Per-and polyfluoroalkyl substances (PFAS) in Drinking Water: A Socio-economic Perspective

Nabin B. Khanal[1] and Levan Elbakidze[1]

[1]West Virginia University

Abstract

Contamination of drinking water with PFAS poses a significant public health threat. We examine the spatial distribution of PFAS in the US public water systems (PWS) and explore the relationship between PFAS contamination, PWS characteristics, and socioeconomic and industrial attributes of the affected communities. Using data from the third Unregulated Contaminant Rule (UCMR3), the Census Bureau of Statistics, and the Bureau of Labor Statistics (BLS), we identify spatial contamination hot spots and find that PFAS contamination is correlated with PWSs size, non-surface water sources, population, and housing density. We also find that non-white communities have lower PFAS in drinking water. Finally, we detect some evidence of PFAS contamination being associated with regional industrial structure.

Key Word: Per-and polyfluoroalkyl substances (PFAS), Public Water Supply system (PWS), Drinking water, Income, Demographics

Introduction

Per- and Polyfluoroalkyl Substances (PFAS) cause a range of serious health problems, including cancer, hyperlipidemia, thyroid disease, immunodeficiency, ulcerative colitis, chronic kidney disease, coronary artery disease, hypertension, and reduced fertility (CDC, 2022; EPA, 2022; Andersson et al., 2019; Steenland et al., 2010). Prior studies show that 95% of US adolescents and adults are exposed to PFAS, primarily through drinking water (Kato et al., 2011; De Silva et al., 2021). Between 16 and 270 million people in the US rely on PFAS-contaminated drinking water daily, which suggests the need for a better understanding of the incidence and distribution of PFAS in public water systems (PWS) (Hu et al., 2016; Andrews and Naidenko, 2020).

To develop effective public health policies aimed at mitigating the impact of PFAS contamination a comprehensive analysis addressing several fundamental questions is imperative. First, how pervasive is PFAS in drinking water, and are there regional clusters of contamination? Second, does PFAS contamination depend on PWS characteristics like size and water sources? Third, are some communities more vulnerable than others? Finally, is PFAS drinking water contamination driven more by industrial production or by final consumption?

PFAS are a group of 9000 synthetic chemicals widely used in industrial processes and consumer goods for their stain, grease, water, and heat-resistant properties (Cordner et al., 2019; National Institute of Environmental Health Science (NIH), 2019; Glüge et al., 2020). The use and production of PFAS dated back to more than 70 years ago when it was used for uranium separation in the Manhattan Project (Department of Energy (DOE), 2022)[1]. Since then, these substances have become ubiquitous due to their widespread use, bioaccumulation, resistance to degradation, water insolubility, and the ability to translocate easily from one system to another through biological or physical means (De Silva et al., 2021). Close to 180 PFAS have been identified as toxic and added to the Toxic Release Inventory list under the National Defense Authorization Act (EPA 2022a).

Detrimental health impacts of PFAS have not been well understood, documented and recognized until recently. It wasn't until the EPA's Health Advisories were revised in 2022 that the safe levels of PFOA, PFOS, and other PFAS were significantly reduced, suggesting that even low exposure can have detrimental health impacts (Federal Register, 2022; Federal Register, 2016). The 2016 health advisories for PFOA and PFOS indicated that less than 70 ppt (Part Per Trillion) posed no health risks, while the 2022 advisory lowered the threshold to 0.004 and 0.02 ppt, respectively. The addition of GenX (Hexafluoropropylene Oxide Dimer Acid and its Ammonium Salt) and PFBS to the list of hazardous PFAS further highlights the growing recognition of the danger that these chemicals pose. As more research is conducted and the risks associated with PFAS become better understood, it is critical that an appropriate public policy is developed to minimize exposure and prevent further contamination of the environment and drinking water sources.

[1] In 2021, the Department of Energy (DOE) issued a Departmental policy which aimed to reduce or eliminate PFAS release from departmental operations (DOE, 2022). Part of the DOE’s objectives is to identify and quantify Cold War era sources of PFAS including uranium processing operations going back to the Manhattan Project.

In response to growing public concerns, the EPA has announced a PFAS strategic Roadmap in October 2021 (EPA, 2021). The roadmap outlines the agency's plans to protect the public and the environment from PFAS contaminants by minimizing their discharge into the environment, identifying and removing them from ecosystems, and designating PFOA and PFOS as hazardous compounds under Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) (EPA, 2021). The EPA has also committed to conducting environmental and health toxicity assessments for additional PFAS, including PFBA, PFGxA, PFGxS (Perfluorohexanesulfonic acid), and PFDA (Perfluorodecanoic Acid) (EPA, 2021). Additionally, the roadmap includes provisions to ensure that disadvantaged communities have access to PFAS mitigation solutions. To make progress towards these goals, one must understand the distribution of exposure to PFAS via drinking water, including regional contamination clusters, differences across large and small public water systems, vulnerability of certain communities, and socioeconomic factors associated with contamination.

Protecting drinking water from PFAS contamination is a complex challenge. First, PFAS substances are unregulated under the Safe Drinking Water Act (SWDA), which means that PWSs are not required to monitor and control PFAS in the water they supply. Second, there are no effective technological solutions for removing PFAS from drinking water. Third, the sources of contamination are not well understood, making it challenging to prevent future contamination. Although research is ongoing, and some technological solutions are emerging, there is currently no available technology for the effective removal of PFAS from drinking water (Trang et al., 2022). Fourth, PFAS substances are persistent, bioaccumulative, and do not break down easily, increasing the risk of exposure and making it more challenging to remove them from the environment.

Environmental Protection Agency (EPA) uses the Unregulated Contaminant Monitoring Rule (UCMR) to assess the presence of contaminants that do not have health-based standards under the Safe Drinking Water Act (SDWA). Every five years, the EPA identifies 30 potentially harmful but unregulated contaminants under the UCMR program and tests all large Public Water Systems (PWSs) that serve more than ten thousand people, as well as a subsample of smaller facilities. In the UCMR3 program, the EPA tested six per- and polyfluoroalkyl substances (PFAS): perfluorooctanesulfonic acid (PFOS), perfluorooctanoic acid (PFOA), perfluorononanoic acid (PFNA), perfluorohexanesulfonic acid (PFHxS), perfluoroheptanoic acid (PFHpA), and perfluorobutanesulfonic acid (PFBS) in 2014-2016. Water samples were collected at the entry points to the distribution system by PWS operators and sent to an EPA-approved lab to test for the presence of each PFAS (EPA, 2012; EPA, 2017).

PWSs deliver drinking water to 95% of the US population (EPA, 2023a) and the UCMR3 program tested PWSs that serve 75% of the US population (EPA, 2016). UCMR3 tested 4,120 large PWSs that serve more than 10,000 consumers and randomly selected 800 representative small PWSs that serve 10,000 or fewer consumers in USA and its territories (EPA, 2017). Cumulatively, this sample represents 79% of the U.S. PWS consumer base. Many PWSs have more than one water supply facility, and UCMR3 tested all 15,195 facilities within the selected PWSs in the USA and its territories.

The EPA established minimum reporting levels (MRLs) for various types of PFAS, ranging from 10 to 90 ng/L (EPA, 2017)[2]. PFAS detection below the MRL is not reported to the EPA and is considered a "no detection". Total of 1152 detections were reported from 33 US states and three territories. PFBS (n=19) and PFNA (n=19), which is a replacement chemical for PFOS, only accounted for 3% (n=38) of the cases, while PFOA (n=379) and PFOS (n=292) accounted for 58% (n=671) of detections. PFHpA (n=236) and PFHxS (n=207) accounted for the remaining 39% (EPA, 2017). Although the recent Health Advisory (HA) by the EPA (EPA, 2022) included PFOA, PFAS, PFBS, and GenX (Hexafluoropropylene Oxide Dimer Acid and its Ammonium Salt), UCMR3 did not have enough positive data on PFBS and did not collect any data on GenX in UCMR3.

Several studies have utilized the UCMR3 PFAS data (Cadwallader et al. 2022; Andrews & Naidenko, 2020; Guelfo & Adamson, 2018; Hu et al., 2016). Hu et al. (2016) and Andrews and Naidenko (2020) estimated the population exposed to PFAS. Hu et al. (2016) estimated 16.5 million people are exposed to PFAS via drinking water. Andrews and Naidenko (2020) used UCMR3 and state level data from Colorado, Kentucky, Michigan, New Hampshire, New Jersey, North Carolina and Rhode Island. Their state level data had lower MRL than UCMR data, which enabled them to augment the population exposure estimates from UCMR. They also included private wells contamination data from Michigan and New Jersey. Extrapolating their findings from the states in their sample to all states in the US, they estimated the exposure rate of between 18 and 80 million people if the MRL was 10 ng/L and higher, and over 200 million if the MRL was at or above 1 ng/L. These estimates are

[2] The MRLs are 10 ng/L for PFHpA, 20 ng/L for PFOA and PFNA, 30 ng/L for PFHxS, 40 ng/L for PFOS, and 90 ng/L for PFBS.

significantly higher than the estimates based on UCMR3 data. Their study did not explore the relationship between contamination and community characteristics.

Guelfo & Adamson (2018) used the UCMR3 data and investigated the co-occurrence of different types of PFAS and relationship between PWS characteristics and contamination. They found six co-occurring PFAS pairs (PFOS/PFOA, PFOS/PFHxS, PFOS/PFHpA, PFOA/PFHxS, PFOA/PFHpA, and PFBS/PFHpA). PFHpA, PFOA, and PFNA were dominant in surface water whereas PFOS, PFHxS, and PFBS were dominant in groundwater. Large PWSs were more vulnerable to contamination than small ones. However, it is important to note that this study was limited only to the UCMR3 data and did not include other variables that may have an impact on contamination. To address this limitation, we expand on this study by including different socioeconomic and industrial variables from various data sources. This gives us a more comprehensive view of the factors that may contribute to PFAS contamination in drinking water.

Some of the potential sources of PFAS contamination in UCMR3 data are examined in Hu et al. (2016). Using 8-digit HUC (Hydrologic Unit Code) level data, they examined the spatial correlation between presence of Major industrial site, MFTAs (Military Fire Training Area), AFFF (Aqueous Film Forming Foam) use certified airports, Wastewater Treatment Plants (WWTPs) and concentration of PFOA, PFAS, PFHpA, and PFHxS in the PWSs. They found that military fire training sites (MFTAs) were strong predictors of PFOS and PFHxS, and wastewater treatment plants (WWTPs) predicted a modest increase in PFOA, PFOS, and PFHxS concentrations. However, none of the factors they examined predicted PFHpA concentration. Their spatial autocorrelation analysis only included 16 industrial sites in the

USA that participated in the EPA's 2010/2015 PFOA Stewardship program. Hence, the industrial contribution to PFAS contamination is examined using a limited set of industrial sites that participated in the PFOA stewardship. We contribute by examining the association with broader industrial categories using county scale GDP proportions. We also extend Hu et al. (2016) analysis by identifying contamination hotspots.

Cadwallader et al. (2022) assessed the impact of limitations in UCMR3 data, including the MRL and limited inclusion of small PWS. Using UCMR3 and state-level PFAS test data from 17 states with lower MRLs requirement, they found that for most PFAS except PFHxS and PFHpA, incomplete representation of small PWSs and higher MRL in UCMR3 had no impact on contamination predictions.

We contribute to previous literature with an examination of the relationship between PFAS contamination and PWS characteristics, socioeconomic factors, and regional industrial composition. Using publicly available data from UCMR3, Bureau of Labor Statistics (BLS), and American Community Survey (ACS) 5-year estimate, we identify geographic contamination hotspots and socioeconomic and industrial characteristics that are correlated with PFAS detection.

Data

County data on population, non-white population, per capita income, poverty, and housing density were collected from the American Community Survey 5-year estimates (US Census Bureau (USCBS), 2023). Poverty percentage is the share of poor[3] relative to total

[3] A household is deemed poor if income, adjusted by family size, falls below the threshold set by the US Census Bureau (Creamer et al., 2022). For instance, if a household consists of one person under 65 years of age and their

people in the county. Non-white population is the difference between county total population and Caucasian (white) population. Number of non-white poor individuals is estimated by subtracting the number of white poor individuals from the total number of poor individuals. Finally, non-white poverty rate is the share of non-white poor individuals relative to total non-white population in the county. On an average there are 23% non-white people in the county which ranges from 1% to 84%. Overall poverty is 16% (4 – 44%), with 13% (3 to 39%) for white population and 24% (0 – 66%) in for nonwhite population (Table 1).

The county Gross Domestic Product (GDP) and shares of GDP from Agriculture, Forestry, Fishing and Hunting; non-durable and durable goods manufacturing; Health Care and Social Assistance; Accommodation and Food Services; and government enterprise were obtained from the U.S. Bureau of Labor Statistics (BLS). Sectoral contribution (in the percentage) is calculated by dividing the total GDP from the selected sectors by total GDP of the county. On an average Government enterprise has a highest contribution to the county GDP (14%) followed by health care (8%) durable good manufacturing (7%), non-durable good manufacturing (6%), food and accommodation (3%), and agriculture (2%) (Table 1).

National PFAS contamination data was retrieved from the US EPA's National Contaminant Occurrence Database (NCOD), which was collected by the UCMR3 program from 2013 to 2016. UCMR3 is considered the most comprehensive national data source on PFAS contamination in Public Water Systems (PWS). Our study focuses on the 48 lower US states, excluding Washington DC. In the lower USA total of 35,589 water samples were

annual income is less than $14,097, then they are considered below the poverty line. However, if the household consists of three people, the threshold increases to $21,559, and this threshold gradually rises as the number of individuals in the household increases.

collected from 1,616 counties, 4,782 PWSs, and 14,607 water supply facilities at the entry point to the distribution system. The selected PWSs were tested quarterly or bi-annually for a year based on their intake water source (EPA, 2016). PWSs that rely on Ground Water (GW) were sampled twice, with a 5- or 7-month interval, while those that use Surface Water (SW), Mixed Water (MX), or Mixed but dominated by Ground Water (GU) were sampled four times, once in each consecutive quarter. Multiple water supply facilities may be present at each PWS, and water samples were collected from each facility as described above. The collected samples were analyzed at an EPA-approved laboratory. UCMR3 database includes detailed information on the PWSs' characteristics, including county, zip code, PWS ID/name, Water Supply Facility ID/name, PWS size, water source, sample point name, sample ID, sampling date, and analytical results for six types of PFAS.

