Lead Pipe Information and Housing Prices: An Analysis in Washington, DC

James Jiang

October 2022

Get the latest version here

Abstract

In 2016, the Water and Sewer Authority of Washington, DC released an online map that contains information on lead service lines (LSLs) for all properties in the district. Using the release as a natural experiment, this paper estimates the effect of the new information on prices of properties with and without LSLs. Recent literature has found that housing lead reduction policies such as remediation mandates have significant price effects. In DC, while the map's release was followed by a marked increase in requests for water lead tests, neither a difference-in-differences model nor a repeat sales model captures a significant divergence between housing prices of the two types of properties after the release, implying the housing market response to the information was limited.

1 Introduction

Ever since Rosen's (1974) seminal work, the hedonic model has been widely used in a number of fields in economics. Applied in the housing literature, this approach considers housing as a differentiated good composed of multiple attributes, and the equilibrium price schedule in the competitive market can be used to derive the marginal value of a given attribute. Numerous studies have been devoted to measure the value of a wide range of attributes including non-market, "environmental amenities", such as school quality (Black, 1999), air quality (Chay & Greenstone, 2005), and local crime risk (Linden & Rockoff, 2008). This paper adopts the framework and focuses on the presence (or lack thereof) of lead hazards in housing, in particular lead service lines (LSLs), more commonly known as lead pipes.

1

LSLs may be of particular interest to homebuyers because of the demonstrated adverse health effects of exposure to the metal lead, with children being the most susceptible group. Links have been established between children's lead exposure to negative impacts on various health and other outcomes, including cognitive function (Needleman & Gatsonis, 1990), development delays (Selevan et al., 2003), violent and criminal behavior (Reyes, 2007), ADHD (Goodlad, 2013), and academic achievement (Aizer et al., 2018). Importantly, many such effects seem to be long-term and persist well into adulthood (Reyes, 2007; Grönqvist et al., 2020). During the past several decades, as the scientific consensus evolved, the acceptable upper limit (later renamed the "reference level") of blood lead concentration in children has been lowered six times, from 60 micrograms per 100 milliliters of blood (μg/dl) in the early 1960s to 3.5 μg/dl since 2021 (American Academy of Pediatrics Subcommittee on Accidental Poisoning, 1961; Pueschel et al., 1996; Centers for Disease Control and Prevention, 2021). Furthermore, while the thresholds are those that "should prompt public health actions", it is commonly acknowledged that there is no known safe level for blood lead in children (Binns et al., 2007).

Few attempts have been made to actually estimate parents' willingness to pay for reducing their children's exposure to lead hazard. In one of these, under fairly restrictive assumptions, Agee and Crocker (1996) use data on enrollment in chelation therapy and estimate parents were willing to pay between $39.01 and $364.23 in 2022 dollars for a one-percent reduction in their children's lead burden, which was significantly higher than the U.S. Environmental Protection Agency's estimate of savings from lead reduction efforts. However, given the overwhelming consensus of the health risk posed to children by even low levels of lead exposure, it is conceivable that in the hedonic framework, homebuyers, especially parents, will place a positive valuation on the absence of potential lead hazards in housing, including lead paint and LSLs. In this paper, I use data on housing prices and presence of LSLs in Washington, DC to examine the extent to which lead hazards affect housing prices.

In the early 2000s, the District of Columbia experienced a public health crisis due to lead contamination of drinking water, an unintended consequence of a change in the District's water disinfection procedure and resulting water pipe corrosion. The episode saw a notable rise in blood lead levels of affected children (Edwards et al., 2009; Jane Brown et al., 2010). The crisis attracted nationwide attention, led to a congressional investigation, and highlighted the potential health hazards from the prevalent use of LSLs in the district. LSLs were widely used in the Unites States before they were banned in new constructions after 1986, and may are still in service in older buildings, including in Washington DC. Since the lead contamination crisis, the District's government has been working on tackling the issue through replacement and awareness campaigns. In June 2016, as part of the effort, the

District of Columbia Water and Sewer Authority published a lead pipe map on their website that allows anyone to identify the types of water pipes used in the service lines for any address in the District. With the explicit goals of increasing transparency, informing residents about LSLs and eliciting them to help the water authority identify and solve related problems, this easy-to-use visual tool represented a massive upgrade from the previously undigitized records of pipe material, and made such data readily accessible to the public for the first time. The release of the map also coincided with a significant and sustained rise in the American public's interest in the issue of lead contamination of drinking water following the initial national media coverage in early 2016 of the water crisis in Flint, Michigan, and the availability of the lead map has made it possible for any DC resident to obtain information related to LSLs if they wish, enabling them to better inform their housing decisions.

Using the release of the lead map as a natural experiment, I try to estimate the resulting information effect on housing prices due to such negative valuation associated with LSLs. I use two different model specifications for identification: a straightforward difference-in-differences model, and a repeat sales model. Based on the recorded service line material, all housing units are divided into three groups: those with LSLs, those without, and those for which the Water and Sewer Authority does not have information on service line material. The grouping is done multiple times, using information on both the public and private sides of the service line. I set the date of the lead map's release, June 6, 2016, as the cutoff for the pre- and post-information-shock periods, and focus on a two-year window centered on that date for the diff-in-diff specification, and a seven-year window for the repeat sales model. Neither method points to a discernible information effect on housing prices.

This paper joins a number of other studies that adopt the hedonic model to estimate the information effect on housing prices related to disclosure of negative environmental amenities. Negative price effects have been found from disclosure of such unpleasantries as airport noise (Pope, 2008), proximity to polluting firms (Mastromonaco, 2015), health hazards from nearby waste sites (Gayer et al., 2002), and flood risk (Votsis & Perrels, 2016). Each of those studies also utilizes disclosure (usually originating from new policy) of information that was previously known only to an limited extent, similar to the release of the lead map. In addition, this paper also contributes to the existing literature on the effectiveness of lead reduction policies as well as their effects on the housing market. Gazze (2021) makes use of data from multiple states and finds that lead hazard mitigation mandates lead to a decrease in prices of old houses, where such hazards are more likely to be present. Theising (2019) finds that a mandate to replace private LSLs in Madison, Wisconsin had a large, positive price effect on post-replacement units that exceeds the cost of replacement, implying homebuyers place high value on the absence of LSLs in addition to the explicit cost saving under the

3

mandate. Both papers point out the potentially important information effect of disclosure. Billings & Schnepel (2017) find similar returns to lead paint remediation from a voluntary program in Charlotte, North Carolina. My study is most similar to Theising (2019) in that I use building-level service line data instead of resorting to a proxy (such as year of construction) for likely presence of LSLs, enabling me to exploit variation between otherwise similar properties that would have been buried in less granular data; but these papers differ in an important manner because I am able to examine the information effect in isolation by studying a city without a mitigation mandate. As far as I am aware of, this paper is the first to examine information effects on housing prices in the context of LSLs.

The rest of the paper is organized as follows. Section 2 provides a brief introduction of the use of LSLs in the United States and the attendant health hazards, along with the background of DC Water's release of the lead map. Section 3 describes the two empirical specifications I use. Section 4 summarizes the data. Section 5 presents the results, and Section 6 discusses possible interpretations. Section 7 concludes.

