WATER QUALITY VIOLATIONS AND AVOIDANCE BEHAVIOR:
EVIDENCE FROM BOTTLED WATER CONSUMPTION

Joshua Graff Zivin
Matthew Neidell
Wolfram Schlenker

Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption
Joshua Graff Zivin, Matthew Neidell, and Wolfram Schlenker
NBER Working Paper No. 16695
January 2011
JEL No. H41,I18,Q53,Q58

ABSTRACT

In this paper, we examine the impact of poor water quality on avoidance behavior by estimating the change in bottled water purchases in response to drinking water violations. Using data from a national grocery chain matched with water quality violations, we find an increase in bottled water sales of 22 percent from violations due to microorganisms and 17 percent from violations due to elements and chemicals. Back-of-the envelope calculations yield costs of avoidance behavior at roughly $60 million for all nationwide violations in 2005, which likely reflects a significant understatement of the total willingness to pay to eliminate violations.

Joshua Graff Zivin
University of California, San Diego
9500 Gilman Drive, MC 0519
La Jolla, CA 92093-0519
and NBER
jgraffzivin@ucsd.edu

Matthew Neidell
Department of Health Policy and Management
Columbia University
600 W 168th Street, 6th Floor
New York, NY 10032
and NBER
mn2191@columbia.edu

Wolfram Schlenker
Department of Economics
School of International and Public Affairs
Columbia University
420 West 118th Street, MC 3323
New York, NY 10027
and NBER
wolfram.schlenker@columbia.edu

Water Quality Violations and Avoidance Behavior:
Evidence from Bottled Water Consumption

Joshua Graff Zivin, Matthew Neidell, Wolfram Schlenker

The provision and public dissemination of information about health hazards has become

an increasingly important part of state and federal programs designed to manage environmental

and health risks. Examples include the US Environmental Protection Agency's (EPA) Toxics

Release Inventory, the US Food and Drug Administration's fish advisories, and California's

'smog alerts' program. The central idea behind each of these programs is that the provision of

information allows the public to engage in behavioral responses to minimize exposure should the

costs of that exposure exceed its benefits. Whether such information is a substitute or

complement to environmental standards is an open question. In either case, understanding

responses to such informational approaches is critical for determining both the costs and the

effectiveness of these programs.

One important area where informational approaches play a key role is in the management

of drinking water quality. The EPA, under the auspices of the Safe Drinking Water Act

(SDWA), places strict limits on roughly 90 chemicals or contaminants in community drinking

water systems, which is accessed by nearly 270 million people in the United States. Despite

these limits, roughly one in ten Americans is served by a drinking water system that exceeds

these limits on at least one dimension (Duhigg, 2009). Such violations must be disclosed to

consumers under the SDWA Amendments of 1996. This paper examines avoidance behavior in

response to these disclosures regarding drinking water violations.

Matching geocoded violations data for Northern California and Nevada from 2001-2005

with sales data from a major supermarket chain, we estimate the change in bottled water

purchases as a result of tap water violations. Since the behavioral response in this case is a market-based one, it is straightforward to calculate the costs of avoidance behavior. We find a statistically significant increase in bottled water sales of 22 percent from violations due to microorganisms and 17 percent from violations due to elements and chemicals.[1] Combining these store-level estimates from Northern California and Nevada with national sales data on bottled water consumption, we compute back-of-the-envelope costs of avoidance behavior at roughly $60 million for all violations in 2005, noting this figure likely reflects a significant understatement of the total costs of avoidance behavior, and thus willingness to pay to eliminate violations.

I. Data

The SDWA Amendments of 1996 require 150,000 community water districts (CWDs) throughout the US to actively monitor contaminants levels. If contaminants exceed maximum contaminant level (MCL) standards, a CWD must notify the EPA of the violation and adhere to the Public Notification Rule. This rule requires CWDs to notify customers within 24-hours if the responsible contaminant poses an immediate health threat (primarily microorganisms and nitrates) and within 30 days for other health threats. CWDs must notify customers through various social media outlets, posting in public places, and personal delivery. Relevant for our analysis, these notifications must include a description of the violation and potential health effects, the population at risk, actions consumers can take, when the violation occurred, when a resolution is expected, and language encouraging broader distribution of information regarding the violation.[2]

---

[1] Nitrates and nitrites violations result in reductions of similar magnitude, but are not statistically significant.

[2] Unfortunately for our analysis, we do not have data on the exact details of the notification provided by districts.

