Testing the water: Drinking water quality, public notification, and child outcomes

Michelle Marcus*

November 16, 2020

Abstract

Health-based drinking water violations affect about 1 in 12 Americans annually, the benefits of drinking water regulation are not well understood. I exploit plausibly exogenous variation in water quality violation timing to estimate the impacts on avoidance behavior and child outcomes. Using purchases of bottled water and common stomach remedies, emergency room visits for gastrointestinal illness, and school absences, I provide a comprehensive calculation of costs associated with poor drinking water quality. Individuals avoid the negative health impacts of coliform bacteria violations only when informed immediately. Timely public notification is a cost-effective way to induce avoidance behavior and protect health.

JEL Codes: Q51, Q53, Q58, I14, I18, I24

Disasters like the water crisis in Flint, Michigan, have highlighted the need for careful monitoring of drinking water supplies in the United States. Estimates suggest that up to 19.5 million cases

*Vanderbilt University, Department of Economics, 415 Calhoun Hall, Nashville, TN 32740. E-mail: *michelle.marcus@vanderbilt.edu*. For their helpful comments, I would like to thank Anna Aizer, Brian Beach, Ryan Brown, Christopher Carpenter, Arnaud Chevalier, Janet Currie, Chloe East, Maria Fitzpatrick, Dennis Guignet, Kanghyock Koh, Angelica Meinhofer, Paulina Oliva, Sheila Olmstead, Emily Oster, Daniel Rees, Jay Shimshack, David Slusky, Will Wheeler, and participants at AERE, NBER Summer Institute, ASSA, SEA, and CEPR-EBRD-EoT-LSE Symposium. Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

of waterborne illnesses associated with drinking water occur each year in the U.S. (Reynolds et al., 2008; Colford et al., 2006). In 2015, there were 12,137 health-based drinking water violations at 5,009 community water systems serving about 27 million people across the country, or nearly 1 in 12 Americans. Furthermore, since disadvantaged children are more often exposed to water quality violations, poor water quality may contribute to existing socio-economic gaps in health and human capital accumulation.

Despite the large number of individuals potentially affected by these violations, relatively little is known about the benefits of drinking water regulation or whether behavioral responses to public information about violations can mitigate negative health effects. To understand the benefits of drinking water regulation, it is critical to quantify the costs of poor drinking water quality, which may include costs of avoidance behavior, human health consequences and lost productivity, and even long-run implications for human capital formation. Doing so is empirically challenging, however, owing to endogenous residential sorting. Parental income, as well as preferences for school quality, neighborhood amenities, and environmental quality, can generate bias in cross-sectional estimates, because many of these parental characteristics and preferences are likely related to underlying health and human capital accumulation.

In this paper, I take a comprehensive approach to studying the consequences of drinking water quality violations for children and adolescents in the US by considering the impacts on both avoidance behavior and several measures of health and human capital. Moreover, I consider whether the health and human capital benefits from reduced exposure to poor drinking water exceed the cost of avoidance behavior, measured by bottled water purchases. To overcome the empirical challenges in this context, I exploit plausibly exogenous variation in the timing of water quality violations to estimate the within-location impacts of poor water quality on avoidance behavior and child outcomes. Specifically, for North Carolina, I combine data on water quality violations and detailed geographic information on community water supply system service areas with several confidential administrative datasets that measure child outcomes. I control for detailed location fixed effects such that the identifying variation comes from changing water quality within a place over time. Data on bottled water purchases provide a measure of avoidance behavior, and data on emergency room (ER) visits for gastrointestinal illness provide a direct measure of health. Because ER visits represent relatively extreme health effects, I also consider the impact of drinking water violations on purchases of over-the-counter remedies for gastrointestinal illness and school absences, which may provide a more sensitive measure of health. School absences are also of interest since they may affect school performance, graduation rates, grade retention, later-life earnings, and the intergenerational transmission of human capital (Aucejo and Romano, 2016; Gottfried, 2011; Currie and Moretti, 2003; Lleras-Muney, 2005; Grossman and Kaestner, 1997). Absences may also reduce parental earnings if parents miss work to take care of a sick child, and may even reduce school funding when funding is determined by student attendance (Ely and Fermanich, 2013).

This study makes several important contributions to the literature. First, although there is growing evidence on the harmful effects of air pollution, research on the impacts of water pollution at modern levels has been much more limited (Keiser and Shapiro, 2019). Existing research has focused only on the impact of water quality on mortality and infant health, especially among infants born to less-educated mothers (Currie et al., 2013; Hill and Ma, 2017).[1] Much less is known about the effects of exposure to modern levels of water pollution during childhood or adolescence.

[1]Historically, the adoption of clean water technologies in the early 20th century sharply reduced typhoid mortality, which also affected human capital formation (Cutler and Miller, 2005; Beach et al., 2016; Anderson et al., 2018).

Exposure to pollution during this critical period of growth and development may harm short- and long-term health, as well as human capital formation and subsequent earnings. This paper focuses on children and adolescents, a new and sensitive population to study, and is the first to study several measures of health and human capital: ER visits, over-the-counter purchases of stomach remedies, and school absences. By studying these outcomes simultaneously, I provide the first comprehensive view of the impacts of modern-day water quality on child health and human capital accumulation and contribute to the full calculation of the benefits of water quality regulation.

Second, this paper also contributes to the literature on information and transparency (Weil et al., 2006; Loewenstein et al., 2014). Research has shown that individuals respond to information about pollution by adopting costly avoidance behaviors to protect themselves from health threats. Much of this research has focused on air pollution (Neidell, 2004, 2009; Moretti and Neidell, 2011). However, Zivin et al. (2011) show that bottled water purchases increased in response to drinking water violations in northern California and Nevada from 2001 to 2005, especially for coliform bacteria violations. I highlight similar avoidance behavior for a new population with a longer and more recent time series. Importantly, I expand on this work by separating violations by the type of public notification required: within 24 hours vs. within 30 days.

I find that violations of the Total Coliform Rule (TCR) – the most commonly violated drinking water standard – significantly harm health and human capital accumulation. Specifically, TCR violations significantly increase purchases of common stomach remedies, ER visits for gastrointestinal illness, and school absences. To quantify avoidance behavior and the effects of differences in public notification rules, I further exploit variation in the type of TCR violation: "acute" violations require the public to be notified within 24 hours, while "monthly" violations require public notice only within 30 days. I find striking differences between the two types of violations: acute violations – violations that require the public to be notified of immediately – are associated with a statistically significant 78 percent increase in bottled water sales during the violation. Moreover, acute violations are not associated with any outcomes that would indicate adverse effects on health. In contrast, monthly violations – violations that require the public to be notified within 30 days – are not associated with any significant changes in bottled water sales, but do result in significant negative health and human capital effects. Monthly violations are associated with an increase in stomach remedy purchases of 23 percent of a standard deviation, an increase in ER visits for gastrointestinal illness of 23 percent among school-aged children, and a 5 percent increase in the absence rate. Taken together, these results strongly suggest that the type of public notification is important in inducing avoidance behavior and reducing the harmful effects of poor drinking water quality. Public notification rules are, therefore, an important policy lever that can be used to encourage avoidance behavior and reduce the harmful effects of exposure.

These findings on the effects of poor drinking water quality, the effects of water pollution regulations, and the effects of public notification rules are important and timely. Recent revisions to the TCR rule, effective as of 2016, removed warnings for “monthly” violations that previously required public notice within 30 days. It was thought that these violations did not have any direct public health effects. The results in this paper provide direct evidence that these monthly violations do harm both health and human capital accumulation: they significantly increased ER visits for gastrointestinal illness, purchases of common stomach remedies, and school absences among exposed children and adolescents. My results suggest that it is important to test for “monthly” violations and to provide immediate public notification of these violations as a low-cost policy tool that allows individuals to act to protect themselves from the harmful effects of exposure to contaminated water. Finally, I address the question of whether information dissemination is socially efficient. Back-of-the-envelope calculations suggest that providing immediate notification of all coliform violations would have increased avoidance cost by about $365,000 during the study period, which is small relative to the estimated $7.7 million in costs attributable to the health and human capital consequences of exposure.

1 Background

1.1 Water quality and health impacts

Specific disease-producing organisms are difficult, expensive, and time consuming to detect in water. For that reason, coliform bacteria are used as "indicator organisms" whose presence indicates that harmful pathogenic organisms may be in the water (see Table A1 in the appendix for pathogens capable of causing waterborne or water-based disease). Coliform bacteria are microorganisms that are often associated with human and animal fecal matter. Contamination of drinking water supplies can occur in several ways: a heavy rain may increase run-off entering the drinking water source, the water treatment process may be ineffective, or a break in the distribution system may allow contaminants to enter. Approximately 26% of the distribution pipes in the U.S. are in poor condition, and there are approximately 237,000 main breaks per year (Reynolds et al., 2008).[2] Samples collected in the field often show higher concentrations of fecal coliforms in surface water, relative to groundwater, especially where run-off from manure is likely (Cox et al., 2005). Water systems with surface water sources or with distribution systems in poor condition may be at higher risk of coliform bacteria contamination.

Potential health effects from exposure to waterborne or water-based diseases associated with

[2]See Ercumen et al. (2014) for a review of the epidemiological literature on the impact of distribution system deficiencies on gastrointestinal illness.

coliform bacteria include gastroenteric infections and diseases. Common symptoms include nausea, vomiting, diarrhea, and stomach cramps. Children, the elderly, and individuals with compromised immune systems are especially vulnerable. Acute gastroenteritis can be deadly to vulnerable individuals. Microorganisms can be removed by boiling tap water for at least one minute before drinking or purchasing bottled water. The costs associated with these avoidance behaviors include inconvenience as well as expenditures of time and money.

1.2 Water quality regulation and public notification

Under the Public Notification Rule of the Safe Drinking Water Act (SDWA), the US Environmental Protection Agency (EPA) classifies violations into 3 tiers, which differ in their notification requirements. Tier 1 violations require immediate notification within 24 hours via radio, TV, hand delivery, posting, or other methods. Tier 2 violations require public notification as soon as possible within 30 days. Tier 3 violations require an annual notice within one year of the violation date.

Each notification must include a description of the violation and contaminant levels, the date the violation occurred, potential adverse health risks, a description of the population at risk, whether alternative water supply should be used, what action consumers should take, what the system is doing to correct the violation, and when the system expects to return to compliance.[3]

1.3 Total Coliforms Rule

The Total Coliforms Rule (TCR) sets federal standards for bacteria in drinking water and is the most commonly violated regulation (Bennear and Olmstead, 2008). Systems are required to take a certain number of routine samples for total coliforms each month. Compliance is determined each month based on the presence or absence of total coliforms and/or fecal coliforms. Appendix B.2

[3] Additional detail on the SDWA and Public Notification Rule can be found in Appendix B.1.

contains additional detail on sampling and testing requirements.

The TCR outlines two types of violations: acute and monthly. An acute violation occurs when any repeat sample is positive for fecal coliform or E. coli, or the system has a positive total coliform repeat sample following a positive fecal coliform or E. coli routine sample (40 CFR 141.63). Acute violations are those that detect the presence of fecal coliforms or E. coli, which is found in large quantities in animal fecal matter. If E. coli is detected along with total coliforms, there is strong evidence that sewage is present and that a health threat exists. Therefore, acute coliform violations are classified as Tier 1 under the Public Notification Rule and require public notification within 24 hours. Acute coliform violations may result in a boil-water notice.

Monthly coliform violations are classified as Tier 2 and require notification within 30 days. A monthly violation occurs when a specified number of samples test positive for total coliform in a given month, even when tests for E. coli and fecal coliforms are negative. In systems that take fewer than 40 samples per month (those that serve 33,000 or fewer people), a monthly violation occurs when two or more samples test positive for total coliforms. In systems that take at least 40 samples per month, a monthly violation occurs when more than 5 percent of samples test positive for total coliforms.[4]

[4]This structure provides some incentive for strategic avoidance of regulatory action. Systems that would incur a violation if over 5 percent of samples test positive for total coliforms can avoid triggering a monthly TCR violation by taking additional water quality samples. Bennear et al. (2009) find evidence of strategic oversampling to avoid triggering a monthly TCR violation among systems serving over 33,000 people. In the sample used for this study, however, there were no violations of the TCR by systems serving over 33,000 people, so all monthly violations in the sample were triggered by the presence of two positive total coliform samples. To the extent that strategic oversampling occurred in our sample, violations that would have been recorded at large

The TCR was effective through March 2016, at which time the Revised Total Coliform Rule (RTRC) became effective. While the RTCR maintains the acute coliform violation (renamed as an "E. coli" violation), the monthly coliform violation has been eliminated. As there is no longer a monthly coliform violation, there are also no public notification requirements for this type of contamination. The RTRC also requires systems with an indication of coliform contamination to initiate assessments to find sanitary defects and take corrective action.

2 Data

This paper combines data from a number of different sources, each of which is described below. Using detailed geographic information, I combine water quality data with several key outcomes of interest: household purchases of bottled water and stomach remedies, emergency room visits for gastrointestinal illness, and school absences. Data on weather, air pollution, and employment are used as controls. Additional data detail is provided in Appendix C.

Water systems and water quality data

The EPA maintains detailed records on all water quality violations in the Safe Drinking Water Information System (SDWIS). These data provide information on both procedural violations (e.g., requirements for timely testing and reporting) and maximum contaminant level (MCL) violations. I obtain all health-based MCL violations at community water systems in North Carolina from 2004 to 2015. The North Carolina Department of Environmental Quality (NCDEQ) also provided public

systems in the absence of oversampling are unrecorded, and estimated effects shown here may be understated. In addition, because of the discrete nature of the tests and the fact that only 2 positive samples are necessary to trigger a violation, it is not possible to utilize a regression discontinuity in this context.

notification distribution dates.[5]

Geographic information

Information about the public community water supply (CWS) service areas comes from the North Carolina Center for Geographic Information and Analysis, available via NC OneMap.[6] Figure A1 shows the community water supply system service areas for North Carolina. Unserved areas tend to be rural and usually supply their own water for domestic use from fresh groundwater wells, which are not regulated under the SDWA.

To measure exposure, I intersect CWS service areas with point locations (e.g., geocoded residences of students and school locations) or geographic boundaries (e.g., zip codes), depending on the outcome of interest.[7] For each geographic boundary, I calculate the percentage of the total area that is served by each water supply system to measure the probability of exposure. For point locations within the boundaries of a CWS area, I define exposure based on that CWS.[8]

Household purchases

Data on household purchases for 2004-2015 are from the Nielsen Company and are maintained by

[5]Thanks to Linda Raynor for sharing and assisting in the interpretation of these data. These data do not contain information on the method of communication.

[6]Geographic service areas were mapped during 2004-2006 to facilitate planning, siting, and impact analysis. Data are available here: www.nconemap.com.

[7]The results are similar when I define exposure based on first observed location, rather than current location, to account for possible endogenous moving behavior.

[8]Note that not all households within a water supply service area necessarily get their water from the CWS. It is possible that some households supply their own water for domestic use from a well. However, the vast majority households (and schools) within a CWS area get their water from the CWS.

the James M. Kilts Center for Marketing at the University of Chicago Booth School of Business. The household-level consumer panel data include information about purchases made by a panel of consumer households from all retail outlets. The data include sales in dollars and quantity sold for 7,281 unique Universal Product Codes (UPCs) for bottled water (product module 1487). I define bottled water sales each month as total dollar sales (before coupons).[9] I also observe purchases of common remedies to treat gastrointestinal illness: diarrhea treatments, antacids, children's liquid pain remedies, and Pedialyte. Using total dollar sales for each category, I create a standardized mean value index. The products are normalized by the mean and standard deviation, as in Kling et al. (2007). I also use purchases of respiratory illness remedies as placebo tests, since poor water quality should primarily increase gastrointestinal illness. Purchases of children's cold remedies, sinus remedies, and medicines such as Sudafed, Nyquil, and Dayquil should be unaffected by poor water quality.

Emergency Department data

Data on emergency department visits for 2007-2015 are from the Cecil G. Sheps Center for Health Services Research at the University of North Carolina. Restricted-access patient-level data include 5-digit zip code of residence, patient characteristics and diagnosis and procedure codes.

Using age-specific population for each zip code from the 2010 Census, I calculate the number of gastrointestinal illness admissions per 1,000 population in each zip code and month.[10] Separate measures are calculated for the following age groups: 5-19, 20-39, 40-59, and 60 and over. School-age children age 5-19 tend to drink more tap water than very young children and are likely to be more vulnerable to exposure than adults (Drewnowski et al., 2013).

