In the $k = -2$ supported sample, the corresponding estimates are approximately +3.10, -0.22 , and -2.12 percentage points. None of the horizon estimates is statistically significant in either specification. This check indicates that the main descriptive conclusion is not driven solely by systems lacking valid $k = -2$ support.

Overall, the selected robustness checks reinforce the restrained interpretation of the chapter's main results. The modest short- and mid-run gains documented in the baseline specification do not appear to be driven by a narrow same-calendar-quarter benchmark definition, thinly supported post-enforcement tails, the choice of horizon cap, or the inclusion of systems without valid $k = -2$ support. At the same time, the robustness exercises do not strengthen the evidence into a clear long-run post-enforcement effect. The most defensible conclusion therefore remains that first formal enforcement is associated with limited and temporary improvement in quarterly compliance progress rather than a large or sustained shift.

4.6 Discussion

4.6.1 Interpretation and Implications

This chapter examined how quarterly compliance progress changes around first formal enforcement in California community water systems and whether those dynamics differ across community contexts relevant to environmental justice. The results show that first formal enforcement is associated, at most, with modest and temporary changes in compliance progress. Estimated gains emerge with some delay, are concentrated in the earlier post-enforcement period, and fade over time. Formal escalation therefore does not appear to produce a large or durable shift in compliance progress across the full set of systems that reach this administrative threshold.

This limited average response is substantively consistent with the institutional setting in which SDWA enforcement operates. Return to compliance often requires more than a formal directive, since corrective progress may depend on engineering design, operational adjustment, managerial capacity, financing, procurement, and administrative verification. Formal enforcement may still matter by clarifying obligations, imposing schedules, and increasing accountability, but its ability to generate rapid or sustained compliance gains is constrained by the feasibility of correction. The results therefore support a restrained interpretation of formal escalation as one step within a longer corrective process rather than a reliably transformative intervention on its own.

The findings also help explain why formal enforcement is used selectively and why strongly punitive or final actions remain comparatively rare in SDWA implementation. Regulators oversee public water systems that are often small, capacity-constrained, and responsible for essential service provision, so they must balance deterrence and public-health protection against the practical need to sustain service delivery and support corrective action. Excessive reliance on punitive action may worsen compliance challenges or undermine cooperative problem solving, while an overly accommodative approach may weaken deterrence. The chapter's results sharpen this tradeoff by showing that formal enforcement remains important, but practically limited where the underlying barriers to compliance are structural. Formal action is therefore best understood as an imperfect component of a broader compliance strategy rather than a sufficient solution on its own.

The heterogeneity results sharpen this interpretation by showing that DAC systems experience smaller post-enforcement gains than non-DAC systems across the short-, mid-, and long-run horizons. This pattern suggests that the same formal enforcement step does not translate

into comparable compliance progress across systems, so formal equality in the enforcement instrument does not necessarily produce equal protection in practice. Where systems face deeper financial, technical, or organizational constraints, formal escalation may impose a more structured accountability pathway without generating the same pace of recorded compliance recovery. The chapter therefore refines the unequal-protection concern in the enforcement context. The evidence does not show that disadvantaged-community systems are systematically excluded from formal escalation; indeed, they are overrepresented in the estimation sample. The concern instead lies in whether formal enforcement produces comparable corrective movement once it occurs.

These findings carry several policy implications. First, the weaker DAC response indicates that formal escalation should be paired more explicitly with capacity-oriented follow-through in disadvantaged settings. If DAC systems experience smaller gains after formal enforcement, policy attention should focus not only on when regulators escalate, but also on what happens after escalation. Formal enforcement should function less as a stand-alone punitive endpoint than as a trigger for structured follow-through, including corrective-action milestones, financing support, technical assistance, managerial intervention, and closer monitoring where progress stalls. Rather than assuming that stronger formalization will generate comparable compliance gains across systems, regulators may need to invest more directly in the corrective supports that make recovery feasible.

The results also indicate that equity concerns should be incorporated into post-enforcement oversight, not only into access to enforcement. Under limited regulatory capacity, an equity-oriented strategy may require more than equal treatment in the issuance of formal actions. It may also require closer monitoring of post-enforcement trajectories in disadvantaged-

community systems, earlier identification of stalled remediation pathways, and more intensive follow-up where formal action does not produce measurable progress. Environmental justice depends not only on whether enforcement occurs, but also on whether enforcement produces effective and timely compliance recovery.

This interpretation is reinforced by a descriptive bridge to Chapter 3. Among overlapping treated systems, median pre-enforcement return-to-compliance duration is not strongly or negatively associated with short-run post-escalation compliance progress. The Spearman correlations are small and positive across the main bridge measures, and systems with the longest pre-enforcement duration are not disproportionately concentrated in the weakest post-escalation response group. These patterns do not support a simple view that the systems with the slowest pre-enforcement recovery are also the systems with the weakest immediate post-escalation response. Instead, they suggest that prolonged recovery reflects broader constraints, including remediation feasibility, organizational capacity, cumulative compliance burden, and administrative closeout, rather than enforcement timing alone.

This interpretation is reinforced by a descriptive bridge to Chapter 3. Among overlapping treated systems, median pre-enforcement return-to-compliance duration is not strongly or negatively associated with short-run post-escalation compliance progress. The main Spearman correlation is small and positive rather than negative, at 0.14 for median pre-enforcement duration and average post-escalation compliance progress over event quarters 1 to 4. A stricter post-minus-pre response measure shows a similarly small association, and systems with the longest pre-enforcement duration are not disproportionately concentrated in the weakest post-escalation response group. These patterns do not support a simple view that the systems with the slowest pre-enforcement recovery are also the systems with the weakest immediate post-

escalation response. Instead, they suggest that prolonged recovery reflects broader constraints, including remediation feasibility, organizational capacity, cumulative compliance burden, and administrative closeout, rather than enforcement timing alone.

4.6.2 Limitations

This chapter has several limitations that qualify how the results should be interpreted. Most importantly, the event-study design does not identify a fully exogenous causal effect of formal enforcement. First formal enforcement is an administrative threshold reached through an evolving regulatory process rather than a randomly assigned intervention. Systems may already be responding to informal pressure, case development, technical assistance, compliance support, or anticipated escalation before the quarter in which formal enforcement is recorded. The analysis also does not observe the timing or intensity of technical assistance, funding navigation, engineering support, or other support activities that may occur before or alongside formal enforcement. The pre-enforcement diagnostics are consistent with these concerns. The estimates should therefore be interpreted as changes in compliance progress around a broader administrative threshold that may combine regulatory attention, compliance support, and formal accountability, not as definitive causal effects of formal escalation alone.

A second limitation concerns measurement in administrative records. The recorded quarter of first formal enforcement may not align precisely with when enforcement pressure becomes behaviorally consequential. Formal action is often preceded by warnings, negotiations, and other informal contacts that are not fully observed in the data. Similarly, the outcome captures recorded quarterly resolution activity, not the full process through which violations are corrected on the ground. The measured timing of both enforcement and resolution may reflect administrative sequencing and recording practices as well as substantive compliance behavior.

This limitation is especially important in a drinking water setting where compliance recovery can depend on engineering, financing, and documentation steps that unfold over uneven time horizons.

The chapter is also limited by the level and construction of the outcome. The analysis uses a system-quarter measure of benchmark-adjusted compliance progress rather than a violation-level outcome. This aggregation is useful for tracing system-level dynamics around formal escalation and for placing treated systems' quarterly resolution activity in a common calendar-time context. However, it compresses variation across violations within the same system. Violations differ in severity, technical complexity, and public health significance, yet the quarterly resolution-rate measure does not fully distinguish among those differences. More generally, the outcome captures the pace of recorded resolution among open violations relative to the same-calendar-quarter benchmark, not the quality or durability of the corrective action that produced closure, direct changes in delivered water quality, or household-level exposure reduction.