The Safe Drinking Water Information System (SDWIS) maintained by the EPA contains detailed records of all violations. After filing a Freedom of Information Act request, we obtained historical information on all violations that occurred in the US from 2001-2005. These data include the start and end date of the violation, the contaminant responsible for the violation, as well as characteristics about the CWD, including the county and population served. We contacted all water districts in California and Nevada that serve at least 100 people and obtain the zip codes they serve. If a water district serves more than one zip code, we split the population served between zip codes based on their overall population. For example, if a water district serves customers in two zip codes, where zip code 1 has twice the population of zip code 2, we assigned two-thirds of the population served to zip code 1 and one-third to zip code 2.

We combine individual violations into three broadly defined groups based on their potential health effects, which is also consistent with the public notification rule. "Microorganisms" pose immediate gastrointestinal health threats to all individuals; this largely consists of coliform bacteria and can be removed by boiling tap water. "Nitrates" pose immediate threat of "blue-baby syndrome" to infants and can not be removed by boiling. "Elements/chemicals," which includes natural occurring elements, such as arsenic, manufacturing chemicals, such as tetrachloroethylene, and disinfection byproducts from removing microorganisms, also can not be removed by boiling. The health effects from elements/chemicals, which include cancer and toxicity to various organs, typically arise from longer-term exposure and do not require immediate notification. Table 1 displays the number of violations for each group along with the mean duration of each violation for both the US and our final sample for the years 2001-2005.

Our data on bottled water consumption uses weekly sales (Wednesday-Tuesday) from a

national grocery chain for their stores in Northern California and Nevada for weeks starting October 31, 2001 until November 2, 2005. The data includes sales in dollars as well as quantity sold for 308 Universal Product Codes (UPC). These UPCs cover still water (e.g., Aquafina, Dasani), sparkling water (e.g., Pellegrino, Perrier), and flavored sparkling water (e.g., Calistoga lemon flavored sparkling water). Different sizes of the same product (e.g., 16oz versus 1 gallon) have distinct UPCs, although the size of a bottle is unfortunately not identified in the UPC database for most codes. Aggregating quantities is complicated by the fact that an increase in demand might be met by switching from smaller to larger bottles while holding the number of units sold constant. Thus, we aggregate sales in dollars for all 308 UPCs by store and week as our dependent variable.

Store level sales were linked to water violations by matching water districts with stores that are located in the zip code that is served by the district. Our baseline model uses zip-codes, and not a distance measure, to match water districts to zip codes. Such an algorithm better captures the relevant customer base of a store because zip codes are much larger in rural areas and customers may drive further to reach a store than customers in urban areas.

Table 2 displays community characteristics by violation frequency. We define "high (low) violations" as being above (below) the median number of violations for the time period studied. In areas with more violations, residents on average consume less bottled water and come from lower socio-economic status (SES). This pattern is consistent with poorer provision of public goods in lower SES areas and that bottled water consumption is a normal good. Both indicate the importance of accounting for the endogeneity of violations.

II. Methods

In a standard utility maximization setup, we can think of the demand for bottled water as

a function of the price of bottled water, the price of tap water, the price of substitutes (all quality-adjusted), income, and both time-varying and time-invariant individual preferences. If a violation occurs, the quality-adjusted price of tap water increases, thus increasing the demand for bottled water.

We empirically model this relationship between violations and bottled water sales according to the following specification:

(1) $y_{swt} = \beta_1 + \beta_2*\textbf{violation}_{swt}*(pop_{wz}/pop_z) + \beta_3*\textbf{weather}_{swt} + \boldsymbol{\alpha}_{sw} + \boldsymbol{\delta}_t + \varepsilon_{swt}$

where y is log(weekly sales of bottled water) at store s in water district w (both located in zip code z) in week t. The fraction of time a store-water district combination was in violation for each of the three types of violations in week t is captured by the vector **violation**. We multiply this by pop_{wz}/pop_z, the estimated fraction of the population in zip code z that is served by water district w, since not all customers in a zip code are faced with a violation. This provides a measure of the zip code exposure to violations, and enables us to interpret β_2 as the percentage change in sales at each store from a violation affecting all people in the zip code. Controls for weekly mean maximum and minimum temperature and total precipitation, which account for potential time-varying preferences for bottled water, are included in **weather**. Store-water district fixed effects $\boldsymbol{\alpha}_{sw}$ capture time-invariant factors that affect the demand for bottled water, such as income and personal preferences. Temporal and seasonal trends in bottled water consumption are accounted for by year-week fixed effects $\boldsymbol{\delta}_t$. Finally, ε_{swt} is an error term that consists of a store specific term, a water district specific term, and an i.i.d. component. This multi-cluster approach allows for arbitrary serial correlation in sales within stores and for correlation between multiple zip codes served by one water district (Cameron, Gelbach and Miller, 2011).