In the lower USA, out of the six PFAS tested, at least one PFAS is detected in 33 states (68.7%), 121 counties (7.49%), 193 PWS (4.04%), and 345 PWS-Facilities (2.36%). Overall, 1107 water samples were found to be contaminated with PFAS which is 0.52% of total sample tested (Figure 1). The highest frequency of contamination was observed for PFOA (n = 377) followed by PFOS (n = 275), PFHpA (n = 228), PFHxS (n = 191), PFNA (n = 19), and PFBS (n = 17). PFOA, PFOS, PFHpA, PFHpX, PFNA and PFBS were detected in 27, 24, 22, 22, 4 and 7 states, respectively affecting around 16 million people in total (5.07% of the contiguous US population). Add The detailed summary statistics of different PFAS detected in the lower USA are presented in Appendix Table A1.

PFAS contamination seems to depend on water sources (Figure 1). Each Public Water Systems (PWS) can have multiple water supply facilities, each with separate water sources. In the lower USA, there are 14,706 water supply facilities and 4,782 PWS. Of these facilities, according to UCMR3 data, 22% rely on surface water (SW), 75.8% on groundwater (GW), 1.4% on mixed water (MX), and 0.8% on mixed water dominated by groundwater (GU). The UCMR3 data show that PFOA, PFOS, PFHxS, and PFNA contamination is higher if the water source includes groundwater (i.e., GW, MX, or GU) (Figure 1 and Appendix Table A2), which is consistent with Guelfo and Adamson (2018) and Hu et al (2016) findings.

PFAS contamination differs depending on the size of PWSs (Figure 2). The larger PWSs, which serve more than 10,000 people, are more likely to be contaminated than smaller PWSs. Specifically, 4.7% of large PWSs showed contamination with at least one PFAS, as opposed to the 0.8% of small PWSs (Appendix Table A 2). These findings corroborate the results obtained by Guelfo and Adamson (2018) and underscore the importance of PWS size in assessing the vulnerability to PFAS contamination.

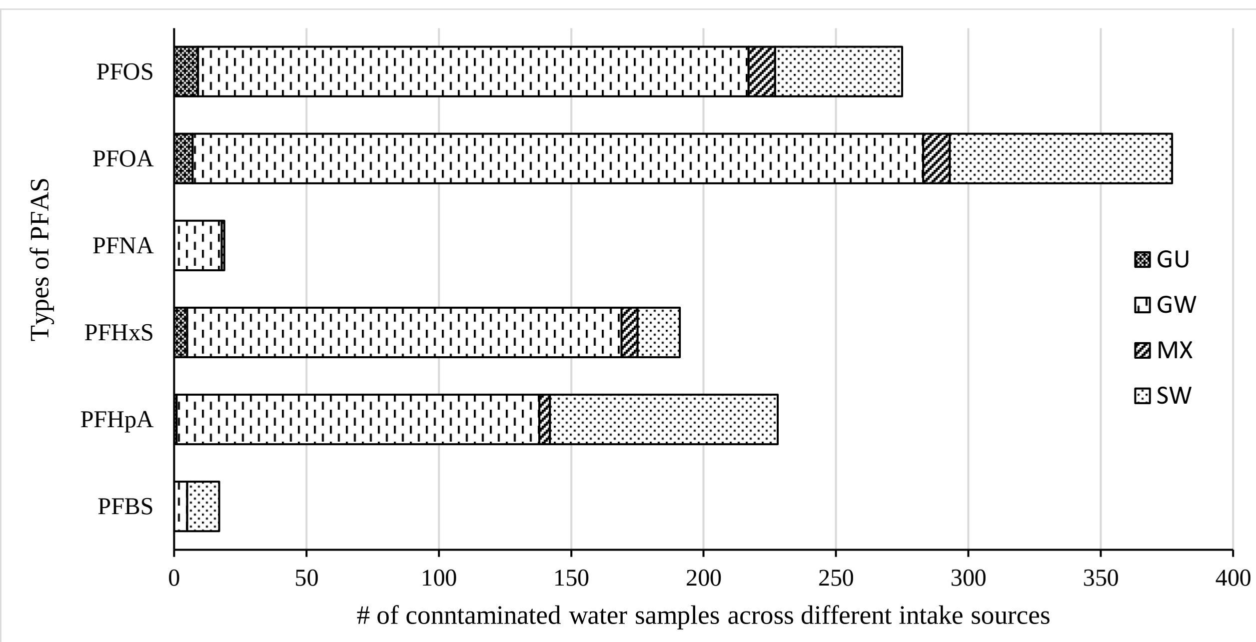


Figure 1: Contaminated samples by water source

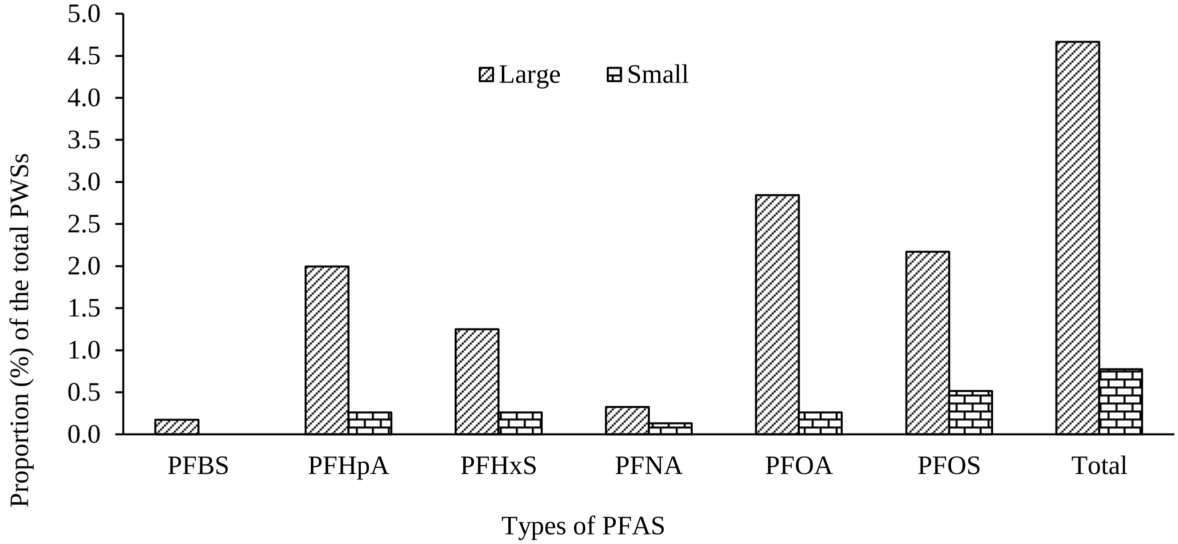


Figure 2: Proportion of the contaminated large and small PWSs

Table 1: Description and summary statistics of the variable used in the regression analysis.

Water System Level variables	Description	Mean	SD	Min	Max
Dependent Variable					
Concentration of Contaminants, measured in mug/L					
PFBS	Perfluorobutane sulfonic acid	0.0001	0.0044	0	0.3700
PFHpA	Perfluoroheptanoic acid	0.0001	0.0022	0	0.0869
PFHxS	Perfluorohexanesulfonic acid	0.0007	0.0137	0	0.7300
PFNA	Perfluorononanoic acid	0.0000	0.0009	0	0.0559
PFOA	Perfluorooctanoic acid	0.0004	0.0058	0	0.3490
PFOS	Perfluorooctanesulfonic acid	0.0010	0.0209	0	1.8000
PFAS	Sum of concentration of all above six PFAS tested	0.0024	0.0381	0	2.7000
Indicator dependent variables					
PFBS	Dummy variable=1, if water system has PFBS in at least one sample	0.0005	0.0219	0	1
PFHpA	Dummy variable=1, if water system has PFHpA in at least one sample	0.0064	0.0798	0	1
PFHxS	Dummy variable=1, if water system has PFHxS in at least one sample	0.0054	0.0731	0	1
PFNA	Dummy variable=1, if water system has PFNA in at least one sample	0.0005	0.0231	0	1
PFOA	Dummy variable=1, if water system has PFOA in at least one sample	0.0106	0.1024	0	1
PFOS	Dummy variable=1, if water system has PFOS in at least one sample	0.0077	0.0876	0	1
PFAS	Dummy variable=1, if water system has PFNA in at least one sample	0.0162	0.1264	0	1
Independent Indicator Variables					
PWS Characteristics					
PWS size: Small	Dummy variable=1, if water system has less than or equal to 10000 consumer	0.0894	0.2853	0	1
Water Source: Surface	Dummy variable=1, if water source to the water system is surface water	0.3549	0.4785	0	1
Water Source: Mixed	Dummy variable=1, if water source to the water system is mixed water	0.0226	0.1488	0	1
Water Source: Mixed but dominated by Ground	Dummy variable=1, if water source to the water system is mixed but is dominated by ground water	0.0121	0.1095	0	1
Population	Total people in the county in which PWS is located	970036.9	1786044.0	1966	10100000.0
Non-White Population percentage	%age of other than White people in the county	23.5673	13.8934	0.983753	83.6226
Total Poverty (%)	%age of the poor people out of the total people sampled	15.4980	5.8766	3.629071	43.9371
White Poverty	%age of Poor white people out of total white sampled	12.9859	5.2090	2.9086	39.3748
Non-white Poverty	%age of Poor non-white people out of total non-white sampled	23.5673	13.8934	0.9838	83.6226
Housing density	Number of housing unit per square mile	286.4846	392.1819	0.5	4832.06
Contribution (%) to the GDP from					
Agriculture		2.2385	5.1479	0	51.6102
Durable goods manufacturing		7.1686	6.4819	0	56.7367
Non-durable good manufacture		5.9150	6.8593	0	94.5604
Healthcare and social assistance		7.7165	3.3096	0	41.9700
Food and accommodation		2.8936	1.8789	0.000	33.5084
Government enterprise		14.0270	7.7341	1.115	75.3341

To provide a visual representation of PFAS spatial distribution, we show cumulative county scale contamination across six PFAS per PWS in Figure 3. We divided the aggregate number of positive samples in each county by the number of PWS in each county. Next, we used the Jenks natural breaks classification method (Jenks, 1967) to classify counties into groups based on contamination per PWS and county. Figure 3 shows that PFAS contaminations are more prevalent in Eastern than in western US counties. Some of the most contaminated counties per PWS are in Colorado, Alabama, Georgia, Delaware, New Jersey, and North Carolina. Appendix Figure A1 shows contamination per PWS in each county for PFOA, PFOS, PFHpA, and PFHxS individually.

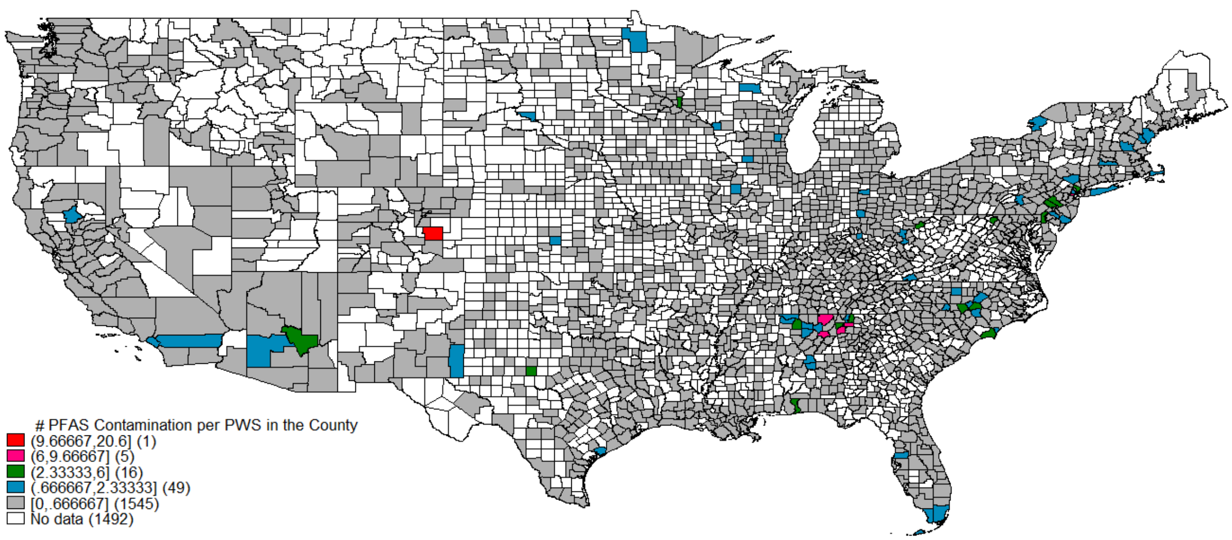


Figure 3. Number of PFAS contaminated water samples per PWS and county.