2 Background

LSLs and Water-Borne Lead

The metal lead has a long history of being used as a piping material for water supply and distribution, dating back to Roman times (Hodge, 1981). Its wide adoption in United States started in mid to late 1800s and was most notable in large cities, appearing in 85 percent of the biggest American cities by 1897 (Troesken, 2006). Compared to alternative piping materials such as steel, iron and cement, lead is both more malleable and more durable, making it ideal for municipal water systems from an engineering perspective (Clay et al., 2006). Around the turn of the twentieth century, however, there were growing public concerns about the potential danger of lead poisoning posed by LSLs, and many cities started placing restrictions on their use by the 1920s (Rabin, 2008). Nonetheless, the relentless and carefully orchestrated lobbying from the lead industry successfully slowed down the abandonment of LSLs and ensured their common and continual use for supplying water to homes and buildings in the United States (Rabin, 2008). It was not until 1986 that their installation was finally banned a set of amendments to the Safe Drinking Water Act. The Act now prohibits the use of LSLs in new constructions, but existing LSLs remain in use in many buildings built before 1986 in large cities. A 2016 survey estimated that 15 to 22 million people (out of a total of 297 million) served by community water systems in the United States had a LSL serving their home (Cornwell et al., 2016).

There is usually little or no lead in water from raw sources or water treatment plants; however, lead in service lines can leach into tap water through the corrosive chemical process that occurs between water and the pipe material, and corrosion control is essential in preventing such contamination (Triantafyllidou & Edwards, 2012). In 1991, the United States Environmental Protection Agency issued the Lead and Copper Rule (LCR), a regulation intended to oversee effective corrosion control measures. The LCR sets an upper limit (of 15 parts per billion) for the concentration of lead in tap water; in the event that the limit is breached, it requires public water utilities to take actions to control plumbing corrosion, inform the public and, if necessary, replace the related LSLs. Despite the majority of public water utilities being in compliance with the LCR,[1] drinking water is still an important source of environmental lead exposure (Triantafyllidou & Edwards, 2012; Brown & Margolis, 2012), and LSLs remain the biggest source of lead of drinking water, accounting for 50%–75% of the metal by mass (Sandvig et al., 2008).

The 21st century has seen two major public health crises in the United States related to lead contamination of drinking water, both of which were results of the release of lead in service lines into tap water. The Washington DC water crisis from 2001 to 2004 was triggered by the water authority's switch of the type of disinfectant, from free chlorine to chloramine. While the former had been used by the water industry for more than a century, not until after the crisis did researchers discover that it had the side benefit of reducing lead solubility in water, and the change in disinfectant effectively increased water corrosivity (Edwards et al., 2009). The Flint, MI water crisis from 2014 to 2016 was the joint work of two factors: the switch to a temporary new water source with a different chemical makeup, and the interruption of corrosion control treatment (Roy & Edwards, 2019). While the scale and severity of both incidents were direct results of government misconduct and oversight, they also highlighted the public health risk in the use of LSLs, especially as the science on related safety standards and acceptable practice seems to be continuously evolving.

The DC Water Lead Map

On June 6, 2016, the District of Columbia Water and Sewer Authority (branded as DC Water) launched an online interactive map tool that allows users to identify the materials of the water service lines for any of the over 120,000 properties in the district and, in particular, find out if those service lines contain lead. The map includes a circular marker for every address, divided in two halves, with colors denoting the materials of the public and private service lines, respectively: green means the line does not contain lead; gray means it does;

---

[1]It has been pointed out that the current sampling protocol under the LCR can fail to detect lead levels in breach of the upper limit and lead to a false conclusion of compliance (Del Toral et al., 2013).

Figure 1: A sample screenshot of DC Water's lead map

white means there is no information for the line. Figure 1 is a sample screenshot of part of the map. Users can either view the overall map or enter an address to search for a particular property, much like they can when using a smartphone map application.

The release of the map meant that DC residents were able to get information on the material of their water service lines for the first time, as the determination of pipe material usually requires excavation and cannot be easily done by individuals. It was also a big leap from the district's previous record-keeping of service line materials, when such information only existed on a variety of physical records in a haphazard fashion. DC Water publicized the map in multiple ways, including via email to users and through social media campaigns (on Facebook and Twitter). The announcement was also covered in a handful of national and local online news outlets including as Vox, DCist (the website of NPR's local station), and Fox 5 DC, as well as popular DC local news blogs including *Petworth News* and the *Georgetown Metropolitan*. While it is hard to gauge the extent to which DC residents were aware of the service immediately following its release, there is limited evidence of some take-up: according to Google Trends, the term "DC lead pipe map" started to be searched the week following the announcement, although the number of queries has been so small for this search term that the plot exhibits a pattern of occasional spikes along a horizontal line at zero, instead of a smooth trend line.[2] On the other hand, the next time the search term

---

[2]Other similar terms directly related to the map, such as "DC lead pipes", "DC lead map" and "DC Water

6

Figure 2: Google Trends results for "DC lead pipe map", 2016–2021

showed up in the data was in March and April 2017 as another two separate spikes. It was not until June 2019 that there started to be more consistent interest in the term, up to today; the reasons for such newfound interest are unclear to me. These are shown in Figure 2. It is worth noting that Google search volume does not directly reflect actual usage of the service.

3 Empirical Methodology

The Hedonic Price Function

My analysis adopts two parallel empirical strategies, both of which builds upon the hedonic model pioneered by Rosen (1974). The model describes housing as a vector of various utility-generating attributes, which can be structural, environmental, neighborhood, etc. Market equilibrium prices of housing are determined by the interactions of consumers and producers, and are a function of those attributes. It can be shown that the partial derivative of price with respect to a given attribute reveals the consumer's marginal willingness to pay for the attribute. In the context of this analysis, assuming a log-linear form, the hedonic price function can be written as

$$\ln P_{it} = \alpha_0 + \beta_0 \times lead_i + X_{it}'\theta_0 + \epsilon_{it}, \tag{1}$$

where P_{it} is the price of housing unit i traded at time t, $lead_i$ is a dummy variable that equals 1 if unit i has a LSL, and X_{it} is a vector that contains other housing attributes; the coefficient β_0 represents consumers' valuation of the presence of LSLs in housing.

Assuming buyers have perfect information about lead pipes, a direct estimation of the price coefficient of LSLs, β_0, using (1) would not be reliable: because it is impossible to include all relevant housing attributes in X_{it} in the empirical analysis, coefficient estimates would be biased due to omitted variables. Hence I explore two alternative identification

lead map", did not produce search data above the threshold for Google Trends to display them.

strategies that address the issue of bias, using the lead map's release as a natural experiment and estimating the interactive price effect of LSLs and the revelation of new information.

Specification #1: Difference-in-Differences

In the first strategy, I employ a simple difference-in-differences approach to identify the information shock's effect on house prices. I estimate the following regression:

$$\ln P_{it} = \alpha + \beta \times lead_i + \gamma \times map_t + \eta \times lead_i \times map_t + Z'_{it}\theta + \epsilon_{it}. \tag{2}$$

Here map_t is a dummy that equals 1 if the time of the transaction, t, was after the release of the lead map, and the control vector Z_{it} includes neighborhood, month of the year and use code fixed effects. The coefficient of interest is η, which captures the change in the effect of presence of LSLs on house prices before and after the map's release, i.e. the information effect. The diff-in-diff design assumes that the trends in sale prices of LSL and non-LSL housing units were parallel before the release of the lead map. Although this might sound restrictive given the high correlation of building age and presence of LSLs, I will demonstrate graphically in Section 5 that the parallel-trend assumption indeed seem to be the case.