Although we do not include prices in this econometric model, we contend that prices are unlikely to change in response to local water warnings. Prices are set weekly by a centralized marketing department of the grocery chain, and hence are unlikely to incorporate local conditions. Using data from the same chain, Schlenker and Villas-Boas (2009) found negligible changes in beef prices when Mad Cow disease resulted in drops of meat purchases by 20 percent. Changes in beef consumption were observed at all stores and hence a coordinated price response would have been much easier than store-specific responses to *local* shocks in water demand. Nonetheless, we create a price index at the store-week level to approximate local prices, and both use it as an independent and dependent variable in our estimation of equation (1).[3]

III. Results

Panel A of Table 3 presents our main results. Column 1 shows cross-sectional correlations, which should be interpreted with caution because violation frequency is correlated with other determinants of bottled-water consumption (as shown in Table 2). Accordingly, we do not find a statistically significant response to water quality violations, with coefficients on two of the violations having counterintuitive signs. Column 2, which includes store-water district fixed effects, indicates that controlling for the endogeneity of violations is essential. We obtain a 22 percent increase in bottled water sales from a microorganism violation, a 26 percent increase in response to nitrate violations, and a 17 percent increase from an element/chemical violation, with only the nitrate violation not statistically significantly different from zero. A larger response to microorganisms and nitrates is consistent with differences in reporting requirements (i.e., within 24 hours), though these differences are not statistically significant.

---

[3] We computed the price index as follows: 1) calculate the total quantity sold for each UPC-store combination; 2) calculate price level: price (at store-UPC-week) divided by average price (store-UPC); and 3) take the weighted average of price levels in (2) using the fixed basket of (1) as weights.

In Panel B, we explore the potential impact of violations on the price of bottled water using our created price index. In both the cross-sectional and fixed effects regressions, we do not find a statistically significant relationship between violations and prices, suggesting that local grocery stores did not change prices in response to the surge in demand from water quality violations. We also included the price index as an explanatory variable in a revised sales regression (not shown), and the violation coefficients remain unchanged.[4]

Heterogeneous responses to these violations may arise for at least two reasons: 1) more vulnerable individuals will have greater incentives to respond to violations to which they have a greater sensitivity; and 2) more forward looking individuals will be more responsive to violations that generate negative health consequence far into the future. As crude proxies for these, we separately include violations interacted with three Census measures of the zip code that a water district serves: median household income, the percent of population under age 5, and the percent of population over age 65. Table 4 shows the estimated responses to each type of violation for the bottom and top quartile of the zip code characteristic, along with a p-value from a t-test for whether the responses are the same across the two quartiles.[5] We find a greater response to microorganisms in communities with a larger elderly population, which is consistent with a greater response by vulnerable populations. However, we do not find a corresponding increase for communities with more young children. Consistent with element/chemical violations posing longer term health risks, we find that communities with wealthier households and a smaller number of individuals over age 65 had a larger response. We find no support for a differential

---

[4] We also explore several other alternative specifications, such as controlling for county-specific seasonality, and found our results to be generally robust to these alternatives. Our baseline regression in column 1 does not include price since it may be endogenous.

[5] We dropped the results showing interactions with percent under age 5 from Table 4 because all interaction terms were statistically insignificant (results are available upon request).

response to nitrates/nitrates, perhaps because they only pose a risk to bottle fed children that may not be well captured by our simple measure of child exposure in each zip code.

A considerable advantage from using bottled water sales as a measure of avoidance behavior is that it reflects a market-based activity that can be used to provide estimates of the cost of avoidance behavior. We perform the following back-of-the-envelope calculation to provide estimates of total expenditures on bottled water sales in the US from all violations in 2005:

$$(2) \qquad \text{Total costs} = \sum_c \sum_t \hat{\beta}_2 \times sales_{ct} \times \{violation_{cwt} \times (pop_{wc} / pop_c)\}$$

where $\hat{\beta}_2$ is the adjusted estimated coefficient from model (1).[6] $Sales_{ct}$ are interpolated sales on week t in county c, where we distribute the yearly US total in 2005 of \$14.9 billion (Datamonitor, 2005) evenly among all 52 weeks of the year, and then distribute the weekly sales volume across counties based on population. The fraction of days in week t that a violation occurred in each water district in each county for each of the types of violations is measured by **violation**$_{cwt}$. Finally, pop_{wc}/pop_c is the estimated fraction of population in county c that is served by water district w. This yields a rough approximation of a county-week time series of population exposure to violations, which when multiplied by the first two terms yields an estimate of weekly county expenditures in response to violations. We sum this across all counties and time periods to yield the total expenditures for the entire nation.

