Data Analysis

Michael Schramm

```
`modelsummary` 2.0.0 now uses `tinytable` as its default table-drawing
  backend. Learn more at: https://vincentarelbundock.github.io/tinytable/

Revert to `kableExtra` for one session:

  options(modelsummary_factory_default = 'kableExtra')

Change the default backend persistently:

  config_modelsummary(factory_default = 'gt')

Silence this message forever:

  config_modelsummary(startup_message = FALSE)
```

```
Loading required package: grid
```

```
Loading required package: Matrix
```

```
Loading required package: survival
```

```

Attaching package: 'survey'
```

```
The following object is masked from 'package:graphics':

    dotchart
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0
──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.5.0     ✔ tibble    3.2.1
✔ lubridate 1.9.3     ✔ tidyr     1.3.1
```

```
✔ purrr     1.0.2
── Conflicts ──────────────────────────────────────── tidyverse_conflicts()
──
✖ tidyr::expand() masks Matrix::expand()
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
✖ tidyr::pack()   masks Matrix::pack()
✖ tidyr::unpack() masks Matrix::unpack()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
Hoisted 7 variants:
Manrope Bold
Manrope ExtraBold
Manrope ExtraLight
Manrope Light
Manrope Medium
Manrope Regular
Manrope SemiBold
```

Methodology

Survey Instrument

The online survey was administered and distributed through the Qualtrics online panel in April 2023. Panels were recruited using gender, age, and race/ethnicity quotas to obtain approximately representative samples of the U.S. population. Panels were continously sampled until a representative ssample was obtained, with a total sample size of 1,100 respondent and estimated ±3% margin of error at 95% confidence (Dillman, 2011). Quality checks, performed by Qualitrics, excluded responses that failed attention or speed checks, IP address checks, or failed to meet representative demographic criteria. All participants were provided and agreed to a written informed consent agreement. The Texas A&M University Institutional Review Board reviewed the study protocol prior to distribution and deemed the study exempt from formal review.

To assess the the drinking water sources and issues of public concern, we asked participants for their main sources of household tap water and their main source of drinking water. We also asked if they had experienced any of 12 different potential drinking water issues. To assess concerns about contamination sources, we also asked how concerned they were about 10 different potential sources of contamination on a five point scale from *"not at all concerned"* to *"extremely concerned."*

Levels of responsibility and trust in different entities involved in drinking water management were assessed with 2 questions. First, participants were asked what level of responsibility (on a five point scale from *"no responsibility"* to *"full responsibility"*) they believed each of six different entities had for ensuring drinking water is safe. We also asked what the participants level of trust was in each of those enties (on a five point scale from *"do not trust at all"* to *"fully trust*").

To assess the public's perceived safety of drinking water, we asked participants to rate their drinking water safety on an ordinal scale from 0 through 10.

Analysis

Survey responses were weighted to match marginal sex/gender, age group, race/ethnicity, and education proportions to national level benchmarks from the 5-year 2021 American Community Survey (ACS). We chose to weight gender using "male" and "non-male" binary responses because the ACS only provides binary responses for sex. This approach groups "female" and "other" respondents under the same marginal weight. Race and ethnicity were recoded at white or Caucasian and non-white categories due to small subpopulaiton sample sizes. Weights were developed by poststratification raking using the American National Election Study (ANES) weighting algorithm implemented in the *anesrake* R package (DeBell & Krosnick, 2009; Pasek, 2018). Weighted summary statistics were calculated using the *survey* R package (Lumley, 2004).

Table 1: Marginal survey and target population proportions and marginal weighted values.

Value	Unweighted N	Unweighted %	Target %	Weighted N	Weighted %
Age					
18:24	125	11.4	11.9	130.6	11.9
25:34	192	17.5	17.7	195.1	17.7
35:44	204	18.5	16.6	183.1	16.6
45:54	198	18.0	16.3	179.2	16.3
55:64	171	15.5	16.8	184.4	16.8
65+	208	18.9	20.7	227.6	20.7
No answer	2	0.2	NA	NA	NA
Education					
Some high school	47	4.3	7.8	85.8	7.8
High school graduate or GED	418	38.0	49.4	543.7	49.4
Associate degree	178	16.2	8.3	91.3	8.3
Bachelor's degree	246	22.4	19.4	213.7	19.4
Master's degree	132	12.0	8.3	91.3	8.3
Doctorate or terminal degree	28	2.5	1.3	14.7	1.3
Other	40	3.6	5.4	59.5	5.4
No answer	11	1.0	NA	NA	NA
Race/Ethnicity					
White	723	65.7	62.4	686.3	62.4
Non-white	373	33.9	37.6	413.7	37.6
No answer	4	0.4	NA	NA	NA
Sex/Gender					
Male	529	48.1	49.0	539.1	49.0
Not Male	569	51.7	51.0	560.9	51.0
No answer	2	0.2	NA	NA	NA