Methods

Hot Spot analysis

Spatial autocorrelation analysis is used to identify county-scale hot spots for PFOA, PFOS, PFHpA, PFHxS individually and all PFAS cumulatively (Kondo et al., 2016; Ord and Getis, 1995; Getis and Ord, 1992). This analysis is not provided for PFBS and PFNA because of limited number of positive samples. Getis-Ord ($G_i^*(d)$) z- statistic is used to identify the spatial clusters. The intuition of this method is that a county is in a statistically significant hot spot if it has higher contamination and is surrounded by other contaminated counties. Higher z-values indicate more intense clustering. Counties with z-value at or above 2.58 and 1.96 are in hot spots at 1%, and 5% level of significance, respectively. Following Getis and Ord (1992) and Ord and Getis (1995) the Getis-Ord $G_i^*(d)$ statistic (z-value) for PFAS contamination per PWS in county i is estimated as

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}}{s(i)\ \{[nS_{1i}^* \ - W_i^{*2}]/(n-1)\}^{\frac{1}{2}}} \quad \ldots\ldots\ldots(1)$$

where, x_j is number of contaminated samples per PWS in the county i, n is the number of counties, and $w_{ij}(d)$ is the symmetric one/zero spatial weight matrix with ones for all links defined as being within the distance of given i; all other links are zero. d is the threshold distance between county i and j. $w_{ij}(d)$ is 1 if bilateral distance between county i and county j d_{ij} is less than the threshold distance d and 0 otherwise. We use d=146.2 km threshold distance following Allaire et al. (2018), such that each county has at least one neighbor. Other term in the equation (1) is defined as;

$$\bar{x}(i) = \frac{\sum_j x_j}{(n-1)} \ldots\ldots\ldots..(2)\ and$$

Where, $\sum_j x_j$ is summation of all x_j within d of i.

$$W_i = \sum w_{ij}(d) \ldots\ldots\ldots\ldots(3)$$

$$s^2(i) = \frac{\sum_j x_j^2}{(n-1)} \ldots\ldots\ldots\ldots\ldots(4)$$

$$Var(G_i) = \frac{W_i(n-1-W_i)}{(n-1)^2(n-2)} \cdot \left[\frac{s(i)}{\bar{x}(i)}\right]^2 ..(5)$$

For the details of $G_i^*(d)$ derivation, please see Getis and Ord (1992) and Ord and Getis (1995).

Regression analysis

We examine the relationship between drinking water PFAS contamination and various physical, socioeconomic, and industrial characteristics of PWS and surrounding communities using Probit and Tobit models. The UCMR3 dataset limitation is the non-reporting of results below the Minimum Reporting Level (MRL) (Cadwallader et al., 2022). Test results below the MRL are reported as zero, which implies data censoring. None of the earlier studies considered this censoring in previous statistical analysis. Therefore, extending prior literature, we use a Tobit limited dependent variable model to deal with the censoring limitation of

UCMR3 PFAS data[4] (Sigelman & Zeng, 1999; Greene, 2018).

The left-censored Tobit Model (Carson & Sun, 2007; Cameron & Trivedi, 2005; Sigelman & Zeng, 1999) is formulated as follows:

$$y_{it}^{*} = \beta_0 + +\beta_l C_{lit} + \beta_m E_{mit} + X_i + S_{il} + T_t + \mu_{it}$$

where, y_{it}^{*} is the latent variable for concentration of PFAS in the county *i (i= 1..n)* and year *t*, ($\boldsymbol{X_i}$ is the vector of PWS characteristics of county i, $\boldsymbol{C_{lit}}$ is the nt*j matrix of socioeconomic characteristics j (j = population, nonwhite population, poverty, nonwhite poverty, and housing density) for county i in year t , and $\boldsymbol{\beta_l}$ is the vector of coefficient of socioeconomic characteristics. $\boldsymbol{E_{kit}}$ is the nt*k matrix of share of county GDP from k sectors (k = Agriculture, forestry and fisheries; durable goods manufacturing, on durable goods manufacturing; health care and social assistance; food and accommodation; government enterprise) and $\boldsymbol{\beta_m}$ is the vector of coefficient of all sectors that contribute to the county GDP. Indicator variables - state dummies (S_i), and year dummies (T_t), were included to account for state year invariant factors. Population and housing density data are log transformed.

Since our data is left censored,

$$y_{it} = \begin{Bmatrix} y_{it}^{*}\ if & y_{it}^{*} > MRL \\ L\ if & y_{it}^{*} \leq MRL \end{Bmatrix}$$

where, MRL is a nonstochastic constant and y_i^{*} is missing when it is less than MRL.

[4] Ordinary Least Squares (OLS) regression produces biased results in the presence of data censoring. Therefore, we do not present the OLS results but have those, along with Heckman two step model results, available upon request.

We also use Probit regression to examine contamination vulnerability (likelihood of contamination) as follows.

$$\Pr\,(y_{it} = 1|X) = \Phi(\beta_0 + +\beta_l C_{lit} + \beta_m\, E_{mit} + X_i + S_{il} + T_t)$$

where, $\Pr(y_{it} = 1|X)$ is the probability of observing a PFAS-positive sample for i^{th} PWS in year t and Φ is a cumulative density function of normal distribution.

Results

Hot Spot of Contamination

We identified four PFAS contamination hot spots (Figure 3). We also identified hot spots of PFOA, PFOS, PFHpA, and PFHxS contaminations individually (Appendix Figure A2). A list of states and counties in Hot spots along with the z-scores are reported in the Appendix Tables A3 -A7.

Four prominent PFAS hotspots encompassing 10 states and 149 counties have been identified (Figure 4). The largest hotspot spans across Alabama, Georgia, and Tennessee, encompassing 31 counties in Georgia, 18 counties in Alabama, and 17 counties in Tennessee. The second largest hotspot resides in the Northeast USA, with 20 counties in New Jersey, 14 counties in Pennsylvania, 10 counties in New York, and 2 counties each in Delaware and Connecticut. The third largest hotspot is situated on the border of North and South Carolina, comprising 25 counties in North Carolina and 2 counties in South Carolina. The smallest hotspot is in 10 counties in Colorado.

Hotspots exhibit either prevalence of PFAS manufacturing plants or PFAS industrial use sites, or densely populated sites or combination of them. For instance, the hotspot extending to the northeast states (NY, DE, NJ, PA, and CT) is predominantly situated in a densely populated region which indicates the contamination related to consumption of PFAS containing goods and its leakage to the water bodies. The Colorado hotspot includes highly populated counties which also have international and regional airports, along with defense and space infrastructure.

The hotspot in Georgia, Tennessee, and Alabama can be attributed to PFAS production and industrial use in this area and the subsequent leakage of PFAS into waterbody there, thereby contaminating the PWSs reliant on this water system (AP NEWS, 2019; de Amorim et al., 2019). Notably, de Amorim et al. (2019) highlight the carpet industry in Georgia as a major source of PFAS contamination in drinking water. The PFAS chemicals used during carpet production have been found to leach into water sources, resulting in the contamination of the drinking water supply. Furthermore, a striking example of PFAS contamination is evident in the Tennessee River, which subsequently led to the contamination of downstream drinking water in Alabama. A PFAS-producing plant in close proximity to the river was the source of this pollution and pay $98 million to compensate the damage it causes (AP NEWS, 2019). These cases emphasize the role of PFAS production and use activity in creating the contamination hotspot in this area. Similarly, the hotspot in the North and South Carolina can be linked to industrial production and the presence of a substantial consumer base in major cities.

Overall, these hotspots highlight the complex interplay between industrial activities, consumer sites, and population centers, underscoring the geographical concentration of PFAS contamination in specific regions.

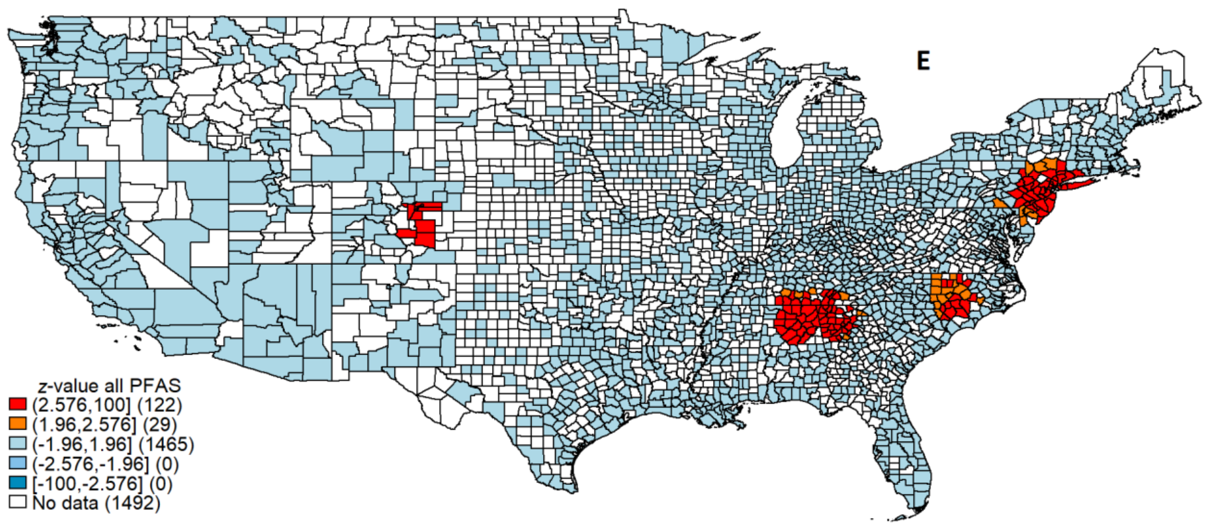


Figure 3. Hot spots of PFAS contamination. Legend intervals in the map represent significant PFAS contamination hot spots based on z-values obtained from standardized Getis Ord statistics. Counties with z-values ranging between 1.96 (-1.96) to 2.576 (-2.56) indicate hot spots (cold spots) at a 5% significance level, while z-values of 2.576 (-2.576) and higher (lower) signify hot spots with even greater significance at 1% level.

Regression results

The combined PFAS results from the Left Censored Tobit and Probit models are presented in table 2. The corresponding PFOA, PFOS, PFHpA, and PFHxS results are provided in the appendix Table A8 -A10. Tobit model shows correlations between cumulative PFAS at PWS scale and the corresponding independent variables, while the Probit model considers the likelihood of a PWS showing a positive PFAS sample. The results show that the

Tobit model provides a better model fit than Probit based on Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) values. However, we present both sets of results because they provide qualitatively similar but technically different interpretations.

PFAS contamination and PWS characteristics

PFAS contamination is less likely in small PWSs. On average, small PWSs have 0.145 ng/L (0.000145µg/L*1000 = 0.145 ng/L) less PFAS than large PWSs. This result is consistent with previous PFAS studies (Hu et al., 2016; Guelfo & Adamson,2018). The result is also consistent with Rahman et al. (2010) who found more SDWA regulated contaminant violations by large PWSs than by small ones in Arizona. However, many other studies on SWDA regulated contaminant violations found small PWSs to have more violations (Allaire et al., 2018; Michielssen et al., 2020; Acquah & Allaire, 2023; Anica and Elbakidze, 2023).

PWSs that rely on surface water sources are less likely to experience PFAS contamination compared to those that depend on ground water. PWSs that use surface water have 0.163 ng/L less PFAS concentration compared to the PFAS that uses GW. This finding is in line with PFAS contamination study by Guelfo and Adamson (2018) but contradicts the SDWA regulated contamination study (Allaire et al., 2018). Hence, surface water may pose greater risk for some of the SDWA regulated contaminants, while groundwater may pose greater risk in terms of PFAS contaminants.

The discrepancy in concentrations between SDWA-regulated contaminants and PFAS in groundwater sheds light on the contrasting behaviors exhibited during the percolation process. The diminished levels of SDWA-regulated contaminants in the ground water as compared to the surface water suggest that they undergo effective filtration through natural processes as they journey through the Earth's surface and reach the aquifer. This natural filtration mechanism acts as a safeguard, reducing the risk associated with these contaminants. However, in stark contrast, the absence of such filtration becomes apparent when considering PFAS. The substantial presence of PFAS in groundwater is a clear indication that these persistent compounds resist degradation through natural processes. Instead, they persist and accumulate wherever they settle, posing an ongoing concern for the quality and safety of groundwater resources.

Contamination and Socioeconomic Characteristics

PFAS contamination is more likely in PWSs located in populous and densely housed areas. For every 1% increase in population, PFAS concentration increases by 0.00057 ng/L (0.00005740μg/L*1000 = 0.0574 ng/L. Since independent variable population is in log, 0.0574/100 = 0.00057 ng/L per 1% increase in population). Similarly, PFAS increases by 0.00085 ng/L for every 1% increase in housing density. The positive effect of population and housing density may be associated to greater consumption of goods that contain PFAS. PFAS from goods such as detergents, cleaning agents, clothes, and others wind up in sewage system and eventually drain into water bodies that serve as sources for PWSs. Since PWSs and wastewater treatment systems lack the capacity to filter PFAS, PWSs in areas with greater population and housing density can experience elevated levels of PFAS.

PFAS contamination is a negatively correlated with non-white populations. Specifically, we find that for every 1% increase in non-white population and poverty, the concentration of PFAS decreases by 0.0063 ng/L (it is 0.2507 n/L in probit estimation). The overall poverty and poverty of disadvantaged community (non-white) has no impact on PFAS concentration. These findings are opposite from what is observed in the case of SDWA regulated contaminants (Allaire et al., 2018; McDonalds and Johns, 2018). In SDWA regulated contamination studies, greater numbers of violations are observed in disadvantaged communities and in areas with higher poverty. PWSs in disadvantaged communities have limited access to financial and other resources, resulting in poorer infrastructure and maintenance, which leads to water quality violations (Elbakidze and Beeson 2021). In the case of PFAS contamination, however, lack of financial resources does not seem to be associated with contamination. One reason for this result may be that no PFAS filtration technology exists to remove these compounds from drinking water. Therefore, affordability does not play a role. Instead, PFAS in drinking water seems to be correlated with some of the characteristics of the surrounding community. Wealthier communities have higher purchasing power and consume more goods, including those that contain PFAS, which can lead to greater PFAS leaching into water bodies that serve as sources for local PWSs. Conversely, disadvantaged communities have lower purchasing power, consume fewer goods, and generate less PFAS leaching into local water bodies.