Clearly these estimates involves several unverifiable assumptions, namely that 1) our estimated responses to violations from Northern California and Nevada are representative of

---

[6] We adjust this coefficient by multiplying it by the exponent of the mean squared residual divided by two to account for Jensen's inequality.

changes throughout the country and at all bottled water retailers, and 2) consumption of bottled water sales are distributed throughout the year and county as we specify. Since 61.3 percent of all bottled water sales occurred at supermarkets (Datamonitor, 2005), we can scale our estimates accordingly if we are concerned that responses at supermarkets are different than responses at other retail establishments. Beyond that, there is unfortunately little more we can do with the data at hand. It nonetheless provides a useful starting point for discussing policy implications from water quality violations.

Our estimates indicate that, in 2005, people spent $11.34 million in response to microorganism violations, $1.77 million in response to nitrate violations, and $47.15 million in response to element/chemical violations. These estimates likely represents a lower bound of the true costs of avoidance behavior because it does not include other responses to violations, such as purchasing alternative beverages (e.g. juice), other actions people may have taken (e.g. boiling water), and more permanent responses (e.g. installing water filters).

IV. Discussion

This paper builds upon a nascent literature that examines the impacts of informational approaches to environmental regulation (Mansfield et al., 2006; Shimshack, Ward, and Beatty 2007; Neidell, 2009) to examine responses to warnings about drinking water violations. Unlike Shimshack, Ward, and Beatty (2007), who find a broad response to mercury warnings regardless of vulnerability, our work provides some evidence for a differentiated response across consumers. Neidell's (2009) examination of responses to smog alerts finds some evidence of a differentiated response based on vulnerabilities, but like Shimshack, Ward, and Beatty (2007) does not assess the costs of avoidance behavior. Mansfield, Johnson, and van Houtven (2006) rely on stated preference methods to provide estimates of the costs of avoidance behavior for

children in response to high ozone-pollution days.[7] Our work combines features of each of

these, but adds a considerable innovation by using market-based responses to information to

provide estimates of avoidance costs. Moreover, unlike many other forms of pollution, drinking

water quality violations are quite common. As a result, our estimates are quite large – US

consumers paid nearly $60 million in response to water violations in 2005 alone.

Even if bottled water purchases captured most of the behavioral response to violations,

this figure clearly understates the willingness to pay to avoid water violations because it ignores

the health consequences faced by those who did not limit their exposure (Harrington and

Portney, 1987).[8] Since nearly 20 million Americans become ill from consuming drinking water

contaminated with parasites, bacteria, or viruses each year (Reynolds, Mena, and Gerba, 2008), a

complete measure of willingness to pay would be considerably larger. Of course, the value of

avoiding violations must be compared to the costs of eliminating them.[9]   Assessing the costs of

violations reductions and how they vary across water systems is an important direction for future

research.

V. References

Cameron, Colin, Jonah Gelbach and Douglas Miller. 2011. "Robust Inference with Multi-way Clustering." *Journal of Business and Economic Statistics*, forthcoming.

Datamonitor. 2005. "Bottled Water in the United States: Industry Profile." Reference Code: 0072-0016. New York, NY.

---

[7] In a related vein, Moretti and Neidell (2011) infer avoidance costs associated with ozone, using an indirect method based on an instrumental variables approach. Deschenes and Greenstone (2007) also estimate avoidance costs in response to weather shocks, though it is not specific to informational regulation.

[8] Furthermore, since bottled water is generally less regulated than tap water, drinking bottled water carries its own set of potential health risks, and may not be an ideal substitute for tap water from a health perspective.

[9] Indeed, the Safe Drinking Water Act requires cost-benefit analyses to be used to determine safety standards.

Deschenes, Olivier and Michael Greenstone. 2007. "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the U.S." *National Bureau of Economic Research Working Paper 13178.*

Duhigg, Charles. 2009. "Toxic Waters: Clean Water Laws are Neglected, at a Cost in Suffering." *New York Times*, September 13, 2009.