[9] Results are similar when measuring bottled water sales in total ounces sold.

[10] See Appendix C for diagnosis codes.

Absences and school data

The two main sources of administrative school data are from the North Carolina Department of Public Instruction (NCDPI): the Principal's Monthly Report (PMR) and the North Carolina Education Research Data Center's (NCERDC) student-level data.

First, the PMR dataset contains grade-level enrollment and attendance information, which is retrieved at the end of each of 9 PMR intervals from 2007 to 2015. I calculate the percentage of absent days as total absent days divided by total days in membership. I drop observations with a percentage of absences above the 99th percentile, or 22.1 percent. Table A2 shows that the average rate of absences is 4.5 percent in the PMR data.

Each of the 9 PMR intervals covers approximately 20 school days, which is not necessarily a calendar month.[11] To combine these data with the water quality violations reported by calendar month, I limit the sample to schools with a traditional school calendar (i.e., I remove year-round schools) and set the first PMR interval for each school year as September.[12] The 9 PMR intervals are therefore treated as representing the months September through May.[13] As there is no systematic collection of water quality for private wells, which serve primarily rural households, the analysis focuses on non-rural schools and violations by community water systems.

[11]Policy directs that the school systems can only report 20 days each for months 1 and 2 for the purpose of allocating funds. The remaining intervals (3-9) can range from 16 days to 26 days, but are usually 20 days.

[12]Under North Carolina law the public school opening date was August 25 from 2005 to 2013 and was the Monday closest to August 26 from 2013 onward. About 6% of schools in the data have non-traditional calendars.

[13]In appendix C.2, I use school calendar information for the period 2012-2015 to test this assumption. On average, there is 82% overlap between the PMR interval and the assumed month.

Second, the NCDPI also provides the NCERDC with information on absences and days in membership for all students at the time of testing each year. Unlike the PMR, the NCDPI does not indicate when absences occur within the school year. However, the student-level data for 2009-2015 contain a unique code for each student that can be linked over time and to the latitude and longitude of residence. Individuals are linked to water quality information based on the intersection of their precise home locations and community water supply service areas. Table A2 shows summary statistics for the NCERDC student-level data. The average absence rate is 4.4 percent, and the average days absent is about 7.

I also use the residential information in the student-level data to measure exposure to water quality violations at the school level. I calculate the percentage of students for each grade-by-year-by-school cell that is served by each public water supply system. I then use the violation data for each system to calculate the percentage of the class that is exposed to a water quality violation in each month. I merge this information with the grade-by-school-by-month absence data using a unique school identifier that is available in both the student level and the school level data.[14] Results are weighted by the number of students per cell that can be linked to a community water supply system based on their residential information.

Weather and air pollution

Daily weather data for each 2.5 by 2.5 mile square in North Carolina come from Schlenker and Roberts (2009) and are based on the PRISM weather dataset. I calculate the average daily minimum temperature, maximum temperature, and precipitation for each school using grid squares

[14] I drop observations where the difference between the number of students in the two datasets is larger than 200. Because of data limitations prior to 2009, I assume residential location is the same in school year 2007-08 as in 2008-09.

whose centroids fall within the school's attendance boundary. School attendance boundaries come from the National Center for Education Statistics for the 2013-2014 school year.[15] I calculate monthly and yearly measures of total precipitation, average precipitation, and the percentage of days with a maximum temperature that falls within 7 temperature bins in degrees Celsius: below 0, 0-5, 5-10, 10-15, 15-20, 20-25, and over 25. I focus on maximum temperature because it is more representative of temperature during daytime hours. However, the results, available upon request, are robust to using minimum temperature or average temperature as well. I also include controls for 7 precipitation bins in millimeters as a robustness check: 0-1, 1-5, 5-10, 10-15, 15-20, 20-25, and over 25. In the regressions, I omit the lowest temperature and precipitation bins.

School absences may also be driven by exposure to air pollution (Currie et al., 2009). As a robustness check, I include controls for ozone and particulate matter measured by air quality monitors. I include the percentage of days within each of five categories to allow for nonlinear effects: 0-25%, 25-50%, 50-75%, 75-100%, and over 100% of the relevant EPA threshold for each pollutant. For the school-level analysis, I measure air pollution at the school-month attendance boundary level. For the individual analysis, I use inverse distance-weighted measures of air pollution over the past 12 months from monitors within 20 miles of a student's residence.

3 Methods

Given the existing evidence of extensive residential sorting, cross-sectional estimates of the impact of water quality on child outcomes would likely be biased. Parental income, as well as preferences for school quality, neighborhood amenities, and environmental quality are likely related to water quality as well as to underlying health and human capital. Table A3 shows summary statistics

[15] Figure A2 shows primary, middle, and high school attendance boundary areas.

for students exposed and unexposed to any coliform violation at home and at school. Students exposed to a coliform violation have higher absence rates, but are also 7-8 percentage points more likely to be economically disadvantaged and about 8-9 percentage points more likely to be black. This finding is broadly consistent with the environmental justice literature, which consistently finds higher pollution exposure among disadvantaged populations (Banzhaf et al., 2019). The empirical strategies used here control for time-invariant location characteristics by including zip code, school, or individual fixed effects, thus accounting for potential bias from cross-sectional sorting. I exploit plausibly exogenous variation in the timing of water quality violations to estimate the within-location impacts of water quality violations on avoidance behavior and child outcomes.

3.1 Household purchases analysis

I explore the effect of violations on household purchases using data recording all purchases made by a sample of households over time. Bottled water purchases are a measure of avoidance behavior. Purchases of common remedies for gastrointestinal illness are a proxy for health status. The general specification for the household-level data is as follows,

$$Purchases_{h,z,m,y} = \pi_1 Acute_{z,m,y} + \pi_2 Monthly_{z,m,y} + \omega_{z,m,y} + X_{h,z,m,y} + \eta_h + \zeta_z \times \mu_m + \zeta_z \times \psi_y + \psi_y \times \mu_m + \varepsilon_{h,z,m,y} \quad (1)$$

where h indexes household, z indexes zip code, m indexes month, and y indexes school year. The variables of interest, $Acute_{z,m,y}$ and $Monthly_{z,m,y}$, indicate acute and monthly coliform violations, where exposure is measured as the percentage of the zip code exposed to a violation each month. The specification includes a vector of weather controls, $\omega_{z,m,y}$, a vector of household fixed effects, η_h, as well as all two-way interactions between zip codes, years, and months. I also include controls for household size and employment from the Local Area Unemployment Statistics in

$X_{h,z,m,y}$. Standard errors are clustered at the zip code level.

These controls account for many potential sources of bias. Zip code fixed effects control for cross-sectional differences in demographics and neighborhood characteristics that might be related to both water pollution exposure and household purchases. The main specification includes a vector of flexible weather controls, $\omega_{z,m,y}$, to account for fluctuations in weather over time that may influence both exposure and bottled water purchases. For example, when it is hotter, people may increase consumption of both bottled and tap water, thus increasing their potential exposure.

Two-way interactions control for several potential sources of bias. For example, zip code-specific trends or shocks are captured by the zip code-by-year fixed effects, $\zeta_z \times \psi_y$. Abnormal shocks to all zip codes in a particular month (e.g., if a hurricane causes an increase in bottled water purchases throughout the state) are captured by year-by-month fixed effects, $\psi_y \times \mu_m$. Finally, zip code-by-month fixed effects, $\zeta_z \times \mu_m$, account for zip code-specific seasonality. The coefficients of interest, π_1 and π_2, would be biased if there are any unobserved shocks that influence purchases in the same zip code, in the same year, and in the same month as water quality violations.

Bottled water purchases are measured as the inverse hyperbolic sine of total dollar sales. This functional form can be interpreted as a standard logarithmic transformation, but also preserves zeros. As the stomach remedies do not share common units and differ in cost, I use a standardized mean value index. Figure A3 shows the distribution of bottled water and stomach remedy purchases. There are many months with zero purchases in the household data. Therefore I also consider the extensive and intensive margins separately. For the extensive margin, I use binary indicators for any sales. For the intensive margin, I limit the sample to months with non-zero sales.

3.2 ER visit analysis

As a direct measure of health, I explore the effect of violations on emergency room visits for gastrointestinal illnesses. The general specification for the zip code-level data is as follows,

$$GI_{z,m,y} = \lambda_1 Acute_{z,m,y} + \lambda_2 Monthly_{z,m,y} + \omega_{z,m,y} + \lambda_3 Flu_{z,m,y} + \zeta_z + \tag{2}$$
$$\zeta_z \times \mu_m + \zeta_z \times \psi_y + \psi_y \times \mu_m + \varepsilon_{z,m,y}$$

where z indexes zip code, m indexes month, and y indexes school year. $GI_{z,m,y}$ measures the number of gastrointestinal admissions per 1,000 population in the month for each zip code. Exposure is measured as the percentage of the zip code exposed to a violation in each month. The specification includes a vector of weather controls, $\omega_{z,m,y}$, a vector of zip code fixed effects, ζ_z, as well as all two-way interactions between zip codes, years, and months. I also include controls for influenza and pneumonia ER admissions in $Flu_{z,m,y}$ and weight the regression by population. Standard errors are clustered at the zip code level.

3.3 School-level analysis

Next, I explore the effect of poor water quality on school absences. Many factors may influence absence rates, including family structure, opportunity costs, school environment, and school policies. However, most school absences are illness-related and can be attributed to either respiratory infections or gastroenteritis, which is a common symptom associated wtih exposure to many waterborne pathogens (Currie et al., 2009; Gilliland et al., 2001; Kearney, 2008).[16] To show whether the timing of school absences coincides with the timing of violations, I estimate the following regression specification,

[16] Currie et al. (2009) show that some school absences for respiratory conditions can be attributed to air pollution.

$$PctAbsent_{s,g,m,y} = \alpha_1 Acute_{s,g,m,y} + \alpha_2 Monthly_{s,g,m,y} + \omega_{s,m,y} + \gamma_g + \sigma_s \\ + \sigma_s \times \mu_m + \sigma_s \times \psi_y + \psi_y \times \mu_m + \varepsilon_{s,g,m,y} \tag{3}$$

where s indexes school, g indexes grade, m indexes month, and y indexes school year. $PctAbsent_{s,g,m,y}$ measures the percentage of absent days in the month for each grade in each school. Exposure is measured as the percentage of students exposed to a violation each month based on student-level residential information. The specification includes a vector of weather controls, $\omega_{s,m,y}$, a vector of grade indicators, γ_g, a vector of school fixed effects, σ_s, as well as all two-way interactions between schools, years, and months. The regression is weighted by the number of students matched to a water supply system in each grade-school-year-month cell. Standard errors are clustered at the school level.

School fixed effects control for cross-sectional differences in demographics and neighborhood characteristics that might be related to both water pollution exposure and absence rates. Grade indicators control for differences in absenteeism that are related to age. Absence rates tend to increase with age (USDOE, 2019; Allison et al., 2019). Absences for young children are more likely to be health-related, as young children are more reliant on parental permission to miss school than older students.

Two-way interactions control for additional potential sources of bias. For example, school-specific trends, changes in attendance policy, and shocks in demographic composition are captured by $\sigma_s \times \psi_y$ controls. Abnormal shocks to all schools in a particular month, such as an unusually bad flu season, are captured by $\psi_y \times \mu_m$ controls. Finally, $\sigma_s \times \mu_m$ controls account for school-specific seasonality that might arise from school-specific attendance policies or variation in the timing or fraction of students missing school around holidays, for example. The coefficients of interest, α_1

and α_2, would be biased if there are any unobserved shocks that affect absences in the same school, in the same year, and in the same month as water quality violations.

3.4 Student-level analysis

Next, the student-level data allow me to define individual-level exposure, $Acute_{i,t}$ and $Monthly_{i,t}$, based on residential and school latitudes and longitudes. Exposure is equal to one if a violation occurred within the past year. Individual fixed effects, ϕ_i, control for all time-invariant individual, family, and neighborhood characteristics. As before, it is important to include grade level, year, and weather controls. The main specification is as follows,

$$Y_{i,t} = \beta_1 Acute_{i,t} + \beta_2 Monthly_{i,t} + \omega_{i,t} + \gamma_i + \psi_t + \phi_i + \varepsilon_{i,t} \tag{4}$$

where i indexes individuals and t indexes school year and the outcome of interest is the student-level absence rate. Standard errors are clustered at the individual level. The coefficients of interest, β_1 and β_2, capture the impact of exposure to an acute or monthly coliform violation during the past year on an individual's absence rate. The estimate would be biased if there are time-varying unobserved factors correlated with absence rates that affect students in specific areas with water quality violations during the year of violation.

4 Results

4.1 Avoidance behavior: Bottled water results

While acute coliform violations require public notice within 24 hours and often include a boil-water advisory, monthly coliform violations only require notice within 30 days. Using bottled water sales to measure avoidance behavior, we can see how individuals respond to acute and monthly coliform violations.

Table 1 shows estimates from equation 1 for bottled water in column 1. Identification comes from variation in violations at the zip code-by-year-by-month level after controlling for household fixed effects, year-by-month fixed effects, zip code-by-year fixed effects, zip code-by-month fixed effects, weather, employment, and household size.[17] Bottled water sales increase only during acute coliform violations when public notice is required within 24 hours. Acute coliform violations increase bottled water purchases by about 78 percent.[18]

Figure 1 shows these effects graphically. For each variable of interest, I create a placebo violation in the months immediately before and after the actual violation. Violations in the data typically last one month before returning to compliance. If there are pre-existing trends in purchases, these should show up in the months before the violation. If the avoidance behavior or health effects are sustained beyond the violation month, this will show up in the months after the violation. There is an increase in bottled water purchases during the acute coliform violation, but no increase during monthly coliform violations. The increase in purchases during acute violations is even more apparent for the intensive margin, while the intensive margin shows no increase during monthly

[17]Table A4 in the appendix builds up to this fully specified model by sequentially adding in additional controls to show how the point estimates change. Columns 6-8 also show that the results are robust to including region-by-year-by-month fixed effects. Table A5 in the appendix shows the results for household purchases on the extensive and intensive margins. These results show positive and significant effects of acute coliform on bottled water sales overall and for the intensive margin, and positive and significant effects of monthly coliform on stomach remedy purchases overall and for the intensive margin. The negative coefficient for acute coliform on the intensive margin of stomach remedies is not robust to the inclusion of region-by-year-by-month fixed effects, as shown in Panel B of Table A5, and should be interpreted with caution.

[18]The percentage change is based on the approximation $(e^{(\hat{\beta})} - 1) \times 100$. Using the small-sample bias correction, $(e^{(\hat{\beta} - 0.5\widehat{Var}(\hat{\beta}))} - 1) \times 100$, yields a percentage change of 72 percent (Bellemare and Wichman, 2020).

coliform violations. For either type of violation, there are no significant effects for months before or after the violation month.

Using additional data on public notification timing, column 3 in Table 1 explores whether the delay in public notification for a monthly violation matters for bottled water purchases. As monthly violations require notification only within 30 days, there is variation in how quickly public information is delivered. Figure A4 in the appendix shows the distribution of weeks between the determination date and public notification date for monthly coliform violations. About 60 percent of public notifications occur within the first week. However, this distribution has a long right tail with about 30 percent of public notifications occurring more than 6 months after the determination date. Using this variation in the timing of actual public notification, I compare purchase behavior when the notification is provided within one day of the violation (similar to the required 24-hour notice for acute violations) with purchase behavior when the notification is provided later. If the immediacy of public notification matters, we might expect to find increases in bottled water purchases during monthly coliform violations when the public notification occurred within 24 hours rather than later on. In fact, column 3 shows bottled water purchases during monthly violations with quick notification are statistically indistinguishable from bottled water purchases during acute violations, although not statistically different from zero. When notifications are delivered within one day, the point estimates indicate a 78 percent and 40 percent increase in bottled water sales for acute and monthly violations, respectively. For monthly violations with delayed notification, there is no increase in bottled water purchases and the point estimate is actually negative.

An additional benefit of the household-level data is the ability to observe certain demographic characteristics of each household. I explore demographic heterogeneity in bottled water purchases during acute coliform violations in Figure A5 and Table A6 for the intensive margin. All demographic groups increase bottled water purchases, but effects are statistically significantly larger for Hispanic and white households and for those with more education. These findings are generally consistent with existing evidence showing that avoidance behavior is more common among higher socio-economic status individuals (Marcus, 2020; Currie, 2011).