Contamination and Industrial composition

PWSs situated in areas where agriculture, forestry, and fishery sectors are significant components of the local economy are less prone to contamination. A 1% increase in the share of agriculture, forestry, and fishery in the county GDP results in a 0.01 ng/L decline in PFAS concentration. This trend prevails notably in rural regions of the United States, where agriculture holds a prominent economic position (USDA, 2022). There is no evidence of PFAS use of in agricultural production. The occurrence of PFAS contamination through agriculture arises solely when municipal biosolids are employed as fertilizer, and such cases are infrequently reported within the country (MCDCP (Maine Department of Agriculture, Conservation and Forestry), 2023; Kim Lazcano et al., 2020 Choi et al., 2019). Moreover, the lower demand for consumer products in sparsely populated agriculturally dominant rural areas reinforces the validity of these findings. Consequently, the relative risk of PFAS contamination in PWSs within agriculturally driven communities remains low

The correlation between county GDP from non-durable goods manufacturing industry and PFAS contamination in local PWSs is positive and statistically significant. A1% increase in the share of GDP from non-durable goods manufacturing leads to a 0.01 ng/L increase in drinking water PFAS concentration. Non-durable goods, which have an average life of less than three years, include a wide range of products such as textiles, food packaging material, carpet, clothing, cosmetics, hygiene products and more. The manufacturing of these goods often involves the use of various PFAS compounds (EPA, 2023b), which may explain the positive correlation between elevated contamination and the prominence of non-durable goods manufacturing industry in the local economy.

The healthcare and social assistance industry exhibits a notable positive correlation with PFAS contamination. With each 1% increase in the healthcare and social assistance industry's share of the county GDP, there is a corresponding 0.01 ng/L increase in PFAS concentration. Numerous products utilized in hospitals and healthcare facilities, including MRI imaging, ultrasound, positron emission tomography (PET), cell abnormality tests, medicines, surgical gowns, drapes, flooring, and walls, contain PFAS (3M, 2019).

Ensuring a contaminant-free environment is crucial in hospitals and care facilities to prevent infections, necessitating specific cleaning protocols. PFAS are incorporated into the construction of these establishments and product manufacturing (e.g. surgical gown) due to their resistance to heat, water, and chemical degradation which facilitate rigorous cleaning procedures. Moreover, PFAS are commonly employed in medical implants and devices such as vascular grafts, surgical meshes, catheter tubes, filters, needle retrieval systems, tracheostomies, inhalers, catheter guide wires, and imaging products (Gaines, 2022). The consistent positive correlation between PFAS concentration and the proportion of the health sector in the local economy suggests that the healthcare industry serves as a significant user and emitter of PFAS into the environment.

We observe that PFAS contamination is more likely for PWSs located in counties with greater role of government enterprises in the local economy. For every 1% increase in the government enterprises' share of county GDP PFAS concentration increases by 0.0073 ng/L PFAS. Government enterprises such as military, firefighting, and government-operated airports have been found to use firefighting foams and heat-resistant equipment that contain PFAS (Hu et al., 2016). Additionally, federal and state-owned or partnered hospitals also use PFAS-containing products, which could contribute to the elevated levels of PFAS in the surrounding environment.

Table 2: PWS characteristics and socioeconomic factors affecting the overall PFAS contamination in the USA.

	Model			
	Tobit		Probit	
	Marginal effect (y^*)	Marginal effect for Censored Sample (y)	Likelihood	Marginal effect
PWS size (1=small, 0= large)	-0.203*** (0.0508)	-0.00014550*** (0.00002710)	-0.513*** (0.120)	-0.0070471*** (0.000989)
Water source (Surface water =1)	-0.131*** (0.0208)	-0.00016320*** (0.00003360)	-0.272*** (0.0468)	-0.0060083*** (0.0010733)
Water source (Mixed=1)	-0.0452 (0.0560)	-0.00005630 (0.00007000)	-0.0736 (0.127)	-0.0016275 (0.002817)
Water source (Mix but dominated by Ground water =1)	0.0792 (0.0614)	0.00009870 (0.00007750)	0.259* (0.139)	0.0057402** (0.0030902)
Population (log)	0.0461*** (0.0145)	0.00005740*** (0.00001930)	0.0776** (0.0331)	0.0017164** (0.0007303)
Non-White population (%)	-0.00503*** (0.00118)	-0.00000626*** (0.00000168)	-0.0113*** (0.00267)	-0.0002507*** (0.0000596)
Poverty (%)	-0.00515 (0.00434)	-0.00000642 (0.00000545)	-0.0118 (0.00993)	-0.0002619 (0.0002196)
Nonwhite Poverty (%)	0.00213 (0.00227)	0.00000265 (0.00000283)	0.00666 (0.00523)	0.0001473 (0.0001157)
Log Housing density (house/sq mil)	0.0679*** (0.0150)	0.00008460*** (0.00002150)	0.182*** (0.0339)	0.0040218*** (0.0007697)
Percentage Contribution to the GDP from				
Agriculture	-0.00882** (0.00405)	-0.00001100** (0.00000512)	-0.0224** (0.00934)	-0.0004947** (0.0002042)
Durable goods manufacturing	-0.00592*** (0.00227)	-0.00000737** (0.00000298)	-0.0150*** (0.00518)	-0.0003328*** (0.0001142)
Non-durable good manufacture	0.0105*** (0.00161)	0.00001300*** (0.00000249)	0.0237*** (0.00359)	0.0005242*** (0.0000789)
Healthcare and social assistance	0.00820*** (0.00308)	0.00001020** (0.00000403)	0.0158** (0.00717)	0.0003494 (0.0001585)
Food and accommodation	0.00628 (0.00433)	0.00000783 (0.00000546)	0.0125 (0.00999)	0.0002766 (0.0002204)
Government enterprise	0.00592*** (0.00134)	0.00000738*** (0.00000188)	0.0116*** (0.00312)	0.0002577 (0.0000694)
Constant	-1.569*** (0.181)		-3.312*** (0.400)	
Sigma	.4345775 .0151659			
Observations	35,589		30,777	
Log likelihood	-1995.8965		-2464.5555	
LR chi2(44)	997.68		811.01	

Standard errors in parentheses

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

Discussion

Understanding the socio-economic and industrial factors associated with PFAS contamination in PWS is critical for developing effective mitigation strategies. Hot spot analysis identifying regional clusters of contamination helps pinpoint vulnerable communities and prioritize mitigation efforts to contain and eliminate PFAS from the water system. In this study, intense hot spots of PFOA, PFOS, and overall PFAS have been identified in the southeastern US, including - Alabama, Georgia, and North Carolina; western US in Colorado; and northeast US, including New Jersey, New York, and Massachusetts.

The finding that large PWSs have more PFAS is consistent with the results of previous PFAS studies (Hu et al., 2016; Guelfo & Adamson, 2018). This result differs from non-PFAS-SWDA regulated contaminant literature. Allaire et al. (2018) found that the majority of violations were occurring in small PWS which is also supported by many other studies including Switzer and Teodoro (2018), Michielssen et al. (2020), and Smith et al. (2023).

We find that PWSs that rely on groundwater intake experience greater PFAS contamination. This result is consistent with all PWS-PFAS contamination studies. The inferiority of groundwater in terms of PFAS contamination is in stark difference relative to the SWDA-regulated contaminants. Allaire et al. (2018), for instance, found that violations of water quality regulations were significantly higher for PWSs that rely on surface water. The rationale is that groundwater goes through natural filtration processes before being pumped for treatment and delivery as drinking water. However, since PFAS do not biodegrade and are much more persistent (Cousins et al., 2020; Saez et al., 2008), the groundwater filtration processes may not be effective in removing these compounds. Furthermore, since PFAS can bioaccumulate, groundwater can actually contain more PFAS than free flowing surface water. Hence, particular focus on aquifers may be justified in terms of remediating PFAS in drinking water.

PFAS concentration is greater in PWSs that are located in more densely populated counties. This result is likely due to the greater consumption of PFAS-containing goods in densely populated regions (Glüge et al., 2020). PFAS can leach from everyday products like detergents, cleaning agents, and clothing, ultimately making their way into the wastewater treatment system and contaminating water sources (Stoiber et al., 2020). Since drinking water treatment facilities lack the capability to remove PFAS from drinking water, more PFAS contaminated intake implies more PFAS in drinking water. Hence, unless appropriate technology is developed and deployed to remove PFAS from drinking water, remedying drinking water PFAS will require addressing PFAS leakage from consumer goods.

In general, larger PWS in wealthier counties have a higher credit rating and better financial and technical capabilities (Anica and Elbakidze 2023; EPA, 2001). However, we do not observe a significant relationship between PFAS contamination and poverty. We also do not find a statistically significant effect of nonwhite poverty in contrast to Beeson and Elbakidze (2021) who find that SDWA violations are more common in communities with poor nonwhite residents and other studies that document disproportionate SDWA violations in disadvantaged communities with, which often lack access to policymaking processes or federal and state funding (Allaire, Wu, and Lall, 2018; McDonalds and Johns, 2018; EPA, 2001). These limitations put minority communities at greater risk of contamination as they have limited access to necessary technology or funding for proper water treatment. Our results show that limited access to resources and lack of representation in decision-making are not the primary reasons for PFAS contamination. Instead, our results actually show that non-white communities experience statistically less PFAS in drinking water than their white counterparts. These results support our previous interpretation that the PFAS in drinking water does not depend on community wealth or representation in government environmental initiatives because no technology is currently available to remove PFAS from drinking water. Instead, drinking water PFAS contamination seems to depend on population and housing density. Greater population implies greater consumption of consumer goods that contain PFAS, which can leak into local waterways. If PWSs source water from contaminated waterbodies and are unable to remove PFAS from water, then areas with greater populations will experience greater PFAS in drinking water.

PFAS in drinking water also depends on the surrounding area's industrial composition. The observed negative correlation between PFAS in drinking water and the agriculture, forestry, and fishery sector may be due to the limited or no use of PFAS-containing production inputs in these industries. Counties where these industries represent a significant share of local economy also tend to be rural with lower population densities (USDA, 2022) and lower demand for goods that contribute to PFAS in waterways. On the other hand, significantly higher concentration of PFAS is observed in counties with non-durable goods manufacturing industry. Non-durable goods, which have an average life of less than three years, include a wide range of products including textiles, shoes, food packaging material, carpet, clothing, cosmetics, hygiene, and more (BEA, 2023; EPA, 2023b). The nondurable goods manufacturing often involves the use of various PFAS compounds (EPA, 2023c), which may explain the observed positive correlation between PFAS contamination the non-durable goods manufacturing.

The healthcare and social assistance industry is also positively associated with PFAS contamination. Many hospital/health care products such as surgical gowns, drapes, flooring, and walls, contain PFAS. PFAS are also commonly used in medical implants and devices such as vascular grafts, surgical meshes, catheter tubes etc. to enhance longevity (Gaines, 2022). As a result, the health industry may be a significant source of PFAS in local waterbodies. We also observe a positive correlation between PFAS and government enterprises including military, firefighting, and government-operated airports. These government operations often use fire-retardants that contain significant quantities of PFAS. This finding is consistent with Hu et al. (2016) how detect higher PFAS levels in areas surrounding AFFF-certified airports and military bases.

Conclusion

We examine PFAS geographical distribution across the US PWSs, document several regional hotspots, and investigate potential communities and PWS attributes that may be correlated with elevated PFAS in drinking water. Major findings are that large PWSs that are located in densely populated areas and rely on groundwater as intake source have greater PFAS concentrations. Drinking water PFAS contamination is also correlated with non-durable goods manufacturing, healthcare and social assistance, and government enterprises. Conversely, we find lower PFAS contamination in communities with more non-white populations and in areas with larger agricultural industry.

Cumulatively, these results suggest a strong association between PFAS contamination and some industrial activities and consumption. Hence, drinking water PFAS contamination is a negative externality of production as well as consumption. Densely populated areas consume greater quantities of goods like textiles, detergents, personal care products, paints, and food packaging materials that contain PFAS. Inadequate post consumption handling, disposal, and recycling methods and facilities may be contributing to PFAS contamination in these regions.

Hence, our results suggest that future PFAS mitigation efforts ought to pay particular attention to not only industrial production activities but also to consumption as a source of PFAS. Until appropriate technologies are developed to remove PFAS from drinking water, policies and programs may be needed to address PFAS leakage from consumption.

References

3M. 2019. PFAS in the Medical Industry. Online: https://news.3m.com/PFAS-in-the-Medical-Industry.

Acquah, S., & Allaire, M. (2023). Disparities in drinking water quality: evidence from California. Water Policy, 25(2), 69-86.

Allaire, M., Wu, H., & Lall, U. (2018). National trends in drinking water quality violations. Proceedings of the National Academy of Sciences, 115(9), 2078-2083.

Andersson, E. M., Scott, K., Xu, Y., Li, Y., Olsson, D. S., Fletcher, T., & Jakobsson, K. (2019). High exposure to perfluorinated compounds in drinking water and thyroid disease. A cohort study from Ronneby, Sweden. Environmental research, 176, 108540.

Andrews, D. Q., & Naidenko, O. V. (2020). Population-wide exposure to per-and polyfluoroalkyl substances from drinking water in the United States. Environmental Science & Technology Letters, 7(12), 931-936.

Anica, S. T., & Elbakidze, L. (2023). Financial Assistance and Environmental Compliance: Evidence from the Clean Water Act and the Clean Water State Revolving Fund. Land Economics, 99(2), 222-244.

BEA (Bureau of Economic Analysis) (2023). Glossary: Nondurable goods. Available at: https://www.bea.gov/help/glossary/nondurable-goods

Cadwallader, A., Greene, A., Holsinger, H., Lan, A., Messner, M., Simic, M., & Albert, R. (2022). A Bayesian hierarchical model for estimating national PFAS drinking water occurrence. AWWA Water Science, 4(3), e1284.

Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge University Press.

Carson, R. T., & Sun, Y. (2007). The Tobit model with a non-zero threshold. The Econometrics Journal, 10(3), 488-502.

AP News (2021). 3M, Alabama communities reach $98M pollution settlement. Available online at: https://apnews.com/article/business-environment-and-nature-alabama-lawsuits-environment-a4dda1162d018987d691064b02873557

CDC (2022). Per- and Polyfluoroalkyl Substances (PFAS) and Your Health. Retrieved from https://www.atsdr.cdc.gov/pfas/.

Choi, Y. J., Kim Lazcano, R., Yousefi, P., Trim, H., & Lee, L. S. (2019). Perfluoroalkyl acid characterization in US municipal organic solid waste composts. Environmental Science & Technology Letters, 6(6), 372-377.

Cordner, A., De La Rosa, V. Y., Schaider, L. A., Rudel, R. A., Richter, L., & Brown, P. (2019). Guideline levels for PFOA and PFOS in drinking water: the role of scientific uncertainty, risk assessment decisions, and social factors. Journal of exposure science & environmental epidemiology, 29(2), 157-171.

Cousins, I. T., DeWitt, J. C., Glüge, J., Goldenman, G., Herzke, D., Lohmann, R., ... & Wang, Z. (2020). The high persistence of PFAS is sufficient for their management as a chemical class. Environmental Science: Processes & Impacts, 22(12), 2307-2312.

Creamer, J., Shrider, E. A., Burns, K., & Chen, F. (2022). Poverty in the United States: 2021. US Census Bureau. Retrieved from https://www.census.gov/library/publications/2022/demo/p60-277.html

de Amorim, B. O. F., Harbin, N., & White, P. S. (2019). Hazardous PFAS chemicals in Georgia's drinking water. J. Sci. Policy Gov, 14, 1-6.

De Silva, A. O., Armitage, J. M., Bruton, T. A., Dassuncao, C., Heiger-Bernays, W., Hu, X. C., ... & Sunderland, E. M. (2021). PFAS exposure pathways for humans and wildlife: a synthesis of current knowledge and key gaps in understanding. Environmental toxicology and chemistry, 40(3), 631-657.

DOE (2022). PFAS Strategic Roadmap: DOE Commitments to Action 2022-2025. Retrieved from https://www.energy.gov/sites/default/files/2022-08/DOE%20PFAS%20Roadmap%20August%202022.pdf?mc_cid=b0c93efb3c&mc_eid=cfa324a07b.

Elbakidze, L., & Beeson, Q. (2021). State regulatory heterogeneity and compliance with the clean water and safe drinking water acts. Water Resources Research, 57(5), e2020WR028952.EPA (2016). The Third Unregulated Contaminant Monitoring Rule (UCMR 3) Fact Sheet for Assessment Monitoring, online at: https://www.epa.gov/sites/default/files/2016-05/documents/ucmr3-factsheet-list1.pdf

EPA (2001). 1999 Drinking Water Infrastructure Needs Survey Modeling the Cost of Infrastructure (US Environmental Protection Agency, Washington, DC).

EPA (2012). The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Searching for Emerging Contaminants in Drinking Water. Retrieved from https://www.epa.gov/sites/default/files/2015-10/documents/ucmr3_factsheet_general.pdf.

EPA (2017). The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Data Summary, January 2017. Retrieved from https://www.epa.gov/sites/default/files/2017-02/documents/ucmr3-data-summary-january-2017.pdf.

EPA (2021). PFAS Strategic Roadmap: EPA's Commitments to Action 2021-2024, online at: https://www.epa.gov/pfas/pfas-strategic-roadmap-epas-commitments-action-2021-2024.

EPA (2022). Increasing Our Understanding of the Health Risks from PFAS and How to Address Them. Retrieved from https://www.epa.gov/pfas/increasing-our-understanding-health-risks-pfas-and-how-address-them.

EPA (2022a) Addition of Certain PFAS to the TRI by the National Defense Authorization Act, On line at https://www.epa.gov/toxics-release-inventory-tri-program/addition-certain-pfas-tri-national-defense-authorization-act.

EPA (2023). Occurrence Data from the Unregulated Contaminant Monitoring Rule. UCMR 3 (2013-2015) Occurrence Data. Available online:

https://www.epa.gov/dwucmr/occurrence-data-unregulated-contaminant-monitoring-rule#3.

EPA (2023a). Population Served by Community Water Systems with No Reported Violations of Health-Based Standards. Retrieved from https://cfpub.epa.gov/roe/indicator.cfm?i=45#:~:text=Community%20water%20systems%20(CWS)%20are,drinking%20water%20from%20a%20CWS.

EPA (2023c). Our Current Understanding of the Human Health and Environmental Risks of PFAS. https://www.epa.gov/pfas/our-current-understanding-human-health-and-environmental-risks-pfas

EPA (2023b). Nondurable Goods: Product-Specific Data. Available at: https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/nondurable-goods-product-specific-data

EPA 2016. The Third Unregulated Contaminant Monitoring Rule (UCMR3)- Fact Sheet for Assessment Monitoring (List 1 Contaminants). Available online at chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.epa.gov/sites/default/files/2016-05/documents/ucmr3-factsheet-list1.pdf.

Federal Register (2016). Lifetime health advisories and health effects support documents for perfluorooctanoic acid and perfluorooctane sulfonate. Federal Register, 81(101), 33250-33251. Retrieved from https://www.federalregister.gov/documents/2016/05/25/2016-12361/lifetime-health-advisories-and-health-effects-support-documents-for-perfluorooctanoic-acid-and.

Federal Register (2022). Lifetime Drinking Water Health Advisories for Four Perfluoroalkyl Substances. US Federal Register, 87(118), 36848-36849. Retrieved from https://www.federalregister.gov/documents/2022/06/21/2022-13158/lifetime-drinking-water-health-advisories-for-four-perfluoroalkyl-substances; February, 2023.

Gaines, L. G. (2022). Historical and current usage of per-and polyfluoroalkyl substances (PFAS): A literature review. American Journal of Industrial Medicine.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. Geographical analysis, 24(3), 189-206.

Glüge, J., Scheringer, M., Cousins, I. T., DeWitt, J. C., Goldenman, G., Herzke, D., ... & Wang, Z. (2020). An overview of the uses of per-and polyfluoroalkyl substances (PFAS). Environmental Science: Processes & Impacts, 22(12), 2345-2373.

Greene, W. H. (2018). Econometric analysis (Eight Edition).

Guelfo, J. L., & Adamson, D. T. (2018). Evaluation of a national data set for insights into sources, composition, and concentrations of per-and poly-fluoroalkyl substances (PFASs) in US drinking water. Environmental Pollution, 236, 505-513.

Hu, X. C., Andrews, D. Q., Lindstrom, A. B., Bruton, T. A., Schaider, L. A., Grandjean, P., ... & Sunderland, E. M. (2016). Detection of poly-and perfluoroalkyl substances (PFASs) in US drinking water linked to industrial sites, military fire training areas, and wastewater treatment plants. Environmental science & technology letters, 3(10), 344-350.

Jenks, G. F. (1967). The data model concept in statistical mapping. International yearbook of cartography, 7, 186-190.

Kato, K., Wong, L. Y., Jia, L. T., Kuklenyik, Z., & Calafat, A. M. (2011). Trends in exposure to polyfluoroalkyl chemicals in the US population: 1999− 2008. Environmental science & technology, 45(19), 8037-8045.

Kim Lazcano, R., Choi, Y. J., Mashtare, M. L., & Lee, L. S. (2020). Characterizing and comparing per-and polyfluoroalkyl substances in commercially available biosolid and organic non-biosolid-based products. Environmental Science & Technology, 54(14), 8640-8648.

Kondo, K. (2016). Hot and cold spot analysis using Stata. The Stata Journal, 16(3), 613-631.

MCDCP (Maine Department of Agriculture, Conservation and Forestry) 2023. Per- and Polyfluoroalkyl Substances (PFAS). Available online at: https://www.maine.gov/dacf/ag/pfas/index.shtml

McDonald, Y. J., & Jones, N. E. (2018). Drinking water violations and environmental justice in the United States, 2011–2015. American Journal of Public Health, 108(10), 1401-1407.

Michielssen, S., Vedrin, M. C., & Guikema, S. D. (2020). Trends in microbiological drinking water quality violations across the United States. *Environmental Science: Water* Research & Technology, 6(11), 3091-3105.

NIH. (2019). Perfluoroalkyl and Polyfluoroalkyl Substances (PFAS). Retrieved from https://www.niehs.nih.gov/health/materials/perfluoroalkyl_and_polyfluoroalkyl_substances_508.pdf; January, 2023.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. Geographical analysis, 27(4), 286-306.

Rahman, T., Kohli, M., Megdal, S., Aradhyula, S., & Moxley, J. (2010). Determinants of environmental noncompliance by public water systems. *Contemporary Economic Policy*, *28*(2), 264-274.

Sáez, M., de Voogt, P., & Parsons, J. R. (2008). Persistence of perfluoroalkylated substances in closed bottle tests with municipal sewage sludge. Environmental Science and Pollution Research, 15, 472-477.

Sigelman, L., & Zeng, L. (1999). Analyzing censored and sample-selected data with Tobit and Heckit models. Political analysis, 8(2), 167-182.

Smith, E., Krometis, L. A. H., Czuba, J. A., & Kolivras, K. (2023). Land cover and community water system characteristics as predictors of Safe Drinking Water Act violations in Central Appalachia, USA7. *Science of The Total Environment*, *858*, 159726.

Steenland, K., Fletcher, T., & Savitz, D. A. (2010). Epidemiologic evidence on the health effects of perfluorooctanoic acid (PFOA). Environmental health perspectives, 118(8), 1100-1108.

Stoiber, T., Evans, S., & Naidenko, O. V. (2020). Disposal of products and materials containing per-and polyfluoroalkyl substances (PFAS): A cyclical problem. Chemosphere, 260, 127659.

Switzer, D., & Teodoro, M. P. (2018). Class, race, ethnicity, and justice in safe drinking water compliance. *Social Science Quarterly*, *99*(2), 524-535.

Trang, B., Li, Y., Xue, X. S., Ateia, M., Houk, K. N., & Dichtel, W. R. (2022). Low-temperature mineralization of perfluorocarboxylic acids. Science, 377(6608), 839-845.

US Census Bureau (2023). Advanced Search. Available online at: https://data.census.gov/advanced

US Census Bureau (2023a). How the Census Bureau Measures Poverty. Available online at: https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html

USDA/ERA. (2022). Rural Poverty & Well-Being. Retrieved from https://www.ers.usda.gov/topics/rural-economy-population/rural-poverty-well-being.

APPENDIX

Table A1: PFAS in the US PWS

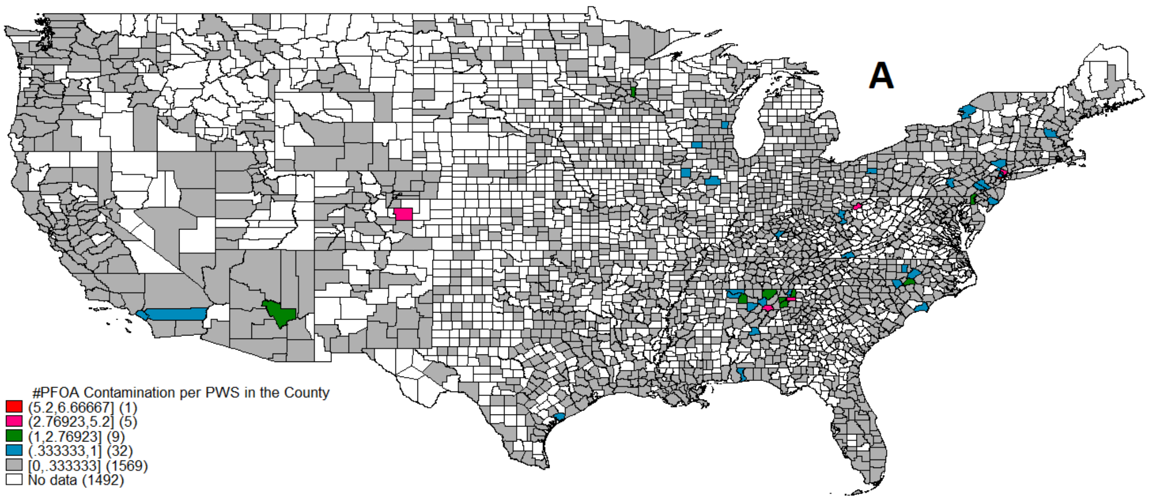
	Total Sampled (n)	Detected						
		PFBS	PFHpA	PFHxS	PFNA	PFOA	PFOS	At least one PFAS
States (n)	48	4(8.33)	22(45.83)	22(45.83)	7(14.58)	27(56.25)	24(50.0)	33(68.75)
Counties (n)	1616	7(0.43)	61(3.77)	43(2.66)	12(0.74)	78(4.83)	60(3.71)	121(7.49)
PWSs (n)	4782	7(0.15)	82(1.71)	52(1.09)	14(0.29)	116(2.43)	91(1.90)	193(4.04)
Facilities (n)	14,607	9(0.06)	136(0.93)	107(0.73)	14(0.10)	227(1.55)	165(1.13)	345(2.36)
Samples (n)	35,589	17(0.05)	228(0.64)	191(0.54)	19(0.05)	377(1.06)	275(0.77)	578(1.62)
Population, 10^6 (n)	314.38	0.28 (0.09)	8.76 (2.79)	5.50 (1.75)	0.52 (0.17)	7.73 (2.46)	10.28 (3.27)	15.95 (5.07)

Note: figures in the parenthesis indicate the percentages of the total(n)

Table A2: PFAS contamination by large and small PWSs and by intake water source.

