# ESSAYS ON HEALTH ECONOMICS

A Dissertation
Presented to
The Academic Faculty

By

Amanda Loyola Heufemann
B.A., University of Chile, 2013
M.A., University of Chile, 2014

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in the Department of Economics at Brown University

Providence, Rhode Island
Febraury 2020

This dissertation by Amanda Loyola Heufemann is accepted in its present form

by the Department of Economics as satisfying the

dissertation requirements for the degree of Doctor of Philosophy.

Date _____ _____

Emily Oster, Advisor

Recommended to the Graduate Council

Date _____ _____

Anna Aizer, Reader

Date _____ _____

Bryce Steinberg, Reader

Approved by the Graduate Council

Date _____ _____

Andrew G. Campbell, Dean of the Graduate School

# VITA

Amanda Loyola Heufemann was born in Santiago, Chile in 1990. She received her Bachelor's degree in economics and business from the University of Chile in 2013, and her Master's degree in Economics from the same university in 2014. Amanda received a Ph.D. in Economics from Brown University in 2020. While at Brown, she was the recipient of the Abramson Award for her third-year research paper. After completing her Ph.D., Amanda became an Economist at Amazon.

## ACKNOWLEDGEMENTS

I would like to thank my advisors, Anna Aizer and Emily Oster, for their continuous support and encouragement. I have learned so much from both of you and I truly appreciate everything you have done for me over the years. I am grateful to Angelica Spertini for all her invaluable work and dedication.

I am incredibly grateful to my friends Carolina Lopez, Catalina Ravizza, Felipe Bruges, Julia Tanndal, Samuele Giambra, Gabriele Borg, Matteo Iudice, Juan Pablo Uribe, and Diego Ramos-Toro for cheering me up and supporting me. You all provided many laughs and friendships that transcend this academic journey. I am also grateful to my high school and college friends Celeste Carreno, Zoe Montero, Tatiana Reyes, and Bernardita Munoz.

To my parents. I am beyond grateful for all the opportunities you provided for me. This academic achievement is a reflection of your hard work and love.

Finally, I would like to thank my husband Alonso Bucarey. I am forever grateful for your unconditional love and for always believing in me.

## Preface

This dissertation explores different topics on health economics. The **first chapter** quantifies the supply response of primary care physicians to a reimbursement bonus, the Physician Scarcity area bonus, introduced by Medicare between 2005 and 2008. The bonus represented a 5% increase in reimbursement and it was paid to healthcare providers working in underserved areas of the U.S. I test whether this incentive attracted more physicians to these areas, if it affected the number of patients served, and if it changed the quantity of care provided by physicians. This is a relevant question as the country faces a shortage of primary care physicians that is expected to increase in the future. Using Medicare claims data between 2000 and 2010 combined with a difference-in-difference strategy, I find that the bonus had either a weak or undetectable impact on these outcomes. Descriptive studies and surveys suggest that the results could be explained by lack of information about the bonus and its small relative size.

In the **second chapter** I explore health effects of the Flint water crisis on the elderly population. The water crisis of 2014-2015 exposed the residents of the city to high levels of lead and other contaminants. Much of the existing evidence of lead exposure comes from studies focused on children, as brains during childhood are more vulnerable than during adulthood. In April 2014, the city of Flint changed its water source from Lake Huron to Flint River, exposing its residents to high levels of bacterial and chemical contaminants, including lead. For approximately a year and a half, local authorities reassured residents that the water was safe for human consumption. Finally, in December 2015 the Mayor of Flint declared state of emergency. Using the universe of Medicare beneficiaries living in Michigan between 2012 and 2016, and the CDC's National Detailed Mortality Files between 2010 and 2017, combined with a difference-in-difference strategy and synthetic controls methodology, I estimate the effect of the change in the water supply on hospitalization and mortality rates. My results suggest a weak increase in all-cause and

kidney disease hospitalization rates in Flint after the start of the crisis, and an increase in all-cause mortality rates. However, these results are usually smaller than what has been suggested by the epidemiological literature and the graphical evidence does not strongly support them.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1
# FINANCIAL INCENTIVES AND MEDICAL PRACTICE IN UNDERSERVED AREAS

## 1.1 Introduction

The U.S. faces a shortage of primary care physicians (PCPs). According to an analysis from the U.S. Department of Health and Human Services, the difference between demand and supply of PCPs was equal to 7,880 physicians in 2013, and it is expected to increase to 23,640 physicians by 2025.[1] Across the country this problem manifests itself in a heterogeneous way, with some areas being considered as underserved in terms of health care provision. Many policies attempt to address this issue by increasing the training of new physicians, alleviating student debt if choosing to work in these areas, or by providing financial incentives to health care professionals, like Medicare programs offering bonus payments to physicians working in underserved areas. Although there is evidence of a positive relationship between health care supply and reimbursement rates, the literature has not looked at this relation in the particular context of underserved areas and how this could be affecting the development of federal incentive programs.

In this paper I study if financial incentives shape medical practice in these areas by looking at a particular Medicare program called Physician Scarcity Area (PSA) bonus. This program ran between 2005 and 2008 and implied a 5% increase in payments to providers working in underserved areas. Physician scarcity areas were defined using the ratio of physicians to Medicare providers at the county level. A cutoff in this ratio created two groups of counties, those where providers received the bonus and those where providers did not, providing a setting with a control and treatment group that allows me to study

[1] The report considers the following medical specialties: General and Family Medicine, General Internal Medicine, Geriatrics, and General Pediatrics.

the effects of the bonus.

I use Medicare claims data to construct a measure of health care supply that is comparable across different services. This measure is based on all the resources needed to provide each medical service, from an office visit to a major surgery. The data also allow me to identify the number of providers and patients in different areas. My study population is a 20% random sample of Medicare beneficiaries between 2000 and 2010. I observe every service billed to Medicare and patient characteristics, including the zip code where services were provided.

To obtain an estimate of the causal effect of financial incentives on health care supply in underserved areas, I combine the data with a differences-in-difference strategy. In particular, I compare the quantity of health care supplied in counties that were and were not subject to the PSA bonus, before and after the start of the program in 2005. The absence of differences in the pre trends of outcome variables between the two groups of counties validates the use of this strategy.

Despite previous evidence, I find that financial incentives do not seem to impact the provision of health care in underserved areas. My results suggest that during the period the bonus was in place, the quantity of health care provided in underserved areas slightly increased, while the number of physicians working in those areas, and the number of visits to the doctor did not change. However, I find suggestive evidence of responses concentrated on more elective health procedures.

Previous literature on financial incentives and treatment choice shows a positive relationship between health care supply and reimbursement rates. In particular, Clemens and Gottlieb (2014) exploit a geographic reimbursement shock that affected physicians participating in Medicare in 1997. They find an elasticity of aggregate health care supply of around 1.5 with respect to reimbursement rates. In the context I study this would imply an increase of around 7.5% in aggregate health care after the introduction of the bonus. I am able to reject this result with my estimates.

Two potential mechanisms could be driving my results. Survey and descriptive evidence suggests that the saliency and relative amount of these incentives are key issues impacting providers and their behavior. Understanding how these mechanisms interact with each other is important in order to put my results in the context of the literature and to think about the design of future incentive programs. This evidence points out to the role of information frictions as a determinant of the degree of response to incentives in this context.

This paper contributes to a small and descriptive literature on financial incentives and physician retention in rural areas (Shugarman et al., 2003; Shugarman et al., 2001; Rabinowitz et al., 2001). It also contributes to a large literature about financial incentives and medical treatment choice (Alexander, 2015; Clemens and Gottlieb, 2014; Coey, 2013; Grant, 2009; Hadley et al., 2009; Hadley et al., 2001; Gruber et al., 1999; Yip, 1998; Keeler and Fok, 1996; Dranove and Wehner, 1994). These papers generally find a positive relation between supply of medical services and reimbursement rates.

The rest of this paper is organized as follows. Section 1.2 discusses the background of the PSA bonus. Section 1.3 presents the data. Section 1.4 presents the empirical strategy. Section 1.5 presents the primary results on the changes in health care and labor supply in response to the bonus. Section 1.6 discusses mechanisms for this effect and Section 1.7 concludes.

## 1.2 Background on Physician Scarcity Area Bonus

Section 413(a) of the Medicare Modernization Act put in place a 5 percent bonus payment for physicians (primary care and specialties) practicing in PSAs. According to the Congressional Record "*The new five percent bonus for physicians in either primary care scarcity counties or specialty care scarcity counties will increase financial incentives for physicians to provide care to Medicare beneficiaries in these areas with a shortage of physicians. This bonus payment will make it easier to recruit and retain physicians in these*

*scarcity areas.*" (Cong. Rec. H12043, 2003). In an address to the House, U.S. Representative for Illinois John Shimkus said "Mr. Speaker, there is a shortage of physicians in Illinois because of medical malpractice, and physicians are leaving the State. At least in the Medicare prescription drug bill, there is help for keeping some of those" (Cong. Rec. H1487, 2004).

Shortage of physicians, especially PCPs providing care in underserved areas, is a growing problem in the U.S. that is expected to increase given the current trends in population growth and aging. The lack of professionals providing this type of care is particularly important as primary care practice is a key element of the health care system. It provides the first contact between patients and health care professionals, it provides continuity of care over time, it has a focus on the whole patient rather than one particular organ or illness, and provides coordination with other parts of the system (Starfield, 1998). Moreover, existing literature shows that access and use of primary care services and practitioners is associated with better health outcomes, more efficient use of care and lower spending levels (Baicker and Chandra, 2004; Starfield et al., 2005; American College of Physicians, 2008; Chang et al., 2011).

The PSA bonus is not the first attempt at using financial incentives to influence providers' behavior. Since 1991 Medicare pays the Health Professional Shortage Area (HPSA) bonus, a 10% payment for primary and mental health care providers. Factors affecting the designation of a HPSA, which can change from one year to another, include the ratio of population to full-time equivalent (FTE) physicians (at least 3,500 to 1 FTE for primary care, and 30,000 to 1 FTE for mental health care), whether or not the area is "rational" for the delivery of medical care services (relatively self-contained geographic unit with respect to the provision of primary care services), and whether medical care professionals are overutilized or inaccessible to the area under consideration. Another example is the Primary Care Incentive Payment Program (PCIP), a bonus equal to 10% of the amount paid for primary care services to primary care practitioners if those ser-

vices accounted for at least 60% of the practitioner's total allowed charges in a given year.[2] This program ran between 2011 and 2015 as part of the Affordable Care Act. In 2015, payments for these two programs represented around $821 millions or 0.1% of total Medicare spending, covering approximately 140,000 providers (MedPac 2017).

This paper focuses on primary care PSAs. CMS together with the Health Resources and Services Administration (HRSA) defined PSAs as counties with the lowest ratio of PCPs to Medicare beneficiaries up to covering 20% of the national Medicare population. This bonus covered 7,539,659 Medicare beneficiaries and 24,719 primary care physicians (Graham Center One-Pager # 55, 2008). PCPs rendering services in PSAs received a 5% bonus over the amount paid by Medicare for professional services and it was paid on a quarterly basis between January 1st, 2005 and June 30th, 2008.[3] Primary care specialties eligible for the bonus were General practice, Family medicine, Internal medicine, and Obstetrics/Gynecology.

