

Hyde, Kelly

**Working Paper**

## The Regressive Costs of Drinking Water Contaminant Avoidance

GLO Discussion Paper, No. 703

**Provided in Cooperation with:**

Global Labor Organization (GLO)

*Suggested Citation:* Hyde, Kelly (2020) : The Regressive Costs of Drinking Water Contaminant Avoidance, GLO Discussion Paper, No. 703, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/225634>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The Regressive Costs of Drinking Water Contaminant Avoidance

Kelly Hyde

University of Pittsburgh<sup>1</sup>

## Abstract

Up to 45 million Americans in a given year are potentially exposed to contaminated drinking water, increasing their risk of adverse health outcomes. Existing literature has demonstrated that individuals respond to drinking water quality violations by increasing their purchases of bottled water and filtration avoidance, thereby avoiding exposure to contaminants. This paper demonstrates that poorer households, for whom the costs of avoidance comprise a greater share of disposable income, bear disproportionate costs of water quality violations in the United States. Following a health-based water quality violation, poor households' expenditure on nutritious grocery products in a nationally representative panel differentially decreases by approximately \$7 per month. This is associated with a decrease of about 1,500 calories per household member per day, placing these individuals at a higher risk of food insecurity. This finding suggests that the indirect costs of drinking water contamination through economic channels exacerbate health disparities associated with poverty.

## 1 Introduction

According to the Environmental Protection Agency's budget proposal for fiscal year 2020, over 7% of the United States' population served by community water systems receives water that does not meet at least one health-based drinking water standard established by the Safe Drinking Water Act of 1974—at least the ones we know about. Monitoring and enforcing water quality standards is a massive undertaking, requiring periodic sample collection at the almost 400,000 public water systems registered in the EPA's Safe Drinking Water Information System (SDWIS), testing by certified labs for a broad panel of regulated contaminants, and legal proceedings in the event of an unresolved violation. Given the many layers of coordination required on a national scale, it is

---

<sup>1</sup>4530 Wesley W. Posvar Hall, 230 S Bouquet St, Pittsburgh, PA 15260; email kdh50@pitt.edu.

**Acknowledgements:** I thank Olena Nizalova and the organizers and participants of the Global Labor Organization's 2019-2020 Virtual Young Scholars (VirtYS) program for their helpful comments.

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.

unsurprising that an estimated 26-38% of violations are either unreported or erroneously recorded in SDWIS (Allaire et al. 2018, United States Environmental Protection Agency 2000), and many of those which are accurately recorded take several months to return to compliance. According to the EPA, serious health-based violations are “expected to be resolved within six months” (United States Environmental Protection Agency 2020); however, many are not. For example, during the high-profile violation of water quality standards in Flint, Michigan, when the maximum contaminant level for trihalomethanes was exceeded in December 2014, return to compliance was not achieved until 9 months later.<sup>2</sup>

As customers wait for the supplier to fix the problem, either voluntarily or following legal action by the state or federal government, many purchase bottled water or filtration devices in order to avoid contaminants (Allaire et al. 2019, Zivin et al. 2011). Expenditure on avoidance methods has been used in the environmental economics literature to estimate the willingness to pay (WTP) for water quality improvements (Brouwer et al. 2015, Brox et al. 2003, Johnston and Thomassin 2010, Rodriguez-Tapia et al. 2017) and similarly for air quality improvements (Freeman et al. 2019). Prior literature using survey-based elicitation methods has estimated average WTP for water quality improvements of between \$5 and \$15 per month in 2020 U.S. Dollars (Brox et al. 2003).

While enforcement of water quality standards is publicly funded, the costs of avoidance are privately borne, and do not scale with income in the absence of targeted subsidies. Thus the need to avoid contaminated water is a regressive income shock. The amount of water required for survival is not a function of income or preferences, so a wealthy household and a poor household with the same number of occupants would need to obtain similar amounts of bottled water to replace their tap water consumption. However, the cost of this bottled water has a different impact on each household’s budget—the wealthy household, unlikely to be budget-constrained, can add bottled water purchases onto their existing consumption patterns, while the poor household is much more likely to need to forego other purchases to fit the cost of avoidance into their budget. Thus poorer households face a difficult trade-off: either maintain current consumption and risk exposure to contaminated water, or purchase avoidance and make sacrifices elsewhere.

In this paper, I show that during an active water quality violation in SDWIS, monthly expenditure on groceries declines among households with income below 200% of the federal poverty line by an average of \$7.23 per month, but is unchanged for those with higher income. As a consequence, purchased calories per household member declines by almost 1,500 per month, just under the average recommendation of 2,000 calories per day. An event study specification confirms that this decline accumulates over the duration of an active violation, resulting in larger effects for violations that take longer to return to compliance. The estimated declines in household grocery consumption coincide with increases in demand for bottled water and filtration devices which are proportionally

---

<sup>2</sup>Perhaps surprisingly, a water quality violation associated with lead levels never appeared in SDWIS, even though the Flint water crisis is most commonly associated with lead in the water. This is likely in part because lead typically enters the household water supply through the corrosion of pipes near the household, rather than contamination occurring earlier in the distribution network.

larger for poor households.

This finding contributes to multiple literatures on the economics of natural resources, poverty, and nutrition. Improvements in potable water supply have been linked to increases in happiness (Devoto et al. 2012), decreases in obesity risk (Ritter 2019), decreases in overall mortality risk (Clay et al. 2014, Ferrie and Troesken 2008, Troesken 2004), and increases in human capital accumulation (Ao 2016, Beach et al. 2016, Troesken et al. 2011). These outcomes have also been linked to improvements in nutrition, especially among infants and children (Adhvaryu et al. 2019, Anderson et al. 2016, Frisvold 2015, Hoynes et al. 2011, 2015, Hoynes and Schanzenbach 2009, Kohler et al. 2017). The findings in this paper suggest a potential link between these parallel literatures: in addition to the direct benefits of water supply improvements, poor households may additionally benefit from a positive income effect as improvements to water supply allow them to reallocate their spending on avoidance devices to other nutritious products. Taking this income effect into account increases the long-run value of residential water infrastructure improvements, for which the EPA requested over \$2 billion in fiscal year 2020 citing broad prevalence of outdated water systems.

This paper also contributes to the literature on the spatial and intergenerational nature of poverty in the United States. The notion that “zipcode is destiny” (Chetty et al. 2018) may be partially explained by spatial heterogeneity and persistence in the quality of the residential water supply. For several reasons, including the construction of lead pipes in the early 20th century (Clay et al. 2014), the location of cities either upstream or downstream of major sources of pollution (Jerch 2018), and historical to present-day environmental injustice (Schneider et al. 2019, Switzer and Teodoro 2018), water quality violations are not evenly distributed across space. Particularly for children, since many of the known health and developmental conditions caused by contaminated water supply are associated with early-life exposure (Beach et al. 2016, Clay et al. 2014, Valent et al. 2004), living in an area with persistent water quality violations increases the risk of conditions that in turn reduce the likelihood of getting out of poverty. If poor water quality requires families to choose between avoidance and adequate nutrition, children are placed in a lose-lose situation, as either choice is likely to have negative consequences for their development. While this paper focuses on the short-term impacts of water quality violations to limit threats to causal identification, the results are consistent with correlational disparities in food security based on an area’s history of water quality violations (see Figure 1).

The paper proceeds as follows. Section 2 describes the data used to construct the water quality violation and household consumption measures. Section 3 describes the results. Section 4 discusses the limitations and areas for further research. Section 5 concludes.

## 2 Data and Empirical Strategy

### 2.1 Data

To assess the differential effect of water quality violations on poor households’ budget constraints and resulting nutrition, this paper combines household-level panel data on grocery and department

store purchases, nutritional information on grocery products, and administrative records of water quality violations and enforcement activities. Each data source is described below.

*Household consumption data.* Measures of monthly household consumption were constructed from the Nielsen HomeScan Consumer Panel, a nationally representative panel of households’ retail purchases, from 2004 to 2016. The panel contains 168,772 unique households spanning 2,967 counties in the United States, and includes purchase dates, quantities, and prices paid for about 4.9 million unique Universal Product Codes (UPCs). The demographic information provided about each household includes zip code of primary residence, number of residents, annual income, and the gender, race, age, and occupation of each resident. Whether or not a household’s income exceeds 200% of the federal poverty line was calculated based on the reported household income and number of residents, including children.