Next, I explore whether other types of purchases change during water quality violations. These placebo tests use purchases of medicines that treat respiratory illness rather than gastrointestinal illness, including cold remedies for children, sinus remedies, and products such as Sudafed, Nyquil, and Dayquil. These purchases should not be affected by water quality violations. Columns 1-3 of Table A7 show that the estimated effects of water quality violations on purchases of medicines used to treat respiratory illness are small and statistically insignificant, providing support for the argument that changes in bottled water purchases are related to water quality and not attributable to some other confounding factor.

4.2 Health: Stomach remedies & ER visit results

Although there is no evidence of an increase in bottled water purchases during monthly coliform violations, there is evidence of an increase in the purchase of common stomach remedies. Column 2 of Table 1 shows that monthly violations increase the purchase of stomach remedies by 0.23 standard deviations, with no significant increase during acute violations. This suggests that households are able to avoid harmful health effects by purchasing bottled water during acute, but not monthly, violations.

Figure 2 shows the effect of monthly violations on the purchase of common stomach remedies. There appears to be an increase in such purchases during the violation month, but little evidence of any effects before the violation. The confidence intervals for the months following the violation are wide, possibly due to lingering health effects.

Next, I explore whether the immediacy of public notification matters. I compare purchase behavior when the notification is provided within one day of the violation with purchase behavior when the notification is provided later. We might expect to find increases in stomach remedy purchases during monthly coliform violations when the public notification is delayed beyond one day. Column 4 of Table 1 also shows that stomach remedy purchases increase significantly only during monthly violations with delayed notification, suggesting that the timing of information is especially important for protecting health.

In addition to purchases of over-the-counter stomach remedies, emergency room visits for gastrointestinal illness provide a measure of health that is both direct and extreme. Columns 5-9 of Table 1 show the impact of water quality violations on ER visits from estimating equation 2 for all individuals and for different age groups.[19] Consistent with the previous results, there is an increase in ER admissions for gastrointestinal illness during monthly coliform violations, but there is no statistically significant impact on admissions during acute coliform violations. For a zip code affected by a monthly coliform violation, ER admissions for gastrointestinal illness increase by about 0.6 per 1,000 individuals, or about 14 percent from the mean.

The effect is especially strong for school-age individuals in column 6, which shows about a 23 percent increase in admissions during a monthly coliform violation. Figure 3 shows these results visually for acute coliform in the left panel and monthly coliform in the right panel. As before, I create a placebo violation in the months immediately before and after the actual violation. During

[19]Table A8 in the appendix builds up to this fully specified model by sequentially adding in controls. Columns 6 and 7 also show that the results are robust to including region-by-year-by-month fixed effects.

an acute coliform violation, there is no evidence of an increase in visits to the ER for gastrointestinal illness before or during the violation. While there does appear to be a significant increase in ER visits two periods after the violation, this coefficient is only significant at the 10 percent level when region-by-year-by-month fixed effects are included or when weighted by the population on a community water supply system, and, therefore, should be interpreted with caution. For monthly coliform violations, on the other hand, there is a significant and robust increase in ER visits during the violation month, with no evidence of increased ER visits in the months before or after the violation month.

In Table 1, there is also a statistically significant impact on adults age 20 to 39, but no significant impact on older adults (see columns 8 and 9). Since some waterborne pathogens, such as *Giardia*, are highly contagious, it may be that increased illness among adults age 20 to 39 is related to parents' interactions with infectious children. Across all age groups, the effect of acute coliform violations on ER admissions for gastrointestinal illness is not statistically different from zero.

I also conduct several placebo tests in Table A9 to show that there is no significant effect of either acute or monthly coliform violations on ER visits for a variety of alternate diagnoses that should not be affected by water quality: fractures; burns; poisoning by drugs and medicinal and biological substances; acute respiratory infections; and diseases of the nervous system and sense organs. The coefficients for both acute and monthly violations are small and statistically insignificant for all placebo outcomes.

4.3 School absences: School-level results

Column 1 of Table 2 shows the results from the school-level analysis using the monthly PMR data to estimate equation 3.[20] The outcome is the percentage of days absent in the month for each grade in each school. Exposure to a violation is measured as the percentage of students affected by each type of violation, which varies at the school-grade-year-month level.

The effect of acute coliform exposure is small and statistically insignificant, whereas the effect of exposure to a monthly coliform violation is positive and significant. If all individuals in the grade are exposed to a monthly coliform violation during the month, relative to no exposure, the percentage of days absent is expected to increase by about 13 percentage points. It is important to note, however, that 100 percent exposure is outside of the observed range in the data for monthly coliform. An increase from zero exposure to the maximum exposure to monthly coliform, 14.7 percent, leads absences to increase by 1.9 percentage points, or about 22 percent from the mean.

Figure 4 shows graphical evidence of the timing of absences relative to the violation with placebo violations in the months before and after the actual violation, as before. For both types of violations, acute and monthly coliform, there are no pre-existing trends in the two months before a violation. Although acute violations are not associated with a significant increase in absences in Panel A, monthly violations in Panel B are associated with a significant increase in absences at the

[20]Table A10 in the appendix shows how the point estimates change as controls are added in sequentially. Only the coefficient for monthly coliform violations is consistently positive and significant across all specifications, suggesting that school absences increase when monthly coliform violations occur. The estimated effect of monthly coliform violations is also robust to including region-by-year-by-month fixed effects to account for any regional shocks or events that might coincide with water violations.

time of the violation.

Next, I explore the results by age. Because younger students are more likely to require parental permission to miss school, their absences are more likely to be health-related, and we might expect to find more precise estimates among younger students. Panels C and D of Figure 4 show the effect of monthly coliform violations for grades 1-6 and grades 7-12. There is a clear increase in absences during the violation month for grades 1-6. For grades 7-12, the confidence intervals are wider, as expected. We cannot rule out a similar impact on older students.

Table A11 tests the robustness of the school-level absence results. Column 1 replicates the baseline results. All specifications include controls for grade, school, year-month, school-year, school-month, and weather. Columns 2 and 3 add controls for the county-month employment rate and controls for any monitoring and reporting violations. Column 4 includes controls for influenza and asthma rates. These are measured as the weighted average of the number of ER admissions per person in zip codes intersecting the school attendance boundary, where the weights are the percentage of the school attendance area covered by each zip code. Column 5 includes controls for ozone and particulate matter, measured as the percentage of days within five pollution-level bins to allow for nonlinear effects. In column 6, I exclude schools in Wake and Mecklenburg Counties to test whether the results are driven by Raleigh and Charlotte. Column 7 drops the two schools in North Carolina that ever experience an unplanned school closure in response to water quality violations.[21] Column 8 excludes magnet schools that draw students from a broader geographic area. Column 9 excludes the month of September, since attendance-based funding decisions may drive the unusually low absence rate in this month. Column 10 collapses the grade-level data to

[21]Unplanned school closure data from 2011-2017 were generously provided by CDC/DGMQ/CI-ICU. Please consult Wong et al. (2014) for data collection methods.

the school level. Next, as PMR intervals do not perfectly overlap with calendar months, column 11 drops schools whose average PMR coverage of calendar month days is less than 70% based on limited data from school years 2012-13 to 2014-15. Finally, column 12 includes flexible bins for precipitation: 0-1mm, 1-5mm, 5-10mm, 10-15mm, 15-20mm, 20-25mm, and over 25mm. Across all specifications the coefficients and standard errors remain very similar.

4.4 School absences: Student-level results

Having shown that the timing of absences coincides with the timing of violations using the monthly school-level data, I now turn to the student-level data, which provides residential location and allows students to be observed over time. Columns 2-4 of Table 2 show estimates of the main results from equation 4.[22] Identification of the effect of acute and monthly violations comes from variation at the individual-by-year level.

The magnitude of the estimates shown in column 2 suggests that exposure to an acute coliform violation has no statistically significant impact on school absences. However, exposure to a monthly coliform violation increases absences by about 0.2 percentage points, or 5 percent from the mean. Using the average number of absent days per year, 7.16, this translates to an additional 0.4 days absent for exposure to a monthly coliform violation at home.

[22]Table A12 in the appendix shows that the results are robust to including controls for other types of health-based violations, including radionuclides, disinfectants and disinfection by-products, synthetic organic compounds (SOCs), and volatile organic chemicals (VOCs). Exposure is measured as months of violation in the past year. Although there appears to be some effect of other violations on absences, these results should be interpreted with caution as this research design is not well suited to capturing the effects of other chemical exposures that have more complicated sampling requirements.

One unique feature of these data is the ability to observe two locations in which individuals are potentially exposed to water pollution: home and school. Children spend significant amounts of time at school and are likely to consume tap water at lunch, during recess, and at other times throughout the school day. Most previous research quantifying the effects of pollution on child health has focused on exposure either at home or at school, but here we can consider both. Understanding the effects of contaminated drinking water at school is especially important since some drinking water regulations, such as the EPA's Lead and Copper Rule, do not require any testing at schools served by CWS systems, and about half of Charlotte-Mecklenburg schools recently tested had lead levels above the EPA action level of 15 ppb. (Rumpler and Dietz, 2019).

Column 3 of Table 2 shows the effect of coliform violations at school on student absences. Monthly coliform violations at school increase absences, and the magnitude of the effect is very similar to monthly coliform violations at home, about 0.2 percentage points. However, acute coliform violations at school appear to reduce absences by about 0.2 percentage points. It is important to remember that acute violations require immediate notification and increase avoidance of tap water. School administrators may be especially cautious about protecting their students. If drinking water at schools is generally of poor quality, acute coliform violations may actually improve students' health when schools restrict their consumption of tap water. Column 4 looks at the interaction between home and school violations. Although there is a high correlation (0.6-0.7) between violations at home and at school since many children attend a school that is served by the same system as their home, there are still many children who are served by different systems at home. Acute violations only at school continue to be associated with lower rates of absence, but there are no statistically significant impacts for acute violations only at home or when acute violations occur in both locations. On the other hand, the increase in school absences during monthly violations is statistically significant whether the violations occur only at home, only at school, or at both locations. The point estimate is slightly larger when violations occur in both locations. These findings are consistent with previous ER visit and stomach remedy purchase results that showed evidence of worse health only during monthly coliform violations.

Because the precise residential location allows me to link students to a specific community water system, I can explore heterogeneous effects by system characteristics. Figure A6 and Table A13 show estimates of equation 4 where the coefficient of interest, monthly coliform violation, is interacted with system characteristics. Estimates show a statistically significant impact only for surface water systems, which is consistent with the fact that samples often show higher concentrations of fecal coliforms in surface water than in groundwater, especially where run-off from manure is likely (Cox et al., 2005).. Next, I define repeat offenders as water systems with more than one violation of the TCR since 1990 and non-repeat offenders as systems that violate the TCR for the first time. The effect is statistically significant for both repeat and non-repeat offenders, but statistically significantly smaller for repeat offenders (p-value of 0.0207). When compared by ownership type, the results are statistically significantly larger for publicly owned water supply systems (p-value of 0.001). Finally, point estimates increase with system size, but only the smallest and largest systems are statistically significantly different.[23]

Next, I explore heterogeneous effects by student demographics. Drewnowski et al. (2013) show that among children certain demographic groups consume more tap water, including older children,

[23]The SDWA classifies systems as very small (0-500), small (501-3,300), medium (3,301-10,000), large (10,001-100,000), and very large (100,000+). Nationally, only about 1% of CWS are very large. There were no very large systems that violated the Total Coliform Rule in North Carolina during the study period.

non-Hispanic whites, and children from higher-income families. Figure A7 and Table A14 show estimates of equation 4 where the coefficient of interest, monthly coliform violations, is interacted with demographic characteristics. In general, point estimates are larger among populations that tend to consume more tap water (non-disadvantaged, non-Hispanic, and older children), but they are not statistically significantly different.

Finally, Table A15 tests the robustness of the student-level results. Column 1 replicates the baseline results. All specifications include individual fixed effects and controls for grade, year, and weather. Columns 2 and 3 add controls for the county-month employment rate and controls for any monitoring and reporting violations. Columns 4 and 5 add controls for influenza and asthma ER visits and air pollution, respectively. Columns 6 and 7 drop Mecklenburg and schools that experienced an unplanned closure in response to water quality violations. Columns 8-10 limit the sample to regular schools (excludes alternative education, exceptional children, and vocational education), regular programs (excludes co-op innovative high schools, early college schools, magnet schools, etc.), and schools with traditional calendars (excludes modified calendars and year-round calendars), respectively. Column 11 includes seven flexible bins for precipitation, and column 12 includes temperature and precipitation bins separately for each month. Columns 13 and 14 include region-by-year controls where NCDEQ regions are used in column 13, and column 14 uses the Coastal Plains, Piedmont, and Mountains geographic regions. Across all specifications, the coefficients and standard errors remain very similar.

I also explore Poisson and Negative Binomial models for the student-level results in Table A16.[24] Regardless of the specification, exposure to monthly violations increase absences by about 5 percent, which is consistent with the magnitude of the findings in Table 2.

[24]Figure A8 shows the distribution of both percentage of days absent and number of days absent.

5 Discussion & Conclusions

This paper shows that drinking water violations can have important effects on health. Monthly coliform bacteria violations increase ER visits for gastrointestinal illness, purchases of common stomach remedies, and school absences. The timing of these poor outcomes coincides with the timing of violations, and exposure at both home and school can affect child health.

Importantly, the results show that when information about coliform violations is provided immediately, households avoid exposure by purchasing bottled water. When information is not provided immediately, households cannot respond by avoiding exposure and there are important impacts on health. However, whether immediate information dissemination is socially efficient in this context is still unclear, as households may over-respond to the information. With some back-of-the-envelope calculations, we can compare the cost of purchasing additional bottled water with the health costs of exposure to contaminated water.

Acute coliform violations, which require immediate 24-hour notification, increase bottled water purchases by 78 percent. This translates into about $61,000 in excess purchases of bottled water from 2007 to 2015 in North Carolina.[25] If requiring immediate notification of monthly coliform violations yielded a similar behavioral response, bottled water purchases would have been about $365,000 higher over this period.

In comparison, the health costs incurred during monthly coliform violations are large. Stomach

[25] Based on the percentage of each zip code exposed to a violation over this period and population estimates from the 2010 census, there were approximately 72,000 and 432,000 person-months exposed to acute and monthly coliform violations, respectively. Of these, about 14,000 and 88,000 individuals were of school age. Note that the estimated effects for bottled water and stomach remedy sales are at the household level, so using population-level exposure will overstate the costs.

remedy purchases increase by 23 percent of a standard deviation, which is an additional $441,000 over this time period. ER visits for gastrointestinal illness increase by 14 percent overall and by 23 percent among school-aged children. Using the average charges for gastrointestinal illness visits, monthly coliform violations increase ER charges by about $422,000, with about $80,000 due to school-aged children.[26] The direct health costs alone exceed the cost of avoiding exposure by drinking bottled water, yet there are additional effects on school absences to consider.

Although it is difficult to put a dollar value on the health consequences associated with a missed school day, absences also have a direct impact on parental earnings to the extent that parents stay home to take care of sick children. Using the average hourly earnings in 2015 from the Bureau of Labor Statistics, $25, and assuming each student absence equates to one lost 8-hour work day, the estimates presented above suggest the total cost in lost earnings from monthly coliform violations over this period was about $6.86 million. In addition to the effect on parental earnings, absences have a direct effect on schools' and districts' budgets. For example, school absences in California cost public schools $3.5 billion in state funding based on daily attendance between 2010-11 and 2012-13 (Harris, 2014). School absences can also have important impacts on student outcomes, including test scores, grade retention, and later life earnings (Aucejo and Romano, 2016).

All together, monthly coliform exposure over the study period cost about $7.7 million, which is substantially more than the estimated avoidance cost of providing immediate information about these violations, $365,000. Importantly, the EPA revised the Total Coliform Rule in 2016 and eliminated monthly coliform violations. Although the contaminations may continue to exist, they no longer need to be reported. While the overall health effects of the Revised Total Coliform Rule

[26]Average charges were $1,648 and $1,258 in North Carolina in 2007 for all individuals and school-aged children, respectively.

are unknown, the findings of this study indicate that monthly coliform violations did in fact have a negative impact on health. This suggests not only that testing for monthly coliform violations is important, but that providing immediate public information about monthly coliform violations can allow individuals to avoid negative health impacts. Disseminating information in a timely way can act as a low-cost policy nudge to mitigate damaging health effects by encouraging individuals to avoid exposure by boiling their tap water or finding a new source of drinking water.

