May 2018

Three Essays on Public Policy and Health

Jessica E. Sauve-Syed
Syracuse University

Abstract

This dissertation is comprised of three essays on health and public policy. My research examines how regulations and targeted programs impacted health outcomes.

The first two chapters use the unexpected shock of water borne lead in Flint Community Schools' classrooms to estimate the causal impacts of lead on students' cognitive and behavioral outcomes. The first paper focuses on the short-run academic achievement of elementary school children. At the average level of lead exposure, conservative estimates find that the share of students scoring proficient on statewide standardized exams drops by 6 to 9 points in math and drops by 12 to 14 points for reading. This represents an average of 3 to 5 students for a typical grade within a school. The second paper estimates the impact of lead exposure on student behavior. The analysis finds that during the Flint Water Crisis a typical grade within a school receives 8.6 additional disciplinary actions per year at the average level of lead exposure.

The third chapter focuses on the perennial topic of access to healthcare and health outcomes. Access in this case is measured by the supply of primary care physicians. The federal Health Professional Shortage Area initiative identifies underserved areas and makes them eligible for incentives to attract physicians. Using propensity score matching methods, an average 3 percent decline in mortality rates is found for areas that receive the designation.

Three Essays on Public Policy and Health

By

Jessica E. Sauve-Syed
B.A., Syracuse University, 2009

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Economics

Syracuse University
May 2018

Acknowledgements

I would like to express my deepest appreciation to my primary advisor Jeffrey Kubik for his guidance and support throughout this process. I am also very grateful to Gary Engelhardt for his advice and feedback along the way. I am thankful to my dissertation defense committee: Elizabeth Ashby, Gary Engelhardt, Jerry Evensky, Jeffrey Kubik, Perry Singleton, and the chair, Sarah Hamersma. I also thank Jerry Evensky and Inge O'Connor who have been supportive in achieving my goals.

On a personal note, I thank my close friends and family who have provided words of encouragement and share in my joy at this completion. I am forever grateful to my parents, Brent and Sandy Sauve, who have always valued education and supported my dreams. Finally, the love and support of my husband, Syed Adnan Alli, has been invaluable. I am appreciative of his continued patience and encouragement.

Table of Contents

List of Figures

List of Tables

Chapter 1

Lead Exposure and Student Performance: A Study of Flint Schools

1.1 Introduction

There is a large literature in public health that links childhood lead exposure to negative cognitive impacts. The findings suggest that even low levels of lead exposure may result in permanent deficiencies. Therapies designed to mitigate the effects of lead after exposure appear unsuccessful in reducing the severity of the cognitive declines (Kaufman (2001), Canfield et al (2005), Shih et al (2007)). Lead ingestion often occurs at an early age and may have considerable ramifications for life long success.

This evidence has motivated economic studies related to human capital, including the impacts of early childhood lead exposure on elementary school standardized exams and lifetime earnings (Cohen-Cole (2006), Rau et al (2013), Reyes (2015), Aizer et al (2017)). These studies rely on variation in lead exposure between households or neighborhoods to calculate the effects on long-run outcomes. Unfortunately, student performance and lifetime earnings are also correlated with household and neighborhood characteristics that are not easily observed. One of the greatest challenges in this literature has been disentangling the causal impacts of lead exposure from endogenous household and neighborhood effects.

This study overcomes this challenge with unique data from a quasi-natural experiment in Flint, Michigan that resulted in an exogenous shock of lead exposure at the classroom level. The City of Flint faced a looming financial crisis and was appointed an emergency financial manager by the governor of Michigan from 2011 to 2015. During this time costs were cut across the city to address the large deficit. In early 2014, the City of Flint stopped sourcing its municipal water from the City of Detroit and started locally treating the Flint River to save money. This switch resulted in highly corrosive water being supplied to homes, businesses and public buildings until October of 2015. Lead found in service lines connecting buildings to the municipal water system as well as lead components in water fixtures leached into the water. The localized lead sources resulted in heterogeneous levels of lead-in-water at the fixture for classrooms within school buildings.

An important contribution of this study is the exogenous source of lead variation to estimate the causal impact of lead on student achievement. Students of Flint Community Schools were randomly exposed to different levels of lead-in-water within schools based on their assigned classroom. Fountains within elementary schools appear similar, but the variation in the distribution of lead within fountain components as well as soldered pipe joints contributed to different levels of lead-in-water across classrooms. Information requested from Flint Community Schools is used to link proficiency measures from state standardized exams to classroom lead-in-water levels. A novel measure of lead exposure is created to account for variation in the cumulative stores of lead in students' bodies. Cohort panel data within each elementary school is used to study the same students over time to estimate the short run impacts of lead exposure. This builds on the work of previous studies which relied on variation between cohorts before and after the passage of lead abatement policies to estimate the effect of early childhood blood lead tests on academic achievement in elementary school.

This study also provides valuable insight into the short run effects of lead on the academic achievement of older children. There is very little work that focuses on this age group. The physical symptoms of lead exposure are rare in elementary school children, so there has been less urgency to study these students. The lack of visible symptoms may provide a false sense of security if there are severe cognitive consequences.

Significant, meaningful impacts on student proficiency are found using the exogenous shock in lead exposure with a fixed effects specification. At the mean level of cumulative lead exposure, the estimated impact on share proficient in reading is -12 to -14 percentage points. The share proficient in math dropped -6 to -9 percentage points. The impact on share not proficient in reading is an increase of 13 to 15 percentage points. The magnitudes of these changes reflect an average of three to five students in a typical school cohort. A larger impact on reading proficiency as compared to math is consistent with patterns found in previous studies in the economic literature.

The education literature has historically focused on production variables, such as class size, per pupil spending, teacher quality, and student nutrition. These factors only explain a small portion of the gaps in student achievement. This study provides evidence of an important classroom environmental factor for student success that has not been previously studied. The education production literature suggests that per pupil spending would have to change by 10 to 20 percent to impact proficiency shares by similar magnitudes (Papke and Wooldridge (2008), Kesler and Munkin (2015)).

This study adds to a growing literature that tests the effects of environment and health on academic outcomes. School level analyses suggest that air quality and temperature influence student test scores (Stafford (2015), Cho (2017), Park (2017), Marcotte (2015) (2017)). This study utilizes within school variation to explore another dimension of school environment, classroom water quality. Health shocks and access to early childhood healthcare have traditionally been used to estimate the effects on future attainment (Case et al (2005), Currie (2009), Levine and Schazenbach (2009), Contoyannis and Dooley (2010), Robinson and Coomer (2014)). By comparison, this paper studies a decline in cognitive ability during and shortly after the period of lead exposure.

The findings suggest that school infrastructure can be important for academic success. A small literature has estimated the effects of school bonds on student achievement and found little impact (Callini et al (2010), Hong and Zimmer (2016), Martorell et al (2016)). However, these previous studies do not differentiate between different school improvement initiatives. Targeted maintenance as well as regular water analysis may be valuable to student success. Corrosion of brass components that contain lead can result in elevated lead-in-water for both old and new buildings. A timely response to elevated levels of lead may be important to ensuring a safe and productive academic environment.

The results of this study are measured in the short run, and it is unclear whether the declines in achievement will persist. The previous literature suggests that the impacts from elevated blood lead levels in young children have permanent consequences for human capital development. However, there are developmental differences between young children and elementary school students. Even if the cognitive deficiencies are temporary, it is possible that short-term lags in achievement may impact future attainment if students are not provided the resources needed to catch up.

The paper continues with the following sections. Background information is provided on the Flint water crisis, the regulation of lead and its effects on children, as well as recent efforts to estimate the impacts of lead on student achievement. Then the data are described and details are provided about the construction of the variable of interest, cumulative blood lead exposure. The estimation strategy and main results follow. The robustness of the treatment variable is tested with an alternative measure of lead exposure. Finally, a brief discussion about the results and a conclusion are offered.

1.2 Background

1.2.A The Source of Flint's Water Crisis

In 2013, the City of Flint decided to break its water supply contract with Detroit Water and Sewage in favor of a new pipeline that was under development. This triggered a one-year notice of termination for Flint's current water supply contract[1]. A decision was made to treat the local Flint River while the new supply lines were constructed. The City of Flint's water treatment plant had the capacity to treat water in the interim, but it had not been regularly utilized in nearly 50 years (Davis et al, 2016).

The water supplied from Detroit had been regularly treated with orthophosphate for more than twenty years. This additive to the water builds a protective passive layer inside pipes to prevent corrosion. After the change in water sources the City of Flint chose not to continue with the orthophosphate or utilize other corrosion inhibitors. This made the passive layer susceptible to flaking and exposed the pipes to corrosion (Torrice, 2016).

Surface waters, such as rivers, are naturally corrosive and have more organic materials in them. They also require more treatment to remove particles and microorganisms[2]. To kill off microorganisms chlorine is added to the water as a disinfectant. In the City of Flint the chlorine reacted with iron in the pipes and caused corrosion. The corrosion process consumed the chlorine, leaving bacteria in the water and exposing lead in the pipes.

Killing off microorganisms with the chlorine increased the organic material in the water. To remove the contaminating materials ferric chloride was added as a coagulate to assist in the filtering process (Torrice, 2016). This combined with the already corrosive water to cause

[1] "Flint Water Crisis Fast Facts." *CNN.* http://www.cnn.com/2016/03/04/us/flint-water-crisis-fast-facts/index.html (accessed September 1, 2017).
[2] Olson, T. 2016. "The science behind the Flint water crisis: corrosion of pipes, erosion of trust." *The Conversation,* January 28. https://theconversation.com/the-science-behind-the-flint-water-crisis-corrosion-of-pipes-erosion-of-trust-53776

chloride levels to soar. High chloride-to-sulfate ratios in water are known to be very corrosive to lead (Edwards and Triantafyllidou, 2007). In general, a chloride-to-sulfate ratio of .58 is considered an upper bound for water management. Researchers from Virginia Tech sampled treated water with ratios as high as 1.6 in Flint during the water crisis.

Corrosive water can leach lead from many plumbing sources. Lead service lines and lead solder in older homes are obvious origins, but lead is also used in brass components for water fixtures including faucets, fountains, valves and other components. Analysis provided by the Michigan Department of Environmental Quality reported elevated levels of lead in the Flint Community Elementary Schools' water and cite water fixtures as the probable sources[3].

1.2.B The Dangers of Lead and Its Impact on Children

Lead in the water has been a target of federal regulations for nearly half a century. Under the 1974 Safe Drinking Water Act, the Environmental Protection Agency (EPA) set a non-enforceable maximum containment level goal of zero for lead, signaling that there are no safe levels of lead exposure[4]. Starting in 1986 lead pipes and lead solder were banned from new water and plumbing systems. Lead content in brass plumbing components was limited to eight percent. This followed the restriction of lead in residential paint (1978) and the transition away from leaded gasoline, two other common sources of lead exposure. The Lead and Copper Rule in 1991 addressed issues of corrosion control and set an actionable level for lead in public water at 15 parts per billion (ppb) at the 90th percentile for customer taps.

The regulations expanded in response to a growing literature of the health impacts of lead. Lead is a toxin that can cause numerous medical issues. At high levels, it may even cause

---

[3] Taking Action on Flint Water. 2015-2017. "School Testing." State of Michigan. http://www.michigan.gov/flintwater/0,6092,7-345-76292_76294_76297---,00.html (accessed September 15, 2016).
[4] "Basic Information about Lead in Drinking Water." *United States Environmental Protection Agency.* https://www.epa.gov/ground-water-and-drinking-water/basic-information-about-lead-drinking-water (accessed on August 28, 2017).

death. Once lead enters the body it is absorbed into the blood. Absorption rates vary by age, with young children absorbing up to 50% of consumed lead and adults absorbing as little as 10%. Pharmacokinetic models of lead assume children as young as 8 years old have the same absorption rates as adults. These models attempt to quantitatively model biological relationships and the impacts of lead but have often had to rely on limited observational data to calibrate absorption rates. As a result, the broader medical field has traditionally assumed that older children are not as susceptible to the impacts of lead (Toxicological Profile of Lead, 2007).

Lead mimics calcium and iron, leading to greater absorption in people with nutrition deficiencies. Once absorbed into the blood lead passes through the circulatory system where it becomes deposited in tissues and bone. The half-life for lead in the blood is 30 days, while for tissues and bones the half-life is years or even decades (Toxicological Profile of Lead, 2007). For this reason, lead can continue to cause issues in the body long after blood lead levels have dropped to within a normal range.

Children are especially vulnerable to problems caused by elevated blood lead levels. Their developing bodies absorb lead more easily than adults. Of particular concern is the underdeveloped blood brain barrier. Lead's ability to pose as calcium and pass into the brain presents a high risk during neurological development. This substitution impacts the creation of synapses and neurotransmitters in the brain. Long-term behavioral and intellectual deficiencies have been observed at relatively low levels of lead exposure. The blood brain barrier is still developing into the second decade of a child's life, posing a particularly large risk for growing children with nutrient deficiencies (Lidsky and Schneider, 2003).

Young children also run a greater risk of exposure. Toddlers can mistakenly put dust and chips from lead paint in older homes into their mouths. In 1991, the CDC recommended universal blood lead testing for all young children. Many states have since adopted blood lead screenings and tests during routine children's appointments. In 1991, a blood lead level of concern was greater than 15 micrograms per deciliter, down from 60 in 1960. As scientific studies have found negative impacts from lead at lower levels, the blood lead level of concern has also dropped. In 2012 the CDC decided to discontinue the use of this system and adopted a policy that all blood lead levels are concerning. Now a reference value of five has been adopted as a trigger for health interventions (Advisory Committee on Childhood Lead Poisoning Prevention, 2012).

Public resources have focused on prevention and testing for young children due to their greater risk of exposure. This has resulted in a large volume of data for blood lead levels of this age group, and they have subsequently been the focus of many studies. There is reason to believe that elementary school students will behave differently than the young children since they are biologically more developed. Historically, medical models have treated them as fully developed. However, they are not simply small adults as their bodies and brains are still developing into their second decade (Lidsky and Schneider, 2003). Potential deficiencies caused by lead exposure later in childhood can impact students' learning and their subsequent human capital development. Their loss of opportunity and truncated success have broader implications for societal welfare and issues of equality.

1.2.C Lead and Student Achievement

Early studies look at both long-term and short-term impacts of early childhood lead exposure and intellectual outcomes. Bellinger and coauthors (1990) measured lead exposure through deposits of lead in children's teeth. Their analysis finds negative impacts in elementary school as well as high school academic achievement. For more contemporaneous outcomes, Chen and coauthors (2005) find that for children age two to seven concurrent blood lead levels became more strongly associated with IQ as the children grew older. In recent years access to data from universal lead testing programs and student achievement records has renewed interest and resulted in new estimates.

Zhang and coauthors (2013) claim to be the first to test the relationship between early childhood blood lead levels and classroom achievement. Confidential data from Detroit Public Schools linked to early childhood blood lead tests suggest that high blood lead levels have an odds ratio of 2 or greater for children to be labeled as not proficient on the third, fifth, and eighth grade Michigan standardized exams. The authors admit the study suffers from potential confounding issues with only limited household data available to control for unobserved characteristics.

Third graders were also studied by Reyes (2015) as well as Aizer and coauthors (2017). Large drops in childhood blood lead levels in the late 1990s in Massachusetts were used to test the relationship between childhood lead exposure and standardized exam results. The strongest relationship suggests that a one percentage point increase in the share of students with blood lead levels over 20 micrograms per deciliter is linked to a one percentage point increase in the share of students scoring unsatisfactory (Reyes, 2015). While interesting, a blood lead level of 20 is no longer common with the 97.5th percentile of early childhood blood lead levels less than 5 (Caldwell et al., 2017).

A Rhode Island lead certificate policy targeting leaded paint in rental units caused a sharp decrease in early childhood blood lead levels. Aizer and coauthors (2017) use an instrumental variable strategy to avoid the issues of endogeneity from previous studies. The probability of having a lead certificate at the time of childbirth based on census tract and family characteristics is instrumented for the child's blood lead levels on their early childhood blood test. An increase in blood lead levels by 1 unit is estimated to increase the probability of being below proficient in third grade reading by 3.1 percentage points and in third grade math by 2.1 percentage points.

These studies provide fresh insight and support the broad assertion of an inverse relationship between early childhood blood lead levels and later academic achievement. They are also limited by the cross-sectional nature of the data as well as broad treatment measures. Household and neighborhood factors are correlated with educational outcomes and often children's health. Blood lead levels are often higher for children from disadvantaged backgrounds.

The use of early childhood blood lead levels may not be a good measure for a child's exposure to lead. Lead has a relatively short half-life in the blood, just over 30 days, so annual tests may not accurately reflect exposure from earlier in the year. At the same time, lead that is deposited into soft tissues has a much longer half-life and can continue to impact a child. Aizer and coauthors (2017) mention potential issues with measurement error in blood lead levels, particularly the use of capillary samples. According to a CDC report, the measurement error on a child's blood test is approximately three. Most children who are tested have blood lead levels less than five. The variation due to measurement error may cause issues when studying the low levels of lead common in the post lead abatement program period (Advisory Committee on Childhood Lead Poisoning Prevention, 2012).

This study adds to the existing literature by focusing on lead exposure to elementary school children in third through sixth grades. During this period the absorption and metabolizing of lead is believed to be different from that of young children, but the brain is still undergoing development. This analysis distinguishes itself from most of the academic achievement literature by focusing on both older children and the short run impacts of lead ingestion. It also departs from the more recent literature by focusing on lead-in-water exposure as opposed to lead paint abatement programs. While this difference may impact initial absorption rates into the blood, it does not change how lead diffuses throughout the body and impacts various organs. Finally, it also uses a unique exogenous shock in lead exposure to identify the impact of lead on academic achievement rather than relying on heterogeneous impacts of public policy.

1.3 Data

1.3.A Student Achievement Data

The student achievement data for this study come from the Michigan Department of Education's Center for Educational Performance and Information. Data files can be downloaded from the MI School Data web portal[5]. Proficiency levels are used to measure student performance. For the 2012-2013 and 2013-2014 academic years, the Michigan Educational Assessment Program (MEAP) standardized test results are utilized; these are the same examinations used by Zhang and coauthors (2013) in their analysis of Detroit students. For the 2014-2015 and 2015-2016 academic years, the Michigan Student Test of Educational Progress (M-STEP) assessment is used to measure proficiency. The change in examinations reflects statewide reform in student assessment by the Michigan Department of Education. Standardized exams measuring student proficiency are administered prior, during, and after the water crisis. A timeline showing exposure to lead-tainted water and the standardized testing schedule is available in Figure 1.1.

Math and reading tests are administered to third through sixth grade students under both regimes. The proficiency results are reported within schools at the grade level. Students are categorized into four levels of proficiency based on the number of points earned under both the

[5] The portal can be accessed from: https://www.mischooldata.org/Default.aspx

MEAP and M-STEP standardized tests. Of the four categories, students who are labeled advanced or proficient are grouped together to generate the share proficient for this study. Students who fall in the bottom tier are labeled not proficient and compose the share not proficient for this study. Further information on the State of Michigan's Student Performance Levels is available in Appendix Table A1.1.

1.3.B Classroom Lead-in-Water Data

Information about lead-in-water levels for classroom drinking fountains is taken from the Michigan Department of Environmental Quality (DEQ)'s Outlet Sampling and Plumbing Assessments conducted from late October 2016 to December 2016[6]. Water samples were taken from each of the operating faucets and fountains in the Flint Community Schools. Four samples over a period of 150 seconds were drawn from each source following a 12-hour stagnation period.

A cumulative blood lead exposure measure is used as the treatment variable in this analysis. This variable is different from the blood lead levels typically used in studies of student achievement. The benefits of the traditional measure include large sample sizes and individual level data. However, blood lead level studies suffer from the short half-life of lead in the blood, which is approximately 30 days. The half-life is important because blood lead levels may fall while there is still significant lead deposited in the organs impacting student success. Analyses that focus on low levels of lead in the blood are also susceptible to measurement error, with the 97th percentile blood lead level less than five micrograms per deciliter and the accuracy of most laboratories within a variance of two micrograms (Caldwell et. al, 2017).

[6] Taking Action on Flint Water. 2015-2017. "School Testing." State of Michigan. http://www.michigan.gov/flintwater/0,6092,7-345-76292_76294_76297---,00.html (accessed September 15, 2016).

This study's cumulative blood lead exposure variable reflects total lead exposure throughout the treatment period. By construction, it grows larger over time. This measure is a better reflection of potential lead in the soft tissues and organs during the period of analysis due to the relatively long half-life of lead in the organs and the short run outcomes of this study. The measure also provides real variation over time for school-specific cohorts.

The primary cumulative blood lead exposure variable is generated using all of the water information collected by the Michigan Department of Environmental Quality. The lead-in-water levels are averaged over all four water samples at the classroom level. This is done to most accurately reflect the levels of lead students may have ingested throughout the day. Achievement data is available at the school-specific cohort level, so a weighted average is taken of the classroom lead-in-water values based on the share of students for a cohort in each particular classroom. Assumptions about the volume of water and absorption rate are made to find the blood lead exposure for one day. Examples of daily blood lead exposure calculations are available in Appendix Table A1.2.

A challenge for this study is that classroom water is only tested at the end of the exposure period. Students had access to the water for many months leading up to the water test, and it is unlikely that the lead-in-water values found by the DEQ were static during the duration. One of the exams measuring student performance was proctored in the middle of this exposure period. The water at that point likely had less lead than when it was tested several months later. To capture the variation in exposure to lead-in-water over time, the cumulative blood lead exposure variable is created. Data about the average weekly chemical composition of the treated Flint River from the water treatment plant and the science behind lead leaching into water are used to calculate average weekly lead-in-water values for each classroom. This is then combined with the academic calendar, information about lead absorption, and annual classroom homeroom assignments to create the cumulative blood lead exposure variable for each school-specific cohort. A complete, detailed explanation of how this variable was generated is available in Appendix A1.1 of this paper.

1.3.C Summary Statistics

Classroom water data from the Department of Environmental Quality (DEQ) is matched to student achievement data by classroom number using information obtained from Flint Community Schools through a Freedom of Information Act (FOIA) request. The request included details about primary classroom usage and class size. The classroom information spans four academic years and covers the eight elementary schools that operated continuously throughout the period of study.