Attributes	Total Count (n)	Contamination (n(%))						
		PFBS	PFHpA	PFHxS	PFNA	PFOA	PFOS	At least one PFAS
PWSs	4782	7 (0.15)	82(1.71)	52(1.09)	14(0.29)	116(2.43)	91(1.90)	193(4.40)
Large	4008 (83.81)	7(0.17)	80(2.00))	50(1.25)	13(0.32)	114(2.84)	87(2.17)	187(4.67)
Small	774 (16.19)	0(0.0)	2(0.26)	2(0.26)	1(0.13)	2(0.26)	4(0.52)	6(0.78)
Source of water to the Water supply facility								
Facility	14607	9(0.06)	136 ((0.93)	107(0.73)	14(0.10)	227(1.55)	165 (1.13)	345(2.36)
SW	3213(22.0)	5(0.16)	44(1.37)	7(0.22)	1(0.03)	46(1.43)	27(0.84)	74(2.30)
GW	11074(75.8)	4(0.04)	88(0.79)	95(0.86)	13(0.12)	173(1.56)	130(1.17)	258(2.33)
MX	210(1.4)	0(0.00)	3(1.44)	2(0.95)	0(0.00)	5(2.38)	5(2.38)	6(2.86)
GU	110(0.8)	0(0.00)	1(0.91)	3(2.73)	0(0.00)	3(2.73)	4(3.64)	7(6.36)

Note: figures in the parenthesis indicate the percentages of the total(n)



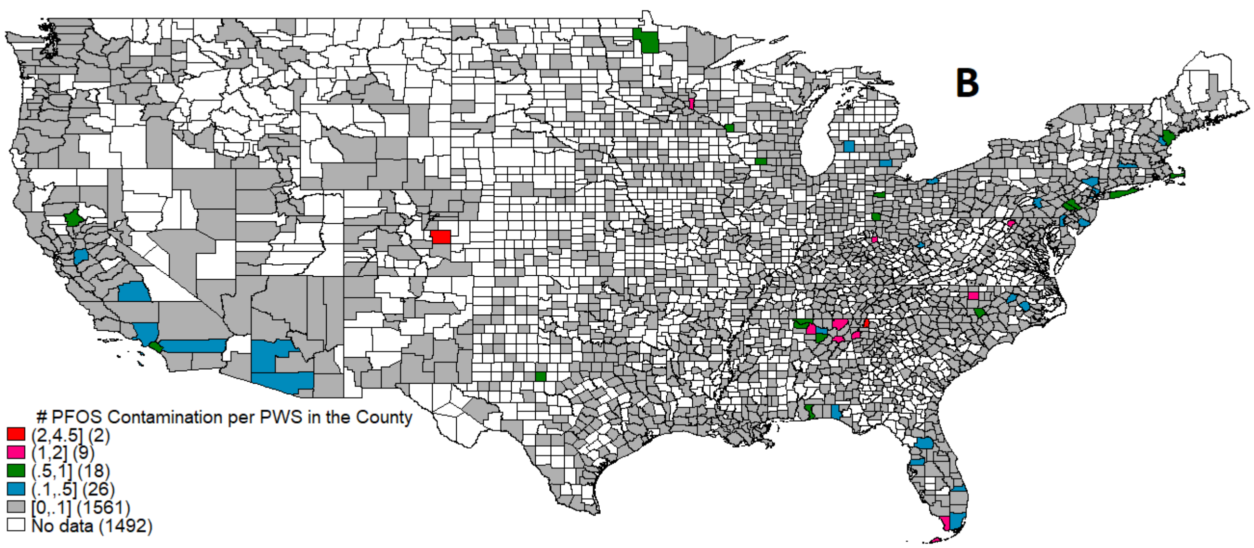


Figure A.1: Number of and PFOA (A) and PFOS (B) contaminated water samples per PWS in the county.

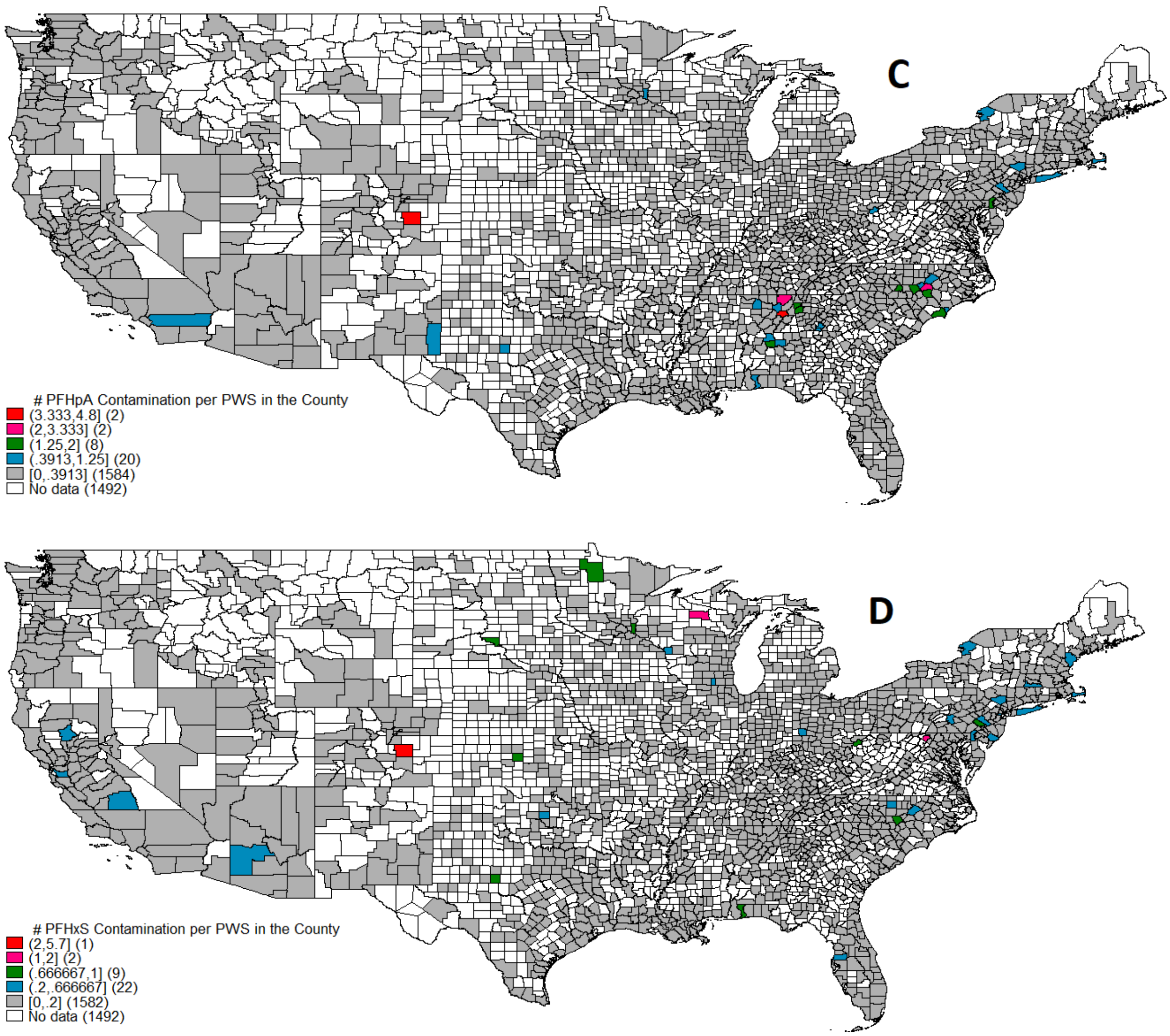


Figure A.1.1: Number of and PFHpA (C), and PFHxS (D) contaminated water samples per PWS in the county.

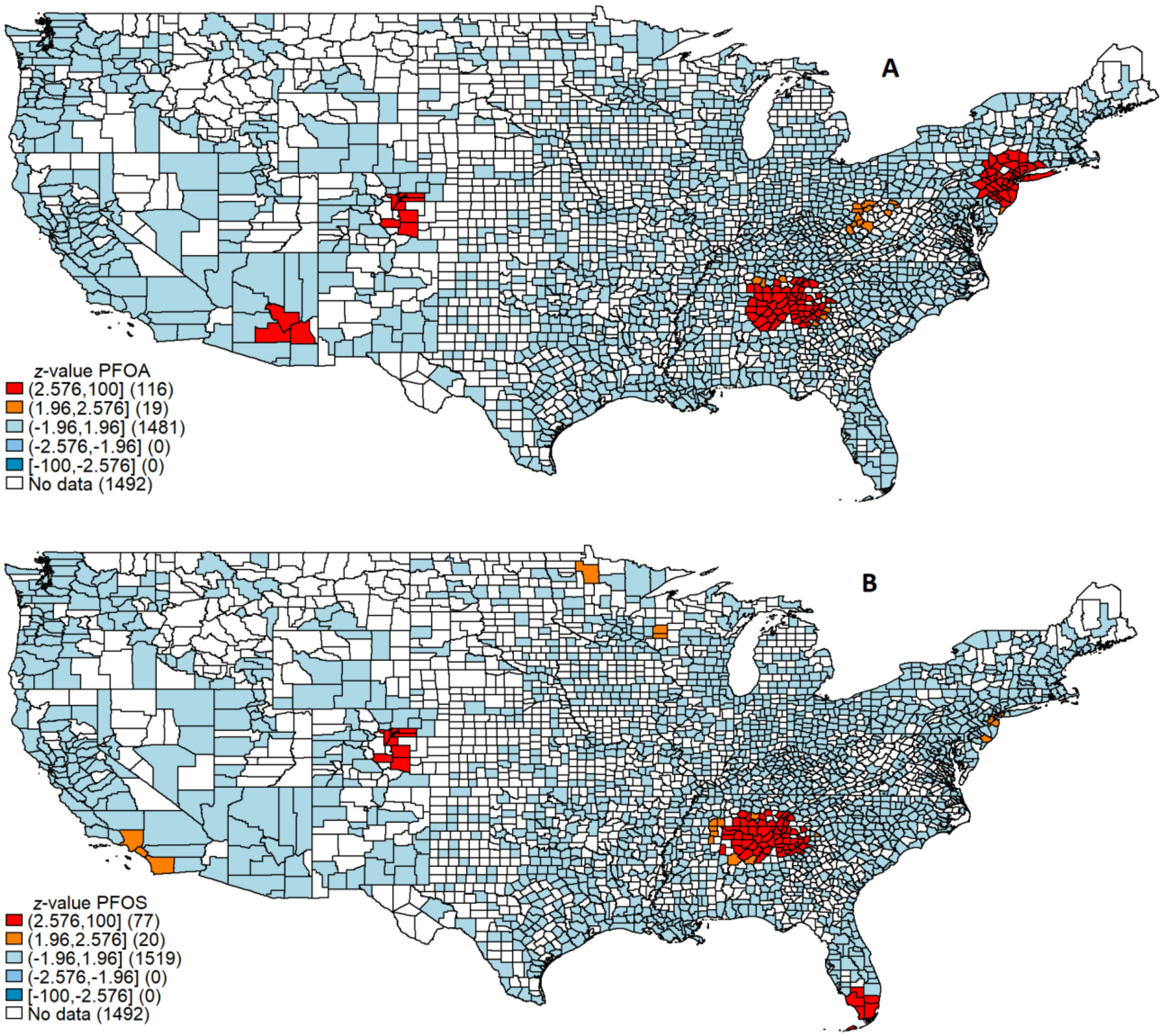
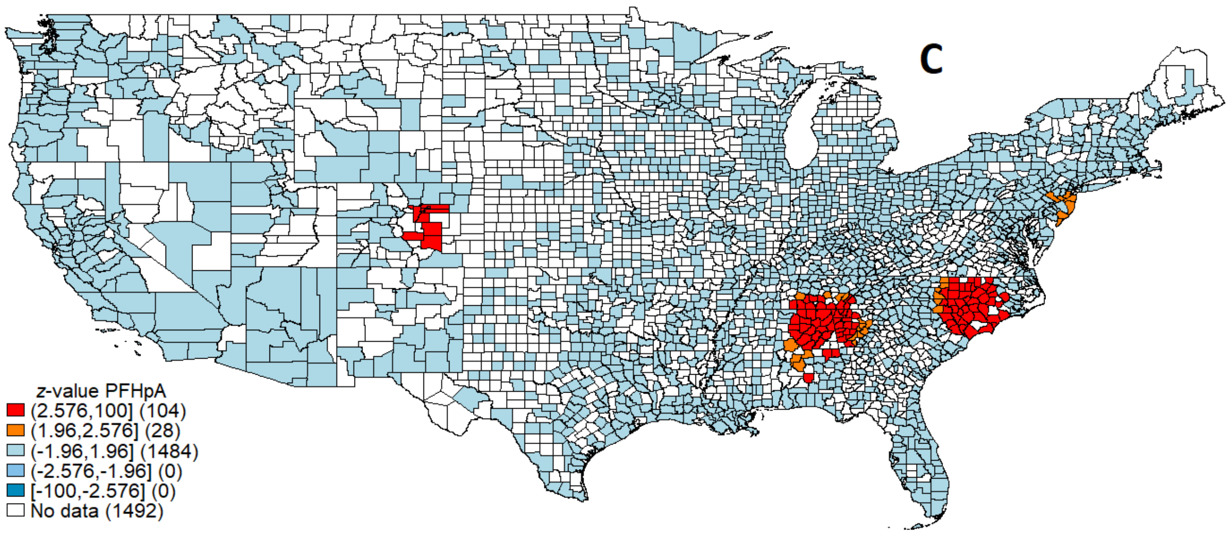


Figure A2. Hot spot of PFOA(A) and PFOS (B) contamination. Hot spot was based on number of contaminated samples per PWS in the county. Intervals in the legends are selected based on 1% and 5% levels of significance.