For my analysis I focus on areas where physicians got an automatic payment of the bonus. At the end of 2004 CMS published a list of all PSA designated ZIP codes and counties. This list remained constant for the whole period the bonus was in place. Payments were automatically made for services provided in ZIP codes areas that (a) fully fell within a PSA designated county, (b) partially fell within a PSA county and were considered to be dominant for that county according to the USPS determination, or (c) fell within rural census tracts of a metropolitan statistical area identified through the latest modification of the Goldsmith Modification that was determined to be a PSA. If the ZIP code where the physician rendered services was not included on the PSA automated pay ZIP code file, but was in the county list, the doctor had to use a modifier to claim the

[2]This bonus was paid to physicians (family medicine, geriatric medicine, pediatric medicine, and internal medicine) and other practitioners (nurse practitioners, clinical nurse specialists, and physician assistants).

[3]If the physician billed for a service that had both a professional and technical component, only the professional component received the bonus payment. The technical component represents the cost of the equipment, supplied and personnel to perform a procedure, while the professional component represents the supervision and interpretation of the service provided by the physician.

PSA bonus payment.

To perform my analyses I use as a control group the set of counties that did not received the PC PSA bonus. Depending on the specification, this group will include counties receiving only the Specialty PSA bonus (and not the PC PSA bonus) plus counties without any scarcity designation, or only the former group of counties. Figure 1.1 shows the geographical distribution of these areas across the country. There are 486 Specialty-only PSA counties, 103 PC-only PSA counties, 1338 Specialty and Primary Care PSA counties, and 1214 counties without scarcity designation.

## 1.3 Data

I use data from claims submitted by providers to Medicare for reimbursement, from a 20 percent random sample of the Medicare Part B beneficiary population for each year from 2000 through 2010. Additionally, I obtain demographic information about the beneficiary sample from the Beneficiary files. Panel A of Table 1.1 shows some basic summary statistics of these characteristics. Overall, the different groups of counties used in the analysis look similar in terms of these measures.

To study the effects of the bonus on the number of physicians serving particular counties, I exploit the fact that health care providers are identified using either the Unique Physician Identification Number (UPIN) or the National Provider Identifier (NPI) depending on the year of the sample. The number of procedures per patient are identified using the Healthcare Common Procedure Coding System (HCPCS). Patients are identified using their beneficiary identification code, while visits are constructed by counting the times a patient appears in the sample in different dates.

To compute the aggregate quantity of health care supplied to this sample of beneficiaries, I use Relative Value Units (RVUs). RVUs are the scaling of individual services used by CMS to reimburse providers, varying across different services. They are supposed to

measure all the resources needed to provide a particular service.[4] These resources include physician work (time, skills, mental effort), practice expense (office space, supplies, staff), and professional liability insurance (malpractice insurance premiums). On average, the first two components account for 52 and 44 percent of total Medicare expenditures on physician services, respectively. For example, a diagnostic colonoscopy has more work RVUs than an intermediate office visit because the former service is estimated to require more time (75 vs 40 minutes, considering pre and post periods), skills, and effort.

The data also include a standard classification system known as Betos categories that allows me to map specific medical services and procedures to broad categories. They include Evaluation and Management (e.g., office visits), Testing, Imaging, and Procedures.

Panel B of Table 1.1 shows basic summary statistics of these outcomes at the county-year level. Counties differ in terms of number of patients and therefore visits, particularly counties with no scarcity designation, however the rest of measures of care look similar between the groups.

Finally, since extra payments were associated with the location where procedures were provided, I assigned services to counties using providers' ZIP codes.

### 1.4 Empirical strategy

To obtain an estimate of the causal effect of financial incentives on health care supply in underserved areas, I use different versions of a differences-in-difference strategy. In particular, I compare the quantity of health care supplied measured using different outcomes in counties that were and were not subject to the PSA bonus during the period it was in place, to a weighted average of the periods before and after its introduction and elimina-

[4] Following Clemens and Gottlieb (2014), the provider's fee for service $j$, supplied in payment area $a$ is approximately

$$\text{Reimbursement}_{a,j,t} = \text{Conversion Factor}_t \times \text{RVUs}_j \times \text{Geographic Adjustment Factor}_a$$

The Conversion Factor is a national adjustment factor, that is updated annually and usually constant across different services, while the Geographic Adjustment Factor corresponds to the federal government's adjustment for differences in input costs across payment regions.

tion. I also present results from an specification that allows me to separately identify the effects of the bonus start and end, and full parametric differences-in-difference estimations to test for absence of pre trends between control and treatment groups.

My main results come from the following estimating equation:

$$Y_{s(i)t} = \alpha + \beta \cdot PSA_i \cdot I\left[2005 \leq t \leq 2007\right] + \delta_t + \gamma_i + \eta_{st} + \rho X_{it} + \varepsilon_{it} \qquad (1.1)$$

where $Y_{s(i)t}$ is the outcome of interest in county $i$ located in state $s$ during year $t$. I look at two sets of outcomes. The first one focuses on the effects of the bonus on the number of providers per patient, number of visits, and my aggregate measure of healthcare supply in county $i$ during year $t$. I believe the first two outcomes are more closely related to the primary goal Congress had in mind when it established the bonus (see Section 1.2), while the last one allows me to compare results with previous literature. Using the second set of outcomes, I disaggregate the measure of healthcare supply into different components, and I explore potential intensive and extensive margin responses. All outcomes are measured in logs, allowing me to interpret the coefficients as elasticities and to compare them to results from the literature.

The dummy variable $PSA_i$ is equal to 1 if county $i$ was designated as Primary Care Physician Scarcity Area, and 0 otherwise. Depending on the specification, non-PC PSA areas will include both non-designated areas and Specialty PSAs, or only the later set of counties. I will present results using both definitions of control group. I also include county fixed effects, $\gamma_i$, year fixed effects, $\delta_t$, and state-by-year fixed effects, $\eta_{st}$. As the bonus ran from January 2005 until the first half of 2008, I consider 2005-2007 the relevant period for the bonus.

I control for county characteristics $X_{it}$ that may be important determinants of the level of care received by Medicare beneficiaries in each county. These variables include indicators for different age ranges, gender, race, ethnicity, and indicators for chronic conditions.

The coefficient of interest is $\beta$, which captures the effect of the bonus in PSA counties between 2005 and 2007, to a weighted average of 2000-2004 and 2008-2010. Throughout the analysis, I cluster the standard errors at the county level, and I weight the county-level observations using population aged 65 and older from the 2000 Census.

In order to capture differential effects of the start and end of the bonus, I estimate the following equation:

$$\begin{aligned} Y_{s(i)t} &= \alpha + \beta_1 \cdot PSA_i \cdot I\left[t \geq 2005\right] + \beta_2 \cdot PSA_i \cdot I\left[t \geq 2008\right] \\ &+ \delta_t + \gamma_i + \eta_{st} + \rho X_{it} + \varepsilon_{it} \end{aligned} \tag{1.2}$$

where $I\left[t \geq 2005\right]$ is an indicator variable equal to 1 for years 2005 and after, and $I\left[t \geq 2008\right]$ is equal to 1 for years 2008 and after. I then interact these indicators with the dummy $PSA_i$.

The interaction with $I\left[t \geq 2005\right]$ measures the effects on the different outcomes in PSA counties, relative to non-PSA areas, during the bonus was in place, relative the period just before that. On the other hand, the interaction with $I\left[t \geq 2008\right]$ will capture any effect on the outcomes on PSA counties, relative to non-PSA ones, after the bonus was eliminated, relative to the period when it was on place. I also present results from the test of the hypothesis that the parameters associated with the two interactions are the same but of opposite sign.

To interpret the effect causally, the key assumption in the context of difference-in-differences-style estimations is that the outcome trends would be the same in both groups of counties in the absence of the bonus. To test for the presence of parallel trends, that I present graphically, I run the following regression

$$Y_{s(i)t} = \alpha + \beta_t \cdot PSA_i \cdot \delta_t + \delta_t + \gamma_i + \eta_{st} + \rho X_{it} + \varepsilon_{it} \tag{1.3}$$

I omit $t = 2004$ so that each $\beta_t$ is estimated relative to the year immediately preceding

the introduction of the bonus.

## 1.5 Results

I first present event study graphs that motivate the regression analyses that follow. My estimates of the effect of the PC PSA bonus on aggregate quantities of healthcare, number of doctors, and number of visits, outcomes more closely related to the goal of the bonus, are shown in Figure 1.2. The different panels show the time path of the respective outcome in PSA counties, relative to non-PSA counties, conditional on the regressors from equation (1.3). The panels provide support for the validity of the design as there is little evidence of differential trends in the outcomes between control and treatment groups before 2005. There is a small and weak increase in the aggregate quantity of healthcare supply after the introduction of the bonus, while doctors and visits were unaffected by the increase in reimbursement.

Tables 1.2 and 1.3 show the baseline estimates for the effect of the PSA bonus on the first set of outcomes. The main difference between these two tables is the composition of the control group. Table 1.2 considers only counties that received the Specialty PSA bonus (and not the primary care one) as control group, while Table 1.3 adds counties that were not designated any type of PSA. I consider the first control group a better control, as all counties are subject to some degree of scarcity of healthcare practitioners, regardless of their specialty. However, as Figure 1.1 shows, scarcity areas were spread roughly evenly across the country. Therefore, in a second specification I include non-PSA counties to asses the robustness of my results.

Panel A of Table 1.2 reports coefficients and standard errors associated with the interaction $PSA_i \cdot I\,[2005 \leq t \leq 2007]$ from equation (1.1). Panel B shows results derived from equation (1.2), reporting coefficients and standard errors associated with the interactions with start and end date of bonus, while the last row shows the results from the test that the start and end date coefficients are equal and opposite in sign. Small counties appear

with gaps in my Medicare sample, which means that I only observe them for a subset of years. To deal with this issue, the first three columns of Table 1.2 present results considering only counties that I observe for the whole 2000-2010 period, larger in terms of population, while the last three columns show results including counties with gaps, in which case I impute zero to the outcome.

The results from Table 1.2 show that, in either specification, the bonus slightly increased the aggregate amount of health care supply, as measured by the log of total RVUs, between 2005 and 2007 in counties subject to the bonus relative to those not exposed to it. The point estimates imply a supply elasticity of around 0.013, which means that the increase of 5% on reimbursement rates generated a 0.065% increase on health care supply during this period. In a related work, Clemens and Gottlieb (2014) study the relation between reimbursement rates and health care supply, using a geographic reimbursement shock that affected Medicare physicians in 1997. They estimate an elasticity of health care supply of 1.5 with respect to reimbursement rates, which in the context I study would imply a 7.5% increase on care provision. My confidence intervals expressed as percentage go from around -0.0003% to 0.145%, rejecting the result from the literature. My results also suggest that the number of physicians per patient decreased by around 0.01%, while the number of doctor visits increased by roughly 0.1%, however these estimates are not statistically significant.

Results from Panel B align with those from the previous panel. They indicate that most of the effect is concentrated around the introduction of the bonus rather than during the phasing out. The final row shows the results from a test that the start and end coefficients are equal and of opposite sign. The hypothesis cannot be rejected in most cases with the exception of the case for aggregate care supply when focusing on counties observed in the data for the whole 2000-2010 period.

Table 1.3 shows the results using every county not subject to the PC PSA bonus as comparison group. Overall, these results are in line with the ones presented in the previous

table, as point estimates indicate either non statistically significant or weak results. The supply elasticity of care ranges from -0.05% to 0.05%, rejecting again the results from Clemens and Gottlieb (2014). The number of physicians per patient decreased by around 0.005%, and results for the number of doctor visits indicate a reduction of 0.15% but these estimates are not statistically significant.

Regarding the comparison with results from the literature, it is important to keep in mind that the authors show that most of the effect they find comes elective procedures being particularly relevant medical services that involve high levels of technology and capital investment like MRIs. The nature of services provided by PCPs is much more labor intensive than other specialties (around 90% of services provided are office visits, Table 1.1), therefore a small effect on health care provision is not surprising in this context, specially given the small size of the bonus.