Models

We used a multinomial logistic regression model (Agresti, 2002) to explore variables associated with the public's choice of drinking water source (unfiltered tap water, filtered tap water, bottled water, or other). We used proportional odds models (Agresti, 2002; McCullagh, 1980) to explore factors associated (1) trust in entities responsible for drinking water safety, and (2) perceptions of drinking water safety. Multinomial logistic regressions were fit using the *svyVGAM* R package (Lumley, 2023) and proportional odds models were fit using the *survey* R package (Lumley, 2004). We calculated the generalized variance inflation factor (GVIF) to evaluate potential impacts of multi-collinearity among model predictors on our interpretation of the models (Fox & Monette, 1992). We report our model results as odds-ratios.

Primary choice in drinking water

Prior research provides evidence that age, gender, perceptions of drinking water safety, category of water supplier, and dissatisfaction with taste/color are commonly associated with the higher likelihoods of bottled water usage (Doria, 2006; Gorelick et al., 2011; Hu et al., 2011). We modeled choice of primary drinking water source (unfiltered tap water, filtered tap water, or bottled water) using demographic variables (Table 1), source of household tap water, and experience with any of (1) odd taste, (2) odd smell, (3) discolored water, or (4) cloudy water. We hypothesize that …

Levels of trust

Although public trust in government institutions is at historically low levels (Pew Research Center, 2022), there is evidence that the public may have higher trust in local public water utilities relative to other state, local, and federal institutions (Weisner et al., 2020). However, this effect might be confounded by experience with specific local drinking water problems. We modeled the level of trust participants had in their local water provider using demographic variables, household tap water source, and experience with any of (1) odd taste, (2) odd smell, (3) discolored water, or (4) cloudy water.

Perception of drinking water safety

We explored factors associated with perceived safety in drinking water using a proportional odds model relating sex/gender, age, race/ethnicity, education, home ownership, and household tap water supply to self-reported rating (on scale 1-10) of drinking water safety.

Results

Drinking water sources and issues

The majority of respondents lived in homes with a municipal public water supply (56%, 1.6% SE) with 13.3% using some type of private water supply (Table 2). A relatively high percentage of respondents stats they did not know what the source of their tap water was (13.3%, 1.2% SE). The majority of respondents (65.5%) used unfiltered or filtered tap water as their main source of drinking water (Table 2). Bottled or pre-packaged water was the primary drinking water sources for 34.2% (1.6% SE) of respondents. Only 3.1% (0.6% SE) of respondents stated that they had not experienced any issues with their tap water (Table 3). Odd taste (27.5%, 1.4% SE), hardness or scale buildup (20.3%, 1.3% SE), and odd smell (18.8%, 1.3% SE) were the most common issues. Notably, 14.9% (1.1% SE) of respondent stated they have experienced a boil water notice.

Table 2: The proportion of respondents with the specified household water source and drinking water source.

Response	Proportion	SE
What is the source of your household tap water?		
Public supply - municipal	0.561	0.016
Public supply - rural water district	0.163	0.012
Private supply - well, river, pond	0.114	0.01
Private supply - rainwater harvest system	0.019	0.004
I don't know	0.139	0.012
Other	0.004	0.002
What is your main source of drinking water?		
Unfiltered tap water	0.28	0.015
Filtered tap water	0.375	0.015
Bottled/prepackaged water	0.341	0.016
Other	0.003	0.001

Table 3: Proportion of respondents that have experienced specified issues with their drinking water.

