Tufts University
School of Engineering
Civil and Environmental Engineering

University of Illinois Chicago
Government Finance Research Center

A Systems Perspective on Contributing Factors to Safe Drinking Water Act Noncompliance

Christelle Khalaf, Ph.D.; Deborah A. Carroll, Ph.D.; Abigail Iuorio; Joseph C. Toland; Lauryn A. Spearing, Ph.D.

Project Reporting Period: August 16, 2024 – February 28, 2026
Contract No. 68HERH23P0157

Submitted to:
U.S. Environmental Protection Agency
Georgette A. Fogle, Environmental Protection Specialist
Office of Enforcement and Compliance Assurance
1200 Pennsylvania Ave., NW, Washington, DC 20460
Phone: (202) 564-6746
Email: fogle.georgette@epa.gov

Submitted by:

Civil and Environmental Engineering
Tufts University
200 College Avenue
Medford, MA 02155
Website: https://engineering.tufts.edu/cee/

Government Finance Research Center
University of Illinois Chicago
412 S. Peoria Street (MC 349)
Chicago, IL 60607
Website: https://gfrc.uic.edu/

Principal Investigator: Lauryn Spearing, Ph.D.
Co-Principal Investigators: Deborah A. Carroll, Ph.D. and Christelle Khalaf, Ph.D.

ACKNOWLEDGEMENTS

A Systems Perspective on Contributing Factors to Safe Drinking Water Act Noncompliance is a project funded by the U.S. Environmental Protection Agency. This report is based on work completed in the School of Engineering at Tufts University and the Government Finance Research Center at the University of Illinois Chicago. Its contents do not necessarily reflect the views of Tufts University, the Government Finance Research Center, the University of Illinois Chicago, or the U.S. Environmental Protection Agency.

The authors thank the representatives of the U.S. Environmental Protection Agency who shared their expertise and guided this project, including: Arlene Anderson, Esha Bhardwaj, Michael Chen, Thomas Crabtree, Leslie Cronkhite, Jill Dean, Dan Feuer, Georgette Fogle, Elsbeth Hearn, Matthew Heberling, Lauren Jones, Leah Kim, Carol King, Connor Lainson, Aaron Margolis, Michele McKeever, Brandon Mediate, Mariya Pak, Sean Paul, Heidi Rauch, Greg Savitske, Prashant Singh, Genevieve Soule, Alison Treglia, Allison West, William Wheeler, and Dane Wilson.

Suggested Citation:

Khalaf, Christelle, Deborah A. Carroll, Abigail Iuorio, Joseph C. Toland, and Lauryn A. Spearing. *A Systems Perspective on Contributing Factors to Safe Drinking Water Act Noncompliance*. Chicago, IL: Government Finance Research Center, 2026.

TABLE OF CONTENTS

Key Takeaways 6

Executive Summary 7

1. Introduction 10

1.1. Study Objectives and Methodology Overview 12

1.2. Research Design and Report Roadmap 13

2. Safe Drinking Water Act Compliance 15

2.1. Data 18

2.2. Methods 31

2.3. System-Level Correlates of Noncompliance 34

2.4. Service Area Correlates of Noncompliance 38

2.5. Government Finance Correlates of Noncompliance 41

2.6. Summary of Quantitative Findings 45

3. Data Collection Tools for Future Noncompliance Assessment 47

3.1. Research Design 47

3.2. Survey Instrument and Interview Guide Development 47

3.3. Implementation Best Practices for Surveys and Interviews 52

4. Recommendations and Conclusions 56

4.1. Evidence-Based Recommendations 56

4.2. Conclusions and Next Steps 57

References 59

Appendix A: Safe Drinking Water Act Data and Robustness Checks 64

The Research Team 77

TABLES

Table 2.1. National Primary Drinking Water Regulations and Violations by Community Water Systems, 2015–2024 17
Table 2.2. Safe Drinking Water Act Violations for Community Water Systems, 2023–2024 19
Table 2.3. Summary Statistics: All Community Water Systems 23
Table 2.4. Comparison of Community Water Systems Based on Compliance in 2018–2022 26
Table 2.5. Summary Statistics: All Systems Compared to Local Govt. Owned Systems 29
Table 2.6. System-Level Regression Results: All Community Water Systems 35
Table 2.7. System and Service Area Regression Results: All Community Water Systems 39
Table 2.8. Regression Results: Local Government Owned Community Water Systems 42
Table 3.1. Survey Content Overview 49
Table 3.2. Interview Guide Content Overview 49
Table A.1. Dataset Construction: Service Area 64
Table A.2. Dataset Construction: Violations and Enforcement 64
Table A.3. Dataset Construction: Service Area Characteristics 65
Table A.4. Data Subset – Local Government Characteristics 66
Table A.5. Regression Results Underlying Figure 2.3 67
Table A.6. Robustness Check: Functional Form 68
Table A.7. Robustness Check: Excluding Service Area Imputations 69
Table A.8. Robustness Check: Including Indicators for Service Area Imputations 70
Table A.9. Robustness Check: Only 2024 Violations 71
Table A.10. Robustness Check: Excluding Ownership 72
Table A.11. Robustness Check: Excluding Government Finance Variables 73
Table A.12. Robustness Check: Excluding Multi-Jurisdictional Systems 74
Table A.13. Robustness Check: Including Government Unit Type 75
Table A.14. Robustness Check: Excluding Employment Imputations 76

FIGURES

Figure 1.1 Research Design Illustration....13
Figure 2.1. Proportion of Noncompliant Systems by State, 2023–2024....20
Figure 2.2. Changes in Community Water System Compliance, 2018–2024....21
Figure 2.3. Probability of Noncompliance Based on Water Provision Interactions....37
Figure 2.4. Model-Based Differences in Noncompliance by State....40
Figure 3.1: Common Sampling Strategies for Surveys and Interviews....52
Figure 3.2. Example Systems Map....55

Key Takeaways

- Nearly 36 percent, or 15,920 of the 44,481 community water systems in the U.S. had at least one violation in 2023 or 2024, and nearly 28 percent had at least one violation in the five years prior (2018–2022). All Systems in 2023–2024 had a mean of three violations while noncompliant systems had a mean of nine violations, with the majority of these being non-health based.

- There is considerable variation in community water system noncompliance across states. In 2023–2024, all systems in Hawaii were compliant, while 71 percent of systems in West Virginia had at least one violation. In several other states, 50+ percent of systems were noncompliant, potentially impacting more than 48.97 million people.

- For oversight, on-site inspections for sanitary surveys are most common, as is statutorily required, with 35 percent of systems receiving one in 2022. Yet, only about one percent of systems required a formal enforcement action for observed noncompliance.

- Community water systems are more likely to have a violation and have a higher number of violations if they have: Native American ownership, surface water as a source, self-produce their water, and had a federal inspection or state technical assistance visit. Noncompliant systems with a prior violation or a formal enforcement action are more likely to have more violations.

- Over half of community water systems are local government owned; the rest are mostly privately owned. More than 80 percent self-produce (rather than purchase) water, 77 percent rely on groundwater (as opposed to surface water), and about 5 percent sell water wholesale.

- Local government owned community water systems have similar compliance rates and numbers of violations compared to all systems but differ by serving larger populations and having fewer self-producers, greater reliance on surface water, and more wholesalers.

- Local governments with community water systems spend about six percent of revenue for interest on outstanding debt, would need to levy $3 in property tax revenue to offset a $1 reduction in intergovernmental revenue, and use 83 percent of total spending to cover operations. Local governments also levy about twice the amount needed to fund their water utilities, which might be used to support unrestricted operating (i.e., general fund) activities.

- Noncompliant local governments may be using intergovernmental revenue to fund water system activities to help reduce violations, while some may be levying additional revenue through their water billing to support unrestricted operating (i.e., general fund) activities, ultimately deferring water system maintenance that helps them to remain compliant.

Executive Summary

The United States (U.S.) Environmental Protection Agency commissioned researchers in the School of Engineering at Tufts University and the Government Finance Research Center at the University of Illinois Chicago to conduct a national study on the primary correlates of noncompliance with the Safe Drinking Water Act by community water systems. This work informs recommendations aimed at increasing compliance. Community water systems are one type of public water system that, by definition, supply drinking water to the same population year-round. Nationwide, there are approximately 51,000 community water systems that serve a little under 300 million people or most of the U.S. population (U.S. Environmental Protection Agency, 2025e). Yet only 4,221, or 8 percent of these community water systems are responsible for providing drinking water to over 246 million people or 82 percent of the U.S. population (U.S. Environmental Protection Agency, 2025e) making monitoring and regulation of these systems crucial for public health protection.

This report is provided as a deliverable of the commissioned study. Throughout the study, the researchers received guidance from subject matter experts employed by the U.S. Environmental Protection Agency. The researchers compiled a comprehensive dataset for quantitative analysis of all community water systems nationwide, as well as developed a survey instrument and semi-structured interview guide to be used for future data collection. The research objectives were to:

1. Identify the system-level, service area, and government finance characteristics correlated with Safe Drinking Water Act noncompliance by community water systems through quantitative regression analysis.
2. Design a research approach, including a survey instrument and semi-structured interview guide, to identify additional contextual factors driving Safe Drinking Water Act violations by community water systems in preparation for future systems analysis to understand the root causes of noncompliance.

Potential factors associated with Safe Drinking Water Act noncompliance were categorized as system-level, service area, and government finance to allow for the quantitative research findings to provide preliminary actionable insights that can inform policy direction and operational recommendations. To assess the correlates of noncompliance through quantitative analysis, the researchers combined Safe Drinking Water Act violations from the U.S. Environmental Protection Agency's Enforcement and Compliance History Online database with relevant characteristics of each community water system's service area obtained from the U.S. Census Bureau's American Community Survey and, for community water systems owned by local governments, with the U.S. Census Bureau's Census of Governments Finance Survey data.

The findings from both prior research and this study are then used to inform the design of the survey instrument and semi-structured interview guide, which are both structured to engage with community water systems across the U.S. These data collection tools focus on potential technical, managerial, and financial factors leading to noncompliance. Once implemented for data collection, the combined findings will allow for a systems analysis to fully understand the root causes of noncompliance and inform the U.S. Environmental

Protection Agency on future policy direction to enhance compliance with the Safe Drinking Water Act.

Preliminary Conclusions

Nearly 36 percent, or 15,920 of the 44,481 community water systems in the U.S. had at least one violation in 2023 or 2024, and nearly 28 percent had a history of noncompliance with at least one violation in the five years prior (2018-2022). All Systems in 2023-2024 had a mean of three violations while noncompliant systems had a mean of nine violations, with the majority of these being non-health based.

Systems with any of these characteristics are more likely to be noncompliant with the Safe Drinking Water Act: Native American ownership, sourcing from surface water, self-producing water, and receiving a federal inspection or state technical assistance visit. Noncompliant systems with a violation in 2018–2022 and those receiving a formal enforcement action have a greater number of violations. Surface water sources for community water systems include natural and manmade bodies of water that are open to the atmosphere (e.g., rivers, streams, lakes, and reservoirs) and therefore require filtration and disinfection to treat sediment and bacteria, which likely contributes to why source water use is often associated with greater noncompliance to the Safe Drinking Water Act. Self-production (sometimes referred to as self-supply) involves community water systems sourcing and treating their own water, rather than purchasing water from a wholesaler as part of a consecutive system that may be exempt from National Primary Drinking Water Regulations, which helps to explain why self-production is often associated with greater noncompliance.

Community water systems operating in more densely populated areas and those with wealthier residents are less likely to have violations, partly due to economies of scale in drinking water provision. Among noncompliant systems, service area characteristics are generally not associated with the number of violations, except in areas with the oldest housing stock, potentially indicating broader built environment issues in these areas. These service area findings are more pronounced for local government owned systems. However, noncompliant local government owned systems that are wholesalers and those receiving technical assistance visits have fewer violations compared to all noncompliant systems. In terms of fiscal capacity, noncompliant local governments may be using intergovernmental revenue to fund water system activities to help reduce violations, while some may be levying additional revenue through their water billing to support unrestricted operating (i.e., general fund) activities, ultimately deferring water system maintenance that helps meet regulations.

Together, the preliminary regression results shed light on the primary factors correlated with noncompliance. Although these generalizable trends are useful, there is a need to engage directly with community water systems to understand contextual details, as well as the technical and managerial challenges, to fully identify the root causes of noncompliance. To do so, this study developed additional data collection tools—a survey instrument and semi-structured interview guide—along with guidance to the U.S. Environmental Protection Agency about how to implement these tools. Pairing the preliminary quantitative findings

with qualitative research is needed to make causal claims regarding Safe Drinking Water Act noncompliance.

Recommendations and Next Steps

Overall, the quantitative research findings lend themselves to the following preliminary recommendations:

- Increase the compliance monitoring capacity of community water systems that are persistent violators (i.e., utilities that have a violation every year from 2018 to 2024) or at greater risk of noncompliance based on results from the quantitative models (e.g., community water systems that rely on surface water or produce their own water).
- Adopt clear criteria for targeting inspections and combine federal inspections of community water systems with technical assistance to achieve greater Safe Drinking Water Act compliance.
- Implement regularly scheduled inspections, in addition to sanitary surveys, for persistent violators that include educational components based on issues that arise.
- Enact outreach about best practices regarding compliance-supporting financial management and utility maintenance to local governments that own and operate drinking water systems.

The U.S. Environmental Protection Agency may find the recommendations in this report helpful when developing or refining policy or practices. It is important to note that these recommendations would require adequate staffing and training, as well as adequate administrative support, technology, contract support, and access to other resources. Although the above findings provide a basis for understanding correlates of noncompliance with the Safe Drinking Water Act, qualitative research should be conducted to capture perspectives of utility personnel. The additional data collection tools developed for this study should be deployed to community water systems throughout the U.S. to better understand the contextual and managerial factors that serve as root causes of violations. Once implemented for data collection, the combined findings will allow for a systems analysis to inform the U.S. Environmental Protection Agency on what targeted technical, managerial, or financial assistance could improve Safe Drinking Water Act compliance by community water systems.

1. Introduction

In the U.S., the Centers for Disease Control (2024) estimated that nine out of every 10 people, more than 301 million total, received their drinking water from public water systems in 2023. The heavy reliance on public water systems for safe, clean drinking water makes monitoring and regulation crucial for public health protection (Centers for Disease Control, 2024). The Safe Drinking Water Act was enacted in 1974 (and amended significantly in 1986, 1996, and 2018, among smaller amendments in other years) to regulate drinking water quality. Through this legislation, the U.S. Environmental Protection Agency regulates public water systems, setting legal limits and enforcing protective standards for over 90 contaminants such as lead, arsenic, and microbial pathogens through the National Primary Drinking Water Regulations. The National Primary Drinking Water Regulations minimize water quality issues, protecting public health and reducing negative financial impacts. Collectively, the Centers for Disease Control estimates waterborne pathogen exposure from all water sources results in $3.33 billion annually in direct health care costs across the country from 7.15 million illnesses, 118,000 hospitalizations, and 6,630 deaths (Centers for Disease Control, 2024). Approximately 3,300 of these deaths are associated with drinking water exposure (Centers for Disease Control, 2024). In addition, contaminated drinking water has been linked to gastrointestinal illnesses, nervous system effects, and chronic diseases (Ashbolt, 2004; Fawell & Nieuwenhuijsen, 2003; Rahman et al., 2009).

Water utilities throughout the U.S. are facing operational challenges such as an aging workforce, lack of funding, and high attrition rates, among other concerns. The Government Accountability Office (2018) estimates there will be more than 9,200 job openings in the water sector annually through 2026. Such workforce shortages are paired with knowledge and training gaps, which were only accelerated by the COVID-19 pandemic due to retirements, leading to a loss of institutional knowledge (Spearing et al., 2020). In turn, utilities can sometimes fall short of meeting drinking water standards. Recent data shows that 28 percent of active public water systems in the U.S. violated at least one drinking water standard in 2023 (U.S. Environmental Protection Agency, 2025b).

Community water systems are one type of public water system that, by definition, supply drinking water to the same population year-round.[1] Of the 156,000 public water systems in the U.S., there are approximately 51,000 community water systems that serve a little under 300 million people or most of the U.S. population (U.S. Environmental Protection Agency, 2025e). Yet, 43.2 percent of community water systems nationwide were noncompliant in 2022 (U.S. Environmental Protection Agency, 2025e), indicating that any violation in these systems can put a large number of people at risk (U.S. Environmental Protection Agency, 2025b). Further, only 4,221 or 8 percent of these community water systems are responsible for providing drinking water to more than 246 million people or 82 percent of the U.S. population (U.S. Environmental Protection Agency, 2025e). As such, it is imperative to understand the primary factors correlated with water utility noncompliance.

[1] The other two types of public water systems are: 1) non-transient non-community water systems that supply water to at least 25 of the same people at least six months per year, such as schools or hospitals with their own water systems, and 2) transient non-community water systems that provide water in places like gas stations or campgrounds where people do not remain for long periods of time. These other two types of public water systems are outside the scope of this study.

The U.S. Environmental Protection Agency divides drinking water violations into two categories—health-based and non-health-based. Health-based violations include reported issues with maximum contaminant levels, maximum residual disinfectant levels, and treatment techniques. Non-health-based violations concern inadequate water quality monitoring and reporting, including public notification and consumer confidence reporting violations. It is estimated that the number of monitoring and reporting violations is 10 times higher than the number of contaminant violations (Wallsten & Kosec, 2008). However, monitoring and reporting violations can shield health-based violations at facilities since these violations may not be found and properly recorded (Young et al., 2023). This historic trend is shown in a U.S. Environmental Protection Agency audit of drinking water data from 2002 to 2004, which found that only 62 percent of health-based violations were reported (U.S. Environmental Protection Agency, 2008). Due to this potential underreporting and the interrelated nature of these two main categories of violations, this study combines both health-based and non-health-based violations together into the violation variables used for analysis.

To understand the primary correlates of noncompliance with the Safe Drinking Water Act, prior researchers have explored links between various measures of compliance (i.e., overall, health-based, non-health-based) and a variety of factors related to a water system and characteristics of its service area. Research has largely highlighted that smaller, more rural utilities tend to have more Safe Drinking Water Act violations (Allaire et al., 2018; Konisky & Teodoro, 2016; Oxenford & Barrett, 2016; Rubin, 2013; Scanlon et al., 2022; Switzer & Teodoro, 2017). Further, characteristics of the area served by systems have been linked to compliance outcomes, with utilities serving lower-income communities tending to have more violations (Fu et al., 2020; Sowby, 2023), likely due to less financial capacity to leverage necessary resources from their tax bases.

The mutually exclusive ownership categories of water systems (e.g., federal government, local government, public/private, Native American, private, or state government) has been widely studied, with mixed results regarding noncompliance. One study found community water systems owned by local governments to have the lowest amount of monitoring and reporting violations per year out of all ownership types, while also experiencing more contaminant violations than privately owned systems (Wallsten & Kosec, 2008). Other evidence suggests public water systems have a higher probability of maximum contaminant level violations (Rahman et al., 2010), as well as higher numbers of health violations and monitoring and reporting violations (Konisky & Teodoro, 2016), with state systems specifically linked to higher numbers of maximum contaminant level violations (Kirchhoff et al., 2019).

Water source is also related to noncompliance. Purchased water has been linked to lower levels of noncompliance in many studies (Allaire et al., 2018; Konisky & Teodoro, 2016; Switzer & Teodoro, 2017; Wallsten & Kosec, 2008), potentially because consecutive systems purchasing water from wholesaler systems are often subject to fewer National Primary Drinking Water Regulations. Studies exploring groundwater versus surface water have had mixed results (Kirchoff et al., 2019; Rahman et al., 2010; Rubin, 2013), but most

work has identified groundwater sourced utilities to have worse compliance outcomes (Switzer & Teodoro, 2017) despite surface water typically requiring more treatment.

Although previous research has explored how these and other factors are related to compliance, there is a lack of focus on the role of oversight activities and financial factors. Anecdotally, managerial and financial challenges hinder community water systems from Safe Drinking Water Act compliance (Allaire et al. 2017; Scanlon et al., 2022). For example, representatives of municipal utilities in Illinois report a lack of overall capacity for appropriate financial management and capital planning, leading to a wide variety of practices that municipalities use for managing their water funds and planning for infrastructure improvements to maintain water quality (Carroll et al., 2023; Carroll et al., 2024). However, existing literature has yet to explore these connections holistically.

1.1. Study Objectives and Methodology Overview

This study attempts to fill the gap from prior research by including oversight information (i.e., visits and inspections) and local government financial data in the analysis to identify the primary correlates of community water system noncompliance with the Safe Drinking Water Act. Ultimately aiming for a mixed-methods approach in line with extant research, this study pursues the following sequential research objectives:

1. Identify the system-level, service area, and government finance characteristics correlated with Safe Drinking Water Act noncompliance by community water systems through quantitative regression analysis.
2. Design a research approach, including a survey instrument and semi-structured interview guide, to identify additional contextual factors driving Safe Drinking Water Act violations by community water systems in preparation for future systems analysis to understand the root causes of noncompliance.

To approach these objectives, the researchers compiled a comprehensive dataset for quantitative analysis of the population of community water systems nationwide. As explained in detail in section two, the researchers combined Safe Drinking Water Act violations from the U.S. Environmental Protection Agency's Enforcement and Compliance History Online database with relevant characteristics of each community water system's service area obtained from the U.S. Census Bureau's American Community Survey and, for community water systems owned by local governments, with the U.S. Census Bureau's Census of Governments Finance Survey data. The potential factors correlated with Safe Drinking Water Act noncompliance were categorized as system-level, service area, and government finance to allow for the preliminary research findings to provide actionable insights that can inform policy direction and operational recommendations to increase Safe Drinking Water Act compliance by community water systems.

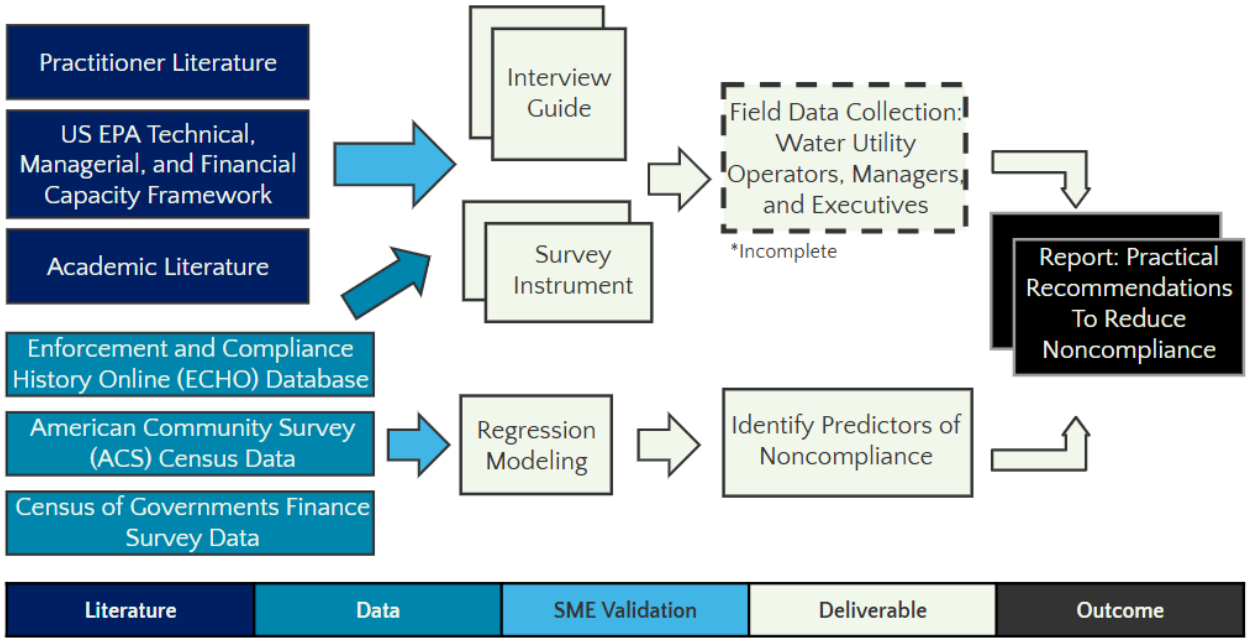
Although statistical models can provide insight into large-scale trends, to fully understand the root causes of noncompliance, there is a need to understand nuanced operational and managerial challenges of community water systems. Such contextual factors can be captured through direct engagement with utility staff, through tools such as surveys and interviews (e.g., Spearing et al., 2020). In turn, the researchers developed a survey

instrument and semi-structured interview guide to be implemented for future primary data collection.