To analyze differential response to incentives by type of service, I divide my measure of care supply into Evaluation and Management, Tests, and Imaging services, using Betos categories. Figure 1.3 shows the results from equation (1.3) using these outcomes, keeping the aggregate measure in Panel A. In general, different services follow similar patterns to those shown on Panel A. The supply of care measure through different types of services slightly increases after the introduction of the bonus, however the magnitudes are not economically relevant. It is worth noting that the response of tests and imagining, as opposed to E&M, shows statistical significance, in line with the importance of elective procedures found by Clemens and Gottlieb (2014).

I further explore the composition of supply responses along different dimensions reporting results in Figure 1.4. Panels B and C divide the log RVUs per patient, presented on Panel A, into log RVUs per procedures and log procedures per patient. The first measure captures the intensity of the average procedure patients receive, while the second one captures the extensive margin of the response. The results suggest that the total effect comes from an increase in the number of procedures provided per patient rather than

from an increase in intensity.

The remaining panels of Figure 1.4 focus on margins such as the number of patients, and RVUs per physician. Both outcomes are unaffected by the bonus, suggesting that financial incentives of this size are not a key determinant of whether or not a Medicare beneficiary receives care in a given year, and of the intensive margin of RVUs per physician.

## 1.6 Mechanisms behind effect

The previous section showed that the PSA bonus had a small or null effect on different health related outcomes such as availability of providers, number of patients getting access to care, and several measures of health care per patient. Behind these results, two main explanations arise. The first one is related to the level of information providers have regarding these incentives, while the second is about the relevance of the payment.

According to the Kaiser Family Foundation/Commonwealth Fund 2015 National Survey of Primary Care Providers, which interviewed a nationally representative sample of 1,624 primary care physicians and 525 midlevel clinicians (nurse practitioners and physician assistants) working in primary care practices between January and March on 2015, 49% of physicians accepting Medicare patients were not aware of the Primary Care Incentive Payment Program that implied a 10% increase in Medicare payments to PCPs. Moreover, 47% of physicians accepting Medicaid patients were not aware of a temporary increase in Medicaid payments for primary care to match Medicare rates which began in 2013. This survey evidence indicates that physicians have limited information about financial incentives that directly affect them and that CMS efforts to spread this information may not be working as expected.

The survey also provides evidence about the perceived relevance of these incentives for the sample of physicians that were aware and actually received them. Only 5% of PCPs accepting Medicare patients said the 10% bonus made a big difference in their ability to serve their patients, while 48% said the bonus made no difference at all. Regarding the

Medicaid payment increase, 41% said it made no difference and 22% said it made a big difference in their ability to treat Medicaid beneficiaries. As this bonus was equal to 10% while the PSA was only 5%, it is be expected that smaller financial incentives would be equally or even more irrelevant in the decision making process of physicians when treating their patients.

Another piece of evidence comes from early studies focusing on the HPSA bonus. Shugarman et al. (2001, 2003) show a declining trend during the 1990s in the use of Medicare bonus payments by physicians serving rural patients, suggesting that they failed to claim these extra payments available to them. The authors argue that some of the underlying factors explaining this trend are the level of information physicians have about these payments, their perceived value, how easy it is to get them, and concerns about the risk of audits.

Results from economic literature indicate a strong relation between financial incentives and different measure of health care supply. Evidence from descriptive literature and from surveys, plus the findings presented in this paper, points out to the importance of awareness and bonus size when studying the effects of financial incentives in underserved regions.

## 1.7 Conclusions

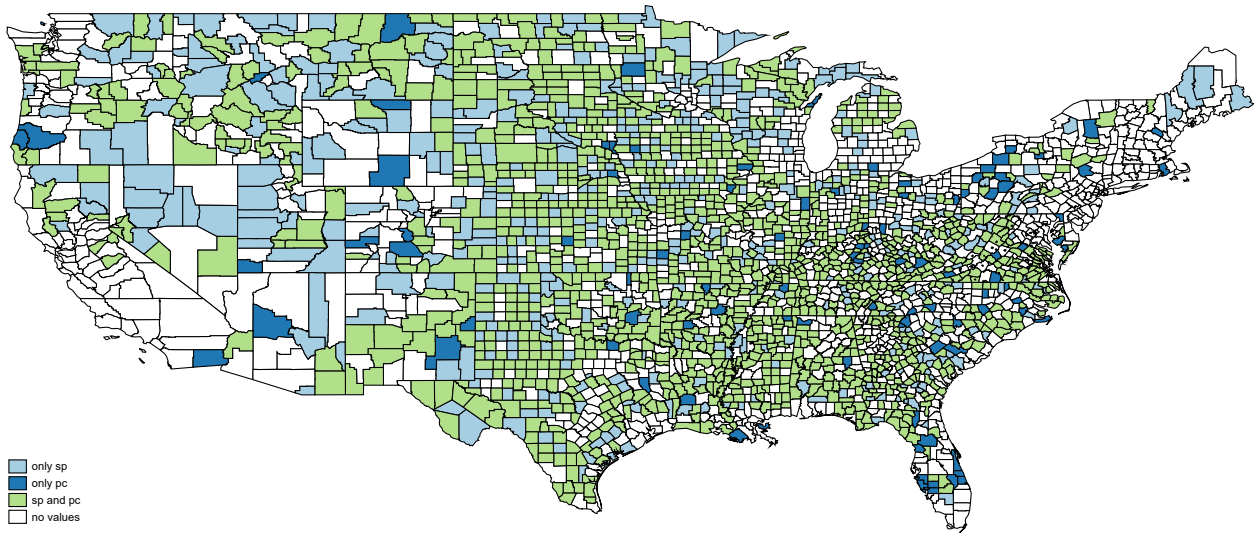
This paper studies how financial incentives designed to address the shortage of primary care physicians affect their labor and health care supply in underserved areas. I focus on the Physician Scarcity Area bonus, a 5% extra payment provided between 2005 and 2008 by Medicare to PCPs rendering services in areas with such designation. Despite the large body of literature showing a positive relation between financial incentives and several measures of care supply in different contexts, I find that the bonus had no impact on the availability of physicians, the number of patients seen in these areas, and the number of visits to the doctor. I also find a positive but weak and small effect on the overall provision

of health care, with estimates that allow me to reject results from previous literature.

The key policy issue motivating this paper is how to deal with the shortage of PCPs across the U.S. The paper shows that providing financial incentives to health care professionals, a policy widely used by Medicare and Medicaid during recent years, may not be the most efficient way to address this problem, at least not with its current design. Physicians surveys and previous literature suggest that the level of awareness regarding these bonus payments is low, and that their relative value is not enough to actually affect providers, in line with the magnitudes of the effects presented in this paper. These design characteristics are therefore crucial and should be addressed in future policy proposals involving the provision of monetary incentives.

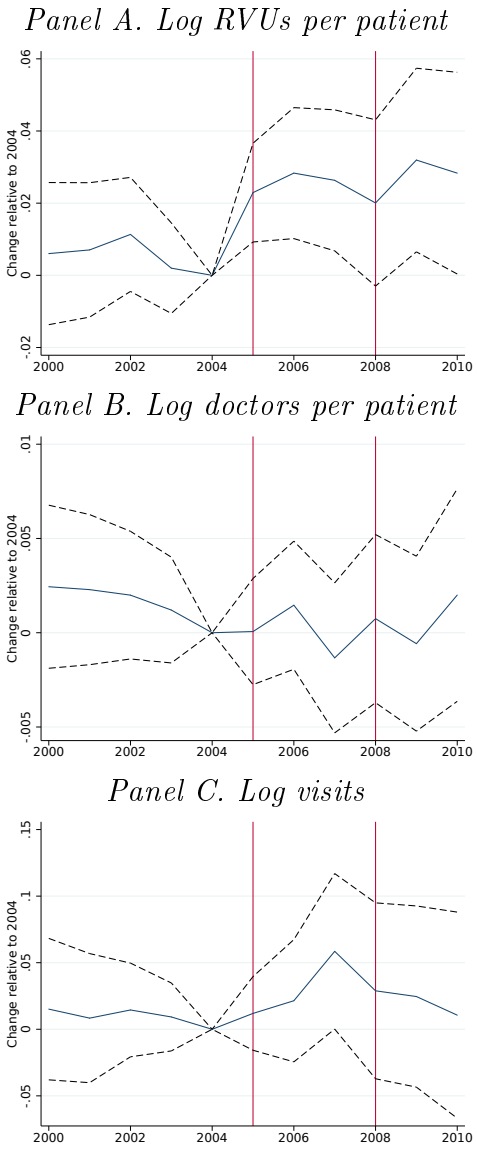
## 1.8 Figures

Figure 1.1: Primary care and Specialty Physician Scarcity Areas, 2005-2008



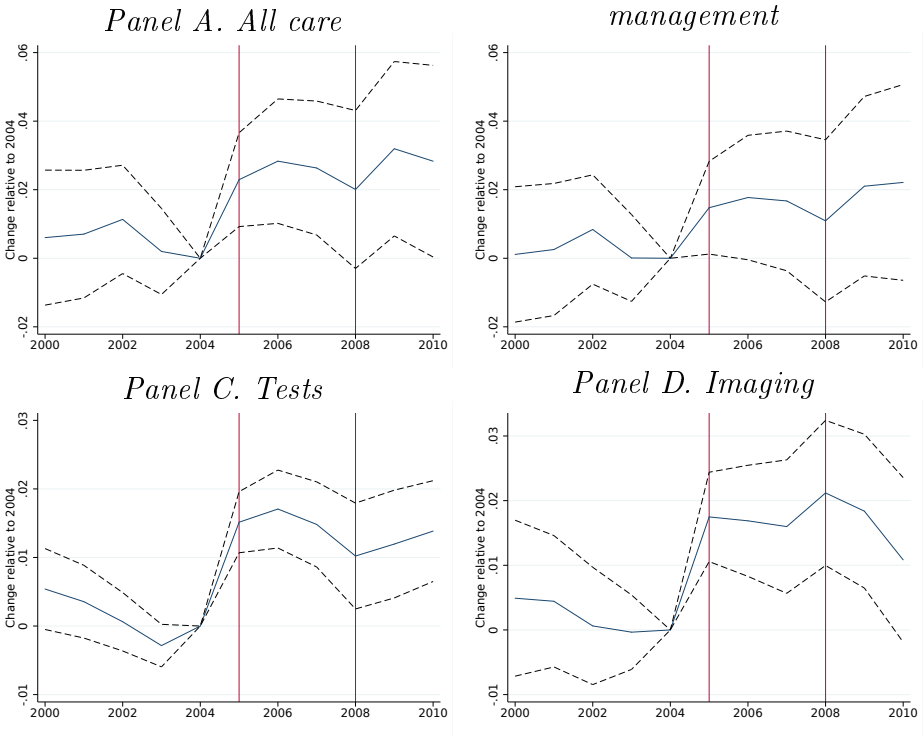

*Notes:* This figure presents the geographic distribution of Primary care and Specialty Physician Scarcity Areas. Counties in light blue received the Specialty PSA bonus, counties in dark blue received the Primary Care PSA bonus, counties in green recevied both the Specialty and Primary Care PSA bonus, and counties in white did not receive any designation.