Issue	Proportion	SE
Odd taste	0.274	0.014
Odd smell	0.187	0.013
Discolored water	0.117	0.01
Cloudy water	0.148	0.011
Staining of clothes, teeth, or skin	0.055	0.007
High salt content	0.031	0.005
Hardness or scale buildup	0.202	0.013
Corrosion of pipes or water fixtures	0.081	0.009
Boil water notice	0.149	0.011
Water shortage	0.094	0.009
None of the above	0.031	0.006
Other	0.031	0.006

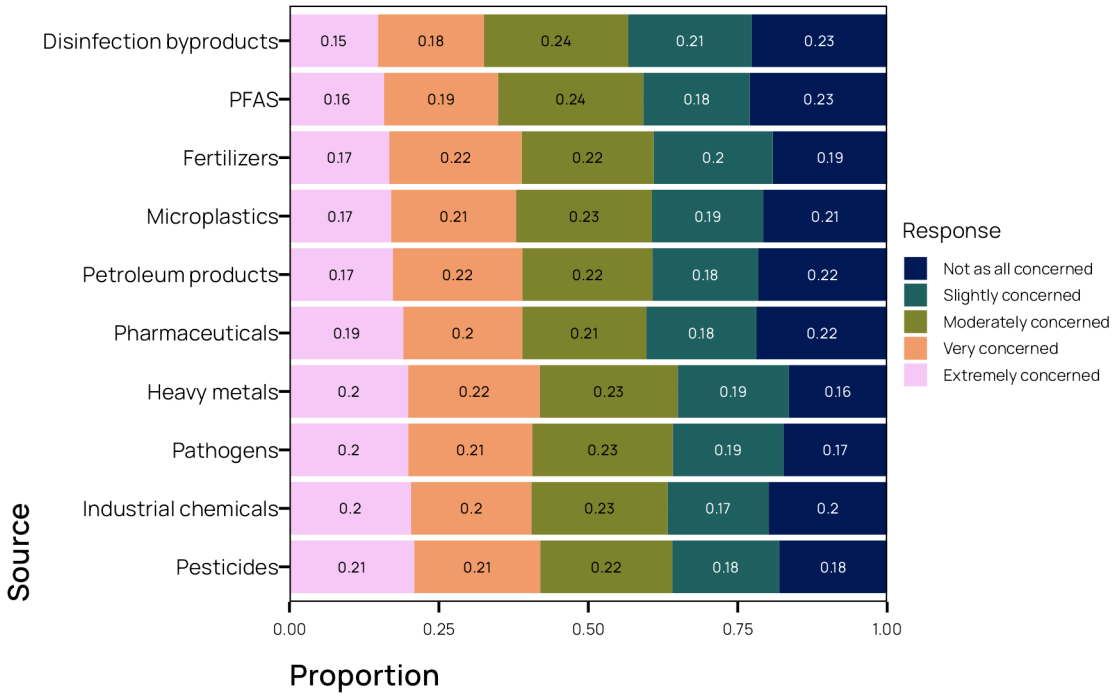

Figure 1: Contaminants of concern

We did not observe strong differences among response levels between different contaminants of potential concern (Figure 1). Pesticides (*extremely concerned* = 21%, 1.3% SE; *very concerned* = 21.1%, 1.3% SE), industrial chemicals (*extremely concerned* = 20.3%, 1.3% SE; *very concerned* = 20.1%, 1.3% SE), pathogens (*extremely concerned* = 19.9%, 1.3% SE; *very concerned* = 20.8%, 1.3% SE), and heavy metals (*extremely concerned* = 19.8%, 1.3% SE; *very concerned* = 22.0%, 1.3% SE) are sources that elicited the highest levels of concern across all respondents. Disinfection byproducts (*Not at all concerned* = 22.6%, 1.4% SE; *slightly concerned* = 20.7%, 1.3% SE) and PFAS (*not at all concerned* = 23.0%, 1.4% SE; *slightly concerned* = 17.8%, 1.2% SE) were sources of least concern.