Specification #2: The Repeat Sales Model

The second empirical strategy uses the repeat sales model first proposed by Palmquist (1982). With this approach, the price of housing unit i at time t is assumed to depend on a time-varying housing price index (B_t) that is true but unknown, the age of the house (A_{it}), a composite measure of time-invariant housing attributes (Z_i) and time-varying environmental variables, in this case the interaction of $lead_i$ and map_t, in the form of

$$\ln P_{it} = \ln B_t + \ln Z_i - \delta A_{it} + \eta \times lead_i \times map_t + \epsilon_{it}, \tag{3}$$

where δ is the coefficient of depreciation. If the same housing unit is traded twice at t and t' respectively, taking the difference between equation (3) at the both points in time will lead to

$$\ln \frac{P_{it}}{P_{it'}} = (\ln B_t - \delta A_{it}) - (\ln B_{t'} - \delta A_{it'}) + \eta \times lead_i \times (map_t - map_{t'}) + (\epsilon_{it} - \epsilon_{it'}). \tag{4}$$

Here, the coefficient of interest, η, can be estimated by regressing the log price ratio on the interaction term and a set of year dummies that equal 1 for year t, -1 for year t' and 0 for other years. This estimation does not yield reliable estimates for the log price indices B_t and

$B_{t'}$ because the change in the house's age $(A_{it'} - A_{it})$ is collinear with the two year dummies, but that does not affect the unbiasedness of the estimate for η; adjustments would have been necessary only if we were interested in the indices themselves. The main identifying assumption here is that other housing characteristics remain constant between sales. This can be a strong assumption, and is more likely to be relatively realistic when a shorter time span is studied. Still, I maintain this assumption due to limitations of my data sets, which, among other things, do not include information on renovations.

The repeat sales method focuses on housing units that are traded at least twice. When the number of transactions is n, we are able to derive $(n-1)$ independent equations in the form of (4). One complication here is that when $n \geq 3$ for any housing unit, the error covariance matrix is no longer diagonal because of error correlation between sales of the same unit: given three sales at time t, t' and t'', the error terms from the two resulting equations are $\epsilon_{it} - \epsilon_{it'}$ and $\epsilon_{it} - \epsilon_{it''}$, and $\text{cov}\left(\epsilon_{it} - \epsilon_{it'}, \epsilon_{it} - \epsilon_{it''}\right) = \text{var}\left(\epsilon_{it}\right) = \sigma^2$. This calls for a generalized least squares estimation, and I will return to the topic when presenting my results.

As the discussion above makes clear, the two specifications used here are similar in spirit but use different identifying assumptions. By presenting results from both models, I hope they will complement each other and help paint a more robust picture about the map's effect on housing prices.

4   Data

The data I use comprises of public records from two sources. Firstly, records of real property transactions were obtained from a public database from the Real Property Tax Administration of the DC Office of Tax and Revenue, and contain such information as premise address, property use code (which details the purpose of use and belongs to broad categories like residential, commercial and office), sale price, sale date, current-year (i.e. 2022) assessment, etc. Secondly, information about materials of water service lines is obtained by scraping DC Water's lead map website, and describes the types of pipes used for both public and private service lines for each property, along with succinct descriptions on the method of determination. For unknown reasons, the latter data set contains a small number of duplicate entries with conflicting service line material records; I dealt with these manually by keeping only the apparently correct, up-to-date information. Both data sets are then merged based on the building address, joining transaction records to pipe information. In the transactions data set, "premise address" contains both the building address and a unit number for housing

Percentage of units/transactions with...	Public line		Private line	
	LSL	non-LSL	LSL	non-LSL
among: All DC units	6.15%	82.50%	13.43%	74.43%
Diff-in-diff sample	5.17%	84.44%	12.72%	75.79%
Repeat-sales sample	5.49%	83.70%	13.42%	74.46%

Table 1: Prevalence of LSLs in the samples vs. all of DC

units in multi-unit premises such as apartment buildings; in order to perform the merge by address, the premise address field is parsed to drop the unit number, so that both data sets contain comparable address fields. As a result, transactions of different units in the same building are matched to the same, unique service line record and remain distinct observations in the merged data set.

The sample selection processes of the two empirical strategies are different in some aspects to account for the different priorities and requirements of each approach. What is common between the two samples is that they both includes only residential property transactions[3] that were carried out at with a non-zero price and categorized as a market sale (as opposed to a transfer of ownership for other reasons such as gifting and divorce).

For the diff-in-diff sample, I include all transactions in a two-year window centered on the release date of the lead map, June 6, 2016. Transactions with extremely high prices (namely, over $20 million; for comparison, the 95th percentile of sale prices is $1.33 million) were dropped; those accounted for approximately 1.16% of the sample. The final sample contains 12,712 observations. One notable feature of the sample is the under-(over-)representation of units with (without) LSLs, on both the public and private sides. Some statistics are summarized in table 1.[4] This could be due to a number of reasons: for example, it may reflect lower-than-average transaction volumes of LSL units; alternatively, it may be a consequence of higher prevalence of LSLs in non-residential buildings, which are present in the full lead-map data set (but cannot be dropped because the lead map does not assign use codes to buildings) but not in my sample.

For the repeat-sales sample, I include all properties that were traded more than once in a seven-year window centered on the map's release date. In total, there were 9,713 such transactions of 4,743 properties: 4,521 were traded twice, 217 thrice, and 5 four times. After taking differences across transactions for the same unit, this yields 4,970 equations, which is the number of observations for the regression analysis. As shown table 1, the proportions of

---

[3]The use codes present in the sample are 001, 011, 012, 013, 016, 017, and 021 through 029. Details about the designation of use code are available at https://otr.cfo.dc.gov/sites/default/files/dc/sites/otr/publication/attachments/Use%20codes.pdf.

[4]The percentages for LSL and non-LSL units do not add up to 100% because the data include units for which no service line material information is available. I code these units separately.

transactions of LSL and non-LSL units are similar to those in the diff-in-diff sample, with non-LSL units overrepresented compared to the population of buildings. Among the twice-traded units, the median gap between the two sales is 1,197 days (3.28 years), and the mean gap is 1,205 days.

One issue with the lead map is that the data at any point only contains the most up-to-date records of the service lines, and DC Water does not release information on changes of service line material. As a result, the scraped data I use, obtained in March 2021, does not reflect the most accurate picture during the relevant period when housing transactions were made. I will present possible remedies for the issue in my later analysis.