External validity is also limited. The event-study estimation sample consists of systems that eventually receive a first formal enforcement action and contribute observations within the event-time window. As shown earlier in the chapter, this sample contains higher shares of disadvantaged and severely disadvantaged community systems than the full system-quarter dataset. The findings therefore generalize most directly to the population of systems that progress to formal escalation, not to all California community water systems. This distinction is consequential, as systems that never receive formal enforcement may differ systematically in compliance history, capacity, and regulatory visibility.

The heterogeneity analyses also require caution. Community-context measures such as DAC, SDAC, and racial and ethnic service-area indicators are system-level contextual proxies rather than direct measures of household exposure or individual risk. They capture the demographic characteristics of the populations served by each system, but they do not measure within-system inequality or individual-level differences in access and vulnerability. In addition, several moderators are constructed as binary thresholds, which improves interpretability but necessarily compresses variation. The weaker and less stable subgroup patterns should therefore not be overinterpreted.

Finally, this chapter is specific to California's SDWA implementation context during the study period. California's institutional setting, primacy structure, and mix of publicly owned community water systems may differ from those of other states. The extent to which these findings travel beyond California is therefore uncertain. California nevertheless remains substantively important given that the state contains a large and diverse population of community water systems and a regulatory environment in which questions of capacity, inequality, and formal escalation are especially salient.

4.6.3 Future Directions

Several directions follow from this chapter. First, future research should pursue designs that better isolate exogenous variation in enforcement timing and severity. Given the endogenous nature of first formal enforcement as an administrative threshold, stronger causal identification will require variation less tightly linked to prior compliance history, informal pressure, and regulator discretion. A related future direction is to extend the design beyond California to compare states with different enforcement styles, regulatory capacities, and reliance on formal enforcement tools. Cross-state variation could provide more informative comparison groups,

including systems with similar violation profiles that do not receive formal enforcement within the same time frame. Such designs would help distinguish changes associated with formal escalation from changes associated with underlying violation severity, state enforcement culture, or broader compliance-support infrastructure.

Second, future work should examine more directly the mechanisms through which formal enforcement translates, or fails to translate, into compliance progress. The present chapter shows that post-enforcement gains are modest on average and weaker in disadvantaged community systems, but it cannot determine precisely whether that pattern reflects technical infeasibility, financial constraint, managerial weakness, differential oversight intensity, or delayed administrative recording. More detailed administrative data on technical assistance, financing access, engineering milestones, and informal regulator-system interaction would help clarify these pathways.

Third, qualitative and mixed-method research would be especially valuable for developing a deeper understanding of the regulatory and compliance behaviors that underlie the quantitative patterns documented here. Interviews, case studies, and archival administrative analysis could clarify how regulators decide when formal escalation is warranted, how water systems interpret and respond to different enforcement signals, and why formal action produces weaker gains in some settings than in others. Such work would be particularly useful for understanding the practical tradeoffs regulators face as both enforcers and problem-solving institutions, and for clarifying how capacity, feasibility, trust, and institutional relationships shape post-enforcement trajectories.

Fourth, future work should examine more explicitly how formal enforcement interacts with corrective assistance. One implication of this chapter is that formal escalation may be an

imperfect tool where the underlying drivers of noncompliance are structural. Research that tracks the joint effects of formal enforcement, technical assistance, consolidation efforts, and financing interventions would be especially useful for identifying which combinations of tools are most effective in different system contexts.

Finally, future research should connect enforcement dynamics more directly to downstream exposure and public health outcomes. This chapter evaluates compliance progress as an administratively meaningful indicator of regulatory performance, but an important next step is to assess whether faster post-enforcement progress reduces exposure duration or health risk in measurable ways. Linking enforcement trajectories to contaminant monitoring or health-related outcomes would help clarify the public health significance of the uneven enforcement response documented here.

4.7 Conclusion

This chapter examined whether first formal enforcement is followed by measurable improvement in quarterly compliance progress and whether those dynamics differ across community contexts. The results support a restrained conclusion. Formal escalation is associated with limited improvement in compliance progress, and that improvement is neither immediate nor durable. More importantly, the average pattern conceals unequal post-enforcement response across systems. Disadvantaged community systems experience smaller gains than non-DAC systems, indicating that the same formal enforcement step does not produce comparable corrective momentum across community contexts.

These findings contribute to a more implementation-centered understanding of SDWA enforcement. Formal escalation is significant, but its practical effects are modest, delayed, and uneven. This pattern is consistent with a regulatory setting in which many compliance barriers

are technical, financial, and organizational rather than purely motivational. In such settings, formal action may clarify obligations and structure accountability, but it may be less capable of generating rapid or sustained compliance gains on its own. The findings therefore help explain why SDWA implementation continues to rely heavily on sequenced oversight, informal engagement, and problem-solving support rather than frequent resort to strongly punitive or terminal enforcement.

The chapter also sharpens the dissertation's broader unequal-protection argument. Unequal protection does not arise only from greater exposure burdens or longer recovery trajectories. It may also arise when formal regulatory escalation does not translate into equivalent compliance progress once enforcement occurs. In that sense, this chapter complements Chapter 3. Chapter 3 showed that compliance recovery is shaped by underlying capacity constraints, especially in disadvantaged and smaller-system settings. Chapter 4 shows that even after systems cross into formal enforcement, the resulting gains in compliance progress remain limited and are smaller in DAC systems. Viewed together, the two chapters suggest that unequal protection operates through both unequal recovery capacity and unequal post-escalation responsiveness.

4.8 References

- Allaire, M., & Acquah, S. (2022). Disparities in drinking water compliance: Implications for incorporating equity into regulatory practices. *AWWA Water Science*, 4(2), e1274.
- Allaire, M., Wu, H., & Lall, U. (2018). National trends in drinking water quality violations. *Proceedings of the National Academy of Sciences*, 115(9), 2078–2083.
- Arguedas, C., & Rousseau, S. (2012). Learning about compliance under asymmetric information. *Resource and Energy Economics*, 34(1), 55–73.
- Association of State Drinking Water Administrators. (2013, December). *Insufficient resources for state drinking water programs threaten public health: An analysis of state drinking water programs' resources and needs*. <https://www.asdwa.org/wp-content/uploads/2017/03/SRNAP-Analysis.pdf>

- Ayres, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. Oxford University Press.
- Baldwin, R., & Black, J. (2008). Really responsive regulation. *The Modern Law Review*, 71(1), 59–94.
- Barreca, A. I., Lindo, J. M., & Waddell, G. R. (2016). Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1), 268–293.
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217.
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.
- Braithwaite, J. (2011). The essence of responsive regulation. *UBC Law Review*, 44(3), 475–520.
- Breyer, S. G. (1982). *Regulation and its reform*. Harvard University Press.
- Callaway, B., & Sant’Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- California State Water Resources Control Board. (2023, December 5). Water quality enforcement policy. https://www.waterboards.ca.gov/water_issues/programs/enforcement/docs/2024/2024-enforcement-policy.pdf
- Clarke, D., & Tapia-Schyte, K. (2021). Implementing the panel event study. *The Stata Journal*, 21(4), 853–884.
- Cory, D. C., & Rahman, T. (2009). Environmental justice and enforcement of the Safe Drinking Water Act: The Arizona arsenic experience. *Ecological Economics*, 68(6), 1825–1837.
- Dahl, G. B., Løken, K. V., & Mogstad, M. (2014). Peer effects in program participation. *American Economic Review*, 104(7), 2049–2074.
- Deily, M. E., & Gray, W. B. (1991). Enforcement of pollution regulations in a declining industry. *Journal of Environmental Economics and Management*, 21(3), 260–274.
- Earnhart, D. (2004). Panel data analysis of regulatory factors shaping environmental performance. *Review of Economics and Statistics*, 86(1), 391–401.
- Earnhart, D., & Jacobson, S. (2025). How does the deterrence effect of regulatory enforcement differ between privately and government-owned facilities? *Environmental and Resource Economics*, 88(3), 553–588.
- Earnhart, D., & Segerson, K. (2012). The influence of financial status on the effectiveness of environmental enforcement. *Journal of Public Economics*, 96(9–10), 670–684.
- Gray, W. B., & Shimshack, J. P. (2011). The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*, 5(1), 3–24.
- Glicksman, R. L., & Earnhart, D. H. (2007). The comparative effectiveness of government interventions on environmental performance in the chemical industry. *Stanford Environmental Law Journal*, 26, 317–371.

- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Gunningham, N., & Sinclair, D. (2017). Smart regulation. In P. Drahos (Ed.), *Regulatory theory: Foundations and applications* (pp. 133–148). ANU Press.
- Hausman, C., & Rapson, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, 10(1), 533–552.
- Kirchhoff, C. J., Flagg, J. A., Zhuang, Y., & Utemuratov, B. (2019). Understanding and improving enforcement and compliance with drinking water standards. *Water Resources Management*, 33(5), 1647–1663.
- Konisky, D. M. (2015). Federal environmental justice policy: Lessons learned. In D. M. Konisky (Ed.), *Failed promises: Evaluating the federal government's response to environmental justice* (pp. 233–257). MIT Press.
- Konisky, D. M., & Reenock, C. (2013). Compliance bias and environmental (in)justice. *The Journal of Politics*, 75(2), 506–519.
- Konisky, D. M., & Reenock, C. (2015). Evaluating fairness in environmental regulatory enforcement. In D. M. Konisky (Ed.), *Failed promises: Evaluating the federal government's response to environmental justice* (pp. 173–203). MIT Press.
- Konisky, D. M., & Reenock, C. (2018). Regulatory enforcement, riskscapes, and environmental justice. *Policy Studies Journal*, 46(1), 7–36.
- Konisky, D. M., & Teodoro, M. P. (2016). When governments regulate governments. *American Journal of Political Science*, 60(3), 559–574.
- Levi-Faur, D., Kariv-Teitelbaum, Y., & Medzini, R. (2021, May 26). Regulatory governance: History, theories, strategies, and challenges. In *Oxford Research Encyclopedia of Politics*. Oxford University Press.
<https://doi.org/10.1093/acrefore/9780190228637.013.1430>
- McKean, R. N. (1980). Enforcement costs in environmental and safety regulation. *Policy Analysis*, 6(3), 269–289.
- Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2), 203–230.
- Polinsky, A. M., & Shavell, S. (2000). The economic theory of public enforcement of law. *Journal of Economic Literature*, 38(1), 45–76.
- Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Rousseau, S. (2007, May). *Economic empirical analysis of sanctions for environmental violations: A literature overview* (Energy, Transport and Environment Working Paper Series ETE-WP-2007-03). KU Leuven, Department of Economics, Research Group Energy, Transport and Environment.
<https://feb.kuleuven.be/research/economics/ces/documents/DPS/ETE/ETE-WP-2007-03.pdf>

- Salzman, J., & Mackay Peltzer, A. (2025). The cleanest water and the dumbest kids: Do small water systems comply with the Safe Drinking Water Act? *Stanford Environmental Law Journal*, 44, 276–338.
- Scheberle, D. (2004). *Federalism and environmental policy: Trust and the politics of implementation* (2nd ed., rev. and updated). Georgetown University Press.
- Schmidheiny, K., & Siegloch, S. (2023). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*, 38(5), 695–713.
- Shimshack, J. P. (2014). The economics of environmental monitoring and enforcement. *Annual Review of Resource Economics*, 6(1), 339–360.
- Shimshack, J. P., & Ward, M. B. (2005). Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*, 50(3), 519–540.
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
- U.S. Environmental Protection Agency. (n.d.). *Basic information on enforcement*. <https://www.epa.gov/enforcement/basic-information-enforcement>
- U.S. Environmental Protection Agency. (2009a, December 8). *Enforcement Response Policy for the Public Water System Supervision (PWSS) Program under the Safe Drinking Water Act and implementation of the Enforcement Targeting Tool*. <https://www.epa.gov/enforcement/enforcement-response-policy-public-water-system-supervision-pwss-program-under-safe>
- U.S. Environmental Protection Agency. (2009b, December 8). *Proposed revision to Enforcement Response Policy for the Public Water System Supervision (PWSS) Program under the Safe Drinking Water Act and implementation of the Enforcement Targeting Tool* [Memorandum]. Office of Enforcement and Compliance Assurance. https://www.epa.gov/sites/default/files/documents/drinking_water_erp_2009.pdf
- U.S. Environmental Protection Agency. (2009c, December 8). *Drinking water enforcement response policy* [Memorandum]. <https://www.epa.gov/sites/default/files/2015-09/documents/drinking-water-erp-2009.pdf>
- U.S. Environmental Protection Agency, Office of Water. (2012, March 30). *Update to the implementation of capacity development & drinking water state revolving fund programs to reflect the new enforcement policy & enforcement targeting tool* [Memorandum]. <https://www.epa.gov/sites/default/files/2015-04/documents/erpcttmemo.pdf>
- U.S. Environmental Protection Agency. (2024, December 5). *Safe Drinking Water Act (SDWA) resources and FAQs*. <https://echo.epa.gov/help/sdwa-faqs>
- U.S. Environmental Protection Agency. (2025, December 9). *SDWA data download summary and data element dictionary*. ECHO: Enforcement and Compliance History Online. <https://echo.epa.gov/tools/data-downloads/sdwa-download-summary#dataelement>
- U.S. Environmental Protection Agency. (2026, February 24). *Criminal provisions of the Safe Drinking Water Act (SDWA)*. <https://www.epa.gov/enforcement/criminal-provisions-safe-drinking-water-act-sdwa>

Viscusi, W. K. (2000). Risk equity. *The Journal of Legal Studies*, 29(S2), 843–871.

Wilson, J. Q., & Rachal, P. (1977). Can the government regulate itself? *The Public Interest*, 46, 3–14.

APPENDIX C: Chapter 4 Supplemental Tables and Figures

Appendix Table 4.A1: Key Variable Definitions for the Chapter 4 Event-Study Analysis

Variable	Type	Definition
Panel and variables		
pwsid_id	Identifier	Public Water System identifier.
tq	Integer	Calendar quarter used as panel time.
gq	Integer	First quarter in which system i receives a formal enforcement action; there is at most one gq per system.
relq (event time k)	Integer (... , -2, -1, 0, 1, 2, ...)	Counts quarters relative to the first formal enforcement action for each system, defined as $tq - gq$. Negative values are pre-enforcement and positive values are post-enforcement.
k_ix	Integer (1-19)	Shifted event time used for factor indicators. It is defined as $relq + 7$. The omitted base is 5, corresponding to $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation.
Outcome and benchmark variables		
resolved	Numeric	Number of violations resolved in quarter tq for a given system.
open_start	Numeric	Number of violations open at the start of quarter tq for a given system.
rate_t	Numeric [0,1]	Treated-system quarterly resolution rate, defined as resolved violations divided by violations open at the start of the quarter when the denominator is defined.
y_c	Numeric [0,1]	Same-calendar-quarter benchmark resolution rate.