1.2. Research Design and Report Roadmap

Prior research is used to inform the research design for the quantitative analysis in this study, as illustrated in Figure 1.1. The findings from both prior research and this study are then used to inform the design of the survey instrument and semi-structured interview guide, which are both structured to engage with community water systems across the U.S. These data collection tools focus on potential technical, managerial, and financial factors leading to noncompliance. Once implemented for data collection, the combined findings will allow for a systems analysis to fully understand the root causes of noncompliance and inform the U.S. Environmental Protection Agency on future policy direction to enhance compliance with the Safe Drinking Water Act.

Figure 1.1. Research Design Illustration



Note: This figure illustrates the research design developed by the authors and used to guide this study.

Section two of this report provides an overview of the Safe Drinking Water Act and how the U.S. Environmental Protection Agency defines compliance with the Act. The section proceeds with a description of the data collection and dataset compilation process, methods used for quantitative analysis, and regression results of the system-level and service area correlates of noncompliance from analysis of the population of community water systems throughout the country. Section two concludes by providing the regression results of government finance variables correlated with noncompliance from analysis of a subset of local government-owned community water systems.

Section three of this report describes the research design used for development of a survey instrument and semi-structured interview guide to be implemented for future data collection. Along with descriptions of these data collection tools, this section also provides a discussion of best practices for implementing surveys and conducting semi-structured interviews for qualitative and quantitative data collection. Section three concludes with a description of how data collected with these tools might be analyzed. Doing so will help to inform the U.S. Environmental Protection Agency and policymakers on the root causes of noncompliance and the most effective approaches for increasing Safe Drinking Water Act compliance by community water systems.

Section four offers a discussion of overall conclusions and next steps to advance this study intended to help reduce Safe Drinking Water Act noncompliance. Additional information pertaining to data collection, dataset compilation, and robustness checks for the quantitative analyses are provided in Appendix A.

2. Safe Drinking Water Act Compliance

The Safe Drinking Water Act, first enacted in 1974 and last significantly amended in 2018, establishes the primary federal regulatory framework for drinking water quality in the U.S. (U.S. Environmental Protection Agency, 2025a). Under the Safe Drinking Water Act, the U.S. Environmental Protection Agency promulgated the National Primary Drinking Water Regulations, which include enforceable maximum contaminant levels and treatment technique requirements or rules for a broad range of microbial, chemical, and radiological contaminants.

Most direct oversight of community water systems' compliance with National Primary Drinking Water Regulations is conducted by state Public Water System Supervision programs (U.S. Environmental Protection Agency, 2024). All territories, most states, and the Navajo Nation are granted primary enforcement authority, or "primacy," meaning the U.S. Environmental Protection Agency allows them to implement the Public Water System Supervision program since their own regulations are at least as stringent as federal standards (U.S. Environmental Protection Agency, n.d.). In addition, there are enforcement, recordkeeping, sanitary survey, reporting, and other requirements imposed on primacy agencies, among other oversight activities.

While federally recognized Tribes may apply to administer eligible U.S. Environmental Protection Agency programs in Indian country through the Treatment as a State process, most Tribes have not done so. When a Tribe lacks authorized or delegated Public Water System Supervision program implementation approval, or where the U.S. Environmental Protection Agency lacks authority to authorize a Tribe to administer a Public Water System Supervision program, the U.S. Environmental Protection Agency carries out oversight responsibilities in Indian country through direct implementation. As the primacy agency, the U.S. Environmental Protection Agency engages with Tribes, using an array of approaches to identify, address, and resolve noncompliance.

Compliance with the National Primary Drinking Water Regulations of the Safe Drinking Water Act entails several activities by primacy agencies including, but not limited to, routine testing of public water systems for contaminants, review of proposed system improvements, on-site inspections and sanitary surveys, training and certification of water system operators, and enforcement actions against systems that fail to meet established standards. A sanitary survey is a proactive public health measure that involves reviewing a public water system to assess their capability to supply safe drinking water.[2] Primacy agencies are responsible for completing sanitary surveys, which provide an opportunity for the primacy agency to visit the water system and educate the operator about proper monitoring and sampling procedures and to provide technical assistance as needed. Community water systems generally receive a sanitary survey visit every three years, or every five years with outstanding performance in prior sanitary surveys.

Primacy agencies collect detailed operational and monitoring data from the community water systems under their jurisdiction and submit a subset of these data to the U.S.

[2] For additional information, please see: https://www.epa.gov/dwreginfo/sanitary-surveys.

Environmental Protection Agency's Safe Drinking Water Information System on a quarterly basis. The data reported to the U.S. Environmental Protection Agency include system-specific details, such as system name, classification (community, transient, or non-transient), population served, and source water type, along with recorded violations of National Primary Drinking Water Regulations and any enforcement actions undertaken to restore compliance with the Safe Drinking Water Act (U.S. Environmental Protection Agency, n.d.).

Enforcement involves informal and formal actions aimed at restoring compliance as quickly as possible. These enforcement actions often begin with informal actions, including reminder letters, warning letters, notices of violation, field visits, and telephone communications. Formal actions include citations, administrative orders with or without penalties, and civil referrals to state attorneys general or the U.S. Department of Justice and may obligate the water system to achieve compliance on a specified schedule. The U.S. Environmental Protection Agency may immediately take formal enforcement action when it serves as the primacy agency (including for new rules where the State, Tribe, or territory does not yet have primacy), or where the appropriate primacy agency refers the water system to the U.S. Environmental Protection Agency. In cases where public health is in imminent and substantial endangerment, the U.S. Environmental Protection Agency and the primacy agency, if state authorities exist, may issue emergency orders requiring the system to take immediate corrective action to protect public health and restore compliance (U.S. Environmental Protection Agency, n.d.).

Despite this regulatory framework, compliance challenges remain. Table 2.1 presents the proportion of violations by community water systems reported to U.S. Environmental Protection Agency's Safe Drinking Water Information System by category and rule groupings from 2015 to 2024. Among these violations, over 67 percent are monitoring and reporting, with the largest proportions of violations for volatile organic chemicals (16.77 percent), synthetic organic chemicals (17.94 percent), and inorganic chemicals (7.38 percent). It should be noted, however, that the majority of regulated contaminants are not for radionuclides or microbials, which explains the low numbers of violations in the table.

Prior research has revealed important trends in Safe Drinking Water Act noncompliance, which is used to inform the analysis in this study. To illustrate, Allaire and colleagues (2018) analyzed a panel of 17,900 community water systems from 1982 to 2015 and identified violation hot spots in the Southwest where repeat violations were also common. Systems with private ownership and those relying on purchased water sources were more likely to remain in compliance (Allaire et al., 2018). Similarly, in another national study examining violations from 1950 to 2017, Fu and coauthors (2020) found that publicly owned systems had significantly higher rates of maximum contaminant level, treatment technique, and health-related violations, but lower rates of monitoring and reporting violations compared with private systems. The study also found that including municipal fixed effects accounts for a substantial portion of the variation in violation behaviors, suggesting that local regulations and economic conditions were driving some of the noncompliance (Fu et al., 2020). More recently, the U.S. Environmental Protection Agency's Interagency Water Workforce Working Group Report to Congress (2024) examined all types of violations

Table 2.1. National Primary Drinking Water Regulations and Violations by Community Water Systems, 2015-2024

Regulation Categories	MCL	TT	MRDL	MR	MON	RPT	Other	Totals
Total Coliform	5,450 0.87%	1,518 0.24%		10,358 1.66%	31,950 5.12%	5,583 0.89%	1,155 0.19%	56,014 8.97%
Surface Water Treatment		5,981 0.96%		18,508 2.97%			171 0.03%	24,660 3.95%
Groundwater		8,024 1.29%		16,185 2.59%			638 0.10%	24,847 3.98%
Stage 1 Disinfectants and Disinfection Byproducts	332 0.05%	2,617 0.42%	145 0.02%	29,295 4.69%			- 	32,389 5.19%
Stage 2 Disinfectants and Disinfection Byproducts	29,062 4.66%			34,723 5.56%			11 0.00%	63,796 10.22%
Volatile Organic Chemicals	202 0.03%			104,655 16.77%			1 0.00%	104,858 16.80%
Synthetic Organic Chemicals	82 0.01%	7 0.00%		111,949 17.94%			1 0.00%	112,039 17.95%
Inorganic Chemicals	15,665 2.51%			46,056 7.38%			9 0.00%	61,730 9.89%
Radionuclides	7,952 1.27%			14,603 2.34%			- 	22,555 3.61%
Lead and copper		1,760 0.28%		33,109 5.30%			14 0.00%	34,883 5.59%
Public Notice							50,345 8.07%	50,345 8.07%
Consumer Confidence							36,010 5.77%	36,010 5.77%
Miscellaneous Other							52 0.01%	52 0.01%
Totals	58,745 9.41%	19,907 3.19%	145 0.02%	419,441 67.20%	31,950 5.12%	5,583 0.89%	88,407 14.16%	624,178 100%

Note: This table presents the number (top) and proportion (bottom) of National Primary Drinking Water Regulation violations committed by community water systems in the U.S. by regulation category and rule groupings. The total number of violations from 2015 to 2024 was 624,178. Maximum Contaminant Level (MCL), Treatment Technique (TT), and Maximum Residual Disinfectant Level (MRDL) are **health-based violations (in red)**. Monitoring and Reporting (MR), Monitoring (MON), and Reporting (RPT) are **non-health-based violations (in blue)**. For additional information on rules, refer to the U.S. Environmental Protection Agency's national primary drinking water regulations: https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets, obtained from https://echo.epa.gov/tools/data-downloads#downloads.

from 2017 to 2021 and found higher noncompliance rates for small and rural systems, Tribal-owned systems, and community water systems relying on surface water (U.S. Environmental Protection Agency, 2024).

Prior research from individual states is largely consistent with the U.S. Environmental Protection Agency's (2024) findings. In Virginia, very small community water systems serving less than 500 people exhibited significantly more monitoring and reporting violations than larger systems, while medium community water systems serving 3,301-10,000 people had significantly more maximum contaminant level violations (Marcillo & Krometis, 2019). Additionally, community water systems in isolated rural areas (i.e., a score of 10 on the U.S. Department of Agriculture's rural-urban commuting area scale) showed significantly higher monitoring and reporting noncompliance compared with those in towns or urban centers (Marcillo & Krometis, 2019). In Alaska, results were similar, indicating that small community water systems (<100 connections) were less likely to be compliant with the Safe Drinking Water Act than larger systems (Tariq et al., 2025). Results from Arizona found larger (serving more than 3,300 people) public water systems to be more likely to violate maximum contaminant level standards than public systems serving school districts or office buildings and privately owned systems (Rahman et al., 2010). Notably, the Arizona results diverged from other research by finding larger public water systems to be more likely to violate maximum contaminant level standards than smaller public systems, highlighting heterogeneity across states.

Extant research also reveals interesting dynamics related to oversight and enforcement. In Connecticut, Kirchhoff and colleagues (2019) found that systems responded differently to formal and informal enforcement actions, with informal measures like a letter and technical assistance more effectively reducing the duration of violations than formal actions. It may be the case, however, that informal enforcement is often a first step for minor violations that are easy, cheap, and less time consuming to fix, whereas formal enforcement is frequently used where difficult, complex, and expensive remediation is required. Examining the relationship between public disclosures (an enforcement tool) and compliance, Grooms (2016) found in California that although public disclosure discouraged systems from committing violations, the Public Notification Rule was ineffective at prompting systems to return to compliance once a violation occurred.

Building on these research findings, this analysis examines Safe Drinking Water Act violations in 2023 and 2024 for the population of community water systems across all 50 states and Washington, DC, modeling these violations as a function of system-level and service area characteristics. For a subset of the data, local government-owned systems, government finance variables were also incorporated and analyzed.

2.1. Data

2.1.1. Safe Drinking Water Act Noncompliance

A binary variable (i.e., coded 1 or 0) was created to indicate whether or not a community water system was noncompliant (i.e., had any type of violation) in 2023 or 2024 (e.g., *Any Violation*). In addition, a variable was created measuring the total number of violations for

each community water system in 2023–2024. This information is available for 44,481 unique community water systems (Table A.1). These violations were calculated in aggregate, as well as by violation type and category as reported in Table 2.2.[3]

Table 2.2. Safe Drinking Water Act Violations for Community Water Systems, 2023–2024

Violation Categories	Percentage	Mean	Minimum	Maximum
Any Violation	35.79%	3.11	0	1,378
Any Health-Based Violation	7.44%	0.56	0	158
Maximum Contaminant Level	4.44%	0.40	0	158
Maximum Residual Disinfectant Level	0.02%	0.0008	0	12
Treatment Technique	3.47%	0.16	0	111
Any Monitoring and Reporting Violation	33.07%	2.54	0	1,378
Monitoring and Reporting	19.71%	1.69	0	1,337
Monitoring	8.74%	0.30	0	67
Reporting	1.58%	0.04	0	47
Other Violation	16.91%	0.52	0	229

Note: Total number of observations is 44,481 community water systems, which represents all systems in the U.S. for which compliance data are available. The percentage column calculates the proportion of community water systems with at least one violation in each category in 2023–2024 of the total 44,481 systems. The mean, minimum, and maximum values are calculated from the numbers of violations in each category from among the total 44,481 systems. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets, obtained from https://echo.epa.gov/tools/data-downloads#downloads.

The main takeaway from Table 2.2 is that nearly 36 percent of the 44,481 community water systems in the U.S. had at least one violation in 2023 or 2024, amounting to 15,920 community water systems in noncompliance with the Safe Drinking Water Act. All systems had a mean of three violations while noncompliant systems had a mean of nine violations. during these two years. The majority of these violations were non-health based.

Although 64.21 percent or 28,561 community water systems complied with the Safe Drinking Water Act during these two years, the number of violations among those that were noncompliant reached 1,378 for a single community water system. Consistent with extant research examining violations in earlier years, the majority of violations were non-health-based, with 33.07 percent of community water systems having at least one monitoring and reporting violation compared to 7.44 percent of community water systems having at least one health-based violation. The mean (0.56) and maximum (158) numbers of health-based violations that occurred during 2023 and 2024 were also much lower compared to the mean (2.54) and maximum (1,378) numbers of monitoring and reporting violations.

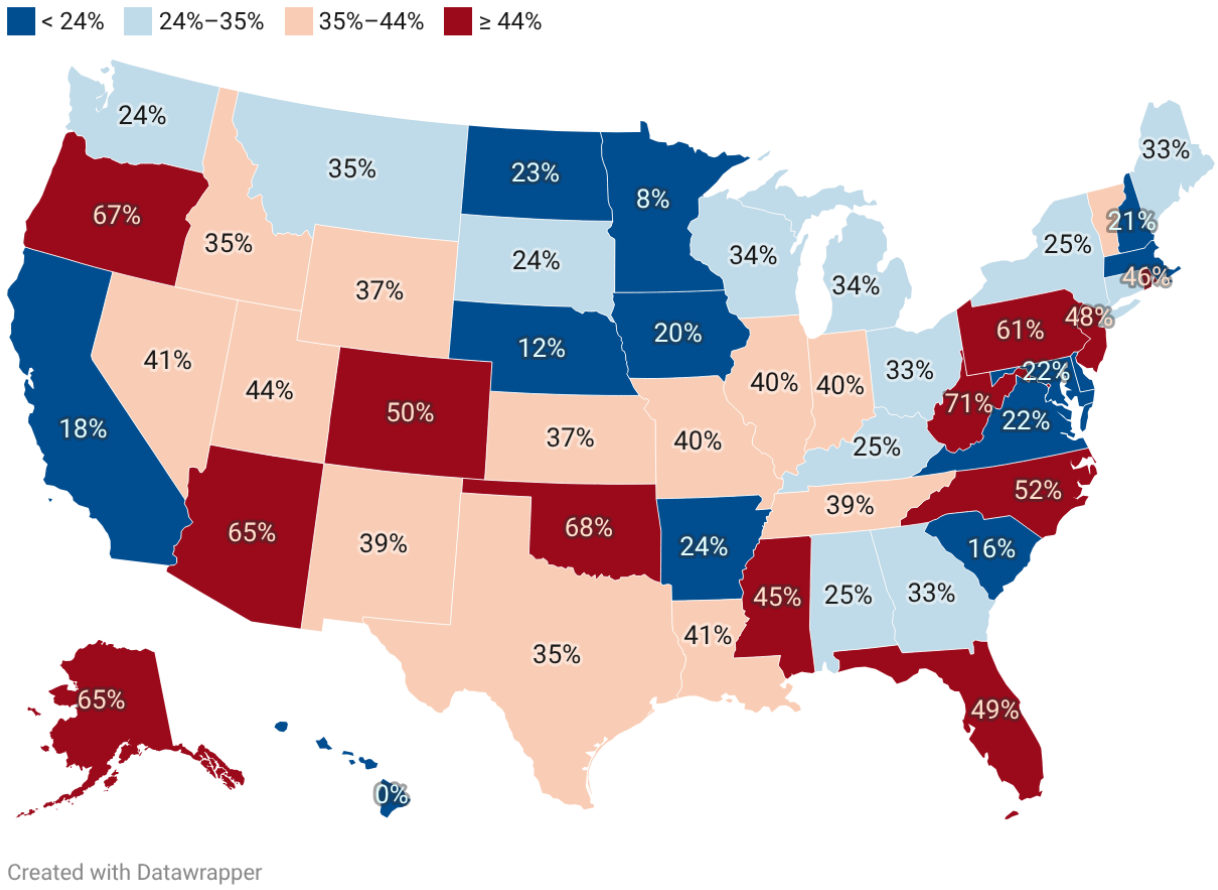
Among health-based violations, noncompliance with maximum contaminant level was most common, with 4.44 percent of community water systems having at least one of these

[3] The dataset for quantitative analysis was constructed by combining the U.S. Environmental Protection Agency's 2025 Community Water System Service Area Boundaries (U.S. Environmental Protection Agency, 2025c) with information on community water system characteristics from the U.S. Environmental Protection Agency's Safe Drinking Water Information System database for 2024 Q4 (U.S. Environmental Protection Agency, n.d.) and the U.S. Environmental Protection Agency's Safe Drinking Water Information System data on violations and enforcement (Table A.2).

violations, followed by treatment technique, with 3.47 percent of community water systems in violation. Maximum residual disinfectant level violations were negligible. Among monitoring and reporting violations, noncompliance with some aspect of monitoring was far more common than those associated with reporting alone.

Since the data presented in Table 2.2 are skewed toward zero (because only 35.79 percent of community water systems have a violation of any kind in 2023 or 2024), the median values for all violation categories are zero. At the 75th percentile, the number of violations is two, both of which are monitoring and reporting. At the 90th percentile, the number of violations is six and reaches 13 at the 95th percentile, illustrating the data skew. Additional summary statistics for community water systems consistently in compliance versus those with violations are presented in Table 2.4.

Figure 2.1. Proportion of Noncompliant Systems by State, 2023–2024



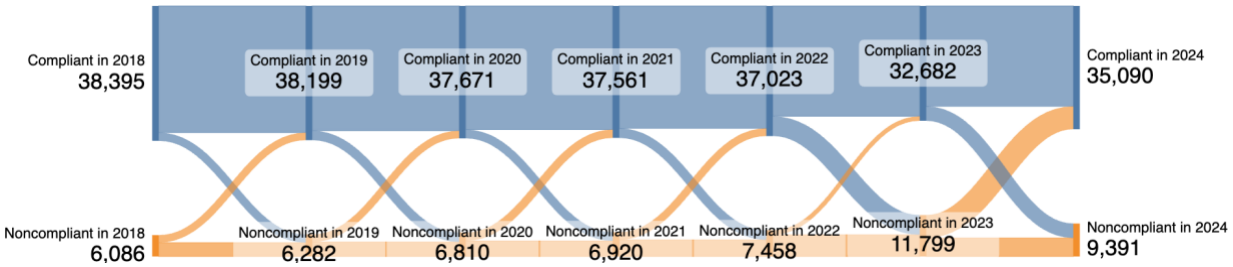
Note: This figure illustrates the proportions of community water systems that had at least one violation (including both health-based and non-health-based) in 2023 or 2024 within each state. The map is color coded to reflect four quantiles of equal state counts. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets, obtained from https://echo.epa.gov/tools/data-downloads#downloads.

Figure 2.1 illustrates the geographical variation in Safe Drinking Water Act noncompliance by mapping the proportions of community water systems that had at least one violation (including both health-based and non-health-based) in 2023 or 2024 within each state. The map is color coded to reflect four quantiles of equal counts (i.e., the same number of states in each quantile), where the darkest blue color indicates less than 24 percent of systems within a state are noncompliant.

The main takeaway from Figure 2.1 is that there is considerable variation in community water system noncompliance across states. All of the community water systems in Hawaii were compliant with the Safe Drinking Water Act in 2023 and 2024, while 71 percent of systems in West Virginia had at least one violation during these two years. In several other states, 50 percent or more of their community water systems were noncompliant in 2023-2024, potentially impacting more than 48.97 million people living there.

As shown in Figure 2.1, Hawaii (0 percent), Minnesota (8 percent), Delaware (11 percent), Nebraska (12 percent), South Carolina (16 percent), California (18 percent), Massachusetts (20 percent), and Iowa (20 percent) have the lowest proportions of systems with observed noncompliance in 2023 or 2024. In contrast, West Virginia (71 percent), Oklahoma (68 percent), Oregon (67 percent), Arizona (65 percent), Alaska (65 percent), Pennsylvania (61 percent), North Carolina (52 percent), and Colorado (50 percent) have the highest proportions of systems with violations, as 50 percent or more of community water systems in each state had observed noncompliance in 2023 or 2024. In addition, DC's two systems both had violations during this time. These states (and DC), with 50 percent or more of their community water systems in noncompliance are home to more than 48.97 million people, according to U.S. Census Bureau population estimates.

Figure 2.2. Changes in Community Water System Compliance, 2018-2024



Note: This figure illustrates changes in all community water systems nationwide between compliance and noncompliance with the Safe Drinking Water Act year-over-year during the 2018-2024 time period. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets, obtained from https://echo.epa.gov/tools/data-downloads#downloads.

Figure 2.2 illustrates changes in all community water systems nationwide between compliance and noncompliance with the Safe Drinking Water Act year-over-year during the 2018-2024 time period.

The main takeaway from Figure 2.2 is that most community water systems were compliant with the Safe Drinking Water Act each year 2018–2024, but a small proportion of systems had at least one violation each year during this time. About one-third of community water systems moved between compliance and noncompliance during these years.

As shown in Figure 2.2, most systems (64 percent) remained consistently compliant throughout 2018–2024, meaning these community water systems never had any violation. A small proportion (3 percent) remained persistently noncompliant during this time, meaning these community water systems had at least one observed violation every year, either due to lagging unresolved compliance from an earlier time or new violations that occurred. A substantial proportion (33 percent) of community water systems oscillated between compliance and noncompliance by having at least one observed violation during the time period that was ultimately resolved and moved the system back into compliance. Generally, the number of community water systems in noncompliance increased year-over-year, with a 3.22 percent increase 2018–2019, 8.4 percent increase 2019–2020, 1.62 percent increase 2020–2021, 7.77 percent increase 2021–2022, and 58.21 percent increase 2022–2023. However, the percentage change in noncompliant systems decreased 2023–2024 by 20.41 percent.

2.1.2. System and Service Area Characteristics

Table 2.3 lists summary statistics of all variables used for analysis of all community water systems. As described above, *Any Violation* is a binary variable equal to one if a community water system is noncompliant, i.e., had any violation in 2023 or 2024, and zero otherwise. The *Number of Violations* is a count of the total number of violations for each community water system in 2023 and 2024. If a system had no violations in either year, the *Number of Violations* equals zero. *Any Violation 2018–2022* is also a binary variable that is equal to one if a community water system had a violation at any point from 2018 to 2022, and zero otherwise. The mean for *Any Violation* is 0.3579, indicating that 35.79 percent of community water systems had a violation in 2023 or 2024. The mean *Number of Violations* in 2023 and 2024 is three, and 27.83 percent of systems had at least one *Any Violation 2018–2022*.

The main takeaway for these noncompliance variables in Table 2.3 is that nearly 36 percent of community water systems had a violation in 2023 or 2024, and nearly 28 percent had a history of noncompliance with at least one violation in the five years prior. The mean number of violations in 2023 and 2024 is three.

In the U.S. Environmental Protection Agency's Safe Drinking Water Information System database, community water systems are coded to indicate the ownership category of the water system: federal government, local government, public/private, Native American, private, or state government. In this analysis, each ownership category is a binary variable coded one to reflect each ownership type and zero otherwise. The means of these variables in Table 2.3 indicate that 0.65 percent of community water systems are owned by the *Federal Government*, 52.11 percent by *Local Governments*, and 0.80 percent by *State Governments*. In addition, 42.87 percent are *Privately* owned, while 2.47 percent are part of a joint *Public/Private* ownership structure. Finally, 1.09 percent of community water systems are *Native American* owned.