References

Allison, Mandy A, Elliott Attisha et al., “The link between school attendance and good health,” *Pediatrics*, 2019, *143* (2).

Anderson, D Mark, Kerwin Kofi Charles, and Daniel I Rees, “Public health efforts and the decline in urban mortality,” Technical Report, National Bureau of Economic Research 2018.

Aucejo, Esteban M and Teresa Foy Romano, “Assessing the effect of school days and absences on test score performance,” *Economics of Education Review*, 2016, *55*, 70–87.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins, “Environmental Justice: The Economics of Race, Place, and Pollution,” *Journal of Economic Perspectives*, 2019, *33* (1), 185–208.

Beach, Brian, Joseph Ferrie, Martin Saavedra, and Werner Troesken, “Typhoid fever, water quality, and human capital formation,” *The Journal of Economic History*, 2016, *76* (1), 41–75.

Bellemare, Marc F and Casey J Wichman, “Elasticities and the inverse hyperbolic sine transformation,” *Oxford Bulletin of Economics and Statistics*, 2020, *82* (1), 50–61.

Bennear, Lori S and Sheila M Olmstead, "The impacts of the "right to know": Information disclosure and the violation of drinking water standards," *Journal of Environmental Economics and Management*, 2008, *56* (2), 117–130.

, Katrina K Jessoe, and Sheila M Olmstead, "Sampling out: regulatory avoidance and the total coliform rule," 2009.

Colford, John M, Sharon Roy, Michael J Beach, Allen Hightower, Susan E Shaw, and Timothy J Wade, "A review of household drinking water intervention trials and an approach to the estimation of endemic waterborne gastroenteritis in the United States," *Journal of water and health*, 2006, *4* (S2), 71–88.

Cox, Stephen E., William Simonds, Llyn Doremus, Raegan L. Huffman, and Rose M. Defawe, "Ground water/surface water interactions and quality of discharging ground water in streams of the lower Nooksack River Basin, Whatcom County, Washington," Technical Report, U.S. Geological Survey 2005.

Currie, Janet, "Inequality at birth: Some causes and consequences," *American Economic Review*, 2011, *101* (3), 1–22.

and Enrico Moretti, "Mother's education and the intergenerational transmission of human capital: Evidence from college openings," *The Quarterly journal of economics*, 2003, *118* (4), 1495–1532.

, Eric A Hanushek, E Megan Kahn, Matthew Neidell, and Steven G Rivkin, "Does pollution increase school absences?," *The Review of Economics and Statistics*, 2009, *91* (4), 682–694.

, Joshua Graff Zivin, Katherine Meckel, Matthew Neidell, and Wolfram Schlenker, "Something in the water: Contaminated drinking water and infant health," *Canadian Journal of Economics/Revue canadienne d'économique*, 2013, *46* (3), 791–810.

Cutler, David and Grant Miller, "The role of public health improvements in health advances: the twentieth-century United States," *Demography*, 2005, *42* (1), 1–22.

Drewnowski, Adam, Colin D Rehm, and Florence Constant, "Water and beverage consumption among children age 4-13y in the United States: analyses of 2005–2010 NHANES data," *Nutrition journal*, 2013, *12* (1), 85.

Ely, Todd L and Mark L Fermanich, "Learning to count: School finance formula count methods and attendance-related student outcomes," *Journal of Education Finance*, 2013, pp. 343–369.

Ercumen, Ayse, Joshua S Gruber, and John M Colford Jr, "Water distribution system deficiencies and gastrointestinal illness: a systematic review and meta-analysis," *Environmental Health Perspectives*, 2014, *122* (7), 651.

Gilliland, Frank D, Kiros Berhane, Edward B Rappaport, Duncan C Thomas, Edward Avol, W James Gauderman, Stephanie J London, Helene G Margolis, Rob McConnell, K Talat Islam et al., "The effects of ambient air pollution on school absenteeism due to respiratory illnesses," *Epidemiology*, 2001, *12* (1), 43–54.

Gottfried, Michael A, "The detrimental effects of missing school: Evidence from urban siblings," *American Journal of Education*, 2011, *117* (2), 147–182.

Grossman, Michael and Robert Kaestner, "Effects of Education on Health," *The social benefits of education*, 1997, *12*, 69.

Harris, Kamala D., "In school + on track: Attorney general's 2013 report on California's elementary school truancy & absenteeism crisis," Technical Report, California Dept. of Justice 2014.

Hill, Elaine and Lala Ma, "Shale Gas Development and Drinking Water Quality," *American Economic Review*, 2017, *107* (5), 522–525.

Kearney, Christopher A, "School absenteeism and school refusal behavior in youth," *Clinical psychology review*, 2008, *28* (3), 451–471.

Keiser, David and Joseph Shapiro, "US Water Pollution Regulation over the Past Half Century: Burning Waters to Crystal Springs?," *Journal of Economic Perspectives*, 2019, *33* (4), 51–75.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

Lleras-Muney, Adriana, "The relationship between education and adult mortality in the United States," *The Review of Economic Studies*, 2005, *72* (1), 189–221.

Loewenstein, George, Cass R Sunstein, and Russell Golman, "Disclosure: Psychology changes everything," *Annu. Rev. Econ.*, 2014, *6* (1), 391–419.

Marcus, Michelle, "Going Beneath the Surface: Petroleum Pollution, Regulation, and Health," *Forthcoming at American Economic Journal: Applied Economics*, 2020.

Moretti, Enrico and Matthew Neidell, "Pollution, health, and avoidance behavior evidence from the ports of Los Angeles," *Journal of human Resources*, 2011, *46* (1), 154–175.

Neidell, Matthew, "Information, Avoidance behavior, and health the effect of ozone on asthma hospitalizations," *Journal of Human resources*, 2009, *44* (2), 450–478.

Neidell, Matthew J, "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma," *Journal of health economics*, 2004, *23* (6), 1209–1236.

Reynolds, Kelly, Kristina Mena, and Charles Gerba, "Risk of waterborne illness via drinking water in the US," in "Reviews of environmental contamination and toxicology," Springer, 2008, pp. 117–158.

Rumpler, John and Emma Dietz, "Get the Lead Out: Ensuring Safe Drinking Water for Our Children at School," Technical Report, Environment America Research and Policy Center 2019.

Schlenker, Wolfram and Michael J Roberts, "Nonlinear temperature effects indicate severe damages to US crop yields under climate change," *Proceedings of the National Academy of sciences*, 2009, *106* (37), 15594–15598.

USDOE, "Chronic Absenteeism in the Nation's Schools," Technical Report, USDOE 2019.

Weil, David, Archon Fung, Mary Graham, and Elena Fagotto, "The effectiveness of regulatory disclosure policies," *Journal of Policy Analysis and Management*, 2006, *25* (1), 155–181.

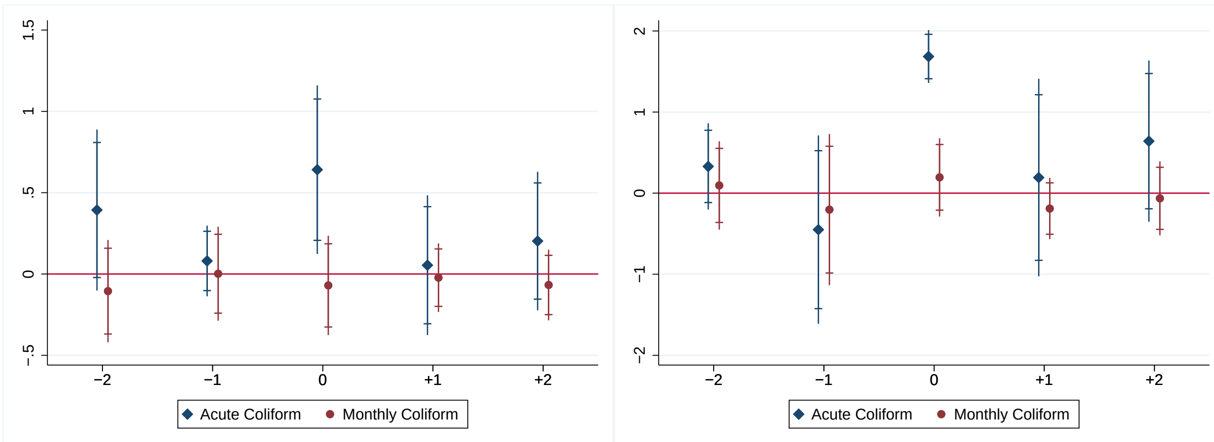
Wong, Karen K, Jianrong Shi, Hongjiang Gao, Yenlik A Zheteyeva, Kimberly Lane, Daphne Copeland, Jennifer Hendricks, LaFrancis McMurray, Kellye Sliger, Jeanette J Rainey, and Amra Uzicanin, "Why is school closed today? Unplanned K-12 school closures in the United States, 2011–2013," *PloS one*, 2014, *9* (12), e113755.

Zivin, Joshua Graff, Matthew Neidell, and Wolfram Schlenker, "Water quality violations and avoidance behavior: Evidence from bottled water consumption," *The American Economic Review*, 2011, *101* (3), 448–453.

6 Figures

Figure 1: Purchases of Bottled Water Relative to Violation Timing

(a) Water (b) Water: Intensive

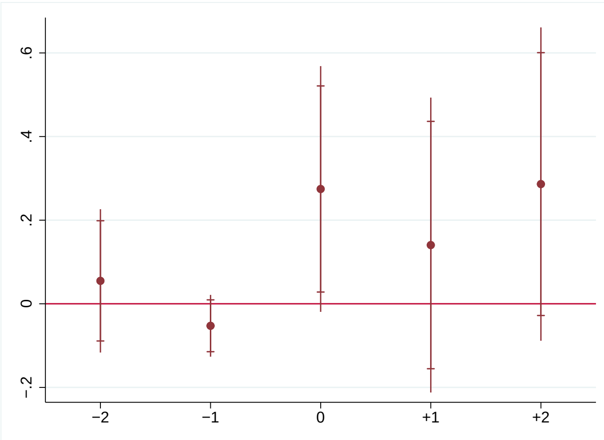


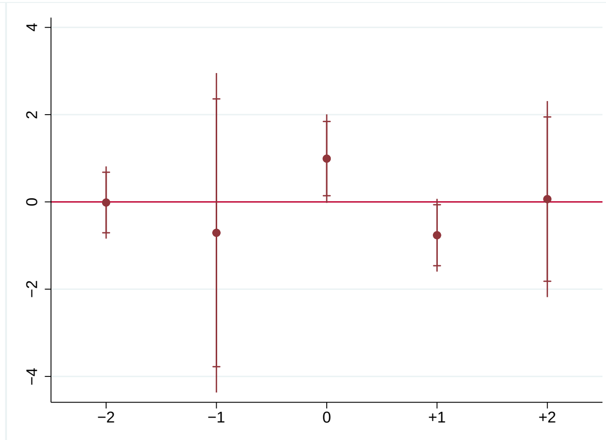
Notes: Water is measured as the inverse hyperbolic sine of total dollar sales of bottled water. Panel B shows results for the intensive margin, where the sample is limited to months with non-zero bottled water sales. Estimated effects for placebo violations are shown for the violation month, at time zero, and for two months before and after the actual violation. Coefficients are shown for the effect of both acute and monthly violations on bottled water purchases. Regressions include controls for weather, employment rate, household size, and the following fixed effects: household, year-by-month, zip code-by-year, and zip code-by-month. Standard errors are clustered at the zip code level.

Figure 2: Purchases of Stomach Remedies Relative to Monthly Coliform Violation Timing

(a) Stomach Remedies

(b) Stomach Remedies: Intensive



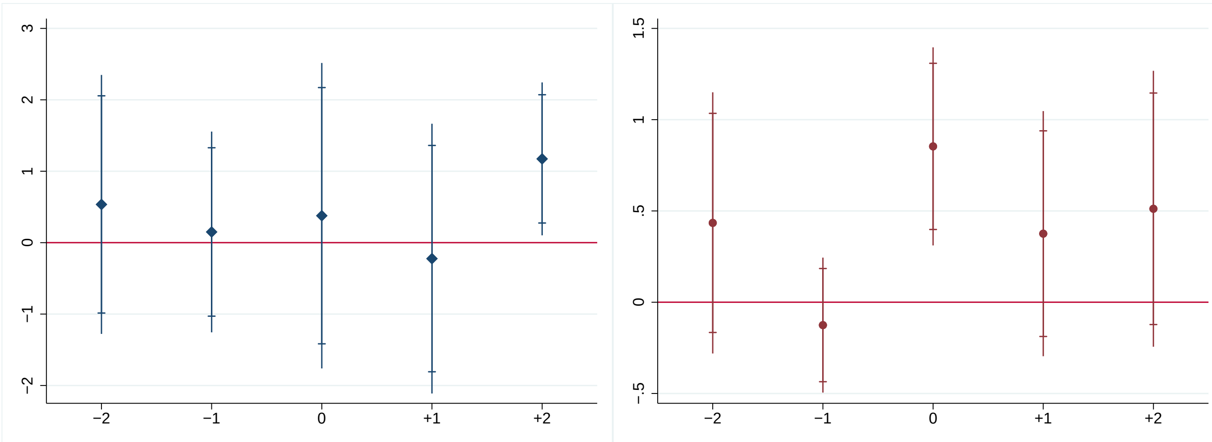


Notes: Stomach remedies is a standardized mean value index of the dollar sales for each of the following categories: diarrhea treatments, antacids, children's liquid pain remedies, and Pedialyte. Panel B shows results for the intensive margin, where the index is based on months with non-zero sales. Estimated effects for placebo violations are shown for the violation month, at time zero, and for two months before and after the actual violation. Coefficients are shown for the effect of monthly coliform violations. Regressions include controls for weather, employment rate, household size, and the following fixed effects: household, year-by-month, zip code-by-year, and zip code-by-month. Standard errors are clustered at the zip code level.

Figure 3: ER Admissions for Gastrointestinal Illness: Age 5-19

(a) Pct Acute Coliform

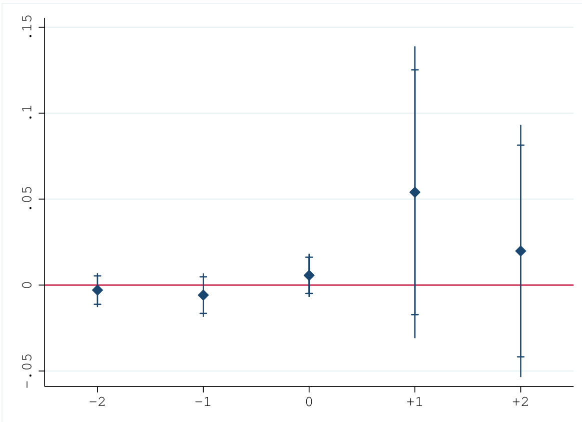
(b) Pct Monthly Coliform



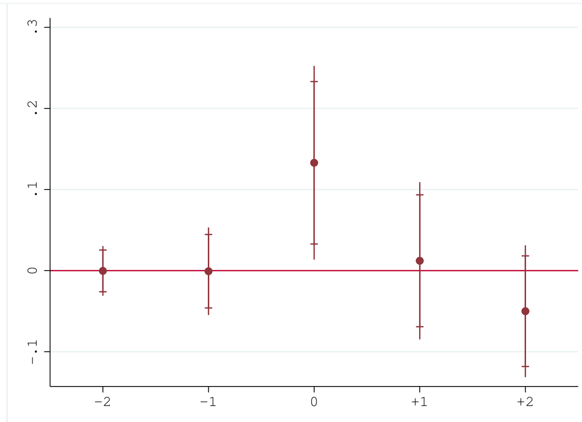
Notes: ER visits per 1,000 individuals for gastrointestinal illness among school-age children, age 5-19, is based on any of the following ICD-9 diagnosis codes: 001-009 (Intestinal Infectious Diseases), 558.9 (Noninfectious Gastroenterit NEC), 787.0 (Nausea and vomiting), and 787.9 (Diarrhea). I exclude newborn admissions and visits with *Clostridium difficile* (008.45). Estimated effects for placebo violations are shown for the violation month, at time zero, and for two months before and after the actual violation. Coefficients are shown for the effect of both acute and monthly coliform violations on ER visits for gastrointestinal illnesses. Regressions are weighted by population and include controls for weather, influenza and pneumonia ER visits, and the following fixed effects: zip code, year-by-month, zip code-by-year, and zip code-by-month. Standard errors are clustered at the zip code level.