Of the eight elementary schools, only six were successful in matching classroom numbers to classroom grade assignments. The room numbers used in the DEQ report for Holmes STEM Academy do not match the room numbers provide by the school district. Based on the DEQ report for Durant Tuuri Elementary drinking fountains were not available in the classrooms occupied by the students of this study. The remaining six elementary schools used in this analysis include Doyle-Ryder Elementary, Eisenhower Elementary, Freeman Elementary, Neithercut Elementary, Pierce Elementary, and Potter Elementary.

Summary statistics are available in Table 1.1. The left two columns provide the statistics for Flint Community Schools before and after the change in the water. Column (1) shows an assumed cumulative blood lead exposure level of two and the average shares proficient and not proficient for math and reading for the two years prior to any change in the water source. Column (2) shows the average values for these variables for the two years following the change in water. In column (1), Flint students tend to do significantly better in reading than math on the standardized test. Comparing these values to column (2), the share proficient drops for both subject areas and the share not proficient rises. The changes are more than twice the magnitude for reading than for math. The average cumulative blood lead exposure level is 49 and has a large standard deviation, reflecting a large variance in the sample.

Columns (3) and (4) provide the summary statistics for the comparison school districts before and after the change in Flint water. The cumulative blood lead water variable is assumed to be two throughout the period of study. Math and reading proficient shares both drop in the post-water period, suggesting the change from the MEAP to the M-STEP impacted the share proficient more broadly. For not proficient, the reading share grew while the math share dropped slightly. Comparing the pre-period for Flint Community Schools to the comparison schools shows that the students in Flint generally performed worse on the exams prior to lead exposure. The gap is consistently about 10 points for all four proficiency measures.

Figure 1.2 shows the distribution of the cumulative blood lead exposure variable for Flint Community Schools at the time of the standardized tests in the 2014-2015 and 2015-2016 academic years. Most of the observations are less than 50 micrograms, but the right tail extends out to 190. Figure 1.3 presents the average cumulative blood lead exposure by year. From the graph, it is clear that the largest change is between the 2013-2014 and 2014-2015 academic years. A slight increase in the exposure level for the 2015-2016 academic year reflects exposure in May 2015 following the M-STEP exams and exposure in September 2015 prior to when the water fixtures were no longer accessible to students.

Control schools are used in the estimation of the model. The schools are carefully selected to ensure that they are comparable to the Flint Community Schools. The Michigan

Department of Education uses an algorithm to generate peer districts for comparison purposes on their Michigan School Data website. Schools are matched at the district level on a handful of characteristics, including the size of the student population, student-to-instructor ratio, instructional spending per pupil, the share of students receiving free lunch, and geographic distance from the other districts. Elementary schools from four school districts were chosen as the comparison group. Appendix Table A1.3 summarizes the matching characteristics for each of the districts as well as Flint Community Schools.

1.4 Empirical Strategy

This study utilizes an unexpected change in water quality to estimate the relationship between cumulative blood lead exposure and student academic achievement. The essentially random allocation of students within schools to classrooms with elevated lead-in-water drinking fountains provides an opportunity to address the issues of endogeneity found in other studies. Figure 1.4 provides an example of the within school, across classroom variation of lead for one of the elementary schools in the study.

Panel data analysis is utilized for within school-specific cohorts. The cohorts are identified by anticipated year of graduation. A fixed effects approach that controls for time invariant characteristics within schools, cohorts, and school-specific cohorts addresses many of the issues that confound other studies. The general model used in this study is highlighted below:

$$Achieve_{c,s,t} = a_0 + \beta_1 Lead_{c,s,t} + \eta_c + \kappa_s + \eta_c{}^*\kappa_s + \tau_t + \mathcal{E}_{c,s,t}$$

The dependent variable, Achieve, is measured as share proficient (not proficient) for cohort *c* in school *s* in year *t*. The variable $Lead_{c,s,t}$ is the cumulative blood lead exposure variable, measured in micrograms. The coefficient β_1 is the estimated impact of lead exposure on academic achievement and is the coefficient of interest. The regression includes cohort fixed effects η_c to control for differences across cohorts, such as test difficulty. There are school fixed effects κ_s to control for differences across schools, not limited to the quality of teachers and the quality of surrounding neighborhoods. The inclusion of cohort by school fixed effects controls for variation in student quality across cohorts within schools. In the basic regression τ_t is year fixed effects to control for annual shocks.

All the estimated standard errors are clustered at the school-specific cohort level. Clustering controls for correlation among groups in the error term and is important for obtaining accurate standard errors. The motivation for clustering comes from the potential autocorrelation between the error terms from consecutive time periods for a school-specific cohort. While there is also potential for error correlation within schools, the small dataset has only six schools to cluster which is too few for accurate statistical inference (Cameron and Miller, 2015).

The variables $Achieve_{c,s,t}$ and $Lead_{c,s,t}$ are winsorized at the 5^{th} and 95^{th} percentiles to help control for outliers. The form of the relationship between lead and academic achievement is unclear in the literature, so estimates are provided in both level and log forms for completeness.

Identification of the effects of cumulative blood lead exposure is driven by variation in the intensity of lead exposure and standardized testing results within each school cohort over time. Fixed effects for schools, cohorts, and school-specific cohorts absorb variation caused by static differences among these groups. Year fixed effects and school-specific linear trends control for broad changes impacting all students or students within a school over time. The estimates subsequently reflect the covariance between intensity of cumulative blood lead exposure and movement in proficiency shares within school cohorts over time.

There are several assumptions necessary for identification in this study. First, the exogenous shock of water quality within the classroom cannot be correlated with the quality of students assigned to those classrooms. There were no visible signs indicating which classroom water fountains would contain higher levels of lead based on the analysis conducted by the Department of Environmental Quality and photos of the drinking fixtures.

The shock of water quality within the classroom also cannot be correlated with the level of exposure outside the classroom. The corrosive waters of the treated Flint River caused elevated lead-in-water levels throughout the city, not just in the elementary schools. This study assumes that school-specific cohorts of children that are exposed to higher than average lead-in-water levels at classroom fountains are not systematically exposed to higher than average lead-in-water at home. In general, students in a public elementary school come from the same neighborhood, and often siblings who share the same home living situations are enrolled in different grades. Based on this reasoning, this study assumes that the lead exposure out of school is a random draw among cohorts within the same elementary school.

A third assumption is that the change in state standardized testing is not correlated with the level of lead exposure within classrooms. Such correlations seem unlikely based on the structure of the data. School fixed effects control for general differences in quality across schools. If the exam has become relatively more difficult for a grade, the cohorts across the six schools in that grade would need to experience relatively higher levels of lead exposure at their respective classroom levels. All evidence suggests that the level of lead-in-water at the classroom level is random within schools.

The results start with the basic regression, which is a two-way fixed effect model estimated with a set of control schools. The inclusion of year fixed effects can control any potential shocks correlated with the treatment variable that are also related to student achievement. The analysis then turns to specifications using only Flint Community Schools data. The two-way fixed effects regression is re-estimated. Next, structure is added to the model to preserve variation with linear trends replacing the year fixed effects. The sensitivity of the treatment variable is tested by re-estimating the model with a maximum lead water measure.

As a preview, the results suggest larger impacts for the reading proficiency measures than for the math. Using the mean post treatment lead-in-water value, the impact on share proficient using level and log treatment is -6 to -9 points for math and -12 to -14 points for reading. Similarly, the impact on the share not proficient is 13 to 15 points for reading.

1.5 Results

1.5.A Primary Results

The basic regression is estimated from academic year 2012-2013 through academic year 2015-2016. The fixed effects regression uses variation from within school-specific cohorts to calculate the impact of cumulative blood lead exposure on proficiency shares in math and reading. This approach utilizes the greatest within cohort variation in lead exposure, occurring between the 2013-2014 and 2014-2015 academic years. The lead exposure variable and share (not) proficient variables are all winsorized at the 5^{th} and 95^{th} percentiles to temper the influence of outliers. To calculate the impact of these estimates the median cumulative blood lead exposure value is 27 micrograms and the mean is 49 micrograms.

Table 1.2 shows the impacts of cumulative blood lead exposure on student proficiency shares using the basic regression. Starting from the left, columns (1), (2), (5), and (6) use level values of the treatment variable and columns (3), (4), (7), and (8) use a logged treatment variable. The cumulative blood lead levels are skewed right, and the logged treatment can reduce the importance of outliers. The even columns – columns (2), (4), (6), and (8) – are estimated with the inclusion of school-specific linear trends.

Panel A of Table 1.2 shows the results for share proficient for both math and reading. All of the estimates are negative as expected, suggesting that an increase in lead exposure causes a decline in student achievement. The inclusion of the school linear time trends results in more significant estimates. The impacts on share proficient in reading in columns (6) and (8) are statistically different from zero and significant at around the five percent level. The estimates for math are noisier, with only the estimate in column (2) statistically different from zero. The impacts on the share proficient in math is -5 points at the median level of cumulative blood lead exposure and -9 points at the mean. For reading, the estimated impacts of level and logged treatment for share proficient are -8 to -10 points at the median and -12 to -15 points at the mean.

Panel B provides the corresponding results for share not proficient. The estimates for math are statistically different from zero while the results for reading are more significant. The impacts of level and logged treatment on share not proficient in math at the median are 5 to 6 points and 7 to 10 points at the mean. For reading the values are 9 to 12 points at the median and 15 to 17 points at the mean.

The estimates in Table 1.2 rely on the assumption that students in the comparison schools are similar to the students of Flint Community Schools. Comparison schools were chosen from a list of peer school districts identified by the state of Michigan. The variables used to select the comparison schools are available in Appendix Table A3.1. To relax the assumptions surrounding the comparison schools, the model is estimated again using achievement data only from Flint Community Schools.

1.5.B Estimating Within Flint

The basic model is run again using data from Flint Community Schools for 2012-2013 through 2015-2016. The results are shown in Table 1.3. The impacts found in Table 1.2 have disappeared. Nearly all of the estimates are not statistically different from zero. Moving from the full specification to estimating within Flint Community Schools resulted in a significant drop in observations. This decreases the power of the regression. The fixed effects use a lot of the useful variation in the observations, so it is not surprising that these estimates are noisy. The flexibility of the year fixed effects is not practical for this sample. A more restrictive specification that replaces the year fixed effects with a linear time trend is estimated.

The new model is estimated in Table 1.4. Panel A shows the results for share proficient in math and reading. The regressions that include the school-specific trends are all significant at the one percent level. The estimates for math are slightly less than half the magnitude of the estimates for reading. The impacts for level and logged treatment at the median are -4 to -13 points on the share proficient in math and -6 to -16 points at the mean. Based on the median, the impacts on the share proficient in reading are -8 to -30 points and at the mean they are -14 to -35 points.

Panel B provides the corresponding results for share not proficient. The estimates for math are insignificant and not statistically different from zero. Columns (5) through (8) provide the estimates for share not proficient in reading, which are all highly significant. At the median, the impacts on share not proficient in reading using level and logged treatment are 7 to 31 points and at the mean are 13 to 37 points.

In general, the coefficients for reading, both shares proficient and not proficient, are highly significant at the one percent level. The math proficient estimates are significant at the one percent level except for column (1), and the share not proficient estimates are not statistically different from zero[7].

Overall, the logged treatment results are larger in magnitude than the results from Table 1.2. The level results are similar across both specifications. The results for reading in both Tables 1.2 and 1.4 tend to be more significant and larger in magnitude than the corresponding coefficients for math. Consistency in the linear estimates from both specifications suggests that lead exposure in schools had meaningful impacts on student proficiency[8].

Taken together, Tables 1.2 and 1.4 suggest a conservative estimated impact at the mean of -6 to -9 points for share proficient in math, -12 to -14 points for share proficient in reading, and 13 to 15 points for share not proficient in reading. These estimates account for approximately half of the observed changes in proficiency between the pre- and post-periods. The mean class cohort size is 38 students, so this suggests that exposure to lead in the classroom caused an average of three students to fall out of proficiency in math, five students to fall out of proficiency in reading, and five students to be labeled not proficient in reading for each school-specific cohort.

The education literature has studied many determinants of academic success. From the education production analyses class size, teacher quality, and spending per pupil have received most of the attention. The class size literature is large and conflicted in its analysis. Most studies have failed to identify a meaningful relationship between class size and overall student achievement in elementary school (Hoxby (2000), Chingos (2012), Bosworth (2014)).

[7] A squared term is added to the regression to ease the linear restriction and the estimated coefficients are nearly identical to those found in Table 4.

[8] The regression from Table 4 was also estimated using a full dataset incorporating the control schools. With the treated variable winsorized to the same level, the results were not statistically different from those found in Table 4.

Teacher quality has also been widely studied. While it is generally understood that a great teacher can have meaningful impacts on student success, measuring what makes a teacher great has been difficult. By and large, the evidence has been meager on teacher qualifications and student outcomes. Teacher productivity has been linked to classroom experience, but studies have found this is only relevant during the first few years of a teacher's career when they are still honing their practice (Rivkin, Hanushek, and Kain (2005), Buddin and Zamarro (2009), Harris and Sass (2011), Garritsen, Plug, and Webbink (2017)).

There is evidence that per-pupil spending influences student achievement. Findings from both Michigan and Florida suggest a similar impact of a five percentage point increase in passage rates for fourth graders with an increase of ten percent in spending (Papke and Wooldridge (2008), Kesler and Munkin (2015)). Using the conservative results from this study, an increase of 10 to 20 percent in per pupil spending is needed to return math proficiency shares to levels prior to the change in water or an increase of 25 to 30 percent in per pupil spending to return reading proficiency shares back to their respective levels.

1.6 Robustness

The main regressions have used a treatment variable calculated from an average of the reported lead-in-water levels for each classroom. The purpose of averaging the water draws collected by the Michigan Department of Environmental Quality was to reasonably approximate the quality of the water students drank throughout the day. There is no reason to believe that this is the only appropriate way to model the variable of interest. The cumulative blood lead exposure treatment variable is unique to this study, and testing its sensitivity is important for understanding the robustness of the results.

The basic model is estimated again using cumulative blood lead exposure levels derived from the maximum value of lead as opposed to the average of four draws. The process is the same as the derivation of the other cumulative blood lead exposure measure. The new median treatment value is 62 micrograms and the mean is 124 micrograms.

Figure 1.5 shows the distribution of the new treatment variable. The right tail has moved even further away from the mean of the distribution. To mitigate the effects of outliers, the treatment is again winsorized at the 95th percentile and specifications using logged treatment are estimated. The new average cumulative treatment values by academic year are shown in Figure 1.6. The proficiency data is the same as the summary statistics used for the previous estimates, available in Table 1.1.

Table 1.5 shows the results from estimating the basic regression using the new treatment variable. The estimates follow the same process as Table 1.2. Panel A shows the results for share proficient. The estimates are the expected direction. The reading estimates are larger and more significant, particularly with the inclusion of school linear trends in columns (6) and (8). For math, the estimated impact is -6 points at the median and -11 points at the mean. The impacts on share proficient in reading for level and logged treatment are -9 to -10 at the median and -12 to -19 at the mean.

Panel B of Table 1.5 shows the results for share not proficient. Similar to Table 1.2, the estimates are mostly statistically different from zero and show the anticipated relationship. The impacts for math not proficient are not significant. Columns (5) through (8) show the estimates for share not proficient in reading which are more significant. Using level and logged treatment, the estimated impacts at the median are 11 to 12 points and at the mean are 15 to 22 points.

The results from using the maximum lead draw as an alternative to the average of the four draws largely support the earlier findings in Table 1.2. Comparing the preferred estimates using lead values from the average and maximum water lead levels, the impacts measured at the median are very similar and estimates at the mean are slightly larger when using the maximum values. Table 1.6 compares the estimated impacts between the two treatment variables based on the estimates in Tables 1.2 and 1.5[9].

1.7 Discussion

The magnitudes found in this study may not accurately reflect the impacts from similar levels of lead exposure in other schools. The absorption of lead into the body and its distribution to the brain and other organs is dependent on student nutrition. Children with deficiencies in calcium and iron are at a greater risk for lead absorption and impacts to cognitive development. More than 80 percent of students in Flint Community Schools qualify for the free lunch program which may signal a greater likelihood for these deficiencies. The impact of similar cumulative lead exposure levels may be smaller for schools whose students are more food secure.

The magnitudes of the impacts may also be affected by the average lead exposure outside of school. In previous studies of younger children, the largest marginal neurological impacts occur at relatively low blood lead levels (Canfield et al (2003), Lamphear et al (2000)). This nonlinearity at low levels of lead exposure suggests that the marginal magnitudes of this study may have been greater if the students had not been exposed to lead outside the classroom.

The main contribution of this study is evidence of meaningful, short run impacts from lead on the achievement of students in elementary school. Until recently, it was assumed that lead exposure in older children was negligible until blood lead levels reached 10 micrograms, the

---

[9] A similar exercise was conducted to compare the results of Table 4 to estimates using the maximum value of lead. The results were very similar across the two specifications.

threshold at which physical symptoms begin to manifest. In 2006, Seattle Public Schools were advised that elevated lead-in-water levels as high as 963 parts per billion presented a very small risk of elevated blood lead levels and subsequent health issues[10]. Other studies have since documented negative cognitive impacts for young children with blood lead levels below 10 micrograms. This study adds to the growing literature by finding important impacts for older children with moderate lead exposure.

It is expensive to raise proficiency rates and provide extra support to students who are struggling. A study commissioned by the State of Michigan found it costs 30 percent more for schools to educate at-risk students. The potential costs for students in special education are even higher[11]. Academic papers have found that per pupil spending needs to increase approximately twenty percent to raise proficiency rates a magnitude equal to the decline found in this study (Papke and Wooldridge (2008), Kesler and Munkin (2015)). These estimated increases in spending are to raise proficiency shares for a typical cohort. It is not known whether the efforts to raise the performance of older children suffering from lead exposure will require additional support.

This study focuses on the short-run impacts of lead exposure and cannot conclude whether the cognitive impacts are permanent or can be easily overcome with remedial education. The studies of early childhood blood lead levels suggest that the impacts of lead continue to affect student achievement years later. If students continue to lag in academic achievement after

---

[10] Heffter, Emily, and Warren King. 2006. "Above-normal levels of lead found in Seattle schools' water." *The Seattle Times*, November 9. http://www.seattletimes.com/seattle-news/above-normal-levels-of-lead-found-in-seattle-schools-water/

[11] Higgins, Lori. 2016. "Report: At-risk students need more Michigan funding." *Detroit Free Press,* June 28. http://www.freep.com/story/news/education/2016/06/28/study-michigan-must-create-equal-school-funding-system/86289694/

the lead exposure, there may be meaningful negative impacts on human capital development that persist into adulthood.

Appropriate plans and actions can be taken with knowledge of the consequences of lead exposure for elementary students. Most schools are not legally required to test their drinking water. However, maintenance of plumbing systems and regular analysis of water quality may now be prudent with evidence that student success is at risk. Officials responding to elevated levels of lead in schools can now make tough decisions with the knowledge that limiting exposure is protecting students. Revisiting what level of lead-in-water is safe for children in school may also be appropriate considering the cumulative impacts on the brain and other organs from extended exposure to low-levels of lead over time.

1.8 Conclusion

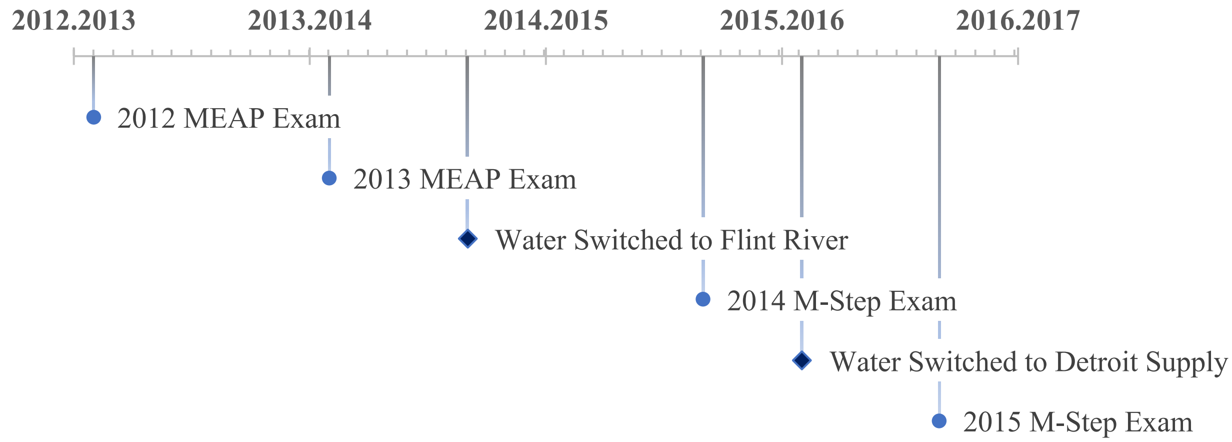
This study utilizes an exogenous shock of lead-in-water levels within schools to estimate the impacts of cumulative blood lead exposure on elementary school students. The analysis of Flint Community Schools suggests that at a mean cumulative blood lead exposure of 49 micrograms, the share proficient in math dropped by -6 to -9 points, the share proficient in reading dropped by -12 to -14 points, and the share not proficient in reading rose 13 to 15 points. These findings were the first to confirm negative, short run impacts on achievement for elementary school children due to lead exposure.

The analysis is limited by a small sample size and limited variation in the treatment variable. Consistent estimates are found from multiple specifications. The magnitudes found in this study may not accurately reflect the response to exposure in other schools due to differences in environmental factors. The results provide evidence of the risks of lead exposure to older children and offer important insight for school policy.

This paper is the first to use a treatment variable that reflects cumulative blood lead exposure. Other applications using this unique data may prove helpful in understanding the impacts of lead on elementary students. Long run studies have found significant impacts on student behavior, and it would be interesting to test if these relationships are salient in the short run.