*UPC-level nutrition facts data.* The UPC-level consumption data in the Nielsen HomeScan Consumer Panel was merged with the FoodData Central database of nutrition facts for over 300,000 grocery products sold in the United States (U.S. Department of Agriculture, Agricultural Research Service 2019). FoodData Central reports all information included in a product’s legally mandated nutrition facts label, including calories, fat, carbohydrates, sugars, and protein per serving. While the majority of the most frequently purchased branded products in the Nielsen HomeScan Consumer Panel are represented in the FoodData Central data, these only comprise about 13% of all unique UPCs purchased by households between 2004 and 2016. Thus to obtain a more complete representation of household nutrition, the average nutritional content per serving for each of the 1,227 product modules specified by Nielsen (e.g. “Fruit Drinks and Juices - Cranberry,” “Soft Drinks - Carbonated,” “Granola & Yogurt Bars”) was calculated and imputed for products not included in the FoodData Central database.

*Water quality violations and enforcement data.* The incidence, timing, and type of water quality violations were determined at the county level<sup>3</sup> from the EPA Safe Drinking Water Information System (SDWIS) database, an administrative dataset that records public water system facility locations and populations served, site visit logs, and various types of water quality standard violations. Health-based violations, which are failures to adhere to established maximum contaminant levels (MCLs) for regulated contaminants such as lead, arsenic, and nitrates or related treatment protocols, are the focus of this paper. SDWIS also separately categorizes some health-based violations as acute when the violated MCL poses an immediate threat to customers’ health.

I classify household  $i$  in county  $j$  as under an “active violation” in month  $t$  if there exists a health-based water quality violation recorded in SDWIS for a water supplier that serves at least 500 customers<sup>4</sup> in county  $j$  for which public notification has been requested by state or federal authorities prior to month  $t$  and compliance has not yet been achieved by month  $t$ . Because each

---

<sup>3</sup>For a zip code based specification which trades off sample size for enhanced geographic specificity, see Appendix Section A.

<sup>4</sup>This eliminates “Very Small” water suppliers, which include restaurants, office buildings, and other public institutions that are required to report to state or federal water quality monitors but do not provide water to households on a large enough scale for households in the panel to plausibly be affected. This cutoff was also used in Allaire et al. (2018), citing different reporting requirements and less reliable data from “Very Small” systems.

individual household’s water supplier is not observed, there is uncertainty regarding whether or not household  $i$  is actually affected by the violation. The results should thus be interpreted as “intent-to-treat” rather than the “treatment on the treated.” Because entirely unaffected households should not respond to these violations, this attenuates estimates, and thus the estimated effects in this paper should be considered a lower bound.<sup>5</sup>

For both the panel fixed-effects and event study specifications, I restrict the sample in three ways. First, I only include households that experience at least one health-based water quality violation during the sample period, which comprise about 35% of the overall consumer panel sample. As is evident in Figure 1, many counties do not have any active violations during the sample period. Those that do are a selected subsample of counties that are more rural and lower-income than average, meaning that comparisons along the extensive margin of water quality violations may be confounded by unrelated differences, especially if those differences are time-varying. Instead, this sample restriction means that the differential timing of water quality violations across counties is used to identify the effect of those violations on household consumption. Second, I restrict to households for which the classification of above or below 200% of the federal poverty line does not change during the sample period to avoid biased estimates from large increases or decreases in a household’s income over time that may coincide with water quality violations. Finally, I only count violations as “active” if they have *both* relevant dates recorded in SDWIS (the date public notification was requested and the date of return to compliance).

## 2.2 Empirical Strategy

### 2.2.1 Panel Fixed-Effects Regression

Using the above data sources, the following panel fixed-effects model was estimated:

$$TotalExp_{it} = \delta_1 ActiveViolation_{it} \times Below200PctFPL_i + \beta_1 ActiveViolation_{it} + i + t + \epsilon_{it} \quad (1)$$

where  $Below200PctFPL_i$  is a dummy that equals 1 when the household’s total income does not exceed 200% of the federal poverty line based on the household’s number of residents,  $ActiveViolation_{it}$  is a dummy variable that equals 1 when household  $i$  is under an active water quality violation defined above,  $i$  and  $t$  are vectors of fixed effects for each household  $i$ <sup>6</sup> and time period  $t$  respectively, and  $\epsilon_{it}$  is an error term. Standard errors are clustered at the county level because the treatment varies at the county level.

To further investigate the potential causes and consequences of differential changes in total expenditure between wealthier and poorer households, the model in equation 1 is re-estimated with

---

<sup>5</sup>It is possible for households which are not directly affected by a water quality violation to respond if they hear about a violation in their area and thus become more skeptical of their own water supply. This would simply be a different channel by which local water quality violations affect household consumption with similar consequences.

<sup>6</sup>Note that the main effect of  $Below200PctFPL_i$  is excluded from equation 1 because it is collinear with the household fixed effects.

two other dependent variables of interest: the share of household expenditure on avoidance products including bottled water and water filtration devices, and the total number of calories purchased per household member per day. Estimating the event study coefficients for avoidance products serves two complementary purposes: first, it provides corroborating evidence that the estimated effects on total expenditure are attributable to the water quality violation if the decreases in overall expenditure are accompanied by relative increases in avoidance expenditures; and second, if the relative share of expenditures on avoidance is higher for poorer households than for wealthier ones, it confirms that the costs of avoidance are regressive.<sup>7</sup> Estimating the event study with total calories purchased as the dependent variable maps the estimated effects on total expenditure onto their nutritional consequences. While total expenditure and total calories purchased are obviously highly correlated, it is possible that poorer households only forego non-nutritive retail purchases, which would lead to total expenditure decreasing but calories purchased remaining constant. It is also possible that households engage in substitution to keep total calories purchased fixed while reducing total expenditure by purchasing cheaper and more calorie-dense foods. Thus it is necessary to estimate the effect on total calories purchased separately in order to infer the effect of water quality violations on food security.

### 2.2.2 Event Study

To complement the panel fixed-effects regression described in section 2.2.1, I conduct an event study using the first recorded water quality violation for each household in the sample. This allows for the visualization of heterogeneous treatment effects over time (e.g. more intense effects for violations that take longer to return to compliance) as well as the assessment of parallel trends pre-violation, which is necessary to confirm that the panel fixed-effects regression results are not driven by the continuation of differential trends pre-treatment that would not be absorbed by unit or time fixed effects. I plot event study coefficients for the 24 months prior and 24 months after the violation. Period 0 represents the month in which public notification of the violation was requested.<sup>8</sup>

In addition to the sample restrictions described in section 2.2.1, I impose two additional restrictions on the sample for the event study specification. First, I only use the earliest recorded water quality violation in each county. This is done to prevent the assignment of single observations to multiple time periods.<sup>9</sup> Second, I drop households after the associated active water quality viola-

---

<sup>7</sup>This depends on the fairly strong assumption that poorer and wealthier households within a county are equally likely to be affected by water quality violations that occur in that county. Although this is plausible for violations in large community water systems if the majority of the county's population is served by the same public utility, it may not hold for contaminants such as lead, which are more likely to affect poorer households because they are more likely to have old and corroded lead pipes.

<sup>8</sup>Without data limitations, it would be preferable to set period 0 as the month in which public notification was *issued*. However, unfortunately, while some violations do have this date recorded in SDWIS, it is missing for the vast majority of recorded violations.

<sup>9</sup>Since counties move in and out of treatment and some counties repeatedly experience water quality violations, it would otherwise be possible for a single observation to be, for example, 4 months into an active water quality violation and 12 months prior to the next violation, putting that observation in both time periods 4 and -12, potentially contaminating either pre- or post-treatment coefficients.

Table 1: Differential Effects of Active Health-Based Water Quality Violations

	Avoidance (Pct. pts.)	Total exp. (USD)	Calories (Cal)	Fat (Grams)	Carbs (Grams)	Sugars (Grams)	Protein (Grams)
Active Violation $\times$ Below 200% FPL	0.03*** (0.01)	-0.24*** (0.05)	-49.64*** (7.30)	-2.48*** (0.33)	-5.69*** (1.02)	-3.18*** (0.60)	-1.15*** (0.19)
Active Violation	-0.00 (0.00)	-0.03 (0.03)	-0.36 (3.56)	0.22 (0.17)	-0.40 (0.49)	-0.12 (0.28)	-0.03 (0.09)
Mean of dep. var.	0.59	10.39	1928.51	77.61	253.04	129.19	47.62
Coef. % of mean	4.7%	2.3%	2.6%	3.2%	2.2%	2.5%	2.4%
$N$	3040007	3040007	3040007	3040007	3040007	3040007	3040007
Adjusted $R^2$	0.255	0.559	0.421	0.449	0.403	0.367	0.442

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

Reported mean of dependent variable is for households below 200% of the federal poverty line.

tion returns to compliance. Since some water systems take longer to return to compliance than others, the coefficients after time 0 would otherwise be mixtures of households still experiencing an active water quality violation and those who were, but are no longer. With this restriction, the coefficient on e.g. month 9 can be unambiguously interpreted as the estimated effect of being in the 9th month of a yet-unresolved water quality violation.

