Supplemental Appendix

Quantifying Threshold Manipulation in the Presence of Rounding: The Case of Lead Monitoring in US Drinking Water

Tihitina Andarge, Dalia Ghanem, David A. Keiser, and Gabriel E. Lade

July 30, 2025

Table of Contents

A Supplemental Information for Section I **2**

A.1 Rounding bias when the direction of manipulation is up 2

A.2 Proof of Proposition 1 2

A.3 Simplification of η_j^p 4

B Supplemental Information for Section II **5**

B.1 Background on the Lead and Copper Rule and the Flint Water Crisis . . 5

B.2 Implementation 6

B.3 Robustness Checks 8

B.4 Spatial Distribution of Threshold Values 8

A Supplemental Information for Section I

A.1 Rounding bias when the direction of manipulation is up

Suppose agents have an incentive to manipulate data to be greater than or equal to some policy threshold c_j, such as in educational settings where funding thresholds depend on average test scores. In these cases, the proportion of manipulation is given by $\Pr(Z = 1) = \Pr(X \geq c_j) - \Pr(X(0) \geq c_j)$. If we use $X^N(0)$ to obtain the counterfactual probability of having a value greater than or equal to c_j absent manipulation instead of $X(0)$, then

$$\begin{aligned} & \Pr(X \geq c_j) - \Pr(X^N(0) \geq c_j) \\ = & \underbrace{\Pr(X \geq c_j) - \Pr(X(0) \geq c_j)}_{\text{Proportion of Threshold Manipulation}} + \underbrace{\Pr(X(0) \geq c_j) - \Pr(X^N(0) \geq c_j)}_{\text{Rounding Bias}}. \end{aligned}$$

The rounding bias will be positive if $X^N(0)$ is rounded up to c_j, while there will be no rounding bias if $X^N(0)$ can only be rounded down to c_j.

A.2 Proof of Proposition 1

Here we show that $\Pr(Z = 1|X \in (\underline{x}_j, \overline{x}_j)) = 0$ implies the testable implication. We first show (1) $\Pr(l_j \leq X(0) < u_j) = \Pr(l_j \leq X < u_j)$ and then (2) $\Pr(l_j \leq X(0) < u_j) = G(u_j; \boldsymbol{\theta}) - G(l_j; \boldsymbol{\theta})$.

To show (1), note that by the law of total probability,

$$\text{(A1)} \quad \Pr(l_j \leq X(0) < u_j) = \Pr(l_j \leq X(0) < u_j, Z = 1) + \Pr(l_j \leq X(0) < u_j, Z = 0).$$

To show (1), we need to show that $\Pr(l_j \leq X(0) < u_j, Z = 1) = 0$ for two possible cases: (a) $l_j \leq \underline{x}_j$ and (b) $l_j > \underline{x}_j$.
Case (a):

$$\begin{aligned} \Pr(l_j \leq X(0) < u_j, Z = 1) &= \Pr(l_j \leq X(0) \leq \underline{x}_j, Z = 1) + \Pr(\underline{x}_j < X(0) < u_j, Z = 1) \\ \text{(A2)} \qquad &= 0 \end{aligned}$$

The second term on the RHS of the first equality is zero since $\Pr(\underline{x}_j < X(0) < u_j, Z = 1) \leq \Pr(\underline{x}_j < X(0) \leq \overline{x}_j, Z = 1) = 0$ under the condition $\Pr(Z = 1|X \in (\underline{x}_j, \overline{x}_j)) = 0$. The first term on the RHS of the first equality is also zero, since in this case $\overline{x}_{j-1} < l_j \leq \underline{x}_j$

for $j \in \{2, \ldots, J\}$.[1] That is $[l_j, \underline{x}_j]$ neither intersects with $(\underline{x}_{j-1}, \overline{x}_{j-1})$ nor $(\underline{x}_j, \overline{x}_j)$. It then follows that $\Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j]) = 0$ by Assumption 2.1, which rules out manipulation outside the manipulation windows. Now note that

$$\Pr(Z = 1, X(0) \in [l_j, \underline{x}_j]) \leq \Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j]),$$

where the inequality follows from Assumption 1 as well as $\underline{x}_j < c_j$.[2] As a result, it follows that $\Pr(Z = 1, X(0) \in [l_j, \underline{x}_j]) \leq \Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j]) = 0$.

We have thus shown equation (A2). Now plugging it into equation (A1) yields $\Pr(l_j \leq X(0) < u_j) = \Pr(l_j \leq X(0) < u_j, Z = 0)$, which implies that $\Pr(Z = 1 | l_j \leq X(0) < u_j) = 0$, and therefore $\Pr(l_j \leq X(0) < u_j) = \Pr(l_j \leq X < u_j)$.[3] We have thus shown (1) for Case (a).

As for Case (b), equation (A2) is immediate from $\Pr(Z = 1 | X \in (\underline{x}_j, \overline{x}_j)) = 0$, since in this case $[l_j, u_j) \subset (\underline{x}_j, \overline{x}_j)$. Therefore, $\Pr(Z = 1 | X(0) \in (\underline{x}_j, \overline{x}_j)) = 0$ implies $\Pr(Z = 1 | X(0) \in [l_j, u_j)) = 0$, which implies $\Pr(Z = 1, X(0) \in [l_j, u_j)) = 0$ by Bayes' rule. It follows that (1) holds in Case (b) by similar arguments to Case (a).

(2) follows from equations (2) and (3), which hold by Assumptions 1 and 2.

---

1. For the case where $j = 1$, the same argument applies, since there is no manipulation below $\underline{x}_1$ by construction.

2. To see this, note that

$$\begin{aligned} \Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j]) &= \Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j], R = 1) + \Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j], R = 0) \\ &= \Pr(Z = 1, X^N(0) \in [l_j, \underline{x}_j], X(0) = c_j) + \Pr(Z = 1, X(0) \in [l_j, \underline{x}_j]). \end{aligned}$$

The first equality follows from the law of total probability. The first term on the RHS of the second equality follows from the equivalence of the two events: $\{R = 1, X^N(0) \in [l_j, \underline{x}_j]\}$ and $\{X(0) = c_j, X^N(0) \in [l_j, \underline{x}_j]\}$. The second term on the RHS of the second equality follows from the equivalence of the two events $\{X^N(0) \in [l_j, \underline{x}_j], R = 0\}$ and $\{X(0) \in [l_j, \underline{x}_j]\}$ under Assumption 1.

3. To see this, note that

$$\begin{aligned} \Pr(l_j \leq X < u_j) &= \Pr(l_j \leq X < u_j, Z = 1) + \Pr(l_j \leq X < u_j, Z = 0) \\ &= \Pr(l_j \leq X < u_j, Z = 1) + \Pr(l_j \leq X(0) < u_j, Z = 0) \\ &= \Pr(l_j \leq X \leq \underline{x}_j, Z = 1) + \Pr(\underline{x}_j < X < u_j, Z = 1) + \Pr(l_j \leq X(0) < u_j, Z = 0) \\ &= \Pr(l_j \leq X(0) < u_j, Z = 0) = \Pr(l_j \leq X(0) < u_j). \end{aligned}$$

The first equality follows from the law of total probability. The second equality follows from $X(0) = X$ when $Z = 0$. The third equality follows from the law of total probability, since $[l_j, \overline{x}_j) = [l_j, u_j) \cup [u_j, \overline{x}_j)$. The penultimate equality follows from two observations: (1) $\Pr(\underline{x}_j \leq X < u_j, Z = 1) \leq \Pr(X \in (\underline{x}_1, \overline{x}_1), Z = 1) = 0$, since $(\underline{x}_j, u_j) \subset (\underline{x}_j, \overline{x}_j)$, (2) $\Pr(l_j \leq X \leq \underline{x}_j, Z = 1) = 0$ by Assumption 2.1. The last equality follows from the law of total probability and equation (A2).

A.3 Simplification of η_j^p

Recall that η_j^p in Section IB is given by,

$$\eta_j^p = \Pr(X = c_j) - \sum_{r \in \mathcal{R}_j} \underbrace{\frac{\Pr(l_r \leq X < u_r)}{\sum_{r \in \mathcal{R}_j} \Pr(l_r \leq X < u_r)} \Pr(X = r | l_r \leq X < u_r)}_{\equiv \omega_r^p \Pr(X = r | l_r \leq X < u_r)} [G(u_j; \boldsymbol{\theta}) - G(l_j; \boldsymbol{\theta})].$$

Here we demonstrate how this probability-weighted excess rounding estimand admits a convenient simplification. For $r \in \mathcal{R}_j$,

$$\begin{aligned}
& \omega_r^p \Pr(X = r | l_r \leq X < u_r) \\
= \; & \frac{\Pr(l_r \leq X < u_r) \Pr(R = 1 | l_r \leq X < u_r)}{\Pr(X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r))} = \frac{\Pr(R = 1, l_r \leq X < u_r)}{\Pr(X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r))} \\
= \; & \frac{\Pr(R = 1, l_r \leq X < u_r, X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)) + \Pr(R = 1, l_r \leq X < u_r, X \notin \cup_{r \in \mathcal{R}_j} [l_r, u_r))}{\Pr(X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r))} \\
= \; & \Pr(R = 1, l_r \leq X < u_r | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)).
\end{aligned}$$

The first equality follows from the assumption that there is no manipulation in the roundable intervals $[l_r, u_r)$ as well as the continuity of $X^N(0)$ by Assumption 2.3, and therefore it follows that $\Pr(X = r | l_r \leq X < u_r) = \Pr(R = 1 | l_r \leq X < u_r)$. The second equality follows by Bayes' rule. The third equality follows from the law of total probability, and the last equality follows from the second term in the numerator on its LHS being zero as well as Bayes' rule.

As a result, it follows that

$$\eta_j^p = \Pr(X = c_j) - \sum_{r \in \mathcal{R}_j} \Pr(R = 1, l_r \leq X < u_r | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)) [G(u_j; \boldsymbol{\theta}) - G(l_j; \boldsymbol{\theta})].$$

Since $\{R = 1, l_r \leq X < u_r\}_{r \in \mathcal{R}_j}$ is a sequence of mutually exclusive events, it follows that

$$\begin{aligned}
& \sum_{r \in \mathcal{R}_j} \Pr(R = 1, l_r \leq X < u_r | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)) \\
= \; & \Pr(\cup_{r \in \mathcal{R}_j} \{R = 1, l_r \leq X < u_r\} | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)) \\
\text{(A3)} \; = \; & \Pr(R = 1, X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r) | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)) = \Pr(R = 1 | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)).
\end{aligned}$$

The probability-weighted excess rounding estimand then simplifies to the following,

$$\eta_j^p = \Pr(X = c_j) - \Pr(R = 1 | X \in \cup_{r \in \mathcal{R}_j} [l_r, u_r)) [G(u_j; \boldsymbol{\theta}) - G(l_j; \boldsymbol{\theta})].$$

B Supplemental Information for Section II

B.1 Background on the Lead and Copper Rule and the Flint Water Crisis

The Lead and Copper Rule is the primary regulation aimed at monitoring the risk of lead exposure in drinking water. It sets monitoring and remediation requirements for public water systems (PWSs) in the United States.[4] The rule requires that all community water systems (CWS) and non-transient non-community water systems (NTNCWS) periodically collect lead and copper samples from high-risk homes in their service territories.[5] PWSs report the ninetieth percentile of the sample values for both lead and copper. If the values are at or below the "reporting" threshold, 0.005 mg/L for lead and 0.65 mg/L for copper, for two consecutive six-month periods, then systems may be eligible for reduced tap monitoring. If the values exceed a higher "action" level, 0.015 mg/L for lead and 1.3 mg/L for copper, systems must take costly actions to reduce lead and copper levels and educate the public. Exceeding either threshold increases regulatory scrutiny and anticipated costs.

The Flint water crisis began in April 2014 when the city of Flint switched its water supply from the Detroit Water and Sewer Department to the Flint River. The Flint River's water was more corrosive and caused lead to leach into the city's water supply due to improper corrosion control. Despite complaints from residents and businesses, public officials ensured the public that the water was safe for human consumption through the Fall of 2015. A series of revelations by private and public actors in the Fall of 2015 revealed the extent of lead contamination throughout the city, resulting in several health emergency declarations in late 2015 and January 2016 (Christensen, Keiser, and Lade 2023).