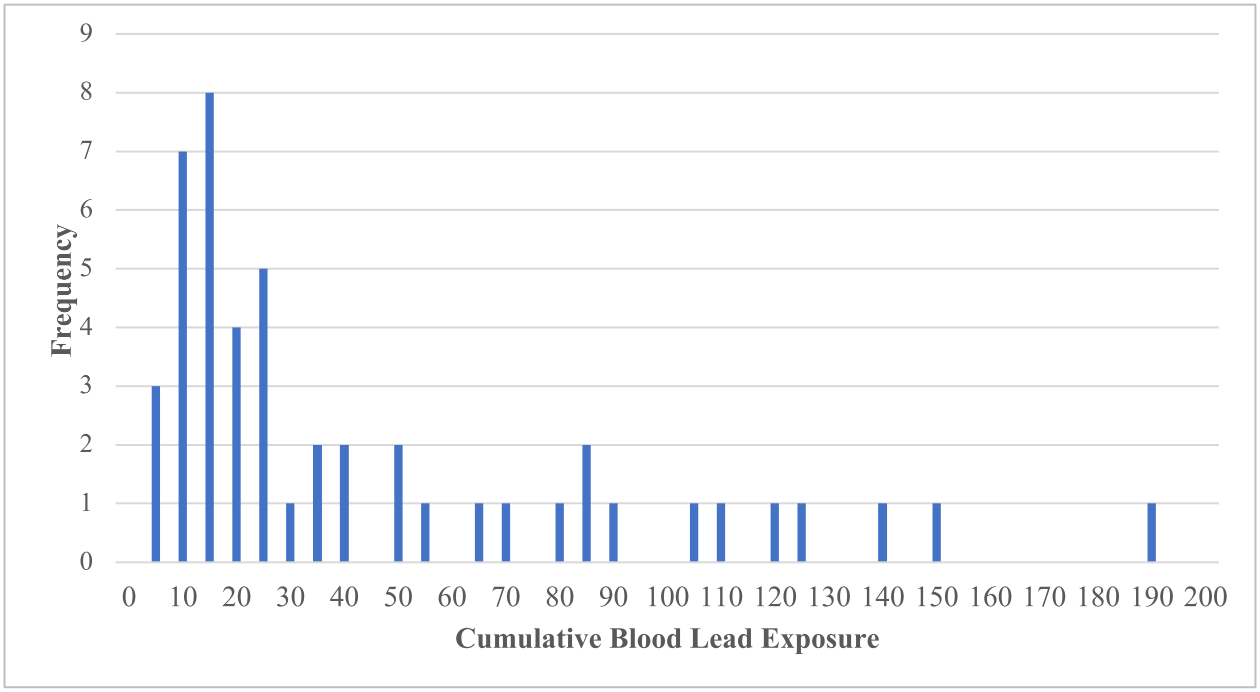
The potential harm from exposure to lead has been of public concern for decades. The findings of this study suggest that older children are vulnerable to exposure to elevated lead-in-water levels. This information can help schools and public policy form appropriate responses to the risks of lead exposure in the classroom.

Figure 1.1: Timeline of Flint Water and School Testing by Academic Year



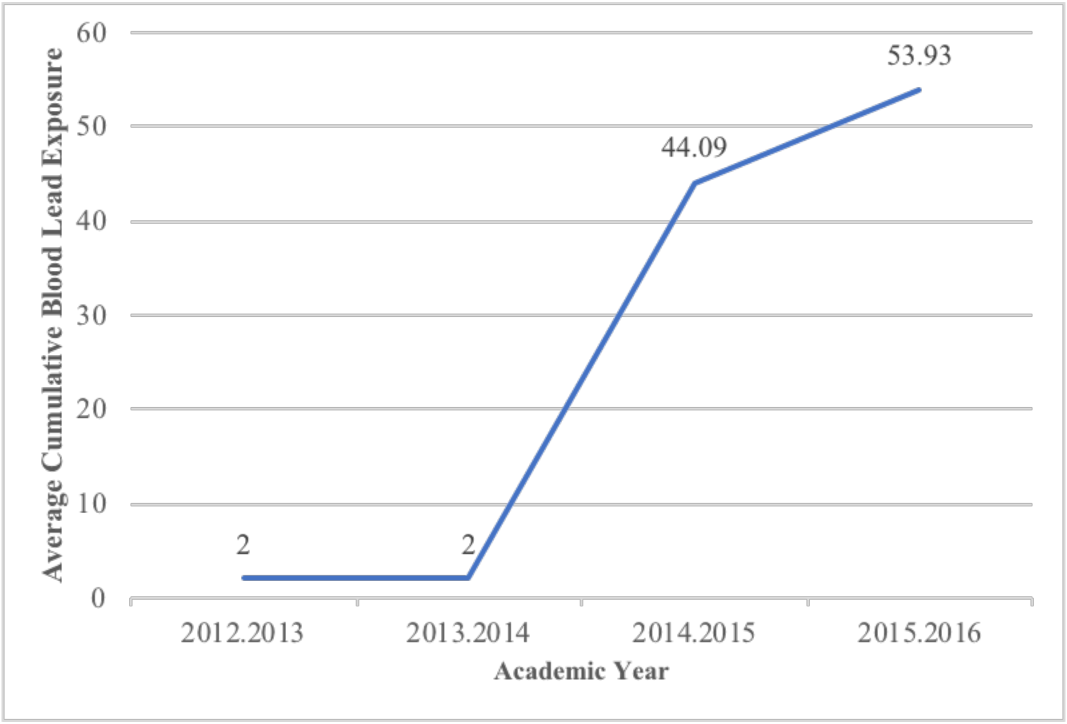
The timeline is presented by academic year. On the timeline, 2012.2013 represents September 2012, the beginning of the 2012-2013 academic year.

Figure 1.2: Blood Lead Exposure from Average Classroom Water Draws



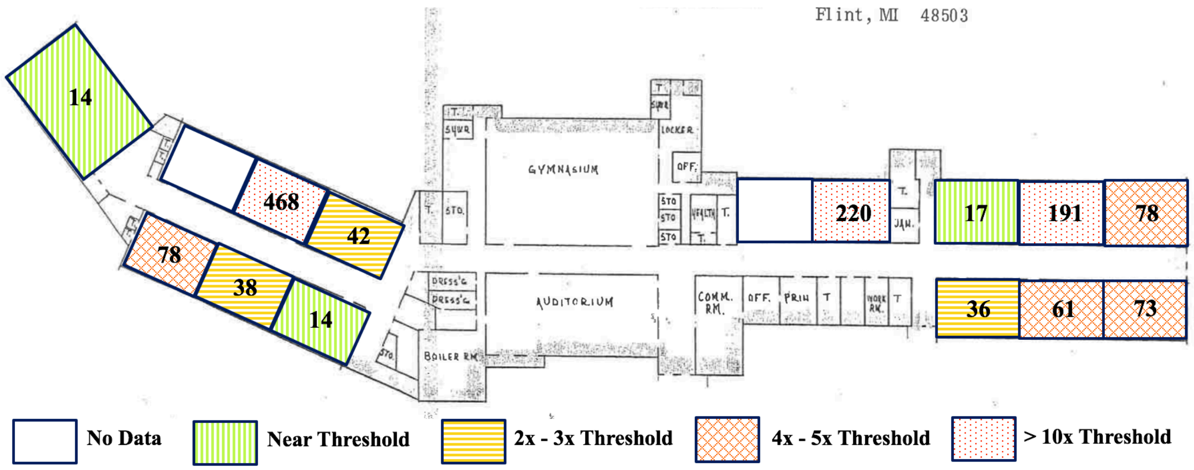
Cumulative blood lead exposure values are pooled from academic years 2014-2015 and 2015-2016. They are calculated using averaged lead-in-water values for Flint Community Schools.

Figure 1.3: Mean Blood Lead Exposure from Average Draws by Year



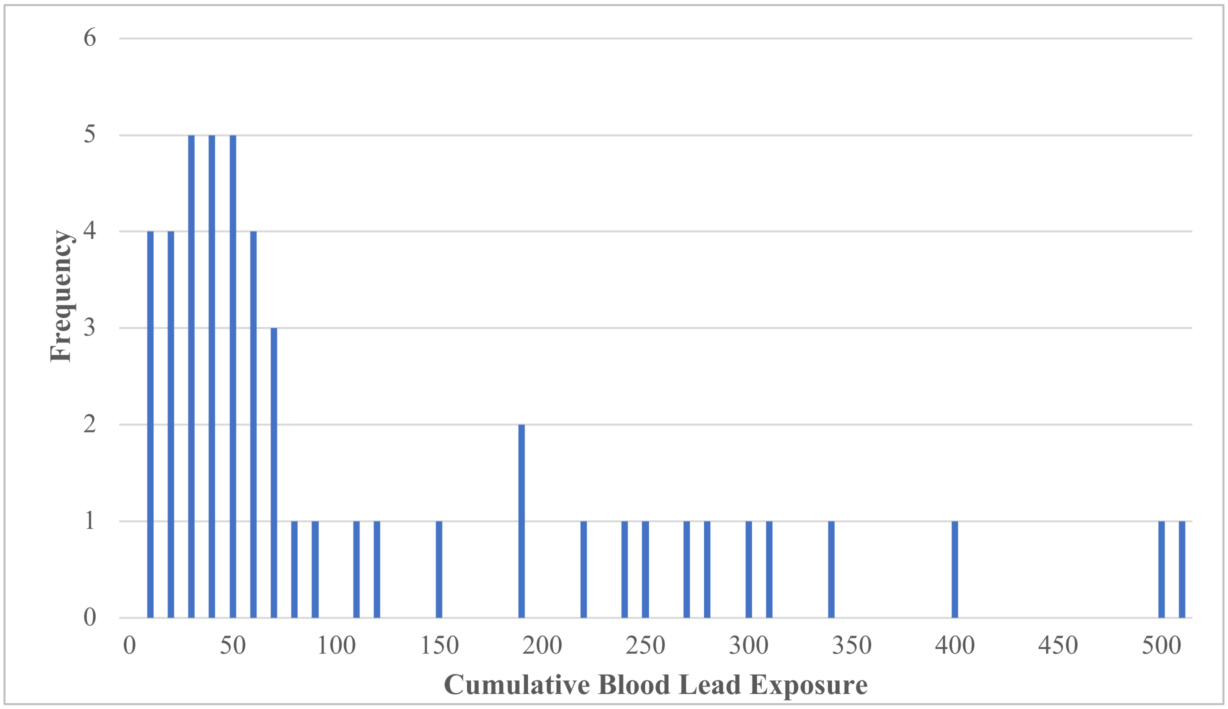
Average cumulative blood lead exposure values are assumed to be two for the 2012-2013 and 2013-2014 academic years. Values are estimated at the time of standardized testing.

Figure 1.4: Heterogeneous Levels of Water Borne Lead Across Classrooms



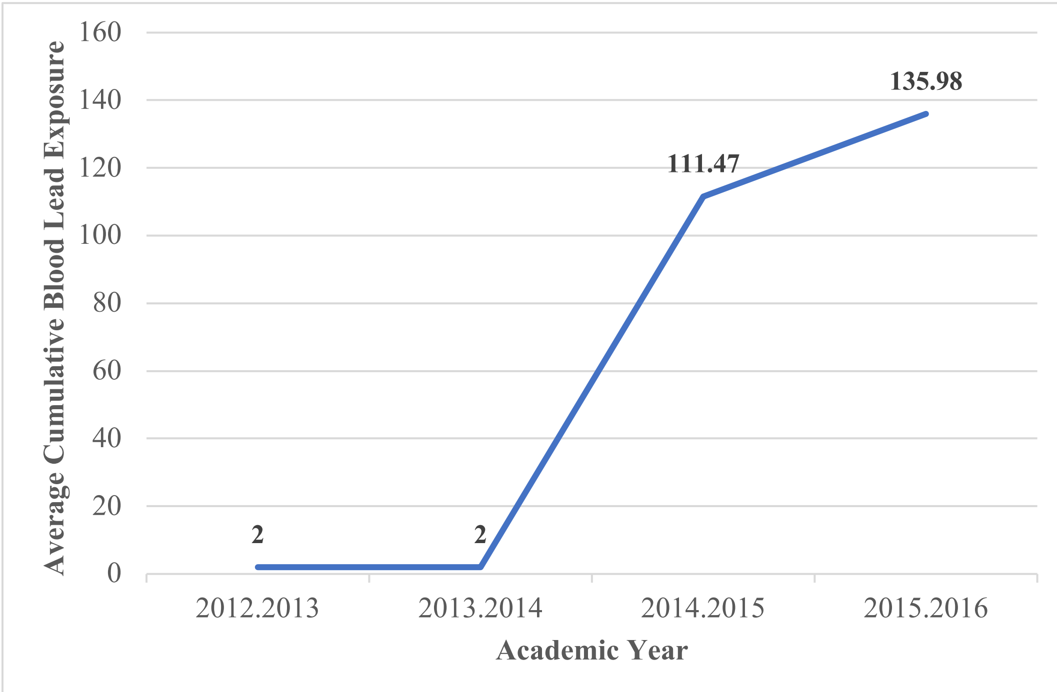
This is a map of Pierce Elementary School. The values show the lead-in-water levels as measured in parts per billion for the largest value collected by the Michigan Department of Environmental Quality. The EPA's threshold is 15ppb.

Figure 1.5: Blood Lead Exposure from Maximum Classroom Water Draws



Cumulative blood lead exposure values are pooled from academic years 2014-2015 and 2015-2016. They are calculated using maximum lead-in-water values for Flint Community Schools.

Figure 1.6: Mean Blood Lead Exposure from Maximum Draws by Year



Average cumulative blood lead exposure values are assumed to be two for the 2012-2013 and 2013-2014 academic years. Values are estimated at the time of standardized testing.

Table 1.1: Summary of Data for Flint and Comparison School Districts

	City of Flint School District		Comparison School Districts	
	2012.2013 - 2013.2014	2014.2015 - 2015.2016	2012.2013 - 2013.2014	2014.2015 - 2015.2016
	(1)	(2)	(3)	(4)
Cumulative Lead from Water	2	49.01	2	2
	(0)	(46.78)	(0)	(0)
Math Proficient	.24	.11	.35	.24
	(.10)	(.08)	(.13)	(.15)
Reading Proficient	.47	.20	.58	.33
	(.12)	(.10)	(.12)	(.13)
Math Not Proficient	.56	.61	.46	.41
	(.12)	(.16)	(.14)	(.17)
Reading Not Proficient	.22	.52	.14	.39
	(.09)	(.15)	(.06)	(.13)
Observations	48	48	133	120

A constant value of 2 micrograms of cumulative lead exposure is assumed for the 2012 and 2013 academic years and for the comparison schools. Standard deviations are presented in parentheses.

Table 1.2: Basic Regression

Panel A: Proficient Shares

	Math Proficient				Reading Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	-.0003	-.0018			-.0011	-.0031*		
	(.0009)	(.0013)			(.0011)	(.0016)		
Ln(Lead Exposure)			-.0006	-.0112			-.0102	-.0310*
			(.0100)	(.0144)			(.0117)	(.0169)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	349	349	349	349	349	349	349	349
Groups	169	169	169	169	169	169	169	169
R-Squared	.4614	.5633	.4612	.5608	.6447	.7107	.6442	.7099

Panel B: Not Proficient Shares

	Math Not Proficient				Reading Not Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	.0019	.0020			.0019*	.0035**		
	(.0014)	(.0018)			(.0010)	(.0017)		
Ln(Lead Exposure)			.0186	.0184			.0170	.0378**
			(.0141)	(.0178)			(.0105)	(.0170)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	349	349	349	349	349	349	349	349
Groups	169	169	169	169	169	169	169	169
R-Squared	.1297	.2844	.1289	.2835	.7284	.7992	.7270	.7993

These estimates are calculated using data for Flint Community Schools and a full set of control schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. All estimates include fixed effects and year fixed effects. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 1.3: Basic Regression within Flint Community Schools

Panel A: Proficient Shares

	Math Proficient				Reading Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	.0003 (.0004)	-.0002 (.0005)			-.0001 (.0005)	.0005 (.0008)		
Ln(Lead Exposure)			.0057 (.0181)	-.0171 (.0260)			-.0074 (.0216)	-.0206 (.0317)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	96	96	96	96	96	96	96	96
Groups	42	42	42	42	42	42	42	42
R-Squared	.6539	.6948	.6509	.6973	.7570	.8002	.7572	.8005

Panel B: Not Proficient Shares

	Math Not Proficient				Reading Not Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	-.0006 (.0004)	-.0003 (.0007)			-.0003 (.0005)	-.0001 (.0007)		
Ln(Lead Exposure)			-.0262 (.0274)	-.0191 (.0432)			.0033 (.0337)	.0008 (.0368)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	96	96	96	96	96	96	96	96
Groups	42	42	42	42	42	42	42	42
R-Squared	.3183	.4287	.3179	.4304	.7296	.8146	.7283	.8145

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. All estimates include fixed effects and year fixed effects. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 1.4: Flint Community Schools with Linear Time Trends

Panel A: Proficient Shares

	Math Proficient				Reading Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	-.0004	-.0013***			-.0015***	-.0028***		
	(.0004)	(.0004)			(.0004)	(.0009)		
Ln(Lead Exposure)			-.0314***	-.0409***			-.0780***	-.0903***
			(.0109)	(.0100)			(.0114)	(.0136)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	96	96	96	96	96	96	96	96
Groups	42	42	42	42	42	42	42	42
R-Squared	.4686	.5560	.5305	.6093	.5183	.6237	.6772	.7588

Panel B: Not Proficient Shares

	Math Not Proficient				Reading Not Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	-.0003	.0001			.0013***	.0027***		
	(.0004)	(.0006)			(.0004)	(.0009)		
Ln(Lead Exposure)			-.0009	.0040			.0864***	.0974***
			(.0104)	(.0127)			(.0130)	(.0154)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	96	96	96	96	96	96	96	96
Groups	42	42	42	42	42	42	42	42
R-Squared	.1775	.2941	.1741	.2947	.5160	.6470	.6606	.7679

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. All estimates include fixed effects and year linear time trends. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 1.5: Basic Regression with Maximum Treatment Variable

Panel A: Proficient Shares

	Math Proficient				Reading Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	-.0001 (.0004)	-.0009 (.0006)			-.0005 (.0005)	-.0015** (.0007)		
Ln(Lead Exposure)			-.0004 (.0077)	-.0078 (.0111)			-.0078 (.0089)	-.0239* (.0129)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	349	349	349	349	349	349	349	349
Groups	169	169	169	169	169	169	169	169
R-Squared	.4615	.5637	.4612	.5606	.6447	.7114	.6443	.7099

Panel B: Not Proficient Shares

	Math Not Proficient				Reading Not Proficient			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lead Exposure	.0008 (.0006)	.0009 (.0008)			.0009* (.0004)	.0018** (.0007)		
Ln(Lead Exposure)			.0139 (.0110)	.0138 (.0139)			.0130 (.0081)	.0302** (.0129)
School Linear Time Trend	No	Yes	No	Yes	No	Yes	No	Yes
Observations	349	349	349	349	349	349	349	349
Groups	169	169	169	169	169	169	169	169
R-Squared	.1284	.2846	.1283	.2834	.7286	.8009	.7270	.7999

These estimates are calculated using data for Flint Community Schools and a full set of control schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. All estimates include fixed effects and year fixed effects. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 1.6: Comparing Estimated Impacts on Shares (Not) Proficient

	Average of Four Draws		Maximum of Four Draws	
	Median Lead	Mean Lead	Median Lead	Mean Lead
Basic Regression				
Math Proficient	-5	-9	-6	-11
Reading Proficient	-8	-15	-9	-19
Math Not Proficient				
Reading Not Proficient	9	17	11	22

These values are the estimated impacts using the average lead-in-water level of four draws of water and the maximum value drawn. These are calculated using the respective coefficients from the basic regression specification found in Tables 1.2 and 1.5. For the average of four draws the median cumulative lead exposure value is 27 micrograms and the mean is 49 micrograms. For the maximum of four draws the median cumulative lead exposure value is 62 micrograms and the mean is 124 micrograms.

Chapter 2

Student Behavior and Lead Exposure: Evidence from School Discipline Data

2.1 Introduction

Lead is a toxin that causes neurophysiological changes in the body. Symptoms may manifest in both neurocognitive disorders as well as psychiatric disturbances (Mason et al, 2014). The first chapter of this dissertation focused on estimating the causal cognitive declines due to lead exposure. This chapter focuses on the psychiatric disturbances, which include changes in mood and behavior. Behavioral response is the latest frontier in the decades of study of lead toxicity. Dr. Herbert Needleman, a distinguished physician and scholar in lead research, suggested affected social behaviors may prove to be a more important impact than the long-held focus on cognitive dysfunction (Needleman, 2004).

There is empirical evidence of this troubling relationship. Elevated blood lead levels have been linked to ADHD behavioral disorders (Cecil et al (2008), Daneshparvar et al (2016)), disciplinary actions (Aizer and Currie (2017), Needleman et al (2002)), displays of strong temper, and teenage pregnancy (Reyes (2015), Nevin (2000)). Rather than dissipating with age, there is evidence that lead exposure in young children has contributed to crimes committed as adults, including violent crimes (Reyes (2007), Nevin (2007)) and homicides (Feigenbaum and Muller (2016), Stretsky and Lynch (2001)).

One of the greatest challenges in estimating the impacts of lead on behavioral outcomes is the issue of endogeneity. Measured individual blood lead levels are highly correlated with personal and neighborhood characteristics that may not be easy to observe and control. The potentially deleterious effects of lead make conducting an experiment unconscionable. Recent studies have used clever spatial differences in exposure and simulated instrumental variables to tackle this challenge, and the results suggest that groups exposed to lead behave differently than those who were not.

The purpose of this study is to use a qausi-natural experiment to estimate the causal relationship between lead exposure and student behavior. A change in the municipal water supply in Flint, Michigan caused an unexpected shock of water borne lead within classrooms in Flint Community Schools. The nature of the shock provides variation in the intensity of lead exposure within groups of children over time as well as cross sectional variation across cohorts. This offers a unique approach to address the perennial issue of endogeneity.

The city of Flint was on the brink of bankruptcy in 2011. An emergency city manager forced a series of budget cuts to address the deficit. One of the financial decisions ended a long-held water supply contract with Detroit Water and Sewage effective April of 2014. The Flint Water Service Center assumed responsibility for supplying municipal water. Water drawn from the local Flint River was treated and used as the primary water source through October of 2015.