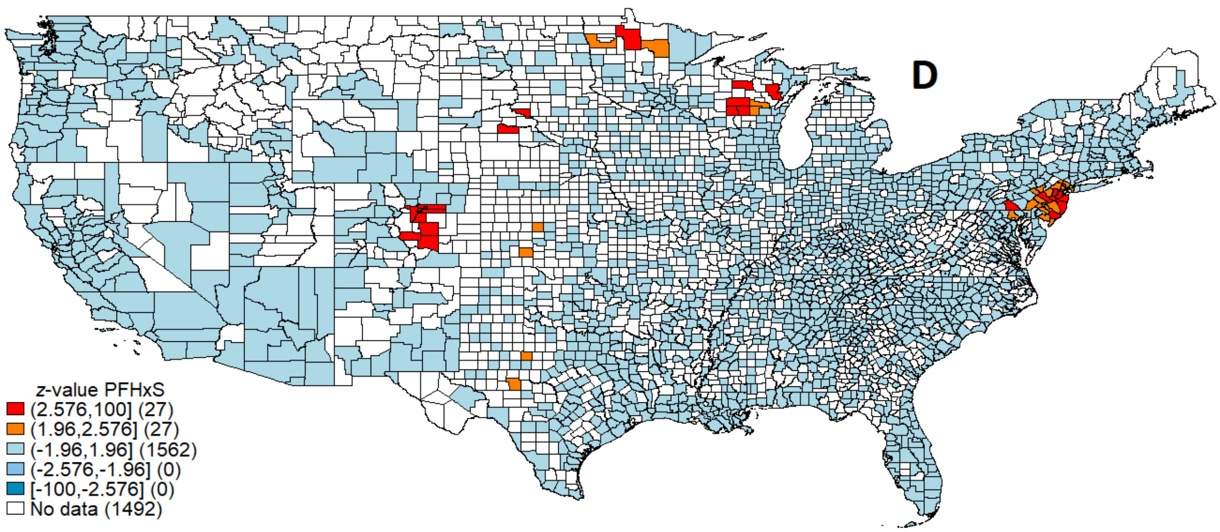


Figure A2. Hot spot of PFHpA (C), and PFHxS (D) contamination. Hot spot was based on number of contaminated samples per PWS in the county. Intervals in the legends are selected based on 1% and 5% levels of significance.

Table A3: Hot spots of PFOA Contamination

Hot Spots	state	County
1	Alabama [18]	Etowah (9.53), DeKalb (9.02), Jackson (8.90), Marshall (8.81), Cherokee (8.02), Calhoun (7.89), Blount (7.03), Madison (7.00), St. Clair (6.89), Cullman (6.76), Morgan (5.78), Winston (5.27), Lawrence (5.06), Limestone (4.94), Walker (4.49), Talladega (4.36), Jefferson (3.83), Shelby (2.86)
	Georgia [33]	Chattooga (7.62), Floyd (7.38), Catoosa (7.38), Haralson (7.38), Walker (7.30), Whitfield (7.28), Dade (7.21), Gordon (7.19), Polk (7.19), Bartow (7.19), Paulding (6.93), Murray (6.85), Carroll (6.21), Cobb (5.80), Douglas (5.65), Cherokee (5.65), Gilmer (5.03), Coweta (4.66), Gwinnett (4.36), DeKalb (4.29), Hall (4.23), Forsyth (4.17), Fulton (4.17), Union (4.00), Dawson (4.00), Rockdale (3.15), Barrow (3.09), Fayette (3.04), Habersham (2.86), Stephens (2.61), Jackson (2.35), Henry (2.30), Walton (2.30)
	Tennessee [16]	Franklin (9.36), Marion (7.75), Hamilton (5.80), Meigs (5.70), McMinn (5.61), Lincoln (5.46), Bradley (5.46), Bledsoe (4.72), Rhea (4.65), Coffee (2.87), Lawrence (2.84), Giles (2.72), Monroe (2.65), Warren (2.40), Marshall (2.36), Maury (2.17)
2	New Jersey (18]	Ocean (7.42), Burlington (6.88), Mercer (6.72), Monmouth (6.72), Middlesex (6.72), Atlantic (6.43), Somerset (6.43), Hunterdon (6.08), Union (5.72), Hudson (5.71), Passaic (5.66), Bergen (5.66), Essex (5.62), Camden (5.39), Morris (5.37), Gloucester (5.21), Warren (4.92), Cape May (2.25), Marion (1.97)
	Pennsylvania [13]	Philadelphia (6.50), Bucks (6.29), Montgomery (6.01), Lehigh (5.90), Northampton (5.89), Lackawanna (5.44), Pike (5.28), Carbon (5.19), Monroe (5.12), Delaware (5.03), Wayne (4.67), Chester (4.39), Berks (4.16)
	New York [10]	Suffolk (6.26), Nassau (6.02), Queens (6.00), Rockland (5.48), Orange (5.10), Westchester (5.09), Putnam (5.00), Dutchess (4.95), Ulster (4.86), Sullivan (4.72)
	Connecticut [2]	New Haven (3.00), Fairfield (2.97)
3	West Virginia [9]	Kanawha (2.47), Mason (2.33), Putnam (2.33), Cabell (2.16), Wood (2.10), Wayne (2.10), Lewis (2.05), Taylor (2.05)
	Ohio [3]	Meigs (2.16), Washington (2.10), Athens (2.00)
4	Colorado [10]	Pueblo (8.76), Fremont (6.79), El Paso (5.99), Arapahoe (5.39), Adams (4.93), Douglas (4.93), Denver (4.93), Broomfield (4.93), Jefferson (4.72), Gilpin (4.72),
5	Arizona [3]	Gila (3.73), Pinal (3.13), Graham (2.77)

Value in [ ] indicates number of counties that fall in the hot spot in respective states.

Value in ( ) indicates the z-score of Standardized Getis Ord statistics.

Table A4: Hot spots of PFOS Contamination

Hotspot	state	county
1	Alabama [23]	Etowah (8.50), DeKalb (8.03), Jackson (7.92), Blount (7.13), Madison (6.79), Cherokee (6.56), St. Clair (6.46), Marshall (6.38), Cullman (6.29), Winston (6.16), Lawrence (5.92), Limestone (5.70), Morgan (5.60), Walker (5.26), Calhoun (4.18), Jefferson (4.00), Lamar (3.70), Talladega (3.49), Colbert (3.34), Franklin (3.18), Lauderdale (3.11), Tuscaloosa (2.57), Shelby (2.42),
	Georgia [27]	Chattooga (6.22), Walker (5.56), Dade (5.49), Floyd (5.39), Catoosa (5.39), Haralson (5.39), Whitfield (5.31), Gordon (5.24), Polk (5.24), Bartow (5.24), Murray (5.17), Paulding (5.03), Carroll (4.32), Cobb (4.20), Douglas (4.08), Cherokee (4.08), Gilmer (3.77), Rockdale (3.22), Barrow (3.16), Gwinnett (3.11), DeKalb (3.06), Hall (3.00), Forsyth (2.95), Fulton (2.95), Union (2.81), Dawson (2.81), Stephens (2.06)
	Mississippi [3]	Lee (2.34), Alcorn (2.20), Prentiss (2.08)
	Tennessee [18]	Franklin (8.60), Marion (5.88), Lincoln (5.22), Meigs (4.32), McMinn (4.25), Hamilton (4.18), Bradley (4.12), Lawrence (4.06), Giles (3.91), Maury (3.29), Coffee (3.16), Marshall (3.16), Bledsoe (2.87), Rhea (2.82), Warren (2.72), Bedford (2.44), Hardin (2.14), McNairy (2.08)
2	Colorado [10]	Pueblo (8.76), Fremont (6.79), El Paso (5.99), Arapahoe (5.39), Adams (4.93), Douglas (4.93), Denver (4.93), Broomfield (4.93), Jefferson (4.72), Gilpin (4.72),
3	Florida [5]	Monroe (4.79), Miami-Dade (4.32), Broward (3.43), Collier (3.43), Hendry (2.76),
4	New Jersey [4]	Middlesex (2.45), Monmouth (2.45), Atlantic (2.05), Union (2.00)
	New York [1]	Queens (2.02)
5	California [3]	Orange (2.46), San Diego (2.12), Los Angeles (2.06),
6	Wisconsin [2]	Chippewa (2.28), Eau Claire (2.09)
7	Minnesota [1]	Beltrami (2.36),

Value in [ ] indicates number of counties that fall in the hot spot in respective states.

Value in ( ) indicates the z-score of Standardized Getis Ord statistics.

Table A5: Hot spots of PFHpA Contamination

Hot Spot	state	County
1	Alabama [25]	St. Clair (11.09), Blount (8.99), Cullman (8.41), Madison (8.28), Calhoun (8.09), Etowah (7.98), Marshall (7.92), Talladega (7.79), DeKalb (7.59), Jackson (7.49), Morgan (7.18), Cherokee (6.96), Winston (6.30), Jefferson (6.30), Lawrence (5.93), Shelby (5.75), Limestone (5.72), Walker (4.66), Chambers (4.41), Tallapoosa (4.01), Butler (2.81), Tuscaloosa (2.55), Perry (2.29), Dallas (2.29), Chilton (2.05)
	Georgia [25]	Floyd (6.98), Haralson (6.83), Polk (6.81), Bartow (6.81), Paulding (6.56), Chattooga (6.50), Gordon (6.50), Catoosa (6.36), Walker (6.35), Whitfield (6.27), Dade (6.26), Murray (5.51), Carroll (4.88), Coweta (4.26), Cherokee (4.17), Cobb (4.14), Douglas (4.03), Gilmer (3.68), Forsyth (2.31), Gwinnett (2.28), Fayette (2.28), DeKalb (2.24), Hall (2.19), Dawson (2.19), Fulton (2.15)
	Tennessee [16]	Franklin (7.70), Lincoln (6.78), Marion (6.69), Hamilton (4.40), Meigs (3.88), McMinn (3.82), Bradley (3.70), Coffee (3.28), Lawrence (3.21), Giles (3.08), Marshall (2.90) Bedford (2.85), Bledsoe (2.53), Rhea (2.48), Warren (2.39), Maury (2.30)
2	North Carolina [42]	Robeson (8.95), Cumberland (7.98), Bladen (7.88), Scotland (7.53), Sampson (7.44), Hoke (7.41), Wayne (7.18), Lenoir (6.84), Alamance (6.23), Pender (6.20), Orange (6.12), Harnett (6.12), Chatham (5.83), Lee (5.74), Moore (5.66), Randolph (5.57), Guilford (5.57), Richmond (5.49), Montgomery (5.41), Franklin (5.28), Granville (5.28), Anson (5.11), Wake (4.98), Johnston (4.89), Durham (4.89), Stanly (4.81), Rockingham (4.80), New Hanover (4.79), Onslow (4.62), Brunswick (4.46), Nash (4.32), Davidson (4.27), Wilson (4.05), Person (3.52), Pitt (3.45), Warren (2.67), Stokes (2.27), Forsyth (2.22), Davie (2.22), Cabarrus (2.21), Union (2.16), Rowan (2.09),
	South Carolina [7]	Dillon (8.81), Marlboro (7.18), Chesterfield (5.27), Marion (5.02), Horry (4.18), Darlington (4.17), Florence (3.84)
3	Colorado [10]	Pueblo (8.80), Fremont (6.83), El Paso (6.03), Arapahoe (5.43), Adams (4.96), Douglas (4.96), Denver (4.96), Broomfield (4.96), Jefferson (4.76), Gilpin (4.76)
4	New Jersey [6]	Atlantic (2.39), Monmouth (2.37), Middlesex (2.37), Ocean (2.11), Mercer (2.07), Cape May (2.02),
	Pennsylvania [1]	Bucks (1.99)

Value in [ ] indicates number of counties that fall in the hot spot in respective states.

Value in ( ) indicates the z-score of Standardized Getis Ord statistics.

Table A6: Hot spots of PFHxS Contamination

Hot Spot	state	county
1	New Jersey [18]	Atlantic (3.26), Monmouth (3.22), Middlesex (3.22), Mercer (2.72), Ocean (2.69), Union (2.68), Somerset (2.58), Cape May (2.50), Camden (2.47), Burlington (2.42), Hunterdon (2.41), Gloucester (2.37), Passaic (2.35), Bergen (2.35), Hudson (2.30), Essex (2.25), Morris (2.13), Cumberland (2.03)
	Pennsylvania [8]	Bucks (2.69), York (2.61), Delaware (2.48), Chester (2.32), Philadelphia (2.28), Montgomery (2.28), Northampton (2.17), Lehigh (2.17)
	Maryland [1]	Baltimore (2.01)
	New York [1]	Queens (2.49)
2	Colorado [10]	Pueblo (13.96), Fremont (10.94), El Paso (9.73), Arapahoe (8.83), Adams (8.12), Douglas (8.12), Denver (8.12), Broomfield (8.12), Jefferson (7.82), Gilpin (7.82)
3	Wisconsin [7]	Oneida (3.94), Wood (3.21), Portage (2.97), Marinette (2.87), Marathon (2.87), Shawano (2.49), Waupaca (2.04)
4	Minnesota [3]	Beltrami (3.04), Itasca (2.26), Polk (2.26)
5	South Dakota [2]	Mellette (3.04), Hughes (3.04)
6	Kansas [2]	Ellis (2.02), Ford (2.02)
7	Texas [2]	Tom Green (2.02), Jones (2.02)

Value in [ ] indicates number of counties that fall in the hot spot in respective states.

Value in ( ) indicates the z-score of Standardized Getis Ord statistics.