5   Results and Discussion

Difference-in-Differences

My analysis does not seem to suggest there exists an information effect of the lead map's release on the real property market in Washington DC. I first present some visualized results in the diff-in-diff spirit. Figure 3 shows the weekly average log transaction price for each week starting from 10 weeks before the release till 10 weeks after for different groups of housing units, controlling for neighborhood and property use code. Although the figure focuses on the 21-week window, the estimates and confidence intervals are derived from the full sample for statistical power. Panel (a) groups observations based on the pipe material of the public service line, creating three groups of housing units: those with LSLs, those without, and those for which information is missing. Panel (b) groups the observations based on the private service line material instead. Panels (c) through (f) are created by replicating the first two panels but bunching the missing-information group with either the LSL or the non-LSL group, creating two pairs of comparison: the former way of grouping compares units that *may* have LSLs (red) with those that definitely don't (blue); the latter compares those that definitely have LSLs (red) with those that *may not* (blue). Such regrouping is done in an attempt to narrow the confidence intervals. Finally, panels (g) and (h) group observations taking into account both the public and private sides: (g) is the counterpart of (c) and (d), where units in one group *may* have LSLs on at least one side, and those in the other group do not have LSLs on either; (h) is the counterpart of (e) and (f), where units in one group definitely have LSLs on at least one side, and those in the other are not known to have LSLs for sure on either side.

First of all, the figures show that the parallel trend assumption needed for a diff-in-diff analysis is satisfied, as there was no significant difference between the price trends of different
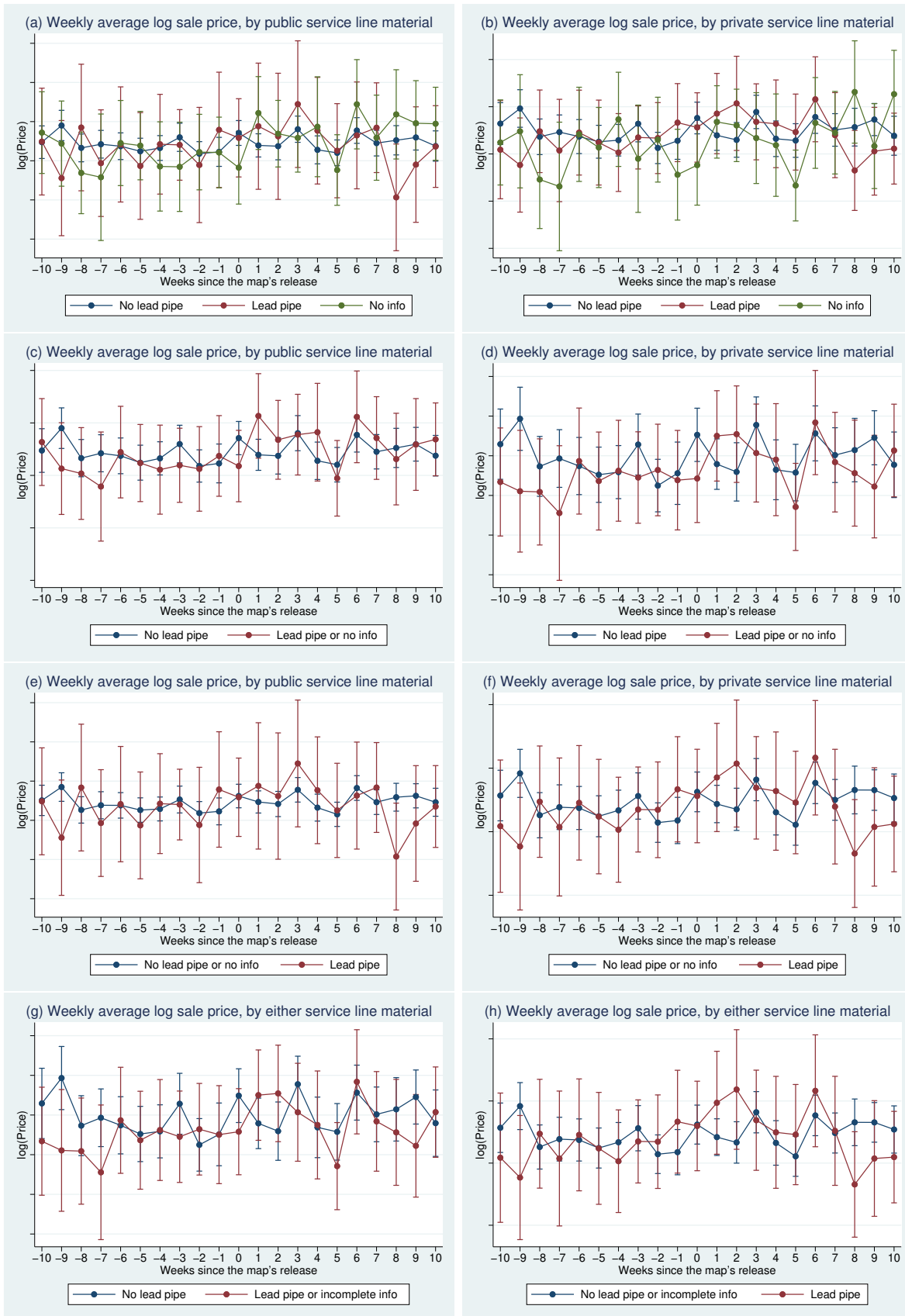
Figure 3: Weekly average log sale price by service line material

Variable	Specification			
	(a)	(b)	(c)	(d)
map_t	.069***	.072***	.072***	.072***
	(.007)	(.008)	(.008)	(.007)
$lead_i$	.020	.006	−.021**	.013
	(.018)	(.013)	(.010)	(.013)
$lead_i \times map_t$	−.026	−.020	−.011	−.020
	(.025)	(.017)	(.015)	(.017)
Service line included	Public	Private	Public & private	Public & private
Missing-information units	Separate group	Separate group	Together with LSL units	Together with non-LSL units

$$n = 12,719$$

Table 2: Estimation results for the main diff-in-diff specification

groups before the release of the lead map. But they also suggest that there was no divergence in the trends after the release. These figures, while controlling for neighborhood and use code, do not look qualitatively different from those plotted with simple averages by group. One takeaway from the figures is that the relatively small number of transactions conducted each week (about 100 on average, of which about 10 or fewer were for LSL units depending on the definition) probably resulted in the high fluctuations, lack of clear trends as well as wide confidence intervals for the average sale price estimates. Using cruder time measures such as fortnights, months and quarters does not appear to alleviate such problems to an extent enough for different conclusions to be drawn.

Now I present my estimates for η, the coefficient of interest in the diff-in-diff regression (2), in columns (a) through (d) of table 2, along with robust standard errors. I similarly estimate the regression multiple times, changing between the public and private lines and different grouping methods. Columns (a) and (b) correspond to panels (a) and (b) of figure 3, respectively, grouping housing units based on either public or private service line material and including an extra dummy variable to indicate units with missing information; columns (c) and (d) correspond to panels (g) and (h), combining information on the public and private sides and grouping missing-information units with LSL and non-LSL units respectively. While all the estimates of γ all have the expected negative sign, none of the specifications produce a statistically significant estimate.

