Invited presentation at the 2018 Southern Agricultural Economics Association Annual Meeting, February 2-6, 2018, Jacksonville, Florida

WILLINGNESS TO AVOID LEAD RISK IN WATER QUALITY: ARE THERE INFORMATION ASYMMETRIES?

Willie B. Bedell[1]
University of Kentucky

February 2018

Abstract
Most Americans have access to water, yet America faces water challenges including lead contamination. Our research asked questions about potential lead risk and lead risk information asymmetry among homeowners. Using data from the Property Valuation Administrator in Fayette, Kentucky, we answer the questions through hedonic analysis. We find an implicit positive value to avoiding lead-risk and buyers in high-lead risk neighborhoods might be less informed. Our policy implications recommend to States and local authorities to periodically communicate lead risk to the public.

Keywords
Water Quality, Lead Exposure Risk, Hedonic Price, Information Asymmetries

---

I. Introduction

The United States access to water is rated at 99.26%, yet America is constrained with water quality assurance, pointing to risks of lead in the water supply systems. Accordingly, a quality water supply should be freed from meaningful physical, chemical, biological, or harmful radiological substances, including lead contamination. The Environmental Protection Agency (EPA) regulations set lead concentration to be less than 15 parts per billion (ppb) in a sampled 10% of a water zone. An EPA disclosed puts nine States to be reporting safe lead levels in their water supply. However, about 5,300 States' water systems might be in violation of the lead rules, and there is a lack of residents' trust when it comes to reporting, testing, and treatment for lead contamination in water supplies, CNN (2016). Residents could possess elementary knowledge about lead contamination, but they might not be certain in determining the level of lead in their drinking water. The specific problem lies in the identification of lead contamination in a community water supply. The risks of lead in a water system can be certainly identified or challenging to detect. In a certain case of lead risk identification, some counties in America report with confidence that their geographical subdivisions have no known lead risks; other counties confirm lead contents in their water supply. On the gray front, some counties cannot point out if their homes, water meters, water treatment and distribution systems contain lead materials.

Given this background, we ask the following questions: How does lead exposure risk in a water supply system affect housing values? Is there a presence of information asymmetries among

---

[1] This paper is prepared for the Annual Meeting of the Southern Agricultural Economics Association, Jacksonville, Florida, February 2018. Bedell: Graduate student, Department of Agricultural Economics, University of Kentucky. willie.bedell@uky.edu

homeowners relative to their water supplies? We hypothesized that residential homeowners will pay more for lead-free water. We expect that homeowners in the high lead risk water areas may have little information on their water supply compared to low lead exposure neighborhoods. We assume asymmetric information in high-risk areas because water lead risk is not required to be reported during home sale transactions. Besides, independent observers see that a water problem is frequently concealed above all other home disclosures, Gassett (2016). In this study, we hope to measure the influence of lead risk level on residential home sale transaction and to communicate policy implications for homeowners, home sellers, and policymakers to address the concern of lead risk in the water systems.

This paper's contribution to the water quality literature lies in assessing the hedonic valuation of a probable lead risk in drinking water and searching for the possibility of asymmetric information. The literature on water hedonic pricing ranges from valuing surface water to pricing water through the land market. Other studies, not using hedonic analysis, have also evaluated the willingness to pay for municipal water supply and drinking water quality. Leggett & Bockstael (2000) used hedonic analysis to demonstrate the effect of water quality on waterfront properties in the Chesapeake Bay area. Using eight empirical specifications, Leggett & Bockstael reports show that as bacteria level increase in the Bay, properties values reduce by 5%. Buck *et al.* (2014) research also contribute to this paper through their supportive argument that a stable market can be used to infer the value of an environmental good. Piper (2003) evaluates the impact of water quality on municipal water price and residential water demand. Piper's results support the argument that households appreciate an improved water quality system. Powell & Allee (1990) results, from a contingent valuation survey carried out in four towns in Massachusetts, show that people are willing to pay more for drinking water quality, especially when they have experienced contamination harms. Jordan & Elnagheeb (1993) find that 23% of households were uncertain of their drinking water quality. There are also trending studies which are addressing the general course of lead in the national water supply. Theising (2017) is discussing "Lead pipes, prescriptive policy and property values, lead pipes and prescriptive policy"; Irwin (2017) on "Homebuyer risk perception in the face of potential lead exposure", and Grooms *et al.*, (2017) on "Drinking water and lead: Evidence from local treatment changes in North Carolina".

The theoretical framework of our paper is underpinned by the hedonic theory. Rosen (1974) presents that goods are valued for their utility bearing attributes. A class of differentiated products has a vector of measurable characteristics which define a set of hedonic prices. We assume the housing market is a differentiated product market, which has a bundle of characteristics. Residential homes host characteristics like square feet, structural design of the home, being close to an environmental amenity or being linked to a certain water quality zone. Overall, the theory calls for information on the structural characteristics, neighborhood and environmental attributes, and other controls to capture implicit price effects. We depart from Rosen on the assumptions of no second-hand market and the perception of identical characteristics. Akerlof (1970)'s theory debated on quality in. He called out good and bad qualities in the market mechanisms and argued that rational consumers will demand better quality. To carry out the hedonic analysis, our dataset merges information from Fayette County Property Valuation Administrator, Vox Media, and other sources like Census Bureau. The study period covered a 17-year period (2000 to 2016) and included 70,619 sales transactions that occurred in the County. To evaluate the willingness to pay for lead risk in water quality supply, we applied empirical specification of ordinary least square, two-stage least squares, and propensity score matching. To check for asymmetric information, we constructed a deterministic model, advised by Kurlat and

Stroebel (2015), to test for differences between appreciating rates in the low and high lead water risk zones.

In this paper, Section 2 presents the problem of lead contamination in water supplies and the specific problem of identifying lead risk in water communities. Section 3 provides a review of the hedonic literature, water quality papers, and hints on trending studies in the case of lead in the United States. Section 4 describes the theories that support the study and the data in Section 5. We examine the research questions in Section 6 through the empirical specifications, including a deterministic test. Section 7 shows the findings from the willingness to pay evaluations and reports the test of the deterministic model on asymmetric information. Finally, Section 8 gives a summary, provides implications, and suggest a way to improve the study's gaps.

II. Background

Global water supply is very important to every region of the world, as water supply issues sprout from health to environmental concerns. For example, despite the North America, Canada as well, water accessibility above the global average of 90.7%, the United States face challenges of water allocation and quality, counting alarming risks of lead level in drinking water. According to the Safe Drinking Water Act (SDWA), a quality water supply is free from meaningful physical, chemical, biological, or harmful radiological substances. Relative to the risk of lead contamination, the United States Environmental Protection Agency (EPA) regulates that the lead concentration in public water systems must not exceed an action level of 15 parts per billion (ppb) in more than 10% of customers' tap samples.

In March 2016, CNBC News article published: "America's water crisis goes beyond Flint, Michigan…impacting millions of lives and costing billions of dollars in damages." The new wave of lead awareness reverberates the need for financing and investing in the Nation's water infrastructures to protect citizens from serious public health dangers. EPA disclosed that only nine States are reporting safe levels of lead in their water supply. These States include Alabama, Arkansas, Hawaii, Kentucky, Mississippi, Nevada, North Dakota, South Dakota and Tennessee. 5,300 States' water systems, supplying about eighteen million Americans, are believed to be in violation of lead rules. From a residential viewpoint, homeowners perceive a lack of trustworthiness in the reporting, testing, and treatment for lead contamination in water supplies CNN (2016).