Harrington, Winston, and Paul R. Portney. 1987. "Valuing the Benefits of Health and Safety Regulations." *Journal of Urban Economics*, 22(1): 101-112.

Mansfield, Carolyn, F Reed Johnson, and George Van Houtven. 2006. "The Missing Piece: Valuing Averting Behavior for Children's Ozone Exposures." *Resource and Energy Economics*, 28(3): 215–228.

Moretti, Enrico and Matthew Neidell. 2011. "Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles." *Journal of Human Resources*, 46(1): 154-175.

Neidell, Matthew. 2009. "Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations." *Journal of Human Resources*, 44(2): 450-478.

Reynolds, Kelly, Kristina Mena, and Charles Gerba. 2008. "Risk of waterborne illness via drinking water in the United States." *Reviews of Environmental Contamination and Toxicology* 192: 117-158.

Schlenker, Wolfram, and Sophia B. Villas-Boas. 2009. "Consumer and Market Response to Mad-Cow Disease." *American Journal of Agricultural Economics*, 91(4): 1140-1152.

Shimshack, Jay, Michael Ward, and Timothy Beatty. 2007. "Mercury Advisories: Information, Education, and Fish Consumption." *Journal of Environmental Economics and Management*, 53(2): 158-179.

Table 1. Number and mean duration of violations, November 2001- November 2005

	US		Our sample	
	# of violations	duration (days)	# of violations	duration (days)
microorganism	37,645	44.7	239	36.9
nitrates	3,798	153.9	25	65.7
elements/chemicals	13,261	131.5	21	223.2

Table 2. Community characteristics by violation frequency

	high	low	p-value
average weekly sales of bottled water ($)	3,500	4,227	0.00
median house price ($)	220,362	284,499	0.00
median household income ($)	49,120	55,820	0.00
population density	391	1442	0.00
% white, not Hispanic	70.89	74.83	0.03
% less than high school degree	17.68	13.05	0.00
% high school degree only	60.95	55.07	0.00
% college educated	21.37	31.88	0.00
% population under age 5	0.06	0.06	0.01
% population over age 65	0.12	0.14	0.00
January maximum temperature (F)	54.71	54.57	0.83
July maximum temperature (F)	84.71	87.05	0.03
January precipitation (cm)	0.26	0.29	0.03
July precipitation (cm)	0.00	0.00	0.51
microorganism violations	9.12	2.31	0.00
nitrates violations	2.06	0.08	0.00
elements/chemicals violations	1.46	0.10	0.00
observations	277	255	

Notes: "high (low)" violations defined as being above (below) than the median number of violations for the time period studied. "p-value" is from a t-test for the difference in values across the two columns.

Table 3. Cross-sectional and fixed effect estimates of water violations on bottled water

	1	2
Panel A. Dependent variable = log(sales)		
Microorganism	-1.031	0.219
	[0.666]	[0.076]**
Nitrates	-14.652	0.257
	[12.403]	[0.587]
Elements/chemicals	0.533	0.174
	[1.050]	[0.060]**
Panel B. Dependent variable = log(price index)		
Microorganism	-0.012	-0.011
	[0.008]	[0.007]
Nitrates	-0.025	-0.052
	[0.067]	[0.095]
Elements/chemicals	-0.007	-0.011
	[0.016]	[0.020]
Fixed effects	N	Y
Observations	41534	41534

Notes: * significant at 5%; ** significant at 1%. Standard errors clustered on both water district and store in brackets. All regressions controls for year-week indicators, average weekly minimum and maximum temperature, and average weekly precipitation. Regressions in column 2 include store-water district fixed effects.

Table 4. Heterogeneity of estimates of water violations on bottled water

	1	2	3
	Lowest Quartile	Top Quartile	p-value of difference
Panel A. Household Income			
Bugs	0.2492**	0.1811**	0.225
	[0.079]	[0.060]	
Nitrates	-0.6888	0.7751*	0.458
	[1.842]	[0.363]	
Elements	0.2714**	1.6893**	0.001
	[0.047]	[0.464]	
Panel B. Percent Over 65			
Bugs	0.1391**	0.3139**	0.026
	[0.044]	[0.067]	
Nitrates	3.4816	-2.3178	0.261
	[2.333]	[2.910]	
Elements	0.4602**	0.2805**	0.009
	[0.100]	[0.036]	

Notes: * significant at 5%; ** significant at 1%. See notes to Table 3.