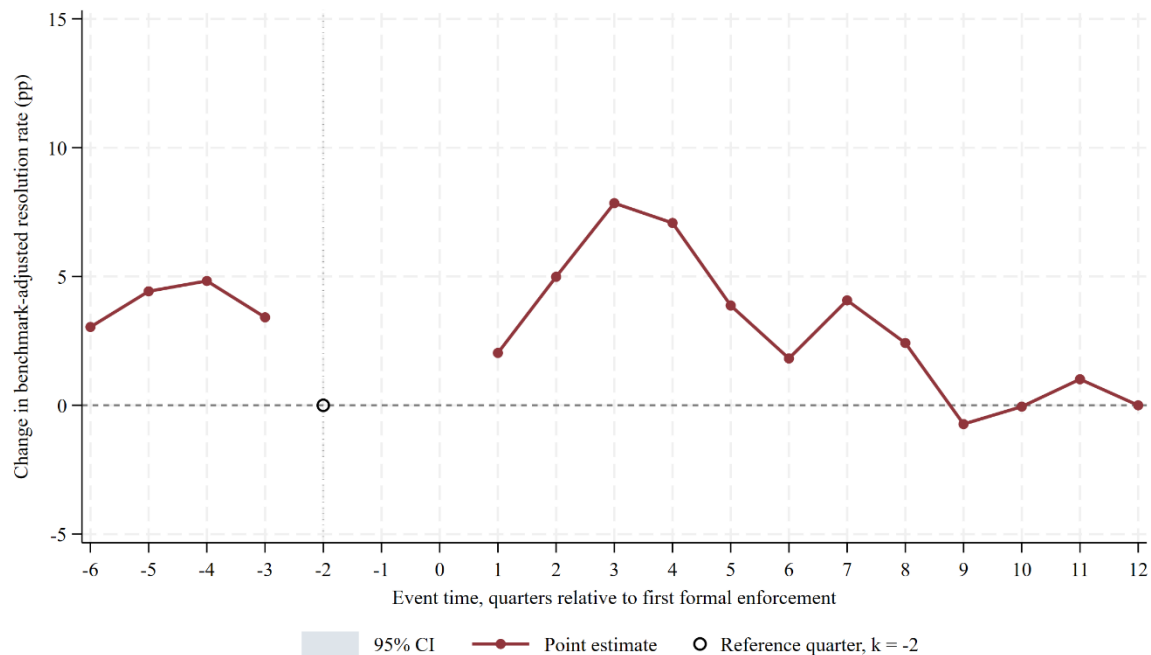
		Operationalized as the mean resolution rate among systems that do not receive first formal enforcement during the panel.
diff_es_pooled	Numeric (percentage points)	Benchmark-adjusted quarterly resolution outcome, defined as the treated system's quarterly resolution rate minus the same-calendar-quarter benchmark, expressed in percentage points.
Community-context variables		
african	Binary	Equals 1 if the system is an African American-serving system, defined as a system whose served-population African American share exceeds the corresponding statewide average benchmark over the study period.
hispanic	Binary	Equals 1 if the system is a Hispanic-serving system, defined as a system whose served-population Hispanic share exceeds the corresponding statewide average benchmark over the study period.
DAC	Binary	Equals 1 if the system serves a disadvantaged community (DAC).
SDAC	Binary	Equals 1 if the system serves a severely disadvantaged community (SDAC).
DACxhispanic	Binary	Equals 1 if the system is both DAC and Hispanic-serving.
Weighting variable		
w_open_norm	Numeric	Normalized analytic weight proportional to the number of violations open at the start of the quarter.

Note. This appendix table lists the principal variables used in this chapter. Missingness is reported in Appendix Table 4.A2.

Appendix Table 4.A2: Missingness and Event-Study Sample Coverage for Key Variables

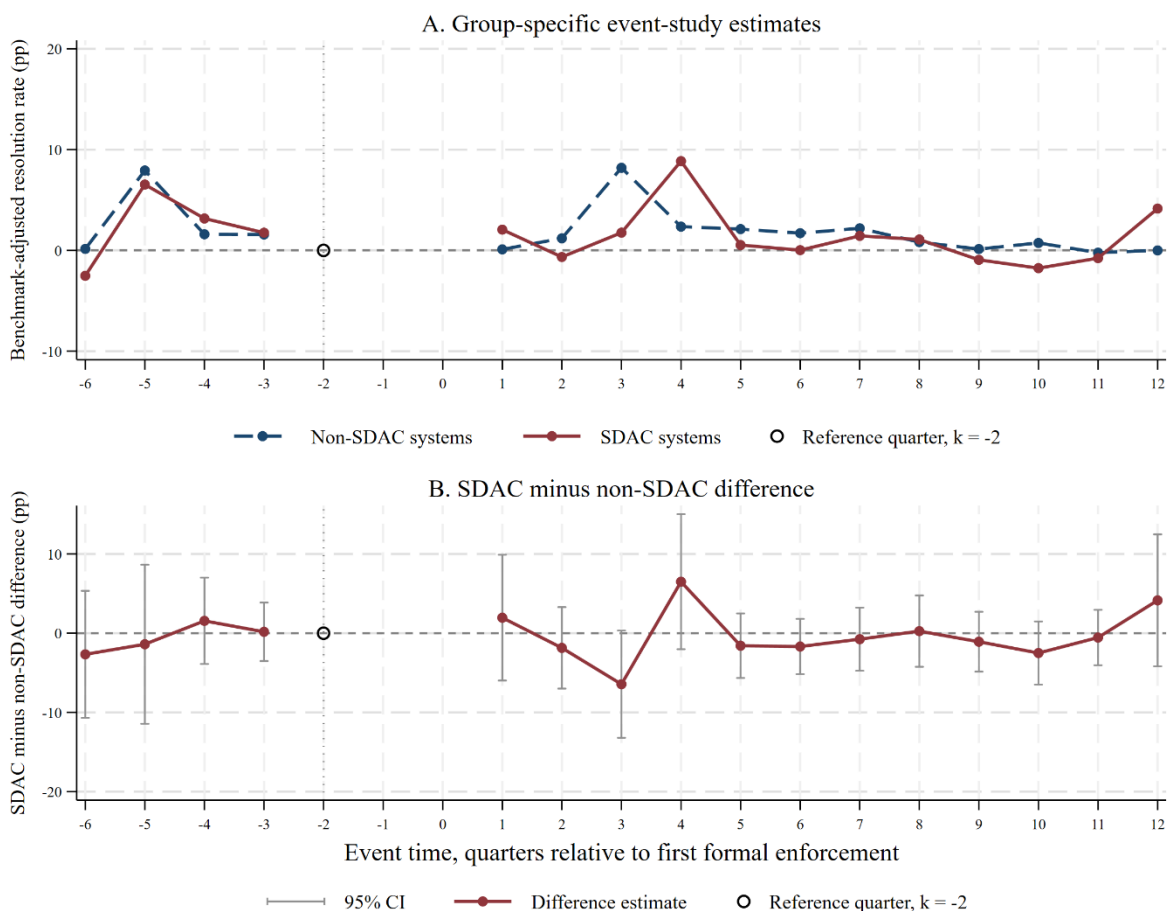
Variable	Full sample			Non-donut event-study window sample		
	Total N	Missing n	Missing %	Total N	Missing n	Missing %
diff_es_pooled	286,770	257,073	89.6	6,961	2	0.03
w_open_norm	286,770	250,448	87.3	6,961	0	0.0
gq	286,770	117,612	41.0	6,961	0	0.0
relq	286,770	117,612	41.0	6,961	0	0.0
k_ix	286,770	260,915	91.0	6,961	0	0.0
rate_t	286,770	250,448	87.3	6,961	0	0.0
y_c	286,770	4,740	1.7	6,961	2	0.03
open_start	286,770	0	0.0	6,961	0	0.0
resolved	286,770	250,448	87.3	6,961	0	0.0

Note. The full panel contains 286,770 system-quarters. The basic event-study window sample contains 6,961 treated system-quarters in the non-donut event window, excluding the donut quarters $k = -1$ and $k = 0$. The main benchmark-adjusted event-study regression uses 6,959 observations because two treated system-quarters are missing the same-calendar-quarter benchmark y_c and therefore also missing the constructed outcome diff_es_pooled . Variables that depend on active violations, event-time assignment, or treated-system-quarter outcome construction are missing for many full-panel rows by design because many system-quarters have no open violations, are not assigned event time, or do not contribute to the treated event-study outcome.