Figure 4: Absences Relative to Violation Timing

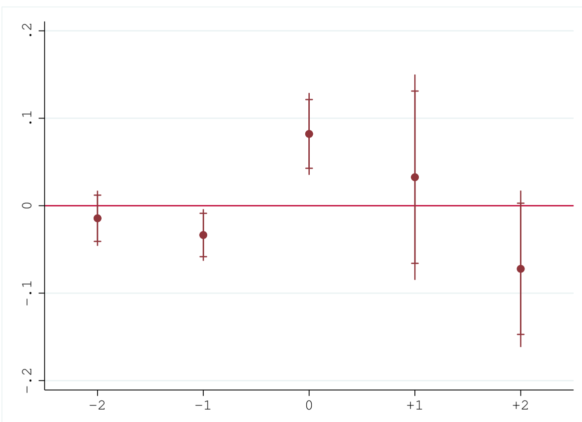
(a) Pct Acute Coliform



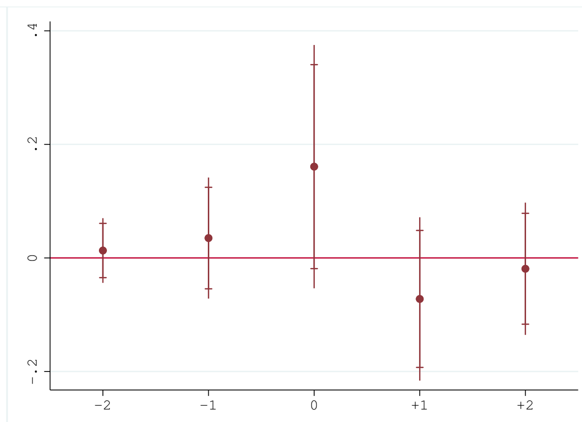
(b) Pct Monthly Coliform



(c) Pct Monthly Coliform: Grades 1-6



(d) Pct Monthly Coliform: Grades 7-12



Notes: School absences are measured as total absent days divided by total days in membership. Estimated effects for placebo violations are shown for the violation month, at time zero, and for two months before and after the actual violation. Coefficients are shown for the effect of both acute and monthly coliform violations in Panels A and B, respectively. Panels C and D show the estimated effects of monthly coliform violations on school absences for grades 1-6 and grades 7-12, respectively. Regressions are weighted by the number of students matched to a water supply system and include controls for weather, grade, and the following fixed effects: school, year-by-month, school-by-year, and school-by-month. Standard errors are clustered at the school level.

7 Tables

Table 1: Effect of violations on avoidance and health

	Bottled	Stomach	Bottled	Stomach	Gastrointestinal ER visits				
	Water	Remedies	Water	Remedies	All	Age 5-19	Age 20-39	Age 40-59	Age 60+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pct Acute Col	0.577**	0.0444	0.575**	0.0456	0.0966	0.237	0.608	-0.686	0.807
	(0.259)	(0.0623)	(0.259)	(0.0632)	(1.008)	(0.956)	(1.090)	(1.335)	(0.723)
Pct Monthly Col	-0.0525	0.234*			0.593**	0.717***	1.377**	0.234	0.0618
	(0.151)	(0.134)			(0.234)	(0.262)	(0.534)	(0.307)	(0.363)
Pct Monthly Col $\times$			0.338	-0.0366					
Notice $\leq$ 1 Day			(0.227)	(0.0476)					
Pct Monthly Col $\times$			-0.174	0.318*					
Notice $>$ 1 Day			(0.180)	(0.165)					
Observations	245,632	245,632	245,632	245,632	82,339	82,242	82,339	82,025	81,571
R-squared	0.382	0.177	0.382	0.177	0.874	0.622	0.805	0.760	0.643
Household	yes	yes	yes	yes					
Zip					yes	yes	yes	yes	yes

Year-month	yes	yes	yes	yes	yes	yes	yes	yes	yes
Zip-yr	yes	yes	yes	yes	yes	yes	yes	yes	yes
Zip-month	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mean					4.329	3.165	6.048	3.836	3.771

Source: Nielsen household-month level data (Columns 1-4) and ER visit zip code-month level data (Columns 5-9).

Notes: All regressions include controls for weather. Columns 1-4 include employment rate and household size. Columns 5-9 include influenza and pneumonia ER visit controls and are weighted by population. Standard errors clustered at the zip code level are shown in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: Absences and the Impact of Water Quality Violations

	(1)	(2)	(3)	(4)
Acute Coliform: Home	0.00696	-0.0159		
	(0.00545)	(0.0681)		
Monthly Coliform: Home	0.131**	0.235***		
	(0.0571)	(0.0438)		
Acute Coliform: School			-0.230***	
			(0.0705)	
Monthly Coliform: School			0.247***	
			(0.0424)	
Only Acute School				-0.547***
				(0.155)
Only Acute Home				0.0610
				(0.138)
Both Acute				-0.136
				(0.0839)
Only Monthly Home				0.179***
				(0.0646)
Only Monthly School				0.174**
				(0.0748)
Both Monthly				0.305***
				(0.0564)

Observations	150,594	2,967,567	2,810,428	2,728,698
R-squared	0.760	0.724	0.729	0.729
Year		yes	yes	yes
Student FE		yes	yes	yes
School FE	yes			
Year-month	yes			
School-yr	yes			
School-month	yes			

Source: PMR school-grade-month level data (Column 1) and NCERDC student-year level data (Columns 2-4).

Notes: All specifications include grade and weather controls. Column 1 is weighted by the number of students matched to a water supply system. Standard errors are clustered at the school level in column 1 and at the individual level in columns 2-4. Coefficients and standard errors are scaled by 100 in columns 2-4. *** p<0.01, ** p<0.05, * p<0.1

For Online Publication

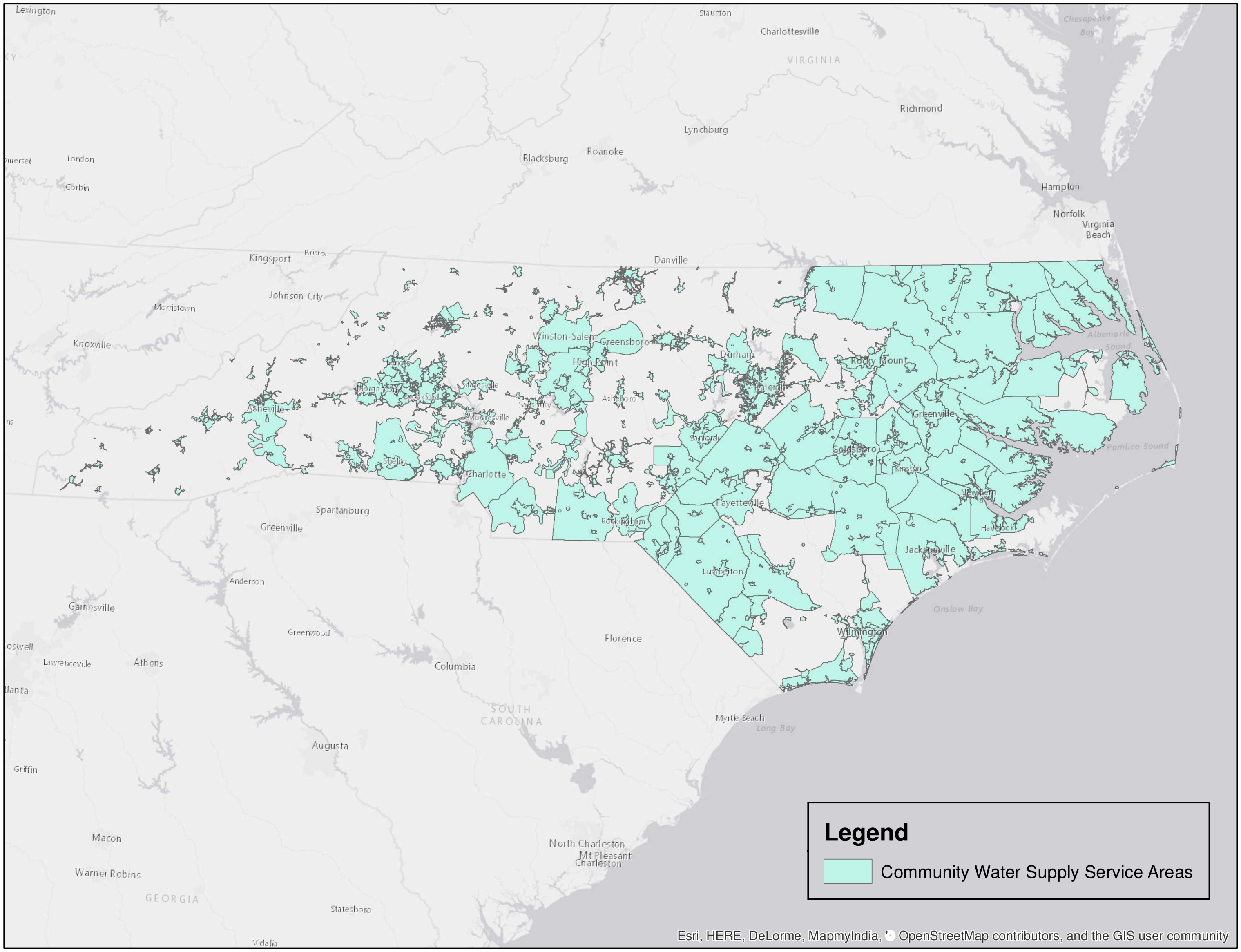
A Appendix Tables & Figures

Table A1: Agents of Waterborne or Water-based Disease

Bacteria	Protozoa	Viruses
Vibrio cholerae	*Giardia lamblia*	Norovirus
Salmonella spp.	*Cryptosporidium parvum*	Sapprovirus
Shigella spp.	*Entamoeba histolitica*	Poliovirus
Toigenic *Escherichia coli*	*Cyclospora cayetanensis*	Coxsackievirus
Campylobacter spp.	*Isospora belli*	Echovirus
Yersinia enterocolitica	Microsporidia	Paraechovirus
Legionella	*Ballantidium coli*	Enteroviruses 69-91
Helicobacter pylori	*Toxoplasma gondii*	Reovirus
	Naegleria fowleri	Adenovirus
		Hepatitis A & E
		Rotavirus
		Astrovirus
		Picobirnavirus
		Coronavirus

Source: Reynolds et al. (2008)

Figure A1: NC Community Water Supply Areas

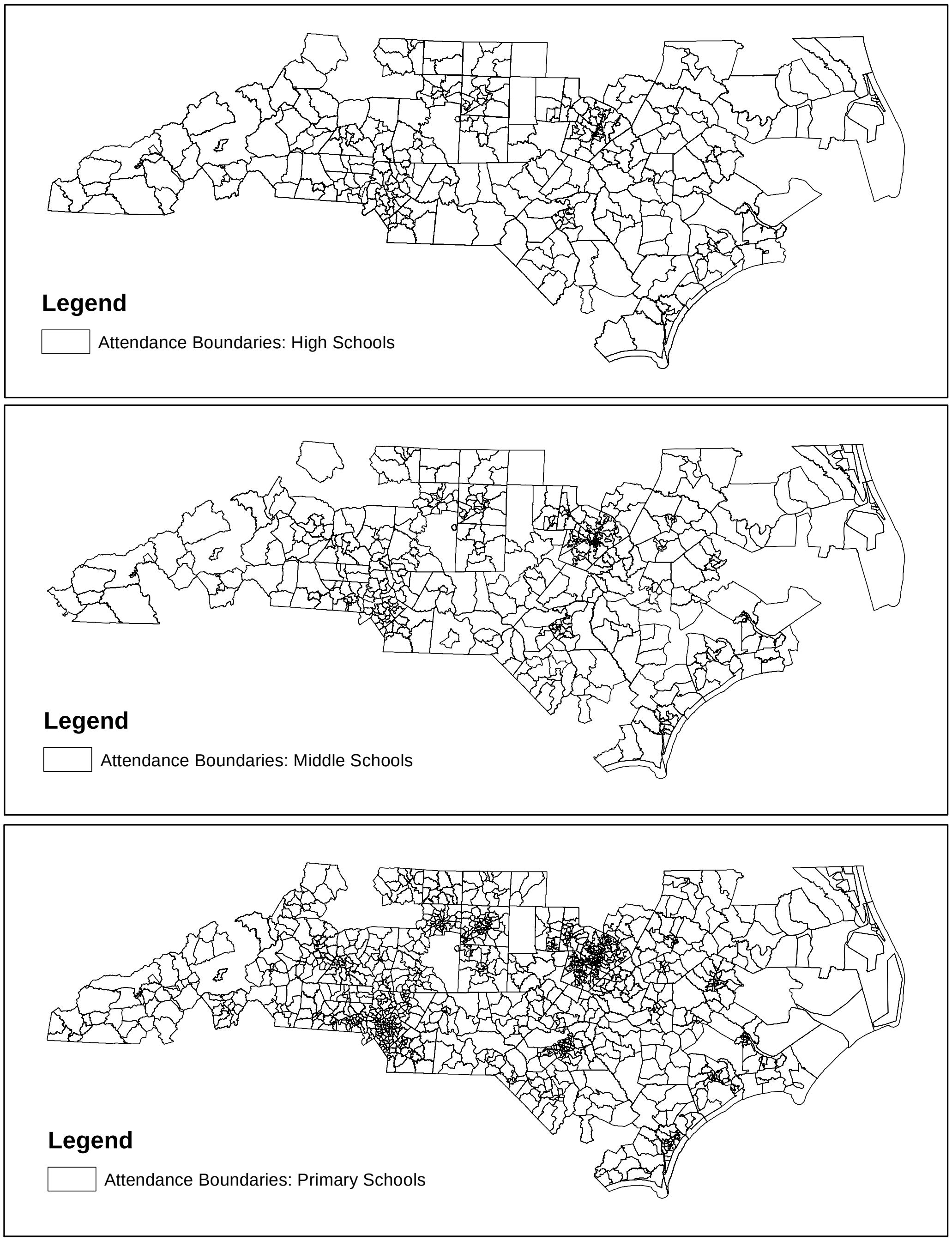


Notes: Figure shows community water supply service areas in North Carolina based on data from the North Carolina Center for Geographic Information and Analysis, available via NC OneMap.

Table A2: Summary Statistics

	Mean	Std. Dev.	Min	Max	N
Purchases data (household-level)					
Bottled Water (oz)	115	384	0	13312	219,666
Bottled Water Sales ($)	1.27	4.23	0	282	219,666
Any oz	.197	.398	0	1	219,666
Any sales	.197	.398	0	1	219,666
Bottled Water (oz) $\mid$ oz >0	585	687	4.05	13312	43,315
Bottled Water Sales ($) $\mid$ sales >0	6.43	7.59	.03	282	43,287
Pct Acute Coliform	.0000666	.00632	0	.986	219,666
Pct Monthly Coliform	.000254	.0112	0	.996	219,666
Household Size	2.39	1.23	1	9	219,666
$<$ $30,000	.221	.415	0	1	219,666
$30,000-$49,999	.267	.442	0	1	219,666
$50,000-$69,999	.196	.397	0	1	219,666
$\geq$ $70,000	.317	.465	0	1	219,666
Any Kids	.229	.42	0	1	219,666
$\geq$ College Grad	.531	.499	0	1	219,666
White	.764	.424	0	1	219,666
Black	.192	.394	0	1	219,666
Asian	.0137	.116	0	1	219,666
Other	.03	.171	0	1	219,666
Hispanic	.0224	.148	0	1	219,666
ER data (zip code-level)					
GI (per 1000 pop)	4.33	6.23	0	256	82,339
GI: age 5-19	3.16	11.9	0	1000	82,242
GI: age 20-39	6.05	12.2	0	500	82,339
GI: age 40-59	3.84	8.17	0	500	82,025
GI: age 60+	3.77	9.27	0	429	81,571
Flu/Pneumonia	1.18	3.49	0	302	82,339
Pct Acute Coliform	.0000766	.00775	0	.986	82,339
Pct Monthly Coliform	.000659	.0221	0	1	82,339
PMR data (school-level)					
Pct Absent	.0449	.0192	0	.22	150,615
Pct Acute Coliform	.0000894	.00638	0	.647	150,615
Pct Monthly Coliform	.0000278	.00150	0	.147	150,615
NCERDC data (student-level)					
Pct Absent	.0442	.0537	0	1	2,967,567
Days Absent	7.16	8.39	0	174	2,967,567
White	.442	.497	0	1	2,967,565
Black	.332	.471	0	1	2,967,565
Hispanic	.141	.348	0	1	2,967,565
Other	.0844	.278	0	1	2,967,565
Male	.508	.5	0	1	2,967,566
Econ. Disadvantaged	.533	.499	0	1	1,980,554
Limited English	.0957	.294	0	1	1,980,581
Acute Coliform: Home	.00314	.056	0	1	2,967,567
Monthly Coliform: Home	.00552	.0741	0	1	2,967,567
Acute Coliform: School	.00278	.0526	0	1	2,728,698
Monthly Coliform: School	.00625	.0789	0	1	2,728,698

Figure A2: NC School Attendance Boundary Areas



Notes: Figure shows school attendance boundaries for high schools, middle schools, and primary schools for the 2013-2014 school year from the National Center for Education Statistics. Boundaries may overlap and may be non-contiguous.