Table 2.3. Summary Statistics: All Community Water Systems

Variables	Mean	Standard Deviation	Minimum	Maximum
Dependent Variables				
Any Violation	0.3579	0.4794	0	1
Number of Violations	3	16	0	1,378
System-Level Characteristics				
Ownership				
Federal Gov.	0.0065	0.0805	0	1
Local Gov.	0.5211	0.4996	0	1
Public/Private	0.0247	0.1552	0	1
Native American	0.0109	0.1040	0	1
State Gov.	0.0080	0.0894	0	1
Private	0.4287	0.4949	0	1
Service Connections	2,357	15,972	0	958,265
Self-Produced	0.8058	0.3955	0	1
Surface Water	0.2271	0.4190	0	1
Wholesaler	0.0503	0.2186	0	1
Regularly Scheduled Visits	0.0186	0.1351	0	1
Technical Assistance Visits	0.0101	0.1000	0	1
Sanitary Survey Visits	0.3562	0.4789	0	1
State Inspections	0.0165	0.1273	0	1
Federal Inspections	0.0027	0.0523	0	1
Any Violation 2018-2022	0.2783	0.4482	0	1
Proportion Unreported Data	0.3563	0.0408	0.02	0.5778
Formal Enforcement Action	0.0140	0.1175	0	1
Service Area Characteristics				
Population Density (per km^2)	217	506	0	18,143
Median Household Income ($)	77,062	29,814	2,499	250,001
Median Age	43	7	12	78
Proportion of Housing Stock				
Built Before 1960	0.2260	0.1662	0	1
Built 1960-1969	0.0850	0.0560	0	1
Built 1970-1979	0.1554	0.0722	0	0.7837
Built 1980-1989	0.1377	0.0731	0	0.9141

Note: The number of observations is 44,481 community water systems. Missing data for service area variables (Median Household Income, Median Age, and Proportion of Housing Stock) were imputed using the median of the observed values. The imputation impacts 2.37% of observations for the income variable, 0.95% for age, and 1.34% for the housing stock proportions. Robustness checks reported in the appendix assess the sensitivity of the regression results to the imputations by (i) excluding imputed observations and (ii) including indicator variables for imputed values. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Service Connections is a measure of community water system size, with systems having a mean of 2,357 connections. *Self-Produced* is a binary variable that captures how water is procured. Self-production (sometimes referred to as self-supply) involves community water systems sourcing and treating their own water, rather than purchasing water from a wholesaler as part of a consecutive system. This distinction is important because the National Primary Drinking Water Regulations apply to community water systems that collect and treat their own water but may not apply to systems that purchase water. If a community water system obtains all of its water from (but is not owned or operated by) a

public water system, only maintains storage and distribution facilities (not collection and treatment facilities), and does not sell water to any person, the system may be exempt from National Primary Drinking Water Regulations and therefore not required to comply (U.S. Code § 300g–Coverage). Of the 44,481 community water systems included in the analysis, 80.58 percent self-produce their water, while the remaining 19.24 percent purchase water.

Surface Water is a binary variable that indicates the water source, where 22.71 percent of community water systems rely on surface water (while the remaining 77.29 percent rely on groundwater).[4] Surface water sources for community water systems include natural and manmade bodies of water that are open to the atmosphere (e.g., rivers, streams, lakes, and reservoirs) and therefore require filtration and disinfection to treat sediment and bacteria to ensure safety and compliance. Groundwater is typically accessed through wells drilled into underground aquifers where water is generally naturally filtered through sand, gravel, or bedrock to provide a higher-quality natural water supply than surface water. *Wholesaler* is a binary variable indicating that a community water system sells water, with a mean of 0.0503 indicating that 5.03 percent of systems wholesale water.

The main takeaway from these system-level variables in Table 2.3 is that over half of community water systems are local government owned with the remaining majority privately owned. More than 80 percent of community water systems self-produce (rather than purchase) their water, 77 percent rely on groundwater (as opposed to surface water), and only about five percent sell water wholesale.

Ensuring compliance with the Safe Drinking Water Act requires state site visits that are conducted for different purposes, which may be accomplished simultaneously or through multiple visits. The U.S. Environmental Protection Agency's Safe Drinking Water Information System database includes information about site visits, which are coded for the following reasons: sanitary survey, regularly scheduled, technical assistance, state inspection. In this analysis, each of these site visit reasons are coded as mutually exclusive binary variables to indicate whether a community water system received a visit for each reason in 2022. While it is entirely possible that one site visit could fulfill more than one of these needs, these variables provide insight into the extent to which primacy agencies provided oversight during this time. As Table 2.3 shows, around 1.86 percent of community water systems received a *Regularly Scheduled Visit* in 2022, while 1.01 percent received a *Technical Assistance Visit*. About 35.62 percent received a *Sanitary Survey Visit*, and 1.65 percent received a *State Inspection*.

Federal Inspections is a binary variable that captures evaluations by the U.S. Environmental Protection Agency conducted in 2022 to assess compliance with regulatory standards. Inspections are conducted for several reasons, including to: 1) evaluate a system's compliance with Safe Drinking Water Act, National Primary Drinking Water Regulations, and primacy agency requirements, 2) assess for emergency orders whether a system poses a public health risk by providing unsafe drinking water, 3) follow-up on previous inspections,

[4] Six community water systems did not have information on provision and source of water in the Safe Drinking Water Information System data. These variables were re-coded based on publicly available data.

sanitary surveys, or enforcement actions, 4) respond to public complaints, 5) assist a primacy agency when it is requested to do so, and 6) to conduct routine inspections as a check-in on compliance if the system has not been assessed in a long time (U.S. Environmental Protection Agency, 2024). The mean for *Federal Inspection* is 0.0027, indicating that 0.27 percent of community water systems received a federal inspection in 2022, while 1.40 percent received a *Formal Enforcement Action*.

The U.S. Environmental Protection Agency's Safe Drinking Water Information System database presents data reported to the U.S. Environmental Protection Agency by primacy agencies and U.S. Environmental Protection Agency regions, as statutorily required for Public Water System Supervision programs. However, there is considerable variation in how much information is reported, which could occur for a number of unknown reasons, including the extent of work conducted by each state. Thus, a measure of information completeness is included in the quantitative analysis, which captures the proportion of missing information out of 45 characteristic fields (under *Public Water Systems*) in the dataset. This measure is included as a proxy for reporting quality, which may reflect oversight intensity or integrity. The *Proportion Unreported Data* has a mean of 35.63 percent, indicating more than one-third of data information fields in the U.S. Environmental Protection Agency's Safe Drinking Water Information System data contained missing values.

The main takeaway from these system-level oversight and enforcement variables in Table 2.3 is that on-site inspections related to sanitary surveys are most common, as they are statutorily required, with 35 percent of community water systems receiving one in 2022. Yet, only about one percent of systems required a formal enforcement action for observed noncompliance.

Population Density reflects how concentrated the population is in a given community water system service area. A system's service area has a mean of 217 individuals per km^2, a *Median Household Income* equal to $77,062, and a *Median Age* of 43.[5] The final set of variables in the *Proportion of Housing Stock* category focuses on the age of housing stock in a community water system service area. A system's service area has a mean of 22.60 percent of houses *Built Before 1960*, 8.50 percent *Built 1960–1969*, 15.54 percent *Built 1970–1979*, and 13.77 percent *Built 1980–1989*.

2.1.3. Historical Noncompliance

Table 2.4 groups community water systems into those that were *Consistently Compliant* (i.e., never had any violation), *In/Out of Compliance* (i.e., had a violation at least once), and *Persistent Violators* (i.e., had a violation every year) between 2018 and 2022 to provide comparisons among these three groups with all variables of interest. Column (4) in Table 2.4 reports the p-values from one-way Analysis of Variance (ANOVA) tests.[6] Each p-value

[5] The community water system service area variables are spatially weighted means of information available at the Census tract level.

[6] One-way Analysis of Variance is a statistical method used to determine if there are statistically significant differences in the means of three or more groups by comparing the variance between groups to the variance within groups using an F-statistic. Specifically, ANOVA tests the null hypothesis that all group means are equal against the alternative hypothesis that at least one group mean differs from the others. Reported p-values

reflects the result of testing for statistically significant differences in the means of each variable among the three groups. So, the numbers listed in columns (1)-(3) are the mean values for community water systems within each group.

Table 2.4. Comparison of Community Water Systems Based on Compliance in 2018-2022

Variables	Consistent Compliance (1)	In/Out of Compliance (2)	Persistent Violators (3)	P-value (ANOVA) (4)
Dependent Variables				
Any Violation, 2023-2024	0	1	1	-
Number of Violations, 2023-2024	0	7	29	<0.001
System-Level Characteristics				
Ownership				
Federal Gov.	0.0067	0.0065	0.0038	0.4371
Local Gov.	0.5261	0.5129	0.5034	0.0143
Public/Private	0.0214	0.0315	0.0225	<0.001
Native American	0.0064	0.0163	0.0495	<0.001
State Gov.	0.0078	0.0086	0.0083	0.1601
Private	0.4317	0.4242	0.4126	0.6790
Service Connections	2,604	2,027	674	<0.001
Self-Produced	0.7977	0.8185	0.8425	<0.001
Surface Water	0.2211	0.2323	0.2986	<0.001
Wholesaler	0.0576	0.0368	0.0435	<0.001
Regularly Scheduled Visits	0.0180	0.0208	0.0075	0.0011
Technical Assistance Visits	0.0080	0.0130	0.0233	<0.001
Sanitary Survey Visits	0.3557	0.3559	0.3698	0.5698
State Inspections	0.0112	0.0232	0.0555	<0.001
Federal Inspections	0.0023	0.0033	0.0060	0.0128
Any Violation 2018-2022	0	0.7572	1	<0.001
Proportion Unreported Data	0.3560	0.3574	0.3505	<0.001
Formal Enforcement Action	0	0.0315	0.1223	<0.001
Service Area Characteristics				
Population Density (per km^2)	239	186	93	<0.001
Median Household Income	$79,434	$73,457	$65,714	<0.001
Median Age	43.16	43.40	42.99	0.0028
Proportion of Housing Stock				
Built Before 1960	0.2261	0.2265	0.2196	0.3520
Built 1960-1969	0.0848	0.0852	0.0875	0.1834
Built 1970-1979	0.1544	0.1565	0.1649	<0.001
Built 1980-1989	0.1363	0.1391	0.1515	<0.001
Number of Observations	28,561	14,587	1,333	

Note: The *Consistent Compliance* category includes all community water systems that never had any violation 2018-2022. *In/Out of Compliance* includes all systems with at least one violation 2018-2022. *Persistent Violators* include systems with a violation each year 2018-2022. The values in columns (1)-(3) reflect the means of each variable for community water systems within each group. The p-values in column (4) are from one-way Analysis of Variance (ANOVA) tests comparing the means of each variable across the three compliance groups. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

below conventional statistical thresholds ($p<0.05$ or $p<0.01$) allow for the null hypothesis to be rejected, thereby leading to the conclusion that there are observed differences among group means that are unlikely to be due to random sampling variation alone.

Since the first p-value listed in Table 2.4 is statistically significant (i.e., $p<0.01$), the test result indicates there are differences among the three groups in terms of the mean number of violations in 2023–2024. Specifically, community water systems that are persistently noncompliant have a statistically significantly ($p<0.01$) higher number of violations, with 29 compared to seven for those in and out of compliance and none for those consistently in compliance.

The main takeaway from Table 2.4 is that persistently noncompliant community water systems with at least one violation each year 2018–2022 have a higher number of violations in 2023–2024, higher proportion of Native American ownership, fewer service connections serving smaller populations, higher rates of self-production, greater reliance on surface water, and receive more technical assistance visits and formal enforcement actions. Community water systems in full compliance 2018–2022 have the highest proportions of ownership by local governments.

In terms of ownership, a notably higher proportion of community water systems with Native American ownership are persistent violators, at 4.95 percent, compared to 1.63 percent of Native American-owned community water systems in flux, and 0.64 percent of Native American owned systems consistently in compliance ($p<0.01$). Community water systems that are persistent violators also tend to be smaller systems with a mean of only 674 service connections ($p<0.01$), self-produce water ($p<0.01$), have greater reliance on surface water ($p<0.01$), receive more technical assistance visits ($p<0.01$) and formal enforcement actions ($p<0.01$), and have service areas with lower population density ($p<0.01$) and lower median household income ($p<0.01$). On the other hand, community water systems that are persistent violators had a smaller proportion of systems receiving regularly scheduled visits ($p<0.05$) and a higher proportion of systems receiving state ($p<0.01$) and federal ($p<0.05$) inspections.

Community water systems in full compliance with the Safe Drinking Water Act between 2018 and 2024 include the highest proportions of systems owned by local governments ($p<0.05$) and the lowest proportions of systems owned by Native American tribes ($p<0.01$). Fully compliant systems are also largest, with a mean of 2,604 service connections ($p<0.01$), the lowest proportions self-producers ($p<0.01$), least reliant on surface water ($p<0.01$), and have a greater proportion of systems that serve as wholesalers ($p<0.01$). These community water systems also have the lowest proportions of systems receiving technical assistance visits ($p<0.01$) and both state ($p<0.01$) and federal ($p<0.05$) inspections. In terms of their service area characteristics, fully compliant community water systems operate in the most densely populated areas ($p<0.01$) with a noticeably higher median household income ($p<0.01$).

Notably, there are no statistically significant differences in mean values among the groups in terms of community water systems owned by the federal or state governments and private entities, those systems receiving sanitary survey visits, or systems operating in communities with the oldest housing stock.

2.1.4. Government Finance Characteristics

Table 2.5 provides comparisons of all variables for analysis between all community water systems (N=44,481) and a subset of only those owned by local governments (N=23,180), which comprise 52.11 percent of all community water systems. The mean for *Any Violation* indicates that 35.79 percent of community water systems had a violation in 2023 or 2024 for the full sample compared to 35.17 percent for local government owned systems. Although the mean *Number of Violations* is the same for both groups at three, local government owned systems portray less variability around this mean with a standard deviation of 12 compared to 16.

The main takeaway from Table 2.5 is that local governments owned about half of all community water systems, and these systems have very similar compliance rates and numbers of violations compared to all systems. However, local government owned systems tend to have more service connections serving larger populations, a lower proportion of self-producers, greater reliance on surface water, and a higher percentage of systems that sell water wholesale. In addition, a greater proportion of local government owned systems receive technical assistance visits as part of compliance oversight.

Table 2.5 shows that local government owned systems tend to be larger, with a mean of 3,903 *Service Connections* compared to a mean of 2,357 for all systems. However, there is also greater variation in size among community water systems owned by local governments, with a standard deviation of 21,147 compared to 15,972 for all systems.

In terms of how water is procured, although 80.58 percent of all community water systems have *Self-Produced* water, only 70.73 percent of local government owned systems self-produce their water. The remaining 29.27 percent of local government owned community water systems purchase water. The way in which local government owned systems source their water also differs compared to all community water systems. While only 22.71 percent of community water systems rely on *Surface Water*, 35.41 percent of local government owned systems rely on surface water. As such, 77.28 percent of all community water systems and 64.59 percent of those owned by local governments source from groundwater. Similarly, a higher percentage of local government owned systems (8.05 percent) act as a *Wholesaler* compared to 5.03 percent of all community water systems.

By comparison, local government owned systems received more site visits as part of oversight compliance compared to all community water systems. Specifically, 1.86 percent of community water systems received a *Regularly Scheduled Visit* in 2022 compared to 1.95 percent of local government owned systems. Among all community water systems, only 1.01 percent received a *Technical Assistance Visit*, whereas 10.66 percent of local government owned systems received such a visit. About 35.62 percent of all community water systems received a *Sanitary Survey Visit* compared to 37.32 percent of local government systems, while 1.65 percent of all community water systems received a *State Inspection* compared to 2.24 percent of local government owned systems. In terms of formal evaluations by the U.S. Environmental Protection Agency to assess compliance with regulatory standards, 0.27 percent of all community water systems received a *Federal Inspection* in 2022 compared to 0.38 percent of local government owned systems.

Table 2.5. Summary Statistics: All Systems Compared to Local Govt. Owned Systems

	All Systems		Local Govt. Owned Systems	
Variables	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Dependent Variables				
Any Violation	0.3579	0.4794	0.3517	0.4775
Number of Violations	3	16	3	12
System-Level Characteristics				
Ownership				
Federal Gov.	0.0065	0.0805		
Local Gov.	0.5211	0.4996	1	0
Public/Private	0.0247	0.1552		
Native American	0.0109	0.1040		
State Gov.	0.0080	0.0894		
Private	0.4287	0.4949		
Service Connections	2,357	15,972	3,903	21,147
Self-Produced	0.8058	0.3955	0.7073	0.4550
Surface Water	0.2271	0.4190	0.3541	0.4782
Wholesaler	0.0503	0.2186	0.0805	0.2720
Regularly Scheduled Visits	0.0186	0.1351	0.0195	0.1384
Technical Assistance Visits	0.0101	0.1000	0.1066	0.1027
Sanitary Survey Visits	0.3562	0.4789	0.3732	0.487
State Inspections	0.0165	0.1273	0.0224	0.1480
Federal Inspections	0.0027	0.0523	0.0038	0.0618
Any Violation 2018-2022	0.2783	0.4482	0.2795	0.4487
Proportion Unreported Data	0.3563	0.0408	0.3475	0.0394
Formal Enforcement Action	0.0140	0.1175	0.0138	0.1167
Service Area Characteristics				
Population Density (per km^2)	217	506	240	556
Median Household Income ($)	77,062	29,814	73,670	28,503
Median Age	43	7	43	6
Proportion of Housing Stock				
Built Before 1960	0.2260	0.1662	0.2718	0.1738
Built 1960-1969	0.0850	0.0560	0.0893	0.0521
Built 1970-1979	0.1554	0.0722	0.1531	0.0650
Built 1980-1989	0.1377	0.0731	0.1261	0.0653
Government Finance				
Debt Service Burden			0.0631	4.8706
Risk Exposure Factor			3	102
Operating Expenditure Ratio			0.8346	0.2124
Water Revenue Expenditure Ratio			2	61
Full Time Employees			1,399	6,607

Note: The number of observations is 44,481 for all systems and 23,180 for local government owned systems. Missing data for Full Time Employees were recoded as zeros. The recoding impacts 0.62 percent of observations. Robustness checks reported in the appendix assess the sensitivity of the regression results to the recoding by excluding recoded observations. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html, and U.S. Census Bureau's Census of Governments Finance Survey data for 2022 obtained from: https://www.census.gov/programs-surveys/cog.html.

Trends in inspections are somewhat surprising considering local government owned systems are a little less likely to be in violation of the Safe Drinking Water Act and also do not exhibit a greater number of violations. These systems are also no more likely to have a history of noncompliance. Among all community water systems, 27.83 percent had *Any Violation 2018–2022* compared to 27.95 percent of local government owned systems. Further, local government owned systems are also less likely to experience a *Formal Enforcement Action*, with 1.38 percent receiving one compared to 1.40 percent of all community water systems. Finally, the *Proportion Unreported Data* among all community water systems has a mean of 35.63 percent compared to 34.75 percent of local government owned systems, indicating better reporting practices among this subset of community water systems.

Population Density has a mean of 217 and 240 individuals per km^2 for all systems and local government owned systems, respectively. The *Median Household Income* is equal to $77,062 for all community water systems compared to $73,670 for local government owned systems, while the *Median Age* is 43 for both. In addition, local government owned systems have service areas with a mean 27.18 percent (compared to 22.60 percent for all systems) *Proportion of Housing Stock Built before 1960*, 8.93 percent (compared to 8.50 percent) *Built 1960–1969*, 15.31 percent (compared to 15.54 percent) *Built 1970–1979*, and 12.61 percent (compared to 13.77 percent) *Built 1980–1989*.

The final group of variables included in Table 2.5 focus on local government financial condition. In the Census of Governments Finance survey data, total revenues and total expenditures reflect all four sectors of government activity: general government, utilities, liquor stores, and social insurance trust. General revenue reflects revenue earned by the general government sector only (excluding utility, liquor store, and insurance trust sectors) from external sources and consists of: 1) taxes, 2) intergovernmental revenue, 3) current charges, and 4) miscellaneous general revenue, net of internal or interfund transfers.[7] This classification is based on the nature of the revenue source and not the fund or administrative unit that accounts for or controls a particular activity.

Among the four categories of general revenue, intergovernmental revenue comprises monies from other governments and is therefore the only category outside of a government's own sources. Further, all intergovernmental revenue is reported by the Census of Governments in the general government sector, even if it is used to support activities in other sectors, like community water systems. The remaining portion of general revenue (i.e., taxes, current charges, and miscellaneous revenue, net of internal or interfund transfers) is considered own-source revenue. In addition, the Census of Governments Finance survey data include information on indebtedness, which is also only reported in the general government sector, regardless of its use. Debt information includes total debt outstanding and interest expenditures paid on outstanding debt, among other variables.

[7] See U.S. Bureau of the Census, Government Finance and Employment Classification Manual, available at: https://www2.census.gov/govs/pubs/classification/2006_classification_manual.pdf.

The main takeaway for the government finance variables in Table 2.5 is that local governments that own community water systems spend about 6 percent of their total revenue to cover interest costs on outstanding debt, would need to levy an additional $3 in property tax revenue to offset a $1 reduction in intergovernmental revenue, and dedicate 83 percent of total spending to cover operations. In addition, these local governments levy about twice the necessary revenue to fund their water utilities; however, evidence suggests some of this revenue is used to support unrestricted operating (i.e., general fund) activities.

In this analysis, *Debt Service Burden*, calculated as interest expenditures on debt (from all sectors) divided by total revenue (from all sectors), has a mean value of 0.0631, suggesting that local governments are required to expend a mean of 6.31 percent of total revenue to cover interest costs on outstanding debt. The *Risk Exposure Factor* is defined as intergovernmental revenue (reflecting the amount of revenue that might decline due to factors totally outside of the local government's control) divided by property tax revenue (reflecting the largest source of controllable revenue for most local governments). The mean value of 3 indicates that a $1 decline in uncontrollable intergovernmental revenue would require a mean of $3 more levied in property tax revenue to make up the difference.

The *Operating Expenditure Ratio*, measured as total current operating expenditure (for all sectors) divided by total expenditure (for all sectors), has a mean of 0.8346, indicating that most spending (83.46 percent) is devoted to ongoing operations and must therefore be funded by revenue collected in the same fiscal year. However, the *Water Revenue Expenditure Ratio*, defined as water utility sector revenue divided by water utility sector expenditure has a mean of two, indicating local government water utilities levy twice the amount needed for interest on debt, current operations, and capital outlays. However, recent research suggests, at least anecdotally, that some local government water systems levy additional revenue through their water billing and use it to support unrestricted operating (i.e., general fund) activities (Carroll et al., 2025). So, the literal interpretation of this ratio could be misleading in terms of actual practice in the field. Finally, local governments employ a mean of nearly 1,399 full time employees; however, the number of full-time local government employees varies considerably depending on the size of the population served with a standard deviation of 6,607.

2.2. Methods

2.2.1. Binary Dependent Variable

As previously mentioned, one of the dependent variables used for analysis is binary and equal to one if a community water system had any Safe Drinking Water Act violations in 2023 or 2024. As such, this model is estimated using Probit, with Logit and Linear Probability estimation results reported for robustness in Table A.6. It is important to note that the estimated coefficients from the Probit model do not directly reflect changes in outcome probabilities. Thus, marginal effects are calculated and also reported to translate the coefficients into meaningful changes in the probability of a community water system having a violation. All estimations also include state fixed effects to account for unobserved, time-invariant state-level factors that may also correlate with noncompliance.

Supplemental Technical Documentation

For binary dependent variables, commonly used estimation methods include the Linear Probability Model (LPM) (Lee et al., 2025), Logit (Demaris, 1992), and Probit (Allaire et al., 2018). The LPM estimates the probability of compliance using ordinary least squares, which allows for a straightforward interpretation of the coefficients as percentage-point changes in the likelihood of compliance. However, because the LPM assumes a linear relationship and can yield predicted probabilities outside the 0–1 range, the use of nonlinear models, e.g. Logit and Probit, that constrain predicted probabilities between 0 and 1, is preferred.

The Logit model uses a logistic cumulative distribution function, while the Probit model uses the standard normal distribution to estimate the probability of noncompliance. Both Logit and Probit provide nearly identical substantive results, though the Probit estimation assumes normally distributed errors and typically produces smoother probability transitions. The Probit model is estimated using the following functional form:

$$Pr\ (Y_i = 1|X) = \Phi\ (B_0 + B_1 S_i + B_2 A_i + \phi_s)$$

where Y_i is a binary indicator of whether community water system i had a violation in 2023 or 2024, and $\Phi\ (\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution. The probability of a community water system having a violation is modeled as a function of time-invariant system characteristics (S_i) and service area characteristics (A_i).