### 3 Results

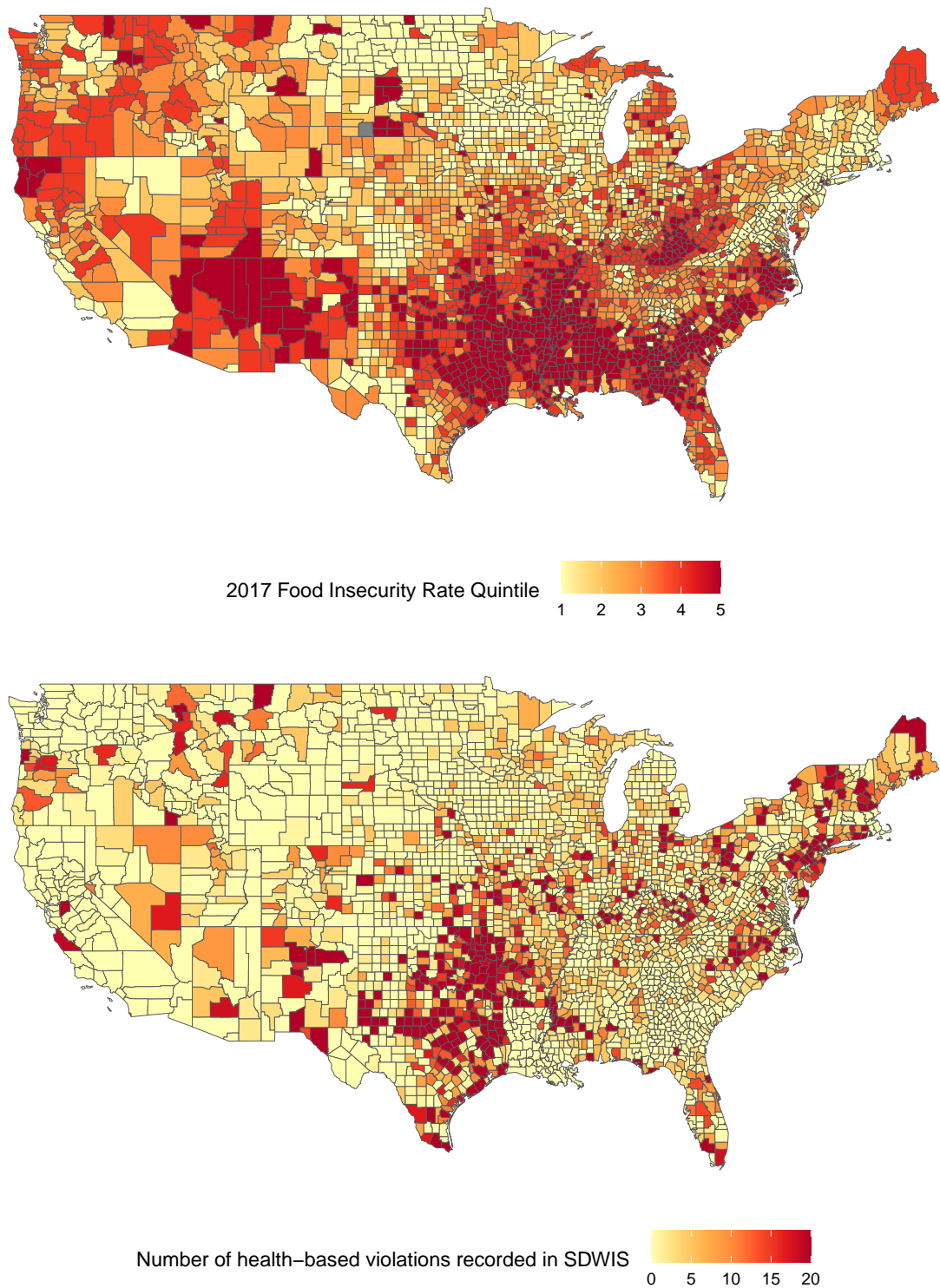
Figure 1 maps the number of health-based violations and the food insecurity rate in 2017, as measured in Feeding America’s 2019 *Map the Meal Gap* report. The average food insecurity rate is positively correlated with the number of health-based violations since 2010, from an average of 12.7% among counties with no recorded violations to 14.2% for counties with 3 violations or more (the 90th percentile), and this relationship is statistically significant ( $p < 0.01$ ). A similar pattern holds for the child food insecurity rate, for which the average among counties with no violations is 18.6% and the average among counties with 3 violations or more is 20.9% ( $p < 0.01$ ).

While this strong positive correlation between food insecurity and water quality violations motivates the research question of this paper, there are many reasons it may not be causal. Both the history of water quality violations and the rate of food insecurity are correlated with income-based poverty rates and themselves could be considered dimensions of a broader definition of poverty. Thus this correlation could be attributable to many other things associated with poverty, including the possibility that wealthier households move away from or avoid moving to an area with a history of water quality violations. To causally identify the effect of water quality violations on household budget constraints and resulting consequences for nutrition, it is necessary to observe consumption within a household before and after a water quality issue emerges.

Table 1 reports the results of the panel fixed-effects regression which estimates the differential



Figure 1: Maps of Food Insecurity Rate in 2017 and Number of Health-Based Violations Since 2010



Note: A health-based violation is included in a county's total if it affects at least 500 customers, based on the population served indicated in SDWIS.

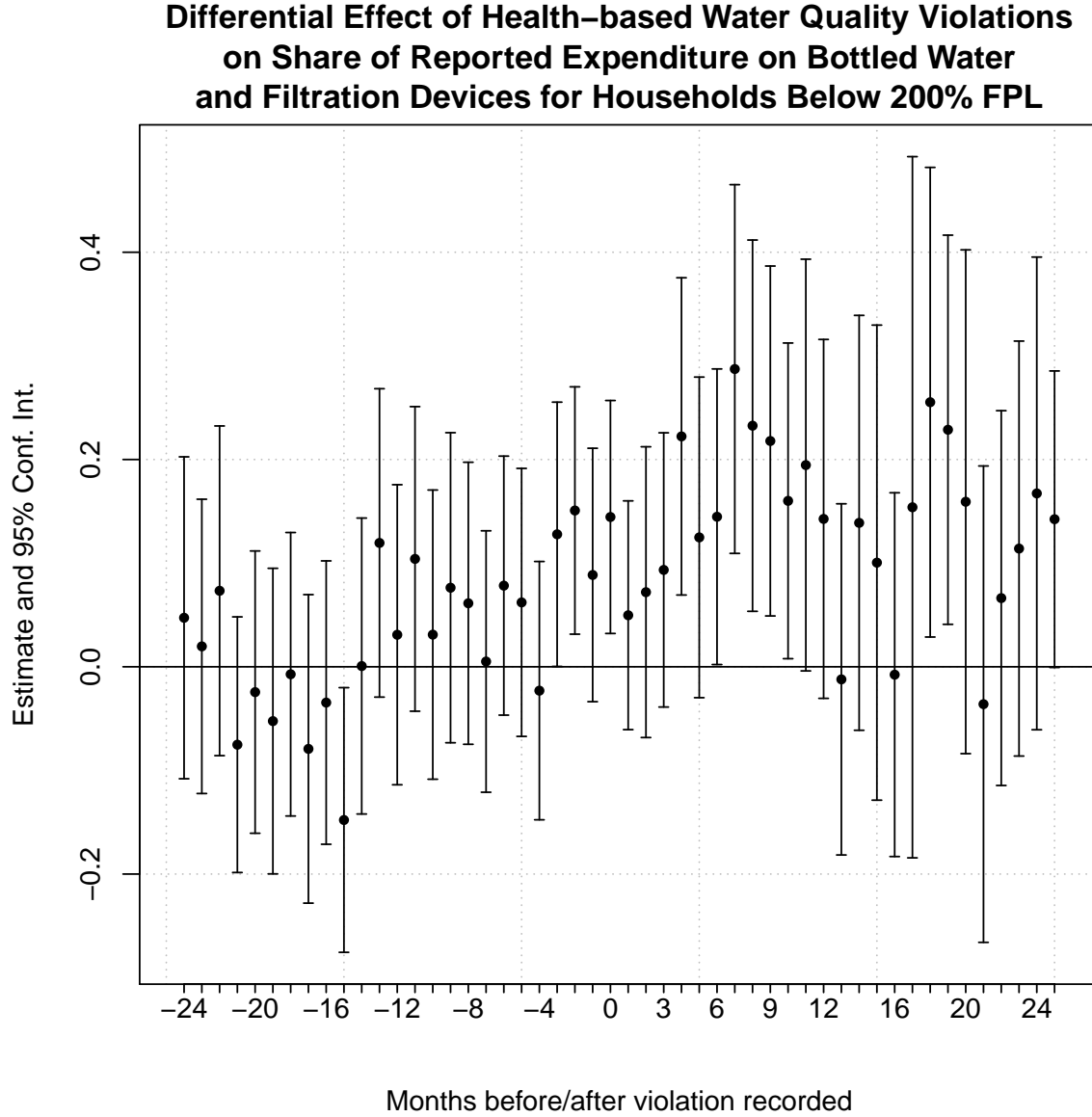
effect of an active water quality violation for households below 200% of the federal poverty line relative to their wealthier counterparts. The interaction coefficient in Column 1 of Table 1 is positive and statistically significant (albeit small), confirming that the increase in avoidance expenditures (bottled water and filtration devices) during an active water quality violation is proportionally larger for poorer households. The interaction coefficient in column 2 estimates a \$0.24 decline in total expenditures per day (about \$7 per month) among poorer households relative to wealthier households in the same area during an active water quality violation. As a consequence, purchased macronutrients per household member per day decline across the board. A decline in calories purchased of about 50 per day translates to about 1,500 per month, or about 0.75 days' worth of the recommended 2,000 calories per day. In all columns of Table 1, the coefficient on the main effect of an active violation is not statistically significant, suggesting that wealthier households are unaffected by the water quality violation. This is consistent with the hypothesis that the effect of water quality violations on the nutritional content of household grocery purchases is driven by budget constraints rather than preference changes or other channels unrelated to wealth.