Table A3: Student Demographics by Total Coliform Exposure

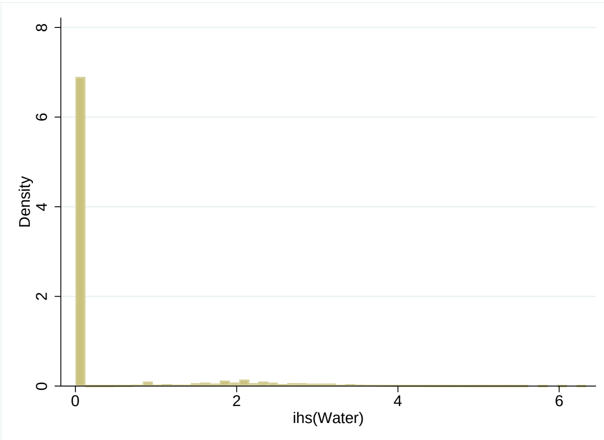
	Home location		School location	
	Unexposed	Exposed	Unexposed	Exposed
Econ. Disadvantaged	0.533	0.613***	0.531	0.599***
White	0.443	0.419***	0.440	0.433**
Black	0.331	0.424***	0.333	0.409***
Hispanic	0.141	0.103***	0.141	0.105***
Other	0.085	0.054***	0.086	0.053***
Limited English	0.096	0.090**	0.095	0.089**
Disabled	0.123	0.116**	0.122	0.114***
Pct Absent	0.044	0.048***	0.044	0.048***

Source: NCERDC student-level data.

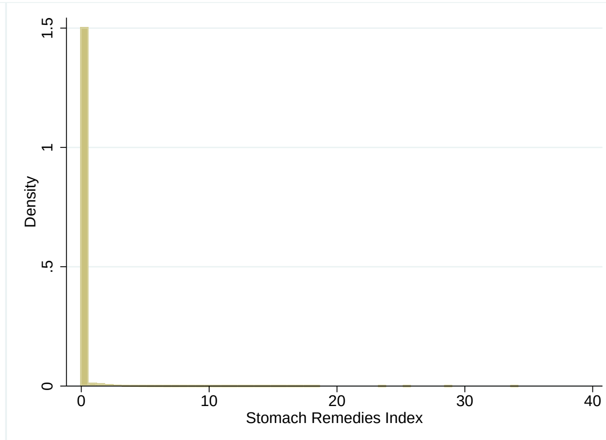
Notes: Exposure at home or at school is measured as any exposure to a violation of the Total Coliform Rule and includes both acute and monthly violations. Stars indicate the p-value associated with a test of the equality of the means between exposed and unexposed characteristics. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure A3: Distribution of Household-level Purchases

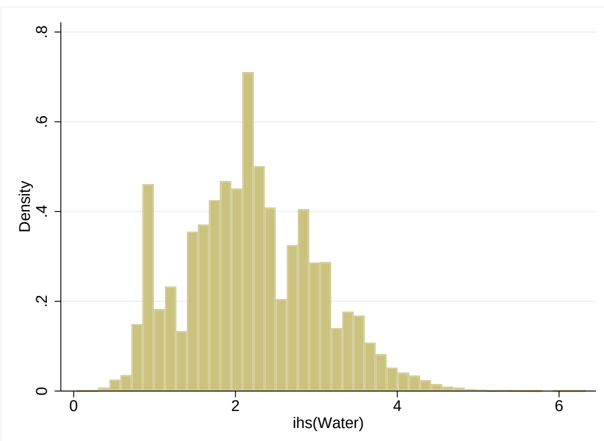
(a) Water



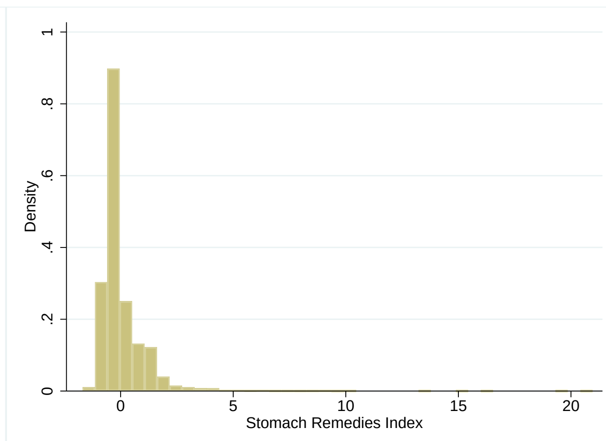
(c) Stomach remedies



(b) Water: Intensive



(d) Stomach remedies: Intensive



Notes: Figure shows the distribution of monthly household-level purchases. Bottled water purchases are measured as the inverse hyperbolic sine of total dollar sales. Purchases of common remedies to treat gastrointestinal illness include diarrhea treatments, antacids, children's liquid pain remedies, and Pedialyte. Using total dollar sales for each category, I create a standardized mean value index. Panels a and c show the distribution water and stomach remedy purchases overall. Panels b and d show the distribution for the intensive margin after excluding months without any sales.

Table A4: Household-level Purchases of Bottled Water and Stomach Remedies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Bottled Water								
Pct Acute Coliform	0.764**	0.519**	0.464**	0.453**	0.577**	0.583**	0.578**	0.615**
	(0.307)	(0.230)	(0.230)	(0.194)	(0.259)	(0.285)	(0.262)	(0.285)
Pct Monthly Coliform	0.201	0.0989	0.0987	0.0918	-0.0525	-0.0215	-0.0655	-0.0528
	(0.196)	(0.113)	(0.112)	(0.110)	(0.151)	(0.167)	(0.149)	(0.168)
Observations	246,409	246,236	246,236	246,086	245,632	245,632	245,632	245,632
R-squared	0.014	0.329	0.330	0.366	0.382	0.384	0.383	0.384
Panel B. Stomach Remedies								
Pct Acute Coliform	-0.0155	0.0332	0.0297	0.0487	0.0444	0.0308	0.0225	0.0504
	(0.0430)	(0.0669)	(0.0702)	(0.0475)	(0.0623)	(0.0572)	(0.0492)	(0.0685)
Pct Monthly Coliform	0.226	0.229	0.233	0.219	0.234*	0.244*	0.233*	0.225*
	(0.172)	(0.157)	(0.156)	(0.137)	(0.134)	(0.139)	(0.137)	(0.132)
Observations	246,409	246,236	246,236	246,086	245,632	245,632	245,632	245,632
R-squared	0.002	0.121	0.122	0.156	0.177	0.180	0.178	0.180
Year	yes	yes	yes	yes	yes	yes	yes	yes
Month	yes	yes	yes	yes	yes	yes	yes	yes
Household		yes	yes	yes	yes	yes	yes	yes
Year-month			yes	yes	yes	yes	yes	yes
Zip-yr				yes	yes	yes	yes	yes
Zip-month					yes	yes	yes	yes
DEQ Region-yr-month						yes		yes
Geo Region-yr-month							yes	yes

Source: Nielsen household-month level data.

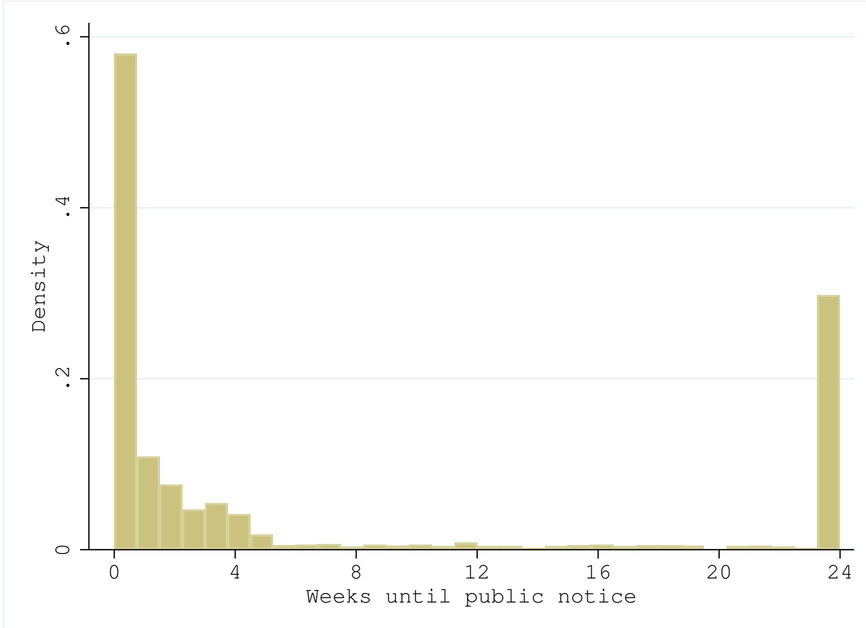
Notes: All regressions include controls for weather, employment rate, and household size. Column 6 uses regions defined by North Carolina's Department of Environmental Quality and column 7 uses the Coastal Plains, Piedmont, and Mountains geographic regions. Standard errors clustered at the zip code level are shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A5: Household Purchases: Extensive and Intensive Margins

			Extensive		Intensive	
	Water	Remedies	Water	Remedies	Water	Remedies
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.						
Pct Acute Coliform	0.577**	0.0444	0.0987	0.0151	1.622***	-1.344***
	(0.259)	(0.0623)	(0.0949)	(0.0554)	(0.0960)	(0.403)
Pct Monthly Coliform	-0.0525	0.234*	-0.0433	0.0164	0.218	1.169***
	(0.151)	(0.134)	(0.0683)	(0.0273)	(0.188)	(0.433)
Observations	245,632	245,632	245,632	245,632	40,758	11,735
R-squared	0.382	0.177	0.322	0.195	0.575	0.629
Panel B.						
Pct Acute Coliform	0.615**	0.0504	0.0946	0.0269	1.872***	-0.902
	(0.285)	(0.0685)	(0.107)	(0.0601)	(0.216)	(0.850)
Pct Monthly Coliform	-0.0528	0.225*	-0.0295	0.00645	0.123	1.937**
	(0.167)	(0.132)	(0.0746)	(0.0309)	(0.229)	(0.757)
Observations	245,632	245,632	245,632	245,632	40,758	11,692
R-squared	0.384	0.180	0.324	0.198	0.585	0.659

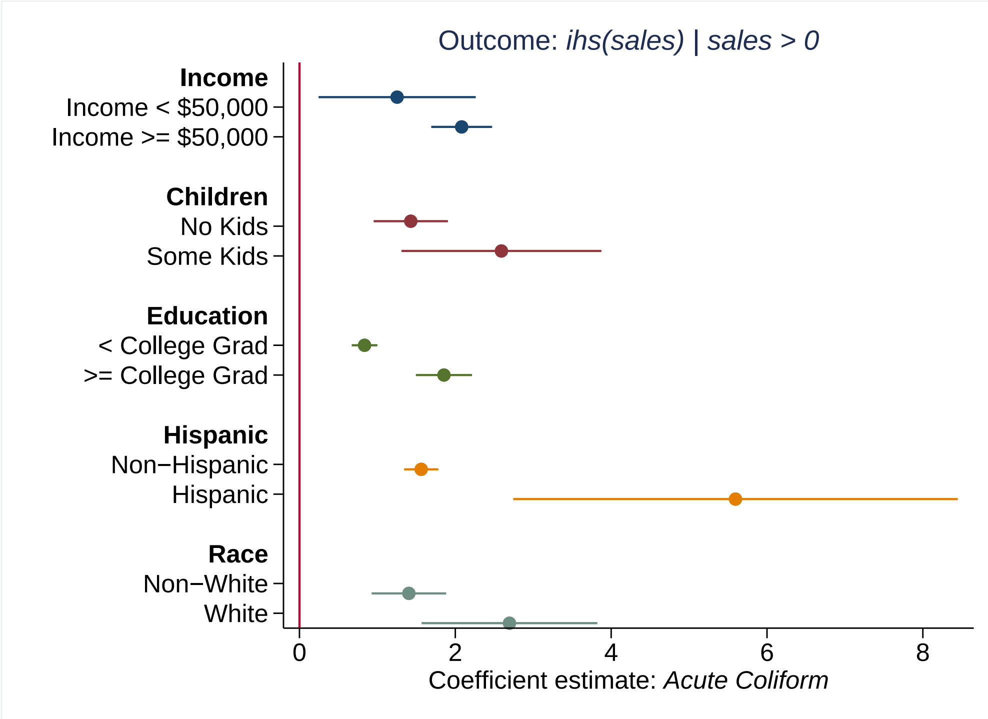
Source: Nielsen household-month level data.
Notes: All regressions include controls for weather, employment rate, and household size. All regressions also include fixed effects for year, month, year-by-month, zip code-by-year, and zip code-by-month. Panel B also includes NCDEQ and Geographic region-by-year-by-month fixed effects. Columns 3 and 4 are binary outcomes equal to one if the household purchases any dollar amount of bottled water or stomach remedies in that month, respectively. Columns 5-6 drop months without any bottled water or stomach remedies purchases. Standard errors clustered at the zip code level are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Figure A4: Time Until Public Notification: Monthly Coliform



Notes: Figure shows the distribution of weeks between the violation and the dissemination of public information. The final bin includes notifications that occurred 24 or more weeks from the time of the violation.

Figure A5: Heterogeneous effects of acute coliform violations on bottled water sales



Notes: Water is measured as the inverse hyperbolic sine of total dollar sales of bottled water. Results are shown for the intensive margin, where the sample is limited to months with non-zero bottled water sales. Estimated effects of acute coliform violations on bottled water purchases are shown separately by household income, number of children, the highest education of a household head, Hispanic ethnicity, and race. Regressions include controls for weather, employment rate, household size, and the following fixed effects: household, year-by-month, zip code-by-year, and zip code-by-month. Standard errors are clustered at the zip code level.

Table A6: Household Bottled Water Purchases: Heterogeneous Results

	$ihs(sales) \mid sales > 0$				
	(1)	(2)	(3)	(4)	(5)
Income $< \$50,000 \times$Pct Acute Coliform	1.254**				
	(0.513)				
Income $\geq \$50,000 \times$Pct Acute Coliform	2.082***				
	(0.199)				
No Kids × Pct Acute Coliform		1.429***			
		(0.243)			
Some Kids × Pct Acute Coliform		2.592***			
		(0.653)			
$<$ College Grad × Pct Acute Coliform			0.835***		
			(0.0839)		
$\geq$ College Grad × Pct Acute Coliform			1.855***		
			(0.183)		
Non-Hispanic × Pct Acute Coliform				1.563***	
				(0.112)	
Hispanic × Pct Acute Coliform				5.596***	
				(1.451)	
Non-white × Pct Acute Coliform					1.405***
					(0.243)
White × Pct Acute Coliform					2.695***
					(0.574)
Observations	40,994	40,994	40,994	40,994	40,994
R-squared	0.576	0.576	0.576	0.576	0.576
P-value	0.159	0.139	0.000	0.006	0.061

Source: Nielsen household-month level data.
Notes: All regressions include household fixed effects and controls for year, month, year-month, zip-year, zip-month, weather, employment rate, and household size. The sample excludes observations without any bottled water purchases. P-values are for tests of the equality of estimated coefficients. Standard errors clustered at the zip code are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table A7: Placebo Tests: Respiratory Medicine

	Child Cold Remedies (1)	Sinus Remedies (2)	Sudafed, Nyquil/Dayquil (3)
Pct Acute Coliform	0.00502	-0.0492	-0.00602
	(0.0151)	(0.0482)	(0.0180)
Pct Monthly Coliform	-0.0319	-0.0241	-0.0194
	(0.0223)	(0.0623)	(0.0120)
Observations	245,632	245,632	245,632
R-squared	0.191	0.186	0.149
Household FE	yes	yes	yes
Year-month FE	yes	yes	yes
Zip-yr FE	yes	yes	yes
Zip-month FE	yes	yes	yes

Source: Nielsen household-month level data.