Table A7: Hot spots of PFAS Contamination

Hot Spot	state	county
1	Georgia [31]	Chattooga (7.77), Floyd (7.63), Haralson (7.58), Catoosa (7.45), Polk (7.43), Walker (7.43), Bartow (7.43), Whitfield (7.35), Gordon (7.34), Dade (7.33), Paulding (7.15), Murray (6.89), Carroll (6.09), Cobb (5.65), Cherokee (5.55), Douglas (5.50), Gilmer (5.18), Coweta (4.42), Gwinnett (4.26), DeKalb (4.20), Hall (4.13), Forsyth (4.11), Fulton (4.07), Dawson (3.93), Union (3.76), Rockdale (3.31), Barrow (3.26), Fayette (2.84), Henry (2.27), Habersham (2.27), Stephens (2.10)
	Alabama [18]	Etowah (9.74), DeKalb (9.22), Jackson (9.09), Marshall (8.83), Cherokee (8.23), St. Clair (8.15), Calhoun (7.96), Blount (7.75), Madison (7.61), Cullman (6.89), Morgan (5.93), Winston (5.66), Lawrence (5.39), Limestone (5.21), Talladega (5.18), Walker (4.44), Jefferson (4.29), Shelby (3.34)
	Tennessee [17]	Franklin (9.55), Marion (7.86), Hamilton (5.92), Meigs (5.81), McMinn (5.72), Lincoln (5.56), Bradley (5.55), Bledsoe (4.27), Rhea (4.20), Lawrence (3.15), Coffee (3.07), Giles (3.01), Marshall (2.62), Warren (2.51), Maury (2.42), Bedford (2.19), Monroe (2.04)
2	New Jersey [20]	Middlesex (5.00), Monmouth (5.00), Ocean (5.00), Atlantic (4.85), Mercer (4.67), Burlington (4.56), Somerset (4.43), Hunterdon (4.15), Union (3.96), Camden (3.88), Hudson (3.86), Essex (3.78), Gloucester (3.73), Bergen (3.64), Passaic (3.64), Morris (3.58), Warren (2.92), Cape May (2.80), Cumberland (2.32), Salem (1.98)
	Pennsylvania [14]	Bucks (4.48), Philadelphia (4.28), Montgomery (4.09), Northampton (4.09), Lehigh (4.09), Delaware (3.77), Carbon (3.36), Chester (3.33), Lackawanna (3.21), Monroe (3.16), Berks (3.11), Pike (3.05), York (2.07), Wayne (2.02)
	New York [10]	Queens (3.92), Nassau (3.52), Suffolk (3.45), Rockland (3.16), Orange (2.92), Westchester (2.69), Putnam (2.63), Dutchess (2.62), Ulster (2.20), Sullivan (2.04)
	Connecticut [2]	New Haven (3.00), Fairfield (2.97)
	Delaware [2]	New Castle (2.10), Kent (2.04)
3	North Carolina [25]	Cumberland (3.19), Robeson (3.00), Scotland (2.96), Hoke (2.90), Sampson (2.86), Bladen (2.79), Wayne (2.72), Harnett (2.67), Alamance (2.66), Franklin (2.65), Granville (2.65), Orange (2.60), Wake (2.45), Chatham (2.43), Johnston (2.39), Durham (2.39), Lee (2.38), Moore (2.23), Guilford (2.18), Randolph (2.18), Rockingham (2.16) Lenoir (2.15), Richmond (2.13), Montgomery (2.09), Person (1.97),
	South Carolina [2]	Dillon (2.90), Marlboro (2.29)
4	Colorado [10]	Pueblo (10.61), Fremont (8.23), El Paso (7.27), Arapahoe (6.55), Adams (5.98), Douglas (5.98), Denver (5.98), Broomfield (5.98), Jefferson (5.74), Gilpin (5.74)

Value in [ ] indicates number of counties that fall in the hot spot in respective states.

Value in ( ) indicates the z-score of Standardized Getis Ord statistics.

TableA8: PWS characteristics and socioeconomic factor affecting the PFOA contamination in the USA.

VARIABLES - PFOA	Models			
	Tobit		Probit	
	Marginal effect (y^*)	Marginal effect for Censored Sample (y)	Likelihood	Marginal effect
Size of the PWS – Small (1=small, 0= large)	-0.0668*** (0.0173)	-0.0000093*** (0.0000024)	-0.725*** (0.181)	-0.0045854*** (0.0006519)
Source of water to the PWS- Surface (SW)	-0.0301*** (0.00580)	-0.0000092*** (0.0000028)	-0.293*** (0.0567)	-0.0037098*** (0.0007919)
Source of water to the PWS- Mixed (MX)	0.00437 (0.0144)	0.0000013 (0.0000044)	0.0592 (0.143)	0.0007486 (0.001814)
Source of water to the PWS- Mix but dominated by Ground water (GU)	-0.00704 (0.0186)	-0.0000022 (0.0000057)	-0.0306 (0.184)	-0.0003867 (0.0023285)
Population log	0.00522 (0.00404)	0.0000016 (0.0000013)	0.0537 (0.0404)	0.0006792 (0.0005107)
Non-White population (%)	-0.00120*** (0.000328)	-0.0000004*** (0.0000001)	-0.0120*** (0.00323)	-0.0001513*** (0.0000425)
Poverty (%)	0.00115 (0.00123)	0.0000004 (0.0000004)	0.00787 (0.0124)	0.0000996 (0.0001565)
Nonwhite Poverty (%)	-0.00102 (0.000647)	-0.0000003 (0.0000002)	-0.00577 (0.00656)	-0.000073 (0.0000832)
Log Housing density (house/sq mil)	0.0194*** (0.00421)	0.0000059*** (0.0000019)	0.204*** (0.0410)	0.0025839*** (0.0005679)
Percentage Contribution to the GDP from				
Agriculture	-0.00244* (0.00126)	-0.0000007** (0.0000004)	-0.0223* (0.0123)	-0.0002824** (0.0001537)
Durable goods manufacturing	-0.000480 (0.000621)	-0.0000001 (0.0000002)	-0.00395 (0.00619)	-0.00005 (0.0000784)
Non-durable good manufacture	0.00276*** (0.000459)	0.0000008*** (0.0000002)	0.0275*** (0.00448)	0.0003479*** (0.0000597)
Healthcare and social assistance	0.00184** (0.000853)	0.0000006* (0.0000003)	0.0184** (0.00872)	0.0002333** (0.0001113)
Food and accommodation	0.000494 (0.00120)	0.0000002 (0.0000004)	0.00396 (0.0120)	0.0000501 (0.0001514)
Government enterprise	0.00145*** (0.000388)	0.0000004*** (0.0000002)	0.0146*** (0.00385)	0.0001847*** (0.0000508)
Constant	-0.336*** (0.0514)		-3.472*** (0.496)	
Sigma	.1009417 (.0046164)			
Observations	35,589		28,908	
Log likelihood	-825.49792		-1657.6946	
LR chi2(44)	866.11		705.75	
Prob > chi2	0		0	
Pseudo R2	0.3441		0.1755	
AIC	1740.996		3403.389	
BIC	2122.586		3767.352	

Standard errors in parentheses

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

Table A9: PWS characteristics and socioeconomic factors affecting the PFOS contamination in the USA.

VARIABLES - PFOS	Models			
	Tobit		Probit	
	Marginal effect (y^*)	Marginal effect for Censored Sample (y)	Likelihood	Marginal effect
Size of the PWS – Small (1=small, 0= large)	-0.126** (0.0571)	-0.000009** (0.00000376)	-0.328** (0.148)	-0.00244*** (0.0007582)
Source of water to the PWS- Surface (SW)	-0.151*** (0.0277)	-0.000018*** (0.00000629)	-0.382*** (0.0690)	-0.00404*** (0.0007939)
Source of water to the PWS- Mixed (MX)	-0.0159 (0.0567)	-0.000002 (0.00000669)	-0.00229 (0.143)	-0.00002 (0.00151)
Source of water to the PWS- Mix but dominated by Ground water (GU)	0.138** (0.0685)	0.000016* (0.00000951)	0.397** (0.173)	0.00419** (0.0018761)
Population (log)	0.103*** (0.0194)	0.000012*** (0.00000424)	0.277*** (0.0485)	0.00293*** (0.0005201)
Non-White population (%)	- 0.00751*** (0.00171)	-0.000001*** (3.32E-07)	- 0.0204*** (0.00429)	-0.00022*** (0.0000455)
Poverty (%)	-0.00377 (0.00537)	-4.44E-07 (6.43E-07)	-0.0103 (0.0138)	-0.00011 (0.0001452)
Nonwhite Poverty (%)	-0.000242 (0.00275)	-2.85E-08 (3.24E-07)	0.000405 (0.00711)	0.00000 (0.0000751)
Log Housing density (house/sq mil)	0.0534*** (0.0179)	0.000006** (0.00000282)	0.146*** (0.0450)	0.00154*** (0.0004776)
Percentage Contribution to the GDP from				
Agriculture	-0.00138 (0.00430)	-1.63E-07 (5.08E-07)	-0.00314 (0.0110)	-0.00003 (0.0001161)
Durable goods manufacturing	-0.00179 (0.00292)	-2.11E-07 (3.51E-07)	-0.00338 (0.00752)	-0.00004 (0.0000792)
Non-durable good manufacture	0.0117*** (0.00214)	0.000001*** (4.78E-07)	0.0325*** (0.00527)	0.00034*** (0.0000575)
Healthcare and social assistance	0.0172*** (0.00360)	0.000002*** (7.43E-07)	0.0471*** (0.00903)	0.00050*** (0.0000997)
Food and accommodation	0.0266*** (0.00497)	0.000003*** (0.00000111)	0.0728*** (0.0122)	0.00077*** (0.0001371)
Government enterprise	0.00814*** (0.00177)	0.000001*** (3.53E-07)	0.0231*** (0.00444)	0.00024**** (0.0000484)
Constant	-2.316*** (0.261)		-6.199*** (0.604)	
Sigma	(0.3881836 (0.0202158)			
Observations	35,589		27,667	
Log likelihood	-1051.62		-1315.18	
LR chi2(44)	587.53		453.07	
Prob > chi2	0		0	
Pseudo R2	0.2184		0.1469	
AIC	2187.249		2712.366	
BIC	2543.401		3049.714	

Standard errors in parentheses

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

Table A10: PWS characteristics and socioeconomic factors affecting the PFHpA contamination in the USA.

VARIABLES - PFHPA	Model			
	Tobit		Probit	
	Marginal effect (y^*)	Marginal effect for Censored Sample (y)	Likelihood	Marginal effect
Size of the PWS – Small (1=small, 0= large)	-0.0277*** (0.00943)	-0.00000060** (0.00000028)	-0.576*** (0.184)	-0.00251*** (0.0004785)
Source of water to the PWS- Surface (SW)	0.00383 (0.00348)	0.00000017 (0.00000017)	0.0813 (0.0658)	0.00065 (0.0005254)
Source of water to the PWS- Mixed (MX)	0.00342 (0.0110)	0.00000015 (0.00000049)	0.0992 (0.204)	0.00079 (0.001628)
Source of water to the PWS- Mix but dominated by Ground water (GU)	-0.0211 (0.0185)	-0.00000093 (0.00000090)	-0.391 (0.347)	-0.00312 (0.0027917)
Population (log)	0.0167*** (0.00319)	0.00000074** (0.00000034)	0.314*** (0.0581)	0.00251*** (0.0004814)
Non-White population (%)	- 0.000967*** (0.000221)	-0.00000004** (0.00000002)	-0.0183*** (0.00406)	-0.00015*** (0.0000336)
Poverty (%)	0.000120 (0.000782)	0.00000001 (0.00000003)	0.000150 (0.0148)	0.000001 (0.0001182)
Nonwhite Poverty (%)	-0.000503 (0.000429)	-0.00000002 (0.00000002)	-0.00762 (0.00814)	-0.00006 (0.0000648)
Log Housing density (house/sq mil)	-0.00196 (0.00290)	-0.00000009 (0.00000013)	-0.0282 (0.0546)	-0.00023 (0.0004366)
Percentage Contribution to the GDP from				
Agriculture	-0.000275 (0.000694)	-0.00000001 (0.00000003)	-0.00410 (0.0130)	-0.00003 (0.0001035)
Durable goods manufacturing	-0.00170*** (0.000474)	-0.00000008** (0.00000004)	-0.0323*** (0.00878)	-0.00026*** (0.0000698)
Non-durable good manufacture	0.00108*** (0.000327)	0.00000005* (0.00000002)	0.0214*** (0.00608)	0.00017*** (0.0000491)
Healthcare and social assistance	0.00147*** (0.000511)	0.00000006** (0.00000004)	0.0288*** (0.00962)	0.00023*** (0.0000786)
Food and accommodation	0.000923 (0.000909)	0.00000004 (0.00000004)	0.0180 (0.0171)	0.00014 (0.0001358)
Government enterprise	0.00126*** (0.000216)	0.00000006** (0.00000003)	0.0247*** (0.00388)	0.00020*** (0.0000369)
Constant	-0.279*** (0.0382)		-5.297*** (0.660)	
Sigma	.0532621 .0032024			
Observations	35,589		25,844	
Log likelihood	-402.85		-1051.33	
LR chi2(44)	660.28		508.42	
Prob > chi2	0		0	
Pseudo R2	0.4504		0.1947	
AIC	885.7002		2180.663	
BIC	1224.892		2498.897	

Standard errors in parentheses

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

Table: Getis-Ord ($G_i(d)$) Statistics for Hot-Spot analysis

Variables	z≤2.58	-2.58z≤1.96	-1.96z<1.96	1.96≤z<2.58	2.58≤z
Total PFOA contamination per PWS in the county	0	0	1481	19	116
Total PFOS contamination per PWS in the county	0	0	1591	20	77
Total PFHpA contamination per PWS in the county	0	0	1484	28	104
Total PFHxS contamination per PWS in the county	0	0	1562	27	27
Total of all PFAS contamination per PWS in the county	0	0	1465	29	122