Figure 1.2: Impact of PSA bonus on aggregate care supplied, doctors, and visits



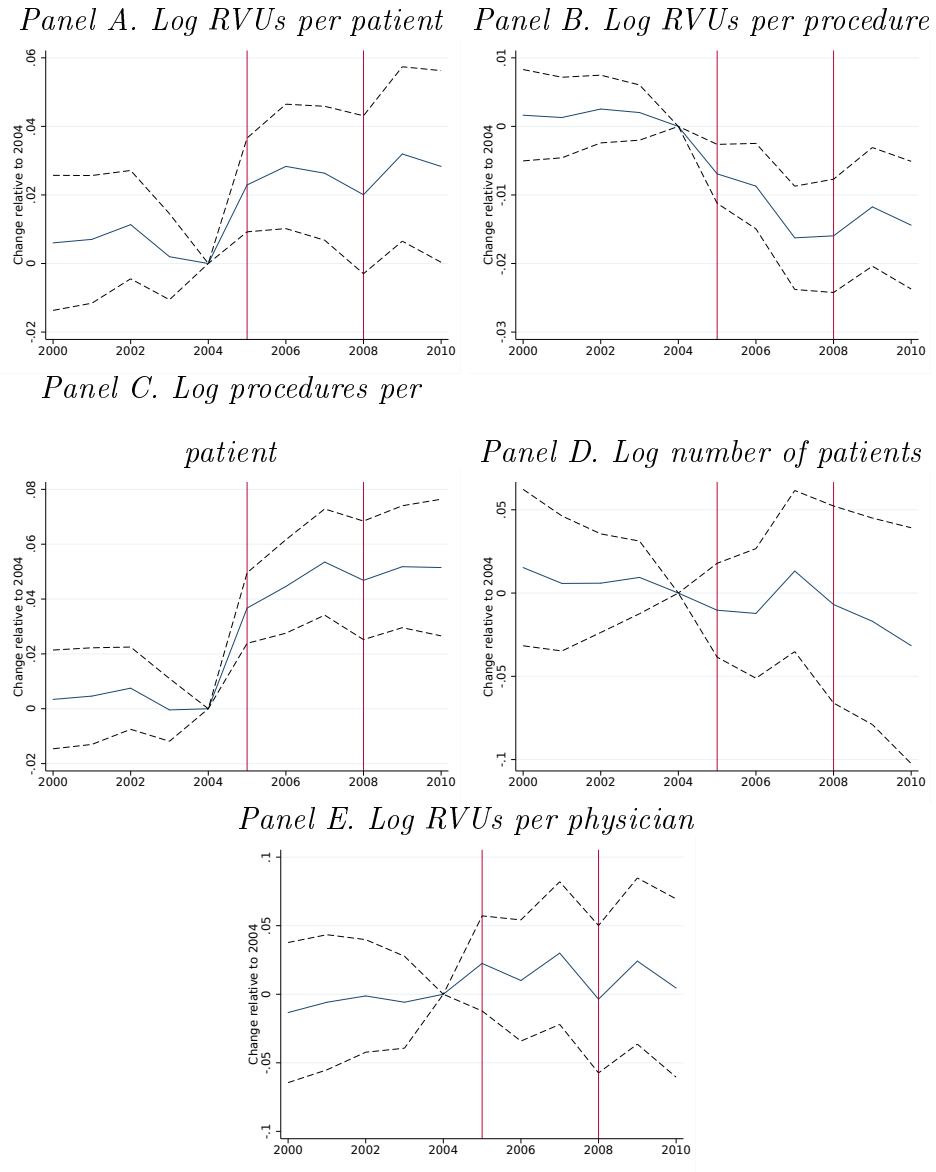

*Notes:* These graphs show coefficients and standard errors from OLS regressions at the county level in which log of RVUs per patient, log of doctors per patient, and log of visits to the doctor are the dependent variables. These outcomes are regressed on a dummy variable equal to 1 if the county received the PSA bonus and 0 otherwise, interacted with indicator variables for each year. These regressions control for county fixed effects, year fixed effects, and state-by-year fixed effects. They also control for the fraction of beneficiaries aged 65-69, 70-74, 75-79, and 80-84, black, Hispanic, female eligible for Medicare due to end-stage renal disease or disability, with 2 or more, 3 or more, 4 or more, and 6 or more chronic conditions. Regressions are weighted using the population aged 65 and older from the 2000 Census. Standard errors are clustered at the county level.

Figure 1.3: Supply response by service category




*Notes:* These graphs show coefficients and standard errors from OLS regressions at the county level in which the quantities of health care supplied in different service categories are the dependent variables, measured in Relative Value Units. These outcomes are regressed on a dummy variable equal to 1 if the county received the PSA bonus and 0 otherwise, interacted with indicator variables for each year. These regressions control for county fixed effects, year fixed effects, and state-by-year fixed effects. They also control for the fraction of beneficiaries aged 65-69, 70-74, 75-79, and 80-84, black, Hispanic, female eligible for Medicare due to end-stage renal disease or disability, with 2 or more, 3 or more, 4 or more, and 6 or more chronic conditions. Regressions are weighted using the population aged 65 and older from the 2000 Census. Standard errors are clustered at the county level.

Figure 1.4: Potential margins of response




*Notes:* These graphs show coefficients and standard errors from OLS regressions at the county level in which different measures of healthcare supply are the dependent variables. These outcomes are regressed on a dummy variable equal to 1 if the county received the PSA bonus and 0 otherwise, interacted with indicator variables for each year. These regressions control for county fixed effects, year fixed effects, and state-by-year fixed effects. They also control for the fraction of beneficiaries aged 65-69, 70-74, 75-79, and 80-84, black, Hispanic, female eligible for Medicare due to end-stage renal disease or disability, with 2 or more, 3 or more, 4 or more, and 6 or more chronic conditions. Regressions are weighted using the population aged 65 and older from the 2000 Census. Standard errors are clustered at the county level.

## 1.9 Tables

Table 1.1: Summary statistics

| | PC PSA counties | | SP PSA counties | | Non-PSA counties | |
|---|---|---|---|---|---|---|
| | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* |
| **Panel A: Demographics, county-by-year** | | | | | | |
| Fraction aged 65-69 | 0.28 | 0.04 | 0.27 | 0.04 | 0.28 | 0.03 |
| Fraction aged 70-74 | 0.23 | 0.03 | 0.23 | 0.02 | 0.23 | 0.02 |
| Fraction aged 75-79 | 0.19 | 0.02 | 0.19 | 0.02 | 0.19 | 0.01 |
| Fraction aged 80-84 | 0.14 | 0.02 | 0.14 | 0.02 | 0.14 | 0.01 |
| Fraction female | 0.56 | 0.03 | 0.56 | 0.03 | 0.57 | 0.02 |
| Fraction black | 0.06 | 0.12 | 0.04 | 0.09 | 0.06 | 0.09 |
| Fraction Hispanic | 0.01 | 0.03 | 0.01 | 0.03 | 0.01 | 0.02 |
| Fraction eligible due to ESRD or DI | 0.09 | 0.04 | 0.08 | 0.03 | 0.08 | 0.02 |
| Fraction with 2 or more chronic conditions | 0.68 | 0.06 | 0.66 | 0.07 | 0.68 | 0.06 |
| Fraction with 3 or more cc | 0.53 | 0.07 | 0.51 | 0.08 | 0.53 | 0.07 |
| Fraction with 4 or more cc | 0.39 | 0.07 | 0.37 | 0.08 | 0.39 | 0.07 |
| Fraction with 6 or more cc | 0.17 | 0.05 | 0.16 | 0.05 | 0.17 | 0.04 |
| **Panel B: Outcomes, county-by-year** | | | | | | |
| Total RVUs per patient | 5.19 | 2.58 | 5.99 | 2.12 | 6.43 | 1.74 |
| E&M RVUs per patient | 4.70 | 2.38 | 5.38 | 2.00 | 5.78 | 1.61 |
| Test RVUs per patient | 0.05 | 0.08 | 0.06 | 0.07 | 0.07 | 0.06 |
| Imaging RVUs per patient | 0.09 | 0.19 | 0.09 | 0.14 | 0.13 | 0.13 |
| Physicians per patient | 0.07 | 0.12 | 0.08 | 0.12 | 0.05 | 0.05 |
| Visits | 2271.34 | 5273.28 | 2219.82 | 2444.12 | 20461.77 | 39387.04 |
| RVUs per procedure | 0.94 | 0.36 | 1.04 | 0.29 | 1.06 | 0.19 |
| Procedures per patient | 5.18 | 2.33 | 5.77 | 1.87 | 6.05 | 1.30 |
| Patients | 413.84 | 932.90 | 407.04 | 456.46 | 3484.43 | 5672.46 |
| RVUs per physician | 118.59 | 102.97 | 114.40 | 87.28 | 132.09 | 66.39 |
| Observations | 15059 | | 5302 | | 13035 | |

*Notes:* This table reports summary statistics on demographics (Panel A) and outcome variables (Panel B) at the county-year level for Primary Care PSA counties, Specialty PSA counties, and counties with no scarcity designation.

Table 1.2: Effect of PSA bonus, selected outcomes - other scarcity counties as control group

| | Counties observed whole period | | | All counties | | |
|---|---|---|---|---|---|---|
| | RVUs | Doctors | Visits | RVUs | Doctors | Visits |
| *A. Estimated effect of bonus* | | | | | | |
| $PSA \cdot I[2005 \leq t \leq 2007]$ | 0.0125* | -0.0012 | 0.0167 | 0.0148** | -0.0025* | 0.0258 |
| | (0.0064) | (0.0012) | (0.0167) | (0.0072) | (0.0014) | (0.0202) |
| *B. Estimated effect of bonus start and end* | | | | | | |
| $PSA \cdot I[t \geq 2005]$ | 0.0206** | -0.0015 | 0.0212 | 0.0147* | -0.0032* | 0.0087 |
| | (0.0081) | (0.0016) | (0.0240) | (0.0088) | (0.0018) | (0.0270) |
| $PSA \cdot I[t \geq 2008]$ | 0.0009 | 0.0006 | -0.0092 | -0.0151 | 0.0013 | -0.0541* |
| | (0.0090) | (0.0017) | (0.0240) | (0.0112) | (0.0019) | (0.0301) |
| $H_0$*:Bonus start=-Bonus end (p-value)* | | | | | | |
| | 0.061 | 0.715 | 0.738 | 0.974 | 0.469 | 0.264 |
| *Observations* | 18073 | 18073 | 18073 | 20361 | 20361 | 20361 |

*Notes:* This table shows coefficients and standard errors from OLS regressions at the county level in which log of RVUs per patient, log of doctors per patient, and log of visits to the doctor are the dependent variables. The control group are counties that received only the Specialty PSA bonus. The first three columns focus on counties that are observed without gaps during the 2000-2010 period, while the last three columns include counties with gaps. In Panel A, the outcomes are regressed on a dummy variable equal to 1 if the county received the PSA bonus and 0 otherwise, interacted with indicator variables for years between 2005 and 2007. In Panel B, the outcomes are regressed on the PSA dummy variable interacted with indicators of the start and end year of the bonus. This panel also includes $p$-values from tests that the two coefficients are equal, but of opposite sign. These regressions control for county fixed effects, year fixed effects, and state-by-year fixed effects. They also control for the fraction of beneficiaries aged 65-69, 70-74, 75-79, and 80-84, black, Hispanic, female eligible for Medicare due to end-stage renal disease or disability, with 2 or more, 3 or more, 4 or more, and 6 or more chronic conditions. Regressions are weighted using the population aged 65 and older from the 2000 Census. Standard errors are clustered at the county level.
* $p < 0.10$, * $p < 0.05$, *** $p < 0.01$.