The financial decisions surrounding the change in municipal water resulted in highly corrosive water being supplied throughout the city. Lead leached from the services lines connecting buildings to the public water mains as well as from the plumbing and fixtures within buildings. These highly localized sources resulted in elevated levels of lead in the drinking water.

Reports from the Michigan Department of Environmental Quality's formal investigation reveal heterogeneous levels of water borne lead across classrooms within Flint Community Schools. The elementary schools included in this study have water fountains located within each classroom that serve as the primary source of drinking water for students. The fountains look nearly identical, but variation in exposed lead from the fixtures and solder resulted in significant differences in the lead-in-water levels across the classrooms. Classroom usage is highly persistent over time, with grades usually taught in the same rooms year after year. As a result, students were randomly exposed to different levels of water borne lead based on their classroom assignment.

A contribution of this study is the exogenous source of classroom lead exposure to estimate the causal impacts of lead on student behavior. In the regression analysis groups of students within schools are observed across academic years. Variation in the intensity of lead exposure over time drives the identification.

This study finds significant and meaningful impacts on discipline. At the mean level of lead exposure, 27 micrograms, the impact is an additional 8.6 disciplinary actions per grade within each school. The results are strongest for students in second through fourth grades. Most of the disciplinary actions are coming from short out-of-school suspensions rather than expulsions or long out-of-school suspensions. The results are robust to alternative measures of water borne lead exposure as well as the treatment of missing data.

A unique contribution of this study is its focus on the short-run relationship between lead and behavior. Previous work has found lagged effects of lead and adverse behaviors, with the timing between cause and effect spanning five to fifteen years. Very little work has looked at the contemporaneous impacts of lead on behavioral outcomes. This paper focuses on student behavior during and shortly after lead exposure.

This study also distinguishes itself by testing whether older children are vulnerable to the psychiatric impacts of lead. Most the work on lead and behavior has focused on exposure occurring during the very young, formative years. The focus on young children comes from their physical vulnerability to exposure as well as their increased propensity to consume things that may contain lead, such as dirt and paint chips. These reasons have prompted public resources to target lead prevention and blood lead screenings for small children. A consequence of these policy decisions is that the available data for empirical work is almost exclusively on young children.

Understanding the impacts of lead on older children is valuable for public policy. Prior to 2016 there were no state or federal laws requiring regular tests of water quality within schools. In the last two years, six states have adopted new guidelines for oversight of water borne lead in schools[12]. Recognizing behavior as a symptom of lead exposure in older children may contribute to better identification of lead issues. It may also help schools tailor responses that address the behavioral issues in a way that is productive rather than the use of disciplinary sentences that may be ineffective at curbing the lead related behavior.

The results of this paper also contribute to recent work that studies the impacts of behavior and discipline in schools. Studies have estimated whether the use of exclusionary discipline impacts the academic achievement of the offenders and their peers (Kinsler (2013), Cobb-Clark et al (2015)). The efficacy of the policies as well as the external validity of the studies may be dependent on whether there are environmental health factors, such as lead exposure. Furthermore, it is well-documented that low-income students and minority students are more likely to receive disciplinary actions (Jordan and Anil (2009), Kinsler (2011), Bekkerman and Gilpin (2015, 2016)). Lead exposure is often correlated with both observed characteristics and may be an underlying issue contributing to the disparities.

This paper continues with the following sections. Background information is provided on the Flint water crisis, the theoretical relationship between lead and behavior as well as previous empirical work. This is followed by details about the data and empirical strategy. Next

[12] Nunez, Elissa, Amy Molloy, and News21. 2017. "Schools fail lead tests while many states don't require testing at all." *The Center for Public Integrity*, August 15. https://www.publicintegrity.org/2017/08/15/21076/schools-fail-lead-tests-while-many-states-don-t-require-testing-all

the main results and robustness checks are provided. Finally, the paper ends with a brief discussion and concluding remarks.

2.2 Background

2.2.A The Flint Water Crisis

In 2011, the city of Flint was in a financial crisis resulting from years of economic decline. On the verge of bankruptcy, Flint was appointed an emergency city manager by the governor of Michigan. The purpose of the city manager was to make difficult budgetary decisions to push the struggling city back into financial solvency. Among the initiatives pursued was a change in municipal water sources. The city of Detroit had supplied Flint with water for several decades and costs had steadily risen. A new water authority that would supply water from Lake Huron was under construction and would be available for municipal use as early as 2017. A contract with the new water supplier was signed in 2013 under advisement of the city manager. Faced with an interim period of several years, a decision was made to locally treat the Flint River until the new water pipeline was established[13].

The city of Flint's water treatment plant had not been regularly utilized in nearly 50 years (Davis et al, 2016). A year was taken to make sure the plant was updated and prepared to be the primary supplier. In April 2014, responsibility for municipal water transitioned from Detroit Water and Sewage to the city of Flint.

The provision of water appeared seamless to consumers, but there were important differences between the water provided by Flint and the water that had been supplied from Detroit. First, Detroit Water and Sewage had regularly treated the drinking water with orthophosphate for more than twenty years. This additive built a passive layer of protection on

[13] "Flint Water Crisis Fast Facts." *CNN*. http://www.cnn.com/2016/03/04/us/flint-water-crisis-fast-facts/index.hteml (accessed September 1, 2017).

the inside of pipes to prevent the corrosion of plumbing materials. The city of Flint chose not to continue the use of orthophosphate due to budgetary constraints nor utilize an alternative corrosion inhibitor to protect the water infrastructure. As a result, the passive layer inside the pipes was susceptible to flaking and exposed pipes to corrosion (Torrice, 2016).

Second, the water entering the Flint water treatment plant was of poor quality and required extra treatment to make it safe for consumption. The city of Flint drew water from the Flint River while Detroit sources water from Lake Huron. In general, rivers tend to be naturally more corrosive and contain more organic materials. This is particularly true in urban areas where rain can carry pollutants into the water from local streets. To remove particles and microorganisms additional treatment is needed[14]. In Flint, chlorine was added to the water as a disinfectant to kill the microorganisms. As the passive layer of orthophosphate in the pipes flaked off the chlorine reacted with iron in the pipes. The corrosion process consumed the chlorine and left bacteria in the water. This forced the city to raise the chlorine levels even higher to prevent E. coli outbreaks.

Third, the microorganisms killed by chlorine increased the organic material in the water. To remove those contaminating materials ferric chloride was added as a coagulate to assist with the filtering process (Torrice, 2016). This additive to the already corrosive water caused chloride levels to soar. High chloride-to-sulfate ratios in water are known to be very corrosive to lead (Edwards and Triantafyllidou, 2007). In general, a chloride-to-sulfate level of .58 is considered an upper bound for water management. Researchers from Virginia Tech sampled treated water with ratios as high as 1.6 in Flint during the water crisis.

---

[14] Olson, T. 2016. "The science behind the Flint water crisis: corrosion of pipes, erosion of trust." *The Conversation,* January 28. https://theconversation.com/the-science-behind-the-flint-water-crisi-corrosion-of-pipes-rosion-of-trust-53776

Corrosive waters can leach lead from many plumbing sources. Lead service lines connecting older homes to the municipal water mains are an obvious source as is lead solder connecting pipes in older buildings. Brass components often include lead and are commonly used in water fixtures such as faucets, fountains, and valves. New "lead-free" brass is permitted to have .25% lead by weight; prior to 2014 brass was considered "lead-free" if lead composed less than 8% of the material by weight. Lead is a favored material for plumbing because it has a relatively low melting point which makes it malleable and effective at combating pinhole leaks[15].

The low melting point also means that it solidifies after the other metals in the brass components and often has relatively more surface exposure to water than other materials. Analysis provided by the Michigan Department of Environmental Quality reported elevated levels of lead in the Flint Community Elementary Schools' water and cite water fixtures as the probable sources. Heterogeneity in the distribution of lead on the surface area of the plumbing components within classrooms may have also contributed to the variation in the levels of lead-in-water across classrooms within school buildings.

2.2.B Lead and Behavior

Lead has historically served as a valuable additive to consumer products by improving both durability and performance in such common items as paint, car engines, and plumbing components. Unfortunately, it is also a neurotoxin that can cause severe and permanent physical and neurological effects. Lead is most dangerous when ingested. Environmental lead exposure is often accidental and undetectable by sight, taste, or smell. Several common sources include breathing lead exhaust prior to the leaded gasoline phase out, accidental ingestion of dust or chips from peeling leaded paint, or consumption of water with elevated levels of lead.

---

[15] Dickey, Kirk. 2014. "Lead-free brass – what is it? Why buy it?" *Direct Material,* October 24. https://www.directmaterial.com/knowledge/lead-free-brass-buy/

After lead enters the body it is absorbed into the blood stream. Once absorbed, it travels through the circulatory system and becomes deposited in soft tissues, organs, and bones. Lead is filtered relatively quickly from the blood with an approximate half-life of 30 days. As a result, blood lead levels drop to within normal ranges after a few months. Lead deposited in organs has a longer half-life of years, and lead can be encapsulated in bones for decades.

As a neurotoxin, lead impedes the proper development of the brain and nervous system. On a molecular level it mimics calcium, causing issues with neuronal signaling and neurotransmitter release. The development of synapses in the brain may be compromised as well as the integrity of the blood brain barrier which continues to develop throughout childhood into the second decade (Lidsky and Schneider, 2003). Furthermore, lead inhibits important enzymes which may lead to behavioral disorders (Needleman, 2004).

Regulations have become more stringent as evidence of the potential adverse effects of lead have grown. In the 1970's the Safe Drinking Water Act was passed along with restrictions of lead in residential paint and the transition away from leaded gasoline. The current regulation for municipal water authorities was set by the Lead and Copper Rule of 1991 with an actionable level for lead in public water at 15 parts per billion (ppb) at the 90^{th} percentile for customer taps. Oversight on drinking water within schools has been nearly nonexistent. Prior to 2016 there were no state or federal regulations requiring schools to test for water borne lead. As of January 2018 that number had grown to six states.

Most of the legislative focus has been on preventing lead exposure to very young children, the group long held as most vulnerable to the impacts of lead. In 1991, the CDC recommended universal blood lead testing for all young children. Many states have adopted blood lead screenings and routinely test at children's appointments up to their fifth birthdays.

For child recipients of Medicare blood lead screening is compulsory. The blood lead level of concern was also lowered in 1991 from 60 mg/dL to 15 mg/dL. The previous level was set based on the physical manifestation of symptoms in toddlers, such as stomach aches and seizures. Growing evidence of a relationship between lead and cognitive development suggested that there was permanent neurological damage at the new, lower threshold.

In the following decades medical research continued to find negative cognitive impacts at lower blood lead levels. In 2012 the Center for Disease Control acknowledged that there is no safe level of lead exposure for children. As a result, the blood lead level of concern of 15 mg/dL was replaced with an intervention trigger level of 5 mg/dL (Advisory Committee on Childhood Lead Poisoning Prevention, 2012).

Empirical analyses of the relationship between lead and behavior has periodically been undertaken by economists with clever approaches or as new data have become available. This work has supplemented the body of literature developed by medical professionals. The general shift of focus in the lead literature toward behavioral impacts comes as issues regarding mental health and discipline in schools are receiving greater attention.

2.2.C Previous Literature

The economic literature has grown in the last decade with new work attempting to identify and quantify a causal relationship between lead and adverse behavior. Reyes (2007) uses state measures of crime and the phase out of leaded gasoline to find mixed evidence for the link between lead and violent crime. A clever approach using historical city-level data is employed in Feigenbaum and Muller's paper (2016). The growth of municipally supplied water and use of lead pipes in the second half of the 19th century is used for the variation in exposure to water borne lead. The estimates suggest the use of lead pipes contributed to significant increases in city-level homicide rates. Both papers use a twenty year lag between exposure to lead and criminal behaviors based on the assumption that young, underdeveloped brains are most at risk to lead exposure and that criminal activity primarily occurs in the young adult years.

Recent work has attempted to quantify the impacts of early childhood blood lead levels on the behaviors of school age children. Reyes (2015) measures many potential consequences, including behavior, pregnancy, and aggression. The impacts are estimated using predicted childhood blood lead concentrations and the results suggest elasticities between .1 and 1.0. With instrumental variables, Aizer and Currie (2017) use individual level data to estimate the impact of leaded gasoline on school discipline. They find an increase in the probability of suspension of 6.4 to 9.3 percent.

This paper continues this important work by estimating the impacts of lead on behavior in primary school. The unexpected change in water quality in Flint, Michigan provides a quasi-natural experiment to estimate a causal relationship, overcoming the common identification challenge of endogeneity in lead studies.

This study is also among the few to consider the impacts of lead exposure in older children. While the previous focus has been on young children, medical studies have found that the brain is still developing through primary school. Older children may be at risk for the deleterious effects of lead exposure and understanding this relationship can help guide intervention and prevention measures.

Finally, this study is also among the few to consider the very short run effects of lead on behavior. Data limitations have focused the previous works on lags of five to twenty years between exposure and behavioral outcomes. Understanding this relationship may provide guidance for schools that are concerned about potential water borne lead or that observe distinct changes in student behavior.

2.3 Data

This study makes use of publicly available data. Water borne lead levels come from data reports commissioned by the Michigan Department of Environmental Quality. School data are supplied by Flint Community Schools following several Freedom of Information Act (FOIA) requests. Data is aggregated at the grade within school level and cohorts of students within schools are tracked over time.

The period of study begins one year prior to the change in water sources with the 2012-2013 academic year and continues through the 2015-2016 academic year. There are eight elementary schools operating during this period in the Flint Community School District. Availability of data and matching issues limits the analysis to six of the eight schools. Included in this study are Doyle Ryder Elementary School, Eisenhower Elementary School, Freeman Elementary School, Neithercut Elementary School, Pierce Elementary School, and Potter Elementary School. These elementary schools serve students in kindergarten through sixth grade and all grades are retained for analysis.

2.3.A Water Borne Lead Data

The Michigan Department of Environmental Quality launched a formal investigation following the water emergency declaration in Flint. The purpose of the investigation was to understand the severity of the water crisis. An objective was to thoroughly review water borne lead exposure within Flint Community Schools. Every water fixture in the public schools was tested for lead as part of this investigation.

Classrooms in Flint Community Schools each have their own drinking fountain. These fountains are the primary source of water for students throughout the school day. Members of the Michigan Department of Environmental Quality collected four water samples from each fountain between October and December of 2016. Dispersion of lead in the fixtures and plumbing components contributed to heterogeneous levels of lead across classrooms within schools, providing the valuable exogenous source of variation used for identification. The results of these water tests are used to generate the cumulative blood lead exposure measure, the focal explanatory variable. This variable takes into consideration the relatively long half-life of lead deposited in the organs and soft tissues. Details on the construction of this measure are available in Appendix A1.

There were eight classrooms without functioning water fountains during the investigation. In the main analysis those observations are dropped. A robustness check retains the observations and replaces the missing values with zero. The results are broadly consistent under both data specifications.

2.3.B Student Behavior Data

Previous studies have linked lead in children with impulsive or aggressive behavior. A challenge when studying behavior is finding a reliable approach for measuring it. Previous studies have relied on either observed behaviors or disciplinary records. The observed behaviors are usually reported by the individual or caregiver and can suffer from personal bias, lapses in memory, and small sample size. Formal disciplinary or criminal records are more official but mask issues of bias in leniency and discrimination in the system.

This study focuses on reported school disciplinary actions. During the period of study Michigan had a zero tolerance school discipline law. The zero tolerance approach prescribed suspensions and expulsions as both the appropriate and mandatory responses for a long list of problematic behaviors. By the 2012-2013 academic year Michigan suspended and expelled more students annually than any other state in the region. Reform on the zero tolerance policy was not enacted until December of 2016.

The zero tolerance law compelled teachers and administrators to rely on formal disciplinary action. Due to these legislative constraints, formal disciplinary actions are likely to be the margin of response to changes in student behavior which would be captured in the data for this study.

It is not possible to observe whether there is discrimination in the handling of behavioral issues in the data. However, the study follows the same groups of students over time and uses fixed effects to control for static traits such as race and socioeconomic status. Furthermore, while it is possible that individuals may experience discrimination in discipline, the level of observation is at the grade level within schools and the composition of those groups is likely to be similar across grades.

The school disciplinary data is provided by Flint Community Schools following a Freedom of Information Act Request (FOIA). The data includes information about the number of disciplinary actions taken disaggregated by type for each grade within each school. The types of disciplinary actions include detentions, in- and out-of-school suspensions, and expulsions. These annual counts make it possible to track changes in the number of disciplinary actions for a group of students as they progress from one grade to the next.

2.3.C. Summary Statistics

The water sample data from the Department of Environmental Quality is matched to classroom usage with information collected by a separate FOIA request to Flint Community

Schools. The request provided information such as school maps, school rosters, and redacted class lists. These documents were used to determine primary usage of classrooms and class size. Room numbers from the FOIA were matched to the room numbers used in the Department of Environmental Quality's reports.

The classroom lead values are aggregated to the cohort within school level by using a weighted average based on the share of students in a classroom. The counts of disciplinary actions at the cohort within school level are divided by the number of students in the cohort to control for differences in the size of cohorts. As a result, the dependent variable is the average number of disciplinary actions for a student within a school-specific cohort. The variable of interest is the average cumulative blood lead exposure level for a student within that cohort.

Summary statistics are available in Table 2.1. The average cumulative blood lead exposure is much higher during and following the period of exposure. A cumulative blood lead exposure of 2 micrograms is assigned for the 2012-2013 academic year to underscore the point that environmental lead exposure is common[16]. The average number of disciplinary actions per student for each cohort are averaged across the entire sample. These mean average actions show a steady increase over the period of study. Total actions are also increasing over the period, while the population size of the elementary school children remains stable.

2.4 Empirical Strategy

As described previously, this study utilizes the quasi-natural experiment precipitated from the unexpected change in water quality in Flint, Michigan. Focusing within elementary schools, the plausibly exogenous variation across classrooms is used to estimate the causal impact of cumulative blood lead exposure on student disciplinary actions.

---

[16] The regressions were also run using an assumed cumulative blood lead level of 0, suggesting no lead exposure in the pre-period, and the results were consistent.

The primary specification uses a fixed effects approach to follow cohorts of students within schools over time. This provides the added benefit of controlling for unobserved static characteristics of the groups of students within schools. The general model used in this study is highlighted below:

$$Discipline_{c,s,t} = a_0 + \beta_1 Lead_{c,s,t} + \eta_c + \kappa_s + \eta_c{}^{*}\kappa_s + \sigma_g + \tau_t + \iota_t + \mathcal{E}_{c,s,t}$$

The dependent variable, *Discipline*, is the average number of disciplinary actions per student for a cohort *c* within a school *s* in year *t*. The focal explanatory variable, *Lead*, is a cumulative measure of blood lead exposure measured in micrograms. The coefficient of interest, β_1, measures the mean marginal impact of one microgram of lead on average disciplinary actions per student. Also in the regression are fixed effects, such as school fixed effects, cohort fixed effects, school by cohort fixed effects (the unit of analysis), grade fixed effects, year fixed effects, and linear time trends. The standard errors are robust and clustered at the school-specific cohort level. The cumulative blood lead exposure variable is winsorized at the 5 and 95 percentiles to control for outliers.

Identification is driven by variation in the growth of the cumulative blood lead exposure variable within each cohort of students within a school over time. Fixed effects absorb variation driven by the static difference between schools, cohorts, and grades. Year fixed effects pick up common shocks to the Flint Elementary School students. School specific linear trends control for changes within schools over time.

The main assumption underlying this empirical approach is that the level of classroom lead is not systematically correlated with alternative sources of lead exposure after controlling for differences between schools and grades. Another potential source of lead for children during this period is from their homes. Students within an elementary school come from fairly homogenous neighborhoods surrounding the school and siblings often attend the same school but different classrooms. For these reasons, it is unlikely that students who have higher levels of lead exposure at home would be systematically assigned to classrooms with high levels of water borne lead.

Another assumption is that students who behave poorly are not systematically assigned to classrooms with higher levels of lead exposure. First, the classrooms were constructed at the same time and the water fountains provide no visible indication that they may contain higher levels of water borne lead. Second, grade assignment to classrooms has a lot of inertia, with rooms often serving the same purpose year after year. As such, it also seems unlikely room assignment is related to lead levels.

A third concern may be that students who are better behaved and have strong home support may be leaving the school system during the four years of the study and driving up the mean discipline actions per student. The summary statistics in Table 2.1 show that the total number of disciplinary actions are increasing rapidly over the four years of observations. At the same time, the size of the student population remains relatively constant over the period.

As students become older, expectations in the classroom are often held to a higher standard. It is likely that the same is true for student behavior. As a result, it would not be surprising if students in the later grades were more likely to receive discipline for poor behavior. The included grade fixed effects help control for this. Classroom usage rarely changes from year to year as mentioned previously. If an older grade with higher behavioral standards within a school coincidentally had relatively high or low levels of lead, this could potentially impact the estimate. Analysis that includes grade by school fixed effects is included in one of the regressions to address this potential issue.

The magnitudes of the results may not be generalizable to other schools with water borne lead. The absorption of lead into the body and its distribution to the brain and other organs is dependent on student nutrition. Lead can mimic iron and calcium in the construction of the neural synapses in the brain. Students with deficiencies in either of these nutrients may face greater impacts for the same level of lead exposure. Within Flint Community Schools, more than 80 percent of the students qualify for the free lunch program which may signal a greater risk for these deficiencies. Furthermore, it is unclear whether the relationship between lead and discipline is linear. Previous work has suggested that the largest marginal impacts of lead are at relatively low blood lead levels. Considering many Flint students had exposure from their home environments, the magnitudes of the estimated marginal impacts of this study may have been greater if there hadn't been exposure from other sources.

The paper continues with the results. The general model is estimated for all students and all disciplinary actions. The relationship is then further investigated by first disaggregating the types of disciplinary actions and then disaggregating the students by age. This is followed by additional regressions that test the resilience of the estimates to alternative data decisions.

2.5 Results

2.5.A Lead and Disciplinary Action

The basic results confirm a relationship between water borne classroom lead and student behavior. The coefficient of interest measures how the average number of disciplinary actions per student responds to a marginal increase of cumulative blood lead exposure of 1 microgram. The regression analyses include fixed effects for cohorts within schools, grades, schools, and years as well as interactions. Standard errors are clustered at the within-school cohort observation level to control for potential serial correlation. The unexpected change in water quality in Flint, Michigan provides the exogenous shock used to identify the causal impact of lead exposure on groups of students within schools over time.