Figures 2, 3, and 4 present the event study coefficients for avoidance, total expenditure, and calories per household member per day, respectively. In all three, it is evident that the coefficients in Table 1 are driven by longer-lasting violations. While just under half of water quality violations return to compliance within 6 months, with a mode of 1 month (see Figure 5 for a histogram), almost half (47%) are not, and these long-term violations may be especially problematic for poor households. There are multiple potential explanations for this: longer-term violations may receive more coverage, leading to higher awareness of the water quality problems and thus stronger incentives to substitute avoidance expenditures for grocery purchases, especially if legal action is taken against the water system; longer-lasting violations may mean greater cumulative exposure to contaminants, creating health problems in affected customers that cannot afford avoidance, increasing medical expenses and sick time away from work; or longer-lasting violations may culminate in other forms of economic disruption (e.g. business closures) that effectuate disproportionate negative income shocks on poorer households. More research with more granular data is needed to determine the underlying mechanism of the observed effect heterogeneity with certainty.

## 4 Discussion

While the data used in this paper offers many unique advantages, including the ability to analyze within-household effects of water quality violations, broad spatial coverage of the United States, and rich product-level data on household consumption, it also comes with limitations. Several of these limitations have already been discussed, including ambiguity over the exact time in which a water quality violation could be expected to change household behavior, the need to impute categorical averages for retail grocery products not represented in the FoodData Central nutrition facts dataset and the inability to determine for sure whether or not a particular household receives water from a source that committed a quality violation. While this paper attempts to address

Figure 2: Event Study of Effect on the Share of Monthly Household Expenditure on Bottled Water and Filtration Devices

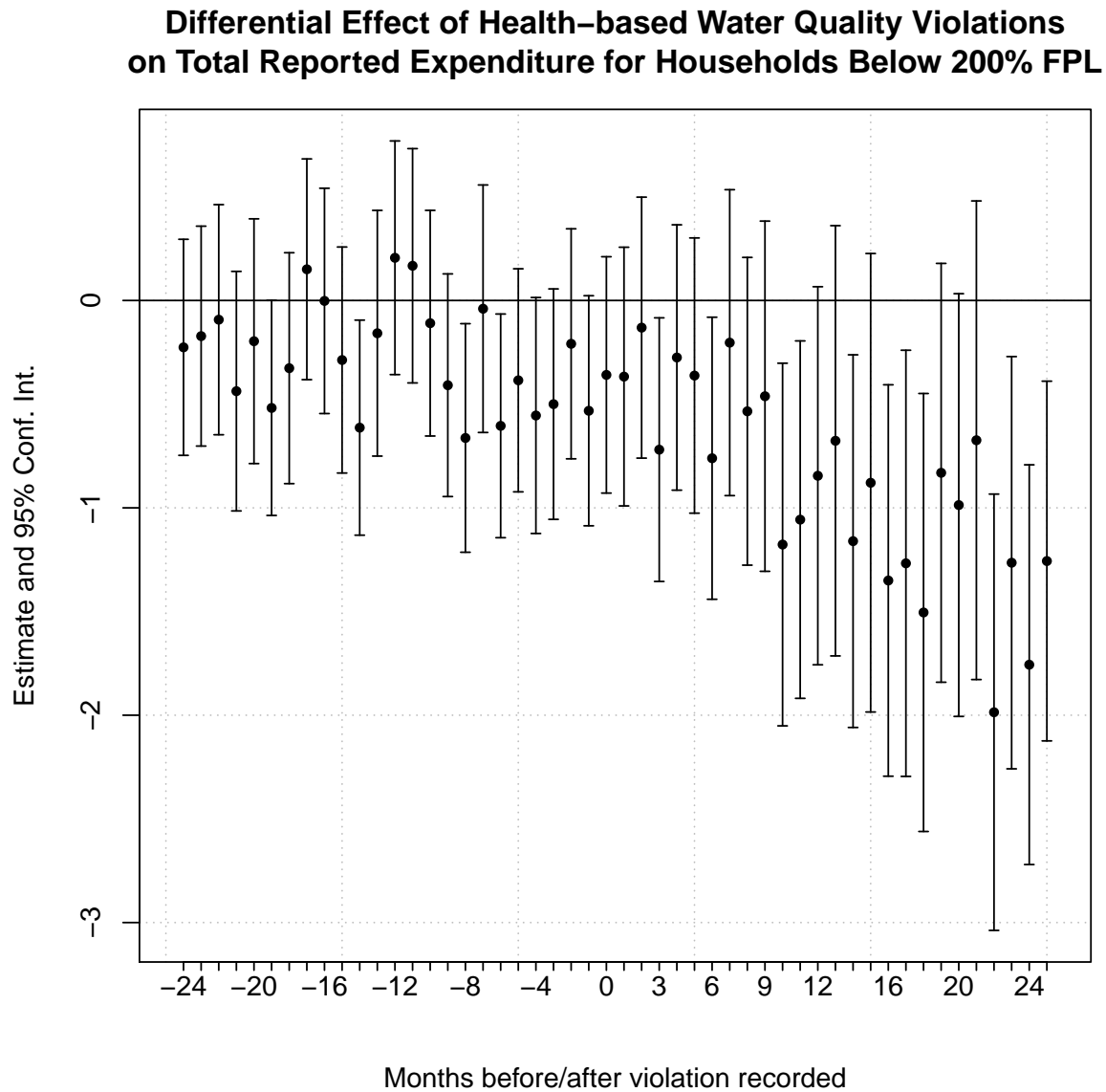


Note: To avoid overlapping pre- and post-periods in counties with multiple health-based violations since 2010, the figure above only includes the first observed health violation in the sample for each county.

these limitations by employing an event study approach that does not require ex-ante specification of a pre- and post-period and showing that the effects hold for categorical expenditure alongside the nutrient measures, further research with more precise data is needed to confirm these findings. This is especially true of the effect magnitudes, since most of the limitations introduce uncertainty that is likely to bias the estimated effects toward zero.