Residents may have basic knowledge about how one can be exposed to lead contamination, but these residents might not know the level of lead in their drinking water. Sometimes, through macro-observation, residents can easily point out that their water supply is contaminated. For example, when the city of Flint switched to the Flint River for supply, residents noticed and complained about the discolored water. Besides, pediatricians and independent studies respectively noticed the high level of lead in children and the local water supply, Jordan (2016). Yet, locating lead pipes, which cause contaminations in communities' water supply, can be a challenge. The central problem lies in identifying water services or supplies that is contaminated by lead. Lead interactions with water supply can be a certain case or a gray area of identification between certainty and uncertainty. In the case of certainty, some counties in the Nation report with confidence that their geographical subdivisions have no known lead contents in their water services, while others confirm lead contents in their water services. Illustrating the former on the certainty of no lead contamination, a few counties in Southwest Ohio, including Butler, Englewood, Fairborn, Fairfield, and Green counties, report no lead contamination in their water services. These counties are certain on the grounds that meters were replaced; communities contain only iron, copper, and plastic pipes; new developments and buildings were constructed after 1998, and their water treatment and distribution systems were created after 1957. Regarding the latter on confirmed lead risk, some counties in the region have validated the presence of lead risks in their areas. Herein, counties such as Miamisburg, New Carlisle, Oakwood, Oxford, and Sidney county confirmed lead services on the basis that supply connections and service lines in public and private properties in these counties contain lead in their water distribution systems. That is, there are proven lead lines, solders, fixtures, or goosenecks; and most homes were built between 1900 and 1950, and as far back as 1895. Counties like Piqua and Franklin in Southwest Ohio report their water systems as probable lead areas because of old developments, while lead pipes are being replaced, Driscoll (2017).

This study asks these research questions: What is the hedonic valuation of lead exposure risk in the water supply? Is there a presence of information asymmetries among homeowners relative to their water supplies? We expect that residential homeowners will pay more price for good water quality. We also expect, on average, homeowners in high lead water exposure areas may have little information on their water supply compared to low lead exposure neighborhoods. The latter hypothesis assumes that asymmetric information is present, mostly for high-risk than low-risk areas, see Figure 1.

Depicting the assumption of water quality information asymmetries

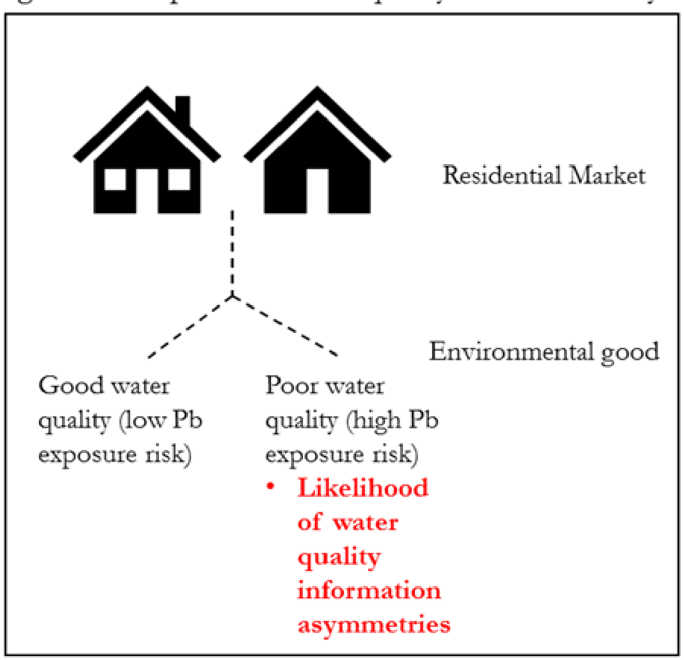

Figure 1: Visually illustrating homes that may be prone to concealed lead risk information in the residential market.

In terms of disclosing lead information, there is a legal obligation to disclose lead-paint, while there are not rigid regulations on the disclosure of a probable lead material that could contaminate the water supply. Nonetheless, EPA provides information on lead in drinking water and important steps families can take to reduce lead contamination in their drinking water. For example, EPA has advised homes to contact their water company in order to determine if their water system or service line has a lead material.

The first objective of this paper is focused on applying an empirical evaluation, the hedonic price analysis, to investigate the marginal willingness to pay for a quality water system. The second objective in this paper aims to discuss the likelihood of market failure of asymmetric information relative to lead risk in the water supply via the residential market. Through these objectives, this study will provide findings, implications, and suggestions for homeowners, residential home sellers, and policymakers to address lead risk in the water systems.

It is important to consider the effect of a probable lead risk on housing valuation due to the health concerns, environmental engagements, and decisions that are needed to finance the replacement of lead service lines and lead components in US communities. Lead contamination in the water supply is an aged infrastructure problem which is caused by older pipes that contain lead. Lead may enter water systems as it dissolves through lead-pipes when water passes through the distribution channels. One of the main sources of lead contamination is lead service lines that connect water mains to residential properties. Also, pipes within homes, soldered with lead, might contribute to lead contamination, as water sits idle in these pipes while the system is not in use in the home, Kentucky Division of Water (2017). In 1986, amendments to the drinking water act prohibited the use of the not-lead-free pipe, plumbing fitting, fixture, solder, or flux in public, residential, and nonresidential buildings. Not until 1996, it became unlawful for any person to introduce into the market any pipe, plumbing fitting or fixture which is not lead-free, EPA and Cornell Law School (2017). Lead toxicity gives rise to serious health defects in the human body. Most especially, its harm is in severity to little children. 10% to 20% of a lead intake in children is caused when water, which is poisoned by lead, is consumed, Rabin (2008). Lead intake and accumulation promote weakened cognitive development in children, damage kidney function, produce cardiovascular problems, and negatively affect the brain, liver, and bones. In extreme cases, lead intake might result in death. In short, the presence of lead in a water supply can engender health risks to the public.

Financing the replacement of lead in public and private properties is a major challenge to mitigating, or better say, eradicating lead poisoning in the United States. There are alternative financing approaches, designed or being proposed. Since 1996, the Drinking Water State Revolving Fund (DWSRF), a federal-state partnership, has promoted the financing of safe water systems in each State. Following the Flint crisis in 2014, which was earlier mentioned, the White House pledged more than five-billion dollars to improve water quality in the nation. Today, Federal agencies such as CDC and EPA, tasked with tackling lead contamination, have experienced a fiscal year (FY) budget cut. For example, CDC encountered a 17% budget reduction in FY 2017/2018, including a cut in areas of prevention, environmental health, and toxic substances, CDC (2018). In addition to the financing mechanisms, a legislative bill is debating the provision of loans or grants to finance the removal of lead pipes. This Lead Act, LRB-1934, calls for authorities to be given to local governments for the provision of an opportunity for a local water utility to provide financial assistance for replacing the lead service lines, Cowles (2017).

III. Literature Review

This paper contributes to the water quality literature by assessing the hedonic valuation of a probable lead exposure risk in drinking water and to search for the possibility of asymmetric information relative to lead risk in the residential market. The literature on water hedonic pricing ranges from the surface water valuation of waterfront properties to pricing water through the land market. Other studies, not using hedonic analysis have also evaluated the willingness to pay for municipal water supply and drinking water quality. Together, these studies have considered and investigated the value of surface and groundwater quality amid water challenges. For instance, valuation papers have covered damages caused by sediments, bacteria, nutrients, and soil erosion-related pollutants in water networks such as streams, lakes, reservoirs, and estuaries. At the time of this research, there are a series of forthcoming papers that are addressing the general course of lead in the national water supply.