Following these revelations, the Flint water crisis received substantial national and international news coverage, the University of Michigan, Flint hosted a presidential debate, and the crisis became the subject of popular documentary films and books (Acosta 2016; Clark 2018; Barnhart 2019; Baxter 2021). Moreover, the EPA issued lead monitoring recommendations that explicitly discouraged some of the "cheats" water systems may use to lower measured lead concentrations. Specifically, it recommended that PWSs (1) refrain from in-

4. Since its promulgation in 1991, there have been three revisions to the rule before our study period (2000, 2004, 2007) and one change after our study period (2021). Due partly to the Flint crisis, the EPA has proposed comprehensive revisions.

5. A CWS serves the same population throughout the year, while a NTNCWS serves at least 25 of the same people for six or more months each year.

structing households to remove or clean aerators before sampling; (2) remove any steps that involve pre-stagnation flushing from the instructions; and (3) stop using wide-mouth bottles to collect samples (Grevatt 2016). Overall, the Flint water crisis greatly increased scrutiny around lead monitoring and reporting, motivating our pre- vs. post-Flint analysis.

B.2 Implementation

Our empirical application has two policy thresholds: the reporting threshold ($c_1 = 0.005$) and the action level ($c_2 = 0.015$). Assumption 1 requires knowledge of the roundable intervals around c_1 and c_2. In our context, PWSs are allowed to round reported ninetieth-perentile lead concentrations to the thousandth decimal place. Thus, we may observe $x = c_j$ for any value of $x \in [c_j - 0.0005, c_j + 0.0005)$. Correspondingly, we define roundable intervals around c_1 and c_2 as $[0.0045, 0.0055)$ and $[0.0145, 0.0155)$, respectively.

Next, in the censored maximum likelihood objective function, we define $(\underline{x}_1, \overline{x}_1) = (0.0035, 0.0075)$ and $(\underline{x}_2, \overline{x}_2) = (0.0135, 0.0175)$ as the manipulation windows around the reporting threshold and action level, respectively.[6] Our selection of manipulation windows is motivated by economic theory. In our setting, water systems have an incentive to manipulate lead values to meet the thresholds; they have no incentive to manipulate lead values to be much lower than the thresholds. However, given that the "cheats" available to water systems may give them imprecise control over their lead values, we choose lower bounds for the manipulation windows that are close but not equal to the thresholds. In choosing the upper bounds of the manipulation windows, we take into consideration that the probability of being detected by the regulator and the severity of the potential subsequent punishment may be higher for water systems that manipulate true values that are far away from the threshold. For this reason, we hypothesize that water systems with lead values above but

6. Selection of the manipulations windows should ensure that $\underline{x}_j$ and $\overline{x}_j$ do not coincide with rounded values since it might lead to a worse fit of the censored component of the maximum likelihood counterfactual. Specifically, $[G(\overline{x};\theta) - G(\underline{x};\theta)]$ would not equal $F_{X(0)}(\overline{x}) - F_{X(0)}(\underline{x})$ if $\overline{x}$ and $\underline{x}$ are rounded values. There are two helpful points to recall to see this point. First, since manipulation occurs inside the manipulation windows $(\underline{x}, \overline{x})$ by Assumption 2.1, the observed CDF equals the unmanipulated CDF at both values. Specifically $F_X(\underline{x}) = F_{X(0)}(\underline{x})$ and $F_X(\overline{x}) = F_{X(0)}(\overline{x})$. Second, under Assumption 2.3, $G(x;\boldsymbol{\theta}) = F_{X^N(0)}(x)$. Because rounding moves mass to rounded values, the distribution of $X^N(0)$ and $X(0)$ may not be equal at those points. For example, in our application, 0.004 is a rounded value. Therefore, $\Pr(X(0) \leq 0.004) = \Pr(X(0) < 0.004) + \Pr(X(0) = 0.004)$. If there is rounding down of $X^N(0)$ from values in $(0.004, 0.0045)$ to 0.004, then $\Pr(X(0) \leq 0.004) > \Pr(X^N(0) \leq 0.004)$, and for that reason, $F_{X(0)}(0.004)$ is not equal to $F_{X^N(0)}(0.004)$. Using 0.0035 is preferable since there is no rounding at that value, and as a result, $\Pr(X(0) \leq 0.0035) = \Pr(X^N(0) \leq 0.0035)$. Similar arguments support choosing 0.0075 and not 0.007.

close to the threshold are more likely to manipulate their data. As a robustness check, in Supplemental Appendix B.3, we assess the sensitivity of our results to this choice.

We estimate the parameter vector $\widehat{\boldsymbol{\theta}}$ via censored maximum likelihood estimation and obtain the counterfactual CDFs. Figures B.9 to B.11 present the empirical CDFs and the counterfactual CDFs for the three system sizes, pre- and post-Flint. The counterfactual CDFs fit the observed distribution quite well outside of the manipulation windows.

To perform the manipulation test using equation (16), we need to estimate the difference between the observed and counterfactual probabilities of being within the roundable interval around c_j, $\widehat{\Pr}(l_j \leq X < u_j)$ - $\widehat{\Pr}(l_j \leq X(0) < u_j)$. The observed probability of being within the roundable interval is just the observed proportion of values in $[l_j, u_j)$ for $j = 1, 2$. To compute the counterfactual probability of being within the roundable interval around c_j for $j = 1, 2$, we use the counterfactual CDF estimated via CMLE and calculate: $G(u_j; \widehat{\boldsymbol{\theta}}) - G(l_j; \widehat{\boldsymbol{\theta}})$. We then estimate the test statistic in equation (16), computing standard errors via nonparametric bootstrap resampling PWSs.

To estimate the excess rounding measures in equation (17), we need to compute two components. The first component is the sample probability of rounding to c_j, $\widehat{\Pr}(X = c_j)$, for $j = 1, 2$. Estimating the second component, the counterfactual (absent manipulation) probability of rounding to c_j, $\widehat{\Pr}(X(0) = c_j)$, requires defining non-policy thresholds that best mimic rounding at the policy threshold absent manipulation. We define $\mathcal{R}_1 = \{0.001, 0.002, 0.003, 0.008, 0.009\}$ and $\mathcal{R}_2 = \{0.01, 0.011, 0.012, 0.013, 0.018, 0.019, 0.02\}$ for c_1 and c_2, respectively. Our choice of $\mathcal{R}_1$ and $\mathcal{R}_2$ is guided by reasoning similar to our reasoning in the choice of manipulation windows. If manipulation occurs in small windows around the thresholds, then we expect only innocuous rounding at points with three decimal digits outside of the manipulation windows, making them suitable candidates for inclusion in $\mathcal{R}_1$ and $\mathcal{R}_2$.

The next step is to define roundable intervals for $r \in \mathcal{R}_j$ that do not overlap with the manipulation windows. Again, since rounding in our context is up to the third decimal point, the roundable interval around $r \in \mathcal{R}_j$ for $j = 1, 2$ is given by $[r - 0.0005, r + 0.0005)$. We then take the simple average and the probability-weighted average of the observed conditional probability of rounding to $r \in \mathcal{R}_j$ and use them to estimate $\hat{\eta}_j^s$ and $\hat{\eta}_j^p$, respectively, for $j = 1, 2$. Again, we compute standard errors via nonparametric bootstrap resampling PWSs.

Lastly, we relax the homogeneous rounding assumption (Assumption 3) and consider a bounded rounding assumption (Assumption 4). We estimate the bounds on η_j by computing the minimum and maximum of the observed conditional probability of rounding to $r \in \mathcal{R}_j$. Under Assumption 4, this provides the bounds on counterfactual rounding to c_j and subsequently bounds for η_j.

B.3 Robustness Checks

We consider a number of robustness checks to assess the sensitivity of our results. First, we assess the robustness of our results to the weighting scheme used in estimating equation (10). In the paper, we present estimates of our simple average excess rounding measure ($\hat{\eta}_j^s$) at c_j obtained using a simple average of the observed conditional probability of rounding to $r \in \mathcal{R}_j$. In Figure B.12, we present estimates of our probability-weighted excess rounding measure ($\hat{\eta}_j^p$). The results are similar, likely because rounding is relatively homogeneous across $r \in \mathcal{R}_j$ in our application. However, in settings where rounding is more heterogeneous, the choice of weighting scheme may lead to different estimates. For this reason, we recommend assessing the sensitivity of results to different weighting schemes.

Second, as the manipulation windows are user-defined, we examine the robustness of our results to a narrower manipulation window. The estimates of excess rounding at the thresholds, presented in Figure B.13, are similar to our baseline estimates. Last, as previously discussed, the incentive to manipulate data should only exist near the thresholds. Thus, estimating excess rounding at non-policy thresholds outside the manipulation windows can serve as a natural placebo test. Figure B.14 presents the results. In general, excess rounding at these non-policy thresholds is statistically indistinguishable from zero with a few exceptions.

B.4 Spatial Distribution of Threshold Values

While we cannot detect which PWSs manipulated their reports, we can explore the spatial distribution of rounding at the reporting threshold. Figure B.15 maps the service population-weighted proportion of PWSs with ninetieth-percentile reports that equal 0.005 in each county within the conterminous US for 2011-2015 (Panel A) and 2016-2020 (Panel B).[7] While several counties have a high proportion of PWSs reporting 0.005, counties in one state, Alabama, were responsible for most reports equaling the reporting threshold pre-Flint. This behavior ended in 2016.

Given this spatial pattern, we exclude Alabama from our sample and re-estimate our simple average excess rounding measure. Figure B.16 presents the results. We find that excess rounding is statistically indistinguishable from zero for small and large PWSs in both periods and at both thresholds. For medium PWSs, pre-Flint, excess rounding at the reporting

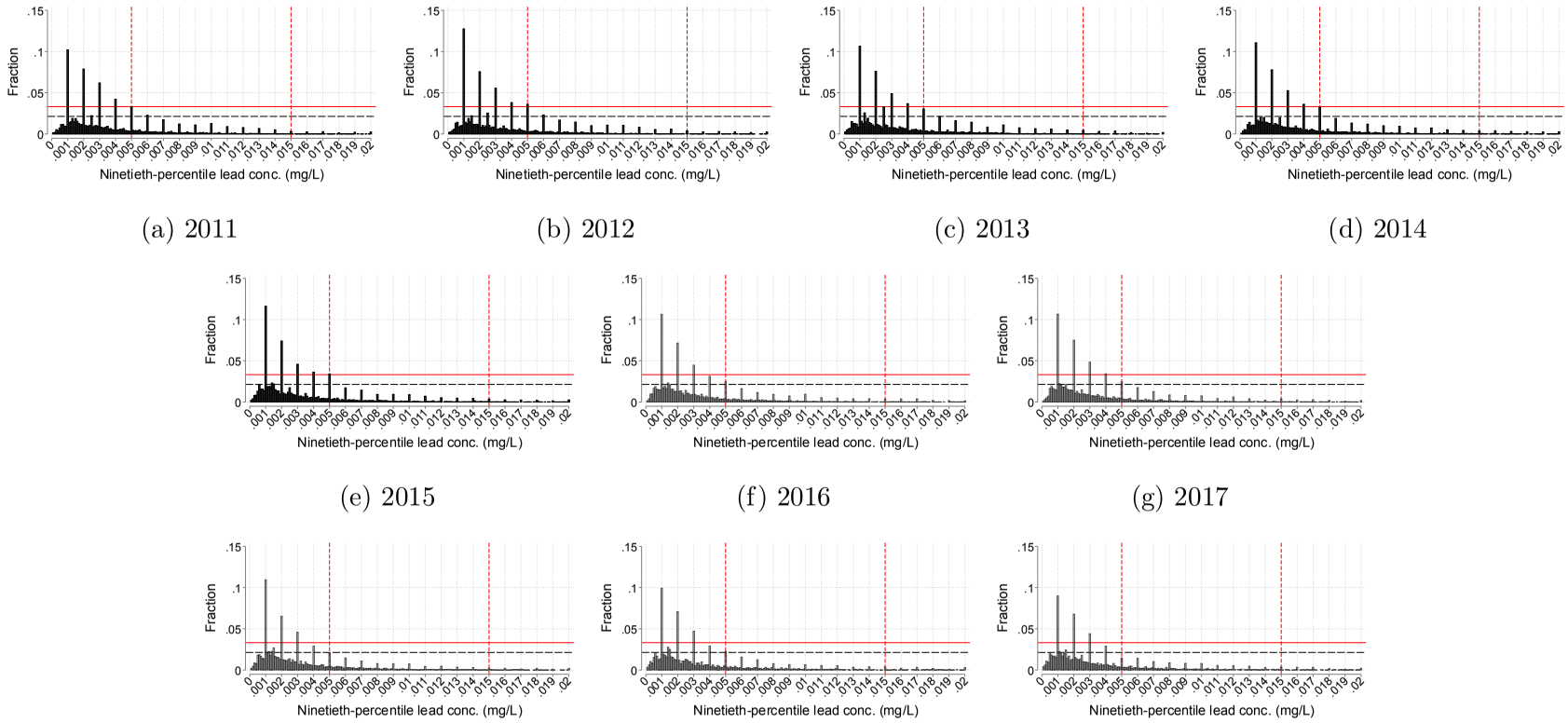
7. We locate PWSs inside the most populated county they serve using 2020 population estimates from the US Census Bureau (US Census Bureau 2021).

threshold is statistically different from zero but much smaller in magnitude compared to our primary estimate. Post-Flint, we find no evidence of threshold manipulation consistent with our primary findings. At the action level, however, the excess rounding estimates excluding Alabama are similar to our primary estimates in both periods. Overall, the results suggest just over 90 pre-Flint reports were excessively rounded outside of Alabama.