Each community water system i is associated with a state s, and the model includes state fixed effects (ϕ_s), which are binary indicator variables for each state to account for unobserved, time-invariant state-level factors. Accounting for unobserved, time-invariant factors is important because the Safe Drinking Water Act is primarily enforced at the state level, and enforcement may vary due to differences in sampling protocols, technical capacity, and financial resources (Allaire et al., 2018).

2.2.2. Count Dependent Variable

The other dependent variable is the count of violations in 2023 and 2024 for those community water systems that were noncompliant in either year. Since the number of violations is measured with whole, non-negative integers representing discrete outcomes, and the variance in the data exceeds the mean indicating overdispersion, this model is estimated using a Negative Binomial model. Here again, the estimated coefficients from the Negative Binomial model do not directly represent changes in the number of violations. Instead, coefficients describe changes in the log of the expected count. Marginal effects are therefore calculated and reported to translate these estimates into meaningful changes in the expected number of violations. All estimations also include state fixed effects to account for unobserved, time-invariant state-level factors.

Supplemental Technical Documentation

For count data as dependent variables, Poisson or Negative Binomial regression is typically used to estimate the correlations of independent variables on the mean count, using maximum likelihood. These models are needed to accurately model the probability distribution because count variables violate the normality and homoskedasticity

assumptions of the Linear Probability Model, likely leading to predictions of negative values, which are impossible for count data. Negative Binomial regression is used in this analysis because the variance in the data exceeds the mean, indicating overdispersion (Cameron & Trivedi, 2013).

The Negative Binomial model estimates the expected number of violations conditional on a set of covariates and assumes a log link function between the conditional mean and the linear correlates. The model is estimated using the following functional form:

$$E\ (Y_i|X_i) = exp\ (B_0 + B_1 S_i + B_2 A_i + \phi_s)$$

where Y_i denotes the number of violations for community water system i in 2023 and 2024. The expected count of violations is modeled as a function of time-invariant system characteristics (S_i) and service area characteristics (A_i). The exponential link function ensures that predicted values are strictly non-negative. Each community water system i is associated with a state s, and the model includes state fixed effects (ϕ_s).

2.2.3. Model Estimations

Covariates included in the Probit and Negative Binomial models were selected based on a review of the extant literature, multicollinearity diagnostics, and expert feedback. These variables are all observed in 2022 and categorized as 1) *System-Level*, capturing ownership, system capacity, water provision, oversight, reporting, prior year violations, and enforcement, 2) *Service Area*, capturing population density, median household income, median age, and the age of the housing stock, and for those community water systems owned by local governments, 3) *Government Finance*, capturing local government fiscal capacity.

Since residual errors are likely correlated within states as they share the same regulatory structure and regional water quality, standard errors are clustered at the state level. Failing to account for this within-state correlation could lead to underestimated standard errors and overstated statistical significance.

Supplemental Technical Documentation

Covariates included in the Probit and Negative Binomial models were selected based on a review of the extant literature, multicollinearity diagnostics, and expert feedback. System Characteristics (X_i) include variables capturing ownership, system capacity, water provision, oversight, reporting, prior year violations, and enforcement. Service Area Characteristics (C_i) include variables capturing population density, median household income, median age, and the age of the housing stock. The covariates are observed in 2022. Thus, noncompliance in 2023–2024 is explained by characteristics determined before the behavior is observed, helping to reduce concerns about potential endogeneity in the model specification (Heberling et al., 2022).

In addition, for a subset of the data, specifically local government owned systems, the models are expanded as follows:

$$Pr\ (Y_i = 1|X) = \Phi\ (B_0 + B_1 S_i + B_2 A_i + B_3 G_i + \phi_s)$$

$$E\ (Y_i|X_i) = exp\ (B_0 + B_1 S_i + B_2 A_i + B_3 G_i + \phi_s)$$

where Government Finance Characteristics (G_i) include variables capturing aspects of local government fiscal capacity using commonly used measures obtained from review of extant research.

2.3. System-Level Correlates of Noncompliance

Table 2.6 presents regression estimates of system-level correlates of Safe Drinking Water Act noncompliance across three models. Columns (1)–(2) estimate the likelihood of any violation during 2023–2024 for all systems using Probit, while Columns (3)–(6) estimate the number of violations among noncompliant systems using a Negative Binomial regression.[8]

The main takeaway from Table 2.6 is that community water systems are more likely to have a violation and, when noncompliant, likely to have more violations if they have the following system-level characteristics: Native American ownership, self-production, and surface water reliance. Systems are also more likely to be found noncompliant if they had a federal inspection or state technical assistance visit. Wholesaler systems and those receiving regularly scheduled state visits are less likely to have violations and, when noncompliant, have fewer violations. Poorer reporting quality, which may reflect less primacy agency oversight intensity is also associated with greater noncompliance. Finally, among noncompliant systems, having a prior violation during 2018–2022 and receiving a formal enforcement action are both correlated with a greater number of violations.

In terms of ownership, results indicate that Native American owned systems are 26.74 percentage points (see Marginal Effect in Table 2.6) more likely to have a violation relative to privately owned systems (i.e., the omitted category). Among noncompliant systems, Native American owned systems also report significantly more violations, nearly four more. However, this is the only ownership category found to be correlated with Safe Drinking Water Act noncompliance.

The results indicate that system size is inversely correlated with noncompliance. An increase in the number of service connections by 100,000, which reflects systems serving larger populations, is correlated with a 9.04 percentage point reduction in the likelihood of a violation. However, this finding does not hold up to robustness checks reported in Appendix A, suggesting system size is not a strong predictor of Safe Drinking Water Act compliance. Further, among noncompliant systems, the number of service connections is not predictive of the number of violations.

[8] All models include state fixed effects and cluster standard errors at the state level. Robustness checks are reported in Appendix A. Table A.6 assesses the sensitivity of the results to the functional form. Tables A.7 and A.8 assess the sensitivity of the regression results to missing data imputations by (i) excluding imputed observations and (ii) including indicator variables for imputed values. Table A.9 examines the sensitivity of the results to a shorter timeframe (only 2024) when defining noncompliance.

Table 2.6. System-Level Regression Results: All Community Water Systems

	Any Violation 2023-2024 All Systems		Number Violations 2023-2024 Noncompliant Systems		Number Violations 2023-2024 Noncompliant Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)	Coefficient (Std. Error) (5)	Marginal Effect (6)
System-Level Characteristics						
Federal Government	0.0334	0.0114	-0.1387	-1.650	-0.1199	-1.0253
	(0.1081)		(0.1325)		(0.1223)	
Local Government	0.0461	0.0158	-0.0908	-0.7812	-0.0919	-0.7966
	(0.0440)		(0.0634)		(0.0587)	
Public/Private	-0.1060	-0.0352	0.0702	0.6543	0.0867	0.8219
	(0.1074)		(0.2110)		(0.1980)	
Native American	**0.7466*****	0.2674	**0.3347***	3.5769	**0.3425***	3.7067
	(0.2460)		(0.1925)		(0.1870)	
State Government	0.1185	0.0409	-0.1974	-1.6120	-0.1815	-1.5065
	(0.0952)		(0.1717)		(0.1581)	
Service Connections (100k)	**-0.2644*****	-0.0904	-0.0908	-0.7838	-0.0627	-0.5456
	(0.0817)		(0.0793)		(0.0731)	
Self-Produced	**0.2095*****	0.0716	**0.5421*****	4.6793	**0.5187*****	4.5149
	(0.0456)		(0.0653)		(0.0658)	
Surface Water	**0.1640*****	0.0561	**0.3139*****	2.7094	**0.2931*****	2.5513
	(0.0418)		(0.0605)		(0.0572)	
Wholesaler	**-0.1983*****	-0.0678	**-0.3596*****	-3.1041	**-0.2992*****	-2.6046
	(0.0486)		(0.0915)		(0.0859)	
Regularly Scheduled Site Visit	0.0376	0.0129	**-0.2573*****	-2.2207	**-0.2337****	-2.0343
	(0.0397)		(0.0941)		(0.0930)	
Technical Assistance Visit	0.0975	0.0333	**0.6708*****	5.7902	**0.5204*****	4.5283
	(0.1986)		(0.2188)		(0.1647)	
Sanitary Survey Visit	0.0087	0.0030	0.0427	0.3684	0.0562	0.4895
	(0.0175)		(0.0409)		(0.0458)	
State Inspection	0.0101	0.0034	**0.0825***	0.7119	0.0082	0.0712
	(0.0331)		(0.0423)		(0.0335)	
Federal Inspection	**0.2434****	0.0832	0.1777	1.5335	0.1783	1.5523
	(0.1133)		(0.1594)		(0.1599)	
Proportion Unreported Data	**2.1846*****	0.7468	**2.8572*****	24.6621	**2.6023*****	22.6510
	(0.7532)		(0.9569)		(1.0055)	
Any Violation 2018-2022					**0.4147*****	3.6097
					(0.0766)	
Formal Enforcement Action					**0.6929*****	6.0310
					(0.1580)	
Mean	0.3587		9		9	
Pseudo R2	0.0801		0.0380		0.0443	
Number of Observations	44,378		15,920		15,920	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(6) present results from a negative binomial model. Coefficients are presented in Column (1), (3), and (5), while mean marginal effects are presented in Columns (2), (4), and (6). In the probit model, two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads.

Self-producing systems are 7.16 percentage points more likely to incur violations and, when noncompliant, tend to have nearly five more violations than systems that purchase water. These results are aligned with the literature where purchased water sources are shown to be less prone to Safe Drinking Water Act violations (Allaire et al., 2018; Switzer & Teodoro, 2017), which likely reflects that these systems are also subject to fewer regulations.

Similarly, systems that rely on surface water are 5.61 percentage points more likely to incur violations and, when noncompliant, have nearly three more violations compared to those relying on groundwater. This aligns with previous literature finding that systems with groundwater sources often experience greater compliance compared to surface water systems (Switzer & Teodoro, 2017). Conversely, wholesaler systems are about 6.78

percentage points less likely to have violations and, when noncompliant, have about three fewer violations. Previous literature has found that utilities purchasing their water have lower levels of noncompliance (Konisky & Teodoro, 2016; Switzer & Teodoro, 2017; Wallsten & Kosec, 2008), so this finding suggests that utilities selling water are more likely to be compliant themselves.

In terms of regulatory oversight, community water systems receiving a federal inspection in 2022 have an 8.32 percentage point higher likelihood of a violation in 2023 or 2024, suggesting that federal oversight through on-site visits is effective at discovering noncompliance as intended. However, state oversight appears to work differently. Although a state inspection in 2022 is correlated with nearly 1 additional violation in 2023 or 2024, this result does not hold up to the inclusion of variables on historical violations and enforcement actions, suggesting having a state inspection is not a strong predictor of Safe Drinking Water Act compliance among all systems. However, among the subset of noncompliant systems, having a regularly scheduled state site visit in 2022 is correlated with more than 2 fewer violations in 2023 and 2024, whereas a technical assistance visit in 2022 is correlated with about 5 more violations in 2023 or 2024. This latter finding is unsurprising since technical assistance visits are designed for smaller systems with pre-existing challenges. At the same time, however, these results suggest that regular visits likely provide effective oversight to motivate community water systems to remain compliant.

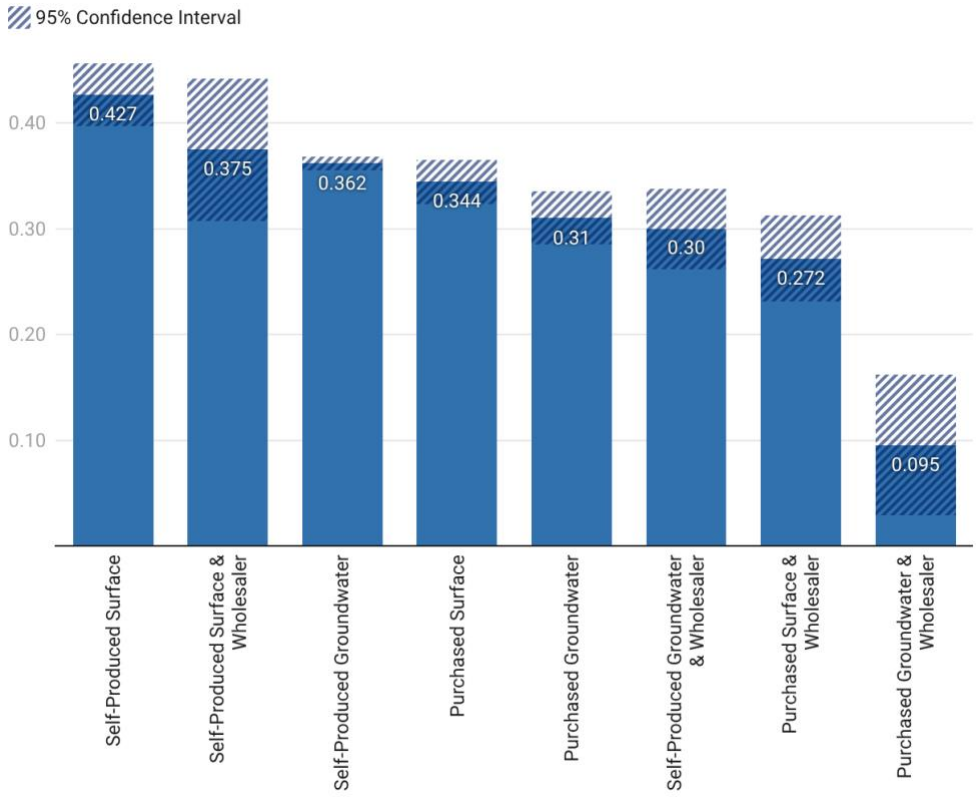
The results also show that a 10-percentage point increase in the proportion of unreported system characteristics is correlated with a 7.47 percentage point increase in the probability of having a violation. Among noncompliant systems, a 10-percentage point increase in the proportion of unreported system characteristics is correlated with a little more than two additional violations. These findings suggest that poorer reporting quality, which may reflect less primacy agency oversight is associated with greater noncompliance with the Safe Drinking Water Act.

In Table 2.6, columns (1)-(4) use the same model specification. Columns (5)-(6) add two binary variables, capturing 1) any historical violation between 2018-2022, and 2) receiving a formal enforcement action in 2022. Since all systems with historical violations also have a violation in 2023 or 2024, the lack of variation causes this variable to get dropped from the probit model reported in columns (1)-(2). Among noncompliant systems, having a prior violation during 2018–2022 is correlated with nearly four more violations in 2023 or 2024, while receiving a formal enforcement action in 2022 is correlated with about six additional violations in the following two years. These findings reiterate the descriptive trends described earlier that some community water systems are persistently in violation with the Safe Drinking Water Act and may need technical assistance to maintain compliance over time.

To further explore the water provision variables, Table A.5 in the appendix presents results from a regression that adds a three-way interaction term to the specification presented in Column 1 of Table 2.6. Water provision variables (source, procurement method, wholesaler status) have been documented in the literature as associated with compliance, treatment costs, and water rates, among other outcomes (Carroll et al., 2023; Piper, 2003; Switzer &

Teodoro, 2017). The inclusion the interaction term in Table A.5 helps us understand whether the correlation of one provision characteristic depends on the presence of another, e.g., does self-producing water correlate with violations similarly for groundwater and surface water systems? For each of the eight possible interactions, representing different water source, procurement method, and wholesaler status combinations, the predicted probabilities of having a violation, based on the results in Table A.5., are calculated for all 44,378 observations, while holding all other covariates constant. Figure 2.3 illustrates the mean probability of noncompliance by each water provision interaction.

Figure 2.3. Probability of Noncompliance Based on Water Provision Interactions



Note: This figure illustrates the mean probability of noncompliance for the 3,256 systems that self-produce surface water, 473 systems that self-produce and wholesale surface water, 30,925 systems that self-produce groundwater, 5,857 systems that purchase surface water, 2,204 systems that purchase groundwater, 1,191 systems that self-produce and wholesale groundwater, 575 systems that purchase and wholesale surface water, and 59 systems that purchase and wholesale groundwater. For each of the eight water provision interactions, representing different water source, procurement method, and wholesaler status combinations, the predicted probabilities (derived from the regression results in Table A.5) of having a violation are calculated for all 44,378 observations, while holding all other covariates at their observed values. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads.

The main takeaway from Figure 2.3 is that self-production of surface water is correlated with higher likelihoods of violations, while wholesale of purchased water is correlated with lower likelihoods of violations.

Interacting the three provision variables, the model-based marginal predictions indicate that systems that self-produce and wholesale surface water have the second highest predicted probability of noncompliance (37.5%), while self-producers of surface water that do not wholesale have the highest predicted probability (42.7%) of having Safe Drinking Water Act violations. Wholesalers that purchase surface water have the second lowest predicted probability (27.2%) of Safe Drinking Water Act violations. Systems that purchase and wholesale groundwater have the lowest predicted probability (9.5%) of Safe Drinking Water Act violations.

2.4. Service Area Correlates of Noncompliance

Table 2.7 presents the results from regressions that add service area characteristics as additional control variables to the models. The coefficients on system characteristics are robust to the inclusion of these service area control variables and are consistent with the results presented in Table 2.6.

The main takeaway from Table 2.7 is that the addition of service area characteristics into the model does not change the results of system-level correlates of noncompliance. Community water systems operating in more densely populated areas and those with wealthier residents are less likely to have violations, partly due to economies of scale in drinking water provision. Among noncompliant systems, service area characteristics are generally not associated with the number of violations, except in areas with the oldest housing stock, potentially indicating broader disinvestment in these areas.

Among the service area characteristics, population density and median household income have negative and statistically significant coefficients. A 100-person increase per km^2 is correlated with a one percentage point decrease in noncompliance, while a $1,000 increase in median household income is correlated with a 0.08 percentage point decrease in noncompliance. These findings are aligned with prior research (Allaire et al., 2018; Switzer & Teodoro, 2017) and are indicative of the benefits of economies of scale in the provision of drinking water.

Conversely, a 10-percentage point increase in the proportion of pre-1960 housing, relative to the proportion of 1990+ housing (the omitted category), is correlated with a 1.41 percentage point increase in the likelihood of having a violation. Similarly, a 10-percentage point increase in the proportion of housing stock built in the 1980s, with a corresponding decrease in 1990+ housing, is correlated with a higher likelihood of noncompliance (a 1.22 percentage point increase). This trend might be reflecting the overall age of the built environment and water infrastructure.

Notably, among noncompliant systems, service area characteristics are generally not predictive of the number of violations. The only exception is the proportion of housing built in the 1960s, where a 10-percentage point (i.e., 0.10) increase in this proportion is correlated with nearly one additional violation. These findings suggest that areas experiencing disinvestment in housing supply may face under-investment more broadly that creates challenges for systems to remain compliant with the Safe Drinking Water Act. In general, housing built before 1950 is more prevalent in central cities, while housing constructed in

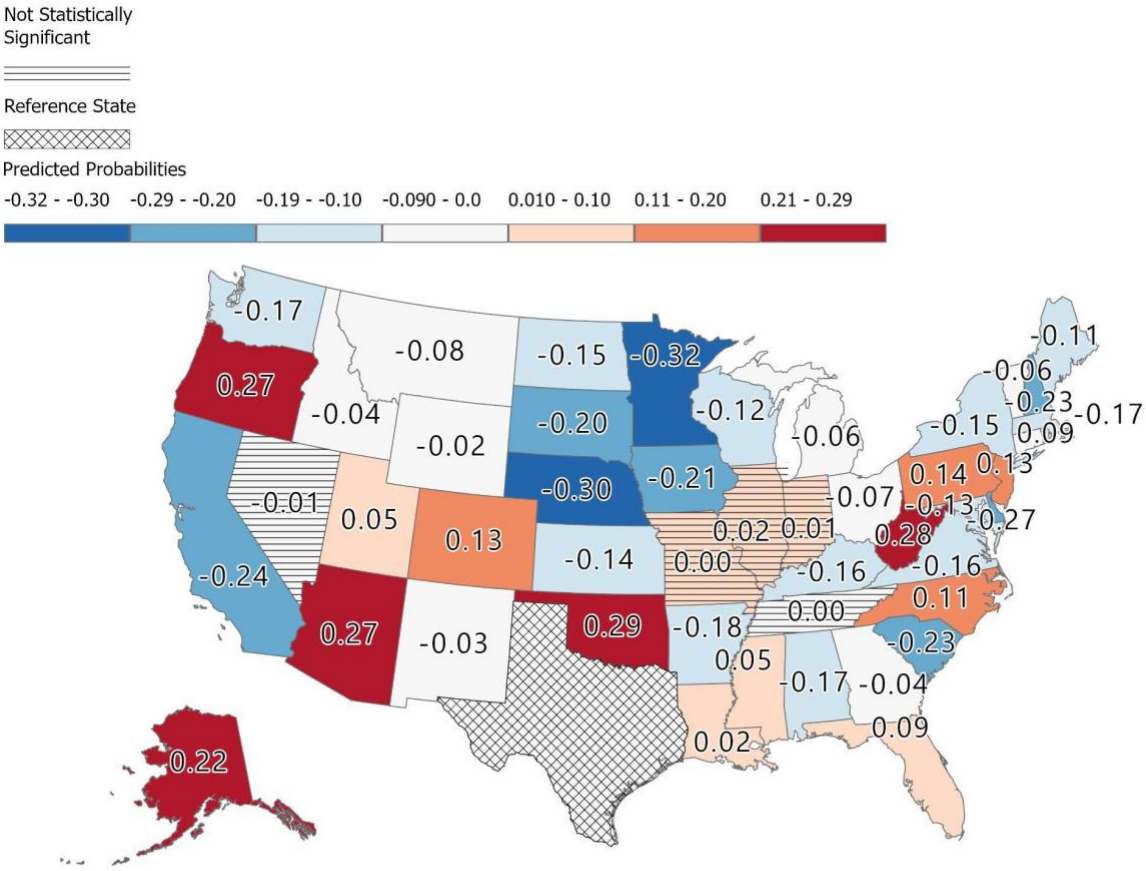
Table 2.7. System and Service Area Regression Results: All Community Water Systems

	Any Violation 2023-2024 All Systems		Number of Violations 2023-2024 Noncompliant Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Federal Government	-0.0251	-0.0085	-0.1093	-0.9313
	(0.1097)		(0.1366)	
Local Government	-0.0050	-0.0017	-0.0784	-0.6787
	(0.0435)		(0.0711)	
Public/Private	-0.1151	-0.0383	0.0950	0.8964
	(0.1052)		(0.1975)	
Native American	**0.6577*** **	0.2347	**0.3272***	3.4826
	(0.2431)		(0.1885)	
State Government	0.0851	0.0293	-0.1722	-1.4235
	(0.0922)		(0.1815)	
Service Connections (100k)	**-0.1294***	-0.0439	-0.0446	-0.3876
	(0.0687)		(0.0761)	
Self-Produced	**0.1721*** **	0.0584	**0.5173*** **	4.4970
	(0.0362)		(0.0646)	
Surface Water	**0.1764*** **	0.0599	**0.2909*** **	2.5292
	(0.0429)		0.0577	
Wholesaler	**-0.1554*** **	-0.0527	**-0.3006*** **	-2.6131
	(0.0406)		(0.0871)	
Regularly Scheduled Site Visit	0.0139	0.0047	**-0.2308** **	-2.0062
	(0.0403)		(0.1049)	
Technical Assistance Visit	0.0601	0.0204	**0.5145*** **	4.4726
	(0.2027)		(0.1647)	
Sanitary Survey Visit	0.0062	0.0021	0.0530	0.4610
	(0.0163)		(0.0449)	
State Inspection	0.0024	0.0008	0.0124	0.1081
	(0.0315)		(0.0341)	
Federal Inspection	**0.2980** **	0.1011	0.1831	1.5919
	(0.1157)		(0.1605)	
Proportion Unreported Data	**2.1445*** **	0.7276	**2.5182*** **	21.8933
	(0.7712)		(0.9819)	
Any Violation 2018-2022			**0.4118*** **	3.5797
			(0.0717)	
Formal Enforcement Action			**0.6934*** **	6.0287
			(0.1578)	
Service Area Characteristics				
Population Density (per km2)	**-0.0002*** **	-0.0001	-0.00002	-0.0001
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	**-0.0025*** **	-0.0008	-0.0005	-0.0045
	(0.0007)		(0.0010)	
Median Age	-0.0021	-0.0007	0.0012	0.0106
	(0.0016)		(0.0033)	
Built Before 1960	**0.4154*** **	0.1409	-0.2151	-1.8703
	(0.1245)		(0.1845)	
Built 1960-1969	0.3522	0.1195	**0.6796***	5.9084
	(0.2293)		(0.3848)	
Built 1970-1979	0.0224	0.0076	0.3281	2.8528
	(0.1267)		(0.2488)	
Built 1980-1989	**0.3591****	0.1218	0.1770	1.5389
	(0.1479)		(0.2486)	
Mean	0.3587		9	
Pseudo R2	0.0865		0.0447	
Number of Observations	44,378		15,920	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

the 1950s and 1960s reflects the first ring of suburban development; housing built in the 1990s and later is more commonly correlated with second-ring suburbs (Aviles, 2022).