A. Hedonic Valuation of Water Quality

The attempt to understand the implicit valuation of environmental goods in a nonmarket scenario is a path we aim to build our analysis on. There may not be an explicit market to price whether homeowners place an appreciation or depreciation on their water quality. However, given this nonmarket scenario, the hedonic analysis is a tool that can tease out the willingness to pay for lead exposure risk. Leggett & Bockstael (2000) used hedonic analysis to demonstrate the effect of water quality on waterfront properties in the Chesapeake Bay area. Leggett & Bockstael influence the approach and analysis of this paper through their argued hypothesis that good water quality positively affects the value of residential property. Accordingly, homeowners are expected to bid for prices of residential units which have a desirable level of characteristics, including water quality. It is expected that locals will be willing to pay an appreciated price for low exposure risk neighborhoods. The low-lead risk is certainly a higher environmental quality. The authors point out to a robust empirical work as a convincing factor in considering the significance of the environmental result. In this case, a robust empirical work means cleaning the analysis of ambiguities, such as functional form, and addressing market segmentation and multicollinearity. So, Leggett & Bockstael (2000) measured water quality, referencing waterfront amenity to properties, on the level of fecal coliform bacteria existing in the water. In their study, it was assumed that information on coliform bacteria is spatially and explicitly available to the public. That is, residential homeowners in the study context had symmetric information on the level of fecal in their surface water. This assumption is vital to this paper because it sets the stage to picture the argument of information asymmetries. Leggett & Bockstael (2000) constructed eight empirical specifications to estimate the hedonic price using an ordinary least square (OLS) estimator. Within the scope of their study, Anne Arundel County, Maryland, Leggett and Bockstael find that an increasing level of bacteria in the Bay significantly reduces property values by 5%.

Buck *et al.* (2014) research also influence this paper through their supportive argument by using a stable market to infer the value of an environmental good; and a reminder that the hedonic procedure is common to influencing environmental policies. Buck *et al.* use evidence from the land market to infer the value of irrigation water, an environmental nonmarket good. Unlike using the OLS cross-sectional data estimator, the authors support the application of a hedonic model that uses a fixed effect estimator. Besides, Buck *et al.* highlight the use of an instrumental variable model, as used by Kuminoff & Pope (2012) and Bishop & Timmins (2013), as an alternate approach which could consistently and comparably estimate the willingness to pay for an environmental good. This advice is adhered to, in our econometric section, as one of the robust estimators. Despite the uniqueness of the hedonic literature, other empirical models can be used to estimate the implicit value of water quality. For example, Bockstael *et al.* (1987) applied models of systems of demands, discrete choice, and hedonic travel cost to validate the willingness to pay for water quality.

B. Non-Hedonic Valuation of Residential Water Supply

It is almost inevitable to argue against the premise that high-quality water supply is valued above low-quality water supply. This argument is also necessary to set the expectation when measuring the impact of lead exposure risk in water quality on the residential market. An array of non-hedonic studies has evidently proven the vertical structure of water valuation. Piper (2003) evaluates the impact and implications of water quality on municipal water price and residential water demand. Piper work assessed the extent to which a water quality influences residential water supply expense system and impacts the households' prices. Piper (2003) supports the arguments that households have a higher willingness to pay for improved domestic water quality. So, in the water-use model, piper conclusion holds that poor water quality leads to higher treatment cost and higher water rates. Jordan & Elnagheeb (1993), Powell & Allee (1990), and Schultz & Lindsay (1990)'s hypotheses also agree on the willingness to pay for an improved water quality. Jordan & Elnagheeb (1993) surveyed people's willingness to pay for improved drinking water quality and the perception of water impurity in their area. Jordan & Elnagheeb, using the contingent valuation method in Georgia, find that 23% of households were uncertain of their drinking water quality. This finding is essential to communicate the presence of asymmetric information among residents in a water community. Powell & Allee (1990) results, from a contingent valuation survey carried out in four towns in Massachusetts, show that people are willing to pay more for drinking water quality, especially when they have experienced contamination harms. Finally, demonstrating homeowners' willingness to pay for water quality, Schultz & Lindsay (1990) results showed that both residents and the community were willing to pay a higher price for a hypothetical groundwater plan.

Forthcoming studies are equally addressing the environmental constraint of lead risks. Some of the impending studies include discussion on lead pipes, prescriptive policy and property values, Theising (2017); lead pipes and prescriptive policy: Estimating homebuyer risk perception in the face of potential lead exposure, Irwin (2017); and drinking water and lead exposure: Evidence from local treatment changes in North Carolina, Grooms *et al.,* (2017).

IV. Economic Model

The core of the analysis in this paper revolves around the hedonic theory. In accordance with Rosen (1974), the hedonic hypothesis presents that goods are valued for their utility bearing attributes. Therein, the theory draws that a class of differentiated products, which has a vector of measurable objective characteristics, define a set of hedonic prices. The housing market meets the assumption of a differentiated product market, which has a bundle of characteristics. Residential homes hold different characteristics such as square feet, structural design of the home, being close to an environmental amenity, or proximity to a certain water quality zone. Paramount to the theory is the argument of spatial economic equilibrium: A consolidated set of implicit prices guides both the consumers and producers' locational decisions in a characteristic space. Analyzing water systems exposure to lead risk through this framework, the result will yield a hedonic price for lead exposure area, as well as empirical implications. Market equilibrium, a price clearing force which guides the decisions of both buyers and sellers, coordinate the implicit prices from a set of characteristics. The bundle of characteristics in the housing market includes structural, neighborhood, environmental and time attributes.

$$P = f(H, N, T\ldots, E) \qquad (1)$$

Where, P, the market price is a function of the vectors of housing characteristics H, neighborhood attributes N, time effect T, and other utilities, including a vector of environmental amenities, E. Although homeowners may subjectively value of lead risk, according to Rosen, it is assumed that all homeowners perceive identical characteristics. Put in another sense, homeowners are knowledgeable about their water quality. It is also assumed that differentiated homes may also be sold in a separate, yet highly interrelated market.

On the producers' front, Rosen (1974) presents that producers carefully consider the package of characteristics to assembly in a locational decision. Residential home sellers want to equally minimize their factor costs and produce optimal utilities. Given the latter producer motive and arguing in favor of asymmetric information, some home sellers could conceal information like the probable level of lead contamination in the water supply. Not intentionally, this action could be carried out in the spirit of minimizing cost and presenting optimal utilities to potential residential buyers.

$$P(z) = CM\ (M, H, N, \ldots, E) \qquad (2)$$

Equation 2 manifests that the *P(z)*, the total cost of all attributes, is dependent on the function of positive and increasing cost *CM*, relative to the number of homes produced by a residential seller *M* and the assembled attributes, the vectors of *H*, *N*, and *E*. For a market equilibrium to be satisfied, Equation 2 must be equal to a consumer bid function, as shown in Equation 3. Equation 3 is a value function where *P(z)* is the amount that consumers are willing to pay for attributes (*H, N, …, E*) at a fixed utility which is optimally chosen. *a*, depicting optimally chosen, indicates that the utility bundles differ from household to household and utility dependents on budget constraints, the household income *y*.

$$P(z) = U\ (M, H, N, \ldots, E; a, y) \qquad (3)$$

Despite many inferences that can be achieved through the hedonic theory, the theory is clouded with assumptions and faced with empirical challenges. To name a few, empirical challenges related to the hedonic model include omitted variable bias, multicollinearity, choice of functional form, market segmentation, and attitude to risks, Leggett & Bockstael (2000) and Hanley *et al.* (2002). First, our empirical evidence of lead exposure could be potentially bias if our model omits a variable which is important to explain either the housing price or the lead risk variable. The variables incorporated in this study, relative to the structural, environmental, and neighborhood attributes, might be highly collinear. Thirdly, potential homeowners attitude to risks, such as avoiding crimes, less energy efficient homes, and environmental disamenities, could also introduce a biased estimate. Finally, choosing either a parsimonious or flexible functional form is affiliated with a benefit and cost. For instance, traditional parsimonious forms like linear and logarithmic functions may have economic interpretations and be more robust to misspecification, Cropper *et al.* (1988) and Kuminoff *et al.* (2010).