When presented with our results, the Alabama Department of Environmental Management (ADEM) stated that before 2016, they calculated ninetieth percentiles "by hand." If the value fell below the detection limit of the testing device, known as the method detection limit (MDL), they would report the MDL. This contrasted with LCR guidance, which stipulates that values below the MDL should be reported as zeros. In 2016, the state adopted a computerized reporting system, which correctly reported zeros when ninetieth-percentile values were below the MDL. Histograms of reported ninetieth-percentile lead values in Figure B.17 are consistent with this claim.

We draw two lessons from these findings. First, regulatory complexity can lead to reporting errors. Second, this situation occurred because Alabama regulators used water testing laboratories with high MDLs. When the LCR was proposed in 1988, the EPA estimated that, with existing technologies, MDLs for lead would be 0.001 mg/L (USEPA 1991). Over 20 years later, Alabama used labs with an MDL five times this amount and one corresponding to a critical policy threshold, 0.005 mg/L. This situation would not have occurred had the EPA mandated all states use labs with a lower detection limit. Technology mandates and clearer regulatory guidance may have prevented most of the threshold manipulation we detected. In the 2021 Lead and Copper Rule Revisions (LCRR), the EPA now requires labs to achieve an MDL of 0.001 mg/L as part of the certification process (USEPA 2021). We leave examining how these changes impact reporting behavior to future work.

Figure B.1: Histograms by Year: Small Systems



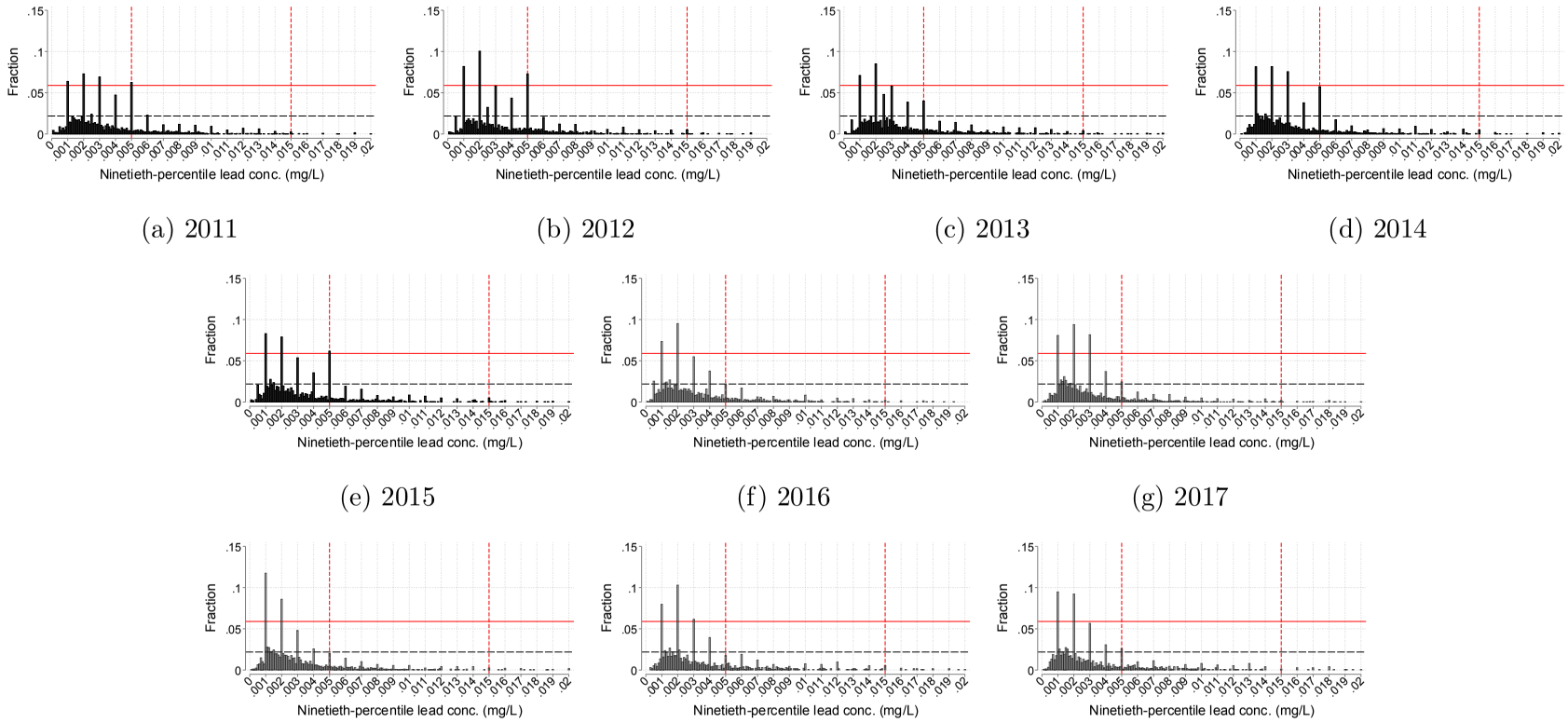
(a) 2011 (b) 2012 (c) 2013 (d) 2014

(e) 2015 (f) 2016 (g) 2017

(h) 2018 (i) 2019 (j) 2020

Notes: The figures display histograms of ninetieth-percentile lead values by year for small PWSs. Pre-Flint histograms are displayed in black and post-Flint histograms are displayed in grey. Solid red horizontal lines indicate the average pre-Flint proportion of reports equal to the reporting threshold (0.005). Dashed horizontal black lines indicate the average post-Flint proportion of reports equal to the reporting threshold.

Figure B.2: Histograms by Year: Medium Systems



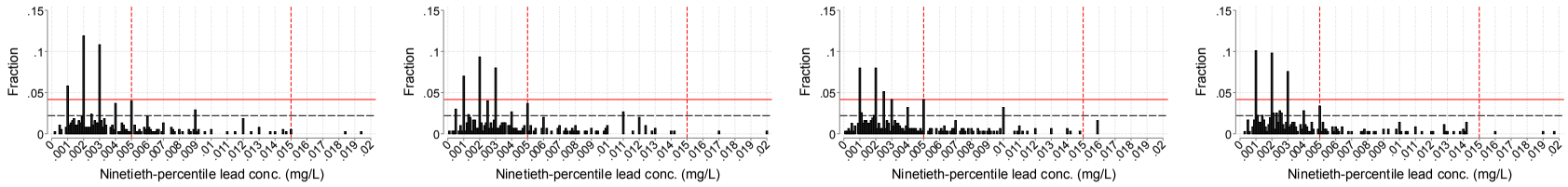
(a) 2011 (b) 2012 (c) 2013 (d) 2014

(e) 2015 (f) 2016 (g) 2017

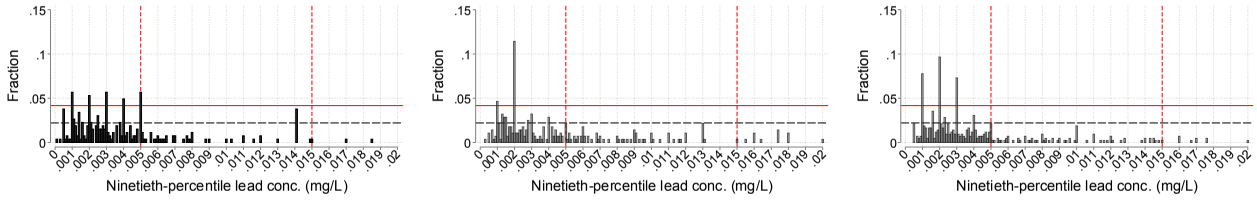
(h) 2018 (i) 2019 (j) 2020

Notes: The figures display histograms of ninetieth-percentile lead values by year for medium PWSs. Pre-Flint histograms are displayed in black and post-Flint histograms are displayed in grey. Solid red horizontal lines indicate the average pre-Flint proportion of reports equal to the reporting threshold (0.005). Dashed horizontal black lines indicate the average post-Flint proportion of reports equal to the reporting threshold.

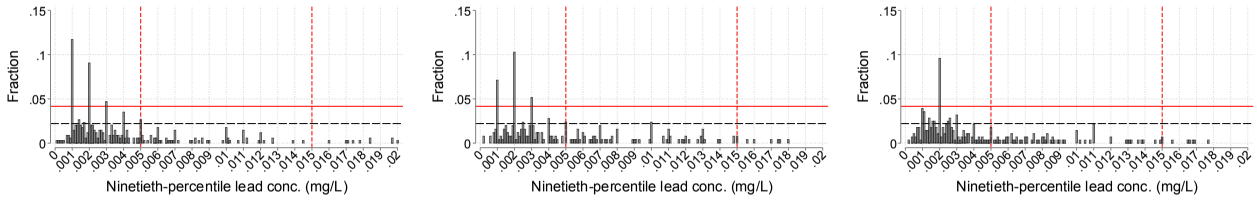
Figure B.3: Histograms by Year: Large Systems



(a) 2011 (b) 2012 (c) 2013 (d) 2014



(e) 2015 (f) 2016 (g) 2017



(h) 2018 (i) 2019 (j) 2020

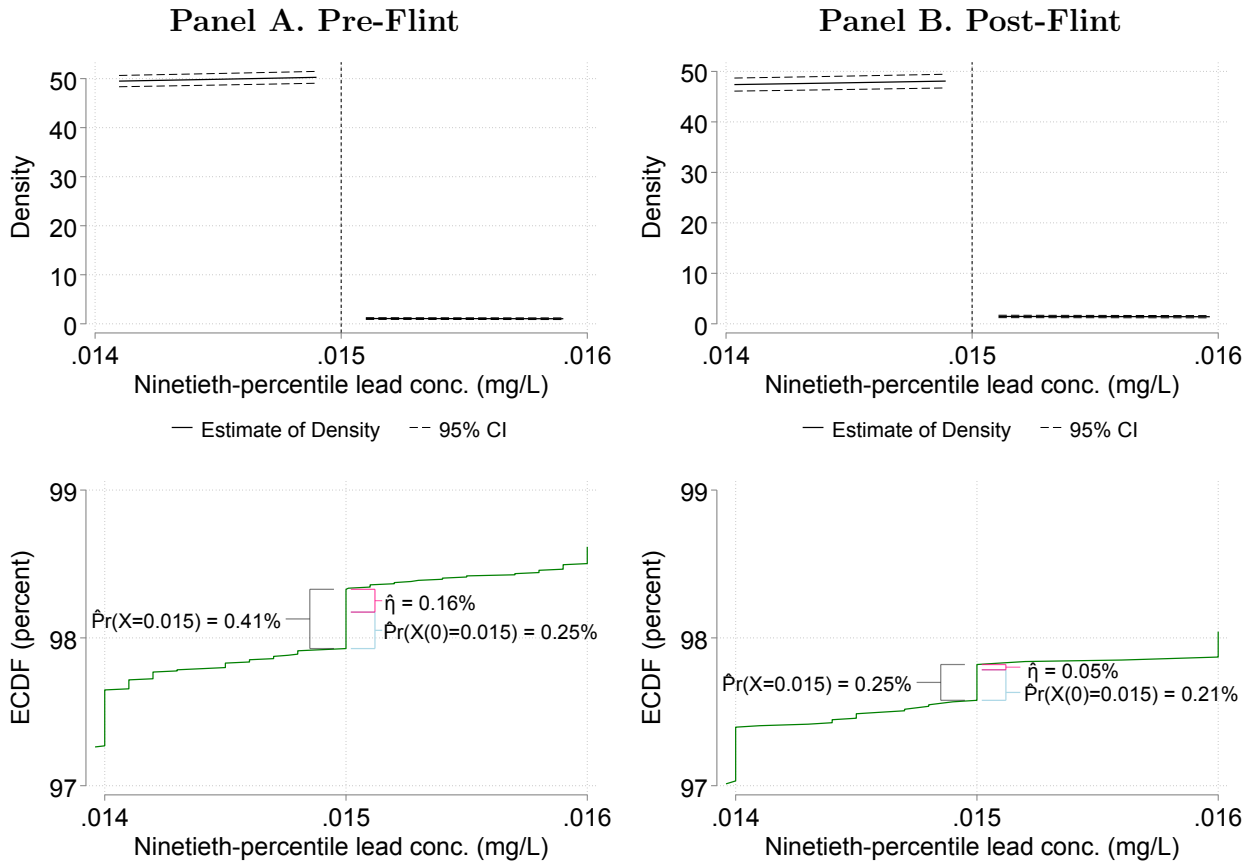
Notes: The figures display histograms of ninetieth-percentile lead values by year for large PWSs. Pre-Flint histograms are displayed in black and post-Flint histograms are displayed in grey. Solid red horizontal lines indicate the average pre-Flint proportion of reports equal to the reporting threshold (0.005). Dashed horizontal black lines indicate the average post-Flint proportion of reports equal to the reporting threshold.