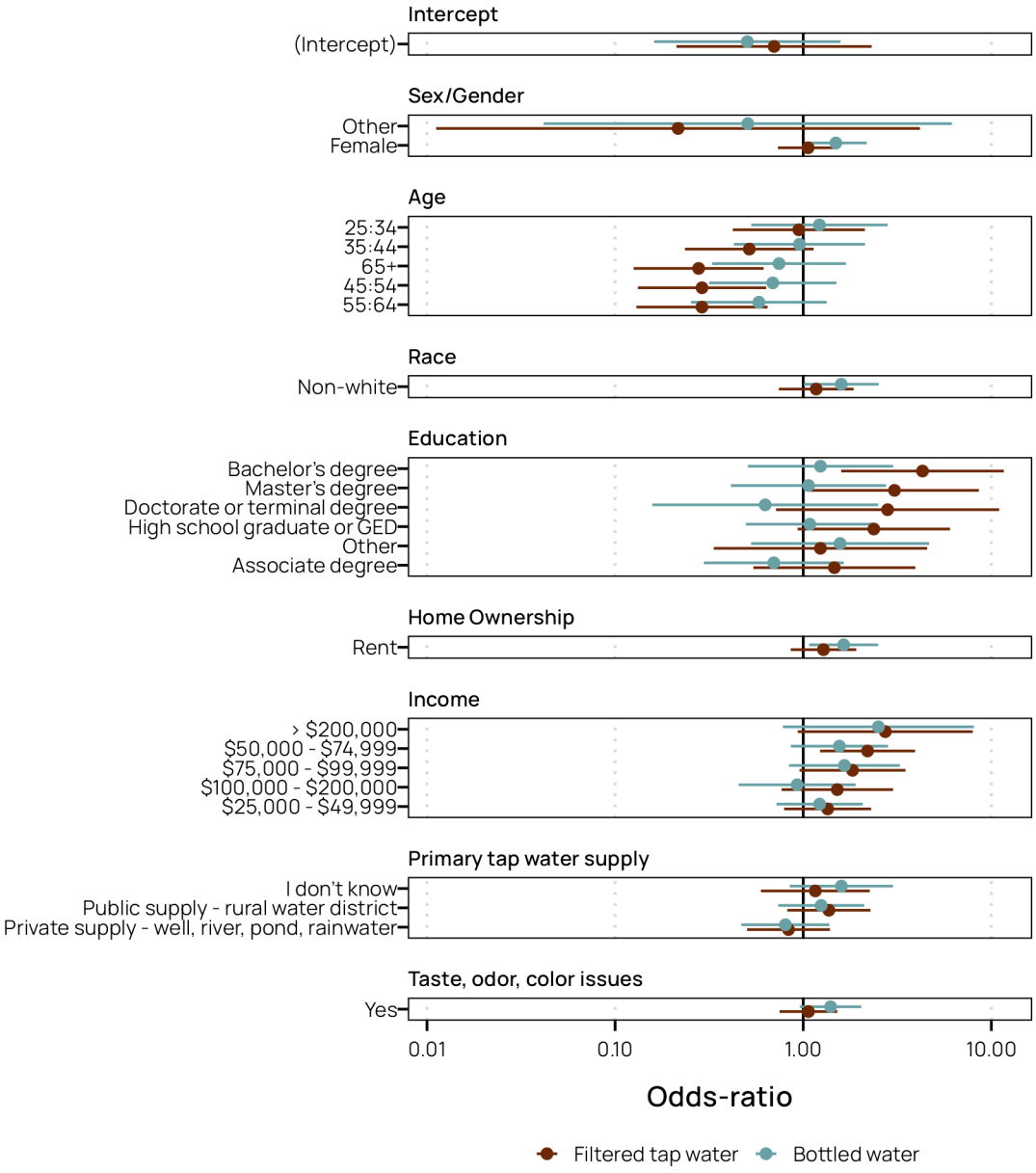

Figure 2: Effects of demographic variables, drinking water source, and experience with drinking water issues on drinking water choice. Odds-ratios are relative to unfiltered tap water.

Based on the multinomial regression model, we found limited effects of demographic variables, drinking water sources, and experience with odd taste, color, or smell on the odds of using bottled drinking water compared to unfiltered drinking water. Notably, females had 1.49 times the odds (95% CI [1.02, 2.17]) of using bottled drinking water instead of unfiltered tap water compared to males. Non-whites had 1.59 times the odds (95% CI [1.01, 2.52]) of using bottled drinking water instead of unfiltered drinking water compared to white respondents. Renters had 1.64 times the odds (95% CI [1.07, 2.50]) compared to homeowners.

Conversely, older respondents were more likely to use unfiltered drinking water compared to filtered water (45:54 age group: OR = 0.29, 95% CI [0.132, 9.636]; 55:64 age group: OR = 0.29, 95% CI [0.130, 0.647]; 65+ age group: OR = 0.278, 95% CI [0.002, 1.25]) but there was not strong evidence for differences in odds between bottled water and unfiltered drinking water among any of the age groups. Education also showed strong effects on the use of filtered drinking water. Having a bachelor's degree resulted in 4.31 times the odds of using filtered drinking water compared to those not graduating high school. A master's degree had a similarly large effect (OR = 3.06, 95% CI [1.09, 8,59]).