Table 1.3: Effect of PSA bonus, selected outcomes - all counties as control group

| | Counties observed whole period | | | All counties | | |
|---|---|---|---|---|---|---|
| | RVUs | Doctors | Visits | RVUs | Doctors | Visits |
| *A. Estimated effect of bonus* | | | | | | |
| $PSA \cdot I[2005 \leq t \leq 2007]$ | -0.0006 | -0.0006 | -0.0346*** | 0.0003 | -0.0013** | -0.0321*** |
| | (0.0048) | (0.0005) | (0.0096) | (0.0050) | (0.0006) | (0.0112) |
| *B. Estimated effect of bonus start and end* | | | | | | |
| $PSA \cdot I[t \geq 2005]$ | -0.0108 | 0.0008 | -0.1140*** | 0.0169** | -0.0001 | -0.1310*** |
| | (0.0066) | (0.0007) | (0.0146) | (0.0068) | (0.0008) | (0.0165) |
| $PSA \cdot I[t \geq 2008]$ | -0.0158** | 0.0030*** | -0.0936*** | -0.0280*** | 0.0033*** | -0.1270*** |
| | (0.0067) | (0.0007) | (0.0141) | (0.0072) | (0.0008) | (0.0170) |
| $H_0$*:Bonus start=-Bonus end (p-value)* | | | | | | |
| | 0.004 | 0.000 | 0.000 | 0.000 | 0.004 | 0.000 |
| *Observations* | 31009 | 31009 | 31009 | 33396 | 33396 | 33396 |

*Notes:* This table shows coefficients and standard errors from OLS regressions at the county level in which log of RVUs per patient, log of doctors per patient, and log of visits to the doctor are the dependent variables. The control group are all counties that did not received the PC PSA bonus. The first three columns focus on counties that are observed without gaps during the 2000-2010 period, while the last three columns include counties with gaps. In Panel A, the outcomes are regressed on a dummy variable equal to 1 if the county received the PSA bonus and 0 otherwise, interacted with indicator variables for years between 2005 and 2007. In Panel B, the outcomes are regressed on the PSA dummy variable interacted with indicators of the start and end year of the bonus. This panel also includes $p$-values from tests that the two coefficients are equal, but of opposite sign. These regressions control for county fixed effects, year fixed effects, and state-by-year fixed effects. They also control for the fraction of beneficiaries aged 65-69, 70-74, 75-79, and 80-84, black, Hispanic, female eligible for Medicare due to end-stage renal disease or disability, with 2 or more, 3 or more, 4 or more, and 6 or more chronic conditions. Regressions are weighted using the population aged 65 and older from the 2000 Census. Standard errors are clustered at the county level.
* $p < 0.10$, * $p < 0.05$, *** $p < 0.01$.

# CHAPTER 2
# HEALTH EFFECTS OF THE FLINT WATER CRISIS ON THE ELDERLY POPULATION

## 2.1 Introduction

The dangers of lead exposure and poisoning are well known and understood within the general and specialized public. However, most of the evidence comes from studies focusing on infant health as brains during childhood are more vulnerable than during adulthood.

Correlational and causal evidence draw from epidemiological and economic studies has found a strong relation between levels of lead in blood or measures of lead exposure during childhood and outcomes such as attention-deficit/hyperactivity disorder (ADHD) and intellectual impairment (Nigg et al., 2010; Chandramouli et al., 2009; Lanphear et al., 2005; Canfield et al., 2003; Wasserman et al., 1997); cognitive, math, and language test scores (Aizer et al., 2018; Rau et al., 2015; Ferrie et al., 2012); and behavioral and criminal problems (Aizer & Currie, 2019; Reyes, 2015; Reyes, 2007).

In adults, epidemiological studies have found a positive relationship between blood or bone lead levels, even at low magnitudes, and health problems such as high blood pressure and cardiovascular disease, kidney disease, and cognitive deterioration (Vaziri, 2008; Jain et al., 2007; Weisskopf et al., 2007; Ekong et al., 2006; Menke et al., 2006; Navas-Acien et al., 2006; Schober et al., 2006; Weisskopf et al., 2004; Muntner et al., 2003; Lustberg & Silbergeld, 2002). In a recent paper, Lanphear et al. (2018) use a nationally representative cohort for the US to quantify the relative contribution of environmental lead exposure to all-cause and cardiovascular disease mortality. They find that around 18% of all deaths every year are attributable to lead exposure, and 28% of all cardiovascular disease deaths. Elderly adults have been alive for a longer period of time, which means their chances of lead exposure while working in unregulated occupations or of having encountered high levels in the environment on a daily basis are higher. This is in line with findings from epidemiological studies that have found higher blood and bone lead levels in older adults compare to younger ones (Vig & Hu, 2000).

In this paper I use the Flint water crisis in Michigan to study the causal effect of lead exposure on the health of elderly population. In April 2014, the city of Flint changed its water source from Lake Huron to Flint River, exposing its residents to high levels of bacterial and chemical contaminants, including lead. For approximately a year and a half, local authorities reassured residents that the water was safe for human consumption. Finally, in December 2015 the Mayor of Flint declared state of emergency. Adult

Flint residents have reported to the press of new high blood pressure diagnoses, among other health problems (Maher, 2016; Glenza, 2018).

The characteristics of the water crisis also allow me to study other health effects. Around the same time of the water switch, Flint also experienced an outbreak of Legionnaires' disease that previous scientific evidence has linked to the change in water source (Zahran et al., 2018). Authorities have set the official death toll for the outbreak at 12 people. This disease manifests as a type of pneumonia, and the final diagnostic is only possible if specific tests are run. Therefore, I study how hospitalizations and deaths for pneumonia changed after the crisis, to study if some Legionnaires' disease cases were diagnosed as pneumonia.

I exploit the within-state variation in exposure to contaminated water. I use a difference-in-differences model and synthetic control methods to compare outcomes in Flint after the start of the crisis to outcomes in Flint before the start and to other cities in Michigan. Contaminated water only reached households within the city limits of Flint, leaving the rest of the cities in Michigan and their residents unaffected. Additionally, the scarce information provided by authorities and the relative short period of contamination, 18 months, provides a context where I do not expect large avoidance responses. In order to get bottled water, older Flint residents have manifested the existence of barriers, including lack of transportation and lack of ability to carry water packages (AARP, 2016).

Using the universe of Medicare beneficiaries living in Michigan between 2012 and 2016, and the CDC's National Detailed Mortality Files between 2010 and 2017, I estimate the effect of the change in the water supply on hospitalization and mortality rates. Results from difference-in-differences and synthetic controls specifications suggest a weak increase in all-cause and kidney disease hospitalization rates in Flint after the start of the crisis, and an increase in the all-cause mortality rates. However, these results are usually smaller than what has been suggested by the epidemiological literature and the graphical evidence does not strongly support them. Results for hospitalizations or deaths for pneumonia are not significant, suggesting no misdiagnoses of Legionnaries' disease.

An important caveat of my analysis is the lack of objective measures of lead exposure in the data, such as levels of lead in blood or bones, and the short period of time after the crisis that I have available. This post period of only 2 or 3 years could be too short to see effects, specially for an outcome as extreme as mortality. Future work should address these questions again when new data become available.

My paper contributes to the previously mentioned group of literature that has established a correlation between lead exposure and negative health effects in adult populations. Additionally, it more directly contributes to a growing body of literature studying health and economic effects of the Flint crisis. Researchers have documented increased blood lead levels in children (Hanna-Attisha et al., 2016; Zahran

et al., 2017), negative impacts on maternal and birth outcomes (Abouk & Adams, 2018, Grossman & Slusky, 2019, Jenkins & Danagoulian, 2019), and an increase in bottled water sales and decline in home values (Christensen et al., 2019).

### 2.2 Flint Water Crisis

In 2011, the Governor of Michigan appointed Darnell Earley as the first of a series of emergency managers for the city of Flint. This appointment was made after a Michigan Treasury Department review found that the city had a severe structural deficit in its finances.

Emergency managers have the power to make unilateral changes in order to reduce financial distress in cities. One of the cost cut measures contemplated by the authority was to switch Flint's water source from the Detroit Water and Sewerage Department (DWSD) to an independent supply from Lake Huron, as DWSD rates were rising. The construction of the pipeline began in March 2013, which would take more than two years to complete. In the meantime, the city decided to terminate the DWSD contract and in April 2014 it started sourcing its drinking water from the Flint River until the completion of the new pipeline. The switch was done despite warnings from county officials and plant supervisors (Fleming, 2018; Taddonio, 2019).

Changes in the water quality did not take long to become apparent. Flint residents voiced concerns about water color, taste, and odor, but during May and June 2014 local authorities, including the Michigan Department of Environmental Quality (MDEQ), the Mayor of Flint, and the emergency manager, assured them the water was safe to drink (Fonger, 014a; Fonger, 014b). Between August and September 2014, three water boil advisories were issued due to high levels of fecal coliform, while authorities increased chlorine level in treatment to deal with this issue. A month later, General Motors stopped using Flint's water citing high corrosion levels due to chlorine was damaging engine parts (Fonger, 014c; Fonger, 014d). In January 2015, the MDEQ issued a notice of violation of the Safe Drinking Water Act for maximum contaminant levels for trihalomethanes, chemicals formed as byproduct of disinfecting water.

In February 2015 high lead levels in water started to be reported to authorities, including the U.S. Environmental Protection Agency (EPA). Some residential water samples contained more than 100 times the EPA lead limit (Semuels, 2015). By September 2015, two reports from independent research groups had shown high residential water lead levels, and an increase in the number of children with elevated blood lead levels (Hulett, 2015). After these events, local authorities declared a public health emergency on October 1, 2015, and the city reconnected to the DWSD shortly after. In January 2016, the Governor

of Michigan and President Obama issued states of emergency, authorizing federal and state aid to the city.

In January 2016, it was announced by the Governor of Michigan that cases of Legionnaries' disease, a deadly pneumonia caused by *Legionella* bacteria, had spiked in Flint and Genesee County around the time of the water switch. Between June 2014 and November 2015, 87 cases were reported, with 10 cases resulting in death. In previous years, cases reported were less than one fourth of those reported between 2014 and 2015. However, by the time of the Governor's announcement, no government agency had tested the water supply for *Legionella* bacteria (Bernstein & Dennis, 2016; Ford, 2016). Legionnaries' disease is transmitted from the environment, when contaminated water vapor is inhaled or when the water is used in invasive medical procedures. Schwake et al. (2016) found high levels of *Legionella* bacteria in large buildings in Flint during 2015, including hospitals, while Zahran et al. (2018) used difference-in-differences models and linked the timing of the disease outbreak with the switch in water supply. A recent investigation from FRONTLINE, a PBS journalist team, has suggested that the actual death toll of the outbreak could be around 70 deaths, by analyzing pneumonia deaths during the same period. The medical community was notified of the outbreak only halfway through it, and as Legionnaries' disease manifests as a severe case of pneumonia, healthcare providers without the right information or experience could have misdiagnosed patients with the disease (Bellware, 2019; Childress, 2019).

### 2.3 Data

To implement my evaluation strategy, I use individual-level hospitalization and mortality data to construct city-level inpatient rates for the 2012-2016 period and mortality rates for the 2010-2017 period for the state of Michigan. I focus on the 15 largest non-Flint cities in Michigan: Ann Arbor, Dearborn, Detroit, Farmington Hills, Grand Rapids, Kalamazoo, Lansing, Livonia, Rochester Hills, Southfield, Sterling Heights, Troy, Warren, Westland, and Wyoming.

Hospitalization variables come from Medicare administrative data. My sample includes all beneficiaries 65 and older living in the state of Michigan, accounting for around 98 percent of state elderly residents. Age and zip code of residence are obtained for all beneficiaries from the 2012-2016 Medicare enrollment files. Health care use is derived from the Medicare Provider Analysis and Review (MedPAR) File, which reports information on each inpatient stay in a hospital or skilled nursing facility for any beneficiary enrolled in fee-for-service (FFS) Medicare.[1] MedPAR observations are derived from facility (Medicare Part A) service claims corresponding to that stay and include the date of admission, length of

[1] I exclude beneficiaries enrolled in Medicare Advantage plans as I don't observe their whole healthcare utilization.

stay, and total cost of the stay. They also include a principal diagnosis code that allows me to distinguish between different causes of inpatient admission.