Table B.1: McCrary Tests

	Small PWS		Medium PWS		Large PWS	
	Pre-Flint	Post-Flint	Pre-Flint	Post-Flint	Pre-Flint	Post-Flint
Reporting Threshold ($c_1 = 0.005$)	-1.018	-1.133	-1.404	-1.540	-1.414	-1.444
	(0.008)	(0.010)	(0.022)	(0.026)	(0.103)	(0.102)
Action Level ($c_2 = 0.005$)	-2.817	-2.723	-3.843	-3.492	-0.736	0.149
	(0.016)	(0.018)	(0.085)	(0.084)	(0.418)	(0.286)

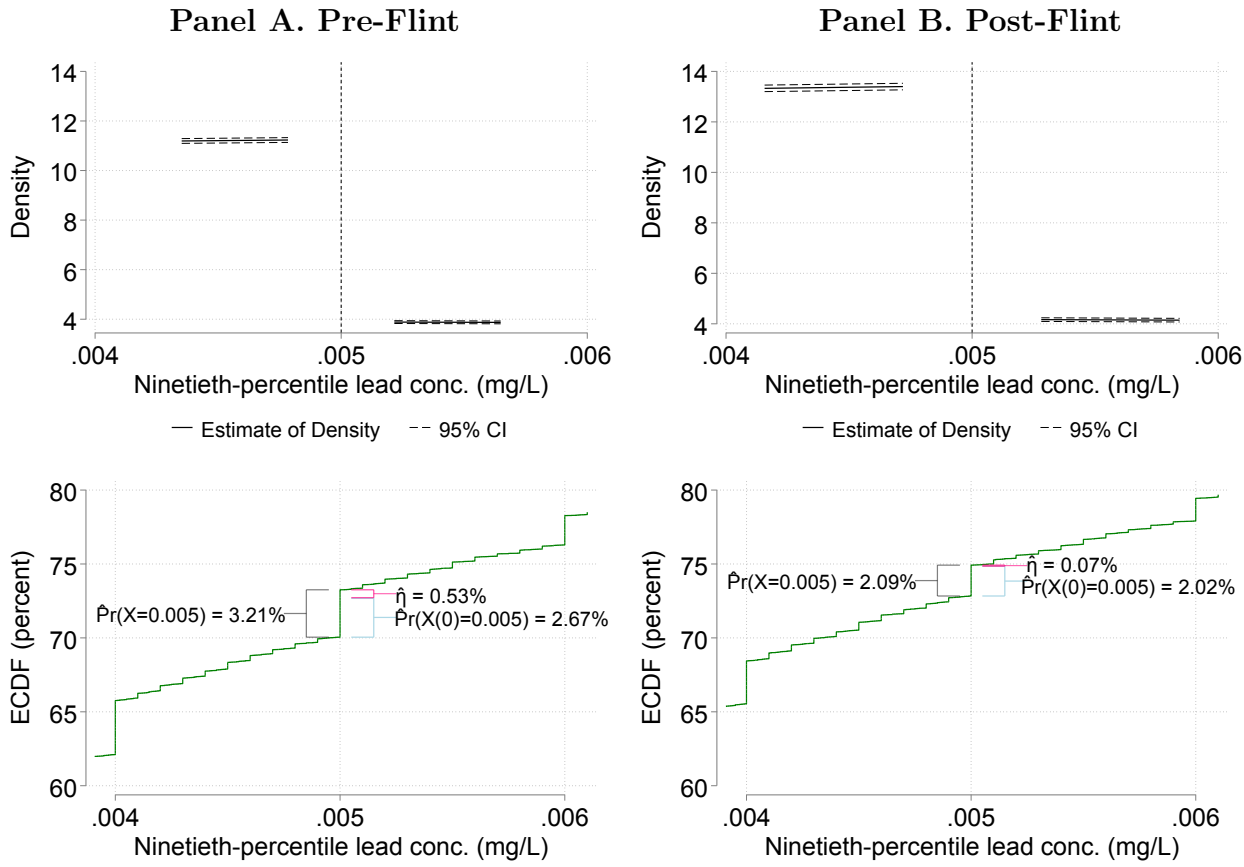
Notes: This table displays discontinuity estimates at the reporting threshold (0.005) and action level (0.015) from applying a McCrary (2008) test for manipulation for small, medium, and large PWSs. We report standard errors in parentheses. The approach in McCrary (2008) selects bin size as follows: $\hat{b} = 2\hat{\sigma}n^{-1/2}$ where $\hat{b}$ is the bin size, $\hat{\sigma}$ is the sample standard deviation of the running variable, and n is the sample size. Using this bin size, we are not able to obtain an estimate of the discontinuity for large water systems, pre-Flint, at the action level. For this reason, we instead use the smallest bin size for which we are able to obtain an estimate of the discontinuity and standard error ($\hat{b} = 0.0004501$).

Figure B.4: McCrary Test and Illustration of Excess Rounding at Action Level - Medium Public Water Systems



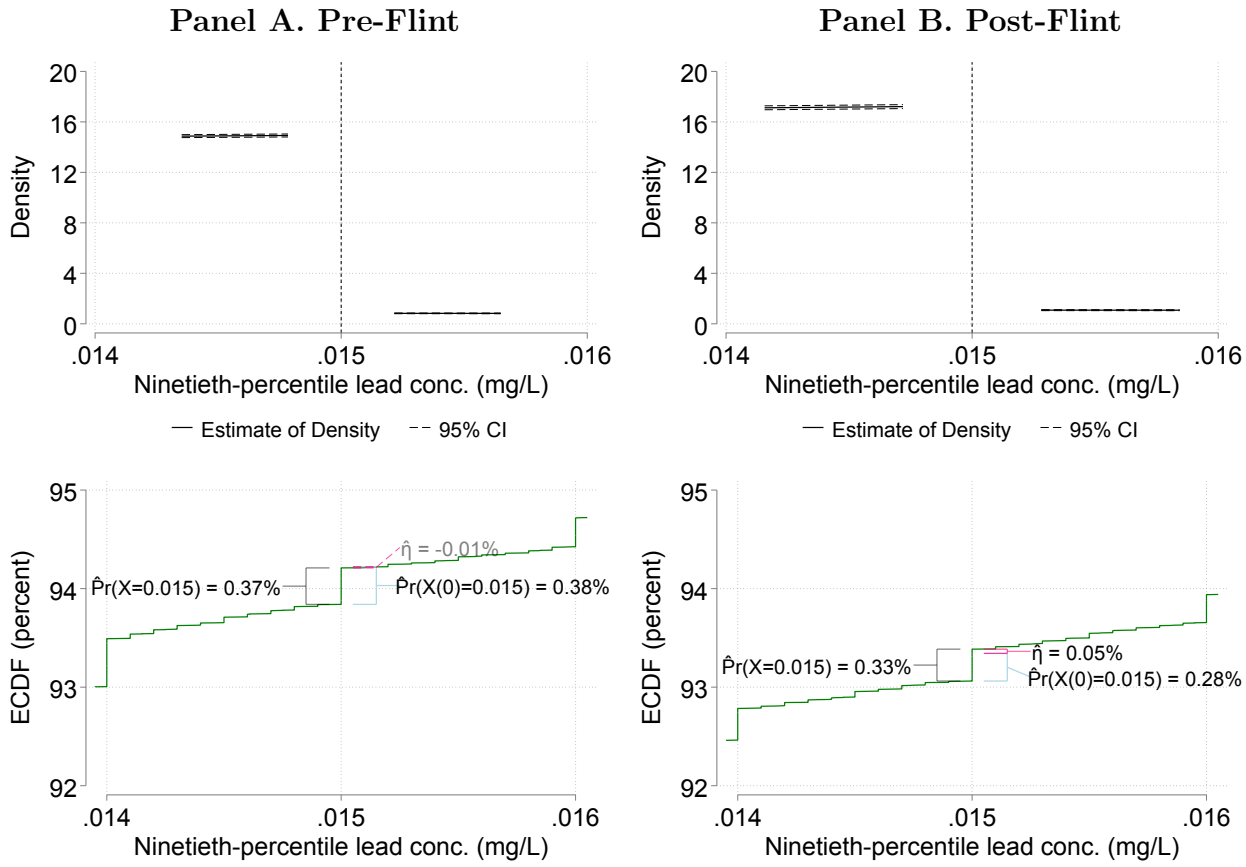
Notes: The top row of figures displays estimates of the density of ninetieth-percentile lead concentrations to the left and to the right of the action level (0.015) for medium-sized PWSs pre-Flint (Panel A) and post-Flint (Panel B). The dashed lines depict the 95% confidence interval around the estimate of the density. The bottom row of figures depicts the pre-Flint (Panel A) and post-Flint (Panel B) empirical CDFs (green line), the observed proportions of rounding at the action level ($\widehat{\Pr}(X = 0.015)$), the counterfactual proportions of rounding (absent manipulation) at the action level ($\widehat{\Pr}(X(0) = 0.015)$, and estimates of the simple average excess rounding measure at the action level ($\hat{\eta}^s$).