Levels of responsibility and trust

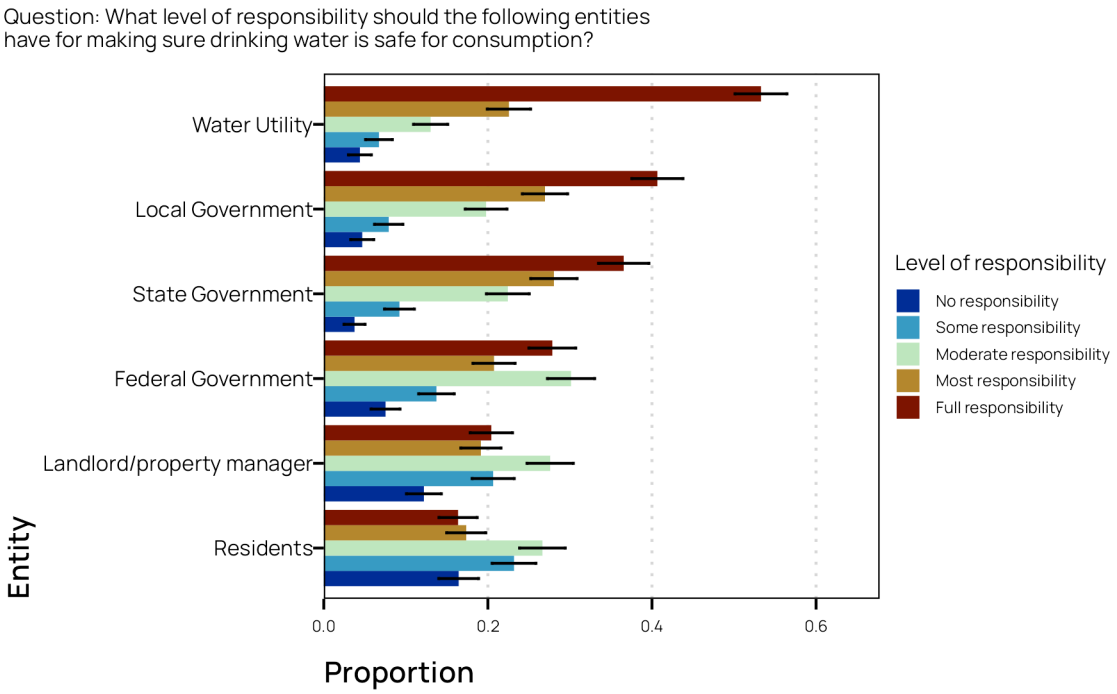

Figure 3: Weighted proportional responses for the level of responsibility people believe different entities have for ensuring drinking water safety. Bar lengths represent the mean proportional response with eror bars representing the 95% confidence interval.

In general, higher proportions of survey respondents assigned water utilities (53.3%, 95% CI [50.1%, 56.5%]), local government (40.7%, 95% CI [37.5%, 43.8%]), and state government (36.6%, 95% CI [33.5%, 39.6%]) with the highest levels of responsibility for ensuring safety of drinking water (Figure 3). Moderate reponsibilty was the most common response for the level of responsibility attributed to household residents (26.6%, 95% CI [23.9%, 29.4%]) and landlords/property managers (27.6%, 95% CI [24.8%, 30.4%]). A relatively high proportion of responses indicated residents have no responsibilty for drinking water quality (16.4%, 95% CI [14.0%, 18.9%]). Relative to other govenrmental institutations, fewer people felt the federal government had full responsibility (27.9%, 95% CI [25.0%, 30.7%]) for drinking water safety.