In order to check if the insignificant estimate conceals any heterogeneity across different types of homebuyers and housing units, I perform two additional exercises. First, I consider a triple-difference model, and expand the sample to include commercial- and office-use

Variable	Specification			
	(a)	(b)	(c)	(d)
map_t	.286**	.236**	.072***	.057***
	(.103)	(.107)	(.011)	(.012)
$lead_i$	−.042	−.376*	−.000	−.002
	(.248)	(.200)	(.036)	(.032)
$lead_i \times map_t$	.031	.166	−.077	−.021
	(.305)	(.264)	(.047)	(.043)
$lead_i \times map_t \times res_i$	−.057	−.185	—	—
	(.305)	(.265)		
$lead_i \times map_t \times avglevel_n$	—	—	.046	−.047
			(.369)	(.174)
Service line included	Public	Private	Public	Private
n	12,913	12,913	12,719	12,719

Table 3: Heterogeneity checks

property. I estimate

$$\ln P_{it} = \alpha + \beta \times lead_i + \gamma \times map_t + \tau \times res_i + \eta \times lead_i \times map_t + \xi \times lead_i \times map_t \times res_i + Z'_{it}\theta + \epsilon_{it},$$

where the new indicator res_i is equal to one if unit i is residential, and zero otherwise (the resulting new two-way interaction terms are included in the regression but omitted from the equation above). The coefficient of interest is ξ, and the aim is to see if there was any divergence in the effect captured by η in the previous model between the two types of units, which could be the case if LSLs started to become more of a concern for homeowners than commercial or office users following the release of the lead map. The results are reported in columns (a) and (b) of table 3 (they correspond to columns (a) and (b) table 2 in terms of grouping no-information units). Similarly, estimates of ξ have the expected sign but are not statistically significant enough to suggest the existence of diverging effects. It is worth noting that the majority of real property transactions that took place during the time period were for residential properties, and the expanded sample did not increase in size by a lot.

In the second exercise, I revert to my original sample, but include $avglevel_n$, a measure of the prevalence of LSLs in neighborhood n, and interact it with the original interaction term:

$$\ln P_{int} = \alpha + \beta \times lead_i + \gamma \times map_t + \nu \times avglevel_n + \eta \times lead_i \times map_t + \rho \times lead_i \times map_t \times avglevel_n + Z'_{it}\theta + \epsilon_{it}.$$

The coefficient of interest is ρ, and the goal is to check whether there were any divergent effects depending on the existing number of LSLs in the neighborhood: for example, higher

(a) Weekly sale volume, by public service line material

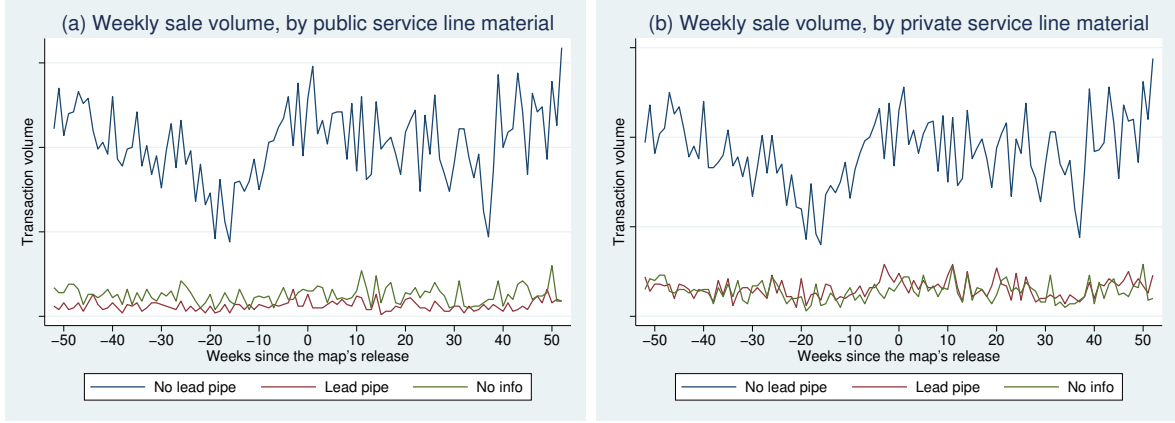(b) Weekly sale volume, by private service line material

Figure 4: Weekly transaction volumes by service line material

availability of non-LSL units may make switching away from LSL units easier, and hence lead to a bigger price effect. The estimation results are reported in columns (c) and (d) of table 3; again, no evidence of such divergence is found.

In addition, apart from price effects, I also look at transaction volumes in the year before and the year after the map's release. Because of a general upward time trend in the total volume of transactions, the second one-year period saw about 20% more transactions than the first. On the other hand, the shares of units with and without LSLs remained consistent between both periods, suggesting the absence of any selection response from property buyers after the map's release. The two panels of figure 4 show the number of transactions per week in each group of housing units, based on the public and private service lines, respectively. No noticeable divergence of trends seem to exist in either panel. The same conclusion can be reached from a diff-in-diff analysis using specifications similar to the one described earlier in the section, which I do not report here.

At the end of section 4, I pointed out that the data from the lead map may suffer from accuracy issues. In particular, data used in my analysis reflects the up-to-date records as of March 2021. Because of constant updating of the records as well as replacement of LSLs, the service line materials displayed on the map for any address can be different from those displayed when the map was released or when a potential homebuyer would check the map to look up a property of interest. With historical records unavailable, I have to resort to other less-than-perfect remedies. One feature of the lead map is that it specifies the method DC Water used to determine the service line material for each property. Determination was based on a variety of records including permits, meter records, maintenance records, replacement records, etc. When the service line material is determined because of replacement work, the map lists the date when the replacement was carried out. Therefore, I am able to omit from my main sample observations where replacement work was done on either the public or the

Variable	Specification					
	(a)	(b)	(c)	(d)	(e)	(f)
$lead_i \times map_t$	.001	$-.003$	.003	$-.000$	.001	$-.003$
	(.004)	(.004)	(.005)	(.004)	(.005)	(.005)
Service line included	Public	Private	Public	Private	Public	Private
Missing-information units			Separate group			
n	4,970	4,970	4,657	4,657	4,970	4,970

Table 4: Estimation results for the repeated-sales specification

private service line after the date of transaction (so that what I observe may not be what the buyer observed at the time), thus eliminating the main source of inaccuracy due to record updates.[5] After removing 476 such observations (with 12,236 remaining in my main sample), I repeat all the preceding exercises in this section, and all of them give virtually the same results as before. Therefore, it seems unlikely that data issues related to record updating are the main reason for my findings.

Repeat Sales

Next, I present estimation results from the repeat sales model. OLS point estimates and robust standard errors from the main specification are reported in columns (a) and (b) table 4 (coefficients on the year dummies are not reported because they don't have the simple interpretation as housing price indices). Neither estimate of η is significant, implying the map's release did not have differential effects on LSL and non-LSL units. As a partial remedy to the aforementioned pipe replacement problem, I repeat the estimation after removing from the sample observations where replacement work was done after the earlier transaction, so that it is less likely that the LSL status changed either between transactions or between the latest transaction and map data collection. This procedure removes 313 observations from the sample (with 4,657 remaining), and estimation results are reported in columns (c) and (d).

In addition, as mentioned in section 3, the existence of units that were traded more than twice causes the error covariance matrix to be non-diagonal and calls for a GLS estimation approach. I performed this using the main sample, and report the new standard errors in columns (e) and (f). Unsurprisingly, these results do not change the conclusion derived from the OLS estimates.