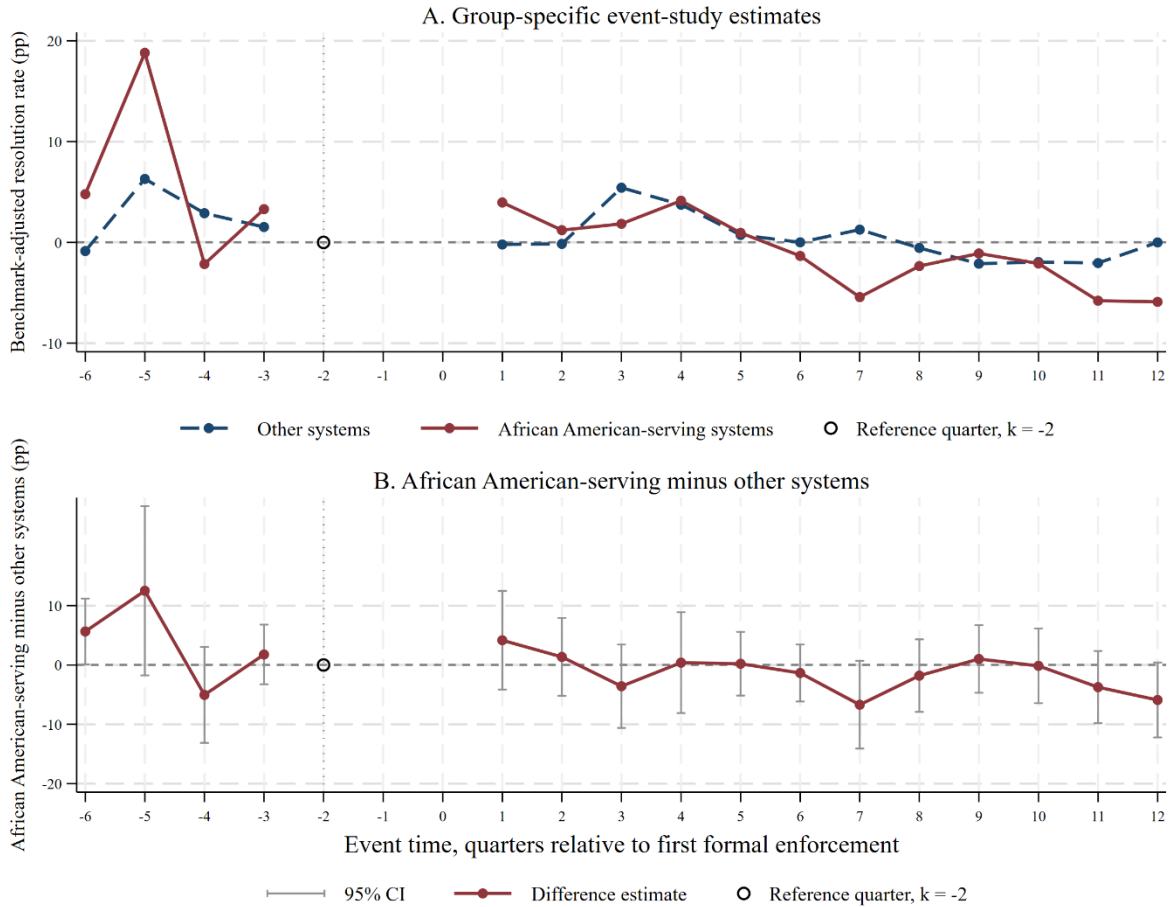
Appendix Figure 4.A1: Unweighted Event-Study Estimates of Quarterly Compliance Progress Around First Formal Enforcement

Note. Estimates are from the unweighted event-study specification. The outcome is benchmark-adjusted quarterly compliance progress, measured in percentage points as the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark among systems that do not receive first formal enforcement during the study panel. The omitted reference period is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation. Vertical bars show 95 percent confidence intervals. Models include system and calendar-quarter fixed effects, with standard errors clustered by system.



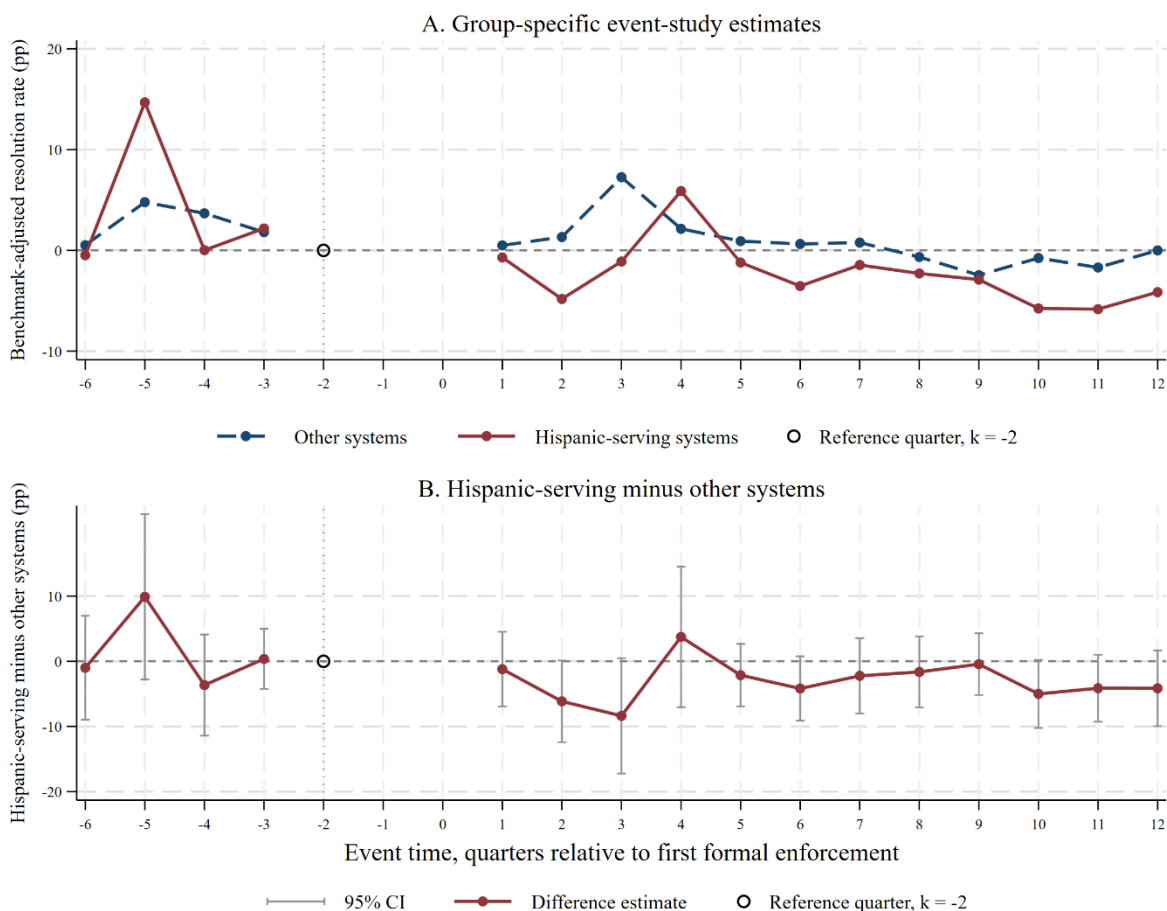
Appendix Figure 4.A2: Heterogeneity by Severely Disadvantaged Community Status in Weighted Event-Study Estimates of Quarterly Compliance Progress

Note. Estimates are from the weighted event-study specification for the indicated community-context comparison. The outcome is benchmark-adjusted quarterly compliance progress, measured in percentage points as the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark among systems that do not receive first formal enforcement during the study panel. The omitted reference period is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation. Vertical bars show 95 percent confidence intervals. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Analytic weights are based on normalized open violations at the start of the quarter.