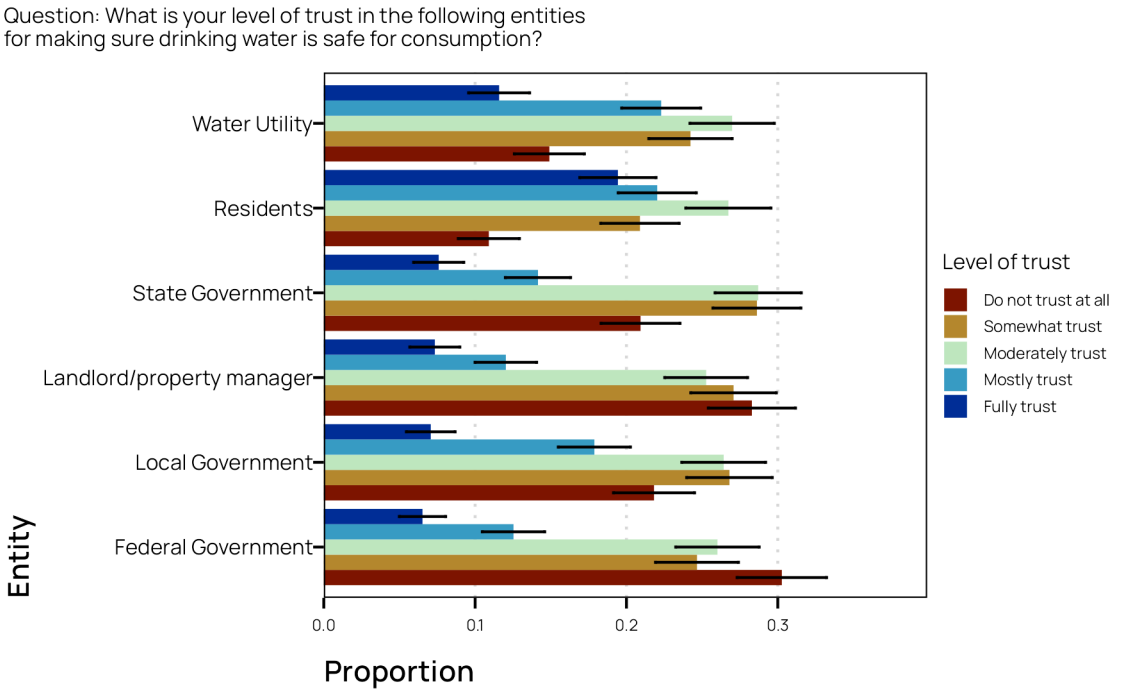

Figure 4: Weighted proportional responses for the level of trust people have in the entities responsible for ensuring drinking water safety. Bar lengths represent the mean proportional response with eror bars representing the 95% confidence interval.

When asked for their level of trust in these entities for ensuring safety of drinking water, 30% (95% CI [27.3%, 33.2%]) of respondents expressed they do not trust the federal government at all (Figure 4). Only 6.5% (95% CI [4.9%, 8.1%]) indicated they fully trust the federal government for ensuring drinking water safety. State government fared slightly better, with 28.7% (95% CI [25.9%, 31.6%]) and 28.6% (95% CI [25.7%, 31.5%]) of respondents indicating they moderately trust or somewhat trust the state government for ensuring safety. Local and state government had rates of full trust (7.1%, 95% CI [5.5%, 8.7%] and 7.6%, 95% CI [5.9%, 9.3%], respectively) comparable to the low rate observed for the federal government (6.5%, 95% CI [5.0%, 8.1%]).

Water utilities appear to be more trusted by the public than other governmental institutions. While the proportion of responses indicating moderate trust (27.0%, 95% CI [24.2%, 29.8%]) and mostly trust (22.3%, 95% CI [19.7%, 24.9%]) are comparable to local, state, and federal government, much fewer people indicated that the "do not trust at all" the water utility (14.9%, 95% CI [12.6%, 17.2%]).