Figure 2.4. Model-Based Differences in Noncompliance by State



Note: This Figure maps the state fixed effect coefficients derived from probit estimation results reported in Table 2.7 to illustrate geographical differences in the likelihood of community water system noncompliance that are not explained by the variables in the model. The color-coded coefficients represent how each state's mean probability of noncompliance differs from that of Texas (the reference category because it has the highest number of water systems), holding all other variables constant. Two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations.

Highlighting the role that state characteristics play in ensuring compliance, Figure 2.4 maps the state fixed effect coefficients, representing differences in the likelihood of Safe Drinking Water Act noncompliance across states that are not explained by the variables in the model. The color-coded coefficients represent how each state's mean probability of noncompliance differs from that of Texas (the excluded state), holding all other variables constant.[9]

[9] The illustrated coefficients are derived from the probit model (Table 2.7), so two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations.

As shown in Figure 2.4, community water systems in Minnesota are 32 percentage points less likely to have a violation than those in Texas, followed by Nebraska at 30 percentage points. Conversely, water systems in Oklahoma are 29 percentage points more likely to have a violation than systems in Texas, followed by West Virginia at 28 percentage points. Notably, the state fixed effects coefficients for Nevada, Missouri, Illinois, Indiana, and Tennessee all fail to reach statistical significance, indicating community water systems in those states are no more or less likely to have a violation than the systems in Texas.

2.5. Government Finance Correlates of Noncompliance

The association between fiscal capacity and Safe Drinking Water Act noncompliance is assessed through analysis of the subset of local government owned community water systems, because this group is the only one for which financial data are publicly available and can be matched to system service areas. Table 2.8 presents regression estimates of all correlates of Safe Drinking Water Act noncompliance across two models. Columns (1)–(2) estimate the likelihood of any violation during 2023–2024 for all local government owned systems using Probit, while Columns (3)–(4) estimate the number of violations among noncompliant local government owned systems using a Negative Binomial regression. Table 2.8 uses the same model specification as that used for analysis of all community water systems, except the system ownership types are no longer included as control variables, since the subset analysis only includes local government owned systems. In addition, this subset analysis adds local government fiscal capacity variables.[10]

The main takeaway from Table 2.8 is that the system-level and service area correlates of local government owned systems are largely consistent of those for all systems with some important differences. Noncompliant local government owned systems that are wholesalers and those receiving technical assistance visits have fewer violations compared to all noncompliant systems. The wealth of the service area and age of housing stock have more pronounced correlations with noncompliance, likely because local governments heavily rely on their residents and property taxation to generate revenue for public service provision. Noncompliant local governments may be using intergovernmental revenue to fund water system activities to help reduce violations, while some may be levying additional revenue through water billing and using it to support unrestricted operating (i.e., general fund) activities, ultimately deferring maintenance on their water systems that help them to remain compliant.

Texas serves as the omitted or reference category because it has the highest number of water systems. It is important to note that it is not possible to directly compare two non-reference state coefficients to each other without conducting a formal statistical test. To do so, a Wald test is typically used to assess whether the difference between two coefficients is statistically significant.

[10] Tables A.10 and A.11 can be used to better understand how the results from Table 2.7 compare to those in Table 2.8. Table A.10 presents results for all systems excluding ownership types, and Table A.11 presents results for local government owned systems prior to the inclusion of local government fiscal capacity variables. Tables A.12-A.14 assess the sensitivity of the regression results to the (1) exclusion of multijurisdictional government units, (2) controlling for government types, and (3) excluding imputed observations.

Table 2.8. Regression Results: Local Government Owned Community Water Systems

	Any Violation 2023-2024 All Local Govt. Systems		Number of Violations 2023-2024 Noncompliant Local Govt. Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Service Connections (100k)	-0.0627	-0.0205	-0.0652	-0.5698
	(0.0455)		(0.0616)	
Self-Produced	**0.1671*****	0.0547	**0.4742*****	4.1418
	(0.0403)		(0.0676)	
Surface Water	**0.2032*****	0.0665	**0.2169*****	1.8948
	(0.0377)		(0.0511)	
Wholesaler	**-0.1211****	-0.0396	**-0.1072****	-0.9367
	(0.0516)		(0.0545)	
Regularly Scheduled Site Visit	0.0799	0.0261	**-0.2686***	-2.3457
	(0.0666)		(0.1436)	
Technical Assistance Visit	-0.0288	-0.0094	**0.3729****	3.2568
	(0.0819)		(0.1675)	
Sanitary Survey Visit	-0.0006	-0.0002	0.048	0.4195
	(0.0176)		(0.0353)	
State Inspection	-0.0336	-0.011	**0.1370***	1.1964
	(0.0255)		(0.0818)	
Federal Inspection	**0.3296*****	0.1078	0.1733	1.5135
	(0.1269)		(0.2125)	
Proportion Unreported Data	**1.2154***	0.3976	0.5673	4.9549
	(0.6373)		(0.9319)	
Any Violation 2018-2022			**0.4090*****	3.5726
			(0.067)	
Formal Enforcement Action			**0.4954*****	4.3267
			(0.1434)	
Service Area Characteristics				
Population Density (per km2)	**-0.0002*****	-0.0001	-0.00003	-0.0003
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	**-0.0036*****	-0.0012	0.0005	0.0041
	(0.0007)		(0.0014)	
Median Age	-0.0019	-0.0006	-0.0052	-0.0455
	(0.0022)		(0.0032)	
Built Before 1960	**0.5848*****	0.1913	0.005	0.0438
	(0.1513)		(0.2809)	
Built 1960-1969	0.3794	0.1241	0.8318	7.2657
	(0.3093)		(0.5371)	
Built 1970-1979	**0.2876***	0.0941	**0.5848***	5.1076
	(0.1741)		(0.3546)	
Built 1980-1989	**0.4000****	0.1308	-0.1059	-0.9253
	(0.1778)		(0.4808)	
Government Finance				
Debt Service Burden	-0.357	-0.1168	**-2.0505***	-17.9103
	(0.4074)		(1.0787)	
Risk Exposure Factor	0.0097	0.0032	**-0.0143*****	-0.1251
	(0.0076)		(0.0006)	
Operating Expenditure Ratio	-0.0074	-0.0024	-0.0107	-0.0938
	(0.0104)		(0.0164)	
Water Revenue Expenditure Ratio	**0.0173***	0.0057	**0.0080*****	0.0696
	(0.0097)		(0.0011)	
Full Time Employees (1,000)	**-0.0104***	-0.0034	-0.0062	-0.0539
	(0.0056)		(0.0051)	
Mean	0.3525		9	
Pseudo R2	0.1131		0.0455	
Number of Observations	23,129		8,153	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, one observation from DC and 50 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html, and U.S. Census Bureau's Census of Governments Finance Survey data for 2022 obtained from: https://www.census.gov/programs-surveys/cog.html.

The results in Table 2.8 for the system-level and service area characteristics are largely consistent with the results from the full analysis (Table 2.7), with a few exceptions. Similar to the results from the full dataset, Table 2.8 shows that self-produced and surface water are correlated with higher probabilities of noncompliance, respectively 5.47 and 6.65 percentage points more likely, while wholesaler status is correlated with a lower probability of noncompliance (3.96 percentage points less likely) for systems owned by local governments. When local government owned systems are noncompliant, they are correlated with four more violations if they self-produce their water and nearly two greater violations if they rely on surface water. Conversely, noncompliant systems that serve as wholesalers are correlated with nearly one less violation, which is half the magnitude of this finding from the analysis with all systems.

The results pertaining to regulatory oversight are also rather consistent with the analysis of all systems. Noncompliant community water systems receiving a regularly scheduled site visit in 2022 are correlated with a little over two fewer violations in 2023 or 2024. Among noncompliant systems, a technical assistance visit in 2022 is correlated with three more violations in 2023 or 2024 among local government owned systems, which is one less violation from the full analysis (Table 2.7).

Also similar to the analysis of all systems, receiving a federal inspection in 2022 is correlated with a 10.78 percentage point increase in the likelihood of a government owned system having a violation in 2023 or 2024. However, the likelihood of any community water system having a violation increases to 12.70 percent when excluding ownership indicators (Table A.10) and to 11.24 percent when examining local government owned systems without the fiscal capacity variables (Table A.11).

Further, among noncompliant systems, receiving a state inspection is correlated with a little more than one greater violation for local government owned systems. This finding differs from the analysis of all systems (and Table A.10) in which the state inspection variable fails to reach statistical significance. However, this finding is not robust to the exclusion of the fiscal capacity variables when examining local government owned systems (Table A.11). This variable also fails to reach statistical significance when multijurisdictional systems are excluded from analysis (Table A.12). Taken together, these findings suggest state inspections have little impact on noncompliance by local government owned systems.

Table 2.8 shows that local government owned systems with a 10-percentage point increase in the proportion of unreported data have a 3.975 percentage point greater likelihood of having a violation. However, this finding is less robust in terms of statistical significance ($p<0.10$) than the analysis of all systems and Table A.10 in which the proportion of unreported data is statistically significant ($p<0.01$ or $p<0.05$) in both the probit and negative binomial model estimations. In addition, this variable fails to reach statistical significance in the negative binomial model estimation in Table 2.8, suggesting unreported data, as a proxy for the level of state oversight, is unrelated to the number of violations among local government owned systems. Notably, the mean proportion of unreported data is 34.75% for local government owned systems compared to 36.58% for other systems, a statistically significant difference of 1.8 percentage points ($p<0.01$).

The results in Table 2.8 for the service area characteristics are also consistent with the full analysis results (Table 2.7), where population density and median household income have negative and statistically significant coefficients in the probit model. Specifically, a 100-person increase per km^2 is correlated with a 0.01 percentage point decrease in the probability of local government owned community water systems having a violation, while a $1,000 increase in median household income is correlated with a 0.12 percentage point decrease in the likelihood of noncompliance. This latter finding is more pronounced than the analysis of all systems, which showed a 0.08 percentage point decrease (0.09 when excluding ownership in Table A.10), which is unsurprising since local governments heavily rely on the wealth of their residents to generate revenue for public service provision. In addition, these findings are aligned with prior research (Allaire et al. 2018; Switzer & Teodoro, 2017).

Further, Table 2.8 shows that local government owned systems located in areas with higher proportions of housing stock built prior to 1960 are 19.13 percentage points more likely to have a violation, which is rather consistent with the result in Table A.11 when excluding the fiscal capacity variables. In addition, higher proportions of housing stock built in the 1970s and 1980s are correlated with respectively 9.41 and 13.08 percentage points greater likelihood of noncompliance by local government owned systems. For those community water systems that are noncompliant, a greater proportion of housing stock built in the 1970s is also correlated with about five greater violations. However, neither of the variables measuring housing stock in the 1970s reach statistical significance in Table A.11 when fiscal variables are excluded from the analysis. Still, the findings correlating community water systems' housing stock before 1960 and during the 1980s to noncompliance are more pronounced than those in the analysis of all systems, which is expected considering local governments' high reliance on property tax revenue to finance service provision.

In terms of fiscal capacity, Table 2.8 shows that only the water revenue expenditure ratio and full-time employees are correlated with the likelihood of noncompliance and only at the 90 percent confidence level, although this latter finding is not robust when controlling for government (county or township) ownership (Table A.13). Debt service burden is correlated with the number of violations for noncompliant community water systems but also only at the 90 percent confidence level, and the correlation not robust to the exclusion of multijurisdictional systems (Table A.12).[11]

More closely aligned with the very limited extant research that does incorporate local government financial variables to examine Safe Drinking Water Act violations (Hughes, 2022), the more robust findings related to fiscal capacity in this analysis are the risk exposure factor and the water revenue expenditure ratio. For noncompliant local government owned systems, a one standard deviation increase in the risk exposure factor (calculated as intergovernmental revenue divided by property tax revenue) is correlated

[11] It should be noted that these specific government finance variables, to the authors' knowledge, have not been tested in prior research, so the lack of robust statistical significance is still quite informative for determining which types of potential financial assistance for local government owned systems might be most influential at enhancing Safe Drinking Water Act compliance.

with 0.12 fewer violations.[12] Holding property tax revenue constant, a local government with a higher risk exposure factor is more reliant on state and/or federal intergovernmental revenue. This finding suggests that local governments in noncompliance may be using intergovernmental revenue to fund water system activities in such ways that help to reduce their Safe Drinking Water Act violations.

This finding is especially noteworthy when considered with the water revenue expenditure ratio. Among noncompliant local government owned community water systems, a one standard deviation increase in water utility revenue relative to water utility expenditure is correlated with nearly 0.07 greater violations. Again, this ratio measures the amount of revenue levied during the fiscal year to cover expenditures in the same fiscal year. So, this finding may be indicative of the practice of local governments levying additional revenue through their water billing and using it to support unrestricted operating (i.e., general fund) activities, ultimately deferring maintenance on their water systems that help them to remain compliant. However, there is only scarce anecdotal evidence of this financial management practice (Carroll et al., 2025), making supplemental data collection directly from utilities necessary to fully explain this finding.

2.6. Summary of Quantitative Findings

Nearly 36 percent, or 15,920 of the 44,481 community water systems in the U.S. had at least one violation in 2023 or 2024, and nearly 28 percent had a history of noncompliance with at least one violation in the five years prior (2018-2022). All Systems in 2023-2024 had a mean of three violations while noncompliant systems had a mean of nine violations, with the majority of these being non-health based.

There is considerable variation in community water system noncompliance across states. All of the community water systems in Hawaii were compliant with the Safe Drinking Water Act in 2023 and 2024, while 71 percent of systems in West Virginia had at least one violation during these two years. In several other states, 50 percent or more of their community water systems were noncompliant in 2023-2024, potentially impacting more than 48.97 million people living there.

Over half of community water systems are local government owned with the remaining majority privately owned. More than 80 percent of community water systems self-produce (rather than purchase) their water, 77 percent rely on groundwater (as opposed to surface water), and only about 5 percent sell water wholesale. On-site inspections related to sanitary surveys is most common, as is statutorily required, with 35 percent of community water systems receiving one in 2023 or 2024. Yet, only about one percent of systems required a formal enforcement action for observed noncompliance.

Local government owned community water systems have very similar compliance rates and numbers of violations compared to all systems. However, local government owned systems tend to have more service connections serving larger populations, a lower proportion of self-producers, greater reliance on surface water, and higher percentage of

[12] The financial ratio variables in Table 2.8 were standardized to allow for interpretation as the effect of a one standard deviation increase from the mean.

systems that sell water wholesale. In addition, a greater proportion of local government owned systems receive technical assistance visits as part of compliance oversight.

Local governments that own community water systems spend about six percent of their total revenue to cover interest costs on outstanding debt, would need to levy an additional $3 in property tax revenue to offset a $1 reduction in intergovernmental revenue, and dedicate 83 percent of total spending to cover operations. In addition, these local governments levy about twice the necessary revenue to fund their water utilities; however, evidence suggests some of this revenue is used to support unrestricted operating (i.e., general fund) activities.

Community water systems are more likely to have a violation and, when noncompliant, likely to have more violations if they have the following system-level characteristics: Native American ownership, self-production, and surface water reliance. Systems are also more likely to be found noncompliant if they had a federal inspection or state technical assistance visit. Wholesaler systems and those receiving regularly scheduled state visits are less likely to have violations and, when noncompliant, have fewer violations. Poorer reporting quality, which may reflect less primacy agency oversight is also associated with greater noncompliance. Finally, among noncompliant systems, having a prior violation during 2018–2022 and receiving a formal enforcement action are both correlated with a greater number of violations.

Community water systems operating in more densely populated areas and those with wealthier residents are less likely to have violations, partly due to economies of scale in drinking water provision. Among noncompliant systems, service area characteristics are generally not associated with the number of violations, except in areas with the oldest housing stock, potentially indicating broader disinvestment in these areas. These findings are more pronounced for local government owned systems likely because local governments heavily rely on the wealth of their residents and property taxation to generate revenue for public service provision. Other system-level and service area correlates of local government owned community water systems are largely consistent of those for all systems with some important differences. Specifically, noncompliant local government owned systems that are wholesalers and those receiving technical assistance visits have fewer violations compared to all noncompliant systems.

A novel contribution of this study was the inclusion of fiscal capacity variables to assess noncompliance among local government owned systems. In terms of fiscal capacity, noncompliant local governments may be using intergovernmental revenue to fund water system activities in such ways that help to reduce violations, while some may be levying additional revenue through their water billing and using it to support unrestricted operating (i.e., general fund) activities, ultimately deferring maintenance on their water systems that help them to remain compliant.

3. Data Collection Tools for Future Noncompliance Assessment

While the quantitative regression analysis provides essential insight about the system-level and service area characteristics influencing noncompliance with the Safe Drinking Water Act, some factors cannot be captured by publicly available datasets. For example, operator training or managerial challenges might not be accounted for. In turn, there is a need to directly engage with utility staff to fully understand all factors that lead to noncompliance and how they might be interrelated. A complementary research approach is needed that can be used to engage with utilities across the U.S. and collect both qualitative and quantitative data through two data collection tools: 1) a survey, and 2) a semi-structured interview guide. These tools were developed using the outputs from the quantitative findings, which helped to 1) inform the sampling approach for the survey and interviews, and 2) identify key constructs and hypothesized drivers to be measured in the instruments.

3.1. Research Design

To develop the primary data collection tools, this study employed a mixed-methods research design that combines 1) a survey instrument to capture breadth, through comparable responses across many utilities, and 2) a semi-structured interview guide focused on depth to capture narratives, contextual nuance, and perspectives on systemic interactions. Together, these tools are designed to support a socio-technical systems analysis (Baxter & Sommerville, 2011; Ottens et al., 2006) that identifies the factors driving Safe Drinking Water Act noncompliance among water utilities. Both instruments are organized around technical, managerial, and financial capacity (U.S. Environmental Protection Agency, 2012), and include open-ended questions that can capture institutional knowledge on how issues originate within a utility (endogenous factors) and how pressures from outside the utility (exogenous factors—such as regulatory, financial, or environmental conditions) may interact to create barriers to compliance.

Data collected through the survey and interviews can be analyzed to reveal overall trends and relationships between factors. Tools like causal-loop diagramming and cognitive mapping (Gray et al., 2014; Meadows, 2008) can be used to develop systems maps that capture how factors interact and cascade to lead to noncompliance. This approach provides a system-level perspective, offering a deeper understanding of the socio-technical interactions that underpin water utility noncompliance and the connections between utilities and the communities they serve.

3.2. Survey Instrument and Interview Guide Development

An analysis of publicly available data was conducted to identify potential risk factors for water utility noncompliance. Based on this analysis, a survey and interview guide were developed to collect contextual and operational data from utilities and, in turn, increase understanding of the factors correlated with noncompliance. While any employee of a U.S. community water system can complete the survey, it is designed for managerial staff, recognizing that not all participants may have expertise on every topic.

These instruments were developed using the U.S. Environmental Protection Agency's Technical, Managerial, and Financial Capacity Framework, introduced in the 1996 amendments to the Safe Drinking Water Act. The Framework has been widely applied in

the literature to assess water system capacity (El-Khattabi et al., 2024) and is considered the gold standard for evaluating utility capacity by the U.S. Environmental Protection Agency (U.S. Environmental Protection Agency, 2012). Within the Framework, technical capacity refers to topics such as source water adequacy, infrastructure, treatment, distribution, storage, technical knowledge, and implementation. Ownership accountability, staffing and organization, and effective external linkages fall under managerial capacity, while revenue sufficiency, credit worthiness, and fiscal management are all addressed with financial capacity. These elements make up the foundation for short- and long-term planning at a utility, with the expectation that proficiency in these metrics will contribute to compliance and improve public health outcomes.

Part of the survey and interview guide was constructed based on the U.S. Environmental Protection Agency resource *Assessing Water System Managerial Capacity*, released in March 2012. This resource includes metrics on the following topics: governing body transparency, accountability, and training; water system staff training; water system planning; asset management programs; budgeting; rates; reserve accounts; water system policy compliance; water loss; and customer education and support. Each metric addresses at least one aspect of technical, managerial, or financial capacity and falls under the broader topics of ownership accountability, staffing and organization, and effective external linkages.

Beyond this framework, the survey and interview guide each touch on factors that contributed to noncompliance and the success of certain actions to help remedy violations when they do occur. The survey and interview guide were developed using prior literature to inform the overarching structure, topics addressed, and the way in which questions were framed. While previous research has not included survey or interview guides specifically focused on water utility noncompliance (to the authors knowledge), studies on related drinking water topics informed the overall design. For example, Palmer et al. (2025) employed a survey on public perceptions of drinking water, while LaPatin et al. (2024) used a semi-structured interview guide to interview water sector experts in rural Alaska. Such studies collectively informed the survey and interview guide.

The content of the survey is summarized in Table 3.1, and the interview guide is summarized in Table 3.2. Both the survey and interview guide include introductions to provide participants with information about the study, what to expect, and human-subjects research information. More details about consent and using these tools are in Section 3.3.

The semi-structured interview guide and survey were designed to be complementarity to each other, specifically using interviews to deepen the understanding of survey questions (Van Thiel, 2014). Both instruments are structured similarly with sections on 1) utility and participant background, 2) technical capacity, 3) managerial capacity, 4) financial capacity, and 5) regulatory compliance. The survey has an additional section on challenges to ensuring compliance. Both the interview guide and survey conclude by asking participants if there is anything else they would like to share about their utilities' experiences with compliance or system performance.

Table 3.1. Survey Content Overview

Section	Topics Covered	Survey Questions
Utility Information	Customer base	4
	System size/complexity	10–May
Participant Information	Role/activities	11–13, 17
	Experience/certifications	14–16
Technical Capacity	Treatment strategy	18–20
	Maintenance	21–26
	Current operation vs capacity	27–28
Managerial Capacity	Staff	29–31, 37, 39
	Governing body	32–36
	Communication strategies	38, 68
	Training	40–42
	Operations and planning programs	43–66
	General policies	67
	Data	69
	Organizational structure/ partnerships	70–71
Financial Capacity	Revenue, assets, and liabilities	72–78
	Billing rates/structure	79–83
	External financial support	84
	Financial health	85
Challenges	Source water concerns	86
	Technical challenges	87–88, 93
	Non-revenue water	89–92
	Workforce planning and optimization	94–95
	Water utility management	96
	Customer complaints	97–100
	Receivership	101
	Capital planning	102
	Bill delinquency and rate increases	104–105
	Funding struggles	106
Regulatory Noncompliance	Past noncompliance	108–113
	Efforts to maintain compliance	114,116
	Compliance support	115, 117–118

Note: This table provides a summary of the various sections, topics covered within each section, and survey question numbers pertaining to the topics covered for the survey.

In each section, beginning with technical capacity, the questions were informed by prior research. For example, treatment techniques have been shown to relate to noncompliance, with variables such as coagulant dosing, backwash frequency, chlorine dosing, and storage and pumping levels each linked to changes in the probability of a system experiencing a violation (Pike, 2004). Technology use can also impact compliance, although availability varies. For example, prior research has found that a variety of environmental and economic factors can impact technology use, adoption, and sustainability at water utilities (Wells et al., 2022). These findings motivated the inclusion of questions on technical capacity and treatment techniques.

Table 3.2. Interview Guide Content Overview

Section	Topics Covered
Utility and Participant Background	Role/activities Experience Source Treatment strategy
Technical Capacity	Technical challenges Maintenance Source water quality Current operation vs capacity
Managerial Capacity	Gaps in data and knowledge of assets Asset management Operational challenges Receivership Standard Operating Procedures (SOPs) Capital improvement planning Workforce management Sustainability and resilience programs Communication strategies Organizational structure/ partnerships Emergency planning
Financial Capacity	Financial challenges Financial health Capital improvement budgeting Long term financial planning Billing rates/structure and assistance
Regulatory Noncompliance	Past noncompliance Reasons for maintaining/falling out of compliance Efforts to maintain compliance
Conclusion	Additional thoughts

Note: This table provides a summary of the various sections and topics covered within each section for the in-depth interview guide.