Table 2.2 provides the estimates for the general model. Column (1) includes standard fixed effects for the cohort by school unit of observation, schools, and years. The coefficient is small and not statistically different from zero, providing little evidence of a relationship between lead and observed disciplinary actions.

The elementary schools are in different areas of the city and serve different student populations. The assumption of similar trends across schools over time is relaxed in column (2) with the inclusion of school specific linear trends. The coefficient is much larger and becomes significant with a marginal impact of .0066 disciplinary actions per student for a 1 microgram increase in cumulative blood lead exposure. At the average cumulative blood lead exposure level of 27 micrograms, a typical cohort of 48 students within a school would have an average of 8.6 additional disciplinary actions per year due to classroom lead exposure.

Column (3) includes the less restrictive school by year fixed effects. While the coefficient is still large and statistically different from zero, it no longer as significant. Together, columns (2) and (3) indicate a relationship between classroom lead exposure and student disciplinary actions.

The specification in column (2) offers a balance between allowing unobserved changes within schools over time while demanding less of the small dataset than the inclusion of school by year fixed effects. The specification does include year fixed effects to control for district-wide shocks during the period of analysis. The remaining results will use the preferred specification of column (2), but the robustness checks will also include the specification of column (3).

Total disciplinary actions are a broad measure of the behavioral impacts of lead on elementary school students. There is an implied difference in the severity of a student's behavior between receiving a short out-of-school suspension versus an expulsion from the school. Generally, minor infractions garner lesser punishments than violent actions. Next the analysis is extended to separately estimate the impact of cumulative blood lead exposure on different types of discipline.

2.5.B Impacts by Discipline Type

Students receive four types of disciplinary action during the period of study. These include detentions, in-school suspensions, out-of-school suspensions, and expulsions. Prior to the change in water quality, ninety-nine percent of the disciplinary actions in the 2012-2013 academic year were out-of-school suspensions. Among the out-of-school suspensions are different levels of severity, with the shortest being SNAP decisions that last less than 24 hours, short suspensions lasting fewer than 10 days, and longer out-of-school suspensions.

Regression analysis using the preferred specification from Table 2.2 is conducted for all out-of-school suspensions, short out-of-school suspensions lasting fewer than 10 days (including SNAP decisions), and other disciplinary actions (detentions, expulsions, and in-school suspensions). Table 2.3 shows the results of the regressions. Column (1) provides the coefficient for all disciplinary actions, taken from Table 2.2 to ease comparison. In column (2), out-of-school suspensions have an estimated mean marginal impact of .0049. Column (3) suggests that most of the changes in out-of-school suspensions are being driven by short suspensions with a coefficient of .0046. The impact on other disciplinary actions is .0016, which is smaller in magnitude but still significant. Altogether, out-of-school suspensions continued to be an important form of discipline in response to classroom behavior with most of the impact coming from short out-of-school suspensions. The results show that alternative forms of discipline are also being utilized in connection to classroom level lead exposure.

The results found in Table 2.3 suggest that most of the behavioral responses resulted in less severe penalties in the form of short out-of-school suspensions. Another dimension of interest is whether there are disparate effects for elementary students of different ages. The physical, emotional, and social development of a fifth grader is different from that of a first grader. The impacts of cumulative of blood lead exposure are further examined by separating the data into more homogenous age groups.

2.5.C Impacts by Age Groups

Children develop rapidly during primary school. Physically, the brain grows rapidly at age eight to reach nearly adult-size. Psychologically, students are quickly moving through three of the four Piaget stages of development[17]. Students age seven and younger typically do not fully understand concepts such as logic, concrete reasoning, and cause and effect. Students between the ages of seven and eleven have the capacity to think logically but still struggle with abstract and hypothetical situations. At age eleven, students reach the ability to think logically, formulate hypotheses and consider multiple possibilities (Wood et al, 2001). Emotionally, puberty usually starts at age ten and impacts the social interactions of students.

These developmental guideposts are used to identify three age ranges in the data: kindergarten through second grade, second grade through fourth grade, and fourth grade through sixth grade. The results for regression analysis of each of these groups is in Table 2.4. A large impact is found for students in second through fourth grade in column (3). The highly

---

[17] Jean Piaget was among the first psychologists to study cognitive development in children. His systematic approach identified concrete stages of development that are not dependent upon formal learning. This structure has informed basic psychology for decades.
McLeod, S. 2015. "Jean Piaget." *Simply Psychology.* https://www.simplypsychology.org/piaget.html.

significant estimated mean marginal impact is .0158, more than twice the magnitude found for the pooled sample. The results for the other age ranges reflect the expected positive relationship, but in column (2) the impact on kindergarten through second grade is very small and insignificant. The impact for older children in fourth through sixth grades is larger in column (4) but not statistically different from zero.

It is possible that teachers may be more strict or lenient depending on the age of the students that they are teaching. Within a school, if the classrooms for an older grade with stricter discipline also experience, on average, lower or higher than average lead exposure it is possible that the coefficient of interest may be impacted.

To help control for this potential effect, an interaction of grade and school fixed effects is introduced into the regression. Table 2.5 shows the new results broken down by age range. The estimates in columns (1), (2), and (3) are very similar to the impacts found in Table 2.4. The estimate for older students in column (4) shows an increase in the magnitude of the estimate and is now statistically different from zero. These results suggest that the older students may also have behavioral responses to the lead exposure.

2.6 Robustness

The robustness of the estimates is examined in several ways. First, the decision to use the average lead-in-water level found in classrooms to construct the variable of interest was made as an approximation of the level of lead consumed throughout the day. However, it may be the case that the highest dose of lead may be a more appropriate choice for analysis. The basic results found in Table 2.2 are replicated using the maximum lead-in-water levels found in classrooms.

Table 2.6 shows the basic results from regressions that used the maximum lead-in-water values. The results follow a similar pattern of significance as those of Table 2.2. The magnitudes are smaller due to the higher average measure of lead exposure. The preferred estimate in column (2) finds a marginal impact of .0026 disciplinary actions per student for a 1 microgram increase in cumulative blood lead exposure. At the new average cumulative blood lead exposure level of 67 micrograms, a typical within school cohort of 48 students would have an average of 8.4 additional disciplinary actions per year due to lead exposure. This is very similar to the 8.6 additional disciplinary actions found in the basic results from Table 2.2.

Second, several of the classrooms did not have working water fountains during the period of study. As a result, it is not possible to calculate the cumulative lead exposure for those students since the primary source of water is not identified. In the previous results of this study, the missing values are dropped from analysis. Retaining those observations and assigning them a value of zero, the minimal level of lead exposure possible, does not substantively change the results. Estimates are available in Table 2.7.

2.7 Discussion

The results from the empirical analyses find significant impacts for cumulative blood lead exposure on disciplinary actions. This provides important insight for a period during which disciplinary actions were increasing rapidly. The summary statistics in Table 2.1 show total disciplinary actions more than triple in the four years of data. To appreciate the magnitudes of the results, a back-of-the-envelope calculation is conducted to estimate how much of the change is coming from classroom lead.

The 2012-2013 academic year was unaffected by the change in water quality. This initial year in the study offers insight into discipline levels prior to the water crisis. If the 680 disciplinary actions for 2025 students are assumed to be indicative of future actions in the absence of the water crisis, then the average actions per student should hover around .34. The actual average actions per student for the 2013-2014 through 2015-2016 academic years is .81 actions per student, an increase of .47.

The preferred estimate from Table 2.1 finds that the average marginal impact of 1 microgram of lead is .0066. The average level of lead exposure for the late three years is 37.7 micrograms. At the average level of exposure, lead can explain an average additional .25 actions per student during the last three years of the study or approximately 1,518 additional disciplinary actions in the six elementary schools studied. This is slightly more than fifty percent of the increase in disciplinary actions for this period.

The remaining increase in the number of disciplinary actions may come from several sources. First, the water crisis impacted all of Flint, Michigan and students almost certainly have some lead exposure outside of school. These random levels of exposure are unlikely to be systematically correlated with the levels of lead within classrooms for students within a school catchment area and so they will not bias the estimates of this study. However, the lead exposure is likely to impact student behavior overall and contribute to the fifty percent of the increase that is not explained by the estimates.

The short period of analysis makes it difficult to observe underlying trends in discipline within schools. School linear trends help control for this in the regression, and so this source may explain part of the unidentified increases in disciplinary actions. It is also possible that there may have been shocks in disciplinary policy at the district level. Year fixed effects would absorb such a shock from the regression and result in the changes remaining unexplained by the estimates.

2.8 Conclusion

This study further substantiates the relationship between lead and adverse behaviors. The findings show that there are significant and meaningful increases in disciplinary actions following exposure to water borne lead in the classrooms. The impacts are concentrated among older students, with children in second through fourth grades having the strongest relationship while those in younger grades experiencing little change explained by the classroom lead exposure.

The disciplinary actions are primarily out-of-school suspensions. Prior to the water crisis this was the most common form of punishment and incidences of out-of-school suspensions increased rapidly during the period of exposure. The relatively short lengths of the suspensions, with most lasting fewer than 10 days, suggests that most of the behavioral changes were not especially egregious.

Further study would benefit from identifying whether the increase in disciplinary action is driven by more students behaving poorly or if students who have a propensity for poor behavior are getting into trouble more often. This examination of the extensive and intensive margins could help to further explain the underlying relationship and inform disciplinary actions in the future. The medical literature has also suggested that boys may have a greater behavioral response. Disaggregating the information by gender may also prove an interesting and informative exercise.

Table 2.1: Summary Statistics

	All Years (1)	2012 (2)	2013 (3)	2014 (4)	2015 (5)
Mean Average Cumulative Blood Lead Exposure	27.5 (43)	2 (0)	7 (6)	56.9 (52.9)	53 (50.6)
Mean Average Disciplinary Actions per Student	.72 (.65)	.37 (.30)	.63 (.54)	.83 (.50)	1.06 (.89)
Mean Average Out of School Suspensions per Student	.65 (.52)	.37 (.30)	.61 (.53)	.76 (.46)	.87 (.63)
Total Disciplinary Actions	5611	680	1281	1569	2081
Total Out of School Suspensions	5061	678	1236	1455	1692
Elementary School Student Count		2025	2132	1991	1982

Columns (2) through (5) are by academic year. 2012 represents the 2012-2013 academic year. A constant value of 2 micrograms of cumulative lead exposure is assumed for the 2012-2013 academic year. Standard deviations are presented in parentheses.

Table 2.2: Lead and Student Discipline

	(1)	(2)	(3)
Average Disciplinary Actions per Student	.0015	.0066**	.0047
	(.0022)	(.0027)	(.0031)
Within School Cohort Fixed Effects	X	X	X
Year Fixed Effects	X	X	X
School Fixed Effects	X	X	X
Grade Fixed Effects		X	X
School x Year Fixed Effects			X
School Linear Time Trend		X	
Observations	147	147	147
Groups	51	51	51

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 2.3: Lead and Types of Discipline

	All Discipline (1)	OSS (2)	Short OSS (3)	Other (4)
Average Disciplinary Actions per Student	.0066**	.0049**	.0046**	.0016**
	(.0027)	(.0021)	(.0020)	(.0007)
Within School Cohort Fixed Effects	X	X	X	X
Year Fixed Effects	X	X	X	X
School Fixed Effects	X	X	X	X
Grade Fixed Effects	X	X	X	X
School x Year Fixed Effects				
School Linear Time Trend	X	X	X	X
Observations	147	147	147	147
Groups	51	51	51	51

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 2.4: Lead and All Discipline by Grade Ranges

	All (1)	K-2 (2)	2-4 (3)	4-6 (4)
Average Disciplinary Actions per Student	.0066** (.0027)	.0011 (.0038)	.0158*** (.0036)	.0037 (.0052)
Within School Cohort Fixed Effects	X	X	X	X
Year Fixed Effects	X	X	X	X
School Fixed Effects	X	X	X	X
Grade Fixed Effects	X	X	X	X
School x Year Fixed Effects				
School Linear Time Trend	X	X	X	X
Observations	147	54	67	70
Groups	51	27	33	35

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 2.5: Lead and All Discipline by Grade Ranges with Added Controls

	All (1)	K-2 (2)	2-4 (3)	4-6 (4)
Average Disciplinary Actions per Student	.0065** (.0029)	.0014 (.0046)	.0158*** (.0032)	.0053 (.0050)
Within School Cohort Fixed Effects	X	X	X	X
Year Fixed Effects	X	X	X	X
School Fixed Effects	X	X	X	X
Grade Fixed Effects	X	X	X	X
School x Year Fixed Effects				
School Linear Time Trend	X	X	X	X
School x Grade	X	X	X	X
Observations	147	54	67	70
Groups	51	27	33	35

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 2.6: Lead and Student Discipline with Maximum Lead Exposure

	(1)	(4)	(3)
Average Disciplinary Actions per Student	.0007 (.0008)	.0026** (.0011)	.0018 (.0012)
Within School Cohort Fixed Effects	X	X	X
Year Fixed Effects	X	X	X
School Fixed Effects	X	X	X
Grade Fixed Effects		X	X
School x Year Fixed Effects			X
School Linear Time Trend		X	
Observations	147	147	147
Groups	51	51	51

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 2.7: Lead and Student Discipline with Missing Values

	(1)	(4)	(3)
Average Disciplinary Actions per Student	.0010 (.0018)	.0053** (.0023)	.0029 (.0029)
Within School Cohort Fixed Effects	X	X	X
Year Fixed Effects	X	X	X
School Fixed Effects	X	X	X
Grade Fixed Effects		X	X
School x Year Fixed Effects			X
School Linear Time Trend		X	
Observations	168	168	168
Groups	60	60	60

These estimates are calculated using data for Flint Community Schools. The years of analysis cover academic year 2012-2013 through academic year 2015-2016. The standard errors are clustered at the school cohort level. Standard errors are in parentheses. The lead exposure variable and all outcome variables have been winsorized at the 5 and 95 percentiles. Significant at the 10% (*), 5%(**), 1%(***) levels.

Chapter 3

Health Professional Shortage Area Designations and Mortality

3.1 Introduction

Good health is important for productivity as well as life satisfaction. Access to healthcare is often a crucial component of maintaining good health. Unfortunately, there are barriers to receiving regular care, the least not being costs. The economic literature has many studies that look at how insurance programs impact health outcomes by increasing access. Such work has looked at the impacts of private health insurance (Black et al, 2017), Medicaid and Medicare health benefits (Huh and Reif (2017), Dowd et al (2011), Weathers and Stegman (2012)), and the mandates of the Affordable Care Act (Akosa et al, 2015). While the costs of health care are an important consideration, they are just one part of the larger issue of access to medical care.

Another essential component of medical care is physical access to healthcare professionals. The financial burden of a medical visit is only one of several logistical concerns if the nearest physicians have no availability or practice hours away. Limited work has been done to understand whether people in areas with physician shortages have generally worse health (Robst and Graham, 2004) as well as whether clinics designed for disadvantaged people impact health outcomes (Bailey and Goodman-Bacon, 2015). The issue of physician access is particularly salient in rural areas where there are low rates of health professionals to the local population.

This study adds to the important work on healthcare access by focusing on the availability of primary care physicians. Availability is measured by simple ratios of primary care physicians to the local population. This measure is broader in nature than other programs that target health resources to specific disadvantaged groups. In general, people in rural areas are more likely to suffer from limited access to doctors but these gaps in access can be difficult to identify and remedy. Health Professional Shortage Areas (HPSAs) were designed to tackle this issue.

The federal government created Health Professional Shortage Areas in the 1970s to identify unmet medical needs. The most common geographic designations are county-level. Ratios of one primary care physician per 3,500 people is the qualifying limit. Other federal and state programs tie incentives to HPSA designations to attract primary care doctors to move to high needs areas. Such incentives include student debt forgiveness, reimbursement rate bonuses, and immigration waivers. The number of county-level designations has grown significantly over the past twenty years.

The impact of the program is estimated by matching rural counties that are in many ways similar but differ in their designation as Health Professional Shortage Areas. Characteristics used for matching include measures of how rural the population is, measures of income and poverty, the unemployment rate as well as others. Each treated county is match with replacement with one and then four control counties. Data is compiled from the American Medical Association, Census Bureau, Center for Disease Control, as well as other federal agencies. Propensity score matching methods are used to estimates the impact of county-level designation on mortality rates. The use of federally reported outcomes as well as propensity score matching methods reduce the potential bias found in other studies stemming from self-reported health measures and the use of poor comparison groups.

The main results find a three percent decline in the mortality rate from the baseline. These findings are robust to data from both the Census Bureau and the Center for Disease Control. The impacts are also estimated for cause-specific death. Heart disease and cancer mortality both decline by three percent, while mortality due to stroke declines six percent. The findings suggest that increasing primary care access results in meaningful declines in mortality.

The role of primary care physicians is generally related to preventative and routine health care services. Mortality is an extreme measure of health and improvements on this margin were not the motivation for the HPSA designation program. The results of this study find significant impacts on mortality and provide evidence that physician location decisions have an important role in securing valuable access to healthcare. It is likely that other measures of improved health that may be difficult to observe are also a consequence of this program.

A critique of the Health Professional Shortage Area program is that it purposely does not consider nurse practitioners or physicians assistants when making designations. These professionals may provide adequate primary care support in the absence of physicians. The findings of this study suggest that on average, the HPSA designation does not lead to a redundancy of medical support but results in meaningful improvements in public health.

The paper continues as follows. First, background information is provided about the Health Professional Shortage Areas and the designation process. This is followed by a review of the previous literature. Next, the data and empirical model are described. Then the results are provided with a brief discussion. Finally, a conclusion is offered.

3.2 Background Information

3.2.A The History of Health Professional Shortage Areas

Health Professional Shortage Areas (HPSAs) were introduced during a time of change in the health care industry. There were important shifts in the 1960s and 1970s in the demand for health care services. Medicaid and Medicare were formed in 1965 and were revolutionary in creating insurance coverage for low-income and elderly populations. During this time, a tax subsidy for employers spurred growth in private employer-sponsored health insurance coverage. The coverage offered valuable tax-free benefits to employees who benefited from lower out-of-pocket expenses (Feldstein, 2011).

The shift in demand for medical services raised concerns about physician shortages. The establishment of new medical schools and residency programs were subsidized by the federal government. The number of physicians grew rapidly from 1970 to 1990 (Lohr, Vanselow, and Detmer, 1996). However, with the development of new medical technologies a growing share of physicians chose to pursue specialties that offered more training, higher pay, and placements near large cities[18].

The 1978 Public Services Act included a provision for Health Professional Shortage Area (HPSA) designations to identify areas with unmet medical needs for the National Health Service Corps (NHSC). The population to primary care physician ratio for the bottom quartile of all U.S. counties in 1978 was used as the defining criteria. The rate of 3500 people per primary care doctor remained the standard through the period of this study (Salinsky, 2010). If a county meets a series of other criteria, it may be considered to have unusually high needs and the rate drops to 3000 people per primary care doctor.

The designation process remained stable for decades. The last changes to the methodology occurred in 1993 when the definition of unusually high need was expanded to include areas with large elderly populations (Government Accountability Office, 2006). A scoring system was formulated in 2002 under the Health Care Safety Net Amendments as a way to differentiate the levels of need among the many designated counties (HRSA, 2003). Other

---

[18] National Health Service Corps (NHSC). "Mission and History." *U.S. Department of Health and Human Services.* https://nhsc.hrsa.gov/corpsexperience/aboutus/missionhistory/index.html (accessed May 5, 2016).

changes have been suggested over the years to improve the process, but the recommendations have been postponed and the methodology remained unaffected (Salinsky, 2010).

The HPSA designation is used by federal and state programs to direct resources to areas with high population to physician ratios. National Health Service Corps were the first to use the designation and still offer scholarships or loan repayment benefits to physicians who commit to working in these areas. Foreign medical graduates who complete their medical residencies in the United States are drawn to HPSAs as part of a visa wavier program. More than 30 federal programs and numerous state programs now offer benefits such as medical reimbursement bonuses, visa waivers, scholarships and loans, as well as professional development (Government Accountability Program, 2006).

3.2.B Health Professional Shortage Area Designations

Health Professional Shortage Areas (HSPAs) can be designated for primary care, mental health, and dental. There are three main types of designations: facility, population, and geography. Facility designations identify overburdened correctional institutions and health clinics for vulnerable populations. Population HPSAs identify underserved populations who have difficulty accessing medical help due to barriers such as income, language or culture, or who have an otherwise high need. For example, Native American tribes are automatically recognized as population HPSAs. The most general and common designations are geographic HPSAs[19].

Geographic HPSAs identify areas that are reasonable for the provision of health services and whose local population faces limited access to medical care due to too few primary care physicians. In addition, the resources in adjoining areas must be overused, distant or otherwise inaccessible[2]. The criteria for designation is the population to physician ratio. The threshold is

[19] Health Resources and Services Administration (HRSA). "Primary Medical Care Designation Overview." *U.S. Department of Health and Human Services*. http://bhpr.hrsa.gov/shortage/hpsas/designationcriteria/primarycarehpsaoverview.html (Accessed May 5, 2016)

3500 people per primary care physician. Areas demonstrating unusually high need can qualify for designation at a lower ratio of 3000 people per primary care physician. Such needs are demonstrated by birth rates as high as 100 births per 1000 childbearing aged women, high infant mortality rates of at least 20 deaths per 1000 births, if more than 20% of the population is living below the poverty line, or if 20% or more of the population is over the age of 65.

The Health Resources and Services Administration (HRSA) who oversees the HPSA program defines rational service areas as being a county or a portion of a county that is disconnected from the rest of the population due to topography, transportation patterns or other distinct characteristics. County-level designations are the most common geographic designations and have been the focus of other studies (Government Accountability Office, 2006). The focus of this paper is on county-level primary care geographic HPSA designations.