This paper also cannot address what households are spending money on instead of nutritive

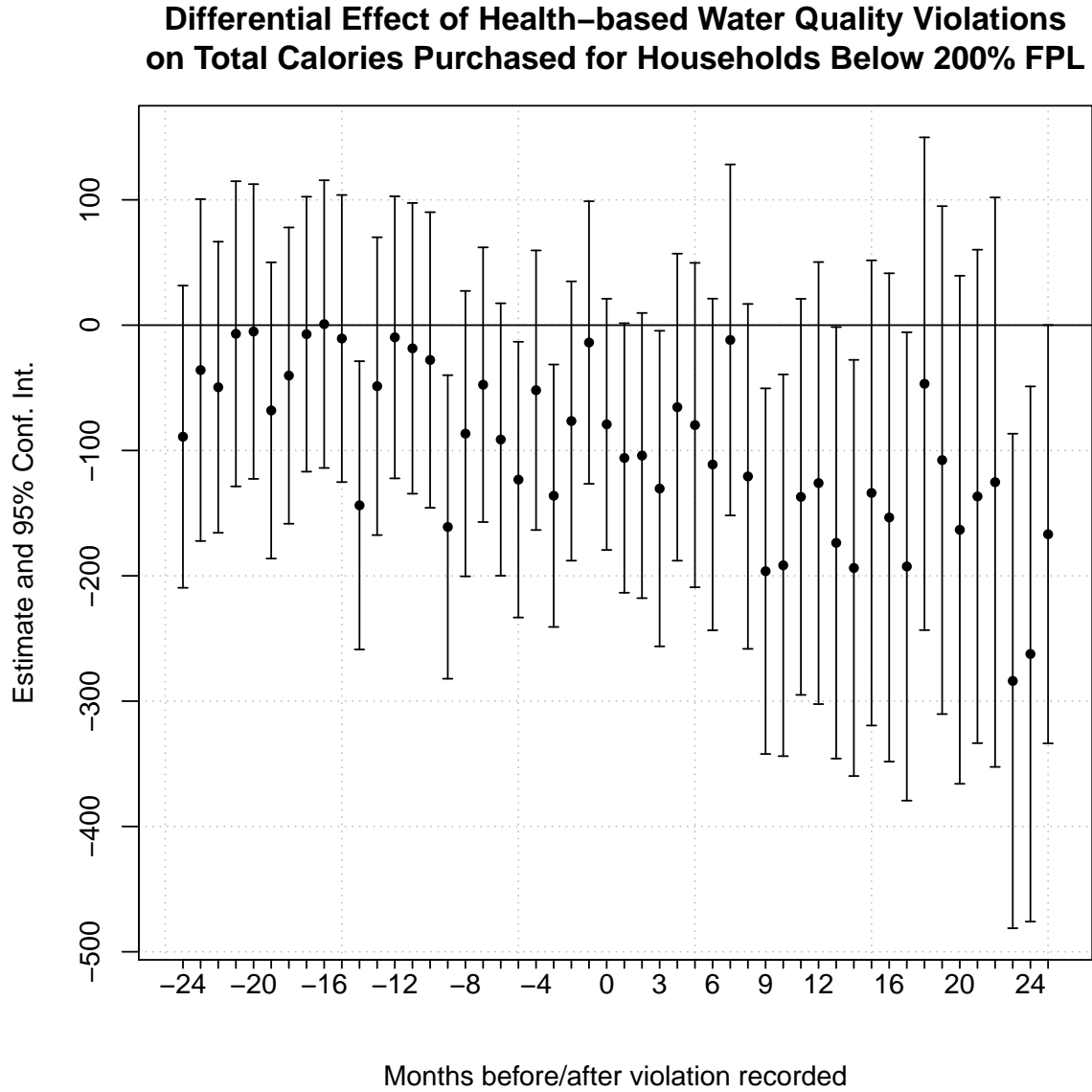
Figure 3: Event Study of Effect on Total Monthly Expenditure



Note: To avoid overlapping pre- and post-periods in counties with multiple health-based violations since 2010, the figure above only includes the first observed health violation in the sample for each county.

products when a water quality problem emerges. There are several possibilities: contaminated water may cause health problems requiring expensive treatment, especially if the impacted poor households are uninsured; reduced trust in household tap water may discourage preparing foods in the home, thus encouraging households to substitute groceries for fast food or other prepared foods that would not be represented in the HomeScan Consumer Panel; or households may be purchasing avoidance in ways that are not reported as consumption in the panel, by installing water filters in the home or purchasing water from retailers other than grocery or department stores. While it is

Figure 4: Event Study of Effect on Total Calories Purchased Per Household Member Per Day

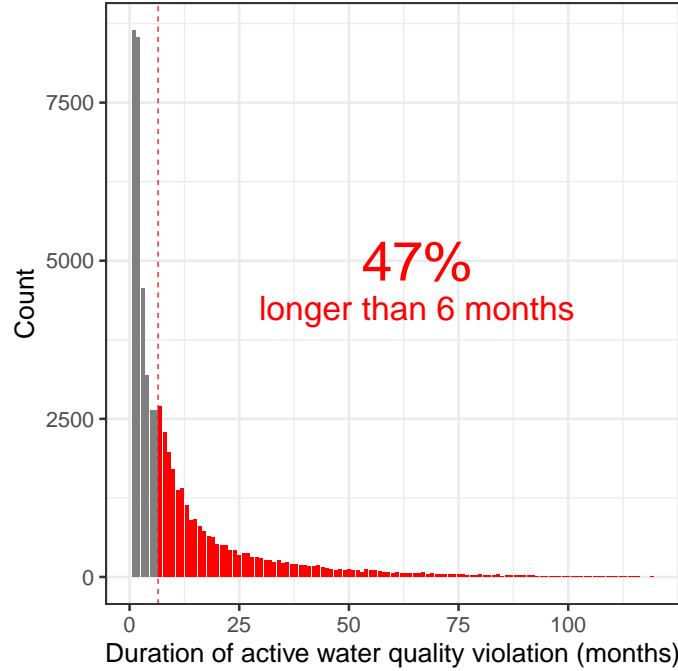


Note: To avoid overlapping pre- and post-periods in counties with multiple health-based violations since 2010, the figure above only includes the first observed health violation in the sample for each county.

unlikely that any of these possibilities would result in a net improvement in household nutrition, the long-run health consequences of the effects presented by this paper depend on which explanation holds for the majority of households.

Further research is also needed to determine the effect of water quality violations on *consumed* calories and other macronutrients rather than purchased quantities at grocery stores. The United States Department of Agriculture estimates total food waste in the U.S. between 30 and 40% of the food supply. This suggests that on average, calories consumed will be a proper subset of calories

Figure 5: Histogram of Active Water Quality Violation Durations



purchased. At the same time, according to USDA ERS, half or more of the food consumed in the United States is prepared outside of the home (e.g. from restaurants) (Saksena et al. 2018). Because the data used in this paper does not include purchases from restaurants, I cannot observe whether households are engaging in substitution of food away from home for groceries (perhaps to avoid cooking with contaminated water) or decreasing their overall food consumption. While food away from home, especially from fast food establishments, tends to be of lower nutritional quality than food prepared at home, this means that more detailed food consumption data would be required to determine which households or particular household members are facing hunger versus those that are consuming more unhealthy foods.

## 5 Conclusion

Using a panel of household retail purchases, nutritional quantities based on UPC-level nutrition facts, and county-level water quality violations, this paper demonstrates a negative effect of water supply contamination on household nutrition for households with income below 200% of the federal poverty line. Event study coefficients suggest that this effect accumulates over the duration of water quality violations that take several months, or years, to return to compliance. In the same time period, demand for bottled water and filtration devices, which facilitate avoidance of contaminated water, increases among poorer households. By virtue of their lower income, comparable absolute increases in demand for avoidance comprise a larger share of the poorer households' total monthly expenditure. However, observed demand for avoidance cannot account for the entire decrease in

household retail expenditure of about \$7 per month, suggesting either an associated income effect or unobserved avenues of purchasing avoidance. This decline in monthly expenditure is associated with a decrease of about 50 calories purchased per household member per day.

These findings suggest multiple potential welfare gains from increased investment in water quality monitoring, enforcement, and improvement. An estimated 26-38% of water quality violations are either incorrectly or never reported to the federal Safe Drinking Water Information System (SDWIS) (United States Environmental Protection Agency 2000). Since the findings of this paper focus on the time leading up to the documentation of a water quality violation rather than the actual public notification of the violation, this suggests that households in areas with persistent undetected water quality violations may experience long-term costs that undermine nutrition and other welfare-improving expenditures without recourse. Even among the violations that are detected, significant delays are likely because testing is costly and there are a large number of water suppliers to monitor; the longer these delays are, the more costs accumulate to affected households. Finally, improvements to residential water supply infrastructure would reduce the likelihood of certain types of violations such as dissolved lead regardless of monitoring frequency. In addition to the many known direct benefits of improved water supply (Ao 2016, Clay et al. 2014, Devoto et al. 2012, Ferrie and Troesken 2008, Ritter 2019, Troesken et al. 2011), this paper adds reduced risk of food insecurity to the list of potential welfare improvements resulting from these investments.