I calculate quarterly hospitalization rates per 1000 population as:

$$HR_{ct} = \frac{TotalAdmissions_{ct}}{MedicarePop_{ct}} \cdot 1000 \tag{2.1}$$

where $c$ indices city, and $t$ the quarter and year. $TotalAdmissions_{ct}$ is calculated considering all inpatient admissions during a particular quarter using Medicare records, while $MedicarePop_{ct}$ corresponds to the total number of Medicare beneficiaries in a given year, derived from the enrollment files. Principal diagnoses included in the inpatient files allow me to construct all-cause, cardiovascular disease, ischemic disease, kidney disease, and pneumonia hospitalization rates.

The mortality data are constructed using the National Mortality Detail Files. These data contain abstracted death certificate information including the city and county of residence, age, sex, and diagnosed cause of death for all deaths in the U.S. To ensure confidentiality, the city is listed only for those individuals who reside in a city of over 100,000 in population. This means I can only observe people living in Ann Arbor, Detroit, Flint, Grand Rapids, Lansing, Sterling Heights, and Warren. I restrict my sample to residents of Michigan who are at least 65 years old.

I calculate quarterly mortality rates per 100000 population as:

$$MR_{ct} = \frac{TotalDeaths_{ct}}{Population65plus_{ct}} \cdot 100000 \tag{2.2}$$

where $c$ indices city, and $t$ the quarter and year. $TotalDeaths_{ct}$ comes from the mortality files, while $Population65plus_{ct}$ comes from US Census Bureau estimates. Diagnosed causes of death included in the mortality files allow me to construct all-cause, cardiovascular disease, ischemic disease, kidney disease, and pneumonia mortality rates.

Additionally, to implement synthetic controls I use socioeconomic characteristics of cities derived from the American Community Survey (1-Year estimates).

### 2.4 Empirical Strategy

To obtain an estimate of the causal effect of the Flint water crisis on hospitalization and mortality rates, I use different versions of a differences-in-differences strategy and synthetic controls. In particular, I compare hospitalization and mortality rates between Flint and other large cities in Michigan, before and after the switch in the water source.

First, I estimate city-by-quarter difference-in-differences regressions of the following form:

$$Y_{ct} = \alpha_c + \delta_t + \beta \cdot I\{c = Flint\} \cdot I\{t \geq 2014q2\} + \epsilon_{ct} \tag{2.3}$$

The variable $Y_{ct}$ represents one of the different hospitalization or mortality rates for city $c$ and quarter $t$. The model includes city fixed effects ($\alpha$), quarter-year fixed effects ($\delta$), and an error term ($\epsilon$) that is assumed to be uncorrelated with other unobserved determinants of the outcome variable. Standard errors are clustered at the city level.

The key coefficient of interest is $\beta$, which is the difference-in-differences estimate of the effect of the Flint crisis. This coefficient is identified by comparing outcomes in Flint after the crisis to outcomes in Flint before the crisis and to other cities in Michigan. The key identifying assumption is that outcomes in Flint would not have evolved differently to other cities in the state in the absence of the crisis.

The state of emergency was declared in December 2015. Therefore, it is expected that any impact on hospitalizations should be concentrated between the water switch and the emergency declaration, given the public health measures undertaken afterwards. To focus on this period, I estimate the following regression for hospitalization rates:

$$HR_{ct} = \alpha_c + \delta_t + \beta \cdot I\{c = Flint\} \cdot I\{2014q2 \leq t \leq 2015q4\} + \epsilon_{ct} \tag{2.4}$$

To test for the key identifying assumption of difference-in-differences, I estimate the following equation:

$$Y_{ct} = \alpha_c + \delta_t + \beta_t \cdot I\{c = Flint\} \cdot I\{t = 2010q1, 2010q2, ..., 2017q4\} + \epsilon_{ct} \tag{2.5}$$

All variables are defined as above in equation (2.3) but $\beta_t$ is a vector which takes a unique value for each quarter from 2010q1-2017q4 (2012q1-2016q4 for hospitalization rates). The base year is 2014q1, the last quarter before the change in the water source.

Additionally, I supplement my previous analysis using synthetic control methods (Abadie & Gardeazabal, 2003; Abadie et al., 2010). The main idea behind this methodology is to construct a synthetic match for Flint by using the other cities in the control group such that the synthetic Flint behaves similarly to actual Flint before the water switch in April 2014. Specifically, let the index $c = (0, 1, ..., C)$ denote cities. The value $c = 0$ corresponds to Flint and $c = (1, ..., C)$ correspond to each of the other $C$ states that are candidate contributors to the control group (the donor pool). Define $F_0$ as a $kx1$ vector with elements equal to the outcome of interest (hospitalization or mortality rate) in Flint in each quarter be-

tween 2010q1 and 2014q1 (between 2012q1 and 2014q1 for hospitalization rates) plus additional year-level covariates predictive of the outcome. Similarly, define the $kxC$ matrix $F_1$ as the collection of comparable data vectors for each of the $C$ cities in the donor pool.

The synthetic control method identifies a convex combination of the $C$ cities in the donor pool that best approximates the pre-treatment data vector for the treated city. Define the $Cx1$ weighting vector $W = (w_1, w_2, ..., w_C)$ such that $\sum_{c=1}^{C} w_c = 1$, and $w_c \geq 0$ for $c = (1, ..., C)$. The product $F_1 \cdot W$ then gives a weighted average of the pre-treatment vectors for all cities omitting Flint, with the difference between Flint and this average given by $F_0 - F_1 \cdot W$. The synthetic control method chooses a value for the weighting vector $W$ that yields a synthetic comparison group (consisting of an average of some subset of donor cities) that best approximates pre-treatment Flint. The weighting vector is chosen by solving the following minimization of the mean squared prediction error (MSPE):

$$W^* = argmin_W (F_0 - F_1 W)' V (F_0 - F_1 W) \quad \text{st} \quad \sum_{c=1}^{C} w_c = 1, \quad w_c \geq 0, \quad \text{for} \quad c = (1, ..., C) \tag{2.6}$$

where $V$ is a $kxk$, diagonal positive-definite matrix with diagonal elements equal to relative weights for the contribution of the square of the elements in the vector $F_0 - F_1 \cdot W$ to the objective function.

After obtaining the weighting vector $W^*$, both pre and post-treatment values for the outcome can be calculated for synthetic Flint using the corresponding weighted average for each quarter using the donor cities with positive weights, $\sum_{c=1}^{C} w_c^* Y_{ct}$ for a given $t$ from the study period. The post-treatment values for synthetic Flint serve as the counterfactual outcomes for Flint. Computing the post-treatment difference between the dependent variables of Flint and synthetic Flint provides the synthetic control estimate of the treatment effect attributable to the water switch in that city for any given quarter $t$, $\hat{\alpha}_{0t} = Y_{0t} - \sum_{c=1}^{C} w_c^* Y_{ct}$. In addition to including all pre-treatment values of the dependent variable in $F_0$ and $F_1$, I also include the share of nonwhite population, share of population 25 years and over with at least high school diploma, unemployment rate, and median household income. These additional covariates are measured yearly between 2010 and 2014 (between 2012 and 2014 for hospitalization rates).

To test the significance of any observed relative effect in Flint's outcomes, I apply the permutation test suggested by Abadie et al. (2010) to the synthetic control estimate $\hat{\alpha}_{0t}$. This tests consists on taking every unit from the donor pool, assuming it was treated at the same time, and finding their synthetic comparison group based on the solution to the minimization problem in equation (2.6). The results from the permutation tests are then used to construct a distribution of placebo estimates. Then, the effect of the treatment on the unit affected by the intervention is considered to be significant if its magnitude is extreme relative to the permutation distribution. Following Galiani & Quistorff (2017), suppose the

distribution of placebo effects is $\hat{\alpha}_{0t}^{PL} = \{\hat{\alpha}_{ct} : c \neq 0\}$. The two-sided $p$-value is then

$$\begin{aligned} p\text{-value} &= Pr\left(\mid \hat{\alpha}_{0t}^{PL} \mid \geq \mid \hat{\alpha}_{0t} \mid\right) \\ &= \frac{\sum_{c\neq 0} 1\left(\mid \hat{\alpha}_{ct}^{PL} \mid \geq \mid \hat{\alpha}_{0t} \mid\right)}{C} \end{aligned} \tag{2.7}$$

If treatment is randomized, then the $p$-value has the standard interpretation from classical randomization inference. Otherwise, it can be interpreted as the proportion of control units with an estimated effect as large as that of the treated unit. In cases where units were not matched well in the pre-treatment period, placebo effects can be too large, causing $p$-values to be too conservative. One way to address this is to divide all estimates by the corresponding pre-treatment match quality, or pre-treatment root mean squared prediction error (RMSPE), obtaining an standardized $p$-value.

In order to assess the joint significance of effects across all post-treatment periods, the authors suggest using the ratio of post/pre-treatment RMSPE for all cities. This ratio should be high for the actual treated city, suggesting a good fit in the pre-treatmet trends (small RMSPE) but a failure at replicating post-treatment trends (large RMSPE). Using the share of cities with a ratio as large as the treated city Flint, allows me to examine whether the effect on the water switch is large relative to the distribution of effects for cities not exposed to the treatment.

### 2.5 Results

Figures 2.1 and 2.2 show raw quarterly trends in hospitalization and mortality rates respectively for different diagnoses, comparing Flint with other large Michigan cities. Figure 2.1 shows a systematic difference between Flint and other Michigan cities with respect to inpatient admissions, with Flint presenting higher hospitalization rates either for all-cause or for specific diagnoses. Also, because Flint is smaller relative to the other cities taken together, hospitalization trends present more volatility. In terms of mortality rates, Figure 2.2 still shows higher volatility for Flint. However, in particular for all-cause, cardiovascular, and ischemic mortality rates, Flint presents similar levels to those from other Michigan cities. Table 2.1 shows summary statistics of the socioeconomic characteristics used in the synthetic controls analysis and of outcomes during the pre-intervention period. There are differences in the characteristics of Panel A between treatment and control groups, which reflects the heterogeneity of the cities in the control group. However, the outcomes reported in Panel B look similar between groups.

It is important to note that the raw data presented in Figures 2.1 and 2.2 do not show any evidence of changes happening in Flint after the crisis. Either before and after the second quarter of 2014,

hospitalization and mortality rates trends appear to behave similarly. To account for seasonality and have a graphical representation of the main trends underlying hospitalization and mortality rates, I fit separate local quadratic regressions on both sides of the event for Flint and the control group.[2] Using theses estimates, I constructed quarter-specific trends which I show against the raw data in Figures 2.3 and 2.4. These graphs show that overall hospitalization and mortality rates present either flat or downward trends across time with the exception of hospitalizations for cardiovascular disease that has an upward trend.

Panel A of Table 2.2 shows the difference-in-differences estimates derived from equations (2.3) and (2.4) for hospitalization rates. I estimate that after the crisis, Flint observed 1.8 more all-cause hospitalizations per 1,000 elderly people per quarter, which represents a 1.7% increase with respect to the pre-intervention mean of the dependent variable. Ischemic and kidney disease hospitalization rates also present significant results. Their magnitudes imply a 16% decrease in ischemic disease hospitalization rates and a 11% increase in kidney disease hospitalization rates. Estimates for cardiovascular disease and pneumonia are not statistically significant. Additionally, results are consistent between the two specifications from equations (2.3) and (2.4). Figure 2.3 is useful to understand these estimates. The increase in all-cause hospitalizations seems to be driven by an increase during 2014q2 and 2015q1, while the increase in kidney disease hospitalizations looks to be driven by a fall in 2014q1, right before the start of the crisis.