Figure B.5: McCrary Test and Illustration of Excess Rounding at Reporting Threshold - Small Public Water Systems



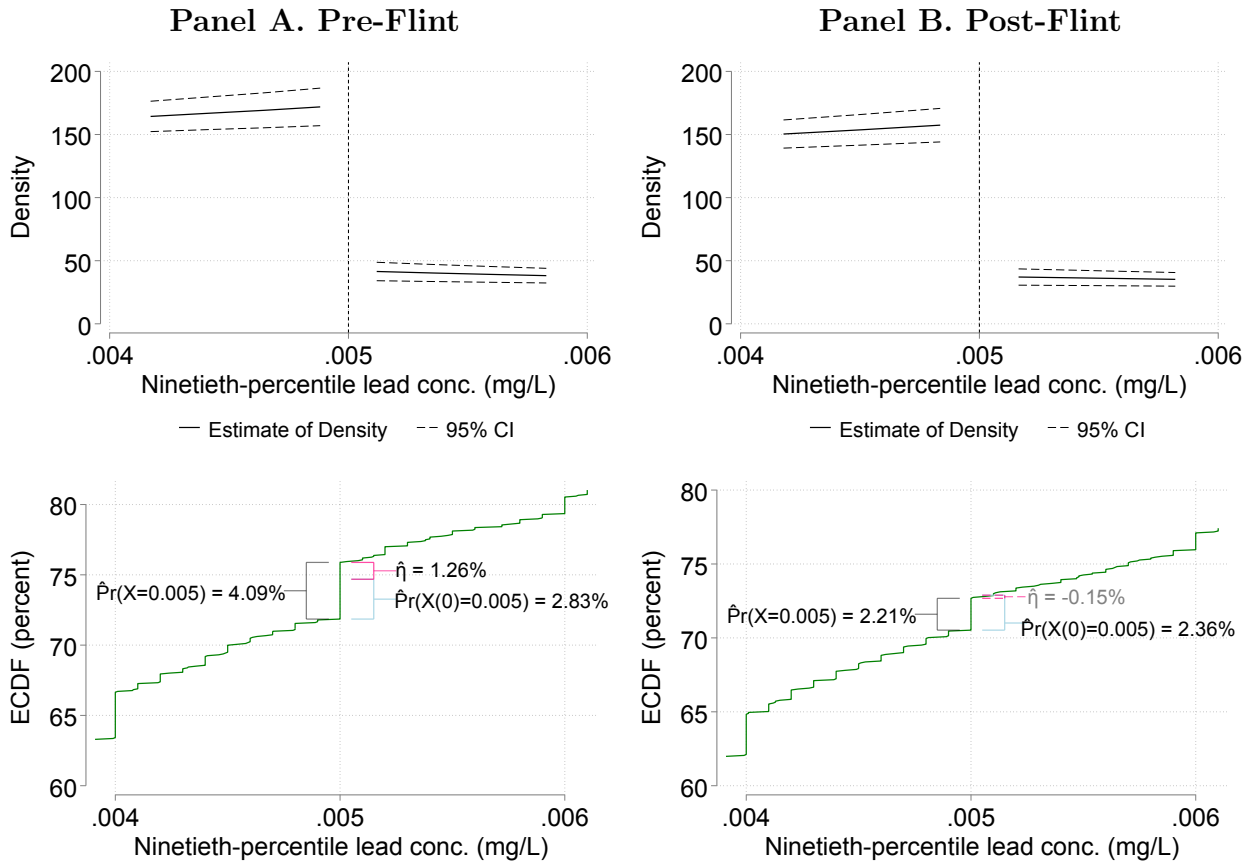
Notes: The top row of figures displays estimates of the density of ninetieth-percentile lead concentrations to the left and to the right of the reporting threshold (0.005) for small-sized community PWSs pre-Flint (Panel A) and post-Flint (Panel B). The dashed lines depict the 95% confidence interval around the estimate of the density. The bottom row of figures depicts the pre-Flint (Panel A) and post-Flint (Panel B) empirical CDFs (green line), the observed proportions of rounding at the reporting threshold ($\widehat{\Pr}(X = 0.005)$), the counterfactual proportions of rounding (absent manipulation) at the reporting threshold ($\widehat{\Pr}(X(0) = 0.005)$, and estimates of the simple average excess rounding measure at the reporting threshold ($\hat{\eta}^s$). To construct the density, we drop ninetieth-percentile lead values greater than 10 mg/L (four observations). Dropping these observations does not impact the conclusion of the McCrary test but facilitates visualization of the density estimates.

Figure B.6: McCrary Test and Illustration of Excess Rounding at Action Level - Small Public Water Systems



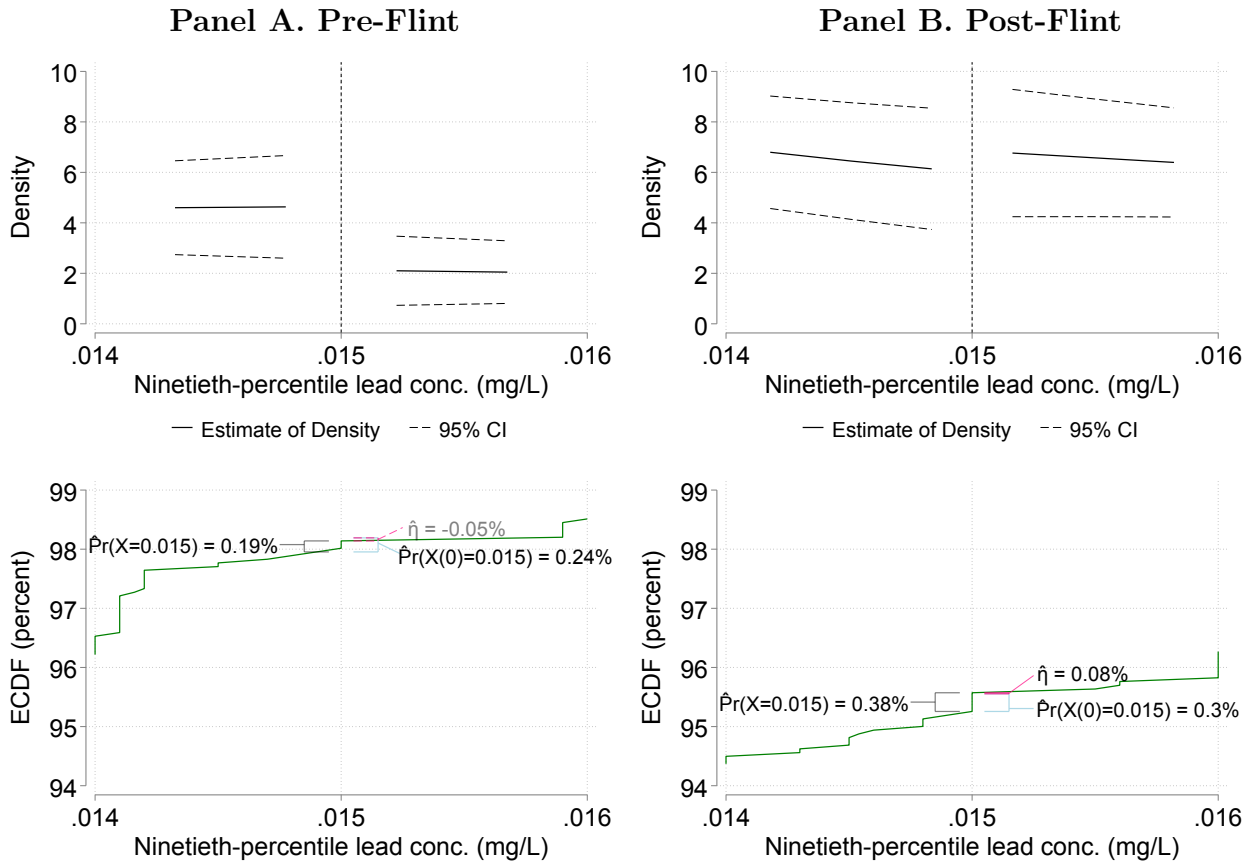
Notes: The top row of figures displays estimates of the density of ninetieth-percentile lead concentrations to the left and to the right of the action level (0.015) for small-sized PWSs pre-Flint (Panel A) and post-Flint (Panel B). The dashed lines depict the 95% confidence interval around the estimate of the density. The bottom row of figures depicts the pre-Flint (Panel A) and post-Flint (Panel B) empirical CDFs (green line), the observed proportions of rounding at the action level ($\widehat{\Pr}(X = 0.015)$), the counterfactual proportions of rounding (absent manipulation) at the action level ($\widehat{\Pr}(X(0) = 0.015)$, and estimates of the simple average excess rounding measure at the action level ($\hat{\eta}^s$). Greyed-out estimates of η^s are negative. To construct the density plots, we drop ninetieth-percentile lead values greater than 10 mg/L (four observations). Dropping these observations does not impact the conclusion of the McCrary test but facilitates visualization of the density estimates.

Figure B.7: McCrary Test and Illustration of Excess Rounding at Reporting Threshold - Large Public Water Systems



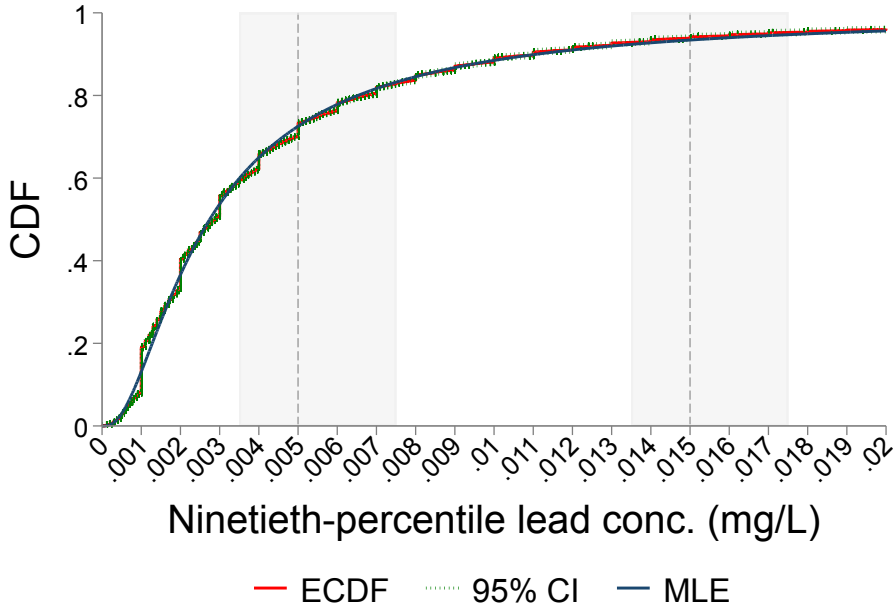
Notes: The top row of figures displays estimates of the density of ninetieth-percentile lead concentrations to the left and to the right of the reporting threshold (0.005) for large-sized PWSs pre-Flint (Panel A) and post-Flint (Panel B). The dashed lines depict the 95% confidence interval around the estimate of the density. The bottom row of figures depicts the pre-Flint (Panel A) and post-Flint (Panel B) empirical CDFs (green line), the observed proportions of rounding at the reporting threshold ($\widehat{\Pr}(X = 0.005)$), the counterfactual proportions of rounding (absent manipulation) at the reporting threshold ($\widehat{\Pr}(X(0) = 0.005)$, and estimates of the simple average excess rounding measure at the reporting threshold ($\hat{\eta}^s$). Greyed-out estimates of η^s are negative.

Figure B.8: McCrary Test and Illustration of Excess Rounding at Action Level - Large Public Water Systems

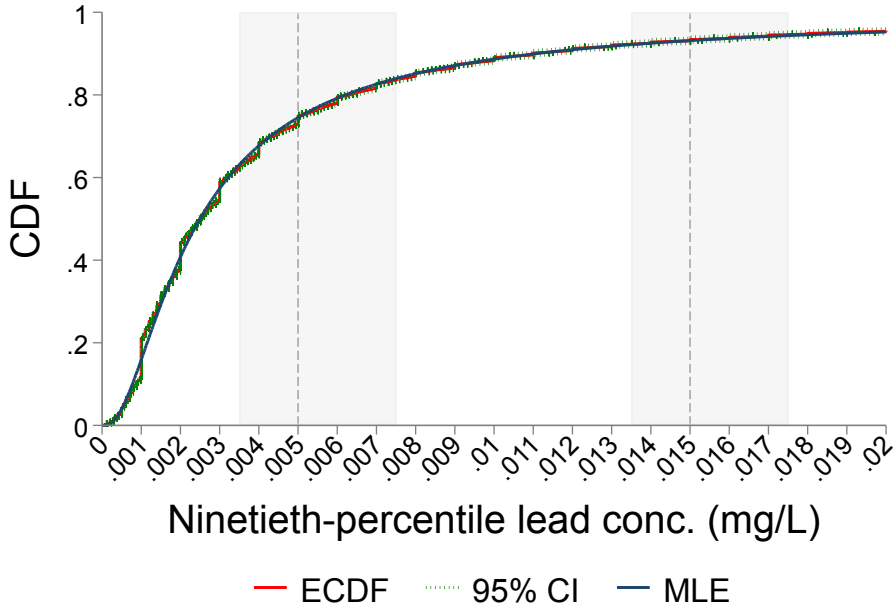


Notes: The top row of figures displays estimates of the density of ninetieth-percentile lead concentrations to the left and to the right of the action level (0.015) for large-sized PWSs pre-Flint (Panel A) and post-Flint (Panel B). The dashed lines depict the 95% confidence interval around the estimate of the density. The bottom row of figures depicts the pre-Flint (Panel A) and post-Flint (Panel B) empirical CDFs (green line), the observed proportions of rounding at the action level ($\widehat{\Pr}(X = 0.015)$), the counterfactual proportions of rounding (absent manipulation) at the action level ($\widehat{\Pr}(X(0) = 0.015)$, and estimates of the simple average excess rounding measure at the action level ($\hat{\eta}^s$). Greyed-out estimates of η^s are negative. The approach in McCrary (2008) selects bin size as follows: $\hat{b} = 2\hat{\sigma}n^{-1/2}$ where $\hat{b}$ is the bin size, $\hat{\sigma}$ is the sample standard deviation of the running variable, and n is the sample size. Using this bin size, we are not able to obtain an estimate of the discontinuity for large water systems pre-Flint. For this reason, we instead use the smallest bin size for which we are able to obtain an estimate of the discontinuity and standard error ($\hat{b} = 0.0004501$).