This paper also provides additional evidence of spatially-determined disparities in the United States. Much of the residential water infrastructure in centuries-old U.S. cities is outdated, hence the EPA’s request of more than \$2 billion in fiscal year 2020 to fund reconstruction projects nationwide. The quality of that infrastructure thus to some extent reflects the historical prosperity of each particular neighborhood. If that infrastructure increases the likelihood of repeated or persistent water quality violations, it will inhibit economic development, as both people and businesses are discouraged from moving in to an area with unsafe water; this in turn makes it more difficult to replace that infrastructure without public funding. Meanwhile, this paper suggests those who live in that neighborhood must choose between exposure to harmful water contaminants or foregoing food on the table to purchase avoidance. Both options carry a risk of developing health conditions, especially for children, that increase the difficulty of social mobility.

## References

- Achyuta Adhvaryu, James Fenske, and Anant Nyshadham. Early Life Circumstance and Adult Mental Health. Journal of Political Economy, 127(4):1516–1549, August 2019. ISSN 00223808.
- Maura Allaire, Haowei Wu, and Upmanu Lall. National trends in drinking water quality violations. Proceedings of the National Academy of Sciences, 115(9):2078–2083, February 2018. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1719805115. URL <https://www.pnas.org/content/115/9/2078>. Publisher: National Academy of Sciences Section: Social Sciences.

- Maura Allaire, Taylor Mackay, Shuyan Zheng, and Upmanu Lall. Detecting community response to water quality violations using bottled water sales. Proceedings of the National Academy of Sciences, 116(42):20917–20922, October 2019. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1905385116. URL <https://www.pnas.org/content/116/42/20917>. Publisher: National Academy of Sciences Section: Social Sciences.
- Patricia M. Anderson, Kristin F. Butcher, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. Beyond Income: What Else Predicts Very Low Food Security Among Children? Southern Economic Journal, 82(4):1078–1105, 2016. ISSN 2325-8012. doi: 10.1002/soej.12079. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/soej.12079>.
- Chon K. Ao. Essays on the impact of clean water on human capital and productivity. PhD Thesis, 2016.
- Brian Beach, Joseph Ferrie, Martin Saavedra, and Werner Troesken. Typhoid Fever, Water Quality, and Human Capital Formation. The Journal of Economic History, 76(01):41–75, March 2016. URL [https://ideas.repec.org/a/cup/jechis/v76y2016i01p41-75\\_00.html](https://ideas.repec.org/a/cup/jechis/v76y2016i01p41-75_00.html).
- Roy Brouwer, Fumbi Crescent Job, Bianca van der Kroon, and Richard Johnston. Comparing Willingness to Pay for Improved Drinking-Water Quality Using Stated Preference Methods in Rural and Urban Kenya. Applied Health Economics and Health Policy, 13(1):81–94, February 2015. ISSN 11755652.
- James A. Brox, Ramesh C. Kumar, and Kenneth R. Stollery. Estimating Willingness to Pay for Improved Water Quality in the Presence of Item Nonresponse Bias. American Journal of Agricultural Economics, 85(2):414–428, May 2003. ISSN 00029092. URL <http://search.ebscohost.com/login.aspx?direct=true&db=ec&AN=0651855&site=ehost-live>.
- Raj Chetty, John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility. Working Paper 25147, National Bureau of Economic Research, October 2018. URL <http://www.nber.org/papers/w25147>. Series: Working Paper Series.
- Karen Clay, Werner Troesken, and Michael Haines. Lead and Mortality. The Review of Economics and Statistics, 96(3):458–470, 2014. ISSN 0034-6535. URL <https://www.jstor.org/stable/43555336>. Publisher: The MIT Press.
- Florencia Devoto, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons. Happiness on Tap: Piped Water Adoption in Urban Morocco. American Economic Journal: Economic Policy, 4(4):68–99, 2012. URL <https://ideas.repec.org/a/aea/aejpol/v4y2012i4p68-99.html>.
- Joseph P. Ferrie and Werner Troesken. Water and Chicago’s mortality transition, 1850-1925. Explorations in Economic History, 45(1):1–16, 2008.



- Richard Freeman, Wenquan Liang, Ran Song, and Christopher Timmins. Willingness to pay for clean air in China. Journal of Environmental Economics and Management, 94:188–216, March 2019. ISSN 0095-0696. doi: 10.1016/j.jeem.2019.01.005. URL <http://www.sciencedirect.com/science/article/pii/S0095069617308902>.
- David E. Frisvold. Nutrition and Cognitive Achievement: An Evaluation of the School Breakfast Program. Journal of Public Economics, 124:91–104, April 2015. ISSN 00472727.
- Hilary Hoynes, Marianne Page, and Ann Huff Stevens. Can targeted transfers improve birth outcomes?: Evidence from the introduction of the WIC program. Journal of Public Economics, 95(7):813–827, August 2011. ISSN 0047-2727. doi: 10.1016/j.jpubeco.2010.12.006. URL <http://www.sciencedirect.com/science/article/pii/S0047272710002082>.
- Hilary Hoynes, Doug Miller, and David Simon. Income, the Earned Income Tax Credit, and Infant Health. American Economic Journal: Economic Policy, 7(1):172–211, 2015. ISSN 1945-7731. URL <https://www.jstor.org/stable/24465947>.
- Hilary W. Hoynes and Diane Whitmore Schanzenbach. Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program. American Economic Journal: Applied Economics, 1(4):109–139, 2009. ISSN 1945-7782. URL <https://www.jstor.org/stable/25760184>.
- Rhiannon L. Jerch. The Local Consequences of Federal Mandates: Evidence from the Clean Water Act. 2018.
- Robert J. Johnston and Paul J. Thomassin. Willingness to Pay for Water Quality Improvements in the United States and Canada: Considering Possibilities for International Meta-analysis and Benefit Transfer. Agricultural and Resource Economics Review, 39(1):114–131, February 2010. ISSN 10682805.
- Ilana V. Kohler, Collin F. Payne, Chiwoza Bandawe, and Hans-Peter Kohler. The Demography of Mental Health among Mature Adults in a Low-Income, High-HIV-Prevalence Context. Demography, 54(4):1529–1558, August 2017. ISSN 00703370. URL <http://search.ebscohost.com/login.aspx?direct=true&db=ecn&AN=1663852&site=ehost-live>.
- Patricia I. Ritter. The Hidden Role of Piped Water in the Prevention of Obesity in Developing Countries. Experimental and Non-Experimental Evidence. Technical Report 2019-02, University of Connecticut, Department of Economics, August 2019. URL <https://ideas.repec.org/p/uct/uconnp/2019-02.html>.
- Lilia Rodriguez-Tapia, Daniel A. Revollo-Fernandez, and Jorge A. Morales-Novelo. Household’s Perception of Water Quality and Willingness to Pay for Clean Water in Mexico City. Economies, 5(2):1–14, June 2017. ISSN 22277099.

- Michelle Saksena, Abigail Okrent, Tobenna D. Anekwe, Clare Cho, Chris Dicken, Anne Effl, Howard Elitzak, Joanne Guthrie, Karen Hamrick, Jeffrey Hyman, Young Jo, Biing-Hwan Lin, Lisa Mancino, Patrick W. McLaughlin, Ilya Rahkovsky, Katherine Ralston, Travis A. Smith, Hayden Stewart, Jessica E. Todd, and Charlotte Tuttle. America's Eating Habits: Food Away From Home, 2018. URL <http://www.ers.usda.gov/publications/pub-details/?pubid=90227>. Library Catalog: [www.ers.usda.gov](http://www.ers.usda.gov).
- Laurel A. Schaider, Lucien Swetschinski, Christopher Campbell, and Ruthann A. Rudel. Environmental justice and drinking water quality: are there socioeconomic disparities in nitrate levels in U.S. drinking water? *Environmental Health*, 18(1):3, January 2019. ISSN 1476-069X. doi: 10.1186/s12940-018-0442-6. URL <https://doi.org/10.1186/s12940-018-0442-6>.
- David Switzer and Manuel P. Teodoro. Class, Race, Ethnicity, and Justice in Safe Drinking Water Compliance. *Social Science Quarterly*, 99(2):524-535, 2018. ISSN 1540-6237. doi: 10.1111/ssqu.12397. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/ssqu.12397>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ssqu.12397>.
- Werner Troesken. *Water, Race, and Disease*. MIT Press, 2004. ISBN 978-0-262-20148-3. Google-Books-ID: GmgYTwfcIMsC.
- Werner Troesken, Joseph P. Ferrie, and Karen Rolf. Cognitive Disparities, Lead Plumbing, and Water Chemistry: Intelligence Test Scores and Exposure to Water-Borne Lead Among World War Two U.S. Army Enlistees. Technical report, National Bureau of Economic Research, 2011.
- United States Environmental Protection Agency. Data Reliability Analysis of the EPA Safe Drinking Water Information System/Federal Version (SDWIS/FED), 2000. URL <https://nepis.epa.gov/Exe/ZyNET.exe/40001BK0.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2000+Thru+2005&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C00thru05%5Ctxt%5C00000016%5C40001BK0.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL>. Library Catalog: [nepis.epa.gov](https://nepis.epa.gov).
- United States Environmental Protection Agency. Safe Drinking Water Act (SDWA) Resources and FAQs, February 2020. URL [, /help/sdwa-faqs](https://www.epa.gov/help/sdwa-faqs). Library Catalog: [echo.epa.gov](https://www.epa.gov).
- U.S. Department of Agriculture, Agricultural Research Service. *FoodData Central*. 2019.
- Francesca Valent, D'Anna Little, Roberto Bertollini, Leda E Nemer, Fabio Barbone, and Giorgio Tamburlini. Burden of disease attributable to selected environmental factors and injury among