---

[5]All replacement work recorded in the data is relatively recent, so the replaced pipes are always non-lead. I am removing these observations instead of assuming the properties had LSLs pre-replacement because a significant number of non-LSLs are also replaced during water main work and emergency repairs, according to DC Water. Such an assumption would introduce a new source of inaccuracy in the data.

6   Interpretation of the Results

The lack of a discernible information effect seems to contradict findings by such studies as Theising (2019) and Billings & Schnepel (2017), where homebuyers are found to place a positive value on the absence of lead exposure hazards. But my results are in line with Bae's (2016) findings on the effect of disclosure alone. The fact that the release of the lead map did not appear to have any price effects in the real property market points to several possibilities, which I will not be able to fully investigate. For example, it could be the case that residents responded to the newly available information only in ways not reflected in the housing market, such as increased water testing, remediation efforts and blood testing for children. Alternatively, it could be that many homebuyers were simply unaware of the availability of the map at the time of their transaction, due to limited media coverage and government publicity given to it.

In order to further examine the price response to concerns about lead hazards in the housing market, I perform the following two exercises. Firstly, I divide housing units into those built before 1986 ("old" units), when the federal lead-pipe ban came into effect, and those built after 1986 ("new" units), and examine their price trends before and after the start of 2016, when the Flint water crisis came to national attention. At the time, because the lead map was not yet available, there was no easy way for residents to obtain information about service line materials, and the best proxy of the likelihood of LSL presence in a building is its year of construction. If health concerns led people to shun units with LSLs, there would be a divergence in the average price trends of old and new units. Figure 5 presents these trends in 2015 and 2016 in a similar way to figure 3; neighborhood and use code are controlled for. Prices for both categories moved broadly in tandem with each other, with new units selling for higher prices than old ones, as expected. While both the monthly and weekly plots seem to show some movements in the price trend around the beginning of 2016, those are in line with seasonal movements observed in other years, and unlikely to be related to the Flint crisis. Thus it seems even renewed concerns about lead exposure in response to a major incident like Flint still had little effect on the housing market. In addition, this exercise also partially address the issue about updated service line material records raised in section 4 and further discussed in section 5, since the results only rely on construction year as a proxy for service line material, instead of accurate records.

A second, related exercise similarly chooses the Flint crisis as a cutoff point, but divide housing units according to their service line material just like before. The rationale is to further check the effect of lead risk salience on the housing market. This exercise would reveal no new information compared to the previous one if people truly had no way of
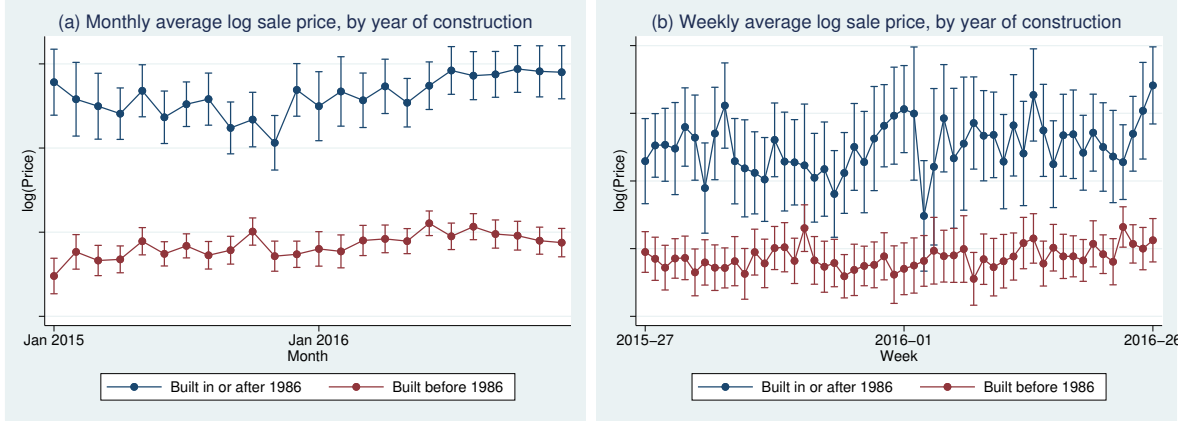
Figure 5: Monthly and weekly average log sale price by year of construction
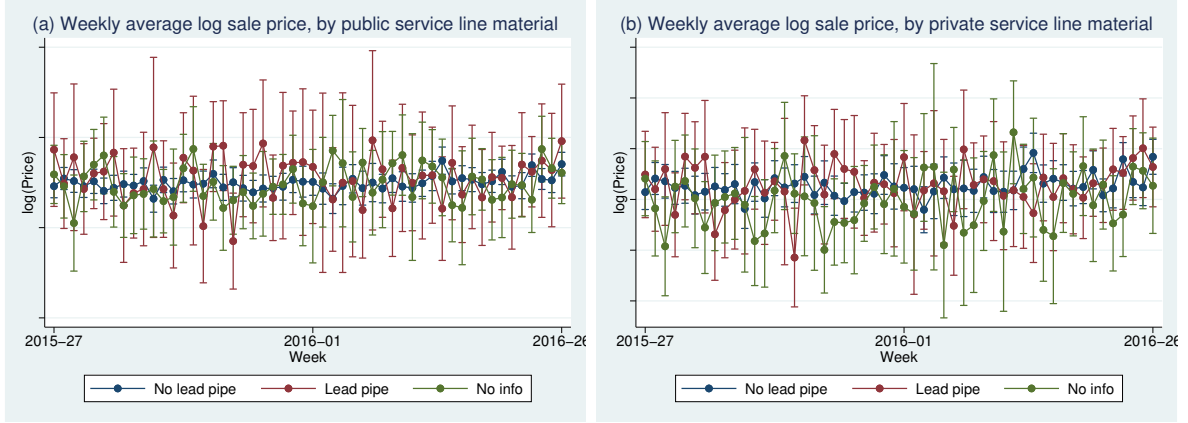

Figure 6: Weekly average log sale price by service line material

learning about their service lines before the map was released; however, if price trends did diverge post-Flint, it would most probably imply people had alternative ways of finding out about LSLs before the map (despite possibly high barriers), and new concerns led to behavioral responses in the market. However, as shown in figure 6, this does not appear to be the case.

To gain more insights into the possible reasons for the lack of significant positive effects, I present findings from several other data sources. First, I examine Google search data for terms related to lead in drinking water. In their analysis for Pew Research Center, Matsa et al. (2017) identified dozens of search terms people used when searching about the Flint water crisis. Interest in such terms reflects people's concerns about lead exposure from water, and can be used to gauge the effect of the map's release on the salience of the topic. Looking at the search frequency of the leading terms identified (including "lead pipes", "lead water", "lead in water", "lead poisoning", etc.) in DC on Google Trends, I find their popularity among DC internet users to broadly mirror the nationwide trends, with elevated interest starting

in January 2016, when the Flint crisis gained national media coverage. However, just like the national trends, there did not seem to be such a spike following the map's release. It is also worth noting that the search trends for these terms do not seem to mirror that for "DC lead pipe map" discussed in Section 2, which experienced a new and sustained rise starting in mid-2019. It seems possible that interest in the map does not fully arise from or lead to concerns about lead in drinking water.