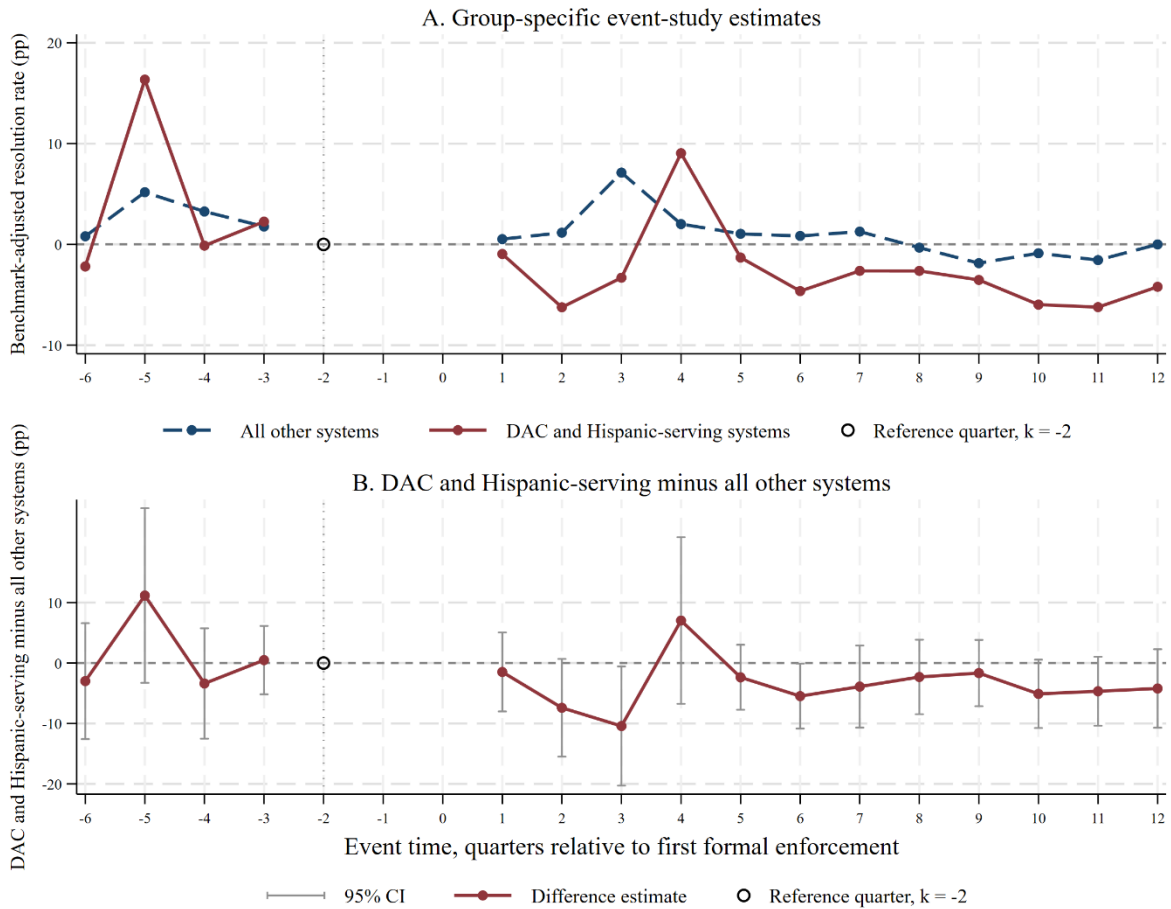
Appendix Figure 4.A3: Heterogeneity by African American-Serving System Status in Weighted Event-Study Estimates of Quarterly Compliance Progress

Note. Estimates are from the weighted event-study specification for the indicated community-context comparison. The outcome is benchmark-adjusted quarterly compliance progress, measured in percentage points as the treated system’s own quarterly resolution rate minus the same-calendar-quarter benchmark among systems that do not receive first formal enforcement during the study panel. The omitted reference period is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation. Vertical bars show 95 percent confidence intervals. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Analytic weights are based on normalized open violations at the start of the quarter.



Appendix Figure 4.A4: Heterogeneity by Hispanic-Serving System Status in Weighted Event-Study Estimates of Quarterly Compliance Progress

Note. Estimates are from the weighted event-study specification for the indicated community-context comparison. The outcome is benchmark-adjusted quarterly compliance progress, measured in percentage points as the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark among systems that do not receive first formal enforcement during the study panel. The omitted reference period is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation. Vertical bars show 95 percent confidence intervals. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Analytic weights are based on normalized open violations at the start of the quarter.



Appendix Figure 4.A5: Heterogeneity by DAC and Hispanic-Serving System Status in Weighted Event-Study Estimates of Quarterly Compliance Progress

Note. Estimates are from the weighted event-study specification for the indicated community-context comparison. The outcome is benchmark-adjusted quarterly compliance progress, measured in percentage points as the treated system's own quarterly resolution rate minus the same-calendar-quarter benchmark among systems that do not receive first formal enforcement during the study panel. The omitted reference period is $k = -2$. Donut quarters $k = -1$ and $k = 0$ are excluded from estimation. Vertical bars show 95 percent confidence intervals. Models include system and calendar-quarter fixed effects, with standard errors clustered by system. Analytic weights are based on normalized open violations at the start of the quarter.

Appendix Table 4.A3: Unweighted Community-Context Differences in Horizon-Averaged Post-Enforcement Compliance Progress

Comparison: focal group minus reference group	Short run, quarters 1–3	Mid run, quarters 4–8	Long run, quarters 9–12
DAC systems minus non-DAC systems	-5.35 [-9.82, -0.87]**	-3.90 [-8.12, 0.32]*	-4.14 [-8.49, 0.20]*
SDAC systems minus non-SDAC systems	-3.61 [-8.52, 1.29]	-1.77 [-6.39, 2.84]	-2.19 [-6.91, 2.54]
Hispanic-serving systems minus other systems	-3.26 [-8.53, 2.01]	-3.41 [-8.39, 1.57]	-3.29 [-8.46, 1.88]
DAC and Hispanic-serving systems minus all other systems	-4.49 [-11.02, 2.04]	-3.61 [-9.85, 2.62]	-4.18 [-10.73, 2.37]
African American-serving systems minus other systems	8.37 [-0.37, 17.12]*	8.24 [-0.14, 16.61]*	8.60 [0.39, 16.80]**

Note. Entries report focal-minus-reference differences in equal-weighted horizon-averaged post-enforcement compliance progress from the unweighted event-study specification, expressed in percentage points. Each row is estimated using the same system-quarter event-study framework as the baseline analysis, not a separate subgroup-only dataset. The focal group is named first in each comparison, and the reference group is named second. Negative estimates indicate weaker post-enforcement compliance progress for the focal group relative to the reference group. The outcome is benchmark-adjusted quarterly compliance progress, measured as the treated system’s own quarterly resolution rate minus the same-calendar-quarter benchmark among systems that do not receive first formal enforcement during the study panel. Short-run, mid-run, and long-run horizons correspond to post-enforcement quarters 1–3, 4–8, and 9–12, respectively. Donut quarters $k = -1$ and $k = 0$ are excluded, and $k = -2$ is the omitted reference period. Hispanic-serving and African American-serving systems are defined as systems whose served-population shares exceed the corresponding statewide average benchmarks over the study period. The unweighted specification treats each observed system-quarter equally. Ninety-five percent confidence intervals are shown in brackets. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table 4.A4: Horizon-Window Sensitivity for Weighted Post-Enforcement Estimates

Horizon	Averaging scheme	Window (k)	Effect (pp)	95% CI (pp)	p-value
Mid run	Equal-weighted	4–6	1.77	[-0.79, 4.32]	0.174
Mid run	Exponential	4–6	2.08	[-0.66, 4.83]	0.136
Mid run	Equal-weighted	4–7	1.57	[-0.80, 3.94]	0.193

Mid run	Exponential	4–7	1.90	[-0.67, 4.47]	0.147
Mid run	Equal-weighted	4–8	1.23	[-1.22, 3.67]	0.325
Mid run	Exponential	4–8	1.67	[-0.91, 4.24]	0.204
Long run	Equal-weighted	9–10	-1.33	[-4.32, 1.65]	0.380
Long run	Exponential	9–10	-1.34	[-4.29, 1.61]	0.373
Long run	Equal-weighted	9–11	-1.48	[-4.47, 1.51]	0.331
Long run	Exponential	9–11	-1.45	[-4.40, 1.50]	0.334
Long run	Equal-weighted	9–12	-1.11	[-3.35, 1.13]	0.331
Long run	Exponential	9–12	-1.21	[-3.67, 1.25]	0.334

Note. Entries report weighted horizon-averaged post-enforcement estimates under alternative mid-run and long-run window caps. The mid-run horizon is recomputed over quarters 4–6, 4–7, and 4–8 after first formal enforcement, while the long-run horizon is recomputed over quarters 9–10, 9–11, and 9–12. The outcome is quarterly compliance progress, measured in percentage points as the treated system’s own quarterly resolution rate minus the same-calendar-quarter benchmark. All estimates come from the weighted two-way fixed-effects event-study specification with system and calendar-quarter fixed effects. Analytic weights are based on normalized open violations at quarter start. Donut quarters $k = -1$ and $k = 0$ are excluded, and the omitted reference period is $k = -2$. Ninety-five percent confidence intervals are shown in brackets.