In each section, beginning with technical capacity, the questions were informed by prior research. For example, treatment techniques have been shown to relate to noncompliance, with variables such as coagulant dosing, backwash frequency, chlorine dosing, and storage and pumping levels each linked to changes in the probability of a system experiencing a violation (Pike, 2004). Technology use can also impact compliance, although availability varies. For example, prior research has found that a variety of environmental and economic factors can impact technology use, adoption, and sustainability at water utilities (Wells et al., 2022). These findings motivated the inclusion of questions on technical capacity and treatment techniques.

In the managerial capacity section, the survey includes questions on community partnership, training for staff, and community engagement in emergency drinking water plans, reflecting the socio-technical dimensions of water management. In situations where systems are struggling and there is risk to public health, community-based organizations can play an important role in increasing health outcomes, indicating their importance in water utility operations (Wallerstein et al., 2019). Many of the questions in the managerial capacity section also address operations and planning, which prior research has identified

as key components in ensuring drinking water infrastructure resilience. These activities include capital improvement planning, climate change planning, supply management, demand management, watershed protection, affordability, and communication, with larger systems generally achieving higher resilience scores across these categories (Friedman et al., 2024). For instance, a study in Saskatoon, Canada investigated successful risk communication from drinking water utilities. The authors found that, in addition to traditional channels, personal networking and social media each played a large role in successful communication, but that official sources were still crucial for alerting customers of disruptions (Bradford et al., 2017). These findings prompted the inclusion of a question on customer notification, as well as questions on customer engagement throughout the survey.

The financial capacity sections of the survey and interview guide each touch on financial challenges that utilities face. Previous effort has been made to increase community water system financial data availability, with researchers creating a Municipal Drinking Water Database that includes all community water systems serving over 8,000 people. This database includes variables such as water utility revenue, water utility total expenditure, water utility intergovernmental expenditure, water utility capital outlay, water utility current expenditure, and water utility construction (Hughes et al., 2023). To build upon this resource, the survey asks questions about revenues, expenditures, and debt to gather insights on the large number of utilities (over 50 percent of total) serving under 8,000 customers, which were not included in the dataset. The survey developed for this study also asks about securing funding, such as loans or grants, as access to these resources has been shown to reduce water pollution (Keiser et al., 2023). Previous studies have found that funding programs are often insufficient, and barriers exist in accessing environmental funds due to limited qualified personnel or insufficient capacity to complete applications for grants and loans (Marcillo & Krometis, 2019). Gathering more information on the experiences of individual utilities in securing funding can improve collective understanding of how to remove barriers to access necessary resources and, in turn, increase compliance.

After addressing technical, managerial, and financial capacity, the next section of the survey presents questions on overall challenges. This section begins with a question on source water quality concerns, informed by a study on the prioritization and risk ranking of regulated and unregulated chemicals in U.S. drinking water (Rosenblum et al., 2024). Rosenblum et al. (2024) identified lead as the highest-priority contaminant, with various types of PFAS also classified as high-priority contaminants. The remaining questions on challenges and overall compliance were largely sourced based on previous research that informed the noncompliance model, as discussed previously. Guided by this framework, the survey and interview guide aim to fill gaps in publicly available data to provide a more comprehensive understanding of drinking water operations and noncompliance risks.

Both the interview guide and survey conclude by asking participants if there is anything else they would like to share about their utilities' experiences with compliance or system performance. It is important to note that the survey instrument includes explicit question skip-logic to facilitate deployment of some or part of the survey through an electronic survey software platform prior to dissemination. The survey was created to cover a broad range of relevant topics, which has led to the inclusion of almost 120 total questions. This

instrument is not intended to be deployed in its full length but was kept as an expansive resource to give more flexibility in including particular topics that are of interest at the time of deployment.

3.3. Implementation Best Practices for Surveys and Interviews

As part of this project, a survey instrument and interview guide were designed to understand the factors driving Safe Drinking Water Act noncompliance at community water systems across the U.S. In this section, best practices for conducting surveys and interviews are outlined, including guidance on how to use the materials shown in Appendices B and C. Both of these instruments are designed to be collected at the community water system level. Within this unit of analysis, the target population of respondents and interviewees includes utility managers, supervisors, or personnel with direct knowledge of community water system operations and compliance history.

A mixed methods research design, as proposed here, requires consideration of the appropriate sequencing—either parallel or sequential. Parallel design is where quantitative and qualitative research are conducted separately and their findings are later compared. Sequential design involves two or more consecutive phases. Sequential designs are commonly classified as explanatory, when the quantitative phase precedes and informs the qualitative phase, or exploratory, where the qualitative phase comes first (Hendren et al., 2023; Mele & Belardinelli, 2019). For this project, the survey will most likely be administered first, using the results to recruit a broad sample for interviews (see Section 3.2.2 for more details).

3.2.1. Survey

It is important to note that the survey questions are intended to provide the U.S. Environmental Protection Agency with options for questions to ask, but this list of questions should be shortened before deployment. This will reduce survey fatigue—a common issue in online surveys (Savage & Waldman, 2008). After the survey is shortened, it should be deployed through an online platform with a mobile-friendly design, such as Qualtrics. As with any human-subjects research, necessary approvals should be completed before data collection (e.g., Information Collection Request). This includes creating a consent form and a process to obtain consent, such as a question in the survey.

Next, the sampling strategy must be selected. Common sampling strategies are shown in Figure 3.1. For the survey, the research team recommends following a stratified purposive sampling approach to ensure a representative sample of utilities. This may be done across geography and region, system size and complexity, type of system ownership, severity and frequency of noncompliance, among other factors deemed important for the research objective. For example, the percent of respondents from each state could be based on the number of utilities in that specific state. The goal is to construct a heterogeneous sample that represents different issues with noncompliance based on pre-specified criteria or strata (Flick, 2018).

Figure 3.1: Common Sampling Strategies for Surveys and Interviews

Note: This figure illustrates common sampling strategies used as best practices for conducting surveys and in-depth interviews, developed by the authors based on the literature cited within.

In terms of recruitment workflow, an initial invitation email should be sent from the U.S. Environmental Protection Agency Office of Enforcement and Compliance Assurance, the research team, or both, with clear branding. This email should include a brief description of purpose, confidentiality, and approximate completion time for the survey. There should also be a clear distinction between Safe Drinking Water Act enforcement and the research objectives. Surveys also necessitate a reminder strategy, typically by scheduling one or two reminders for non-respondents that consist of short, respectful messages emphasizing voluntary participation and the value of their insight.

The survey should be open for a set amount of time, such as two to three months. Once closed, the data should be aggregated from the online platform and checked for quality. This might include removing responses that are incomplete or that were completed in a very short amount of time. After the data are cleaned, the U.S. Environmental Protection Agency can use descriptive statistics or statistical modeling to link survey responses to compliance outcomes. Survey data can be linked to outcomes based on utility name/location (Q2).

3.2.2. Semi-Structured Interviews

The interview guide was designed to be used for semi-structured interviews, meaning the questions are used to guide conversations, but there is flexibility to expand the conversation based on emergent ideas. As with the survey, it is imperative to make sure all human-subjects research approvals are in place and a consent form is developed.

The appropriate sampling strategy will also need to be determined for interviews—see Section 3.2.1. for more details. This decision is based on the desired generalizability of the results. One way to recruit interview participants is to include a supplemental question in the survey asking if the participant is willing to also participate in an interview. Such a convenience sampling approach can allow for more streamlined participant recruitment. Interviews should be conducted until no new information emerges in subsequent interviews, known as theoretical saturation (Strauss, 1998). Prior to starting interviews, the interviewer should conduct at least two practice interviews.

When reaching out to participants, it is important to include a clear explanation of the interview purpose, expected time duration, and confidentiality. The interview process should begin with rapport-building introductions and explanation of roles (i.e., research team vs. the U.S. Environmental Protection Agency). The interviewer should provide interviewees with a plain-language description of the purpose of the study and the voluntary nature of the interview. It is essential to review informed consent, recording permission, and confidentiality. It is also recommended to make explicit statements of the right to decline answering any question or to stop the interview at any time. An example of this language is included at the start of the interview guide.

Interviews should be recorded, when possible, ensuring that procedures from approved research ethics applications are followed. Although many researchers have noted best practices in conducting interviews, best practices are summarized based on work by Mack and Woodsong (2005) below:

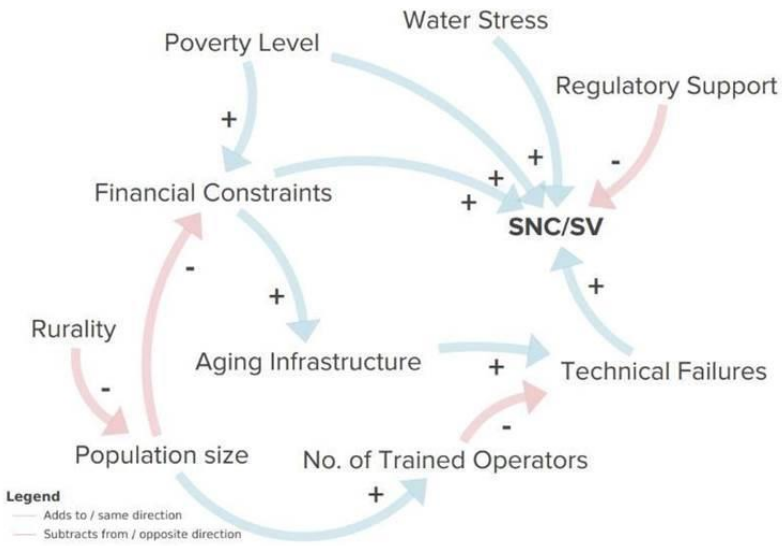
1. Build rapport and trust with a warm, respectful tone. Acknowledge the complexity and pressures utilities face. Start with easier, descriptive questions before moving into sensitive topics.
2. Maintain neutrality and a non-judgmental stance. Avoid leading questions or language implying blame. Use phrases like "What factors contributed…?" Allow respondents to frame issues in their own words.
3. Employ effective probing by using open-ended probes (e.g., "Can you tell me more about…?", "What happened next?"). Probe for connections and mechanisms (e.g., "How did that____affect day-to-day operations?").
4. When handling sensitive content, it is important to recognize that noncompliance events can be stressful and reputationally sensitive. It may be necessary to offer breaks and to remind participants they can skip questions if needed. Avoid sharing personal judgments and focus on understanding experiences.
5. Manage bias and reliability by being aware of social desirability (e.g., tendency to downplay internal shortcomings). Encourage discussion of both successes and failures. Ask for concrete examples rather than general statements.
6. Conduct post-interview debriefing immediately afterwards, by having the interviewer produce notes on main themes, surprises, and systems relationships. Flag emerging insights for iterative refinement of the interview guide or for analysis planning.

If the interviewee did not already complete the survey, they should be asked to do so. This allows for linking between the survey and interview data. After completion, the interview should be transcribed. If software is used, the transcript should be reviewed for quality control. Next, the data should be processed in line with social science research procedures to ensure bias is reduced and findings can be translated. The analysis should be validated based on the analysis technique selected. See Spearing et al., 2022 for more details about analysis and validation approaches.

To create cognitive systems maps (Gray et al., 2014; Meadows, 2008) that outline how factors are interrelated and lead to noncompliance, the researchers can qualitatively code

for connections (i.e., relationships). This means that each time an interviewee mentions a relationship, it is captured through the analysis. Examples of such approaches can be found in the research team's prior work (Spearing et al., 2023; Tiedmann et al., 2023). An example of such a cognitive systems map is shown in Figure 3.2. Note that this figure is conceptual and not based on actual results or empirical data.

Figure 3.2. Example Systems Map



Note: This figure is conceptual and not based on empirical data and is provided for illustrative purposes only.

4. Recommendations and Conclusions

4.1. Evidence-Based Recommendations

The preliminary research findings presented throughout section two of this report are based on extensive quantitative analysis. The qualitative data collection and analysis to be conducted in phase two of this project is crucial to increasing understanding of water system personnel's perspectives on challenges to compliance. In the absence of qualitative findings, the section two results serve as the basis for these preliminary recommendations for enhancing Safe Drinking Water Act compliance among community water systems:

Increase the monitoring capacity of community water systems that are persistent violators or at greater risk of noncompliance. This study's findings suggest there are barriers hindering systems from meeting monitoring requirements. This issue is fairly concentrated among 36 percent of systems and has clear geographic patterns (71% of West Virginia systems experienced noncompliance in 2023 or 2024, Native American owned systems are also more likely to be persistent violators, etc.). Phase two of this project would shed light as to what these barriers are, e.g., lack of technical capacity or low staffing levels, and can help finetune broad recommendations. Nonetheless, broadly, the U.S. Environmental Protection Agency can (1) review the current reporting process seeking to reduce administrative burden while also exploring options to increase collected information that could shed light on additional drivers of non-compliance, e.g., auto-validation to catch missing fields, (2) leverage federal funding programs, such as the Drinking Water State Revolving Fund, to encourage investment in systems identified as at-risk based on the quantitative findings, (3) increase outreach about regional cooperatives such as the consolidation incentivized by SDWA 1414(h) to increase uptake in these programs and/or (4) provide monitoring related training for at-risk systems that can take the form of short courses or visits to exemplary compliant systems with similar characteristics (e.g., similar water source or population served). Such training should be informed by future results from the proposed interview guide and survey instrument.

Adopt inspection selection criteria and couple Federal inspections of community water systems with technical assistance to achieve Safe Drinking Water Act compliance. Community water systems receiving a federal inspection in 2022 had a higher likelihood of noncompliance in 2023 or 2024, suggesting that federal inspections conducted by the U.S. Environmental Protection Agency are likely focused on systems already at risk of Safe Drinking Water Act violations. The U.S. Environmental Protection Agency could design a decision-support tool that predicts systems at higher risk of future compliance using a machine-learning model that builds on the findings in this study where self-producer status, surface water dependence, the proportion of unreported system characteristics, having a history of violations, and receiving a formal enforcement action were predictive of noncompliance. Federal inspections could be coupled with educational or tailored advisory measures for how struggling community water systems might regain and maintain Safe Drinking Water Act compliance. These materials should be co-designed with at-risk community water systems to ensure they meet their needs.

Implement additional scheduled inspections in addition to sanitary survey visits for persistent violators. Among the subset of noncompliant systems, having a regularly scheduled state site visit was correlated with fewer violations, suggesting that regular state visits might provide motivation to remain compliant. Additional visits on top of regularly scheduled sanitary survey visits could incorporate educational components that help utilities respond to any underlying issues that are identified in the visit (e.g., administrative suggestions to reduce monitoring and reporting violations). This might help provide proactive support to utilities. As such, for the systems that are persistent violators, Federal oversight is likely to be more effective if conducted more regularly compared to utilities that are regularly compliant.

Offer guidelines on compliance-supporting reserves to local governments that own and operate drinking water systems. Specifically, the findings suggest that intergovernmental revenue is likely an important contributor to local government fiscal capacity and can be leveraged to reduce Safe Drinking Water Act violations. However, among noncompliant local government systems, greater water utility revenue relative to water utility expenditure is associated with greater violations. Again, this ratio measures the amount of revenue levied during the fiscal year to cover expenditures in the same fiscal year. So, this finding may be indicative of the practice of local governments levying additional revenue through their water billing and using it to support unrestricted operating (i.e., general fund) activities, ultimately deferring maintenance on their water systems that help them to remain compliant. However, there is only scarce anecdotal evidence of this financial management practice (Carroll et al., 2025). While this practice is not recommended as part of sound financial management for local governments, studying this practice broadly is very difficult, because such transfers only appear on a local government's Annual Comprehensive Financial Reports, and there is no publicly available data source with digital access to these documents. Further, reviewing these documents would still prove difficult to determine necessary versus extraneous interfund transfers. As such, supplemental data collection directly from utilities is necessary to fully explain this finding.

4.2. Conclusions and Next Steps

Water utilities throughout the U.S. are facing challenges such as aging infrastructure and workforce limitations. Such challenges have led to noncompliance with the Safe Drinking Water Act. It is crucial to identify the factors that contribute to noncompliance in order to direct funding and interventions and, in turn, increase compliance. This study used system-level and service area characteristics, combined with local government finance data, to estimate both the probability of a Safe Drinking Water Act violation and the number of violations. The regression analysis showed that community water systems with greater noncompliance in 2023 or 2024 were correlated with Native American ownership, reliance on surface water, and self-production of water. Other correlates of noncompliance included serving an area with a higher proportion of pre-1960 housing relative to the proportion of 1990+ housing, a lower median household income, and a higher population density. Additionally, a subset of local government owned community water systems was analyzed separately to explore the role of government fiscal capacity. Results show that community water systems with higher risk exposure factors and lower water revenue expenditure ratios are correlated with a greater likelihood of noncompliance with the Safe

Drinking Water Act. Overall, the quantitative results help identify utilities that are at risk of noncompliance and inform actions that can be undertaken by the U.S. Environmental Protection Agency to achieve greater compliance, as outlined in the policy recommendations section.

Although results stemming from publicly available data are helpful to understand generalizable trends, it is not possible to capture contextual nuances in what contributes to noncompliance. In turn, it is imperative to capture perspectives of managers at community water systems so they can share unique operational and managerial experiences. To do so, a survey and semi-structured interview guide were developed to collect data to capture this essential insight. These tools can be used by the U.S. Environmental Protection Agency to collect data directly from utilities. In parallel with the survey and interview guide, this report includes guidance and recommendations to implement these tools. Those findings could then be used in combination with the quantitative results of this study to inform future policy initiatives focused on increased resources and interventions for noncompliant community water systems as well as operational guidance on emergency drinking water planning and other risk management strategies at the system level.

References

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. https://doi.org/10.1073/pnas.1719805115

Ashbolt, N. J. (2004). Microbial contamination of drinking water and disease outcomes in developing regions. *Toxicology, 198*(1–3), 229–238. https://doi.org/10.1016/j.tox.2004.01.030

Aviles, S. (2022). *How the age of housing impacts affordability*. https://storymaps.arcgis.com/stories/ae7f226a5ffd4466acbe0c7a14deab0e

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers, 23*(1), 4–17. https://doi.org/10.1016/j.intcom.2010.07.003

Bradford, L. E., Idowu, B., Zagozewski, R., & Bharadwaj, L. A. (2017). There is no publicity like word of mouth: Lessons for communicating drinking water risks in the urban setting. *Sustainable Cities and Society, 29*, 23–40. https://doi.org/10.1016/j.scs.2016.11.003

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (No. 53). Cambridge university press.

Carroll, D. A., Albrecht, K., Medwid, L., Khalaf, C., Michnick, J., Huang, D., Wetmore, B., & Li, J. (2023). *Water rate setting in the Lake Michigan service area*. Government Finance Research Center.

Carroll, D. A., Albrecht, K., Medwid, L., Khalaf, C., Huang, D., Senthilkumar, A., Zoh, D., Bonifant, J., & Avelar, D. (2024). *Water rate setting in northwestern, central, and southern Illinois*. Government Finance Research Center.

Carroll, D. A., DuBos, J.P., & Zaka, M. (2025) *Water Affordability Action Plan: Village of Richton Park, Illinois*. Chicago, IL: Government Finance Research Center, 2025.

Centers for Disease Control and Prevention. (2024). *Waterborne disease outbreaks associated with drinking water—United States, 2015–2020*. *Morbidity and Mortality Weekly Report*. https://www.cdc.gov/mmwr/volumes/73/ss/ss7301a1.htm

Centers for Disease Control and Prevention. (2024, April 23). *Drinking water sources: An overview*. https://www.cdc.gov/drinking-water/about/drinking-water-sources-an-overview.html

Demaris, A. (1992). *Logit modeling: Practical applications* (Vol. 86). Sage.

Effective Utility Management. (2024). *A primer for water sector utilities*. https://www.watereum.org/wp-content/uploads/2025/01/FINAL-2024-EUM-PRIMER.pdf

Elbakidze, L., & Beeson, Q. (2021). State regulatory heterogeneity and compliance with the Clean Water and Safe Drinking Water Acts. *Water Resources Research, 57*(4), e2020WR028980. https://doi.org/10.1029/2020WR028980

El-Khattabi, A. R., Mullins, E., & Obenshain, K. (2024). Assessing the viability of water utilities in North Carolina: A principal component analysis approach. *PLOS Water, 3*(8), e0000179. https://doi.org/10.1371/journal.pwat.0000179

Fawell, J., & Nieuwenhuijsen, M. J. (2003). Contaminants in drinking water: Environmental pollution and health. *British Medical Bulletin, 68*(1), 199–208. https://doi.org/10.1093/bmb/ldg027

Flick, U. (2018). *The SAGE handbook of qualitative data collection*. SAGE Publications. https://doi.org/10.4135/9781526416070

Friedman, M. B., Hughes, S., Kirchhoff, C. J., Rauh, E., McOmber, C., Manshardt, D. J., & Prout, J. M. (2024). Broadening resilience: An evaluation of policy and planning for drinking water resilience in 100 U.S. cities. *Global Environmental Change, 84*, 102798. https://doi.org/10.1016/j.gloenvcha.2023.102798

Fu, G., Liu, P., & Swallow, S. (2020). Effectiveness of public versus private ownership: Violations of the Safe Drinking Water Act. *Agricultural and Resource Economics Review, 49*(1), 1–26. https://doi.org/10.1017/age.2019.15

Government Accountability Office. (2018). *Water and wastewater workforce*. https://www.gao.gov/assets/gao-18-102.pdf

Gray, S. A., Zanre, E., & Gray, S. R. J. (2014). Fuzzy cognitive maps as representations of mental models and group beliefs. In E. I. Papageorgiou (Ed.), *Fuzzy cognitive maps for applied sciences and engineering* (pp. 29–48). Springer. https://doi.org/10.1007/978-3-642-39739-4

Grooms, K. K. (2016). Does water quality improve when a Safe Drinking Water Act violation is issued? *The B.E. Journal of Economic Analysis & Policy, 16*(4), 1–31. https://doi.org/10.1515/bejeap-2015-0173

Hansen, K., & Mullin, M. (2022). Barriers to water infrastructure investment: Findings from a survey of U.S. local elected officials. *PLOS Water, 1*(8), e0000039. https://doi.org/10.1371/journal.pwat.0000039

Heberling, M. T., Price, J. I., Nietch, C. T., Elovitz, M., Smucker, N. J., & Schupp, D. A. (2022). Linking water quality to drinking water treatment costs using time series analysis. *Water Resources Research, 58*(4), e2021WR031257. https://doi.org/10.1029/2021WR031257

Hendren, K., Newcomer, K., Pandey, S. K., Smith, M., & Sumner, N. (2023). How qualitative research methods can be leveraged to strengthen mixed methods research. *Public Administration Review, 83*(3), 468–485. https://doi.org/10.1111/puar.13528

Hughes, S. (2022). A multidimensional approach to evaluating the vulnerability of drinking water systems. *Journal of Environmental Policy & Planning, 24*(2), 210–226. https://doi.org/10.1080/1523908X.2021.2012975

Hughes, S., Kirchhoff, C. J., Conedera, K., & Friedman, M. (2023). The municipal drinking water database. *PLOS Water, 2*(4), e0000081. https://doi.org/10.1371/journal.pwat.0000081

Josset, L., Allaire, M., Hayek, C., Rising, J., Thomas, C., & Lall, U. (2019). The U.S. water data gap—A survey of state-level water data platforms. *Earth's Future, 7*(4), 433–449. https://doi.org/10.1029/2018EF001069

Keiser, D. A., Mazumder, B., Molitor, D., & Shapiro, J. S. (2023). *Water works: Causes and consequences of safe drinking water in America*. National Center for Environmental Economics.

Kirchhoff, C., Flagg, J., Zhuang, Y., & Utemuratov, B. (2019). Understanding and improving enforcement and compliance with drinking water standards. *Water Resources Management, 33*, 475–494. https://doi.org/10.1007/s11269-018-2137-4

Konisky, D. M., & Teodoro, M. P. (2016). When governments regulate governments. *American Journal of Political Science, 60*(3), 559–574. https://doi.org/10.1111/ajps.12207

LaPatin, M., Ritsch, N., Armanios, D. E., Albertson, L., Katz, L. E., & Faust, K. M. (2024). A stakeholder-systems analysis of water provision in rural Alaska. *ACS ES&T Water, 4*(2), 575–590. https://doi.org/10.1021/acsestwater.3c00606

Lee, M. J., Lee, G., & Choi, J. Y. (2025). Linear probability model revisited: Why it works and how it should be specified. *Sociological Methods & Research, 54*(1), 173–186.

Linder, K. D., & Brewer, T. F. (2022). Develop standard operating procedures for routine practice and what-if scenarios. *Opflow, 48*(5), 6–7.