Figure B.9: Empirical CDF and Censored Maximum Likelihood Estimate of Counterfactual CDF - Small Systems



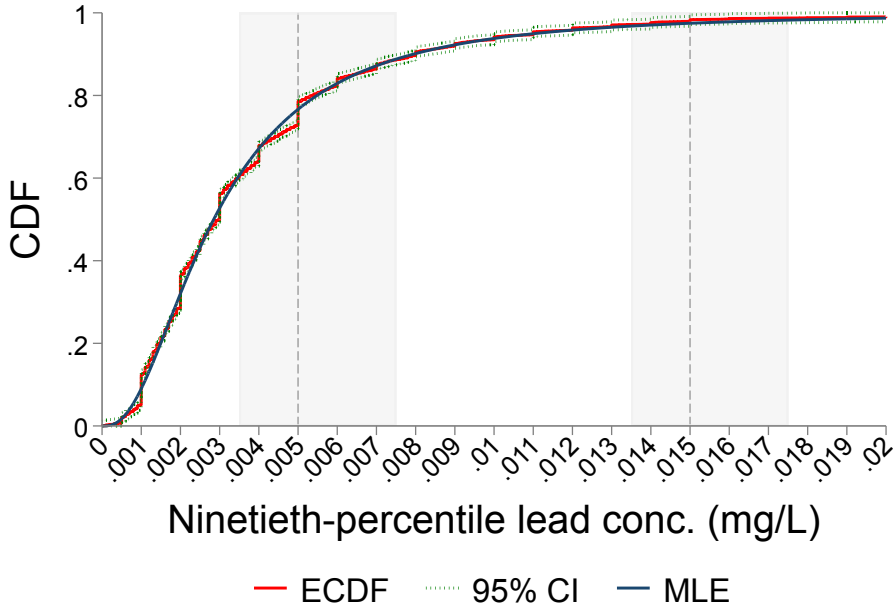
(a) Pre-Flint



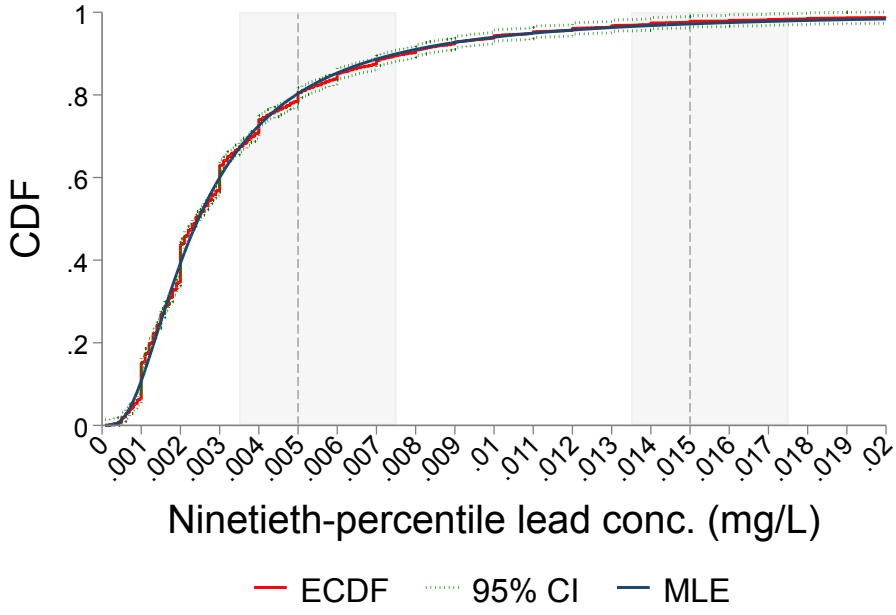
(b) Post-Flint

Notes: The figures display the empirical (red) and counterfactual (navy) CDFs (estimated via censored maximum likelihood) for small PWSs pre- and post-Flint. Uniform 95% confidence intervals around the empirical CDFs are depicted by green dotted lines. The manipulation windows are shaded in gray.

Figure B.10: Empirical CDF and Censored Maximum Likelihood Estimate of Counterfactual CDF - Medium Systems



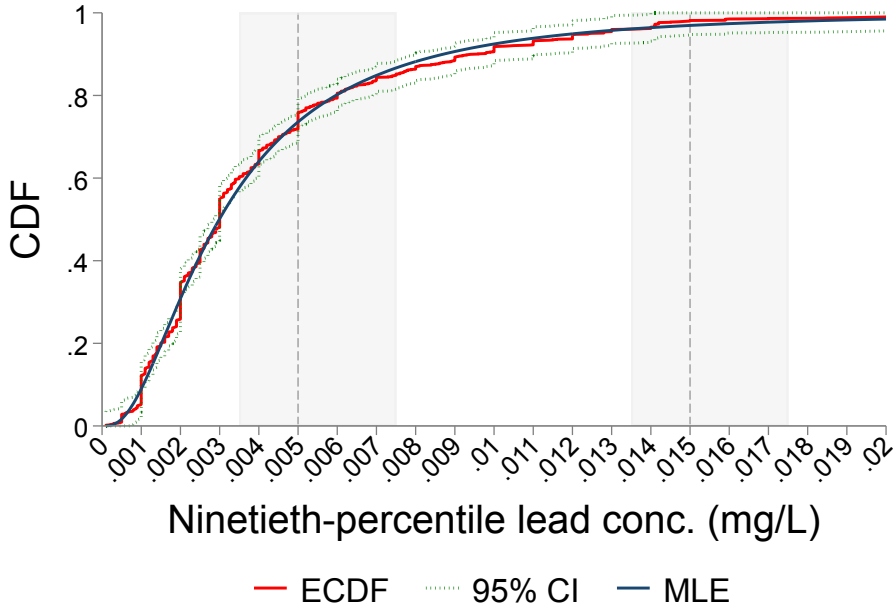
(a) Pre-Flint



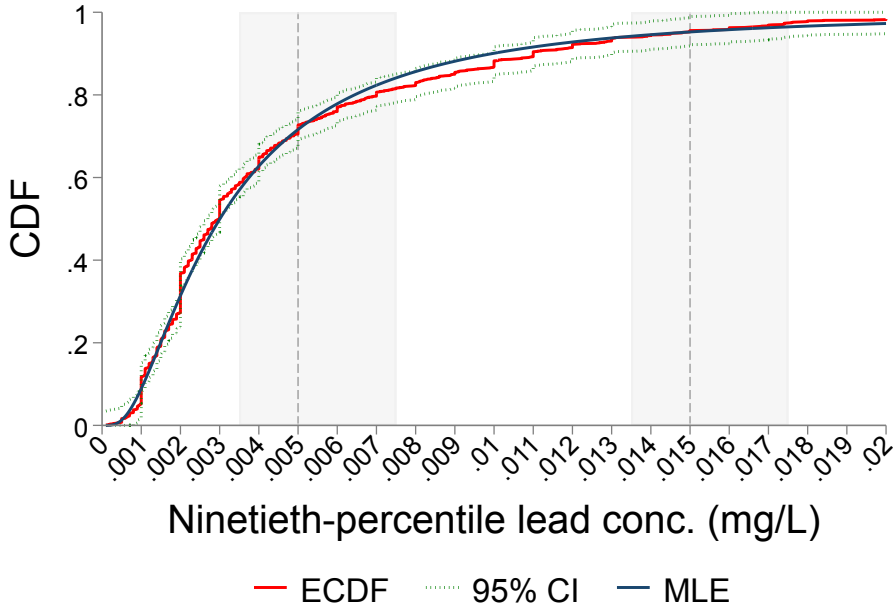
(b) Post-Flint

Notes: The figures display the empirical (red) and counterfactual (navy) CDFs (estimated via censored maximum likelihood) for medium PWSs pre- and post-Flint. Uniform 95% confidence intervals around the empirical CDFs are depicted by green dotted lines. The manipulation windows are shaded in gray.

Figure B.11: Empirical CDF and Censored Maximum Likelihood Estimate of Counterfactual CDF - Large Systems



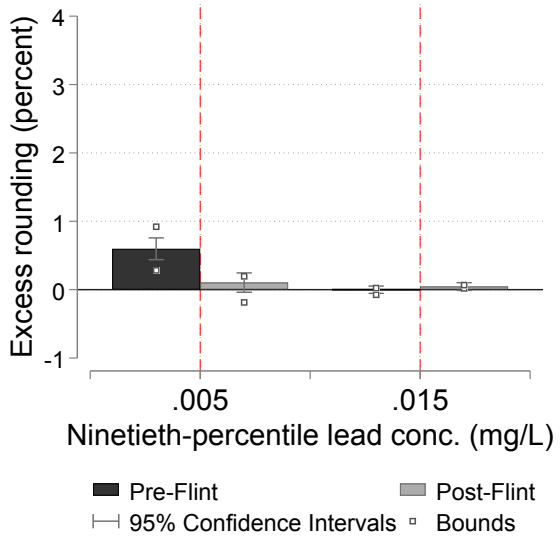
(a) Pre-Flint



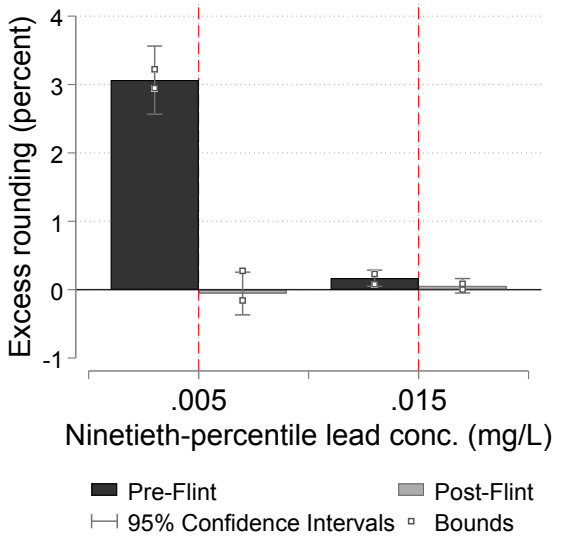
(b) Post-Flint

Notes: The figures display the empirical (red) and counterfactual (navy) CDFs (estimated via censored maximum likelihood) for large PWSs pre- and post-Flint. Uniform 95% confidence intervals around the empirical CDFs are depicted by green dotted lines. The manipulation windows are shaded in gray.