There are clear restrictions about which doctors are considered primary care physicians for the purposes of designation. The doctors must be either allopathic (M.D.) or osteopathic (D.O.) certified. The physicians must be trained and serve in primary care related specialties. These include general/family practice, general internal medicine, pediatrics, and obstetrics and gynecology. The physicians must devote their time to primary care patient services rather than administration, research or teaching. Physicians working in other specialties or serving as federal doctors, National Health Service Corps members, or J-1 Visa Waiver recipients are not included. Other medical staff who provide primary care service, such as physician assistant and nurse practitioners, are also not counted[20] (HRSA, Primary Medical Care HPSA Designation Criteria), (Salinksy, 2010).

---

[20] Health Resources and Services Administration (HRSA). "Primary Medical Care HPSA Designation Criteria." *U.S. Department of Health and Human Services*. http://bhpr.hrsa.gov/shortage/hpsas/designationcriteria/primarycarehpsacriteria.html (Accessed May 5, 2016).

There are three steps to apply for designation. First, local governments coordinate with the State Primary Care office to review current requests for designation and to receive guidance to complete the application. Next, supporting documentation is assembled to prove the area meets the criteria for designation. Data for population size, number of physicians, share of the population in poverty, as well as miles and minutes to surrounding primary care doctors are taken from federal sources and amended with state and local data. Finally, the application is submitted to the Department of Health and Human Services and forwarded to the relevant state departments. Once eligibility is confirmed the designation is granted[21].

Oversight of the designated HPSAs is conducted at both the state and federal levels. States are tasked with annually reviewing the designations. The U.S. Department of Health and Human Services requires data to be resubmitted every three years to confirm the status of designations. A list of the current HPSAs is published annually in the Federal Register.

3.2.C Physician Supply and Health Outcomes

Health Professional Shortage Areas (HPSAs) are unique in that they only identify high needs areas. Other state and federal programs offer physicians incentives to work in the designated counties. For HPSAs to have meaningful impacts on local heath, the incentives must be successful in persuading physicians to locate to these places. Then there needs to be a positive health benefits associated with the arrival of the physicians. The economic and medical literatures have studied these two mechanisms independently.

Locational decisions of physicians have been observed and discussed for many years. In the early 1990s the observed preferences of physicians suggested that they were significantly less likely to locate in rural counties (Goetz and Debertin, 1996). However, young physicians may be

[21] Health Resources and Services Administration (HRSA). "How to Apply for HPSA Designation." *U.S. Department of Health and Human Services*. http://bhpr.hrsa.gov/shortage/hpsas/apply.html (Accessed May 5, 2016).

sensitive to location based incentives. Newly minted doctors in the 1960s were found to be fairly elastic to income in their job choices (Hurley 1990, 1991). Bolduc, Fortin and Fournier (1996) found the supply of physicians to Quebec increased by nearly 34% in response to a government incentive scheme in the 1980s. Newly trained doctors were offered grants and relocation expense coverage in exchange for moving to rural regions. In addition, changes were made to the medical fee structure that favored greater reimbursement for services rendered in rural areas.

The impact of HPSA designation on locational choices has been studied indirectly. Holmes (2005) studied the impacts of the National Health Service Corps, whose membership works in HPSAs, on location choices of physicians. The study checks on physician locations over 5 year intervals in the period following completion of their National Health Service Corp responsibilities. After controlling for selection into the program, the study finds that dropping the program would cause a 10% decline in physician coverage for medically underserved communities. Visa waivers, another program that utilizes HPSAs, was studied by Baer, Ricketts, Konrad and Mick (1997). The authors compare rural areas that were designated as HPSAs to those that were not and found that international graduates constitute a larger share of physicians in the designation areas. This provides suggestive evidence that the waiver program is successful in placing foreign physicians in shortage areas.

Programs incentivizing physicians to locate to shortage areas are only effective if there are health benefits for local residents. It is possible that people living in shortage areas may already receive healthcare services from other primary care providers, such as nurse practitioners, or utilize resources in adjoining areas. Under these conditions the services provided by a new primary care physician may be redundant and ineffective at improving health outcomes.

The medical field has studied the link between primary care supply and health. In general, primary care helps to prevent illness and death (Macinko, Shi, and Starfield, 2005) and counties with greater availability of primary care resources, as measured by being in the top 25th percentile of resources, tend to have lower mortality rates (Macinko et al, 2005). In the economic literature, Robst and Graham (2005) find that people living in designated Health Professional Shortage Areas have worse self-reported health. More recently, Li (2014) looks at how the effects of local doctors attenuate with distance in the early 2000s. She finds that distance may limit access even within counties. In addition, when comparing doctors of equal travel distance those located within state had a greater impact on mortality rates than those located just over the border. Bailey and Goodman-Bacon (2015) find that the establishment Community Health Centers in the 1960s and 1970s led to a decline in mortality of 2% for people over the age of 50. This program was aimed at disadvantaged populations and staffed mainly by nurses and social workers.

This paper is the first to study the relationship between HPSA designations and mortality. It is also the first from the discussed literature to use propensity score matching to estimate the impact of designation. The data for this study come from administrative records of the United States government and the American Medical Association. While not infallible, these sources are less likely to suffer from the same bias and measurement error arising from self-reported health measures and sample selection found in survey results. This study uses recent data and covers a broad sample of 48 states over nearly 20 years to estimate the impact of HPSA designation on mortality rates.

3.3 Data

Data for this study are accessed from the Department of Health and Human Resource's 2014-2015 Area Health Resource Files (AHRF). The data can be downloaded from the Health Resources and Services Administration's Data Warehouse[22]. The resource files are sourced from United States federal departments and the American Medical Association.

Matching data for propensity scores are accessed from the Area Health Resource Files. Rural/urban continuum values come from the U.S. Department of Agriculture's Economic Research Service. Per capita income is sourced from the U.S. Department of Commerce's Bureau of Economic Analysis. The unemployment rate comes from the Bureau of Labor Statistic's Local Area Unemployment Statistics File.

Data from the United States Census Bureau include Census division codes, population size, race variables, population density, share receiving food stamps, poverty variables, and median household income. County-level data for median household income, poverty variables, and food assistance were not collected in 1996. Values are generated by averaging the county-level values from 1995 and 1997[23].

3.3.A Health Professional Shortage Areas Variable

The Health Professional Shortage Area designation variable is compiled from several sources. Designation data from 2007 to 2013 were taken from the AHRF. Federal Register lists were requested from the Health Resources and Services Administration. The federally mandated lists were provided for 1995, 1996, 1997, 1999, and 2000. No lists were compiled in the mid-2000's due to a transitional period within the administration. No records have been recovered

[22] Data Warehouse data is available online at https://datawarehouse.hrsa.gov/topics/ahrf.aspx. Data for this paper was access on 9 November 2015.

[23] The general specification is also estimated with 1996 dropped. The results are statistically significant and smaller than those found in Table 3.3.

for 1998 from the Health Resources and Services Administration or the National Archives. Designation status is at the county by year level from the 48 continental states. The analysis spans 19 years, with data covering 1995-1997, 1999-2000, and 2007-2013.

There are some counties that only receive partial designation in the dataset. These counties may be very large or have institutions with particularly high needs. For the main estimation counties that are only partially designated are dropped from analysis. Estimates that retain the partially treated counties are included in the robustness section.

3.3.B Mortality Data

Census mortality data are used for the basic results because they identify the county of residence as the location of death and are used during the HPSA designation process. Census data are accessed through AHRF. Mortality data from the Center for Disease Control (CDC) are used to test the robustness of the results. These data are more detailed and provide information about cause of death. The location of death in CDC data is taken from the death certificate, which may not be the same as the county of residence used by the Census mortality data.

CDC mortality data are accessed using the CDC WONDER Database. The data are publically available and can be downloaded online[24]. Variables include all-mortality, age-adjusted mortality as well as cause-specific mortality. Separate analysis is conducted for mortality resulting from heart disease, cancer and stroke which are studied in the broader literature. ICD-10 codes are used for 1999-2014 data and are consistent with Li (2014). For 1995-1997, ICD-9 codes are chosen for compatibility with the later time period. A list of the codes used is included in Appendix Table A3.1.

Alaska, Hawaii, and Washington, D.C. are dropped from analysis due to lack of data and compatibility with the county designation. Several counties are dropped from Colorado and

---

[24]Data Warehouse data is available online at https://wonder.cdc.gov/.

West Virginia due to missing data. To make the CDC data compatible with the Census data an additional 41 counties are dropped due to very small populations. In total, 3,065 counties are retained for analysis.

3.3.C Summary Statistics

Table 3.1 shows the summary statistics for the treated and control counties. The treated counties tend to be larger as evidenced by the population size but more rural, with a lower population density and a higher index on the rural urban continuum. People in the treated counties are more likely to be receiving assistance through the Supplemental Nutrition Assistance Program (SNAP) or be below the poverty line. The treated counties are also slightly more diverse, with a greater share of the population represented by African Americans, Hispanics, and Native Americans/Alaskans.

Table 3.2 shows the distribution of treated, partially treated and control counties by year. As discussed before, there is no data available for 1998 or 2001-2006. In the 1990s there are more control counties than treated resulting in a robust control group. In the later time period there are significantly fewer control counties. This may result in poorer matches in the late period. Separate results for the early and late periods are also estimated.

3.4 Empirical Strategy

This study uses propensity score matching to estimate the impact of HPSA designation on county-level mortality rates. This approach is used to generate an appropriate comparison group for the counties that receive designation. There may be meaningful differences on average between counties that receive designation and those who are not eligible or choose not to apply. The propensity scores provide a mechanism to choose counties in the control group that are most similar to the counties that receive treatment conditional on a group of variables. The probability of a county being treated is estimated using a logistic regression and a carefully selected collection of covariates. Counties from the treatment and control groups are then matched with replacement based on these estimated probabilities.

Propensity score matching is conducted to control for potential selection issues involved with the designation process. Designation awards are noncompetitive, but they are not automatically conferred. Counties must apply through their states and not all counties that qualify for designation decide to apply. Under the assumptions of conditional independence and common support the propensity score matching produces unbiased estimates (Rosenbaum and Rubin, 1983).

The assumption of conditional independence requires that outcomes be independent of treatment conditional on the covariates. The guidelines for HPSA designations are stringent about which counties qualify and which do not, making it difficult for those who do not meet the criteria to be approved. The covariates included in the estimation are carefully chosen from the federal regulations that dictate designation, variables that impact physician locational choices, as well as others that are correlated with health outcomes.

Among the counties that qualify for designation it is assumed that the decision to apply is not driven by expected future mortality rates. This initiative is focused on increasing access to primary care physicians. These doctors typically provide preventative health benefits rather than emergency medical interventions. The award is noncompetitive, so there is no prioritization at the federal level based on future health outcomes. State and regional controls are used in the estimation to control for states that are more efficient at receiving designation (Government Accountability Office, 2006). It is possible that some counties are better at applying for designation or receive preferential treatment from the state government. As long as this is not systematically correlated with expected mortality rates at the county level it should not bias the estimates.

The overlap condition requires that there exist similar probabilities of designation in both the treatment and control groups. This ensures that similar counties are being compared. Figure 3.1 shows the overlap between the control and treatment groups for the basic results. There appears to be significant overlap between the two groups. Balancing tests and the standardized bias can also provide visual checks for the goodness of fit (Caliendo and Kopeinig, 2005). A balance plot is shown in Figure 3.2 for the control and treatment groups before and after matching. The groups are clearly more similar after matching. The standardized bias is graphed in Figure 3.3. The matched data greatly reduces the bias. The general specification mean and median bias for the propensity scores are 5.1 and 4.5 for the matched, falling near the five percent threshold for sufficiency (Caliendo and Kopeinig, 2005).

The average treatment effect on the treated is estimated. The interest of this paper is how mortality rates are impacted in counties that receive designation. The counties that do not qualify for designation may be generally different from those that do, and it is likely that increasing primary care physicians in counties that have a sufficient supply may result in a redundancy of services.

The propensity scores in this paper are matched with replacement. This ensures that the order in which treatment and control counties are matched does not impact results. The propensity score is estimated with a logistic function. In addition to single matches, the models are run with multiple matches. Standard errors are calculated by STATA statistical package using the two matches method from Abadie and Imbens (2006, 2011, 2012).

The results follow. Matched propensity scores are used to estimate the impact of designation using Census mortality data, CDC mortality data, and CDC age adjusted mortality data. The main specification consistently estimates a three percent decline in mortality. Cause specific mortality is also estimated. A three percent decline is found for heart diseases and cancer while mortality due to stroke declines by six percent.

3.5 Results

3.5.A Basic Results

The basic results for the impact of HPSA designation on mortality are found in Table 3.3. This specification uses mortality data from the Census Bureau and propensity score matching methods. A logistic function is used to generate the scores. In columns (1) and (3) Census division controls are used. State controls are used in columns (2) and (4). All four estimates are highly significant and signal a decline in mortality with HPSA designation. In general, matching with one control results in slightly larger magnitudes while matching with state controls lead to slightly smaller magnitudes. The coefficients show that a HPSA designation results in an average decline of 34 deaths per 100,000 people. This is a three percent decline in mortality against the baseline.

Census mortality data is used to estimate the basic results because it is used in the HPSA designation process. The Center for Disease Control (CDC) tracks more detailed mortality rates at the county level and offers the ability to separate mortality by cause of death. The basic model is estimated again using CDC mortality data and CDC age adjusted mortality data. The detailed data are then used to estimate the impact of HPSA designation on cause-specific mortality.

Table 3.4 shows the impacts of HPSA designation on mortality using CDC data. The results are all highly significant with slightly smaller magnitudes than those found in Table 3.3.

The two CDC datasets find very similar impacts. This suggests that the results are not being driven by differences in the age distribution between the treatment and control counties. Based on these coefficients, a HPSA designation results in an average decline in mortality of 27 deaths per 100,000 people. This is approximately a three percent decline from the baseline.

The results are consistent across different datasets. The following sections will estimate the impacts of HPSA designation on specific causes of death as well as whether there were disparate outcomes between the early and late periods of the study.

3.5.B Impacts by Cause of Death

Table 3.5 presents estimates first by cause of death and then for people over the age of 65. Heart disease deaths are mostly significant, with an average decline of 8 deaths per 100,000. Compared to the baseline this is a three percent decline. The results for cancer are more significant and slightly smaller at 6 deaths per 100,000. From the baseline this is also a three percent decline. The results for death by stroke are also significant at an average decline of 4 per 100,000. This is a six percent decline from the baseline. Finally, the estimated impacts for the population over 65 are all very significant at a decline of nearly 180 per 100,000 people. Compared to the baseline rate this represents a three percent decline.

3.5.C Impacts by the Early and Late Periods

The lapse in administrative records during the early 2000s separates the study into two distinct periods. The earlier period had fewer counties with HPSA designation and a robust set of control counties. The later period had many more counties with the HPSA designation. As a result, it may be difficult to find appropriate comparison counties in the late period. In Table 3.6 results are shown from estimating the early and late periods separately. Column (1) shows the coefficients that were estimated using the basic model. Columns (2) through (5) show the results for when the data is trimmed from the tails.

The early years are all highly significant and nearly twice the magnitude of the general results of Table 3.3. The coefficients are fairly consistent as the tails are trimmed in columns (2) to (5). The average impact of designation is a drop in mortality of 83 people per 100,000, a decline of approximately eight percent from the baseline.

The late years are all insignificant. The results in columns (1) through (3) are positive, suggesting that the HPSA designation increased mortality rates. In columns (4) and (5) the tails are trimmed at .05 and .10 respectively, and the coefficients turn negative. Comparing column (5) to (4), the magnitude becomes larger and the estimate more significant as more is trimmed. This may be because there are few good matches at the tail of the distribution in the late years. Figure 3.4 shows the overlap for the late period, and there is a concentration of propensity scores for the treated counties in the right tail.

Based on these estimates it appears that the early years are driving the results in the basic model. There are several reasons why the results are not as strong in the later period. First, there may be selection issues for counties that don't receive designation until the late period. The eligibility criteria did not change over the period of study but counties who either waited to apply for designation until the later period or became eligible in the later period may be generally different from those of the earlier period.

Another reason the results are not as strong in the later period may be a result of supply and demand. Many of the incentives tied to designation are targeted toward new physicians who are finishing their medical residences. The number of counties designated as Health Professional Shortage Areas was much higher in the late period, but the number of new primary care physicians looking for jobs did not significantly increase during this time. From 2000 to 2007, the number of physicians in primary care specialties finishing their residency programs was stable at 12,500[25]. HPSA designations in the later period may be less likely to attract a physician and as a result the average impact would decline.

Finally, the aging population may be a factor. Many rural physicians are reaching retirement age, and their departure from the labor market may leave some counties with high population to physician ratios. In addition, the oldest baby boomers are entering retirement in the later period of study. Older people tend to use more medical services. This suggests that new primary care physicians may be in higher demand throughout the country and may be less likely to choose high needs areas (China, Park, and Galloway-Gilliam, 2012).

3.6 Robustness

The main results used a logistic regression to estimate propensity scores for nearest neighbor matching. This specification found a consistent three percent decline in mortality using data from both the Census Bureau and the Center for Disease Control. This section will test how sensitive those results are to decisions made in the empirical model and the inclusion of partially treated counties.

Table 3.7 provides the results for three alternative empirical approaches. First, the logistic regression is replaced by a probit regression to generate the propensity scores for the matching model. Then treated and control counties are matched using a nearest neighbor matching model based on a weighted function of covariates. Finally, the nearest neighbor matching model is used to match counties within state and year. The variables used in these

[25] Data come from the annual National Residency Match Program's Results and Data Reports accessed November 2015 from https://www.nrmp.org/main-residency-match-data/

alternative estimation strategies are the same as those used in the basic model of Table 3.3 for ease of comparison.

The propensity scores generated with the probit regression result in estimates that are smaller and less significant than the basic model in Table 3.3. Matching with four controls in columns (3) and (4) lead to results more consistent with the previous estimates. The estimated decline is a little more than two percent from the baseline.

The nearest neighbor matching based on weighted covariates is estimated next. The results are generally more significant than those found by the probit regression, but the magnitudes are still smaller than those found by the original specification. The last row of results for Table 3.6 are estimated using nearest neighbor matching within state and year. These results are highly significant and of similar magnitude to the results found in Table 3.3. The estimated impact of designation is an average decline of three percent in mortality.

Table 3.8 estimates alternative approaches to preserve the partially treated counties. The partially designated counties are first grouped with counties that do not receive any designation. The results are fairly significant with magnitudes smaller than those found by the basic model. The estimated impact is an average decline in mortality of two percent from the baseline. Next the partially treated counties are grouped with counties that received the geographic HPSA designation. These results are all insignificant and much smaller than the previous estimates. In Table 3.8 it is clear that there are differences between the partially treated counties and the counties receiving geographic HPSA designation. Distinguishing between the two is important for estimating the health impacts.

3.7 Discussion

This study finds that HPSA designations result in significant and meaningful declines in mortality. An average three percent decline from the baseline is robust to different data sources and specifications. This is a drop of approximately 30 deaths per 100,000 people, or based on the average population size of the treated counties, an average of 23 people per county. These magnitudes are similar to the results found by Bailey and Goodman-Bacon (2015), who found a two percent decline in mortality from the development of Community Health Centers.

There are many programs that use the HPSA designation to attract physicians to these areas of high need. Unfortunately, this paper is unable to identify which incentive programs are generating the positive health outcomes. More detailed data would be valuable for policy, particularly since the costs of the incentive programs vary greatly. For details about the costs of the different federal programs, please refer to Appendix Table A3.2.

The results show the average treatment effect on the treated. As a result, the estimated impacts may not accurately reflect how untreated counties would benefit from receiving designation. The strongest results in Table 3.6 are found for the earlier period of study. This may be a result of the insufficient group of control counties in the later period. It may also reflect the increase in demand for physicians in shortage areas being met with an insufficient supply of new primary care physicians.

A potential concern for this empirical work would be if places that receive the designation are systematically doing many other things to improve health at the same time. This seems unlikely given the scale of the program, the strict eligibility guidelines, and the low cost to apply for designation.

It is also important to note that the declines in mortality likely understate the benefits of the program. The HPSA designation program was not designed to lower mortality rates. There are many other valuable health measures that may reflect improvements in welfare that cannot be observed from changes in the mortality rates.

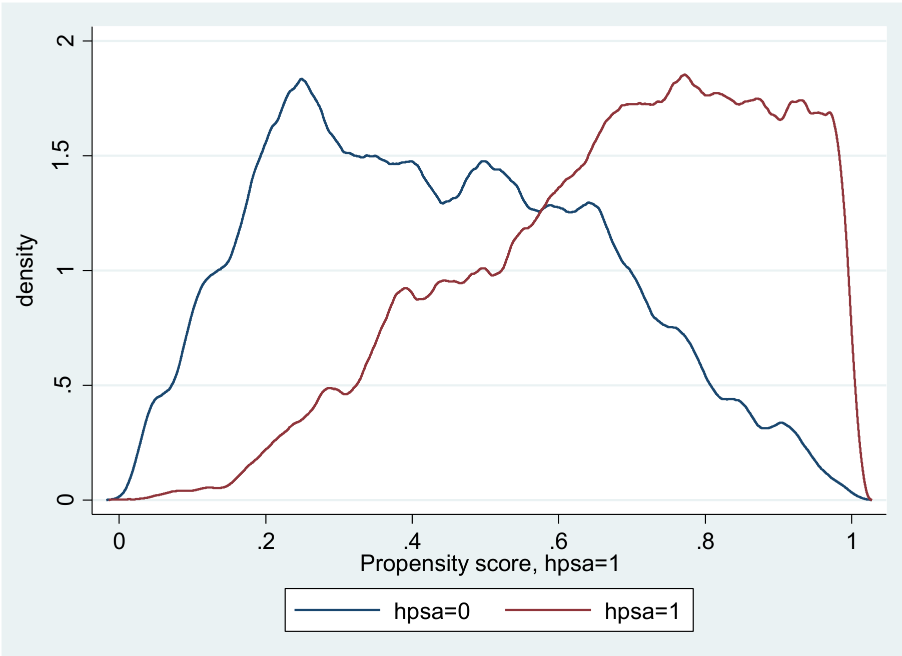
3.8 Conclusion

The importance of access to healthcare has been reflected in legislation and public spending. While insurance and medical cost structures have been the focus of recent public discussion, access is also impacted by the locational decisions of physicians. This study looks at a long-standing program that identifies areas that suffer from a shortage of primary care physicians. The purpose was to increase access to general health services, but this study finds that an unintended result was a meaningful decline in mortality rates.