Our point of departure from Rosen (1974) is on the assumptions of no second-hand market and symmetric information. We disapprove the former assumption of no second-hand market in this paper because the residential market is not purely consumed. There is an overwhelming evidence of resale homes in the residential market. On the latter assumption, Rosen assumes that although consumers may differ in their subjective valuations of differentiated packages, all consumers' perceptions or appraisals of a number of characteristics embodied is identical. We question the latter assumption and align our analysis with the theory of asymmetric information. Akerlof (1970) logically discussed the uncertainty of quality in market mechanisms. Practically observing the housing market through the eye of Akerlof, there are homes with good and bad qualities. Rationally, consumers will demand better quality.

$$Q^d = D(P, \mu) \qquad (4)$$

$$\mu = \mu(P) \qquad (5)$$

Q^d depicts the demand for homes; P depicts price, and μ depicts water quality. In equilibrium, Akerlof assumed that supply equals demand for given average quality. Equation 5 depicts that the demand for water quality depends on price. Yet, Equation 4 and 5 may not hold true in the presence of asymmetric information. Relative to probable lead risk, a buyer may perform a housing transaction without knowing whether the home is located within a relatively low or high lead exposure zone. Even though it is possible to value lead exposure risk through the residential market, there might be a market failure of asymmetric information associated with the housing market. This market inefficiency might be attributed to hidden type model, wherein residential property sellers may have private information on the water supply that potential buyers are not knowledgeable of, Snyder & Nicholson (2008). However, over a length of time, homeowners may acquire a fair knowledge of the quality of their water supply. Knowing then the quality of their water, the home's resale may in part manifest the asymmetric information. Consequently, the null hypothesis establishes that homes in a low-risk area may sell cheaper or at the same price relative to homes in a high-risk area due to information asymmetries.

V. Data

The hedonic function calls for information on the housing characteristics, neighborhood and environmental attributes, and other control variables in order to empirically capture the implicit price effects on residential properties. We collected information from multiple sources and joined the information through a Geo-information system (GIS) technique using the Quantum GIS (QGIS) software. Table 1 gives the summary statistics of the continuous variables of home, neighborhood, and environmental attributes used in this study. Considering the study period, the data cover a 17-year period (2000 to 2016), featuring 70,619 sales transactions that occurred in Fayette County.

A. Lead Risk Data in the Context of Fayette Kentucky

Fayette County, a Commonwealth county in Kentucky, encounters the same concerns of water challenges, including lead risks. Locals frequently ask questions about their drinking water quality like "Why do I have cloudy or milky water? Why do I have brown or yellow water? Is there lead in my water? What is the difference between hard and soft water?", Kentucky American Water (2017).

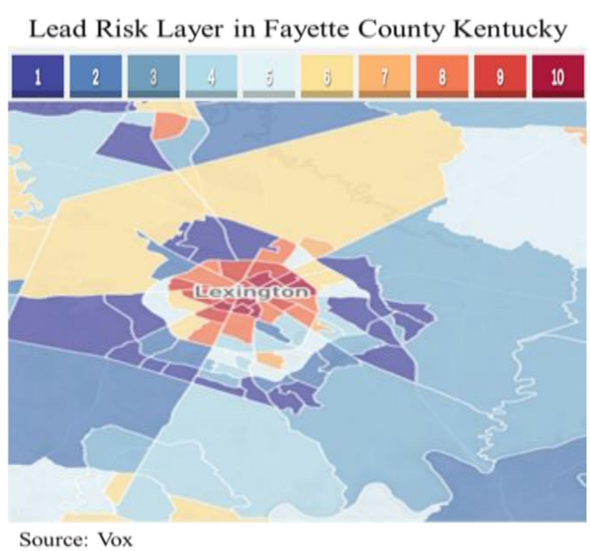

Figure 2: A snapshot of lead risk map zoomed to Fayette, Kentucky. The map is accredited to Rad Cunningham, Sarah Frostenson, and Vox media. Risk increases in ascending order on the scale of 1 to 1

"Because service lines, faucet fixtures, household pipes, and/or solder can contribute significantly to the lead and copper levels in tap water, we ask our customers to collect samples in their homes. These samples are collected on a routine basis (systems begin by monitoring once every six months with reductions in sampling possible that allow for monitoring once every three years) at homes that are considered vulnerable based on when they were constructed, and the materials used. We do this monitoring according to the requirements of the Lead and Copper Rule and use the results to confirm that our corrosion control strategy is operating as intended." This quotation is a statement from Kentucky American Water, revealing that the water system in Fayette County, or State at large, do encounter probable lead challenges. Kentucky has an active program to address lead. The program contains laws and regulations on lead, including trainings, certifications, and investigations of lead complaints. Although the average water quality in Fayette County is great, the County presents an interesting variation in the study of lead contamination. We do not have an actual information on lead contaminants in the Fayette water system. Nevertheless, this study uses an exogenous proxy called lead exposure risk, collected from Vox Media, to account for the lead in the Fayette water supply. Lead-risk zones in Fayette, structured at the Census Tract level, varies from potentially low to high-risk areas, see Figure 2. Older neighborhoods in Fayette County are rated as highly probable exposure zones, while newer residential areas in the County are rated as low probable exposure zones. Homes in the older sections of Fayette County might have lead pipes and solders, yet the search and replace plan is undergoing, WKYT (2016). Vox teamed with epidemiologists from Washington State's Department of Health to estimate the risk levels in every geographic area of the United States. The data is originally calculated from the Census data. Similar to the United States, Fayette County is systematized into 10 lead risk layers. In an ascending order, 1 represents the lowest risk area, while 10 represents the highest risk area. The exogenous proxy, lead exposure risk does not confirm that there is an actual lead contaminant in a water supply. Instead, lead paint, the age of the house, and the poverty rate are attributing factors, as used by the researchers, to construct the lead-risk variable and, as applied by this study, to indicate lead risk in the water supply.

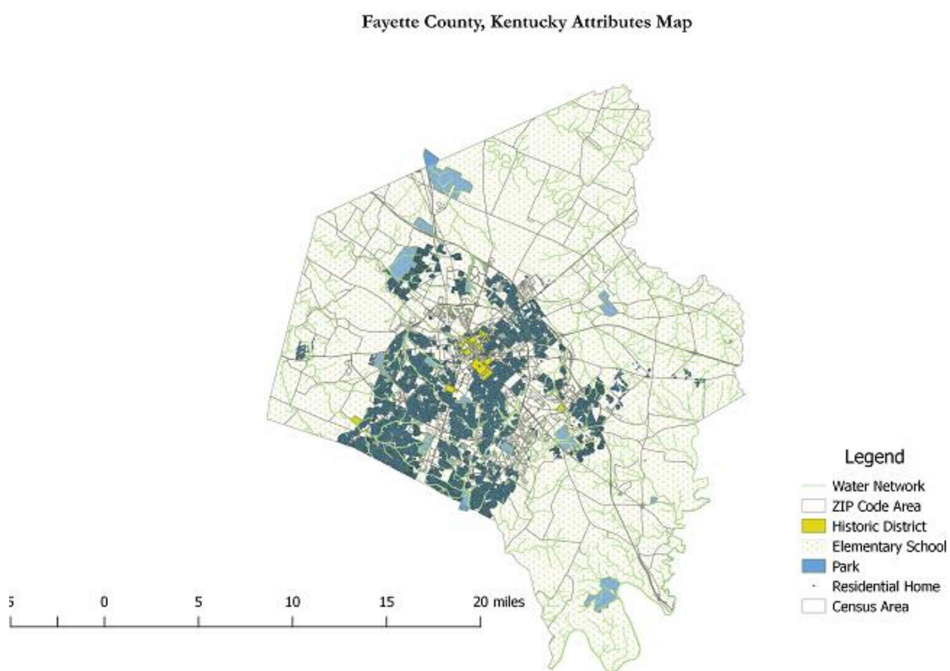

Figure 3: A cluster of residential home transactions in Fayette County over the period 2000 to 2016. Figure 2 graphically illustrates the physical neighborhood characteristics that might influence the price of residential homes

B. Housing and Neighborhood Characteristics Data

Structural attributes data are necessary to control for the housing characteristics influence on residential homes. The Housing transaction data is accredited to Fayette County Property Valuation Administrator (PVA). PVA collects and maintains residential property data, and track ownership changes and update changing characteristics of properties in the County. The PVA data provide information on home price, month and year of the sales transaction, the year the home was built, the building square feet, garage type, fixed and half bathrooms, and the property acreage. The sale price is the study dependent variable. We adjust this price to accurately reflect the current market value of homes, using 2016 consumer price index (CPI 2016 = 240.007). Table 1 presents the summary statistics for the continuous variables of the structural characteristics, including the median household income and lead risk in Fayette County.