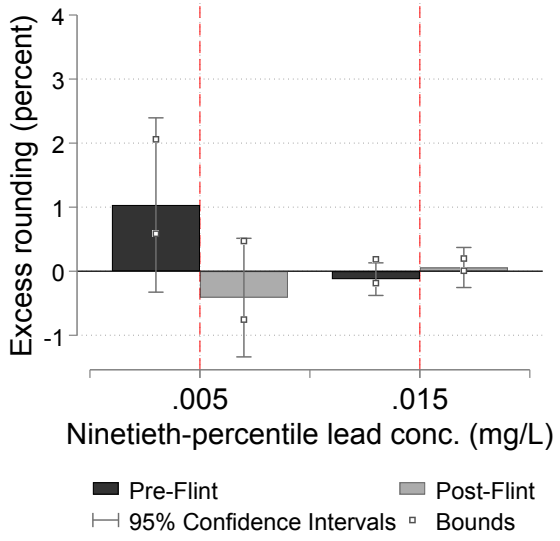
Figure B.12: Probability-Weighted Excess Rounding Measure by System Size



(a) Small Systems



(b) Medium Systems



(c) Large Systems

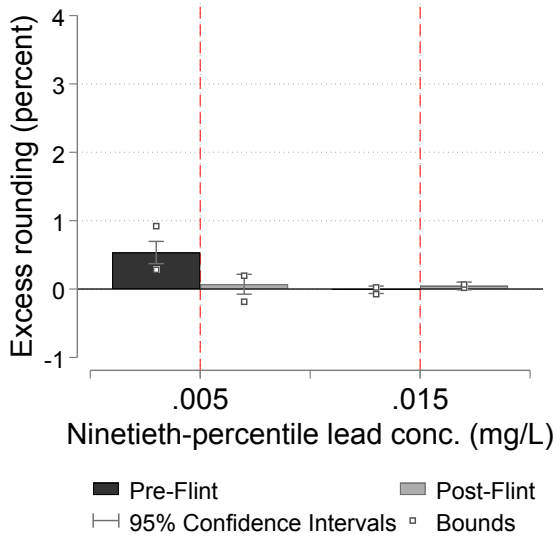
Notes: The figures display the probability-weighted average excess rounding estimates ($\hat{\eta}_j^p$) for small, medium, and large PWSs. We use manipulation windows of (0.0035, 0.0075) and (0.0135, 0.0175) around the reporting threshold and action level, respectively, to estimate the counterfactual distribution via CMLE. Figures B.9-B.11 present the empirical CDFs and the CMLE counterfactual distribution of $X^N(0)$ used to compute our test and η^p. For additional details on the specification choices in our implementation, see Supplemental Appendix B.2. Pre-Flint estimates are depicted in black, post-Flint estimates are depicted in grey, the estimated bounds from equation (18) are indicated by white, square markers, and the 95% confidence intervals for $\hat{\eta}_j^p$ are shown using grey, capped lines.

Table B.2: Probability of Rounding Conditional on Being in Roundable Interval

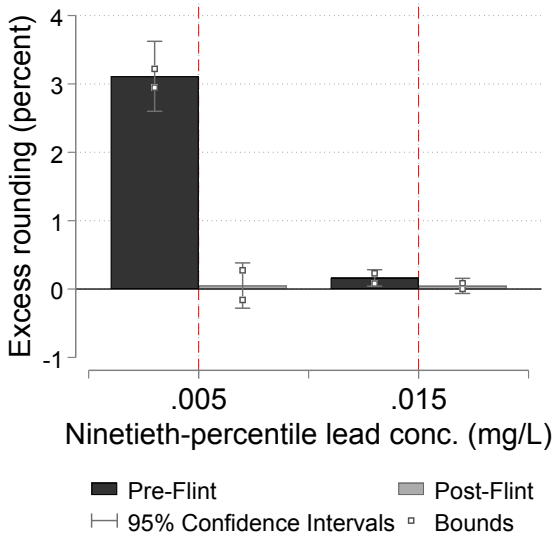
	Small PWS		Medium PWS		Large PWS	
	Pre-Flint	Post-Flint	Pre-Flint	Post-Flint	Pre-Flint	Post-Flint
Panel A: Reporting Threshold Interval ($\mathcal{R}_1$)						
$\frac{\Pr(X=0.001)}{\Pr(l_r \leq X < u_r)}$	0.452	0.369	0.374	0.359	0.384	0.360
$\frac{\Pr(X=0.002)}{\Pr(l_r \leq X < u_r)}$	0.393	0.338	0.355	0.343	0.441	0.401
$\frac{\Pr(X=0.003)}{\Pr(l_r \leq X < u_r)}$	0.363	0.337	0.352	0.369	0.352	0.301
$\frac{\Pr(X=0.008)}{\Pr(l_r \leq X < u_r)}$	0.448	0.347	0.338	0.299	0.256	0.300
$\frac{\Pr(X=0.009)}{\Pr(l_r \leq X < u_r)}$	0.464	0.405	0.343	0.304	0.350	0.235
Panel B: Action Level Interval ($\mathcal{R}_2$)						
$\frac{\Pr(X=0.010)}{\Pr(l_r \leq X < u_r)}$	0.563	0.489	0.524	0.466	0.636	0.615
$\frac{\Pr(X=0.011)}{\Pr(l_r \leq X < u_r)}$	0.610	0.476	0.676	0.506	0.773	0.571
$\frac{\Pr(X=0.012)}{\Pr(l_r \leq X < u_r)}$	0.628	0.484	0.699	0.620	0.692	0.520
$\frac{\Pr(X=0.013)}{\Pr(l_r \leq X < u_r)}$	0.588	0.492	0.615	0.594	0.529	0.458
$\frac{\Pr(X=0.018)}{\Pr(l_r \leq X < u_r)}$	0.613	0.481	0.462	0.696	0.000	0.300
$\frac{\Pr(X=0.019)}{\Pr(l_r \leq X < u_r)}$	0.606	0.450	0.850	0.500	0.333	
$\frac{\Pr(X=0.020)}{\Pr(l_r \leq X < u_r)}$	0.722	0.533	0.636	0.636	0.500	0.500
Observations	57,774	44,858	13,223	9,908	1,613	1,581

Notes: This table presents the observed probability of being rounded to a non-policy threshold conditional on being within its roundable interval. We provide this probability for all non-policy thresholds in $\mathcal{R}_1$ and $\mathcal{R}_2$ in panel A and panel B, respectively. For large PWSs post-Flint, there are zero observations in the roundable interval around 0.019.

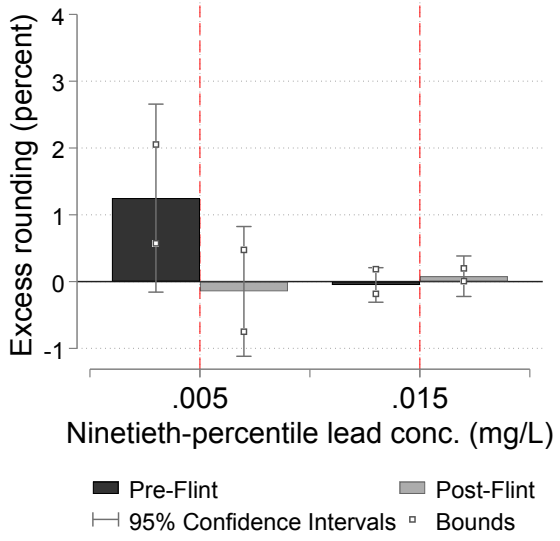
Figure B.13: Excess Rounding by System Size - Alternative Manipulation Windows - (0.0045, 0.0065) and (0.0145, 0.0165)



(a) Small Systems



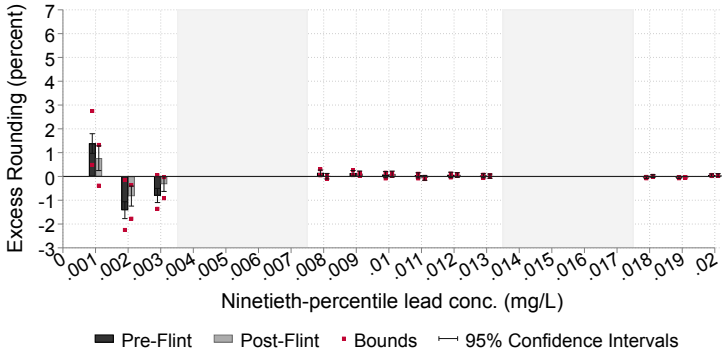
(b) Medium Systems



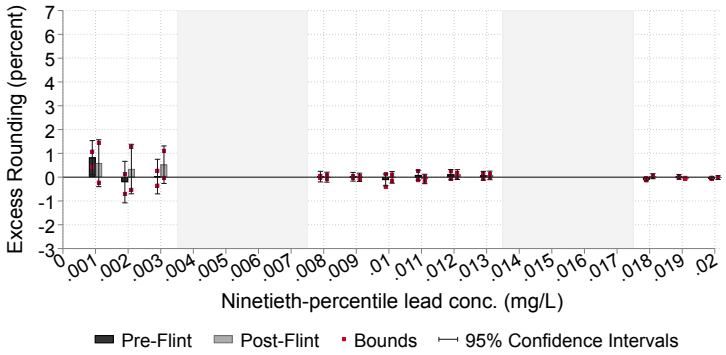
(c) Large Systems

Notes: The figures display the simple average excess rounding estimates ($\hat{\eta}_j^s$) for small, medium, and large PWSs. We use manipulation windows (0.0045, 0.0065) and (0.0145, 0.0165) around the reporting threshold and action level, respectively, to estimate the counterfactual distribution via CMLE. Pre-Flint estimates are depicted in black, post-Flint estimates are depicted in grey, the estimated bounds from equation (18) are indicated by white, square markers, and the 95% confidence intervals for $\hat{\eta}_j^s$ are shown using grey, capped lines.

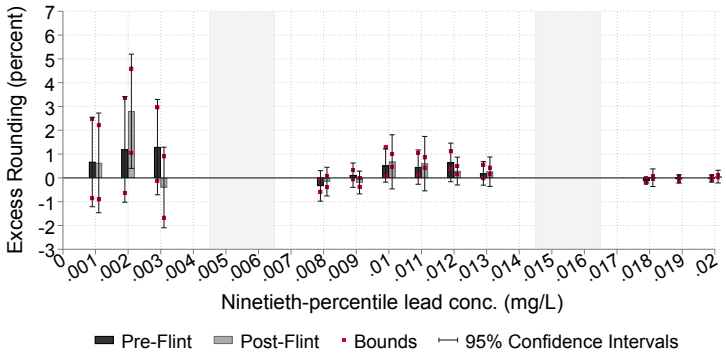
Figure B.14: Excess Rounding - Placebo Tests



(a) Small Systems



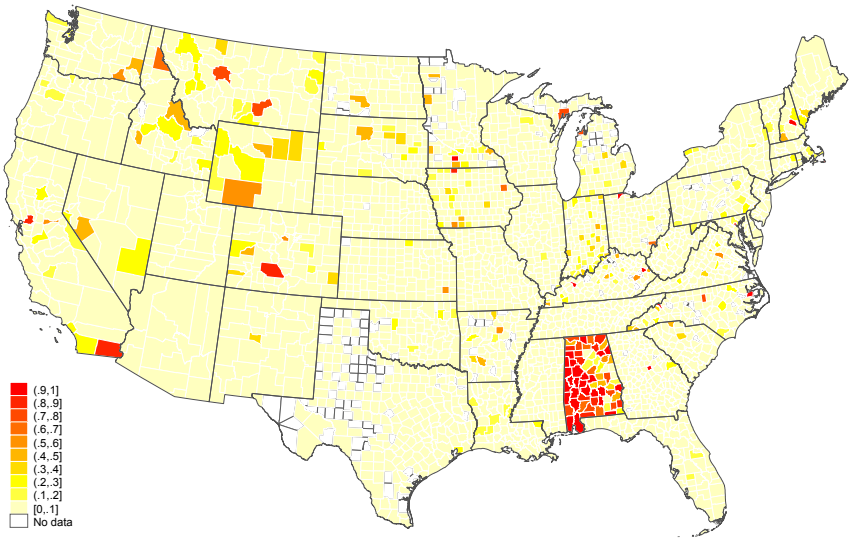
(b) Medium Systems



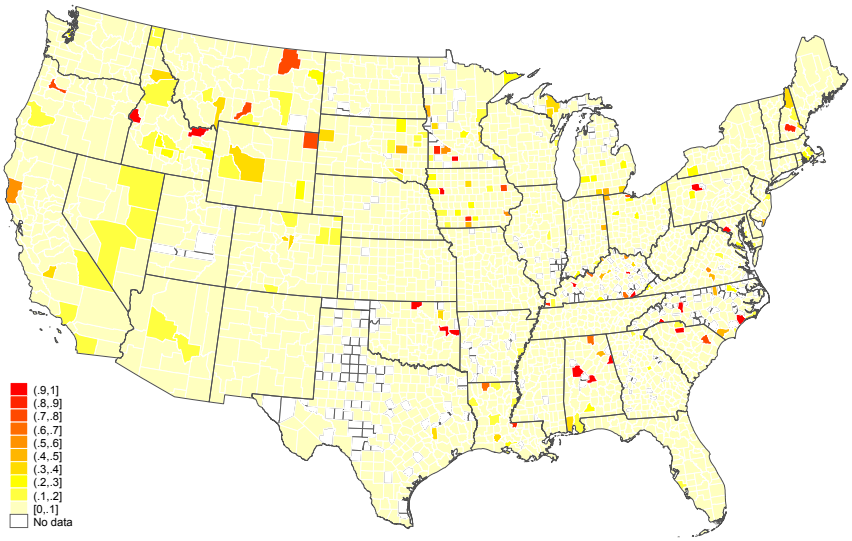
(c) Large Systems

Notes: The figures display the simple average excess rounding estimates ($\hat{\eta}_j^s$) at non-policy cutoffs for small, medium, and large PWSs. We use manipulation windows of (0.0035, 0.0075) and (0.0135, 0.0175) around the reporting threshold and action level, respectively, to estimate the counterfactual distribution via CMLE. The manipulation windows are shaded in grey. Pre-Flint estimates are in black, post-Flint estimates are in grey, the estimated bounds from equation (18) are indicated by red, square markers, and the 95% confidence intervals for $\hat{\eta}_j^s$ are shown with grey, capped lines after implementing a Bonferroni multiple hypothesis testing correction for each period and system size.