A second set of data contains the number of blood lead tests carried out in DC, obtained from the DC Department of Health through a Freedom of Information Act (FOIA) request. It is required by law that every child in DC get tested twice for blood lead by the age of two, but parents also can take initiative and ask their doctor for tests whenever they suspect their child is at risk of lead exposure. Therefore, higher levels of such concern among parents would plausibly lead to more tests conducted. I obtained daily test count data during the three-year period from 2015 to 2017. Figure 7 shows the weekly number of tests, along with the four-week moving average, with the first and twenty-third weeks of 2016 highlighted: these were the weeks of the first national media coverage of the Flint crisis and the release of the lead map, respectively. The figure clearly shows seasonal patterns, but there does not seem to be increases in the number of tests following either event; the hike during late April and early May of 2016 was due to a separate incident, where hundreds of students from three elementary schools were tested because of findings of elevated lead levels in their drinking water. The fact that there was no noticeable response to the Flint crisis suggests that parents may not react to general concerns about lead in drinking water by having their children tested more, possibly because of the time and efforts required.

A third data set turns to a different kind of test that incurs a much lower cost. Via a FOIA request from DC Water, I obtained a data set containing the daily number of water lead test kits requested from DC Water. Similar to other cities like Chicago and New York City, the DC government sends residents such kits to conduct lead tests on request, free of charge. A kit contains two bottles for collecting water samples, along with instructions and a simple questionnaire. Upon receiving the kit in the mail, a resident can fill the bottles with tap water as instructed, and send them back to DC Water using a prepaid shipping label for the water samples to be tested for lead. Results are then delivered electronically after four to six weeks. The simplicity and low cost of requesting a kit and submitting samples mean that such a test is the one of the easiest and most accessible ways for residents to act on their concerns about lead in drinking water, and the number of tests may respond strongly when there are elevated concerns among the public. Figure 8 shows the weekly number of requests from 2015 to 2017, along with the four-week moving average and similarly with the first and twenty-third weeks of 2016 highlighted. Unsurprisingly, it displays a rather different pattern
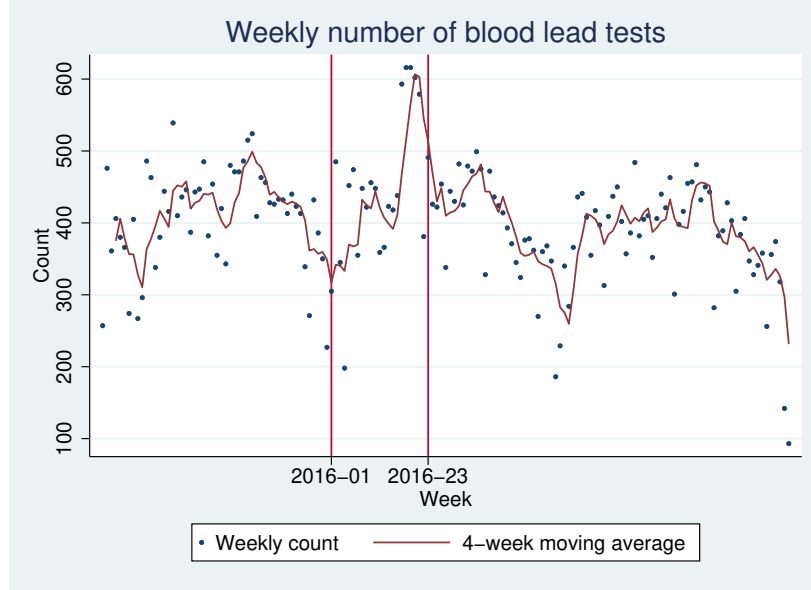
Figure 7: Weekly blood lead test count in DC, 2015–2017

from the blood test counts. Requests started increasing shortly after the Flint crisis gained national attention, reaching a height in late April 2016 about ten times the average count in 2015, and then dropping again through May. Following the map's release, the number of requests shot up in the following two weeks, more than doubling the post-Flint peak, before settling back down to a level not as high as in the immediate aftermath of Flint, but higher than in 2015. These numbers demonstrate that lead map did significantly raise DC residents' interest in getting informed about their tap water quality, even more so than a major national news story like Flint; as a matter of fact, it may have precisely been the local nature of the map that prompted more people to get their water tested. However, as illustrated by the blood test data and housing market results, such interest did not seem to translate into other kinds of actions.

To summarize, based on existing evidence, it appears that the reason for the absence of an information effect is multi-faceted. On one hand, usage of the map service seems low, as suggested by Google Trends data detailed in section 2. On the other hand, it also seems that residents only exhibited a rather limited range of behavioral responses to concerns about lead exposure in general, and not through the housing market, which implies the housing price effect of the map's release would still be limited even if take-up had been higher.
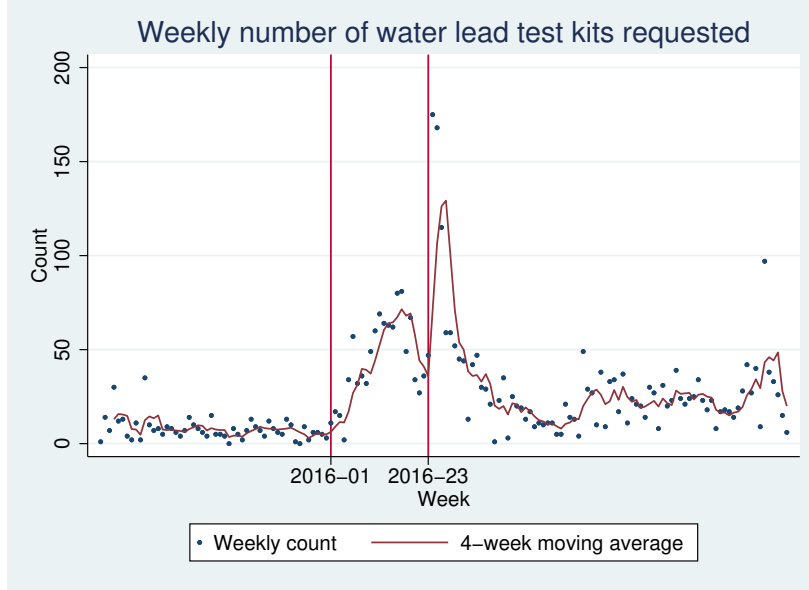
Figure 8: Weekly number of water lead test kit requests in DC, 2015–2017

7   Conclusion

Using real property transaction data, I examine the prices of LSL and non-LSL housing units around the time of the lead map's release, and find no discernible price effect of the newly available information regarding lead service lines. This remains true after I take into account the possibility of heterogeneous price effects between residential and commercial-use property or between neighborhoods with different base-level lead service line prevalence; nor did any differential effect seem to have taken place through transaction volumes.

It is difficult to determine why the map's release seems to have had no price effect. Google search data appear to suggest that few people sought out the map, and hence property buyers and sellers probably did not use it to inform their decisions. It is possible that responses to the new information mostly manifested themselves in subtler ways that were not reflected in the housing market, such as replacement of lead pipes and at-home lead testing of tap water. That points to other potential avenues to further investigate the effect of the new information on the housing market. For example, if records of lead pipe replacements become available, it would be possible and interesting to look at how the number of replacements differed before and after the release, as well as any price effects associated with replacements. Such a study would also give a clearer picture on the extent to which the map has achieved its intended goal as a public health policy.