The data analysis shows that homes are mostly sold at or above the average price during the months of May, June, July, August, and September. June and July are the peak months of sales. Respectively, about 11% of residential homes, over the study period, are sold in June and July and sold above the average inflated price of $219,328. The slowest sales months in the County are January and February. January contributes to the number of homes sold at 5% and February at 6%. Moreover, the peak sales in the county occurred during the years of 2003, 2004, and 2005, at 8%, 8%, and 7%. These sales reveal the year fixed effects, as well as the relative economic times in those years. An additional structural description showed that about 64% of homes in the study period has attached garage, followed by homes with detached garage at 15% and homes with no garage at 14%.

Table 1. Summary Statistics Showing the Continuous Variables of Structural Attributes, including median household income and lead-risk.

Variable	Observation	Mean	Std. Dev.	Min	Max
Real Sale Price (dollars)	70,619	$219,328	$148,676	$16,202	$5,649,844
House Age (years)	70,619	35	25	1	207
Home Size (Square Feet)	70,619	1,846	744	416	10,762
Story (unit)	70,619	1.4	0.5	1	3
Fixed Bathroom (unit)	70,619	2	1	1	10
Half Bathroom (unit)	70,619	0.5	0.5	0	4
Property acreage (acres)	70,619	0.2	0.2	0.002	9.8
Median Income (dollars)	70,619	$57,559	$20,915	$12,288	$168,103
Lead Risk (unit)	70,619	4	3	1	10

Controlling for the neighborhood attributes, we collected data factors from the Lexington government, including Fayette County Public School System (FCPS). These data were collected through the County open access portals for information on neighborhood attributes. Presented in dichotomous variables, information on the neighborhood and environmental include if the home is located within a historic district, if it located within 0.1 miles to a park, and if it is within 0.06 miles to a water network. Other neighborhood factors are the associated elementary school district boundary, Zone Improvement Plan (ZIP) area, and the household median income at a Census Tract level. There are 17 local historic districts included in the data. Being a place of cultural, historical, and environmental attractions, historic districts may provide benefits to residential homeowners in the form of higher property value and tax breaks. Homeowners in Fayette may also enjoy the amenities of about 100 parks. The features in these parks range from the types of parks such as community, golf-course, or neighborhood park with recreational facilities like aquatic centers, and multi-purpose trails. We observe that approximately 10% of the homes in the data are waterfront properties. Waterfronts in Fayette are made up of creeks, runs, tributaries, folks, and branches. In the same spirit of neighborhood factors, we include 14 ZIP areas, about 33 elementary school districts, and the median household incomes in Fayette County to control for potential influences on buyers and seller's decisions. ZIP code areas might give forth the linked socioeconomic factors of a neighborhood; median income averaged at $57,559 might depict homeowner preference, money constraint, and financial ability to address lead risk. The elementary districts control for the associated performance level of a local school district. Measured by the average test score percentiles from 2011 to 2015, this study categorized the elementary districts by their average performance levels into Distinguished (90 – 100 percentiles), Proficient (70 – 89 percentiles), or Needs Improvement (0 – 69 percentiles).

C. Analyzing Neighborhood Characteristics by Relative Lead Exposure Risk

For further data analysis, the lead-risk exposure zones are organized into two groups for matching purposes. As shown in Figure 2, lead-risk level 6 to 10 are identified as neighborhoods having the highest probability of lead exposure. This is our treatment group (37% of residential sales). Lead-risk areas in level 1 to 5 indicate areas with the lowest probability of lead exposure. This zone represents the control group (63% of residential sales). At the 5% significance level, Table 2 finds that variables in the two groups, treatment, and control, are statistically different. For example, the average median income for homes in the high lead exposure zones is $43,637 while the average median income for homes in the low lead-risk areas is $65,722; homes in the high exposure zone have an average age of 26 years, whereas homes in the low exposure zone have an average age of 52 years.

Table 2: Comparative Summary Statistics Belonging to the Lead-Risk Treatment and Control Groups

Variable	Low Lead Risk Zone (Control Group) 44,519 Homes	High Lead Risk Zone (Treatment Group) 26,100 Homes	T-Statistics (Pr(|T| > |t|)
	Mean	Mean	p-value
Real Price (dollars)	246111	173643	0.0000***
Lead Risk (unit)	2	7	0.0000***
House Age (years)	26	52	0.0000***
Home Square Feet (linear feet)	2010	1566	0.0000***
Story (unit)	1.5	1.3	0.0000***
Fixed Bathroom (unit)	2	2	0.0000***
Half Bathroom (unit)	0.6	0.3	0.0000***
Property acreage (acres)	0.23	0.20	0.0000***
Median Income (dollars)	65722	43637	0.0000***

VI. Empirical Estimation

The equilibrium economic phenomena, as presented in Equation 1, underpins our general model for the first phase of the analysis. In the first phase, we incorporate empirical strategies of covariates, two-stage least squares (*2SLS*), and propensity score matching (PSM) strategies. These methodologies allow for a comparable and robust analysis of the influence of lead-risk, relative to the water supply, on the residential market. Our study second phase focuses on a sub-sample of one-time resale of residential homes in Fayette County to check for an asymmetric information relative to lead risk.

A. Hedonic Valuation of Lead Risk Phase

Our first empirical estimator forms a covariate approach to account for the implicit value of lead risk. The inclusion of explanatory variables could determine an unbiased estimate of the variable of interest. A linear hedonic functional form is chosen to estimate the hedonic valuation of lead exposure risk. The linear function form, although susceptible to omitted variable bias, will produce low mean percentage errors, Cropper *et al.* (1988) & Kuminoff *et al.* (2010).

$$P_i = \beta_0 + \beta_E E_i + \beta_H H_i + \beta_N N_i + \beta_Y Y_i + \varepsilon_i \qquad (6)$$

Equation 6 shows the relationship between the linear price P of household i, and the vectors of structural characteristics H, and neighborhood and environmental attributes N. E represents the proxy of lead-risk exposure to a house. Y is the household median income based on the associated Census Tract. Variables in the structural vector include the age of the home, square feet of the home, fixed and half bathrooms, story, property acreage, sale's month and year, and garage type. Variables of the neighborhood and environmental attributes include if the home is located within a historic district, the associated elementary district performance level, and the home's proximity to a water network or park. ε assumes the Gauss-Markov conditions of the idiosyncratic errors.

Our second empirical specification sketches an instrumental variable (*IV*) procedure on Equation 4. Since this study does not actually have a feasible variable for water quality, the two-stage least square (*2SLS*) approach, using strong instruments, is ideal to accurately estimate the influence of water quality on a home price. The first stage regress lead exposure risk E on the instruments of year prohibition dummy θ_i (i.e., if the home was built before 1998), the age of the house A_i, ZIP area Z_i, and the median household income Y_i. The final instrumental term Φ_i is constructed through the interaction of the instrumental variable.

$$E = \gamma_0 + \gamma_1\theta_i + \gamma_1 A_i + \gamma_2 Z_i + \gamma_3 Y_i + \gamma_4 \Phi_i + u_i \qquad (7)$$

In many States, for example, Kentucky and Illinois, lead poisoning is pointed out through the ZIP areas. Campaigns preventing children from lead poisoning do focus on ZIP areas with high risk of lead. These areas are used in formulating the level of lead risk in communities. Income is also instrumental in determining lead risks. Homes or communities with low incomes may be vulnerable to lead poisoning. Living below the poverty line, families and communities cannot finance or facilitate the replacement of lead materials or prevent lead poisoning in their private water systems. These families might hardly purchase water filters to treat their water supply. The age of the home may also contribute to a situation of no, low, or high lead exposure risk. Even though the 1986 amendments to the drinking water system prohibited the use of lead substances in public, residential, and nonresidential buildings, it was in the year 1996 it became illegal to use lead materials. So, homes built up to 1997 may have a high likelihood of being exposed to lead poisoning. In the final instrumental factor, the interaction term points out the interdependencies among the variables which could substantially contribute to the vulnerability of lead risk in the water system. These instruments justify the exogenous decisions of whether a homeowner will purchase a home with a probable low lead water supply or not.

$$P_i = \beta_0 + \beta_{\hat{E}}\hat{E}_i + \beta_H H_i + \beta_N N_i + u_i \qquad (8)$$

Building on the covariate strategy in Equation 6, the predicted $\hat{E}$ in equation 7 is used as a proxy for E in the second stage of the *IV* approach. Our instruments may not be perfect as it may be correlated with price. Even so, we assume that these variables are strongly correlated with the endogenous variable, $\hat{E}$. We also assume that the predicted lead risk variable is uncorrelated with the error term u_i.