children and adolescents in Europe. The Lancet, 363(9426):2032–2039, June 2004. ISSN 0140-6736. doi: 10.1016/S0140-6736(04)16452-0. URL <http://www.sciencedirect.com/science/article/pii/S0140673604164520>.

Joshua Graff Zivin, Matthew Neidell, and Wolfram Schlenker. Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption. American Economic Review, 101(3): 448–453, May 2011. ISSN 0002-8282. doi: 10.1257/aer.101.3.448. URL <https://www.aeaweb.org/articles?id=10.1257/aer.101.3.448>.

# Appendix

## A Zip Code Based Active Violation Measures

In the results reported in the main paper, active violations are assigned at the county level because this is the finest geographic unit reported alongside each violation in SDWIS. (A zip code based table is offered by the EPA, but it simply associates every zip code in a particular county with each violation, whether or not that zip code is served by the violating water system.) The largeness of counties introduces uncertainty regarding whether or not each household is likely to be affected by a water quality violation. While this is likely to attenuate estimates by over-assigning treatment, a potential concern is that water quality violations may be concentrated in areas of counties where lower-income households are likely to be located. If lower-income households are systematically more likely to be affected by water quality violations while higher-income households are more likely to be erroneously assigned the treatment, the estimated null effect of water quality violations on households above 200% of the federal poverty line may be explained by these households never actually experiencing the treatment rather than receiving the treatment and not responding to it. This would change the interpretation of the results, although the main conclusion of water quality violations decreasing expenditure on nutritious grocery products for poor households would still apply.

In this section, I present the results of an alternative measure of water quality violations constructed at the zip code level. I accomplish this by using publicly available service area maps for community water systems to determine the zip codes served by each water utility and matching these zip codes to SDWIS violation records using the unique public water system ID. At the time of writing, service area maps had been obtained for public water systems in Arizona, Arkansas, California, Connecticut, Kansas, Louisiana, Nebraska, New Jersey, New York, North Carolina, Oklahoma, Oregon, Pennsylvania, Texas, Virginia, Washington, and West Virginia. Thus for the results presented in this section, the sample is restricted to households in these states.

Table 2 and Figures 6, 7, and 8 replicate the results of Table 1 and Figures 2, 3, and 4, respectively, with this restricted zip code-based sample. The findings are largely similar (albeit with reduced statistical significance, likely attributable to the greatly reduced sample size), suggesting an increased share of avoidance expenditures, decreased expenditure on grocery purchases, and decreases across the board in purchased nutrients, and accumulating effects for longer-lasting violations. Also similar to the previous findings, there are no significant effects for wealthier households.

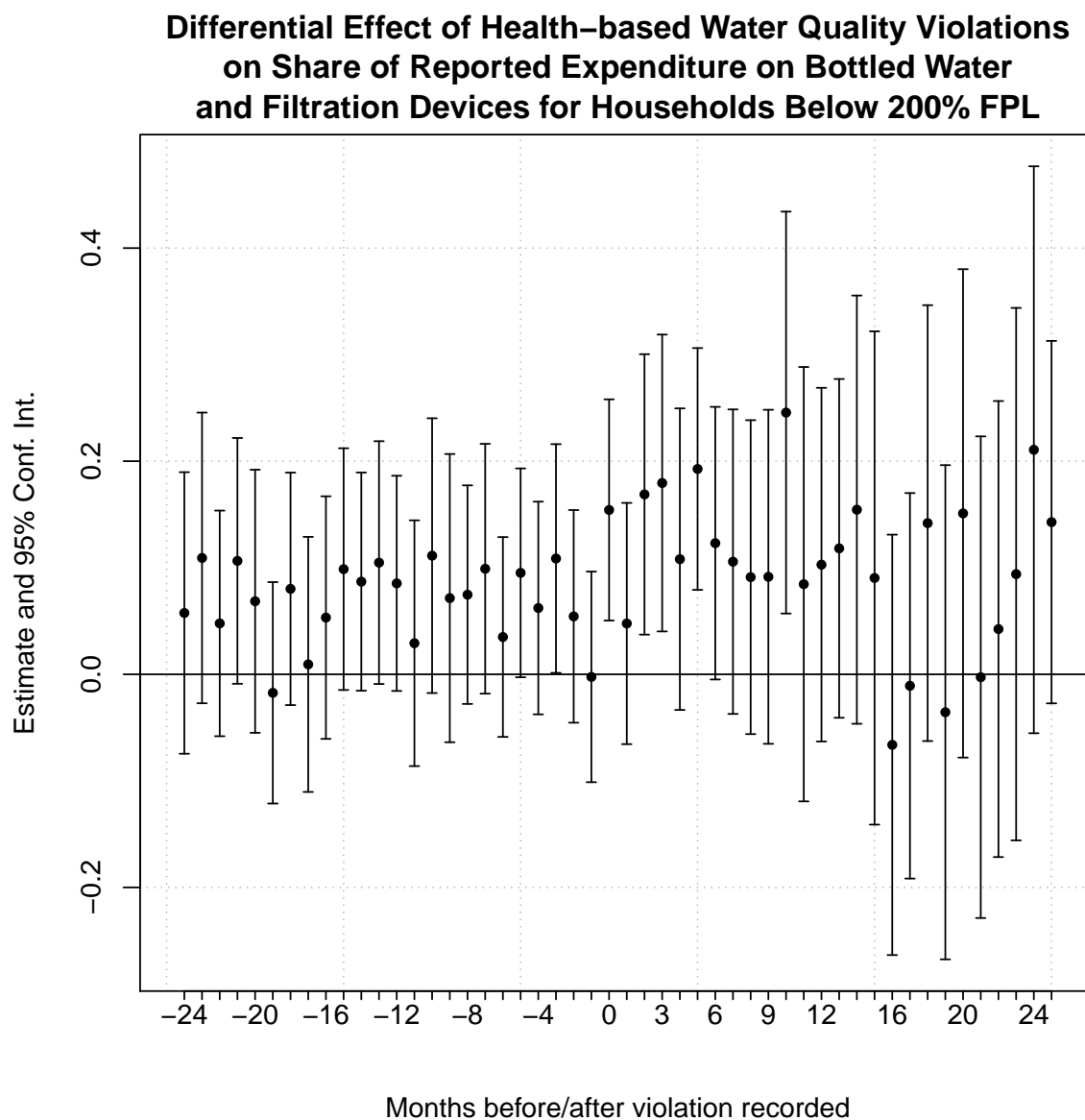
Table 2: Differential Effects of Active Health-Based Water Quality Violations (Zip Code Level Data)

	Avoidance (Pct. pts.)	Total exp. (USD)	Calories (Cal)	Fat (Grams)	Carbs (Grams)	Sugars (Grams)	Protein (Grams)
Active Violation $\times$ Below 200% FPL	0.04** (0.02)	-0.24* (0.15)	-51.47** (21.19)	-3.13*** (0.99)	-5.31* (2.80)	-2.29 (1.58)	-1.27** (0.54)
Active Violation	-0.01 (0.01)	-0.09 (0.07)	-8.36 (8.68)	-0.06 (0.48)	-1.33 (1.11)	-0.65 (0.62)	-0.39 (0.24)
Mean of dep. var.	0.60	10.32	1909.58	76.46	251.97	128.04	47.11
Coef. % of mean	7.2%	2.3%	2.7%	4.1%	2.1%	1.8%	2.7%
$N$	516847	516847	516847	516847	516847	516847	516847
Adjusted $R^2$	0.252	0.560	0.431	0.463	0.411	0.376	0.448