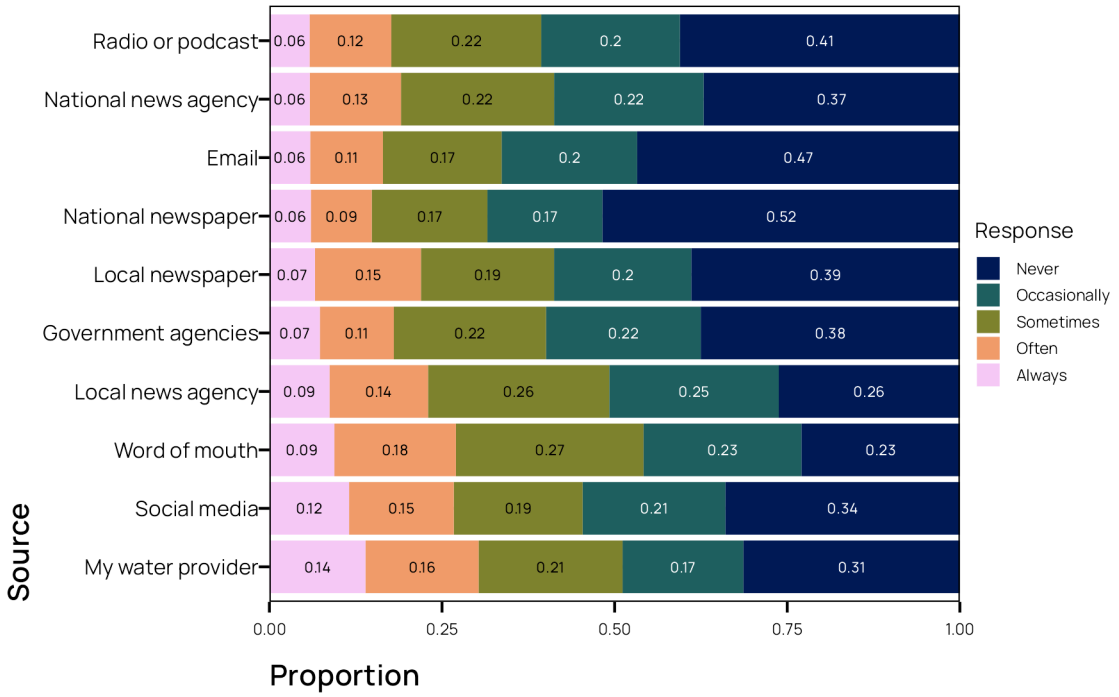

Figure 5: Sources of drinking water quality information.

Water providers appear to be the most relied upon source for information on drinking water with 13.9% (1.1% SE) and 16.4% (1.2% SE) of responses indicating water providers are *always* or *often* used as a source of information (Figure 5). Both social media (*always* = 11.5%, ±1% SE; *often* = 15.2%, ±1.1% SE) and word-of-mouth (*always* = 9.4%, ±1% SE; *often* = 17.6%, ±1.2% SE) are also frequent sources of information with similar or higher usage than local news agencies (*always* = 8.7%, ±0.9% SE; *often* = 14.3%, ±1.1% SE) or local newspapers (*always* = 6.6%, ±0.8% SE; *often* = 15.3%, ±1.1% SE). Some of the least utilized sources included radio or podcasts (*never* = 40.5%, ±1.6% SE; *occasionally* = 20.1%, ±1.3% SE), email (*never* = 46.7%, ±1.6% SE; *occasionally* = 19.6%, ±1.3% SE), national newspapers (*never* = 51.8%, ±1.6% SE; *occasionally* = 16.7%, ±1.2% SE), and government agencies (*never* = 37.5%, ±1.6% SE; *occasionally* = 22.4%, ±1.4% SE).

Perceptions of drinking water safety

Table 4: Odds-ratios and 95% confidence intervals for drinking water quality rating. Higher odds-ratio indicates higher odds of providing a higher rating compared to the reference level rating.

	Variable	Odds-ratio	95% CI	t-statistic	p-value
Sex/Gender	Male	•			
	Female	0.76	[0.59, 0.97]	−2.17	0.03
	Other	0.13	[0.05, 0.32]	−4.49	<0.01
Age	18:24	•			
	25:34	0.94	[0.59, 1.50]	−0.24	0.81
	35:44	0.82	[0.50, 1.35]	−0.77	0.44
	45:54	0.45	[0.27, 0.75]	−3.07	<0.01
	44:64	0.73	[0.44, 1.23]	−1.18	0.24
	65+	1.16	[0.69, 1.96]	0.56	0.58
Race	White	•			
	Non-white	0.80	[0.59, 1.09]	−1.41	0.16
Education	Some high school	•			
	High school or GED	1.17	[0.54, 2.52]	0.39	0.69
	Associates degree	0.96	[0.43, 2.15]	−0.10	0.92
	Bachelor's degree	1.41	[0.63, 3.16]	0.84	0.40
	Master's degree	1.87	[0.82, 4.27]	1.48	0.14
	Doctorate/terminal degree	1.57	[0.64, 3.89]	0.98	0.33
	Other	0.83	[0.32, 2.19]	−0.38	0.71
Home ownership	Own	•			
	Rent	0.67	[0.51, 0.89]	−2.83	<0.01
Income	<$25,000	•			
	$25,000-$49,999	1.19	[0.82, 1.73]	0.93	0.35
	$50,000-$74,999	1.08	[0.73, 1.62]	0.39	0.70
	$75,000-$99,999	0.91	[0.58, 1.43]	−0.42	0.68
	$100,000-$200,000	1.50	[0.96, 2.35]	1.80	0.07
	>$200,000	1.82	[0.89, 3.71]	1.65	0.10
Household Tap Water Supply	Public supply - municipal	•			
	Public supply - rural water district	0.81	[0.58, 1.12]	−1.28	0.20
	Private supply - well, river, pond, rainwater	1.54	[1.10, 2.16]	2.49	0.01
	I don't know	0.98	[0.61, 1.58]	−0.08	0.94
	Other	3.74	[0.01, 1265.58]	0.44	0.66
Taste, Odor, Color Issues	No	•			
	Yes	0.32	[0.24, 0.41]	−8.79	<0.01