Applying the dummies, which were constructed from the relatively low and high-risk lead neighborhoods, our final verification strategy for causality uses a propensity score matching (*PSM*). *PSM* assures that unexpected prediction can be removed from the observations. Homes located in the relatively high-risk area are positioned the treatment group, while homes located in the relatively low lead exposure areas are joined to the control group. Following Dawid (1979) conditional independence notation, $T \perp\!\!\!\perp X \mid U$, T is the treatment group and X and U respectively depict the observed and unobserved covariates. An elementary hypothesis of the PSM states that the assigned treatment group and the observed covariates are conditionally independent given the true propensity score.

$$X \perp\!\!\!\perp T \mid \pi(X) \qquad (9)$$

Equation *9*, from Rosenbaum and Rubin (1983) theorem, assumes that a matched treatment-control pair is homogeneous in the covariates $\pi(X)$. That is, the treatment and control homes in the lead risk zones will be matched based on the same distribution of X. Matching the true propensity score will result in the observed covariates of structural, neighborhood, and environmental characteristics being asymptotically balanced between treatment and control groups.

B. Testing for Asymmetric Information in Water Quality Phase

Addressing the concern of asymmetric information, we subscribe to empirical advice from Kurlat and Stroebel (2015). These authors test for information asymmetries in the real estate markets. Our data is not perfect to conform to the predictions put forth by Kurlat and Stroebel. An ideal dataset would present information on sellers and categorize the sellers by their level of information, more informed versus less-informed. However, we compute and test for asymmetric information through resale information. The resale value is a summation of the structural characteristics, the attractiveness of the neighborhood and environmental attributes, the loading factor of a house to its neighborhood, and idiosyncratic shocks. In conformance with the arguments from Akerlof (1970) and Kurlat and Stroebel (2015), at the time of resale, it is assumed that information about the value of the home is known. Home sellers are likely to acquire better and plentiful information, relative to knowledge on lead risk than potential buyers. For example, assuming that current homeowners and sellers are rational, they use information from their local water utilities and authorities like the EPA to get information on lead risks for their homes or neighborhoods. In the case of an asymmetric information, ceteris paribus, homes in relatively low-lead risk water zones might be better than is commonly known or reflected in the local housing price transactions. The same applies to the reverse. The reverse is: Homes in high lead risk water neighborhoods are worse off than commonly known or reflected in the home sale. Thus, due to hidden-type information, relatively high-lead risk water neighborhoods might be overrated or horizontally valued, compared to homes in the relatively low lead risk water neighborhoods. Problems with water in homes may not disclose by home sellers. These assumptions facilitate the building of a deterministic model, formulated in Equation 8, to check for asymmetric information.

$$Æj = \sum_{i=1}^{n}(V1/V0)^{1/t} - 1 \qquad (10)$$

Æ represents the average appreciating rate of resale homes in low lead risk water neighborhoods if *j* = 0, and average appreciating rate of homes in high lead risk water neighborhood if *j* = 1. V_1 is the resale price and V_0 is the initial or former price excepted by the buyer or offered by the seller. t represents the number of time homeowners occupied the property or engaged by the sellers. If *Æ* for low-risk water area is not statistically different from the high-risk *Æ* or not vertically higher, this would suggest an asymmetric information on water quality. Expressed differently, if the high-lead water risk *Æ* is higher and statistically different from the averaged low-risk *Æ*, this would also imply an asymmetric information. According to Kurlat's predictions, informed buyers in the housing market are able to select better homes at the same prices than uninformed buyers. With these means, informed buyers will be willing to pay for quality structure homes in a better neighborhood. On the contrary, uninformed buyers will be willing to pay more for houses in a relative overrated neighborhood. Notwithstanding, uninformed buyers can buy homes in both underrated and overrated neighborhoods. Again, we expect, on average, homeowners in high lead risk areas may have little information on lead in their water supply compared to low lead risk neighborhoods.

VII. Result

This section provides findings for the lead risk hedonic valuation and gives a report for the possibility of a hidden type information, with respect to water quality in Fayette County, during residential home sales, 2000 to 2017. Initially, in the hedonic valuation phase, we used the unit level of lead risk exposure in the water supply as our variable of interest. The later stochastic and deterministic models used the segmented high (treatment group) and low (control group) risk water zones variables. Results from the hedonic phase are compared to conventional water quality studies. Findings from the asymmetric results are linked to Kurlat & Stroebel (2015) to interpret uninformedness in the market.

A. Water Lead-Risk Influence on Residential Housing Price

Model 1, defined in Equation 6, allows this study to control for factors that might influence house price and our implicit variable of interest, lead risk in water quality. This specification permits us to determine housing values and the implicit water quality, as the lead risk increase or decrease by a unit. The result of the OLS specification is presented in Table 3. We used the robust error treatment to correct for functional and misspecification errors. Given the robust treatment, we do not analyze the percent of variation explained by the model. Still, we find an R-Squared of 72% in Model 1.

We find an unexpected result for water quality in Model 1, as reported in Table 3. Water lead-risk is not statistically significant at the p-value of 0.05. On this account, we reject the hypothesis that residential homeowners will pay a higher price for a water quality lower in lead risk. A background investigation, using a stepwise control technique, show that water quality variable alone, and time fixed effects of months and years, do not have a strong goodness of fit to explain the linear model. Water quality risk alone produced a significant result, but an R-square of 0.05; the inclusion of sales months produced an R-squared of 0.05, and sales year, 0.06. Additionally, controlling for the structural, neighborhood, and environmental attributes increase the explanatory power of the model to 72%, but nevertheless produced a non-significant result for the lead-risk variable.

Table 3. Estimates of Hedonic Model (OLS Estimator): Lead Risk in Water Quality

Dependent variable real price adjusted to year 2016 (mean: $219,328, Std. Dev: $148,676)				
Variable	Estimate	Robust Std. Error	t-stat	p-value
Water Quality (n=10 levels, μ = 4)	.204	.230	0.89	0.373
House age (μ = 35)	-.237***	.045	-5.26	0.0000
Median Income (μ = $57,559)	.32***	.03	10.33	0.0000
Time fixed effects	Yes			
House characteristics controls	Yes			
Neighborhood fixed effects	Yes			
School fixed effect	Yes			
Observation	70,619			
R^2	0.72			

We applied the Two-Stage Least Square (*2SLS*) estimator to examine the robustness of water quality risk. The *2SLS* estimator calculated that the quality of water, relative to lead risk, is identified by exogenous instruments of house age, median household income, lead prohibition fixed effects, and ZIP areas. Initial analysis of the instruments showed a positive correlation between house age and water quality exposure level at 0.69, and a negative correlation between median household income and water quality risk level at 0.66. These findings indicate a respective positive and negative associations of water risk with house age and median household income.