References

Agee, M. D., & Crocker, T. D. (1996). Parental Altruism and Child Lead Exposure: Inferences from the Demand for Chelation Therapy. *The Journal of Human Resources*, *31*(3), 677–691.

Aizer, A., Currie, J., Simon, P., & Vivier, P. (2018). Do Low Levels of Blood Lead Reduce Children's Future Test Scores? *American Economic Journal: Applied Economics*, *10*(1), 307–341.

American Academy of Pediatrics Subcommittee on Accidental Poisoning. (1961). Report Of Subcommittee On Accidental Poisoning : Statement on Diagnosis and Treatment of Lead Poisoning in Childhood. *Pediatrics*, *27*(4), 676–680.

Bae, H. (2016). The impact of the residential lead paint disclosure rule on house prices: findings in the American Housing Survey. *Journal of Housing and the Built Environment*, *31*(1), 19–30.

Billings, S. B., & Schnepel, K. T. (2017). The value of a healthy home: Lead paint remediation and housing values. *Journal of Public Economics*, *153*, 69–81.

Binns, H. J., Campbell, C., & Brown, M. J. (2007). Interpreting and Managing Blood Lead Levels of Less Than 10 μg/dL in Children and Reducing Childhood Exposure to Lead: Recommendations of the Centers for Disease Control and Prevention Advisory Committee on Childhood Lead Poisoning Prevention. *Pediatrics*, *120*(5), e1285–e1298.

Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, *114*(2), 577–599.

Brown, M. J., & Margolis, S. (2012). *Lead in Drinking Water and Human Blood Lead Levels in the United States*. Centers for Disease Control and Prevention.

Brown, M. J., Raymond, J., Homa, D., Kennedy, C., & Sinks, T. (2011). Association between children's blood lead levels, lead service lines, and water disinfection, Washington, DC, 1998–2006. *Environmental Research*, *111*(1), 67–74.

Centers for Disease Control and Prevention. (2021, October). *CDC Updates Blood Lead Reference Value for Children.* (Press release)

Chay, K. Y., & Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, *113*(2), 376–424.

Clay, K., Troesken, W., & Haines, M. (2006). Lead Pipes and Child Mortality. National Bureau of Economic Research Working Paper No. 12603.

Cornwell, D. A., Brown, R. A., & Via, S. H. (2016). National Survey of Lead Service Line Occurrence. *Journal (American Water Works Association)*, *108*(4), E182–E191.

Del Toral, M. A., Porter, A., & Schock, M. R. (2013). Detection and Evaluation of Elevated

Lead Release from Service Lines: A Field Study. *Environmental Science & Technology*, *47*, 9300–9307.

Edwards, M., Triantafyllidou, S., & Best, D. (2009). Elevated Blood Lead in Young Children Due to Lead-Contaminated Drinking Water: Washington, DC, 2001–2004. *Environmental Science & Technology*, *43*(5), 1618–1623.

Gayer, T., Hamilton, J. T., & Viscusi, W. K. (2002). The Market Value of Reducing Cancer Risk: Hedonic Housing Prices with Changing Information. *Southern Economic Journal*, *69*(2), 266–289.

Gazze, L. (2021). The price and allocation effects of targeted mandates: Evidence from lead hazards. *Journal of Urban Economics*, *123*, 103345.

Goodlad, J. K., Marcus, D. K., & Fulton, J. J. (2013). Lead and Attention-Deficit/Hyperactivity Disorder (ADHD) symptoms: A meta-analysis. *Clinical Psychology Review*, *33*(3), 417–425.

Grönqvist, H., Nilsson, J. P., & Robling, P.-O. (2020). Understanding How Low Levels of Early Lead Exposure Affect Children's Life Trajectories. *Journal of Political Economy*, *128*(9), 3376–3433.

Hodge, A. T. (1981). Vitruvius, lead pipes and lead poisoning. *American Journal of Archaeology*, *85*(4), 486–491.

Linden, L., & Rockoff, J. E. (2008). Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *The American Economic Review*, *98*(3), 1103–1127.

Mastromonaco, R. (2015). Do environmental right-to-know laws affect markets? Capitalization of information in the toxic release inventory. *Journal of Environmental Economics and Management*, *71*(C), 54–70.

Matsa, K. E., Mitchell, A., & Stocking, G. (2017). *Searching for News: The Flint water crisis*. Pew Research Center.

Needleman, H. L., & Gatsonis, C. A. (1990). Low-Level Lead Exposure and the IQ of Children: A Meta-analysis of Modern Studies. *Journal of the American Medical Association*, *263*(5), 673–678.

Palmquist, R. B. (1982). Measuring Environmental Effects on Property Values Without Hedonic Regressions. *Journal of Urban Economics*, *11*(3), 333–347.

Pope, J. (2008). Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise. *Journal of Urban Economics*, *63*(2), 498–516.

Pueschel, S. M., Linakis, J. G., & Anderson, A. C. (1996). Lead Poisoning in Childhood. In S. M. Pueschel, J. G. Linakis, & A. C. Anderson (Eds.), (pp. 1–13). P. Brookes Publishing Company.

Rabin, R. (2008). The Lead Industry and Lead Water Pipes: "A Modest Campaign".

American Journal of Public Health, *98*(9), 1584–1592.

Reyes, J. W. (2007). Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime. *The B.E. Journal of Economic Analysis & Policy*, *7*(1), 1–43.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55.

Roy, S., & Edwards, M. A. (2019). Preventing another lead (Pb) in drinking water crisis: Lessons from the Washington D.C. and Flint MI contamination events. *Current Opinion in Environmental Science & Health*, *7*, 34–44. (Drinking water contaminants)

Sandvig, A., Kwan, P., Kirmeyer, G., Maynard, B., Mast, D., Trussell, R. R., ... Prescott, A. (2008). *Contribution of Service Line and Plumbing Fixtures to Lead and Copper Rule Compliance Issues. Research Report 91229.* American Water Works Association Research Foundation.

Selevan, S., Rice, D., Hogan, K., Euling, S., Pfahles-Hutchens, A., & Bethel, J. (2003). Blood Lead Concentration and Delayed Puberty in Girls. *The New England journal of medicine*, *348*, 1527–36.

Theising, A. (2019). Lead Pipes, Prescriptive Policy and Property Values. *Environmental and Resource Economics*, *74*(3), 1355–1382.

Triantafyllidou, S., & Edwards, M. (2012). Lead (Pb) in Tap Water and in Blood: Implications for Lead Exposure in the United States. *Critical Reviews in Environmental Science and Technology*, *42*(13), 1297–1352.

Troesken, W. (2006). *The Great Lead Water Pipe Disaster.* London: The MIT Press.

Votsis, A., & Perrels, A. (2015). Housing Prices and the Public Disclosure of Flood Risk: A Difference-in-Differences Analysis in Finland. *The Journal of Real Estate Finance and Economics*, *53*, 450–471.