Notes: All regressions include controls for weather, employment rate, and household size. The placebo tests using medications for respiratory illness include: "Cold Remedies - Children," "Sinus Remedies," and purchases of "Sudafed", "Nyquil" or "Dayquil." The outcomes are measured as the inverse hyperbolic sine of the total purchase of each category in dollars. Standard errors clustered at the zip code level are shown in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table A8: Zip-code level ER visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pct Acute Coliform	1.357***	0.656	0.476	0.669	0.237	0.0360	-0.0176	0.0907
	(0.502)	(0.438)	(0.892)	(0.948)	(0.956)	(0.875)	(0.908)	(0.867)
Pct Monthly Coliform	0.562**	-0.127	0.400*	0.816***	0.717***	0.738***	0.670***	0.760***
	(0.220)	(0.221)	(0.219)	(0.270)	(0.262)	(0.254)	(0.255)	(0.274)
Observations	82,242	82,242	82,242	82,242	82,242	82,032	82,032	82,032
R-squared	0.212	0.452	0.544	0.604	0.622	0.631	0.626	0.633
Year	yes	yes	yes	yes	yes	yes	yes	yes
Month	yes	yes	yes	yes	yes	yes	yes	yes
Zip		yes	yes	yes	yes	yes	yes	yes
Zip-yr			yes	yes	yes	yes	yes	yes
Zip-month				yes	yes	yes	yes	yes
Year-month					yes	yes	yes	yes
DEQ Region-yr-month						yes		yes
Geo Region-yr-month							yes	yes

Source: ER visit zip code-month level data.
Notes: All regressions include controls for weather and influenza and pneumonia visits. Column 6 uses regions defined by North Carolina's Department of Environmental Quality and column 7 uses the Coastal Plains, Piedmont, and Mountains geographic regions. Regressions are weighted by population and standard errors clustered at the zip code level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A9: ER admission placebos: Age 5-19

	Fractures (1)	Burns (2)	Poisoning (3)	Acute Resp. (4)	Nervous System (5)
Pct Acute Coliform	0.245	-0.0932	-0.00642	-0.0609	-0.699
	(0.884)	(0.132)	(0.156)	(0.932)	(0.721)
Pct Monthly Coliform	-0.143	0.0945	-0.00611	0.410	0.177
	(0.358)	(0.0674)	(0.0426)	(0.288)	(0.320)
Observations	82,242	82,242	82,242	82,242	82,242
R-squared	0.618	0.217	0.216	0.724	0.594
Zip FE	yes	yes	yes	yes	yes
Year-month	yes	yes	yes	yes	yes
Zip-yr	yes	yes	yes	yes	yes
Zip-month	yes	yes	yes	yes	yes
Mean	5.129	0.145	0.147	4.357	3.115

Source: ER visit zip code-month level data.
Notes: All regressions are weighted by population and include controls for weather and influenza and pneumonia ER visits. Outcomes are emergency room visits per 1,000 population. Standard errors clustered at the zip code level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A10: School-level Monthly Absences and the Impact of Water Quality Violations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pct Acute Coliform	-0.00612	-0.0193	-0.0240	-0.0265	-0.0274	0.00696	0.0212***	0.00255	0.0200**
	(0.0215)	(0.0218)	(0.0202)	(0.0204)	(0.0192)	(0.00545)	(0.00623)	(0.00578)	(0.00939)
Pct Monthly Coliform	0.253***	0.234***	0.212***	0.195***	0.222***	0.131**	0.122**	0.136**	0.125**
	(0.0450)	(0.0451)	(0.0458)	(0.0402)	(0.0419)	(0.0571)	(0.0577)	(0.0572)	(0.0584)
Observations	150,678	150,678	150,678	150,678	150,667	150,594	150,594	150,594	150,594
R-squared	0.312	0.362	0.548	0.578	0.678	0.760	0.772	0.763	0.773
Year	yes	yes	yes	yes	yes	yes	yes	yes	yes
Month	yes	yes	yes	yes	yes	yes	yes	yes	yes
School char.		yes	yes	yes	-	-	-		
School			yes	yes	yes	yes	yes	yes	yes
Year-month				yes	yes	yes	yes	yes	yes
School-yr					yes	yes	yes	yes	yes
School-month						yes	yes	yes	yes
DEQ Region-yr-month							yes		yes
Geo Region-yr-month								yes	yes

Source: PMR school-grade-month level data.

Notes: All specifications are weighted by the number of students matched to a water supply system and include grade dummies and weather controls. School characteristics at the school-year level include percent by racial category, title I status, and percent receiving free/reduced price lunch. Column 7 uses regions defined by North Carolina's Department of Environmental Quality and column 8 uses the Coastal Plains, Piedmont, and Mountains geographic regions. Standard errors clustered at the school level are shown in parenthesis.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A11: School-level Absences: Robustness

	Baseline (1)	Employment rate (2)	MR violations (3)	Flu and Asthma (4)	Air Pollution (5)	Drop Wake & Mecklenburg (6)	Drop USCs (7)	Drop Magnet (8)	Drop September (9)	School -level (10)	Drop bad PMR -month match (11)	Rain bins (12)
Pct Acute Coliform	0.00696	0.00843	0.00580	0.00668	0.00717	0.0111**	0.00694	0.00777	0.00785	0.00801	0.00451	0.00687
	(0.00545)	(0.00551)	(0.00499)	(0.00543)	(0.00543)	(0.00529)	(0.00545)	(0.00544)	(0.00541)	(0.00511)	(0.00550)	(0.00536)
Pct Monthly Coliform	0.131**	0.130**	0.133**	0.132**	0.130**	0.138**	0.132**	0.132**	0.135**	0.141**	0.151***	0.132**
	(0.0571)	(0.0571)	(0.0576)	(0.0575)	(0.0574)	(0.0574)	(0.0571)	(0.0569)	(0.0561)	(0.0588)	(0.0561)	(0.0570)
Observations	150,594	150,594	150,594	150,594	150,594	122,253	150,302	145,023	133,518	38,794	129,180	150,594
R-squared	0.760	0.760	0.760	0.760	0.760	0.749	0.760	0.762	0.744	0.868	0.763	0.760

Source: PMR school-grade-month level data.

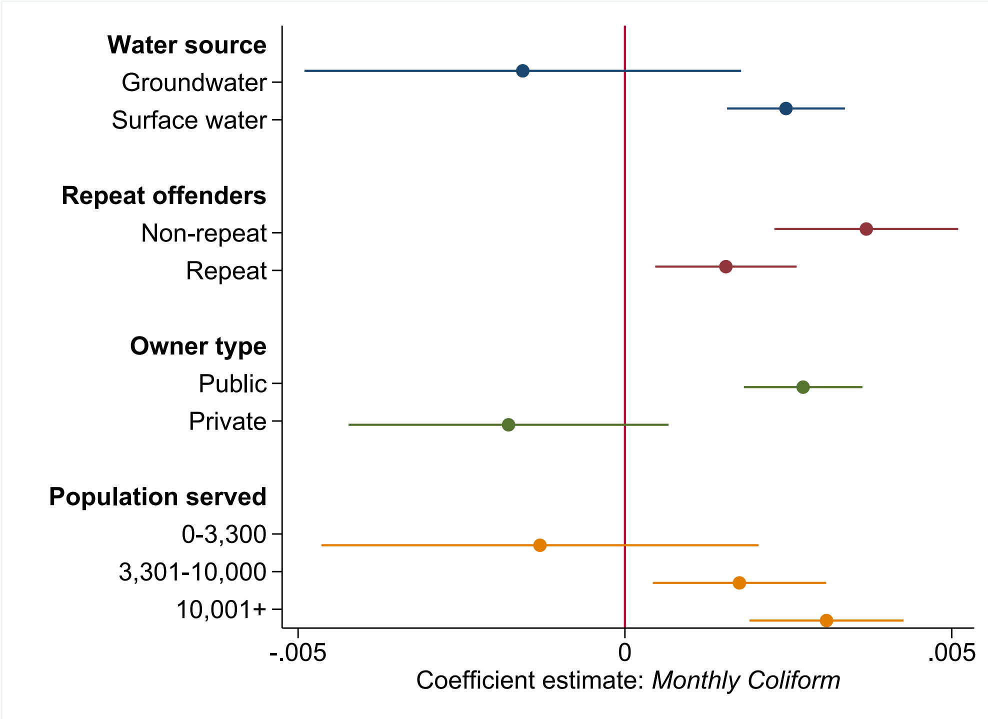
Notes: All specifications are weighted by the number of students matched to a water supply system and include controls for grade, school, year, month, year-month, school-year, school-month, and weather. Columns 2-4 add controls for county-month employment rate, monitoring and reporting violations, and influenza and asthma ER admissions. Columns 5 includes controls for OZ and PM10 as the percent of days within each of five categories to allow for nonlinear effects: 0-25%, 25-50%, 50-75%, 75-100%, and over 100% of the relevant EPA threshold. Columns 6-9 drop schools in Wake and Mecklenburg Counties, schools with water-related unplanned school closures, magnet schools, and September, respectively. Column 10 collapses the data to the school level. Column 11 drops schools with PMR intervals that cover less than 70% of days in the calendar months. Column 12 adds controls for 7 precipitation bins. Standard errors clustered at the school level are shown in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A12: Student-level Absences and the Impact of Water Quality Violations

	Home Exposure			School Exposure
	(1)	(2)	(3)	(4)
Acute Coliform	-0.0308	-0.180***	-0.0174	-0.231***
	(0.0585)	(0.0586)	(0.0543)	(0.0560)
Monthly Coliform	0.378***	0.331***	0.208***	0.239***
	(0.0390)	(0.0401)	(0.0323)	(0.0313)
Radionuclides	-0.0534***	-0.0589***	0.00946	0.0668***
	(0.0139)	(0.0135)	(0.0151)	(0.0219)
Disinfect	0.0372***	0.0220***	-0.00909***	-0.00853***
	(0.00252)	(0.00255)	(0.00262)	(0.00264)
SOCs	0.0525**	0.0523**	0.00646	0.0347*
	(0.0263)	(0.0264)	(0.0185)	(0.0203)
VOCs	-0.0133	-0.00688	0.00213	0.0103
	(0.0169)	(0.0172)	(0.0140)	(0.0158)
Observations	3,053,192	2,967,358	2,564,788	2,420,690
R-squared	0.020	0.027	0.651	0.656
Grade	yes	yes	yes	yes
Year	yes	yes	yes	yes
Demographics		yes	-	-
Weather		yes	yes	yes
Student FE			yes	yes
Cluster			student	student

Notes: Standard errors are clustered at the individual level in columns 3 and 4. Coefficients and standard errors are scaled by 100. *** p<0.01, ** p<0.05, * p<0.1

Figure A6: Student-level absences by CWS characteristics



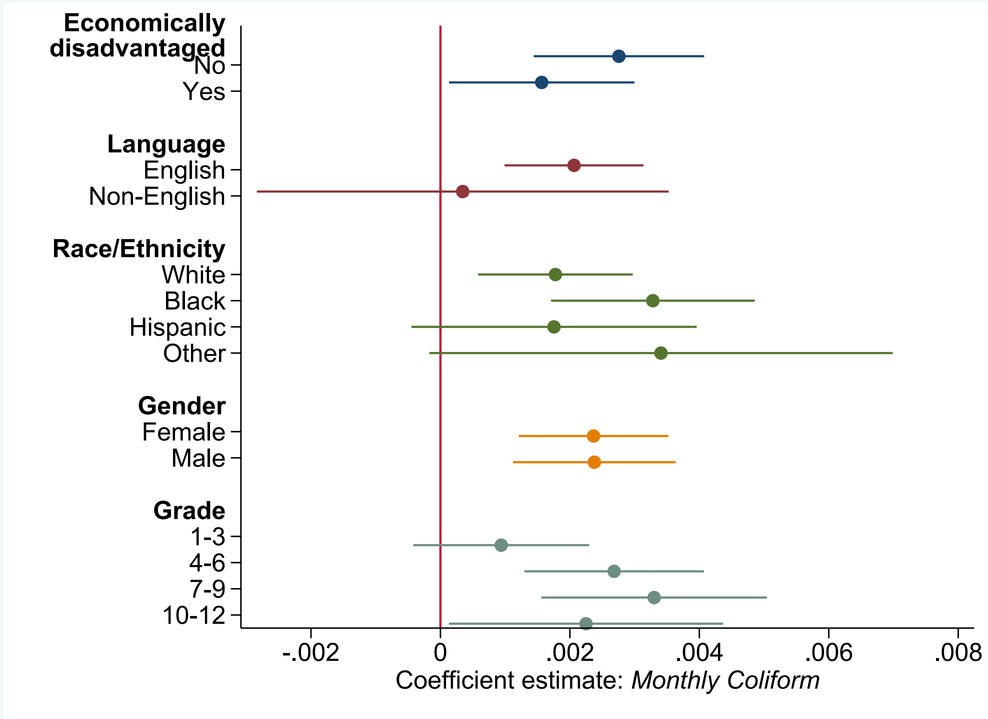
Notes: Estimated effects of monthly coliform violations on absences are shown separately by water source, whether the CWS has multiple violations, ownership type, and size. Regressions include individual fixed effects and controls for acute coliform, grade, year, and weather. Standard errors are clustered at the individual level.

Table A13: Student-level Absences by CWS Characteristics

	(1)	(2)	(3)	(4)
Groundwater × Monthly Coliform	-0.151			
	(0.170)			
Surface water × Monthly Coliform	0.243***			
	(0.0460)			
Non-repeat × Monthly Coliform		0.367***		
		(0.0717)		
Repeat × Monthly Coliform		0.152***		
		(0.0552)		
Public × Monthly Coliform			0.269***	
			(0.0463)	
Private × Monthly Coliform			-0.168	
			(0.126)	
Size: 0-3,300 × Monthly Coliform				-0.140
				(0.171)
Size: 3,301-10,000 × Monthly Coliform				0.174***
				(0.0676)
Size: 10,001+ × Monthly Coliform				0.305***
				(0.0602)
Observations	2,951,020	2,967,567	2,967,567	2,967,567
R-squared	0.723	0.724	0.724	0.724

Source: NCERDC student-year level data.
Notes: All regressions include fixed effects and controls for acute coliform, grade, year and weather. Standard errors are clustered at the individual level. Coefficients and standard errors are scaled by 100. *** p<0.01, ** p<0.05, * p<0.1

Figure A7: Student-level absences by demographics



Notes: Estimated effects of monthly coliform violations on absences are shown separately by economic status, native language, race/ethnicity, gender, and grade. Regressions include individual fixed effects and controls for acute coliform, grade, year, and weather. Standard errors are clustered at the individual level.

Table A14: Student-level Absences by Demographics

	(1)	(2)	(3)	(4)	(5)
Non-Disadvantaged × Monthly Coliform	0.279***				
	(0.0672)				
Economically Disadvantaged × Monthly Coliform	0.152**				
	(0.0731)				
Proficient English × Monthly Coliform		0.202***			
		(0.0549)			
Limited English × Monthly Coliform		0.0392			
		(0.162)			
White × Monthly Coliform			0.177***		
			(0.0610)		
Black × Monthly Coliform			0.325***		
			(0.0803)		
Hispanic × Monthly Coliform			0.171		
			(0.112)		
Other × Monthly Coliform			0.333*		
			(0.183)		
Female × Monthly Coliform				0.234***	
				(0.0591)	
Male × Monthly Coliform				0.236***	
				(0.0641)	
Grades 1-3 × Monthly Coliform					0.0923
					(0.0694)
Grades 4-6 × Monthly Coliform					0.266***
					(0.0709)
Grades 7-9 × Monthly Coliform					0.324***
					(0.0889)
Grades 10-12 × Monthly Coliform					0.225**
					(0.108)
Observations	1,980,554	2,143,522	2,967,565	2,967,566	2,967,567
R-squared	0.733	0.770	0.724	0.724	0.724

Source: NCERDC student-year level data.