A decline of three percent from the baseline is consistent across different data sources and specifications. This decline is similar to other initiatives that work to increase access to health professionals. Smaller impacts are found in the later years, which may be a result of a limited control group of counties and demographic shifts in rural areas.

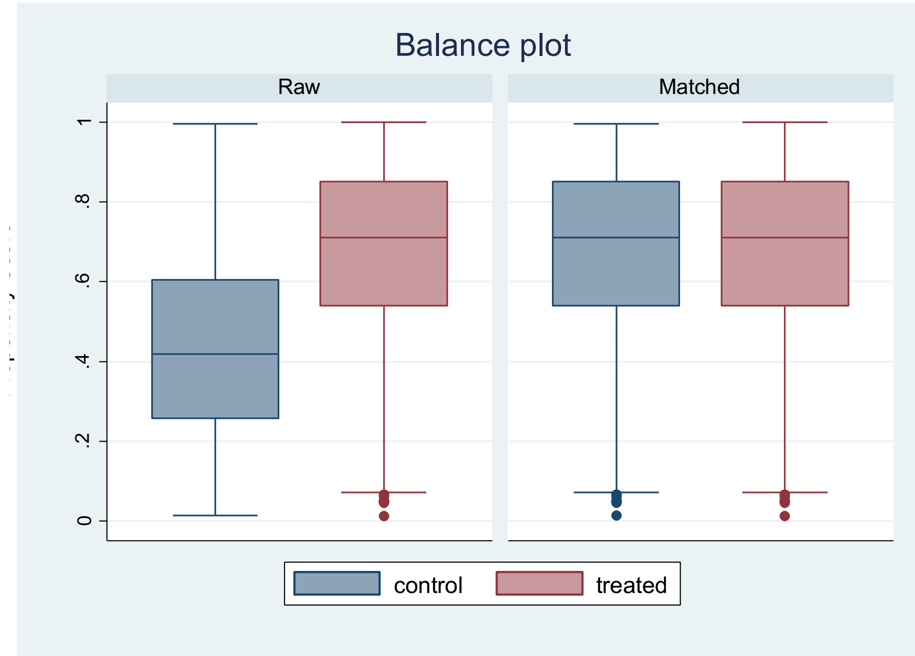
This study is limited by the availability of detailed data. It would be interesting and informative to break down each of the government programs to study which are successfully at placing physicians and whether there are disparate impacts on health. The main results suggest that it is important to continue thinking about health care access broadly, not just for vulnerable populations, and to critically consider barriers to access that extend beyond the financial constraints often at the center of debate.

Figure 3.1: Propensity Score Overlap for Basic Results



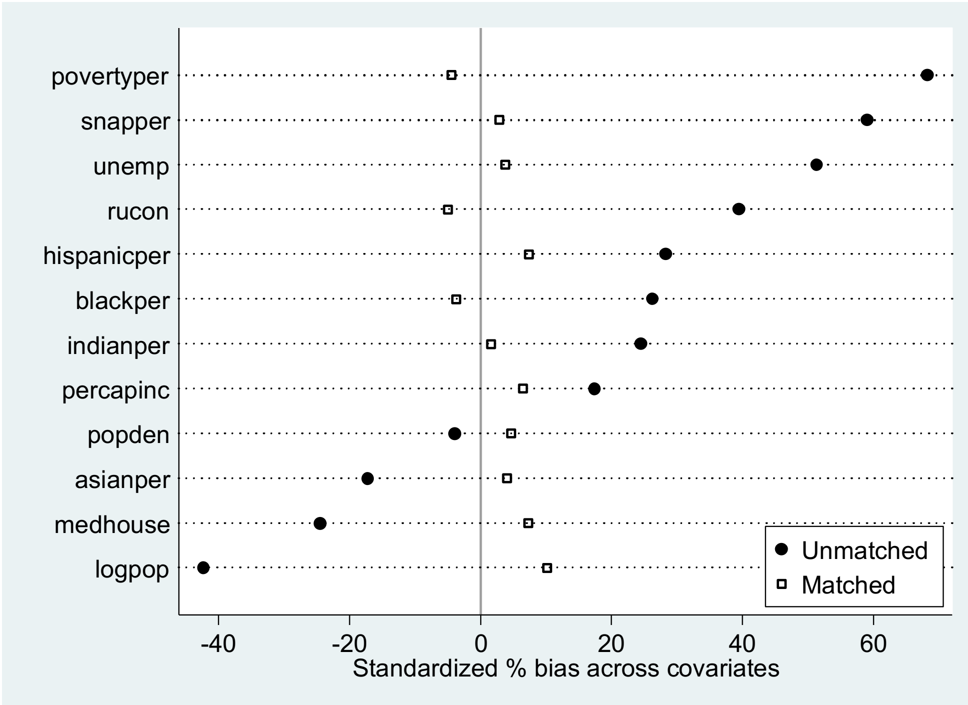
Overlap of propensity scores for the treated and control groups. General specification using one match and controls for year and census division.

Figure 3.2 Box Plot of Basic Results



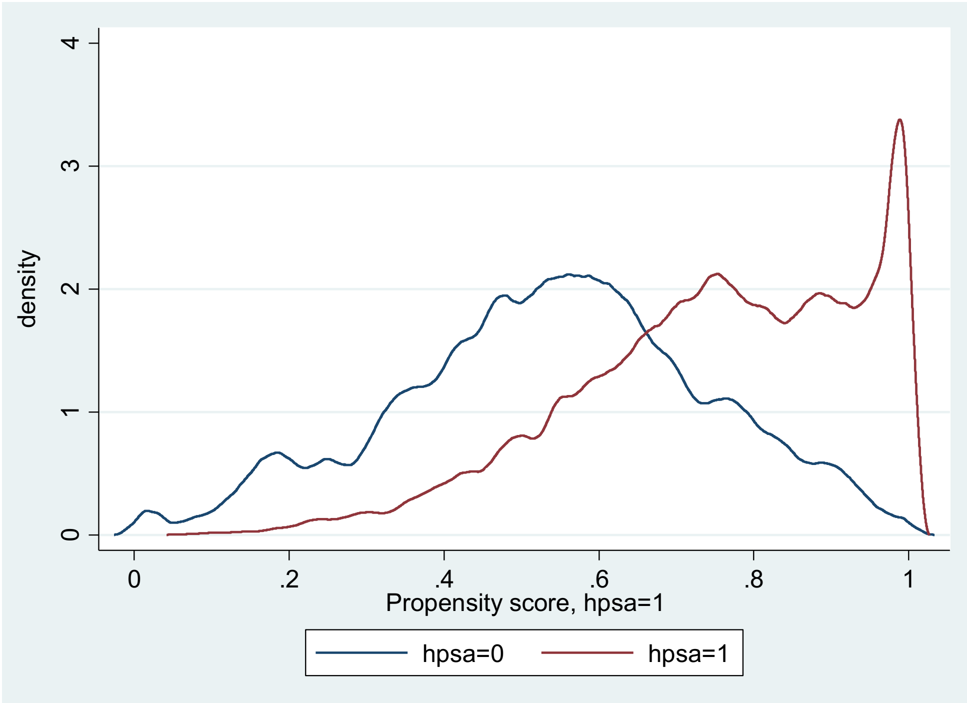
Box plot is the result of general specification one match, with controls for year and census division.

Figure 3.3 Standardized Bias for Matched and Unmatched Counties



Standardized differences of the covariates general specification one match with controls for year and Census division.

Figure 3.4 Propensity Score Overlap for Late Years (2007-2013)



Overlap of propensity scores for treated and control groups from the late period (2007-2013). Graph for the general specification, one match, with controls for year and Census division.

Table 3.1: Summary Statistics with County Level Averages

	HPSA = 0 Control Group (1)	HPSA = 1 Treated Group (2)	Observations (3)
Death Rate	1013 (289)	1021 (288)	22,398
CDC Death Rate	1009 (279)	1046 (272)	22,397
CDC Age Adjusted Death Rate	864 (146)	872 (179)	22,379
CDC Stroke Death Rate	68 (32)	65 (32)	13,636
CDC Heart Death Rate	292 (117)	291 (112)	21,040
CDC Cancer Death Rate	188 (57)	193 (61)	19,933
CDC Old Death Rate	5093 (804)	4890 (907)	22,290
Rural/Urban Continuum	4.69 (2.59)	5.68 (2.73)	22,398
Population Density	235 (660)	183 (1902)	22,398
Population	65,921 (128,344)	76,960 (358,580)	22,398
Poverty Share	.13 (.05)	.17 (.07)	22,398
SNAP Share	.08 (.06)	.13 (.08)	22,398
Median Household Income	40,895 (13,083)	37,860 (10,967)	22,398
Per Capita Income	26,814 (11,016)	28,587 (10,881)	22,398
Unemployment Rate	.06 (.03)	.07 (.03)	22,398
African American Share	.07 (.11)	.11 (.18)	22,398
Hispanic Share	.05 (.09)	.09 (.16)	22,398
Asian Share	.01 (.02)	.01 (.02)	22,398
Native American/Alaskan Native Share	.01 (.03)	.02 (.09)	22,398

Rates are per 100,000 people. Standard deviations are provided in parentheses.

Table 3.2: County Designations by Year

Year	Full HPSA (1)	Partial HPSA (2)	No HPSA (3)
1995	831	1102	1132
1996	858	1147	1060
1997	825	1097	1143
1998	-	-	-
1999	874	1147	1044
2000	803	1152	1110
2007	1278	1034	753
2008	1307	1061	697
2009	1297	1140	628
2010	1279	1255	531
2011	1216	1346	503
2012	1443	1150	472
2013	1139	1458	468

Health Professional Shortage Area county designations for primary care physicians by year. Data is missing for 1998 and 2001-2006.

Table 3.3: Basic Results

Census Mortality Data	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-48.11***	-33.81***	-34.89***	-31.90***
	(7.79)	(7.62)	(6.63)	(6.62)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	22,398	22,309	22,398	22,309

Propensity scores are estimated using logistic functions. Counties are matched with replacement. Standard errors are reported in parentheses (). Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 3.4: Basic Results with CDC Data

CDC Mortality Data	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-30.25***	-27.25***	-29.74***	-24.16***
	(7.51)	(7.38)	(6.44)	(6.40)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	22,379	22,290	22,379	22,290

Age Adjusted CDC Data	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-31.42***	-27.20***	-23.90***	-23.51***
	(5.04)	(6.57)	(4.27)	(5.25)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	22,379	22,290	22,379	22,290

Propensity scores are estimated using logistic functions. Counties are matched with replacement. Standard errors are reported in parentheses (). Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 3.5: Cause of Death with CDC Data

Heart Disease Deaths	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-5.82*	-10.05**	-9.04***	-8.61**
	(3.13)	(4.01)	(2.76)	(3.53)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	21,040	20,951	21,040	20.951

Cancer Deaths	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-9.12***	-5.63**	-6.18***	-5.27***
	(1.77)	(2.23)	(1.53)	(1.94)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	19,933	19,844	19,933	19,844

Stroke Deaths	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-5.57***	-3.82**	-5.84***	-3.83**
	(1.43)	(1.80)	(1.26)	(1.52)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	13,636	13,547	13,636	13.547

Population Age 65+	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-190.75***	-180.08***	-189.74***	-186.96***
	(25.74)	(33.99)	(22.27)	(27.87)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	22,290	22,201	22,290	22,201

Propensity scores are estimated using logistic functions. Counties are matched with replacement. Standard errors are reported in parentheses (). Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 3.6: Late and Early Years with One Match

Early Years 1995-2000	No Trim (1)	Trim .01 (2)	Trim .02 (3)	Trim .05 (4)	Trim .10 (5)
Deaths per 100,000	-88.14***	-79.44***	-82.93***	-83.95***	-82.66***
	(10.66)	(10.69)	(10.32)	(10.61)	(10.13)
Census Division Controls	Yes	Yes	Yes	Yes	Yes
Observations	9,607	9,598	9,565	9,426	9,123

Late Years 2007-2013	No Trim (1)	Trim .01 (2)	Trim .02 (3)	Trim .05 (4)	Trim .10 (5)
Deaths per 100,000	16.44	6.55	13.99	-10.24	-11.12
	(13.84)	(9.93)	(8.83)	(8.38)	(7.87)
Census Division Controls	Yes	Yes	Yes	Yes	Yes
Observations	12,625	11,814	11,631	11,086	9,627

Propensity scores are estimated using logistic functions. Counties are matched with replacement. Standard errors are reported in parentheses (). Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 3.7: HPSA Designation and Mortality Model Robustness Check

Probit Regression	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-21.56***	-15.23	-27.68***	-19.80**
	(8.04)	(10.11)	(6.75)	(8.53)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	22,396	22,299	22,396	22,299

Nearest Neighbor Matching	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-8.72*	-19.34***	-15.02***	-24.15***
	(5.25)	(5.81)	(4.00)	(4.48)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	22,398	22,398	22,398	22,398

Nearest Neighbor Matching Within State and Year	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-34.08***	-32.68***	-56.85***	-53.55***
	(8.20)	(8.07)	(8.37)	(8.07)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	21,689	21,689	19,549	19549

Counties are matched with replacement. Nearest neighbor matching is bias corrected using STATA statistical software. Standard errors are reported in parentheses (). Significant at the 10% (*), 5%(**), 1%(***) levels.

Table 3.8: HPSA Designation and Mortality Data Robustness Check

Partially Treated Counties Included in Control Group	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-11.25**	-10.52**	-18.51***	-16.57***
	(4.72)	(4.97)	(3.86)	(4.12)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	36,780	36,528	36,780	36,528

Partially Treated Counties Included in Treated Group	Matching with One Control (1)	(2)	Matching with Four Controls (3)	(4)
Deaths per 100,000	-8.91	-6.17	-11.50	-6.58
	(6.30)	(9.57)	(5.51)	(7.31)
Census Division Controls	Yes	No	Yes	No
State Controls	No	Yes	No	Yes
Observations	36,780	36,780	36,780	36,780

Propensity scores are estimated using logistic functions. Counties are matched with replacement. Standard errors are reported in parentheses (). Significant at the 10% (*), 5%(**), 1%(***) levels.

Appendices

A1. Calculating the Cumulative Blood Lead Exposure Treatment Variable

The process of lead leaching into the water is not fully understood but studies have identified several factors that are known to impact lead levels in the water. Among these known factors are the water's pH, temperature, conductivity, alkalinity, chloride-to-sulfate ratio, stagnation periods, and use of corrosion inhibitors.

In general, solubility increases when pH decreases, temperature increases, conductivity increases, alkalinity decreases, the chloride-to-sulfate ratio increases, stagnation increases, and with changes to the use of corrosion inhibitors[26] (Giammar et al, 2010).

Monthly water reports from the Flint Water Treatment Plant track many of these variables daily. Among their reports are values for pH, temperature, conductivity, alkalinity, and chloride levels. These monthly reports are available online for the period of study through the Michigan Department of Environmental Quality's reports to the EPA[27]. Periods of stagnation are identified using school calendars and city-wide drinking water warnings.

Cities manage the corrosion in their water pipes several ways. Some cities, such as Boston, use a careful balance of pH and alkalinity to prevent corrosion in old water pipes (Torrice, 2016). Many others use an additive to the water that builds a protective layer in the pipes. These additives are known as corrosion inhibitors. There are different types of corrosion inhibitors, but it is important to be consistent with their use to protect old pipes. Failing to use an inhibitor can put pipes at risk, but even changing from one corrosion inhibitor to another can upset the protective coating in the pipes. In the case of Flint, the corrosion inhibitor

[26] Oram, Brian. "Drinking water issues corrosive water (lead, copper, aluminum, zinc and more)". Water Research Center, http://www.water-research.net/index.php/drinking-water-issues-corrosive-water-lead-copper-aluminum-zinc-and-more (accessed August 28, 2017).
[27] "DEQ Reports to EPA." Taking Action on Flint Water, http://www.michigan.gov/flintwater/0,6092,7-345-76292_76364-376646--,00.html (accessed June 20, 2017).

orthophosphate had been added to their water supply for decades. After the switch in water sources, the City of Flint chose not to add it to their locally treated water (Edwards et al, 2007), (Edwards et al, 2017).

The ex-post values of lead-in-water are used to construct weekly average lead-in-water values. Based on the discontinuation of orthophosphate and the elevated chloride-to-sulfate mass ratios, an increase of 300 to 350% is assumed for the year and half between the change in water and the water tests. Classrooms used in the analysis were separated based on their annual usage. The water change occurred in April 2014, the last quarter of the 2013-2014 academic year. Flint continued to supply its own water through the 2014-2015 academic year and the beginning of the 2015-2016 academic year. Most of the classrooms were occupied for all three years. Some of them were not, and their increased periods of stagnation are accounted for separately.

The weekly average lead-in-water values do not follow a simple linear change. Using the information from the monthly water reports, values fluctuate by units of 5% of the baseline up to 15% weekly. Weeks that school is not in session, such as winter and summer break, are counted as stagnation periods.

To generate lead exposure from lead-in-water values, several assumptions are necessary. From the Department of Agriculture's National Health and Examination Survey, 2/3 of a cup of water is assumed to be consumed from the drinking fountain daily. Of the lead ingested through drinking the water, 20% is assumed to be absorbed into the blood (Toxicological profile for lead, 2007). The volume of water is converted into liters and then multiplied by the absorption rate. This is then multiplied by the weekly average lead-in-water value to find a cumulative blood lead exposure for a typical day that week. Each of these values is then multiplied by the number of instructional days in that week. School calendars and city water warnings are used to identify days that students are not exposed to the drinking fountain water. The resulting product is the cumulative blood lead exposure for that week. The equation below show the calculations conducted.

$$\frac{Avg\ Lead}{Liter\ of\ Water} * \frac{Cups\ Consumed}{Day} * \frac{Liters}{Cup} * Absorption\ Rate * \frac{Instruction\ Days}{Week}$$
$$= Cumulative\ Blood\ Lead\ for\ that\ Week$$

These values are calculated at the within school classroom level, but the study is conducted at the within school cohort level. Since most cohorts are broken into at least two sections, the weekly cumulative blood exposure for the cohort is calculated by taking a weighted average based on class size. The weekly exposure measures are then summed together across time. The treatment variable is the cumulative blood lead exposure prior the test. For the treatment variable in the 2014-2015 academic year the weeks of exposure starting with April 27, 2014 up to April 26, 2015 are summed together. For the 2015-2016 academic year all the weeks of exposure are summed together.

Table A1.1 Student Performance Variable Description

Student Outcome Measure	State of Michigan Performance Level	Description
Proficient	Advanced	The student's performance exceeds grade level content standards and indicates substantial understanding and application of key concepts defined for Michigan students.
Proficient	Proficient	The student's performance indicates understanding and application of key grade level content standards defined for Michigan students.
	Partially Proficient	The student needs assistance to improve achievement. The student's performance is not yet proficient, indicating a partial understanding and application of the grade level content standards defined for Michigan students.
Not Proficient	Not Proficient	The student needs intensive intervention and support to improve achievement. The student's performance is not yet proficient and indicates minimal understanding and application of the grade level content standards defined for Michigan students.

Information is taken from the Michigan Department of Education's M-STEP Guide to Reports: Performance Level Descriptors for Grades 3-8.

Table A1.2 Calculation for Blood Lead Exposure

Lead Water Level microgram/liter	Relevance	Volume of Water Consumed	Absorption Rate	Blood Lead Exposure
5 ppb	FDA bottled water limit	.1577 liters	.20	.1577
10 ppb	Actionable Level for Some Schools	.1577 liters	.20	.3154
15 ppb	EPA Actionable Level	.1577 liters	.20	.4731
50 ppb	EPA Actionable Level Prior to 1991	.1577 liters	.20	1.5770

Example of how blood lead exposure is calculated for one day using water volume and absorption rates based on surveys and the scientific literature.

Table A1.3 Matching Characteristics for Flint and Comparison Schools

	Flint City School District	Carman-Ainsworth Schools	Mt. Morris Consolidated Schools	Roseville Community Schools	Taylor School District
Distance from Flint measured in miles	0	4	7	53	60
Total Students in the School District	6533	5162	2033	4994	7209
Percent Disadvantaged	82.8	70.0	80.3	61.4	73.5
Student to Instructor Ratio	26:1	24:1	30:1	27:1	34:1
Instructional Spending per Pupil	$3465	$6385	$3877	$5287	$5052
Elementary Schools Available for Study	6	3	3	7	8

The comparison schools are system-generated peer districts matched on the above characteristics by the Michigan Department of Education.

Table A3.1: CDC Mortality Codes

	ICD-9 Codes (1995 to 1997)	ICD-10 Codes (1999-2013)
	(1)	(2)
Heart Disease	GR72-320 to GR72-410	GR113-055 to GR113-068
Cancer	GR72-160 to GR72-220	GR113-020 to GR113-036
Stroke	GR72-450 to GR72-470	GR113-070

International Classification of Diseases (ICD) codes are used to identify primary cause of death from CDC mortality data. Codes were chosen from the ICD-9 period to be compatible with the more recent ICD-10 codes.

Table A3.2: Annual Costs per Related Program

	Number of Programs (1)	Cost (2)	Targeted Designations (3)
Indian Health Scholarship Program	1	$9 million	HPSA
J-1 Visa Waivers	3	-	HPSA, MUA, MUP
Medicare Incentive Payment Program	1	$148 million	Geographic HPSA
National Health Service Corps	4	$131 million	HPSA
National Interest Waivers for Immigrant Physicians	1	-	Rural HPSA, Rural MUA
Rural Health Clinic Program	1	$746 million	HPSA, MUA, MUP
Scholarship for Disadvantaged Student Program	1	$47 million	HPSA, MUA, MUP
Title VII Health Professions Education and Training Grant Programs	16	$165 million	HPSA, MUA, MUP
Title VIII Nursing Education Programs	2	$17 million	HPSA, MUA, MUP
	Total Costs:	**$1,263 million**	

All estimates are taken from the Government Accountability Office's 2005 Report.

Bibliography

AAP Council on Environmental Health. 2016. "Prevention of Childhood Lead Toxity." *Pediatrics*, 138 (1): 1-15.