Panel B of Table 2.2 shows the difference-in-differences estimates derived from equation (2.3) for mortality rates. I estimate that after the crisis, Flint observed 61 more all-cause deaths per 100,000 elderly people per quarter, which represents a 5% increase with respect to the pre-intervention mean of the dependent variable. Kidney disease mortality also presents a significant result with an increase equal to 14% of the pre-intervention mean of the outcome. As with hospitalizations, Figure 2.4 shows evidence that these results are driven by particular observations during the post-intervention period, suggesting at most a short-lived effect rather than a permanent one. Cities used in the mortality analysis are a subset of those available when looking at hospitalization rates. Difference-in-differences results are qualitative the same for hospitalizations when using the subset of cities.

[2] Let $Y_{ct}$ be an outcome in city $c$ at quarter $t$, $Flint$ a dummy variable equal to 1 if $c = Flint$, $Post$ a dummy variable equal to 1 if $t >= 2014q2$, and $t$ a variable representing quarters from $2014q2$. I then run the following equation:

$$\begin{aligned} Y_{ct} &= \alpha_0 \cdot (1 - Flint) + \alpha_1 \cdot t \cdot (1 - Post) \cdot (1 - Flint) + \alpha_2 \cdot t \cdot Post \cdot (1 - Flint) + \alpha_3 \cdot Post \cdot (1 - Flint) \\ &+ \alpha_4 \cdot t^2 \cdot (1 - Post) \cdot (1 - Flint) + \alpha_5 \cdot t^2 \cdot Post \cdot (1 - Flint) \\ &+ \beta_0 \cdot Flint + \beta_1 \cdot t \cdot (1 - Post) \cdot Flint + \beta_2 \cdot t \cdot Post \cdot Flint + \beta_3 \cdot Post \cdot Flint \\ &+ \beta_4 \cdot t^2 \cdot (1 - Post) \cdot Flint + \beta_5 \cdot t^2 \cdot Post \cdot Flint \end{aligned}$$

Recent epidemiological evidence has found that increases in lead exposure are associated with a 25 to 46% increase in all-cause mortality (Lanphear et al., 2018; Menke et al., 2006; Schober et al., 2006; Lustberg & Silbergeld, 2002), 39 to 70% increase in cardiovascular disease mortality (Lanphear et al., 2018; Menke et al., 2006; Lustberg & Silbergeld, 2002), 10 to 89% increase in ischemic disease prevalence (Navas-Acien et al., 2006), and 9% increase in kidney disease prevalence (Muntner et al., 2003). Mortality results from Table 2.2 as percentage of the pre-intervention mean of the outcome for all-cause, cardiac, and ischemic disease mortality rates are all smaller than what is expected from previous epidemiological results. Additionally, the literature has found a positive relation between lead exposure and ischemic disease mortality, while both my hospitalization and mortality results suggest a negative relation. For kidney disease hospitalizations, my estimate implies an increase of 11% after the crisis which is in line with the 9% in prevalence found in the epidemiological literature.

Regarding potential misdiagnoses of Legionnaries' disease, hospitalization and mortality results for pneumonia are not significant. This suggests that pneumonia cases did not experience any particular change in Flint after the water crisis, implying no Legionnaries' disease misdiagnoses.

Figures 2.5 and 2.6 plot the coefficients ($\beta_t$) on Flint-specific quarter fixed effects for each group of outcomes (hospitalization and mortality rates) generated from equation (2.5). The solid vertical line denotes the last quarter before the start of the crisis. Results for hospitalization rates from Figure 2.5 are noisy and question the validity of the parallel trends assumption for these outcomes. On the other hand, results for mortality rates from Figure 2.6 provide evidence in favor of the presence of parallel trends.

To address the issue of parallel trends and provide a robustness exercise for the previous analysis, I also implement synthetic control methods. I begin with Figures 2.7 and 2.8 showing a graphical presentation of the Flint hospitalizations and mortality rate trends and the comparable trends in synthetic Flint. Focusing first on hospitalization rates, Figure 2.7 reveals that the synthetic control group does a better job at matching pre-intervention trends in Flint than the raw data used in the difference-in-differences analysis shown in Figure 2.1. For mortality rates, synthetic Flint trends from Figure 2.8 look similar to the raw data from Figure 2.2. Table 2.3 shows the cities that received positive weights to create the synthetic control for different outcomes.[3]

Figures 2.9 and 2.10 show the data needed to conduct the permutation tests of the significance of the relative changes in Flint. Specifically, for each of the cities in the donor pool as well as for Flint, the figures display the quarter-by-quarter difference between the outcome variable for the treated city and the outcome variable for the synthetic control. The differences for each of the donor cities are displayed

[3]Medicare data allow me to identify more cities than the mortality data, therefore I use a larger donor pool when analyzing hospitalization rates. Results are qualitative the same if I use a smaller donor pool.

with gray lines, while the difference for Flint is displayed by the black line. As with the raw data from Figures 2.1 and 2.2, graphical evidence presented in Figures 2.9 and 2.10 does not suggest any noticeable change in Flint after the water crisis.

Tables 2.4 and 2.5 summarize Figures 2.9 and 2.10, and provide information on the significance of the effects estimated using synthetic controls. Table 2.4 shows in each column the magnitude of the difference between Flint and synthetic Flint for different hospitalization rates across all quarters of the post-intervention period, with their corresponding standardized $p$-value derived from equation (2.7). Additionally, the last two columns show the average affect across post-intervention quarters and the $p$-value of the ratio of post/pre-treatment RMSPE respectively. The magnitudes of the average effects are in general close to those from the difference-in-differences analysis, with the average effect for all-cause and cardiovascular disease hospitalization rates falling within the confidence intervals derived from Panel A in Table 2.2. The $p$-values for the ratios of post/pre-treatment RMSPE indicate that synthetic control estimates of the Flint water crisis on hospitalization outcomes are not significant at standard levels. However, the two outcomes with the lowest ratio $p$-values, all-cause and kidney disease, are the ones that present significant quarter-level estimates, and also significant effects in the difference-in-differences analysis from Table 2.2.

Table 2.5 has the same structure as Table 2.4 but for mortality rates. Similar to the previous table, average mortality effects are close to those from the difference-in-differences analysis, and with the exception of kidney disease, they all fall into the confidence intervals derived from Panel B in Table 2.2. The $p$-values for the ratios of post/pre-treatment RMSPE indicate that synthetic control estimates of the Flint water crisis on mortality rates are significant for all-cause mortality, and at higher levels of significance for cardiovascular disease mortality.

Taken together, the results from difference-in-differences and synthetic controls analyses suggest that the Flint water crisis had a weak impact on all-cause and kidney disease inpatient admissions, and on all-cause mortality rates; however, these results are usually smaller than what has been suggested by the epidemiological literature and the graphical evidence does not strongly support them. Effects on cardiovascular and ischemic disease are undetected, as well as effects on hospitalizations or deaths for pneumonia, suggesting no misdiagnoses of Legionnaries' disease.

## 2.6 Conclusion

This paper presents results from a study of the health impacts on the elderly population of the Flint water crisis. I find weak evidence suggesting an increase in all-cause and kidney disease hospitalization

rates, and an increase in all-cause mortality rates.

The main limitation of this study is that previous evidence has shown that lead builds up in the body over time. By focusing on elderly adults I expect they have been exposed to high levels of lead while working in particular unregulated occupations or by having encounter high levels of contamination in the environment. However, I don't have an objective measure of lead exposure before the crisis for this population, and the length of the post-intervention period could be too short to find stronger results.

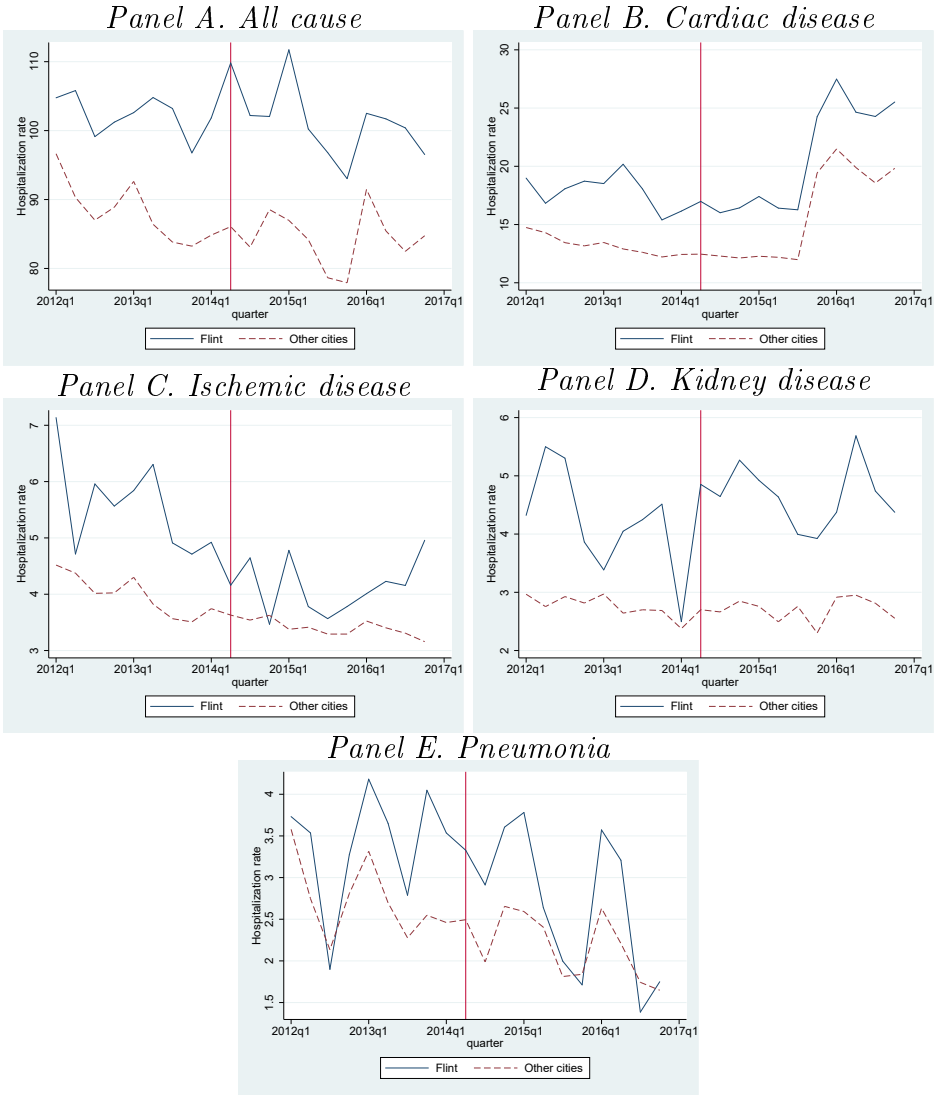
The previous limitation is also relevant to understand the timing of the effects and the relation between hospitalizations and deaths. In the short run, I would expect the sickest patients, as measured by higher levels of lead in blood or bones at baseline, to be the ones driving any change in mortality rates, while those with lower levels of exposure should be driving most of hospitalization rates changes as they start becoming sick. In the medium and long run, lead would accumulate and increase in the bodies of those with lower baseline levels, potentially driving increases in hospitalization and mortality rates.