Chapter 5: Conclusion

5.1 Overview

This dissertation examines whether drinking water governance under the SDWA in California produces patterns consistent with unequal protection, such that populations and places facing greater vulnerability or fewer resources experience higher exposure-related health risks, longer persistence of noncompliance, and weaker improvement following enforcement escalation. This objective is pursued through three complementary essays that analyze distinct but policy-relevant dimensions of risk and implementation performance: associations between nitrate exposure in community water systems and adverse birth outcomes in Chapter 2, persistence in time to return to compliance and its heterogeneity across capacity-linked contexts in Chapter 3, and compliance progress around first formal enforcement escalation and whether that progress differs across disadvantaged communities in Chapter 4. These essays provide coordinated evidence on exposure burden, remediation feasibility, and regulatory responsiveness as key pathways through which environmental justice concerns may emerge in SDWA implementation.

5.2 Key Findings

Chapter 2 shows that nitrate exposure in community water systems is associated with adverse birth outcomes, with the clearest and most consistent pattern appearing for preterm birth. Positive associations appear for preterm birth under continuous, categorical, and binary threshold parameterizations, while the low birthweight results are weaker and more sensitive to specification. The clearest heterogeneity signal appears in the high-exposure threshold comparison for non-Hispanic Black mothers, whereas continuous slopes are broadly similar

across maternal race and ethnicity groups, and place-based heterogeneity by DAC, SDAC, and Central Valley status is limited.

Chapter 3 shows that unequal protection also appears after violations occur, through uneven time to recorded return to compliance. The chapter does not identify a single uniform demographic penalty. Instead, the clearest empirical pattern is the interaction between SDAC status and system size. Among non-SDAC systems, recorded resolution times show minimal variation by size, whereas within SDAC systems the clearest adjusted contrast is between large systems and the small-to-medium system categories. This pattern is best interpreted as a capacity-gradient contrast within SDAC systems rather than as a uniform SDAC penalty among the smallest systems. The results show that unequal recovery is concentrated in capacity-constrained settings rather than expressed as a uniform disadvantage across all systems. Across specifications, prior violation burden and concurrent compliance challenges also emerge as strong predictors of prolonged noncompliance duration.

Chapter 4 further shows that formal enforcement escalation is associated, at most, with modest and temporary improvement in quarterly compliance progress, and that these gains are distributed unevenly across community contexts. Improvement is limited immediately after escalation, becomes more visible several quarters later, and attenuates over time rather than persisting as a durable shift. The most consistent heterogeneity appears by DAC status, with DAC systems showing smaller post-enforcement gains than non-DAC systems, while SDAC and other demographic contrasts are weaker and less stable. These findings indicate that formal escalation marks a meaningful accountability threshold, but formal action alone does not generate comparable corrective momentum across all systems. Figure 5.1 summarizes these

findings across the dissertation’s three stages of unequal protection: exposure burden, compliance recovery, and regulatory responsiveness.

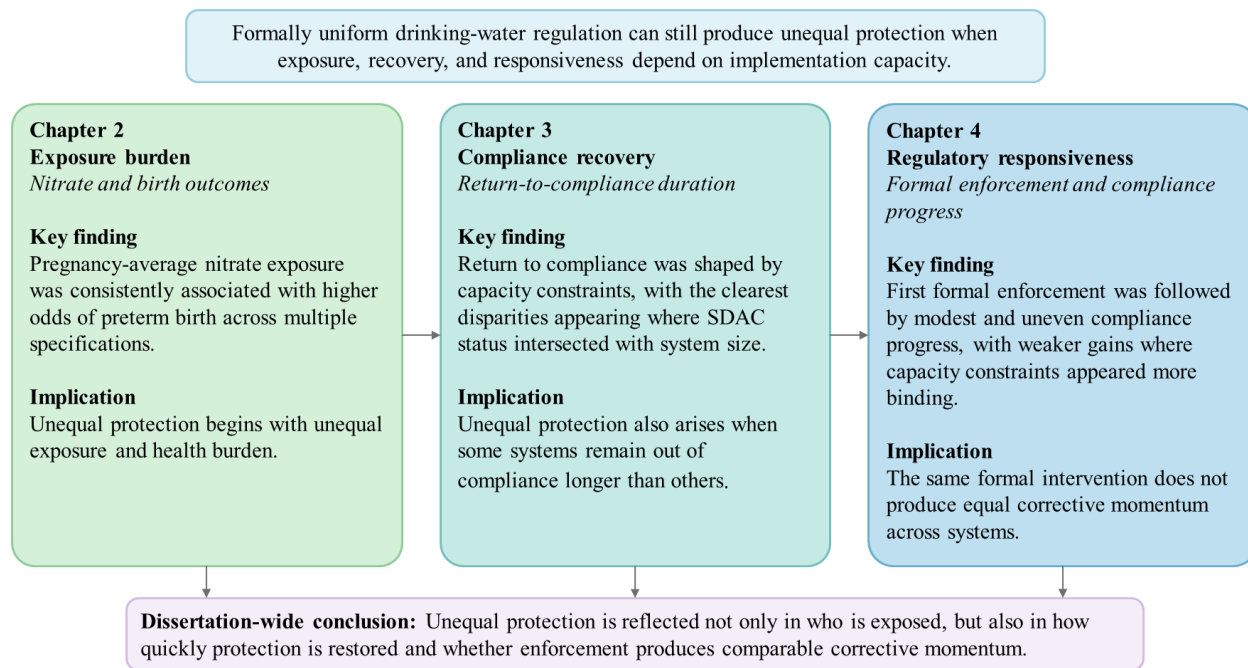


Figure 5.1: Integrated Findings Across Three Stages of Unequal Protection

Note. This figure summarizes how the three empirical chapters jointly support the dissertation-wide argument. Unequal protection under the SDWA appears across exposure burden, compliance recovery, and regulatory responsiveness, showing that formally uniform standards can produce uneven protection when implementation capacity differs across systems and communities.

5.3 Cross-essay Synthesis

Across the dissertation, unequal protection emerges not through a single point of failure, but through multiple stages of governance that shape whether protection is delivered in time. The central conclusion is that inequality in SDWA implementation is produced not only by unequal exposure, but also by unequal capacity to convert monitoring, compliance processes, and enforcement action into timely and durable restoration. Environmental justice concerns therefore

arise not only when contamination burdens are unevenly distributed, but also when communities differ in how quickly risks are corrected, verified, and resolved.

The three essays show that regulatory performance should be understood in terms of restored protection rather than procedural movement alone. Longer noncompliance spells mean longer periods during which protection remains incomplete, and formal escalation matters not simply as an administrative event, but as part of a broader process through which systems and regulators move from recognized risk to completed corrective action. The central cross-essay question is therefore whether governance arrangements shorten incomplete protection where restoration is most difficult. Unequal protection arises not only through who is exposed or who receives regulatory attention, but also through whether systems can recover once problems are identified and whether formal enforcement produces comparable corrective progress across community contexts.

The dissertation's contribution is to link exposure, compliance recovery, and post-enforcement responsiveness as sequential stages in the delivery of regulatory protection. These stages are often treated as separate empirical or policy questions. This dissertation brings them into a single implementation framework and shows that unequal protection can enter when exposure occurs, while systems remain out of compliance, and after formal enforcement has begun. This staged perspective shifts the environmental justice question from whether regulation formally applies to all communities to whether regulation produces timely and durable restoration where protection is most difficult to deliver.