Notes: All regressions include individual fixed effects and controls for acute coliform, grade, year, and weather. Standard errors are clustered at the individual level. Coefficients and standard errors scaled by 100. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A15: Student-level Absences: Robustness

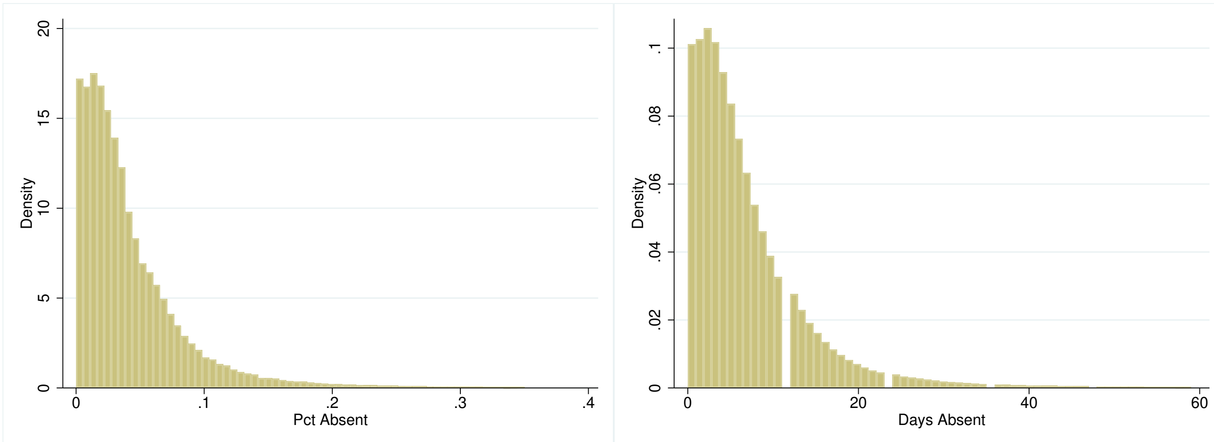
	Baseline (1)	Employment rate (2)	MR violations (3)	Flu and Asthma (4)	Air Pollution (5)	Drop Wake & Mecklenburg (6)	Drop USCs (7)	Regular Schools (8)	Regular Programs (9)	Traditional Calendars (10)	Rain Bins (11)	Monthly Weather (12)	DEQ-Region -by-year FE (13)	Region -by-year FE (14)
Acute Coliform	-0.0159	-0.0251	-0.0410	-0.0193	0.0109	0.00432	-0.0156	-0.0906	-0.0907	-0.0741	-0.0344	0.0188	-0.0343	0.0203
	(0.0681)	(0.0681)	(0.0680)	(0.0680)	(0.0685)	(0.0686)	(0.0681)	(0.0671)	(0.0696)	(0.0712)	(0.0683)	(0.0736)	(0.0563)	(0.0550)
Monthly Coliform	0.235***	0.227***	0.161***	0.217***	0.189***	0.232***	0.235***	0.197***	0.223***	0.228***	0.213***	0.194***	0.181***	0.229***
	(0.0438)	(0.0438)	(0.0440)	(0.0438)	(0.0439)	(0.0442)	(0.0438)	(0.0434)	(0.0455)	(0.0446)	(0.0439)	(0.0449)	(0.0352)	(0.0351)
Observations	2,967,567	2,967,288	2,967,567	2,965,595	2,967,567	2,142,922	2,965,950	2,893,546	2,462,414	2,753,990	2,967,567	2,967,567	2,562,942	2,562,942
R-squared	0.724	0.724	0.724	0.724	0.724	0.734	0.724	0.729	0.746	0.733	0.724	0.724	0.651	0.651

Source: NCERDC student-year level data.

Notes: All specifications include individual fixed effects and controls for grade, year, and weather. Columns 2-4 include controls for the employment rate, monitoring and reporting violations, and influenza and asthma ER admissions. Column 5 includes controls for OZ and PM10 as the percent of days within each of five categories to allow for nonlinear effects: 0-25%, 25-50%, 50-75%, 75-100%, and over 100% of the relevant EPA threshold. Columns 6 and 7 drop schools in Wake and Mecklenburg Counties and schools with water-related unplanned school closures. Columns 8-10 limit the sample to regular schools, regular programs, and schools with traditional calendars, respectively. Column 11 adds controls for 7 precipitation bins and column 12 includes temperature and precipitation bins separately for each month. Columns 13 and 14 add region-by-year fixed effects, where regions are based on North Carolina's Department of Environmental Quality in column 13 and the Coastal Plains, Piedmont, and Mountain regions in column 14. Standard errors clustered at the individual level are shown in parenthesis. Coefficients and standard errors are scaled by 100. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure A8: Distribution of Student-level Absences

(a) Pct Absent (b) Days Absent



Notes: Figure shows the distribution of the absence rate in panel a and the number of days absent in panel b.

Table A16: Student-level Absences: Alternate Specifications

	Days Absent	
	Poisson	Neg. Binomial
	(1)	(2)
Acute Coliform	-0.00515	-0.0133
	(0.0114)	(0.0112)
Monthly Coliform	0.0457***	0.0471***
	(0.00732)	(0.00581)
Observations	2,547,121	2,547,121
Number of students	668,541	668,541
Student FE	yes	yes

Source: NCERDC student-year level data.
Notes: All specifications include controls for grade, year, weather, and days in membership. The outcome of interest is days absent. Bootstrapped standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

B Water Regulation Detail

B.1 SDWA and Public Notification

The Safe Drinking Water Act (SDWA) authorizes the US Environmental Protection Agency (EPA) to set national health-based standards to protect the public against both naturally occurring and man-made contaminants. The SDWA applies to all public water systems in the US. Public water systems are defined as having at least 15 service connections or serving at least 25 people per day for 60 days of the year. There are currently over 170,000 public water systems in the US that provide water to almost all individuals in the country. The study focuses on community water systems, which supply water to the same population year-round. Other types of public drinking water systems include transient non-community water systems (TNCWS) and non-transient non-community water systems (NTNCWS). A public water system that regularly supplies water to at least 25 of the same people at least six months per year is considered a NTNCWS (e.g., water systems that supply schools, factories, office buildings, and hospitals). A public water system that provides water in a place such as a gas station or campground where people do not remain for long periods of time is considered a TNCWS. About 14% of people in the U.S. supply their own water for domestic use, primarily through groundwater wells. There are no federally required monitoring and treatment standards for private wells under the SDWA.

Through the National Primary Drinking Water Regulations the EPA has set enforceable maximum contaminant levels (MCL) for over 90 different contaminants, including microorganisms, disinfectants, disinfectant by-products, inorganic chemicals, organic chemicals, and radionuclides.

Originally, SDWA focused on water treatment as the primary means of providing safe drinking water, but amendments to the law in 1996 introduced source water protection, operator training, funding for water system improvements, and the provision of public information. These amendments introduced the Public Notification Rule, which specifies when and how the public receives information about violations.This rule is required across all states. In North Carolina, the revised Public Notification Rule became effective May 6, 2002, prior to the sample period studied here.

Each notification must include a description of the violation and contaminant levels, the date the violation occurred, potential adverse health risks, a description of the population at risk, whether alternative water supply should be used, what action consumers should take, what the system is doing to correct the violation, and when the system expects to return to compliance. The notification must also include contact information and a statement encouraging recipients to distribute the notice to other individuals served. If 30 percent or more of the public water supply customers are non-English speaking, the system must provide the notification in the appropriate language(s) or provide information on how to get assistance or a translated copy.[27]

B.2 Total Coliform Rule: Sampling and Testing

The number of samples required each month depends on the size of the population served; larger systems are required to sample more often. Samples must be collected at regular time intervals throughout the month at sites that are representative of water quality throughout the distribution system. Sampling locations are based on a written sample siting plan that is subject to state review and revision. Groundwater systems serving 4,900 or fewer people may collect their samples on the

[27] Example notifications are available on the North Carolina Department of Environmental Protection website: `https://deq.nc.gov/public-notification-rule-tier-levels`https://deq.nc.gov/public-notification-rule-tier-levels.

same day.

Samples are tested for the presence or absence of total coliforms or fecal coliforms, rather than the amount or concentration of coliforms present. If a sample tests positive for total coliforms, the system is typically required to take three "repeat" samples: at the same tap, within five service connections upstream, and within five service connections downstream. If the "repeat" samples test positive for total coliforms, another set of samples must be taken, as before, unless a violation has been triggered. In addition, each sample that tests positive for total coliforms must also be tested for the presence of fecal coliforms or E. coli.

C Data Appendix

C.1 Additional Data Detail

Geographic Information

North Carolina Department of Public Instruction (NCDPI) provides the Public School Location data. County and 5-digit zip code boundaries come from the U.S. Census Bureau's Cartographic Boundary Files. There are 100 counties and 1,080 5-digit zip codes in North Carolina.

Emergency Department Data

I record patients as entering the ER with a gastrointestinal illness based on all 18 diagnosis codes provided in the data. Visits to the ER for a gastrointestinal illness are restricted to those with the following ICD-9 codes: 001-009 (Intestinal Infectious Diseases), 558.9 (Noninfectious Gastroenterit NEC), 787.0 (Nausea and vomiting), and 787.9 (Diarrhea). I exclude newborn admissions and visits with a diagnosis of *Clostridium difficile* (008.45), as it is primarily considered to be a hospital-acquired infection.

As a control variable, I calculate the number of admissions to the ER for influenza and pneumonia (ICD-9 codes: 480-488). I also create a number of placebo outcomes using admissions for fractures (ICD-9 codes: 800-869), burns (ICD-9 codes: 940-949), poisoning by drugs and medicinal and biological substances (ICD-9 codes: 960-979), acute respiratory infections (ICD-9 codes: 460-466), and diseases of the nervous system and sense organs (ICD-9 codes: 320-389).

Absences and School Data

Additional school-level characteristics come from Common Core of Data (CCD) Public Elementary/Secondary School Universe Survey Data. These data contain descriptive statistics on students at all schools on a yearly basis, including Title I eligibility, free and reduced-price lunch eligibility, and demographic composition of the student body. Other basic information includes type of school calendar (traditional, year-round, modified, etc.), school program (magnet, charter, etc.), and location type (rural, suburban, etc.).

Employment

Employment rates at the county-month level come from the Bureau of Labor Statistics' Local Area Unemployment Statistics (LAUS) data. For the student-level analysis, employment rates are collapsed to the county-school year level.

C.2 PMR Intervals

The following two figures test the assumption that PMR Intervals 1-9 cover months September-May. Using school calendar information and the reported length of each PMR interval (ranging from 16 to 26 days), I calculate the approximate start and end date of each PMR interval to compare with the assumed start and end dates (first and last day of months September-May).

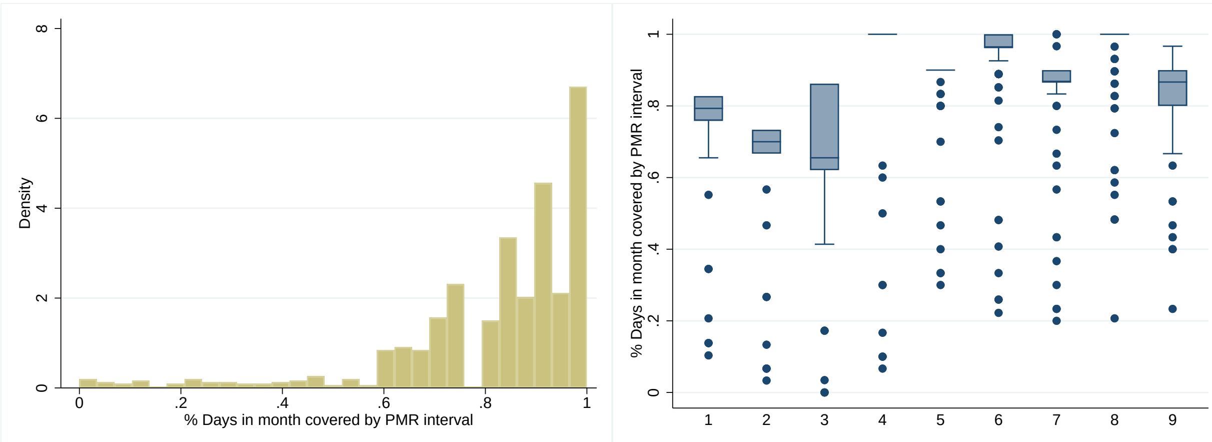
For each Local Education Agency (LEA) in the sample, I identify the dates of the first day of school, spring break, and the last day of school for school years 2012-13 to 2014-15 from the Myrtle Beach Area Chamber of Commerce. Unfortunately, this information is not available for earlier years. For the 4 LEAs where these dates were unavailable, I assumed their spring break fell on the most common dates for the rest of the sample and that the first day of school was August 27, since North Carolina law has required public schools to open on the Monday closest to August 26 since 2013. For all schools, I assume winter break begins December 21 and ends January 1 (inclusive). Finally, I account for the fact that school is not in session each year during the following holidays: Labor Day, Veterans Day, Thanksgiving, and Memorial Day. Note that teacher work days and weather-related school closures are not taken into account here.

Overall, there is a high overlap between PMR intervals and calendar months. Figure A9 below shows the percentage of days in each month covered by the PMR interval associated with that month. On average, PMR intervals cover 82% of the days in their assigned month. Figure A9b shows coverage by each PMR interval. For example, the first PMR interval covers about 76% of the days in September, on average. Coverage is better starting in December, with PMR interval 4 covering 93% of the days in December. The school-grade-month attendance results presented above are robust to excluding LEAs with lower than 70% average coverage.

Furthermore, the misalignment between PMR intervals and calendar months is very low. Figure A10 shows the difference between PMR start/end dates and the first/last day of each month. Error is calculated as the start (end) date of the month minus the start (end) date of the PMR interval. Therefore a positive start (end) error indicates that the PMR interval started (ended) before the first (last) day of the month. On average, the PMR interval started 4 days before the first day of the month and ended 4 days before the last day of the month. Figure A10b shows the start and end error by each PMR interval. As before, we can see that there is greater error in the first 3 months. Over the first 3 months, the average start error is 8 days and the average end error is 9 days, while the average start error and end error for the remaining months is 3 days and 1 day, respectively. These patterns likely reflect the fact that the typical school year starts during the last week of August, rather than the first week of September.

Figure A9: Percent of assumed month covered by PMR Interval

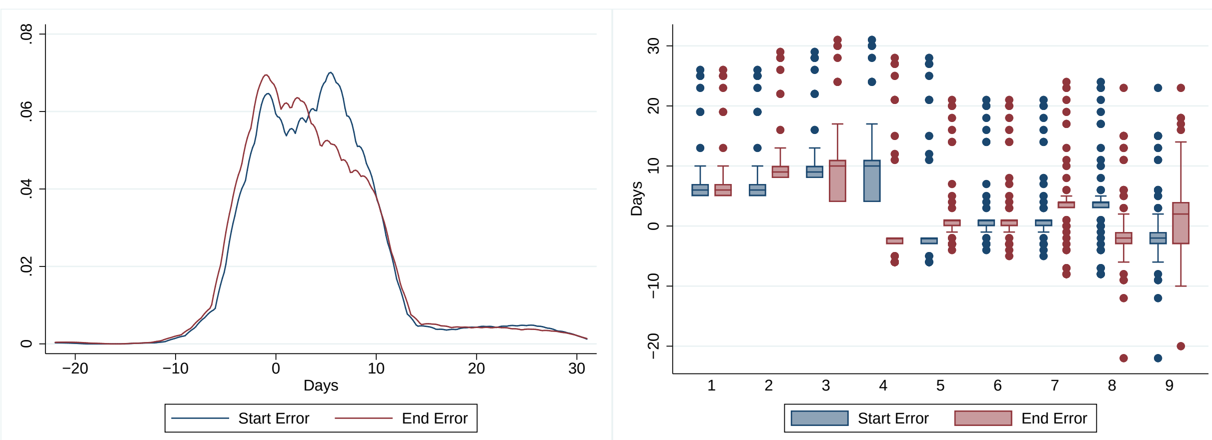
(a) Distribution (b) By PMR Interval



Notes: Percent covered is the percentage of the days in each month (Sept-May) covered by each PMR interval, where PMR interval 1 is assigned to September and other months are assigned sequentially.

Figure A10: Error in PMR interval start/end dates

(a) Distribution (b) By PMR interval



Notes: Error is calculated as the start (end) date of the month minus the start (end) date of the PMR interval. Therefore a positive start (end) error indicates that the PMR interval started (ended) before the first (last) day of the month.