Table 4. Estimates of Hedonic Model (2SLS Estimator): Lead Risk in Water Quality

Dependent variable real price adjusted to year 2016 (mean: $219,328, Std. Dev: $148,676)				
Variable	Estimate	Robust Std. Error	t-stat	p-value
Water Quality (n=10 levels, μ = 4)	.164	.234	.70	0.483
Time fixed effects	Yes			
House characteristics controls	Yes			
Neighborhood fixed effects	Yes			
School fixed effect	Yes			
Observation	70,619			
R^2	0.70			

The coefficients of the instruments are statistically significant at 5% significance level in the first stage of the IV estimator, except for ZIP area 40504. House age is positive and statistically significant from zero. An increase in the age of the house increases in the level of lead risk in the water quality. An increase in the median income of household results in a decrease in water quality risk, a negative and significant result. As shown in Table 4, the result from the second stage also produces an unexpected result of water quality risk. An increase in water risk level increases property values. This estimate of water quality risk level, valued at $164, is positive and non-significant at 5% significant level. Besides interpreting the results from the *2SLS* estimator, we perform post-estimation tests to evaluate the uniqueness of the instrumental variable model. First, we perform the test of endogeneity, where the null hypothesis argues that the instrumental variables are exogenous. The robust chi-squared and regression p-values were 0.6227, thus failing to reject the null hypothesis. This indicates that the instruments were exogenous in nature, and the control specification (the OLS Model) did not suffer from endogeneity problems; it was not meant to treat water lead risk as an exogenous variable and using the *2SLS* was not necessary. Second, we test for the strength of the 1st stage. We hypothesized that the instruments were weak. The partial R-Squared was 0.57, rejecting the null hypothesis that the instruments were weak. Finally, we performed an over-identification (over-id) test for the *2SLS* model on the null that the set of instruments is valid, and the model is correctly specified. Findings from the over-id test, at a p-value of 0.0000, reject the null. This indicates an overidentification of the *2SLS* model.

In the final model application, which aims to evaluate lead risk hedonic price, we measure the consistency of the water lead risk estimate through the PSM model. Looking back on Table 2, the simple ANOVA test showed that the covariates of the treatment and control groups are statistically different. Obviously, these differences alarm the challenge of confounding factors between the treatment and control groups. The second check for confounders employed the standard difference test, using the standard deviations of the covariates means. Results from the standard deviation test also validate the possibility of confounders between the treatment and control groups.

At first, when we used the entire dataset of 70,619, we were unable to find a balanced match between the treatment and control groups. We failed on multiple attempts, despite trying techniques such as changing the functional form from linear to quadratic and cubic function. We applied probit and logistic regressions, and matching algorithm like the nearest neighborhood (NN) and Caliper & Caliper and Radius. For a successful matching, we reassessed the data and chose 4,352 observations for the propensity score. Note that, we used only continuous variables in new data for the PSM. Along with the property acreage control, all dichotomous covariates were dropped because these covariates could not balance the groups, or they violated the matching overlapping assumptions. We applied a probit regression on a common support matching algorithm to estimate the Average Treatment on the Treated effect (ATT). We regressed the treated groups on the covariates of age, age-squared, story, fixed and half bathrooms, and the median income for the household at the Census Tract level.

Table 5. Propensity Score Matching: Average Treatment on the Treated

	Treated	Controls	Difference	S.E.	T-stat
Average Treatment on Treated	182.820	193.921	-11.101	5.286	-2.10
Observation	1,289				
Pseudo R^2	0.32				

Holding all other variables constant, while controlling for the structural characteristics common to homes in the groups, Table 5 shows that there is a probability that residential homes in the high lead water risk neighborhood is devalued by $11,1010. This result, computed via the ATT model, is statistically significant. The conventional expectation for water quality is satisfied: In this model, we fail to reject the hypothesis of a higher willingness to pay for a quality water neighborhood. The ATT findings indicate an implicit negative valuation for neighborhoods or homes that are susceptible to high lead-risk. Table 6, the balance table, and Figure 4, the balance plot, tabularly and graphically represent the balanced matching. Primarily, the balance table ensures and communicates that our covariates for the treatment and controls were not different to promote confounding. Figure 4 visually reinforced Table 6 to show that our structural covariates were not biased, and we can trust the ATT estimate for the water quality valuation, in term of water lead risk. Approximately 29% of the sub-dataset (i.e., 1,289 observations) were matched, while 71% of the data was untreated by the propensity estimator.

Table 6. Balance Table for the Covariate Matching

Variable	Treated	Controls	Difference	Bias Reduction	P-value
Age	53	53	-0.28	97.8%	0.460
Square Feet	1744	1746	2.7	97.7%	0.436
Story	1.2	1.2	-0.001	92.0%	0.574
Fixed Bath	2	2	0.0172	99%	0.917
Half Bath	.4	.4	-0.002	98.4%	0.937
Median Household Income	41.72	40.97	-0.749	95.9%	0.074

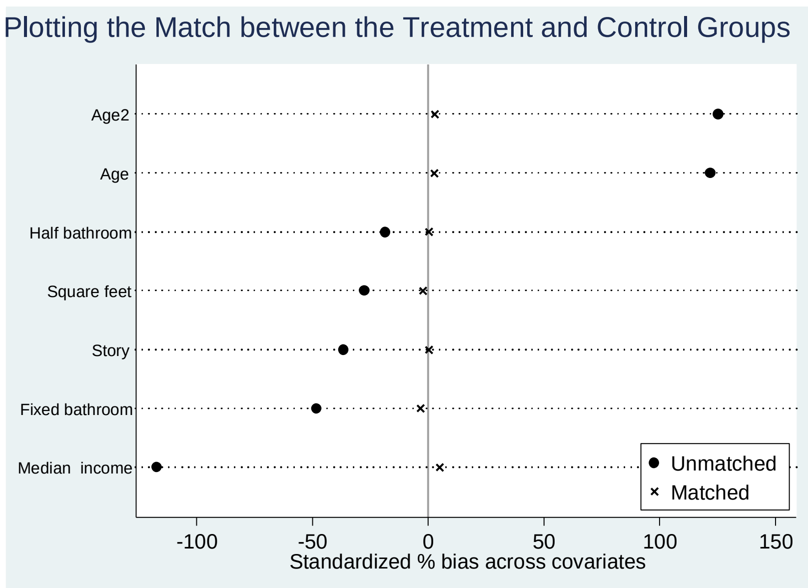

Figure 4: The reference line portrays a region of no or less biasness at the critical value of 0.05. Note: The p>chi² for the Propensity Score Matching in Table 6 has a value of 0.637. This result also portrays and validates a balance between the treatment and control groups.

Unlike the OLS and 2SLS models which are positive and not significant, the PSM estimate coheres with conventional findings on water quality. For instance, Leggett and Bockstael (2000) find a negative value for increased bacteria in water, and Piper (2003) support the argument that poor water quality leads to higher treatment cost and higher water rates. We join and support the argument of a positive willingness to pay for good water quality. Residents will positively value environmental attributes, including water quality which is free from harm or negative externalities.

B. Asymmetric Information in Water Quality

Given that home sellers are not legally responsible to disclose the water quality or problems of a home or neighborhood during a sale or resale, we expect that asymmetric information for hidden lead risk in water quality might be present. This might be true especially for homes in the high-risk water neighborhoods. We constructed a deterministic appreciation rate in Equation 10 to test for any sign of asymmetric information. Hence, we compare the appreciation rates between homes in the high-risk and homes in the low-risk water neighborhoods. To achieve this measure, we first extract a subset of 18,984 observations from the data. These observations are the unique first resales that occurred in Fayette County, considering the data period. During the data cleaning process, we dropped all resales that happened less than months from the former sales. Note, it is evident in the dataset that resale occurred up to eight times for some homes in the county between 2000 and 2018.