Figure B.15: Spatial Distribution of Proportion of PWSs Reporting 0.005 mg/L



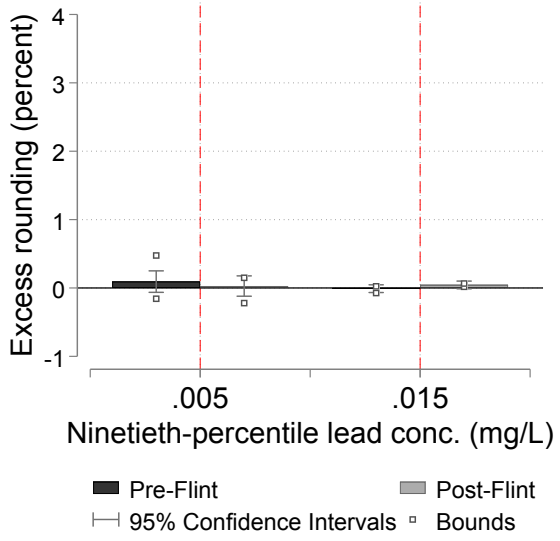
(a) Pre-Flint (2011-2015)



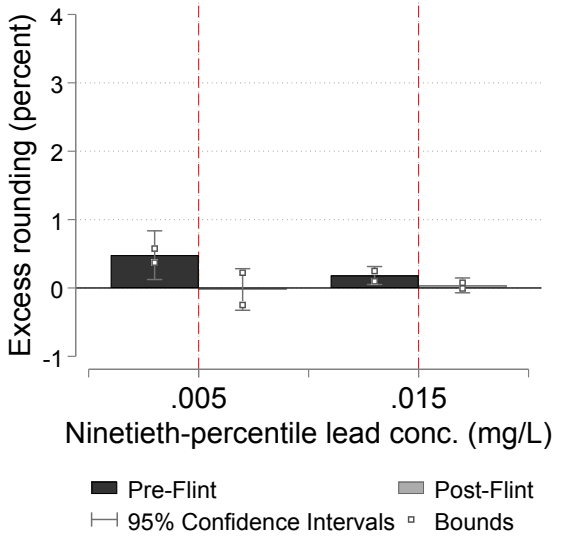
(b) Post-Flint (2016 - 2020)

Notes: This figure displays maps of the proportion of PWSs reporting ninetieth-percentile lead values equal to 0.005 mg/L within each county in the conterminous US. To calculate this, we first determine the proportion of reports equal to 0.005 for each PWS and then take the service population-weighted average of this proportion for each county. Figures B.15a and B.15b display the pre- and post-Flint maps, respectively. Darker colors correspond to a higher proportion. Counties shaded in white have no data available during these periods. We use the 2020 County Cartographic Boundary File and 2020 State Cartographic Boundary File from the Census to create the maps (US Census Bureau 2020a, 2020b). We map PWSs serving tribal areas to counties using the EPA's Tribal Boundary Service (USDOC and USEPA 2023).

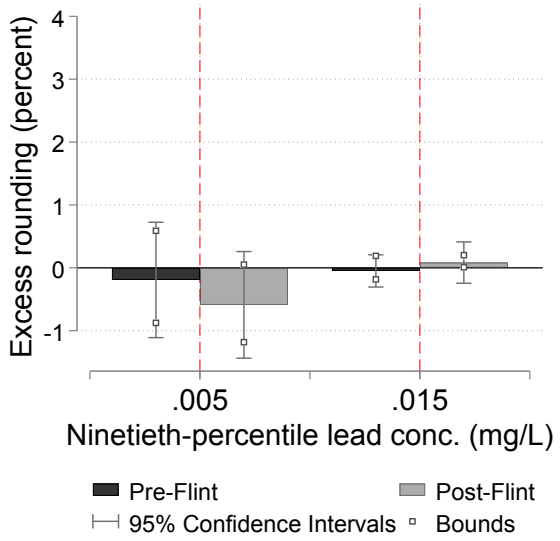
Figure B.16: Excess Rounding by System Size - Excluding Alabama



(a) Small Systems



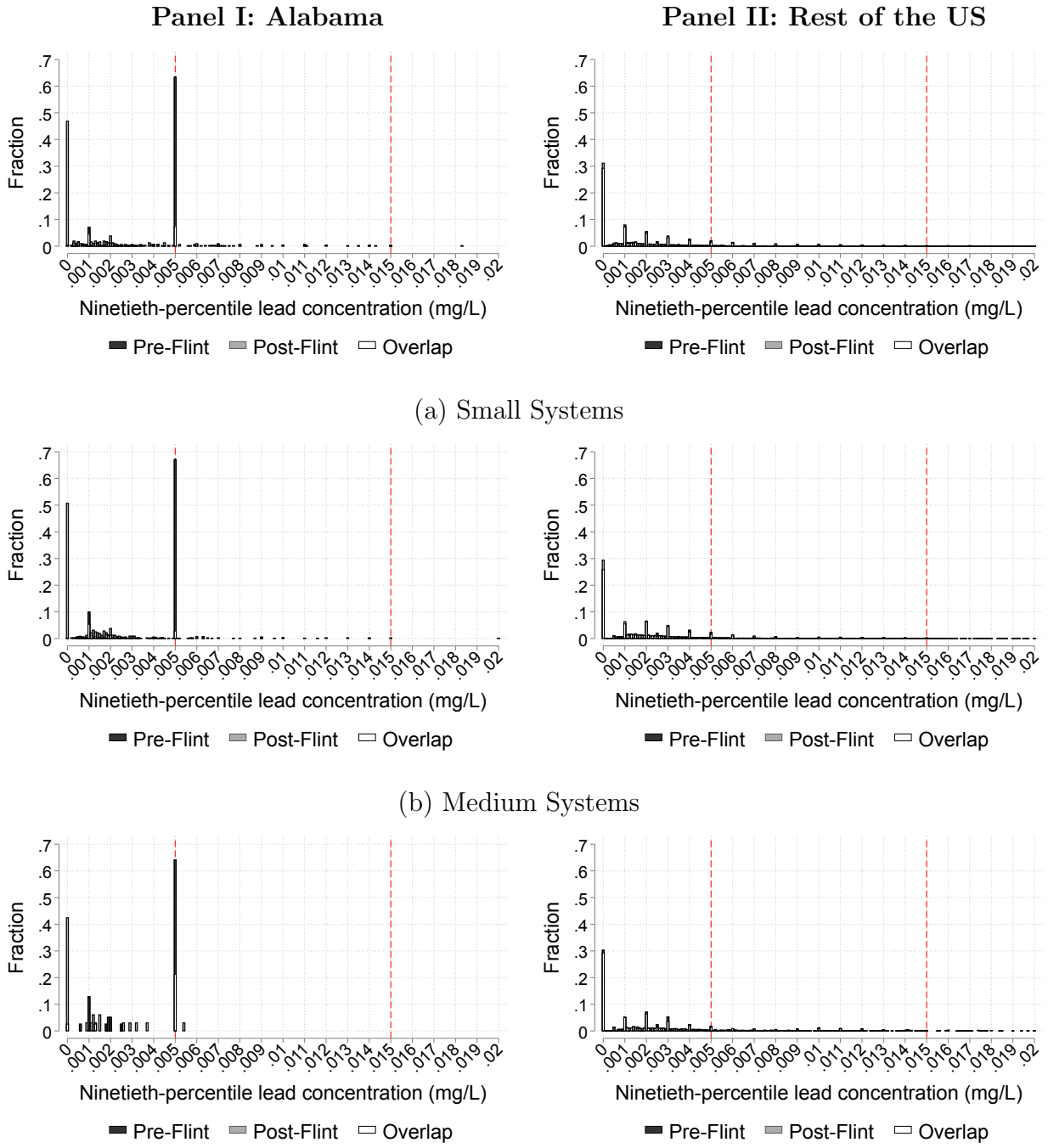
(b) Medium Systems



(c) Large Systems

Notes: The figures display the simple average excess rounding estimates ($\hat{\eta}_j^s$) for small, medium, and large PWSs, excluding systems from the state of Alabama. We use manipulation windows of (0.0035, 0.0075) and (0.0135, 0.0175) around the reporting threshold and action level, respectively, to estimate the counterfactual distribution via CMLE. Pre-Flint estimates are depicted in black, post-Flint estimates are depicted in grey, the estimated bounds from equation (18) are indicated by white, square markers, and the 95% confidence intervals for $\hat{\eta}_j^s$ are shown with grey, capped lines.

Figure B.17: Histograms of Ninetieth-Percentile Lead Values in Alabama vs. Rest of the US



(a) Small Systems

(b) Medium Systems

(c) Large Systems

Notes: The figures display histograms of reported ninetieth-percentile lead values for small, medium, and large PWSs in Alabama (Panel I) and the rest of the US (Panel II). Pre-Flint histograms are depicted in black, post-Flint histograms are depicted in grey, and white indicates the overlap of the pre- and post-Flint histograms. The dashed, red vertical lines indicate the reporting threshold (0.005) and the action level (0.015).

Supplemental Appendix References

Acosta, Roberto. 2016. *UM-Flint hosts CNN on campus for Democratic presidential debate coverage.* https://www.mlive.com/news/flint/2016/03/um-flint_hosts_cnn_on_campus_f.html.

Barnhart, David. 2019. *Flint: The Poisoning of an American City.*

Baxter, Anthony. 2021. *Flint: Who Can You Trust?*

Christensen, Peter, David A. Keiser, and Gabriel E. Lade. 2023. "Economic Effects of Environmental Crises: Evidence from Flint, Michigan." *American Economic Journal: Economic Policy* 15 (1): 196–232. https://doi.org/10.1257/pol.20190391.

Clark, Anna. 2018. *The Poisoned City: Flint's Water and the American Urban Tragedy.* Henry Holt / Company.

Grevatt, Peter C. 2016. *Clarification of Recommended Tap Sampling Procedures for Purposes of the Lead and Copper Rule.* Official memorandum to Water Division Directors, US Environmental Protection Agency, February 29. https://www.epa.gov/sites/default/files/2016-02/documents/epa_lcr_sampling_memorandum_dated_february_29_2016_508.pdf.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142 (2): 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005.

US Census Bureau. 2020a. *County Cartographic Boundary File.* Accessed May 2, 2024. https://www2.census.gov/geo/tiger/GENZ2020/shp/cb_2020_us_county_500k.zip.

———. 2020b. *State Cartographic Boundary File.* Accessed May 2, 2024. https://www2.census.gov/geo/tiger/GENZ2020/shp/cb_2020_us_state_500k.zip.

———. 2021. *Annual County Resident Population Estimates by Age, Sex, Race, and Hispanic Origin: April 1, 2010 to July 1, 2020.* Accessed November 12, 2021. https://www2.census.gov/programs-surveys/popest/datasets/2010-2020/counties/asrh/CC-EST2020-ALLDATA.csv.

USDOC and USEPA. 2023. *American Indian Reservations.* Accessed April 29, 2024. https://epa.hub.arcgis.com/datasets/EPA::epa-american-indian-reservations-air/about.

USEPA. 1991. "Maximum Contaminant Level Goals and National Primary Drinking Water Regulations for Lead and Copper." *Federal Register* 56 (110): 26460–26564. https://www.govinfo.gov/app/details/FR-1991-06-07.

———. 2021. "National Primary Drinking Water Regulations: Lead and Copper Rule Revisions." *Federal Register* 86 (10): 4198–4312. https://www.govinfo.gov/content/pkg/FR-2021-01-15/pdf/2020-28691.pdf.