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

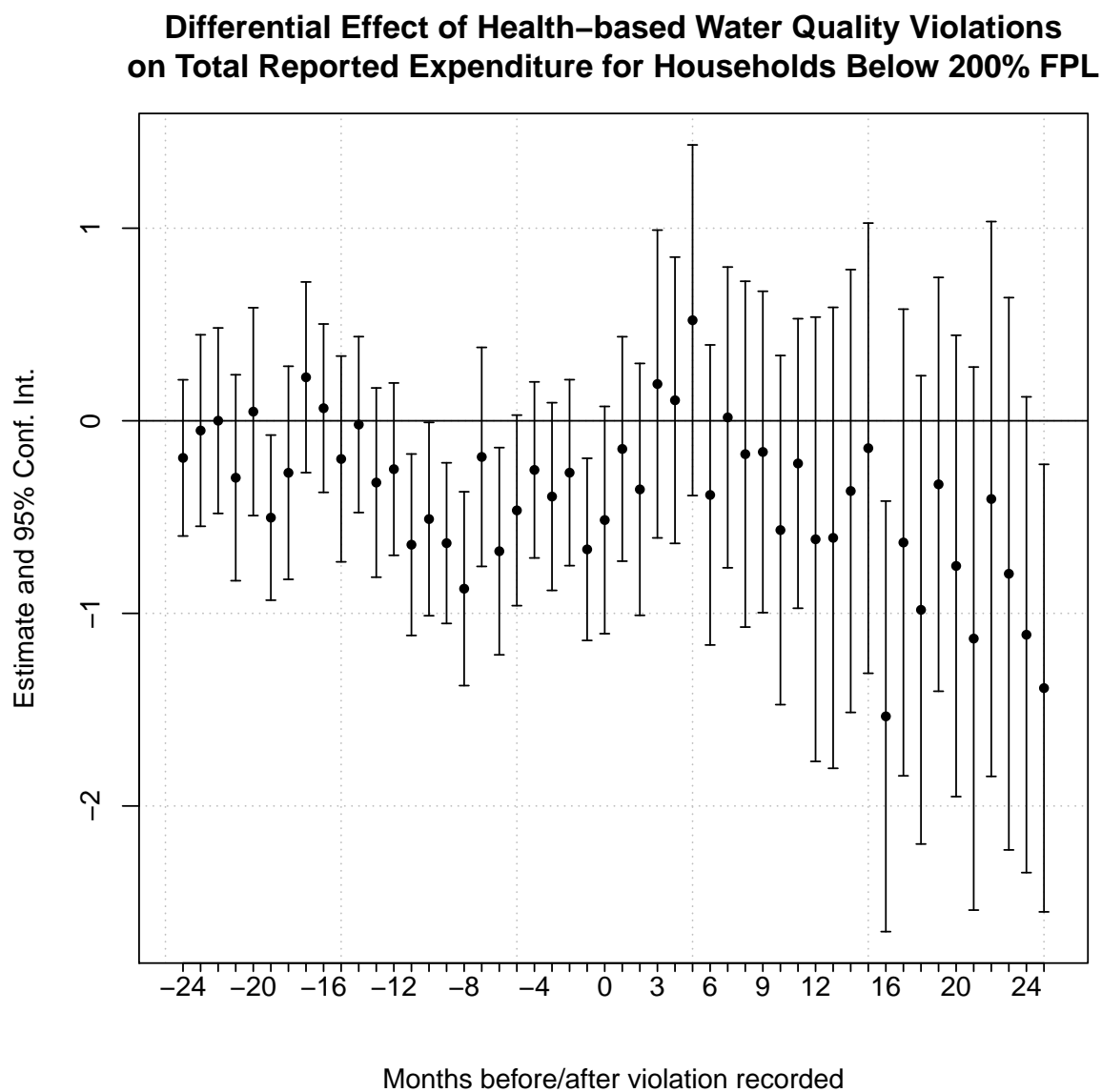
Reported mean of dependent variable is for households below 200% of the federal poverty line.

Figure 6: Event Study of Effect on the Share of Monthly Household Expenditure on Bottled Water and Filtration Devices (Zip Code Level Data)



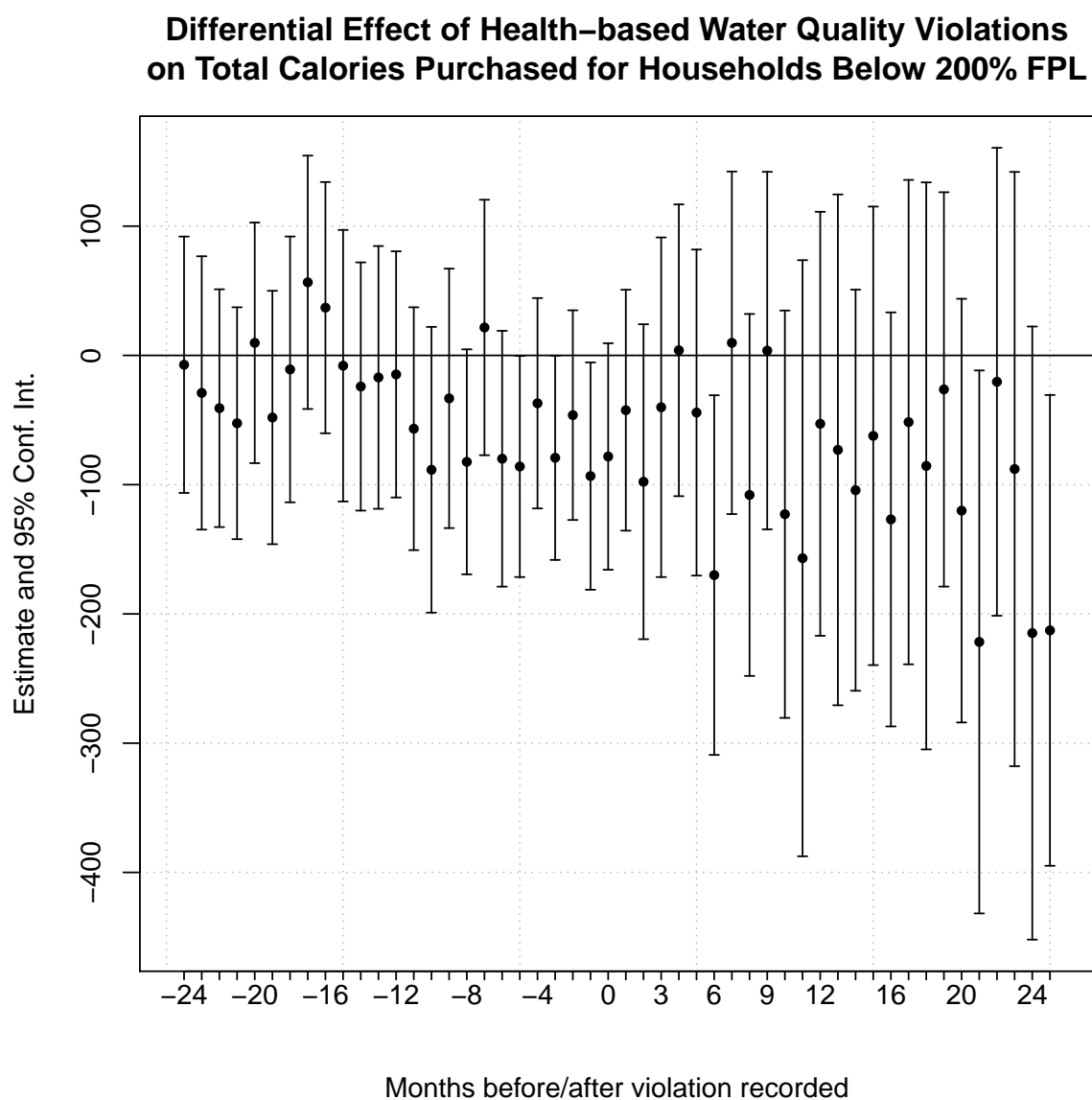
Note: To avoid overlapping pre- and post-periods in counties with multiple health-based violations since 2010, the figure above only includes the first observed health violation in the sample for each county.

Figure 7: Event Study of Effect on Total Monthly Expenditure (Zip Code Level Data)



Note: To avoid overlapping pre- and post-periods in counties with multiple health-based violations since 2010, the figure above only includes the first observed health violation in the sample for each county.

Figure 8: Event Study of Effect on Total Calories Purchased Per Household Member Per Day (Zip Code Level Data)



Note: To avoid overlapping pre- and post-periods in counties with multiple health-based violations since 2010, the figure above only includes the first observed health violation in the sample for each county.