Mack, N., & Woodsong, C. (2005). *Qualitative research methods: A data collector's field guide*. FHI 360.

Marcillo, C., & Krometis, L. (2019). Small towns, big challenges: Does rurality influence Safe Drinking Water Act compliance? *AWWA Water Science, 1*(2), e1121.

Meadows, D. (2008). *Thinking in systems: A primer* (D. Wright, Ed.). Earthscan.

Mele, V., & Belardinelli, P. (2019). Mixed methods in public administration research. *Journal of Public Administration Research and Theory, 29*(2), 334–347. https://doi.org/10.1093/jopart/muy046
Office of Water. (2004). *Understanding the Safe Drinking Water Act*. U.S. Environmental Protection Agency.
Office of Water. (2024). *Compliance evaluation of community water systems in the United States using data from the Safe Drinking Water Information System*. U.S. Environmental Protection Agency.
Ottens, M., Franssen, M., Kroes, P., & van de Poel, I. (2006). Modelling infrastructures as socio-technical systems. *International Journal of Critical Infrastructures, 2*(2–3), 133–145.
Oxenford, J. L., & Barrett, J. M. (2016). Understanding small water system violations and deficiencies. *Journal – American Water Works Association, 108*(3), 31–37.
Palmer, K., Berthold, T. A., McCrary, A., Devilleneuve, S., Schramm, M., & Leggette, H. R. (2025). Perceptions of drinking water and safety. *JAWRA Journal of the American Water Resources Association, 61*(5), e70059.
Pierce, G., Landes, L., Harrison, G., Schlichting, L., & Dunlap, L. (2023). *Considerations for a national drinking water quality compliance assessment*.
Pike, W. A. (2004). Modeling drinking water quality violations with Bayesian networks. *JAWRA Journal of the American Water Resources Association, 40*(6), 1563–1578.
Piper, S. (2003). Impact of water quality on municipal water price and residential water demand. *Water Resources Research, 39*(5), 1121.
Rahman, M. M., Ng, J. C., & Naidu, R. (2009). Chronic exposure of arsenic via drinking water and its adverse health impacts. *Environmental Geochemistry and Health, 31*, 189–200.
Rahman, T., Kohli, M., Megdal, S., Aradhyula, S., & Moxley, J. (2010). Determinants of environmental noncompliance by public water systems. *Contemporary Economic Policy, 28*(2), 264–279.
Ramseur, J. L., Normand, A. E., Cowan, T., Humphreys, E. H., Stern, C. V., Stubbs, M., & Lawhorn, J. M. (2020). *Federally supported projects and programs for wastewater, drinking water, and water supply infrastructure*. Congressional Research Service.
Reibel, M. (2018). *Estimating median household income for community water agencies: A case study of Los Angeles County*.
Rosenblum, J. S., Liethen, A., & Miller-Robbie, L. (2024). Prioritization and risk ranking of regulated and unregulated chemicals in U.S. drinking water. *Environmental Science & Technology, 58*(16), 6878–6889. https://doi.org/10.1021/acs.est.3c08792
Rubin, S. J. (2013). Evaluating violations of drinking water regulations. *Journal – American Water Works Association, 105*(3), E137–E147.
Savage, S., & Waldman, D. (2008). Learning and fatigue during choice experiments. *Journal of Applied Econometrics, 23*(3), 351–371.
Scanlon, B. R., Fakhreddine, S., Reedy, R. C., Yang, Q., & Malito, J. G. (2022). Drivers of spatiotemporal variability in drinking water quality in the United States. *Environmental Science & Technology, 56*(18), 12965–12974. https://doi.org/10.1021/acs.est.2c03178
Sowby, R. B. (2023). The Safe Drinking Water Act at 50. *Water Resources Research, 59*(9), e2023WR035172. https://doi.org/10.1029/2023WR035172
Spearing, L. A., & Faust, K. M. (2020). Cascading system impacts of the 2018 Camp Fire in California. *International Journal of Disaster Risk Reduction, 47*, 101822. https://doi.org/10.1016/j.ijdrr.2020.101822
Spearing, L. A., Thelemaque, N., Kaminsky, J. A., Katz, L. E., Kinney, K. A., Sela, L., & Faust, K. M. (2021). Implications of social distancing policies on drinking water infrastructure. *ACS ES&T Water, 1*(4), 888–899. https://doi.org/10.1021/acsestwater.0c00229

Spearing, L. A., Bakchan, A., Hamlet, L. C., Stephens, K. K., Kaminsky, J. A., & Faust, K. M. (2022). Comparing qualitative analysis techniques. *Journal of Construction Engineering and Management, 148*(7), 04022055. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002313

Spearing, L. A., Tariq, M., Safarpour, H., Abia, T., Mallory, M., Guild, J., Katz, L. E., & Faust, K. M. (2023). Moving beyond techno-economic considerations. *Current Opinion in Chemical Engineering, 42*, 100950. https://doi.org/10.1016/j.coche.2023.100950

Strauss, A. L. 1998. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: SAGE.

Switzer, D., & Teodoro, M. P. (2017). The color of drinking water. *Journal – American Water Works Association, 109*(9), 40–45.

Tariq, M., Spearing, L. A., Albertson, L., Katz, L. E., & Faust, K. M. (2025). Assessing the impact of the operational context of community water systems. *Journal of Water Resources Planning and Management, 151*(9), 04025036.

Tiedmann, H. R., Spearing, L. A., Castellanos, S., Stephens, K. K., Sela, L., & Faust, K. M. (2023). Tracking the post-disaster evolution of water infrastructure resilience. *Sustainable Cities and Society, 91*, 104417. https://doi.org/10.1016/j.scs.2023.104417

U.S. Environmental Protection Agency. (2006). *Distribution Systems: A Best Practices Guide*. U.S. Environmental Protection Agency. https://www.epa.gov/sites/default/files/2015-09/documents/epa816f06038.pdf

U.S. Environmental Protection Agency. (2008). *2006 drinking water data reliability analysis and action plan* (EPA 816-R-07-010).

U.S. Environmental Protection Agency. (2011). *Planning for an emergency drinking water supply* (EPA/600/R-11/054). Office of Research and Development. https://www.epa.gov/sites/default/files/2015-03/documents/planning_for_an_emergency_drinking_water_supply.pdf

U.S. Environmental Protection Agency. (2012). *Assessing water system managerial capacity* (EPA 816-K-12-004). U.S. Environmental Protection Agency. https://www.epa.gov/sites/default/files/2015-07/documents/epa816k12004.pdf

U.S. Environmental Protection Agency. (2017a). *Asset management for water and wastewater utilities*. EPA Sustainable Water Infrastructure. https://19january2017snapshot.epa.gov/sustainable-water-infrastructure/asset-management-water-and-wastewater-utilities_.html

U.S. Environmental Protection Agency. (2017b). *Operations and maintenance guidebook for small drinking water systems* (EPA 816-R-17-003). U.S. Environmental Protection Agency. https://www.epa.gov/sites/default/files/2018-03/documents/o_and_m_guidebook.pdf

U.S. Environmental Protection Agency. (2024). *Interagency Water Workforce Working Group report to Congress* (EPA 810-R-24-004). https://www.epa.gov/system/files/documents/2024-09/interagency-water-workforce-working-group-report-to-congress_august-2024-508-compliant.pdf

U.S. Environmental Protection Agency. (2025, June 5). *Water affordability resources for utilities*. Water Infrastructure and Resiliency Finance Center. https://www.epa.gov/waterfinancecenter/water-affordability-resources-utilities

U.S. Environmental Protection Agency. (2025a). *Summary of the Safe Drinking Water Act*. https://www.epa.gov/laws-regulations/summary-safe-drinking-water-act

U.S. Environmental Protection Agency. (2025b). *Providing safe drinking water in America: National public water systems compliance report*. https://www.epa.gov/compliance/providing-safe-drinking-water-america-national-public-water-systems-compliance-report

U.S. Environmental Protection Agency. (2025c). *Community water system service area boundaries*. https://www.epa.gov/ground-water-and-drinking-water/community-water-system-service-area-boundaries
U.S. Environmental Protection Agency. (2025d). *National primary drinking water regulations*. https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations
U.S. Environmental Protection Agency. (2025e). *National Enforcement and Compliance Initiative: Increasing compliance with drinking water standards at community water systems*. https://www.epa.gov/enforcement/national-enforcement-and-compliance-initiative-increasing-compliance-drinking-water
U.S. Environmental Protection Agency. (2025f). *How does your water system work?* https://www.epa.gov/ground-water-and-drinking-water/how-does-your-water-system-work-text-only
U.S. Environmental Protection Agency. (n.d.). *Safe Drinking Water Act (SDWA) resources and FAQs*. https://echo.epa.gov/help/sdwa-faqs
Van Thiel, S. (2014). *Research methods in public administration and public management* (1st ed.). Routledge. https://doi.org/10.4324/9780203078525
Wallerstein, N., Muhammad, M., Sanchez-Youngman, S., Rodriguez Espinosa, P., Avila, M., Baker, E. A., & Duran, B. (2019). Power dynamics in community-based participatory research. *Health Education & Behavior, 46*(1_suppl), 19S–32S. https://doi.org/10.1177/1090198119852998
Wallsten, S., & Kosec, K. (2008). The effects of ownership and benchmark competition. *International Journal of Industrial Organization, 26*(1), 186–205. https://doi.org/10.1016/j.ijindorg.2006.09.003
Wells, E. C., Vidmar, A. M., Webb, W. A., Ferguson, A. C., Verbyla, M. E., de los Reyes III, F. L., ... & Mihelcic, J. R. (2022). Meeting the water and sanitation challenges of underbounded communities in the U.S. *Environmental Science & Technology*, *56*(16), 11180–11188.
Young, H. A., Kolivras, K. N., Krometis, L. 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. https://doi.org/10.1016/j.envres.2022.114977

Appendix A: Safe Drinking Water Act Data and Robustness Checks

The dataset for quantitative analysis was constructed by combining the U.S. Environmental Protection Agency's 2025 Community Water System Service Area Boundaries with community water system characteristics from the U.S. Environmental Protection Agency's Safe Drinking Water Information System database for 2024 Q4 (Table A.1) and the U.S. Environmental Protection Agency's Safe Drinking Water Information System data on violations and enforcement (Table A.2).

Table A.1. Dataset Construction: Service Area

	Number of Observations
U.S. Environmental Protection Agency's 2025 Community Water System Service Area Boundaries	44,639
Exclude observations with missing PWSID[13]	44,637
Exclude observations from U.S. territories[14]	44,549
Remove inactive systems	44,503
Remove duplicates[15]	**44,481**

Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads.

Table A.2. Dataset Construction: Violations and Enforcement

	Number of Observations
U.S. Environmental Protection Agency Safe Drinking Water Act Violations and Enforcement[16]	14,665,006
Keep information for relevant community water system identified in Table A.1	5,729,246
Keep observations with a violation	5,409,090
Remove enforcement information and related duplicates[17]	4,076,439
Exclude violations that began prior to 2018	730,135
Exclude violations that began prior to 2023[18]	138,235
Remove detailed violation information and related duplicates[19]	24,405
Reshape data so each observation is a unique community water system[20]	**15,920**

Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads.

[13] Three systems were missing PWSIDs. The corresponding PWSID was manually determined for two of these systems (OR4100632 and NH0982010) using Echo's Detailed Facility Report. However, the NH system was still excluded, as it was not included in the U.S. Environmental Protection Agency's Safe Drinking Water Information System data, and the Echo report indicated it was inactive. The third system is in Puerto Rico and, even if it had a PWSID, it would have still been excluded at the next decision point.

[14] A total of 72 observations from Puerto Rico, 13 observations from the Northern Mariana Islands, and three observations from Guam were excluded. In addition, four observations were recoded, where the state was listed as ON, BC, and QC, to their respective states using the information under primacy agency. Similarly, for 69 observations with a missing state code, the information under primacy agency was used. Finally, for two systems with missing state code and primacy agency, the respective state was manually determined by checking Echo's Detailed Facility Report.

[15] The duplicates are a byproduct of service areas being comprised of multiple polygons. All polygons, except two, were included when calculating the spatial overlap variables. The two excluded polygons included one for Cypress Cove (FL3490314), mistakenly located in Miami. Echo's Detailed Facility Report notes that this system is in Kissimmee, which is reflected in the retained polygon. The other excluded polygon was modeled for the Burlington Water District (OR4100644). The polygon retained was the service area from the district.

[16] Systems appear in these data multiple times depending on how many violations and enforcement actions they have. The data include information about 256,220 unique systems.

[17] A count variable was calculated that captures the number of formal enforcement actions a system received in 2022 and then duplicates reflecting multiple enforcement actions per violation were removed.

[18] Count variables that capture the yearly number of violations from 2018 to 2024 were calculated, and violations prior to 2023 were dropped.

[19] Violation variables were calculated and duplicates reflecting detailed violation information were removed.

[20] community water system not included in the 15,920 systems with information from the Safe Drinking Water Act violations data are considered compliant in 2023 and 2024.

Characteristics of each community water system's service area were obtained from the U.S. Census Bureau's American Community Survey, 2018-2022 5-year estimates (Table A.3).[21] These data were collected at the census tract level, using a crosswalk to match census tracts to community water system service areas.

This process began by intersecting the 2025 community water system service area boundaries with 2022 U.S. census tract boundaries, both projected using the *USA Contiguous Albers Equal Area Conic* coordinate system and then calculating the overlap area in square meters followed by the percentage overlap relative to service area. The higher the number of overlapping areas, the more the rounding error compounds. In addition, the service area shapefile overlaps bodies of water while the census tract shapefile does not.

Adjusting for the rounding errors, water overlap, and minor shapefile mismatching was accomplished by rescaling the calculated proportions so that the sum of the adjusted proportions is equal to one. Using areal interpolation to estimate characteristics of water system service areas is a common approach used in prior research on this topic (Pierce et al., 2023; Reibel 2018).

The full dataset combines information from the 2025 U.S. Environmental Protection Agency's Community Water System Service Area Boundaries, Safe Drinking Water Act data for Q4 of 2024, and American Community Survey (ACS) 5-year estimates from 2018 to 2022.

Table A.3. Dataset Construction: Service Area Characteristics

	Number of Observations
Service areas and Census tracts data[22]	256,199
Keep information for relevant community water systems identified in Table A.1	253,316
Aggregate so each system tract match appears only once	170,502
Aggregate by creating areal weighted means of variables so each system appears only once	**44,481**

Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

For a subset of systems, those owned by local governments, data were collected to assess fiscal capacity, which were obtained from the U.S. Census Bureau's Census of Governments Finance Survey data for 2022 (Table A.4). These data are available at the municipality, township, and county level for every government in the U.S.[23]

[21] For additional information about the U.S. Census Bureau's American Community Survey data, please see https://www.census.gov/programs-surveys/acs.html.

[22] These data reflect the overlap between service area and census tract polygons, calculated in ArcGIS Pro. Before calculating the overlap, the service area shapefile geometry was repaired to address self-intersection issues.

[23] For additional information about the U.S. Census Bureau's Census of Governments Finance Survey data, please see: https://www.census.gov/programs-surveys/cog.html.

Fiscal information was matched to water systems by intersecting the 2025 community water system service area boundaries with the U.S. Census Bureau's 2022 Place and County Boundary files, both projected using the *USA Contiguous Albers Equal Area Conic* coordinate system. For systems that intersected with multiple geographies, the name of the system was used to finalize the matching. When system names were not conclusive, data were inspected on a case-by-case basis to ensure the appropriate one-to-one match.

For the 23,180 water systems included in the subset analysis, 60.88 percent were matched to municipal governments, 32.39 percent to county governments, 6.33 percent to township governments, with the remaining 0.41 percent denoted as multi-jurisdictional water systems. Although this latter category has access to the resources of multiple governments, the analysis uses financial information either from the governmental unit that has information available where not all units have available information, or the maximum values where more than one governmental unit has available financial information.

Robustness checks assess the sensitivity of the regression results to the (1) exclusion of multijurisdictional observations, and (2) control for government unit type.

Table A.4. Data Subset – Local Government Characteristics

	Number of Observations
Census of Governments Units	38,736
Merge Census of Government Employment Information	38,736
Merge service areas and municipal/township/county government crosswalk[24]	26,694
Merge Census of Government Financial Information	26,694
Remove duplicates[25]	**23,180**

Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads and U.S. Census Bureau's Census of Governments Finance Survey data for 2022 obtained from: https://www.census.gov/programs-surveys/cog.html.

[24] This crosswalk was created by first overlaying service area and municipal, township, and county polygons, in ArcGIS Pro. Before overlaying, the service area shapefile geometry was repaired to address self-intersection issues. When service area overlapped with multiple governmental units, the name of the system was used to finalize the matching. When system names were not conclusive, data were inspected on a case-by-case basis to ensure the appropriate one-to-one match.

[25] For the 23,180 water systems included in the subset analysis, 60.88 percent were matched to municipal governments, 32.39 percent to county governments, 6.33 percent to township governments, with the remaining 0.41 percent denoted as multi-jurisdictional water systems. Although this latter category has access to the resources of multiple governments, this analysis uses financial information either from the governmental unit that has information available where not all units have available information, or the maximum values where more than one governmental unit has available financial information.

Table A.5. Regression Results Underlying Figure 2.3

	Any Violation 2023-2024 All Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)
System-Level Characteristics		
Federal Government	0.0311	0.0106
	(0.1079)	
Local Government	0.0434	0.0148
	(0.0438)	
Public/Private	-0.1077	-0.0357
	(0.1075)	
Native American	0.7446***	0.2667
	(0.2462)	
State Government	0.1126	0.0389
	(0.0941)	
Service Connections (100k)	-0.2743***	-0.0938
	(0.0793)	
Self-Produced	0.1540***	0.0526
	(0.0430)	
Surface Water	0.1028**	0.0351
	(0.0471)	
Wholesaler	-0.8855***	-0.3026
	(0.2119)	
Self-Produced X Surface	0.0814	0.0278
	(0.0546)	
Self-Produced X Wholesaler	0.6994***	0.239
	(0.1674)	
Surface Water X Wholesaler	0.6605***	0.2257
	(0.1703)	
Self-Produced X Surface X Wholesaler	-0.6209***	-0.2122
	(0.1222)	
Regularly Scheduled Site Visits	0.0365	0.0125
	(0.0403)	
Technical Assistance Visits	0.0945	0.0323
	(0.1989)	
Sanitary Survey Visits	0.0085	0.0029
	(0.0173)	
State Inspections	0.0098	0.0033
	(0.0336)	
Federal Inspections	0.2424**	0.0828
	(0.1136)	
Proportion Unreported	2.2049***	0.7535
	(0.7527)	
Mean	0.3587	
Pseudo R2	0.0803	
Number of Observations	44,378	

Note: Results are from a probit model. Coefficients are presented in Column (1) while mean marginal effects are presented in Column (2). Two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data source: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads.

Table A.6. Robustness Check: Functional Form

	Probit		Logit		LPM
	Coefficient (Std. Error)	Marginal Effect	Coefficient (Std. Error)	Marginal Effect	Coefficient (Std. Error)
Variables	(1)	(2)	(3)	(4)	(5)
System-Level Characteristics					
Federal Government	-0.0251	0.0085	-0.0339	-0.0069	-0.00387
	(0.1097)		(0.1797)		(0.0359)
Local Government	-0.005	-0.0017	0.002	0.0004	0.0007
	(0.0435)		(0.0706)		(0.0144)
Public/Private	-0.1151	-0.0383	-0.1975	-0.0395	-0.0432
	(0.1052)		(0.1717)		(0.0398)
Native American	0.6577***	0.2347	1.1495***	0.2493	0.2350***
	(0.2431)		(0.4199)		(0.0865)
State Government	0.0851	0.0293	0.1491	0.0311	0.0311
	(0.0922)		(0.1516)		(0.0318)
Service Connections (100k)	-0.1294*	-0.0439	-0.2551	-0.0523	-0.0383**
	(0.0687)		(0.1558)		(0.0178)
Self-Produced	0.1721***	0.0584	0.2856***	0.0585	0.0573***
	(0.0362)		(0.0593)		(0.0124)
Surface Water	0.1764***	0.0599	0.2918***	0.0598	0.0579***
	(0.0429)		(0.0702)		(0.0148)
Wholesaler	-0.1554***	-0.0527	-0.2564***	-0.0526	-0.0501***
	(0.0406)		(0.0680)		(0.0135)
Regularly Scheduled Site Visits	0.0139	0.0047	0.0231	0.0047	0.0052
	(0.0403)		(0.0668)		(0.0150)
Technical Assistance Visits	0.0601	0.0204	0.0976	0.02	0.0176
	(0.2027)		(0.3520)		(0.0714)
Sanitary Survey Visits	0.0062	0.0021	0.0099	0.002	0.0020
	(0.0163)		(0.0271)		(0.0056)
State Inspections	0.0024	0.0008	0.0012	0.0002	0.0008
	(0.0315)		(0.0508)		(0.0111)
Federal Inspections	0.2980**	0.1011	0.4817**	0.0987	0.0933**
	(0.1157)		(0.1902)		(0.0387)
Proportion Unreported	2.1445***	0.7276	3.6638***	0.7511	0.744***
	(0.7712)		(1.2576)		(0.2680)
Service Area Characteristics					
Population Density (per km2)	-0.0002***	-0.0001	-0.0003***	-0.0001	-0.00005***
	(0.0001)		(0.0001)		(0.00002)
Median Household Income ($1,000)	-0.0025***	-0.0008	-0.0040***	-0.0008	-0.0008***
	(0.0007)		(0.0012)		(0.0002)
Median Age	-0.0021	-0.0007	-0.0035	-0.0007	-0.0006
	(0.0016)		(0.0027)		(0.0005)
Built Before 1960	0.4154***	0.1409	0.6874***	0.1409	0.1470***
	(0.1245)		(0.2065)		(0.0435)
Built 1960-1969	0.3522	0.1195	0.5942	0.1218	0.1310
	(0.2293)		(0.3729)		(0.0800)
Built 1970-1979	0.0224	0.0076	0.0363	0.0074	0.0133
	(0.1267)		(0.2127)		(0.0418)
Built 1980-1989	0.3591**	0.1218	0.5954**	0.1221	0.1250**
	(0.1479)		(0.2357)		(0.0504)
Mean	0.3587		0.3587		0.3579
Pseudo/Adjusted R2	0.0865		0.0866		0.1078
Number of Observations	44,378		44,378		44,481

Note: In all columns, state fixed effects are included. Clustered Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Table A.7. Robustness Check: Excluding Service Area Imputations

	Any Violation 2023-2024 All Systems		Number of Violations 2023-2024 Noncompliant Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Federal Government	0.0552	0.0189	-0.1072	-0.9156
	(0.1114)		(0.1448)	
Local Government	-0.0027	-0.0009	-0.0819	-0.7077
	(0.0435)		(0.0703)	
Public/Private	-0.1215	-0.0404	0.0998	0.9449
	(0.1054)		(0.1986)	
Native American	0.6439***	0.2298	0.3314*	3.5372
	(0.2368)		(0.1891)	
State Government	0.0809	0.0278	-0.1350	-1.1368
	(0.0962)		(0.1809)	
Service Connections (100k)	-0.2675	-0.0908	-0.0681	-0.592
	(0.2085)		(0.0818)	
Self-Produced	0.1733***	0.0588	0.5096***	4.4303
	(0.0358)		(0.0652)	
Surface Water	0.1835***	0.0623	0.2954***	2.5681
	(0.0397)		(0.0589)	
Wholesaler	-0.1576***	-0.0535	-0.2946***	-2.5613
	(0.0427)		(0.0908)	
Regularly Scheduled Site Visit	0.0068	0.0023	-0.2777***	-2.4141
	(0.0553)		(0.0903)	
Technical Assistance Visit	0.0698	0.0237	0.5312***	4.6178
	(0.2030)		(0.1677)	
Sanitary Survey Visit	0.0089	0.003	0.0470	0.4089
	(0.0159)		(0.0454)	
State Inspection	0.0295	0.01	0.0045	0.0394
	(0.0308)		(0.0331)	
Federal Inspection	0.3182***	0.108	0.1825	1.5866
	(0.1062)		(0.1715)	
Proportion Unreported Data	2.1958***	0.7455	2.4665***	21.4418
	(0.764)1		(0.9778)	
Any Violation 2018-2022		-0.0001	0.4080***	3.5473
			(0.0718)	
Formal Enforcement Action		-0.0008	0.6918***	6.0143
			(0.1607)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0007	0.0001	-0.0002
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	-0.0025***	0.1317	-0.0005	-0.0043
	(0.0007)		(0.0010)	
Median Age	-0.0022	0.1208	0.0013	0.0117
	(0.0016)		(0.0034)	
Built Before 1960	0.3880***	0.0097	0.1854	-1.6116
	(0.1273)		(0.1924)	
Built 1960-1969	0.3559	0.1188	0.6903*	6.0013
	(0.2316)		(0.4138)	
Built 1970-1979	0.0285	0.0076	0.2966	2.5788
	(0.1277)		(0.2623)	
Built 1980-1989	0.3499**	0.1218	0.2713	2.3586
	(0.1557)		(0.2439)	
Mean	0.3594		9	
Pseudo R2	0.0865		0.0449	
Number of Observations	43,329		15,572	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, one observation from DC and 98 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Table A.8. Robustness Check: Including Indicators for Service Area Imputations