Abadie, Alberto, and Guido W. Imbens. 2006. "Large sample properties of matching estimators for average treatment effects." *Econometrica* 74:235-267.

Abadie, Alberto, and Guido W. Imbens. 2011. "Bias-corrected matching estimators for average treatment effects." *Journal of Business and Economic Statistics* 29: 1-11.

Abadie, Alberto, and Guido W. Imbens. 2012. "Matching on the estimated propensity score." *Harvard University and National Bureau of Economic Research*. Accessed from: http://www.hks.harvard.edu/fs/aabadie/pscore.pdf.

Advisory Committee on Childhood Lead Poisoning Prevention. 2012. "Low Level Lead Exposure Harms Children: A Renewed Call for Primary Prevention." *Centers for Disease Control and Prevention*. Available at: https://www.cdc.gov/nceh/lead/ACCLPP/Final_Document_030712.pdf/.

Aizer, Anna, and Janet Currie. 2017. "Lead and juvenile delinquency: new evidence from linked birth, school and juvenile detention records." The National Bureau of Economic Research Working Paper 23392.

Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier. 2017. "Do Low Levels of Blood lead Reduce Children's Future Test Scores?" *American Economic Journal: Applied Economics.* Forthcoming.

Akoa Antwi, Yaa, Asako S. Moriya, and Kosali Simon. 2015. "Access to health insurance and the use of inpatient medical care: evidence from the affordable care act young adult mandate." *Journal of Health Economics,* 39: 171-187.

Baer, Leonard D., Thomas C. Rickets, Thomas R. Konrad, and Stephen S Mick. 1998. "Do international medical graduates reduce rural physician shortages?" *Med Care* 36: 1534-1544.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The war on poverty's experiment in public medicine: community health centers and the mortality of older Americans." *American Economic Review* 105:1067-1104.

Bekkerman, Anton, and Gregory Gilpin. 2015. "On understanding inconsistent disciplinary behavior in schools." *Applied Economics Letters*, 22(10-12): 772-776.

Bekkerman, Anton, and Gregory Gilpin. 2016. "Crime and punishment: the role of student body characteristics in schools' disciplinary behaviors." *Applied Economics,* 48(13-15): 1402-1415.

Bellinger, David, Herbert L. Needleman, Alan Schell, Alan Leviton, and Elizabeth N. Allred. 1990. “The long-term effects of exposure to low doses of lead in childhood.” *The New England Journal of Medicine,* 322 (2): 83-88.

Black, Bernard, Jose-Antonio Espin-Sanchez, Eric French, and Kate Litvak. 2017. “The long-term effect of health insurance on near-elderly health and mortality.” *American Journal of Health Economics,* 3(3): 281-311.

Bolduc, Denis, Bernard Fortin, and Mar-Andre Fournier. 1996. “The effect of incentive policies on the practice location of doctors: a multinomial probit analysis.” *Journal of Labor Economics* 14:703-732.

Bosworth, Ryan. 2014. “Class size, class composition, and the distribution of student achievement.” *Education Economics*, 22 (2): 141-165.

Bound, John, Charles Brown, Greg J. Duncan, and Willard L. Rodgers. 1994. “Evidence on the validity of cross-sectional and longitudinal labor market data.” *Journal of Labor Economics*, 12 (3): 345-368.

Buddin, Richard, and Gema Zamarro. 2009. “Teacher qualifications and student achievement in urban elementary schools.” *Journal of Urban Economics,* 66 (2): 103-115.

Caldwell, Kathleen L., Po-Yung Cheng, Jeffery M. Jarrett, Amir Makhmudov, Kathryn Vance, Cynthia D. Ward, Robert L. Jones, and Mary E. Mortensen. 2017. “Measurement challenges at low blood lead levels.” *Pediatrics*, 140 (2): e20170272.

Caliendo, Marco, and Sabine Kopeinig. 2005. “Some practical guidance for the implementation of propensity score matching.” *IZA Discussion Paper No. 1588.* Accessed from: http://ftp.iza.org/dp1588.pdf.

Canfield, R. L., T. A. Jusko, and K. Kordas. 2005. “Environmental lead exposure and children’s cognitive function.” *Riv Ital Pediatr*, 31 (6): 293-300.

Canfield, Richard R., Charles R. Henderson, Deborah A. Cory-Slechta, Christopher Cox, Todd A. Jusko, and Bruce P. Lamphear. 2003. “Intellectual impairment in children with blood lead concentrations below 10 µg per deciliter.” *The New England Journal of Medicine,* 348 (16): 1517-1526.

Cameron, A. Colin, and Douglas L. Miller. 2015. “A practitioner’s guide to cluster-robust inference”. *Journal of Human Resources* 50 (2): 317-372.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. “The lasting impact of childhood health and circumstance.” *Journal of Health Economics,* 24: 365-389.

Cecil, Kim M., Christopher J. Brubaker, Caleb M. Adler, Kim N. Dietrich, Mekibib Altave, John C. Egelhoff, Stephanie Wessel, Ilayaraia Elangovan, Richard Hornung, Kelly Jarvis, and Bruce P. Lapphear. 2008. "Decreased brain volume in adults with childhood lead exposure." *PLoS Med*, 5(5) e112.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The value of school facility investments: evidence from a dynamic regression discontinuity design." *The Quarterly Journal of Economics*, 125 (1): 215-261.

Chay, Kenneth Y., Daeho Kim, and Shailender Swaminathan. 2011. "Medicare, hospital utilization and mortality: evidence from the program's origins." Unpublished.

Chen, Aimin, Kim N. Dietrich, James H. Ware, Jerilynn Radcliffe, and Walter J. Rogan. 2005. "IQ and blood lead from 2 to 7 years of age: are the effects in older children the residual of high blood lead concentrations in 2-year-olds?" *Environmental Health Perspectives,* 113 (5): 597-601.

China, Casey, Annie Park, and Lark Galloway-Gilliam. 2012. "Health professional shortage areas: a history and guide to proposed revisions." *Community Health Councils, Inc.*

Chingos, Matthew M. 2012. "The impact of a universal class-size reduction policy: Evidence from Florida's statewide mandate." *Economics of Education Review*, 31 (5): 543-562.

Cho, Hyunkuk. 2017. "The effects of summer heat on academic achievement: A cohort analysis" *Journal of Environmental Economics and Management,* 83: 185-196.

Cobb-Clark, Deborah A., Sonja C. Kassenboehmner, Trinh Le, Duncan McVicar, and Rong Zhang. 2015. "Is there an educational penalty for being suspended from school?" *Education Economics*, 23(3-4): 376-395.

Cohen-Cole, Ethan. 2006. "Housing quality, neurotoxins and human capital acquisition." *Applied Economics Letters*, 13 (12): 753-758.

Contoyannis, Paul, and Martin Dooley. 2010. "The role of child health and economics status in educational, health, and labor market outcomes in young adulthood." *The Canadian Journal of Economics*, 43 (1): 323-346.

Currie, Janet. 2009. "Healthy, wealthy, and wise: socioeocnomics status, poor health in childhood, and human capital development." *Journal of Economic Literature*, 47 (1): 87-122.

Daneshparvar, Maryam, Seyed-Ali Mostafavi, Maryam Zare Jeddi, Masud Yunesian, Alireza Mesdaghinia, Amir Hossein Mahvi, and Shahin Akhondzadeh. 2016. "The role of lead exposure on attention-deficit/hyperactivity disorder in children: a systematic review." *Iranian Journal of Psychiatry,* 11(1): 1-14.

Davis, Matthew M., Christ Kolb, Lawrence Reynolds, Eric Rothstein, and Ken Sikkema. 2016. "Final report." *Flint Water Advisory Task Force.* Available from: https://www.michigan.gov/documents/snyder/FWATF_FINAL_REPORT_21March2016_517805_7.pdf

Diamond, James J., Fred W. Markham, Howard K. Robinowitz, and Jeremy R Wortman. 2008. "Medical school programs to increase the rural physician supply: a systematic review and project impact of widespread replication." *Academic Medicine,* 83: 235-243.

Dowd, Bryan, Matthew L. Maciejewski, Heidi O'Connor, Gerald Riley, and Yisong Geng. 2011. "Health plan enrollment and mortality in the Medicare program." *Health Economics,* 20(6): 645-659.

Edwards, Marc, and Simoni Triantafyllidou. 2017. "Chloride-to-sulfate mass ratio and lead leaching to water." *Journal American Water Works Association,* 99 (1): 96-109.

Edwards, Marc, Carolyn Elfland, and Paolo Scardina. 2010."Lead-contaminated water from brass plumbing devices in new buildings." *Journal American Water Works Association,* 102 (11): 66-76.

Edwards, Marc A., Kelsey J. Pieper, and Min Tang. 2017. "Flint water crisis caused by interrupted corrosion control: investigating "ground zero" home." *Environmental Science & Technology*, 51 (4): 2007-2014.

Epstein, Roy J. 1998. "Lead paint, toxic torts, and the housing stock: a case study in risk assessment." *Journal of Real Estate Finance and Economics,* 17 (2): 127-138.

Feigenbaum, James J., and Christopher Muller. 2016. "Lead exposure and violent crime in the early twentieth century." *Explorations in Economic History*, 62: 51-86.

Feldstein Paul J. 2011. "The rise of medical expenditures." Health Policy Issues: An Economic Perspective, Fifth Edition. *Health Administration Press.*

Gazze, Ludovica. 2017. "The price and allocation effects of targeted lead abatement mandates." *Working Paper.* Accessed from http://home.uchicago.edu/~lgazze/webfiles/LGazze_LeadHousing_042017.pdf

Gerritsen, Sander, Erik Pug, and Dinand Webbink. 2017. "Teacher quality and student achievement: evidence from a sample of Dutch twins." *Journal of Applied Econometrics,* 32 (3): 643-660.

Giammar, Daniel E., Katherine S. Nelson, James D. Noel, and Yanjiao Xie. 2010. "Water chemistry effects on dissolution rates of lead corrosion products." Water Research Foundation.

Goetz, Stephan J., and David L. Debertin. 1996. "Rural-urban locational choices of medical doctors: a county-level analysis." *Review of Agricultural Economics* 18: 547-563.

Government Accountability Office, 1995. "Healthcare shortage areas: designations not a useful tool for directing resources to the underserved." *GAO Publication*: *GAO/HEHS-95-200*. Washington, DC: U.S. Government Printing Office.

Government Accountability Office. 2006. "Health professional shortage areas: problems remain with primary care shortage area designation system." *GAO Publication: GAO-07-84.* Washington, DC: U.S. Government Printing Office.

Harris, Douglas N., and Tim R. Sass. 2011. "Teacher training, teacher quality and student achievement." *Journal of Public Economics,* 95 (7): 798-812.

Health Resources and Services Administration (HRSA), US Department of Health and Human Services. 1995. "Lists of designated primary medical care, mental health, and dental health professional shortage areas." *Federal Register,* 60(190): 51518-51655.

Health Resources and Services Administration (HRSA), US Department of Health and Human Services. 2003. "Criteria for determining priorities among health professional shortage areas." *Federal Register,* 68(104): 32531-32533.

Health Resources and Services Administration (HRSA), US Department of Health and Human Services. "Guidelines for primary medical care/dental HPSA designation." Accessed 5 May 2016 from http://bhpr.hrsa.gov/shortage/hpsas/designationcriteria/medicaldentalhpsaguidelines.html.

Health Resources and Services Administration (HRSA), US Department of Health and Human Services. "Medically underserved areas/populations." Accessed 5 May 2016 from http://www.hrsa.gov/shortage/mua/index.html.

Health Resources and Services Administration (HRSA), US Department of Health and Human Services. "Rural health clinic automatic HPSA process." Accessed 5 May 2016 from http://bhpr.hrsa.gov/shortage/hpsas/ruralhealthhpsa.html.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1997. "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme." *Review of Economic Studies,* 64: 605-654.

Holmes, George M. 2005. "Increasing physician supply in medically underserved areas." *Labour Economics,* 12: 697-725.

Hong, Kai, and Ron Zimmer. 2016. "Does investing in school capital infrastructure improve student achievement?" *Economics of Education Review*, 53: 143-158.

Hoxby, Caroline M. 2000. "The effects of class size on student achievement: new evidence from population variation." *Quarterly Journal of Economics*, 115 (4): 1239-1285.

Huh, Jason, and Julian Reif. 2017. "Did Medicare Part D reduce mortality?" *Journal of Health Economics,* 53: 17-37.

Hurley, Jeremiah. 1990. "Simulated effects of incomes-based policies on the distribution of physicians." *Medical Care,* 3: 221-238.

Hurley, Jeremiah. 1991. “Physicians’ choices of specialty, location and mode: a reexamination within an interdependent decision framework.” *The Journal of Human Resources,* 26: 47-71.

Jones, David J. 2012. “Primary prevention and health outcomes: treatment of residential lead-based paint hazards and the prevalence of childhood lead poisoning.” *Journal of Urban Economics*, 71 (1): 151-164.

Jordan, Jeffrey L., and Bulent Anil. 2009. “Race, gender, school discipline, and human capital effects.” *Journal of Agricultural and Applied Economics*, 41(2): 419-429.

Kaufman, Alan S. 2001. “Do low levels of lead produce IQ loss in children? A careful examination of the literature.” *Archive of Clinical Neuropsychology,* 16 (4): 303-341.

Kessler, Lawrence M., and Murat K Munkin. 2015. “Bayesian estimation of panel data fractional response models with endogeneity: An application to standardized test rates.” *Empirical Economics*, 49 (1): 81-114.

Kinsler, Josh. 2011. “Understanding the black-white school discipline gap.” *Economics of Education Review,* 30(6): 1370-1383.

Kinsler, Josh. 2013. “School discipline: a source or salve for the racial achievement gap?” *International Economic Review,* 54(1): 355-383.

Lanphear, Bruce P., Kim Dietrich, Peggy Auinger, and Christopher Cox. 2000. “Cognitive deficits associated with blood lead concentrations <10 μg/dL in US children and adolescents.” *Public Health Reports,* 115 (6): 521-529.

Levine, Phillip B., and Diana Whitmore Schanzenbach. 2009. “The impacts of children’s public health insurance expansions on educational outcomes.” *Forum for Health Economics and Policy: Frontiers in Health Policy Research,* 12 (1).

Li, Jing. 2014. “The influence of state policy and proximity to medical services on health outcomes.” *Journal of Urban Economics,* 80: 97-109.

Lidsky, Theodore I., and Jay S. Schneider. 2003. “Lead neurotoxicty in children: basic mechanisms and clinical correlates.” *Brain*, 126 (1): 5-19.

List, John A., Daniel L. Milliment, Per G. Fredriksson, and W. Warren McHone. 2003. “Effect of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator.” *The Review of Economics and Statistics,* 85: 944-952.

Lohr, K. N., N. A. Vanselow, and D. E. Detmer. 1996. “U.S. physician supply and requirements: match or mismatch?” The Nation’s Physician Workforce: Options for Balancing Supply and Requirements. *Institution of Medicine Committee on the U.S. Physician Supply.* Accessed 5 May 2016 from http://www.ncbi.nlm.nih.gov/books/NBK232541/.

Macinko, James, Robert Politzer, Leiyu Shi, Barbara Starfield, John Wulu, and Jiahong Xu. 2005. "Primary care, social inequalities and all-case, heart disease and cancer mortality in US counties: a comparison between urban and non-urban areas." *Public Health,* 119: 699-710.

Macinko, James, Leiyu Shi, and Barbara Starfield. 2005. "Contribution of primary care to health systems and health." *The Milbank Quarterly,* 83: 457-502.

Marcotte, Dave E. 2017. "Something in the air? Air quality and children's educational outcomes." *Economics of Education Review,* 56: 141-151.

Marcotte, Dave E. 2015. "Allergy test: seasonal allergens and performance in school." *Journal of Health Economics*, 40: 132-140.

Martorell, Paco, Kevin Strange, and Isaac McFarlin Jr. 2016. "Investing in schools: capital spending, facility conditions, and student achievement." *Journal of Public Economics*, 130: 13-29.

Mason, Lisa H., Jordan P. Harp, and Doug Y. Han. 2014. "Pb neurotoxicity: neuropsychological effects of lead toxicity." *BioMed Research International,* 840547.

National Health Services Corps (NHSC), US Department of Health and Human Services. "About us." Accessed 5 May 2016 from https://nhsc.hrsa.gov/corpsexperience/aboutus/index.html.

Robst, John, and Glenn G. Graham. 2004. "The relationship between the supply of primary care physicians and measures of health." *Eastern Economic Journal,* 30: 467-486.

Needleman, Herbert L., Christine McFarland, Roberta B. Ness, Stephen E. Fienberg, and Michael J. Tobin. 2002. "Bone lead levels in adjudicated delinquents: a case control study." *Neurotoxicology and Teratology*, 24(6): 711-717.

Needleman, Herbert. 2004. "Lead poisoning." *Annual Review of Medicine,* 55: 209-222.

Nevin, Rick. 2000. "How lead exposure relates to temporal changes in IQ, violent crime, and unwed pregnancy." *Environmental Research*, 83(1): 1-22.

Nevin, Rick. 2007. "Understanding international crime trends: the legacy of preschool lead exposure." *Environmental Research,* 104(3): 315-336.

Papke, Leslie E., and Jeffrey M. Wooldridge. 2008. "Panel data methods for frational response variables with an application to test pass rates." *Journal of Econometrics*, 145: 121-133.

Park, Jisung. 2017. "Hot temperature, human capital, and adaptation to climate change." *Working paper.* Accessed from https://scholar.harvard.edu/jisungpark/publications/heat-stress-and-human-capital-production-job-market-paper

Rau, Tomas, Loreto Reyes, Sergio S. Urzua. 2013. "The long-term effects of early lead exposure: evidence from a case of environmental negligence." *NBER Working Paper 18915.*

Reyes, Jessica Wolpaw. 2007. “Environmental policy as social policy? The impact of childhood lead exposure on crime.” *The B.E. Journal of Economic Analysis & Policy,* 7(1).

Reyes, Jessica Wolpaw. 2015. “Lead Policy and Academic Performance: Insights from Massachusetts.” *Harvard Educational Review* 85 (1): 75-107.

Reyes, Jessica Wolpaw. 2015. “Lead Exposure and Behavior: Effects on Antisocial and Risky Behavior Among Children and Adolescents.” *Economic Inquiry* 53 (3): 1580-1605.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. “Teachers, schools, and academic achievement.” *Econometrica,* 73 (2): 417-458.

Robinson, Christina, and Nicole M. Coomer. 2014. “Children’s public health insurance and scholastic success.” *Applied Economics Letters,* 21 (7): 459-462.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. “The central role of the propensity score in observational studies for causal effects.” *Biometrika,* 70: 41-55.

Salinsky, Eileen. 2010. “Health care shortage designations: HPSA, MUA, and tbd.” *National Health Policy Forum,* Background Paper No. 75. Accessed 5 May 2016 from https://www.nhpf.org/library/background-papers/BP75_HPSA-MUA_06-04-2010.pdf.

Sebastian, Rhonda S., Cecilia Wilkinson Enns, and Joseph D. Goldman. 2011. “Drinking water intake in the US: what we eat in America NHANES 2005-2008.” *Food Surveys Research Group Dietary Data Brief No. 7.* Available from: http://are.usda.gov/services/docs.htm?docid=19476

Shih, Regina A., Howard Hu, Marc G. Weisskopf, and Brian S. Schwartz. 2007. “Cumulative lead dose and cognitive functions in adults: A review of studies that measured both blood lead and bone lead.” *Environmental Health Perspectives*, 115 (3): 483

Smith, Jeffrey A., and Petra E. Todd. 2005. “Does matching overcome LaLonde’s critique of nonexperimental estimators?” *Journal of Econometrics,* 125: 305-353.

Stafford, Tess M. 2015. “Indoor air quality and academic performance.” *Journal of Environmental Economics and Management*, 70: 34-50.

Stretesky, Paul B., and Michael J. Lynch. 2001. “The relationship between lead exposure and homicide.” *Archives of Pediatrics & Adolescent Medicine*, 155(5): 579-582.

Torrice, Michael. 2016. “How lead ended up in Flint’s tap water.” *Chemical and Engineering News,* 94 (7): 26-29.

“Toxicological profile for lead.” *US Department of Health and Human Serivces: Agency for Toxic Substances and Disease Registry*. (2007) Available from: https://www.atsdr.cdc.gov/toxprofiles/tp13.pdf

Wood, Kay C., Harlan Smith, and Daurice Grossniklaus. 2001. "Piaget's stages of cognitive development." *Emerging perspectives on learning, teaching and technology.* Retrieved March 16, 2018, from http://projects.coe.uga.edu/epltt.

Zhang, Nanhua, Harolyn W. Baker, Michael R. Elliott, Randall E. Raymond, Hamisu Salihu, and Margaret Tufts. 2013. "Early childhood lead exposure and academic achievement: evidence from Detroit Public Schools, 2008-2010." *American Journal of Public Health,* 103 (3): e72-e77.

Vita

Department of Economics	Email: jesauve@syr.edu
Syracuse University	Mobile: 315-530-2944
110 Eggers Hall	Website: jesauve.expressions.syr.edu
Syracuse, NY 13244	

Education

Syracuse University	
B.A. in Public Relations and Public Policy	May 2009

Employment

Primary Instructor	
Intermediate Microeconomics (ECN 301)	Summer Session 2015
Survey of Microeconomic Theory (ECN 601)	Summer Session 2015
Teaching Assistant	
Economic Ideas and Issues (ECN 203)	Fall 2017 – Spring 2018
Econometric Methods (ECN 522)	Fall 2016 – Spring 2017
Survey of Microeconomic Theory (ECN 601)	May 2015
Intermediate Microeconomics (ECN 301)	Fall 2014 – Spring 2015
	Fall 2013 – Spring 2014

Awards and Honors

Certificate in University Teaching	Fall 2017
Outstanding Teaching Assistant Award	Spring 2016
Syracuse University Graduate Fellowship	Fall 2015 – Spring 2016
	Fall 2012 – Spring 2013
Syracuse University Graduate Assistantship	Fall 2016 – Spring 2018
	Fall 2013 – Spring 2015
Maxwell School Summer Fellowship	Summer 2012 – Summer 2017