5.4 Policy Implications

A central policy implication is that drinking-water governance should be judged by whether it restores protection in time, not only by whether standards exist or enforcement actions

occur. Agencies should evaluate performance partly through measures such as time to return to compliance, the share of violations that remain unresolved beyond defined thresholds, recurrence of violations, and whether formal escalation is followed by measurable improvement. A regulatory system can satisfy procedural expectations while still allowing exposure risk to persist when detection, remediation, and follow-through are slow or uneven. Timeliness should therefore be treated as a core dimension of public health protection rather than as a secondary administrative concern.

This perspective also implies that exposure prevention should not be limited to formal exceedance response. Where contaminants such as nitrate show health-relevant variation across a broader exposure gradient, policy attention should include systems and communities facing persistent moderate exposure, recurring vulnerability, or elevated risk among sensitive populations. In practice, this means combining long-term source reduction and treatment investment with near-term protection measures such as targeted testing, risk communication, interim safe-water provision, and household or system-level treatment where appropriate and properly managed. These measures are especially important in places where groundwater contamination, small-system capacity constraints, and community vulnerability make durable restoration difficult to achieve quickly.

A second implication is that formal enforcement should be tied to a defined assistance protocol in disadvantaged and other capacity-constrained settings. When a system receives a first formal enforcement action, or when a violation remains unresolved beyond a defined duration threshold, the case could enter a parallel assistance review within a fixed period. That review could assess technical barriers, managerial weaknesses, financing constraints, need for safe interim drinking water, and the feasibility of longer-term options such as consolidation,

extension of service, shared management, or regional partnership. The agency could then issue a written support plan identifying the main barriers to correction, the assistance pathway, responsible parties, and a timeline for follow-up.

Targeting should also be more specific than broad assistance to disadvantaged systems as a whole. Priority should go to systems where prolonged incomplete protection is most likely, especially those with repeated violations, concurrent compliance burdens, unresolved cases that persist well beyond normal time to recorded resolution, or limited evidence of progress after formal escalation. In practice, this means distinguishing between systems facing isolated and readily correctable problems and systems facing chronic, overlapping, or slow-moving compliance failures. A prioritization framework of this kind would direct scarce regulatory and assistance capacity more clearly toward the systems where incomplete protection is most likely to persist.

Capacity-building should likewise be defined more concretely and more broadly than infrastructure finance alone. Some systems need capital improvements, but others also need operator staffing, managerial continuity, engineering support, help with procurement and project development, administrative capacity to document return to compliance, and case-management follow-through after enforcement begins. Policy responses should therefore distinguish between systems whose main barrier is funding and systems whose main barrier is implementation capacity. In some cases, the most effective intervention may combine financing with direct technical assistance, temporary managerial support, shared management, extension of service, or structured movement toward consolidation.

In California, these recommendations could be integrated more explicitly into the state's human-right-to-water and SAFER framework. Persistent exposure risk, prolonged

noncompliance, repeated compliance burden, and stalled post-enforcement progress could be used as practical triggers for structured review of support eligibility and feasibility options. This would make equity commitments more operational by linking assistance to observable indicators of incomplete protection rather than leaving support decisions to informal judgment alone. In policy terms, environmental justice requires not only equal formal regulation, but also timely and durable restoration of safe service across communities.

5.5 Limitations

This dissertation should be interpreted as evidence on patterns of unequal protection rather than as a set of definitive causal estimates. Across the three essays, the analysis relies on administrative records, contextual proxies, and observational designs that make exposure, compliance recovery, and enforcement dynamics visible at scale, but with unavoidable measurement limits. The findings therefore identify where risk, delay, and uneven responsiveness are concentrated more clearly than they isolate every mechanism that produces those patterns.

The first limitation concerns the measurement of exposure and community context. In Chapter 2, exposure is assigned using residence at delivery and community water system monitoring rather than individual ingested dose, so within-system spatial variation, source blending, residential mobility during pregnancy, and unobserved avoidance behaviors may all contribute to exposure misclassification. More broadly, the dissertation relies on system-level contextual measures that are analytically useful for environmental justice interpretation but necessarily mask within-system heterogeneity. Single-contaminant and system-level models should therefore be read as evidence on patterned risk and inequality rather than as precise measures of individual exposure or vulnerability.

A second limitation is that the governance outcomes analyzed here are administrative measures that do not map perfectly onto on-the-ground correction or behavioral change. In Chapter 3, time to return to compliance reflects not only remediation, but also verification, documentation, and formal closeout, while unresolved spills are right-censored and several relevant capacity mechanisms remain only partially observed. In Chapter 4, first formal escalation is a discretionary institutional threshold that may be preceded by informal pressure and may not align exactly with when enforcement becomes behaviorally consequential. These features indicate that the results are best interpreted as evidence on recorded recovery and compliance dynamics around meaningful administrative thresholds. The external validity of these findings is most direct to California community water systems and, for Chapter 4, especially to systems that eventually enter the formal enforcement process.

5.6 Future Directions

Future research should strengthen measurement and identification across the three dimensions of unequal protection examined in this dissertation: exposure, compliance recovery, and post-enforcement responsiveness. For nitrate exposure and birth outcomes, the most important next steps are improved exposure assignment over pregnancy, including residential mobility, within-system spatial variation, source blending, household avoidance behavior, and contaminant mixtures. Future work could also make stronger use of quasi-experimental variation in delivered water quality, including within-mother comparisons, changes in source water, treatment installation, consolidation, and hydroclimatic variation such as drought and groundwater-level fluctuation. Drought and groundwater change are especially relevant for nitrate because they may alter source reliance, concentration patterns, and delivered exposure, but they are not clean causal shocks by themselves. Future designs would need to separate

nitrate-related pathways from other drought-related risks, including heat, air quality, economic stress, agricultural change, and household coping behavior.

For compliance recovery, future research should move closer to the mechanisms that connect formal regulation to restored protection. This requires better measurement of intermediate remediation milestones, technical and managerial capacity, financing constraints, project development, operator staffing, and administrative closeout. Mixed-methods research is also needed to investigate mechanisms that are difficult to observe in administrative records alone, including how systems diagnose problems, seek funding, retain operators, interact with regulators, manage procurement, and document corrective action. Future studies could then evaluate whether SAFER funding, technical assistance, administrator appointments, shared management, consolidation, extension of service, or other regional partnership strategies reduce time to recorded resolution. Where timing and eligibility rules permit, these questions could be studied using staggered implementation designs, duration models with intervention timing, or event-study approaches around the start of assistance.

For enforcement responsiveness, future work should evaluate not only whether formal enforcement occurs, but also which forms of follow-through make enforcement more effective. A key next step is to compare formal enforcement actions paired with corrective-action milestones, financing support, technical assistance, managerial intervention, or closer monitoring against actions that rely primarily on formal escalation alone. Such research would help distinguish enforcement formalization from the support and oversight that may follow it. Cross-state comparisons would also be valuable because states differ in regulatory capacity, enforcement practices, funding structures, administrative priorities, and approaches to technical assistance. Future work could examine whether similar formal enforcement actions produce

different compliance trajectories across states, and whether those differences reflect regulator-side capacity, agency resources, fiscal conditions, institutional priorities, or broader determinants of regulatory effort.

Future research can also build on this dissertation by connecting drinking-water governance more directly to lived exposure reduction and environmental justice outcomes. Case studies, interviews, household surveys, and administrative process tracing would help explain why recovery is slower and post-escalation gains are weaker in some community contexts, while also showing how regulatory delay is experienced by affected households. A key next step is to ask whether faster recovery and stronger regulatory follow-through actually reduce exposure duration, household coping costs, trust loss, affordability pressure, and health risk. More broadly, future work should integrate exposure burden, remediation feasibility, intervention timing, enforcement follow-through, and lived experience within a single framework. Such research would extend this dissertation's central contribution by showing not only where unequal protection occurs, but also which policy interventions most effectively produce timely, durable, and meaningful restoration across communities.