	Any Violation 2023-2024 All Systems		Number of Violations 2023-2024 Noncompliant Systems	
	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Federal Government	-0.0296	-0.0100	-0.1113	-0.9486
	(0.1107)		(0.1367)	
Local Government	-0.0054	-0.0018	-0.0794	-0.6875
	(0.0435)		(0.0705)	
Public/Private	-0.1150	-0.0382	0.0948	0.8952
	(0.1053)		(0.1971)	
Native American	0.6574***	0.2346	0.3242*	3.4472
	(0.2447)		(0.1887)	
State Government	0.0725	0.0249	-0.1823	-1.5005
	(0.0876)		(0.1792)	
Service Connections (100k)	-0.1242**	-0.0422	-0.0681	-0.5921
	(0.0577)		(0.0754)	
Self-Produced	0.1712***	0.0581	0.5145***	4.4726
	(0.0360)		(0.0649)	
Surface Water	0.1753***	0.0595	0.2880***	2.5035
	(0.0422)		0.0581	
Wholesaler	-0.1556***	-0.0528	-0.3013***	-2.6189
	(0.0406)		(0.0869)	
Regularly Scheduled Site Visit	0.0137	0.0047	-0.2395**	-2.0822
	(0.0402)		(0.0986)	
Technical Assistance Visit	0.0595	0.0202	0.5169***	4.4929
	(0.2027)		(0.1643)	
Sanitary Survey Visit	0.0061	0.0021	0.0527	0.4585
	(0.0163)		(0.0450)	
State Inspection	0.0029	0.0010	0.0127	0.1101
	(0.0315)		(0.0339)	
Federal Inspection	0.2988***	0.1014	0.1782	1.5494
	(0.1158)		(0.1599)	
Proportion Unreported Data	2.1459***	0.7281	2.5078***	21.7997
	(0.7748)		(0.9793)	
Any Violation 2018-2022			0.4107***	3.5704
			(0.0712)	
Formal Enforcement Action			0.6930***	6.0237
			(0.1580)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0001	-0.00002	-0.0002
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	-0.0025***	-0.0008	-0.0005	-0.0045
	(0.0007)		(0.0010)	
Median Age	-0.0020	-0.0007	0.0013	0.0114
	(0.0016)		(0.0033)	
Built Before 1960	0.4170***	0.1415	-0.2075	-1.8037
	(0.1249)		(0.1866)	
Built 1960-1969	0.3522	0.1195	0.6692*	5.8175
	(0.2289)		(0.3910)	
Built 1970-1979	0.0247	0.0084	0.3285	2.8557
	(0.1271)		(0.2518)	
Built 1980-1989	0.3596**	0.1220	0.1874	1.6293
	(0.1497)		(0.2457)	
Mean	0.3587		9	
Pseudo R2	0.0866		0.0448	
Number of Observations	44,378		15,920	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included as well as indicator variables for Median Household Income, Median Age, and Proportion of Housing Stock, which flag observations for which missing service area values were imputed using the median of the observed data. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Table A.9. Robustness Check: Only 2024 Violations

	Any Violation 2024 All Systems		Number of Violations 2024 Noncompliant Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Federal Government	-0.1516	-0.0381	-0.2501	-1.1074
	(0.1417)		(0.2027)	
Local Government	-0.0464	-0.0121	-0.0865	-0.4146
	(0.0407)		(0.0626)	
Public/Private	-0.1266	-0.0321	0.1891	1.042
	(0.1081)		(0.2829)	
Native American	0.5548***	0.1726	0.1660	0.9039
	(0.2512)		(0.1597)	
State Government	-0.0123	-0.0032	0.0419	0.2139
	(0.1047)		(0.2726)	
Service Connections (100k)	-0.1692	-0.0442	-0.0345	-0.1658
	(0.1070)		(0.0638)	
Self-Produced	0.1580***	0.0412	0.4945***	2.3799
	(0.0341)		(0.0693)	
Surface Water	0.1629***	0.0425	0.3446***	1.6583
	(0.0489)		(0.0701)	
Wholesaler	-0.1611***	-0.0421	-0.3412***	-1.642
	(0.0398)		(0.0730)	
Regularly Scheduled Site Visit	-0.0014	-0.0004	-0.4143***	-1.994
	(0.0365)		(0.0787)	
Technical Assistance Visit	0.0116	0.003	0.4105***	1.9753
	(0.2034)		(0.1245)	
Sanitary Survey Visit	-0.0005	-0.0001	-0.0097	-0.0465
	(0.0185)		(0.0367)	
State Inspection	-0.0404*	-0.0105	-0.1304***	-0.6274
	(0.0233)		(0.0327)	
Federal Inspection	0.3481***	0.0909	0.1675	0.806
	(0.1361)		(0.2266)	
Proportion Unreported Data	1.9494***	0.5089	3.4453**	16.5804
	(0.7684)		(1.3550)	
Any Violation 2018-2022			0.2937***	1.4135
			(0.0573)	
Formal Enforcement Action			0.4390***	2.1127
			(0.1434)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.00004	-0.00004	-0.0002
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	-0.0018*	-0.0005	-0.0001	-0.0004
	(0.0009)		(0.0012)	
Median Age	-0.0008	-0.0002	-0.0023	-0.0112
	(0.0018)		(0.0027)	
Built Before 1960	0.3047**	0.0795	0.0574	0.2764
	(0.1388)		(0.1776)	
Built 1960-1969	0.4370*	0.1141	0.7159*	3.4455
	(0.2458)		(0.3781)	
Built 1970-1979	0.0302	0.0079	0.3102	1.4929
	(0.1701)		(0.2643)	
Built 1980-1989	0.3593***	0.0938	0.7821*	3.7637
	(0.1306)		(0.4061)	
Mean	0.2116		5	
Pseudo R2	0.0954		0.0488	
Number of Observations	44,380		9,391	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, 101 observations from Hawaii are excluded as none of these had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Table A.10. Robustness Check: Excluding Ownership

	Any Violation 2023-2024 All Systems		Number of Violations 2023-2024 Noncompliant Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Service Connections (100k)	-0.1333*	-0.0453	-0.0609	-0.5295
	(0.0718)		(0.0728)	
Self-Produced	0.1681***	0.0572	0.5246***	4.5636
	(0.0373)		(0.0643)	
Surface Water	0.1725***	0.0587	0.2740***	2.3831
	(0.0467)		(0.0580)	
Wholesaler	0.1545***	-0.0525	-0.3231***	-2.8102
	(0.0390)		(0.0928)	
Regularly Scheduled Site Visit	0.0055	0.0019	-0.2364**	-2.0562
	(0.0403)		(0.1105)	
Technical Assistance Visit	0.2049	0.0697	0.6167***	5.3641
	(0.2658)		(0.1790)	
Sanitary Survey Visit	0.0073	0.0025	0.0539	0.4692
	(0.0167)		(0.0459)	
State Inspection	-0.0033	-0.0011	0.0019	0.0162
	(0.0316)		(0.0383)	
Federal Inspection	0.3734***	0.1270	0.2290	1.992
	(0.1218)		(0.1603)	
Proportion Unreported Data	1.9544**	0.6646	2.5850***	22.4856
	(0.7845)		(0.9659)	
Any Violation 2018-2022			0.4173***	3.6301
			(0.0716)	
Formal Enforcement Action			0.6849***	5.9572
			(0.1582)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0001	0.0000	-0.0002
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	-0.0026***	-0.0009	-0.0006	-0.0049
	(0.0007)		(0.0010)	
Median Age	-0.0027	-0.0009	0.0014	0.0121
	(0.0017)		(0.0032)	
Built Before 1960	0.3955***	0.1345	-0.2984**	-2.596
	(0.1308)		(0.1613)	
Built 1960-1969	0.3576	0.1216	0.6626**	5.764
	(0.2311)		(0.3749)	
Built 1970-1979	0.0567	0.0193	0.3326	2.893
	(0.1300)		(0.2465)	
Built 1980-1989	0.3825*	0.1301	0.2246	1.9541
	(0.1507)		(0.2606)	
Mean	0.3587		9	
Pseudo R2	0.0844		0.0442	
Number of Observations	44,378		15,920	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, two observations from DC and 101 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Table A.11. Robustness Check: Excluding Government Finance Variables

	Any Violation 2023-2024 All Local Govt. Owned Systems		Number of Violations 2023-2024 Noncompliant Local Govt. Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Service Connections (100k)	-0.0750*	-0.0246	-0.0850	-0.7430
	(0.0449)		(0.0592)	
Self-Produced	0.1716***	0.0562	0.4726***	4.1292
	(0.0416)		(0.0673)	
Surface Water	0.2118***	0.0693	0.2154***	1.8818
	(0.0398)		(0.0520)	
Wholesaler	0.1438***	-0.0471	-0.1103**	-0.9636
	(0.0486)		(0.0549)	
Regularly Scheduled Site Visit	0.0727	0.0238	-0.2736*	-2.3906
	(0.0665)		(0.1407)	
Technical Assistance Visit	0.0271	0.0089	0.3843**	3.3577
	(0.0804)		(0.1637)	
Sanitary Survey Visit	0.0010	0.0003	0.0470	0.4108
	(0.0173)		(0.0355)	
State Inspection	-0.0373	-0.0122	0.1337	1.1686
	(0.0250)		(0.0838)	
Federal Inspection	0.3434***	0.1124	0.1667	1.4565
	(0.1271)		(0.2114)	
Proportion Unreported Data	1.1271*	0.3691	0.5245	4.5832
	(0.6591)		(0.9310)	
Any Violation 2018-2022			0.4083***	3.5679
			(0.0678)	
Formal Enforcement Action			0.4995***	4.3645
			(0.1435)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0001	-0.00004	-0.0003
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	-0.0037***	-0.0012	0.0004	0.0033
	(0.0007)		(0.0014)	
Median Age	-0.0017	-0.0006	-0.0050	-0.0441
	(0.0022)		(0.0033)	
Built Before 1960	0.6073***	0.1988	0.0165	0.1443
	(0.1652)		(0.2844)	
Built 1960-1969	0.4367	0.1430	0.8411	7.3496
	(0.3590)		(0.5684)	
Built 1970-1979	0.2801	0.0917	0.5781	5.0518
	(0.1729)		(0.3544)	
Built 1980-1989	0.3630**	0.1189	-0.1133	-0.9901
	(0.1749)		(0.4837)	
Mean	0.3525		9	
Pseudo R2	0.1121		0.0452	
Number of Observations	23,129		8,153	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, one observation from DC and 50 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, and U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html.

Table A.12. Robustness Check: Excluding Multi-Jurisdictional Systems

	Any Violation 2023-2024 All Local Govt. Owned Systems		Number of Violations 2023-2024 Noncompliant Local Govt. Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Service Connections (100k)	-0.0991	-0.0323	-0.0702	-0.6117
	(0.0612)		(0.0710)	
Self-Produced	0.1564***	0.0510	0.4764***	4.1490
	(0.0401)		(0.0677)	
Surface Water	0.1958***	0.0638	0.2382***	2.0751
	(0.0373)		(0.0518)	
Wholesaler	-0.1175**	-0.0383	-0.1421***	-1.2376
	(0.0565)		(0.0552)	
Regularly Scheduled Site Visit	0.0662	0.0216	-0.2839**	-2.4725
	(0.0550)		(0.1400)	
Technical Assistance Visit	0.0280	0.0091	0.4172**	3.6335
	(0.1026)		(0.1806)	
Sanitary Survey Visit	-0.0032	-0.0011	0.0444	0.3866
	(0.0194)		(0.0401)	
State Inspection	-0.0151	-0.0049	0.0571	0.4976
	(0.0320)		(0.0862)	
Federal Inspection	0.2760**	0.0899	0.2712	2.3624
	(0.1360)		(0.2212)	
Proportion Unreported Data	1.3600**	0.4432	0.4137	3.6033
	(0.6731)		(0.9862)	
Any Violation 2018-2022			0.4111***	3.5805
			(0.0678)	
Formal Enforcement Action			0.5107***	4.4479
			(0.1338)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0001	-0.00003	-0.0003
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	0.0034***	-0.0011	0.0009	0.0077
	(0.0007)		(0.0014)	
Median Age	-0.0019	-0.0006	-0.0047	-0.041
	(0.0021)		(0.0035)	
Built Before 1960	0.6147***	0.2003	-0.0275	-0.2391
	(0.1499)		(0.2900)	
Built 1960-1969	0.4988	0.1626	0.8468	7.3758
	(0.3132)		(0.5425)	
Built 1970-1979	0.3512***	0.1145	0.6718*	5.8514
	(0.1743)		(0.3729)	
Built 1980-1989	0.4373***	0.1425	-0.1584	-1.3794
	(0.1812)		(0.4728)	
Government Finance				
Debt Service Burden	-0.2289	-0.0746	-2.4649	-21.4689
	(0.2223)		(1.5605)	
Risk Exposure Factor	0.0096	0.0031	-0.0147***	-0.1281
	(0.0076)		(0.0006)	
Operating Expenditure Ratio	-0.0081	-0.0026	-0.0052	-0.0452
	(0.0107)		(0.0173)	
Water Revenue Expenditure Ratio	0.0174*	0.0057	0.0080***	0.0694
	(0.0099)		(0.0011)	
Full Time Employees (1,000)	-0.0131**	-0.0043	-0.0051	-0.0443
	(0.0056)		(0.0050)	
Mean	0.3522		9	
Pseudo R2	0.1162		0.0461	
Number of Observations	21,165		7,455	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, 50 observations from Hawaii are excluded as none of these systems had a violation. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html, and U.S. Census Bureau's Census of Governments Finance Survey data for 2022 obtained from: https://www.census.gov/programs-surveys/cog.html.

Table A.13. Robustness Check: Including Government Unit Type

	Any Violation 2023-2024 All Local Govt. Owned Systems		Number of Violations 2023-2024 Noncompliant Local Govt. Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Service Connections (100k)	-0.0646	-0.0211	-0.0607	-0.5301
	(0.0443)		(0.0623)	
Self-Produced	0.1511***	0.0494	0.4818***	4.2088
	(0.0402)		(0.0688)	
Surface Water	0.1985***	0.0648	0.2205***	1.9257
	(0.0373)		(0.0524)	
Wholesaler	-0.1272***	-0.0415	-0.1043*	-0.9112
	(0.0489)		(0.0538)	
Regularly Scheduled Site Visit	0.0845	0.0276	-0.2767*	-2.4168
	(0.0663)		(0.1442)	
Technical Assistance Visit	-0.0302	-0.0099	0.3798**	3.3178
	(0.0809)		(0.1661)	
Sanitary Survey Visit	-0.0006	-0.0002	0.0489	0.4268
	(0.0176)		(0.0359)	
State Inspection	-0.0266	-0.0087	0.1374*	1.2006
	(0.0268)		(0.0808)	
Federal Inspection	0.3246**	0.1060	0.1711	1.4943
	(0.1268)		(0.2101)	
Proportion Unreported Data	1.3769**	0.4497	0.4942	4.3166
	(0.6419)		(0.9517)	
Any Violation 2018-2022			0.4104***	3.5848
			(0.0678)	
Formal Enforcement Action			0.4933***	4.3088
			(0.1440)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0001	-0.00002	-0.0002
	(0.0001)		(0.0001)	
Median Household Income ($1,000)	-0.0035***	-0.0011	0.0004	0.0036
	(0.0007)		(0.0014)	
Median Age	-0.0011	-0.0003	-0.0056*	-0.0487
	(0.0022)		(0.0030)	
Built Before 1960	0.5069***	0.1655	0.0360	0.3148
	(0.1503)		(0.2918)	
Built 1960-1969	0.3823	0.1249	0.8428	7.3621
	(0.3031)		(0.5487)	
Built 1970-1979	0.2779	0.0908	0.5887*	5.1422
	(0.1706)		(0.3553)	
Built 1980-1989	0.3881**	0.1268	-0.1206	-1.0532
	(0.1778)		(0.4824)	
Government Finance				
Debt Service Burden	-0.3606	-0.1178	-1.9815**	-17.3088
	(0.4104)		(0.9773)	
Risk Exposure Factor	0.0091	0.0030	-0.0142***	-0.1236
	(0.0076)		(0.0006)	
Operating Expenditure Ratio	-0.0035	-0.0011	-0.0130	-0.1138
	(0.0108)		(0.0163)	
Water Revenue Expenditure Ratio	0.0163*	0.0053	0.0082***	0.0716
	(0.0093)		(0.0012)	
Full Time Employees (1,000)	-0.0075	-0.0024	-0.0071	-0.0619
	(0.0052)		(0.0060)	
County Government	-0.1025***	-0.0335	0.0515	0.4520
	(0.0321)		(0.0670)	
Township Government	-0.1502**	-0.0487	0.0878	0.7851
	(0.0637)		(0.0703)	
Multi Jurisdiction	-0.5794	-0.1697	-0.3708***	-2.6513
	(0.3737)		(0.1345)	
Mean	0.3525		9	
Pseudo R2	0.1144		0.0456	
Number of Observations	23,129		8,153	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, one observation from DC and 50 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html, and U.S. Census Bureau's Census of Governments Finance Survey data for 2022 obtained from: https://www.census.gov/programs-surveys/cog.html.

Table A.14. Robustness Check: Excluding Employment Imputations

	Any Violation 2023-2024 All Local Govt. Owned Systems		Number of Violations 2023-2024 Noncompliant Local Govt. Systems	
Variables	Coefficient (Std. Error) (1)	Marginal Effect (2)	Coefficient (Std. Error) (3)	Marginal Effect (4)
System-Level Characteristics				
Service Connections (100k)	-0.0578	-0.0189	-0.0647	-0.5642
	(0.0441)		(0.0616)	
Self-Produced	0.1658***	0.0542	0.4731***	4.1265
	(0.0400)		(0.0680)	
Surface Water	0.2014***	0.0658	0.2167***	1.8905
	(0.0379)		(0.0510)	
Wholesaler	-0.1153**	-0.0377	-0.1150**	-1.0032
	(0.0509)		(0.0552)	
Regularly Scheduled Site Visit	0.0769	0.0251	-0.2717*	-2.3698
	(0.0671)		(0.1446)	
Technical Assistance Visit	-0.0272	-0.0089	0.3722**	3.2461
	(0.0814)		(0.1675)	
Sanitary Survey Visit	0.0016	0.0005	0.0475	0.4140
	(0.0176)		(0.0361)	
State Inspection	-0.0337	-0.0110	0.1378*	1.202
	(0.0247)		(0.0771)	
Federal Inspection	0.3121**	0.1020	0.1882	1.6419
	(0.1273)		(0.2135)	
Proportion Unreported Data	1.1579*	0.3784	0.5256	4.5842
	(0.6389)		(0.9535)	
Any Violation 2018-2022			0.4034***	3.5181
			(0.0666)	
Formal Enforcement Action			0.4875***	4.2516
			(0.1445)	
Service Area Characteristics				
Population Density (per km2)	-0.0002***	-0.0001	-0.00003	-0.0003
	(0.0001)		(0.00006)	
Median Household Income ($1,000)	-0.0037***	-0.0012	0.0005	0.0041
	(0.0007)		(0.0014)	
Median Age	-0.0020	-0.0007	-0.0051	-0.0447
	(0.0022)		(0.0032)	
Built Before 1960	0.5733***	0.1874	0.0107	0.0931
	(0.1514)		(0.2822)	
Built 1960-1969	0.4086	0.1336	0.8512	7.4242
	(0.3059)		(0.5383)	
Built 1970-1979	0.2778	0.0908	0.5888*	5.1353
	(0.1743)		(0.3553)	
Built 1980-1989	0.4012**	0.1312	-0.1045	-0.9117
	(0.1776)		(0.4813)	
Government Finance				
Debt Service Burden	-0.3402	-0.1112	-2.0503*	-17.8828
	(0.4865)		(1.0863)	
Risk Exposure Factor	0.0097	0.0032	-0.0143***	-0.1248
	(0.0076)		(0.0006)	
Operating Expenditure Ratio	-0.0074	-0.0024	-0.0107	-0.0936
	(0.0105)		(0.0174)	
Water Revenue Expenditure Ratio	0.0173*	0.0057	0.0080***	0.0695
	(0.0097)		(0.0011)	
Full Time Employees (1,000)	-0.0104*	-0.0034	-0.0062	-0.0545
	(0.0056)		(0.0051)	
Mean	0.3520		9	
Pseudo R2	0.1134		0.0453	
Number of Observations	22,986		8,091	

Note: Columns (1)-(2) present results from a probit model while Columns (3)-(4) present results from a negative binomial model. Coefficients are presented in Column (1) and (3) while mean marginal effects are presented in Columns (2) and (4). In the probit model, one observation from DC and 50 from Hawaii are excluded as all DC systems had a violation and none from Hawaii had any violations. In all columns, state fixed effects are included. Clustered standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01. Data sources: U.S. Environmental Protection Agency, Enforcement and Compliance History Online datasets obtained from https://echo.epa.gov/tools/data-downloads#downloads, U.S. Census Bureau's American Community Survey 2018-2022 5-year estimates obtained from https://www.census.gov/programs-surveys/acs.html, and U.S. Census Bureau's Census of Governments Finance Survey data for 2022 obtained from: https://www.census.gov/programs-surveys/cog.html.

The Research Team

Christelle Khalaf, Ph.D.
Christelle Khalaf is an economist and the associate director of the Government Finance Research Center at the University of Illinois Chicago. Her research uses applied economic and data-driven methods to analyze regional economies, labor markets, and public systems. Khalaf is author or co-author of 25 technical reports and has been quoted in several news media outlets, including NPR, HuffPost, and Energy News Network. She has published peer-reviewed papers in Economic Development Quarterly, Utilities Policy, and Journal of Computational Social Science, among other journals. She holds a Ph.D. in Economics from North Carolina State University. Email: ckhalaf@uic.edu.

Deborah A. Carroll, Ph.D.
Deborah A. Carroll is Professor and Director of the Government Finance Research Center at the University of Illinois Chicago. Her research focuses on financial management and fiscal policy of state and local governments, particularly related to taxation, revenue diversification, and water pricing and affordability, as well as the interconnectedness of the public and nonprofit sectors and the implications for tax policy, tax burden, and public service provision. Dr. Carroll has produced more than 75 scholarly publications and 25 white papers and technical reports reflecting more than $2.75 million in research funding. She has been invited to share her research and expertise with academic and practitioner audiences across the U.S. and in several other countries. Dr. Carroll began her career in public service as a budget and policy analyst for the City of Milwaukee, Wisconsin. Email: deborahc@uic.edu.

Lauryn A. Spearing, Ph.D.
Lauryn Spearing is an Assistant Professor in the Civil and Environmental Engineering Department at Tufts University. She received her PhD in Civil Engineering from the University of Texas at Austin. Dr. Spearing uses a systems thinking approach, involving both qualitative and quantitative methods, to understand complex problems at the intersection of technical, social, and natural systems. Her research focuses on three broad areas: 1) infrastructure management during uncertain contexts; 2) understanding public perceptions towards the built environment; and 3) sustainable water systems. She has worked on various complex engineering problems such as managing infrastructure interdependencies in rural Alaska, understanding community impacts after the East Palestine, Ohio train derailment, and exploring how population dynamics impact the built environment. Email: lauryn.spearing@tufts.edu.

Abigail Iuorio
Abby Iuorio is a PhD student working in the Spearing Lab at Tufts University in the Civil and Environmental Engineering Department. She has a BS in Civil Engineering from the University of Illinois Urbana-Champaign and an MS in Environmental Engineering from Stanford University. Her research focuses on the intersection of built, natural, and social systems, primarily in water infrastructure. Email: Abigail.Iuorio@tufts.edu.

Joseph C. Toland
Joseph Toland is a fourth-year PhD candidate in Civil and Environmental Engineering at Tufts University. His research investigates human-infrastructure interactions in disasters, aiming to improve infrastructure management, strengthen public policy and emergency management practices. His work draws on more than 20 years of practical experience in emergency management with FEMA and DHS, natural hazards research with the USGS, and supporting disaster relief efforts with the Red Cross to steer his research towards practical solutions that strengthen community resilience. Email: Joseph.Toland@tufts.edu.