The proportional odds model provides evidence that sex/gender, home ownership, household tap supply, and experience with taste/odor/color issues are associated with the perception of drinking water safety (Table 4). Females were 1.32 times more likely to rate their drinking water less safe than males (OR = 0.76, 95% CI [0.59, 0.97]). Respondents identifying their gender as other had 7.69 times the odds (OR = 0.13, 95% CI [0.05, 0.32]) of rating their drinking water less safe than male respondents. Home ownership provided a strong signal with renters 1.49 times more likely to provide a lower drinking water safety ranking than home-owners (OR = 0.67, 95% CI [0.51, 0.89]). Respondents on private water supplies had 1.54 times that odds of providing a higher drinking water safety rating than people on municipal supplies (OR = 1.54, 95% CI [1.10, 2.16]). People that experienced taste/odor/color issues with 3.12 times more likely to provide a lower safety rating than those that never experienced an issue (OR = 0.32, 95% CI [0.24, 0.41]).

We identified difference among some age groups, but a clear pattern was not evident. Respondents in the 45-54 age bracket were 2.22 times more likely to provide a lower drinking water safety rating than the 18-24 year old age group.

note- we could add a plot of differences in marginal probabilities, but I think there are too many levels in our factors, plus really long labels.

Discussion

Bibliography

Agresti, A. (2002). *Categorical Data Analysis* (2nd ed). Wiley-Interscience.

DeBell, M., & Krosnick, J. A. (2009). *Computing Weights for American National Election Study Survey Data* (Issue nes12427). http://www.electionstudies.org/resources/papers/nes012427.pdf

Dillman, D. A. (2011). *Mail and Internet Surveys: The Tailored Design Method*. Wiley.

Doria, M. F. (2006). Bottled Water versus Tap Water: Understanding Consumers' Preferences. *Journal of Water and Health*, *4*(2), 271–276. https://doi.org/10.2166/wh.2006.0023

Fox, J., & Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, *87*(417), 178–183. https://doi.org/10.1080/01621459.1992.10475190

Gorelick, M. H., Gould, L., & Nimmer, M. (2011). Perceptions about Water and Increased Use of Bottled Water in Minority Children. *Archives of Pediatrics & Adolescent Medicine*, *165*(10), 928. https://doi.org/10.1001/archpediatrics.2011.83

Hu, Z., Morton, L. W., & Mahler, R. (2011). Bottled Water: United States Consumers and Their Perceptions of Water Quality. *International Journal of Environmental Research and Public Health*, *8*(2), 565–578. https://doi.org/10.3390/ijerph8020565

Lumley, T. (2004). Analysis of Complex Survey Samples. *Journal of Statistical Software*, *9*(8). https://doi.org/10.18637/jss.v009.i08

Lumley, T. (2023). *svyVGAM: Design-Based Inference in Vector Generalised Linear Models*. https://cran.r-project.org/package=svyVGAM

McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, *42*(2), 109–142. https://doi.org/10.1111/j.2517-6161.1980.tb01109.x

Pasek, J. (2018). *anesrake: ANES Raking Implementation*. https://cran.r-project.org/package=anesrake

Pew Research Center. (2022). *Americans' Views of Government: Decades of Distrust, Enduring Support for Its Role*. https://www.pewresearch.org/politics/2022/06/06/americans-views-of-government-decades-of-distrust-enduring-support-for-its-role/

Weisner, M. L., Root, T. L., Harris, M. S., Mitsova, D., & Liu, W. (2020). The Complexities of Trust between Urban Water Utilities and the Public. *Sustainable Water Resources Management*, *6*(3), 50. https://doi.org/10.1007/s40899-020-00407-6