Table 7 reports the findings from the deterministic test. The deterministic test compared the average appreciation rates between the treatment and control groups. Validated by the p-value of 0.05 in Table 7, we fail to reject the null expectation that, on average, homeowners in high lead risk areas may have little information on lead in their water supply compared to low lead risk neighborhoods. Table 7 reports that homes in the high-risk water neighborhoods have a vertical appreciation rate relative to homes in the low-risk water zones, and these groups are statistically different from zero. The average appreciation rate for a probable high-risk area is about 52%, whereas the appreciation rate through resale for homes in the low-risk neighborhood is about 39%. The vertical difference in the appreciation rate is 13%. This finding is consistent with Kurlat and Stroebel's expectations and results. Informed buyers are likely to use their information in an overrated neighborhood, while uninformed buyers are incapable of distinguishing both the bad qualities both neighborhoods and homes. Hence, the deterministic test suggests the presence of information asymmetries relative to residents who live in a high lead-risk water neighborhood. The results imply that these residents are unable to detect, gain perfect information, or pay attention to the revealed information which would them a give gist of the level of lead risk in their water supplies. Although results from the deterministic model may suggest the presence of asymmetric information, the results herein are potentially biased because there might be important independent and explanatory variables we left out when we specified Equation 10. For example, adding the number of years homeowners or sellers occupied residential homes to Equation 10, we could expect the average appreciation rates to shift to the true values.

Table 7. Testing for Significance Difference between Appreciation Rates to Implicate Asymmetric Information for Lead Risk in Water Quality.

Dependent variable Log of real price adjusted to year 2016 (mean: 40.86%, Std. Dev: 75%)

Water Quality Neighborhood	Sample (Obs.)	Mean Appreciation	Std. Dev.	[Confidence Interval at 95%]	T-stat (P-value)
High-risk Pb Exposure (Level 6 – 10)	3,897	52.47	49.65	49.65 – 55.29	
Low-risk Pb Exposure (Level 1 – 5)	4,558	38.97	37.09	37.09 – 40.85	0.0000
Difference		13.50			

VIII. Conclusion and Implications

We argue that residential homeowners will pay more for improved water quality, and homeowners in the high lead risk water areas may have little information on their water supply relative to low lead exposure neighborhoods. Nevertheless, during the time of resale, home sellers are likely to obtain better information, relative to knowledge on lead risk, than potential buyers. We applied the hedonic analysis to empirically measure the effect of lead risk on housing values. The OLS, 2SLS, and PSM specifications were employed to validate the robustness of the implicit lead-risk value.

Accounting only for structural attributes, holding all things constant, we find that homeowners in the relatively low-lead risk water communities are implicitly willing to pay $11,101 to avoid the likelihood of being poisoned by lead. Analyzed in the deterministic model of appreciation rate, we find that the average appreciation rate in the probable low-risk area (52%) is higher than the appreciation rate for homes in the low-risk neighborhood (39%) by 13%. This difference suggests, ceteris paribus, buyers in the low-risk areas are better informed about lead risk than the high-risk buyers. Acquiring a higher appreciation rate for homes in the high-lead risk neighborhood could also imply that potential buyers are uninformed and incapable of distinguishing the quality of high lead risks in water neighborhoods. Findings from our stochastic and deterministic models may be vulnerable to empirical pitfalls and may violate regression assumptions, including omitted variable bias. Future studies could detect omitted variables in this research and add important variables to the models.

Acknowledgements

I thank Dr. Steven Buck, Assistant Professor of Agricultural Economics for his empirical mentorship, comments, and reviews that are being provided for the improvement of this manuscript. Much appreciation is also shown to Dr. Sayed Saghaian, Professor of Agricultural Economics, Dr. Jill Stowe, Associate Professor of Agricultural Economics, Dr. Jack Schieffer, and Dr. Carl Dillon for their technical assistance, comments, and reviews that greatly improved this research.

References

Akerlof, G. A. (1970). The market for" lemons": Quality uncertainty and the market mechanism. The quarterly journal of economics, 488-500.

Avoiding Lead Contamination in Public Drinking Water in Kentucky. Kentucky Division of Water,2017.https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwivm9rAwZjUAhXD6IMKHbsSAoEQFggoMAA&url=http%3A%2F%2Fwater.ky.gov%2FDrinkingWater%2FLEAD%2FConsumers%2FAvoiding%2520Lead%2520Contamination%2520in%2520Water%2520in%2520Kentucky.docx&usg=AFQjCNG2TL14ZJbyAotjuJRahJdSNF2pZw&sig2=MWgOZ-8U17o-8lIF0G-Bjw

Combs, Miranda. WKYT Investigates | Could the Flint water crisis happen here? (2016, 02, 04).http://www.wkyt.com/content/news/WKYT-Investigates-Whats-in-the-water--Lexington-response-to-Flint-MI-lead-disaster-367726981.html

CNBC. (2016, 03, 24). America's water crisis goes beyond Flint, Michigan. http://www.cnbc.com/2016/03/24/americas-water-crisis-goes-beyond-flint-michigan.html.

Buck, Steven, Maximilian Auffhammer, and David Sunding. "Land Markets and the Value of Water: Hedonic analysis using repeat sales of farmland." American Journal of Agricultural Economics (2014): aau013.

CNN. (2016, 06, 28). 5,300 U.S. water systems are in violation of lead rules. http://www.cnn.com/2016/06/28/us/epa-lead-in-u-s-water-systems/

Cropper, M. L., Deck, L. B., & McConnell, K. E. (1988). On the choice of funtional form for hedonic price functions. The Review of Economics and Statistics, 668-675.

Driscoll, Kara (2017, 04, 7). LIST: Does your community have lead pipes? http://www.daytondailynews.com/business/list-does-your-community-have-lead-pipes/DfgG0z5elogIQhRRpEw4cK/

EPA. (2017, 01, 19). Real Estate Disclosure. https://www.epa.gov/lead/real-estate-disclosure

EPA. (2017, 01, 19). Use of Lead Free Pipes, Fittings, Fixtures, Solder and Flux for Drinking Water. . https://www.epa.gov/dwstandardsregulations/use-lead-free-pipes-fittings-fixtures-solder-and-flux-drinking-water

Gassett, Bill. CNN. (2016, 12, 20). What Do You Have to Disclose When Selling a House. http://www.maxrealestateexposure.com/what-disclose-selling-house/

Hanley, N. S., & JF White, B. (2002). Environmental economics in theory and practice (No. 333.7 H241). Palgrave macmillan.

Jordan, J. L., & Elnagheeb, A. H. (1993). Willingness to pay for improvements in drinking water quality. Water Resources Research, 29(2), 237-245.

Jordan. (2016, 03, 11). Q&A: Stanford water expert on lessons of Flint, Michigan, crisis. Michigan.https://news.stanford.edu/2016/03/11/flint-water-luthy-031116/

Kurlat, P., & Stroebel, J. (2015). Testing for information asymmetries in real estate markets. Review of Financial Studies, 28(8), 2429-2461.

Lancaster, K. J. (1966). A new approach to consumer theory. Journal of political economy, 74(2), 132-157.

Leggett, C. G., & Bockstael, N. E. (2000). Evidence of the effects of water quality on residential land prices. Journal of Environmental Economics and Management, 39(2), 121-144.

Palmquist, R. B. (1982). Measuring environmental effects on property values without hedonic regressions. Journal of urban economics, 11(3), 333-347.

Piper, S. (2003). Impact of water quality on municipal water price and residential water demand and implications for water supply benefits. Water resources research, 39(5).

Powell, J. R., & Allee, D. J. (1990). The Estimation of Groundwater Protection Benefits and Their Utilization by Local Government Decisionmakers.

Prohibition on use of lead pipes, solder, and flux. Cornell Law School. https://www.law.cornell.edu/uscode/text/42/300g-6

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. Journal of political economy, 82(1), 34-55.

Shultz, S. D., & Lindsay, B. E. (1990). The willingness to pay for groundwater protection. Water Resources Research, 26(9), 1869-1875.

Rabin, R. (2008). The lead industry and lead water pipes "A Modest Campaign". American journal of public health, 98(9), 1584-1592. https://www.cdc.gov/budget/documents/fy2018/fy-2018-cdc-budget-overview.pdf

Snyder, C., & Nicholson, W. (2008). Microeconomic Theory-Basic Principles and Extensions; internat. student ed. Mason, Ohio (Thomson South-Western).