Another limitation of this study is that it could be subject to a competing risks problem. The main hypothesis that I test in this paper is that lead exposure increases the prevalence of cardiac and kidney disease, which I measure through hospitalizations and deaths caused by those diagnoses. If lead is in fact a risk that affects both diseases, the unobserved durations until hospitalization or death will be dependent between diagnoses, and observing one will likely influence the chance of observing the other one. But, under the assumption the hypothesis is true, I would expect to observe a significant result for at least one of the diagnoses. In particular, taking into account the results from epidemiological studies, I would expect that competing risks imply significant results for cardiovascular disease. However, if anything, the results from difference-in-differences and synthetic controls analyses show the opposite.

Most of the causal evidence on health effects of lead exposure has been focused on children's outcomes. Unfortunately, as lead problems have been recently reported in different cities across the country (Schwartz & Wines, 2016; Corasaniti et al., 2019), it is important to understand how lead contamination affects different population groups.

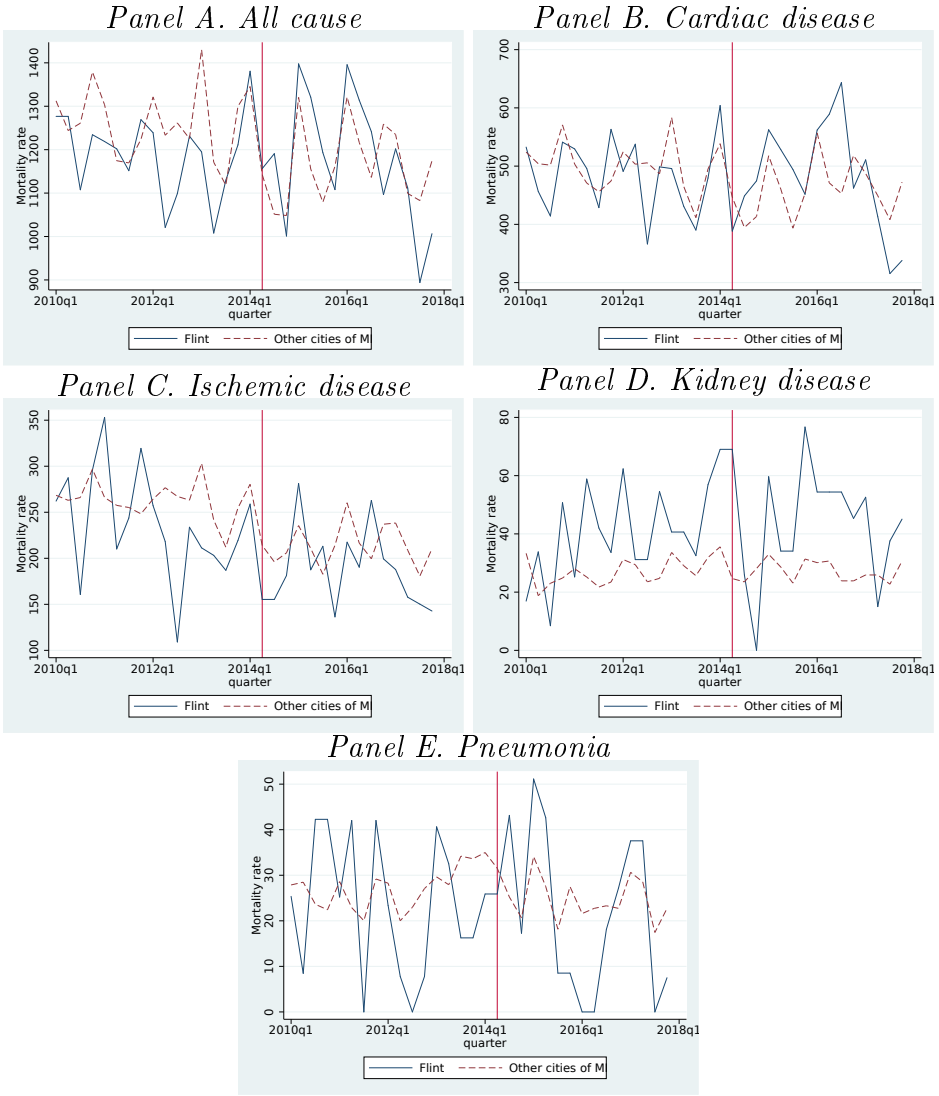
## 2.7 Figures

Figure 2.1: Hospitalization rates in Flint vs other cities in Michigan, quarterly trends



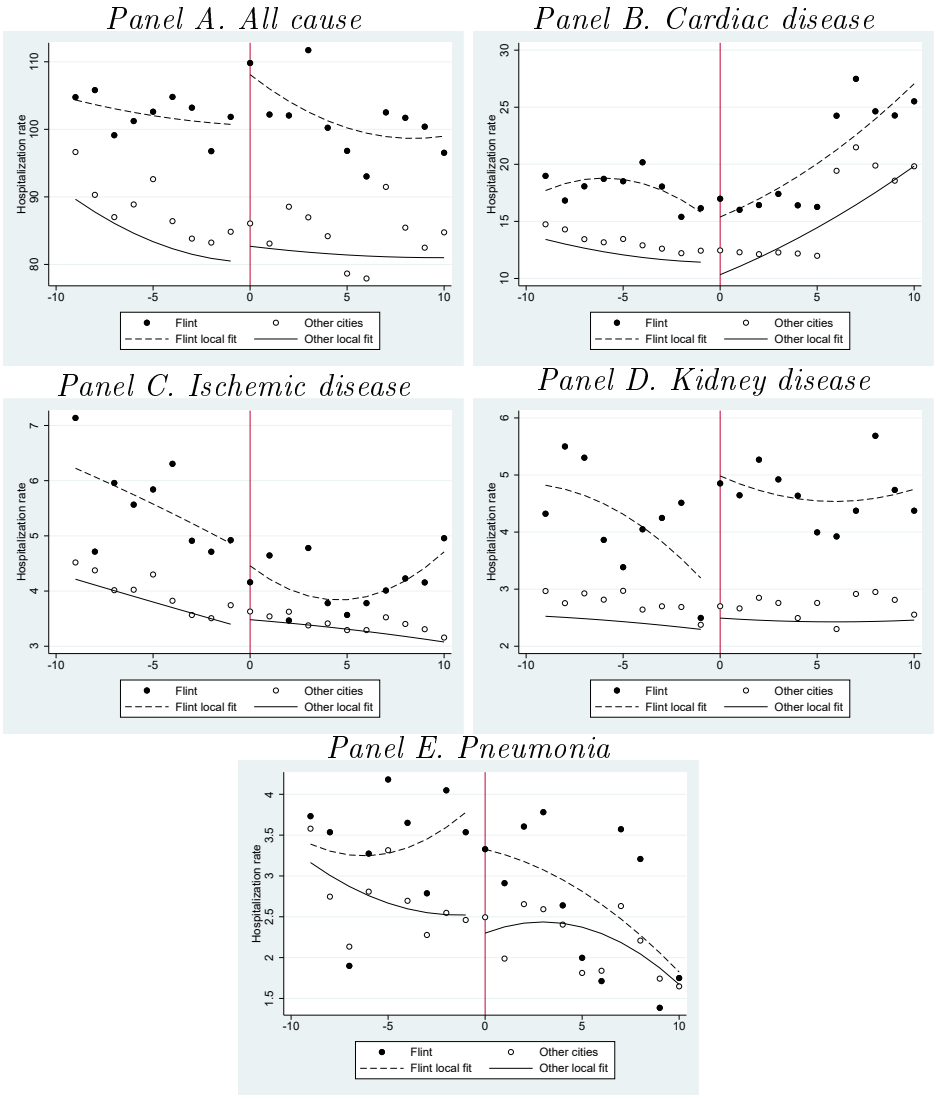

*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the second quarter of 2014 when the crisis started. Other cities include Ann Arbor, Dearborn, Detroit, Farmington Heights, Grand Rapids, Kalamazoo, Lansing, Livonia, Rochester Hills, Southfield, Sterling Heights, Troy, Warren, Westland, and Wyoming.

Figure 2.2: Mortality rates in Flint vs other cities in Michigan, quarterly trends



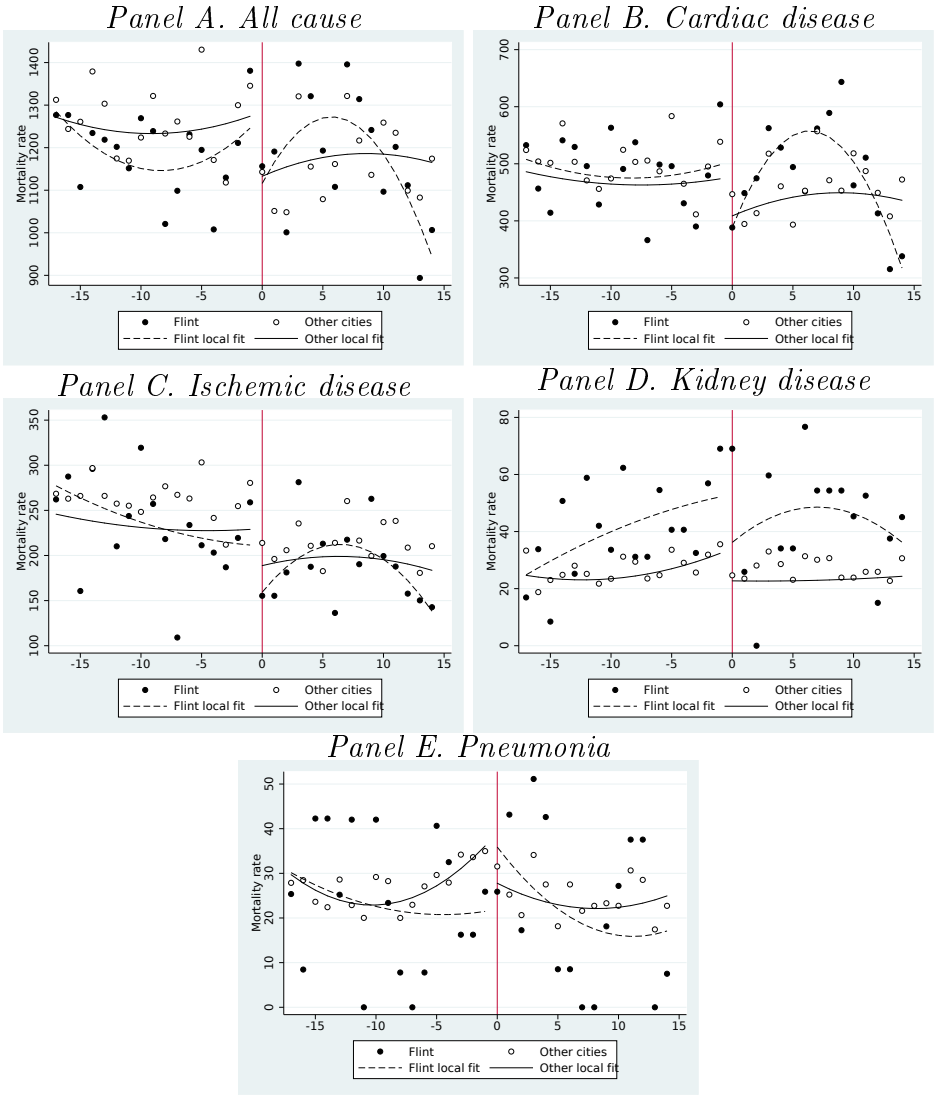

*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the second quarter of 2014 when the crisis started. Other cities include Ann Arbor, Detroit, Grand Rapids, Lansing, Sterling Heights, and Warren.

Figure 2.3: Adjusted hospitalization rates in Flint vs other cities in Michigan, quarterly trends



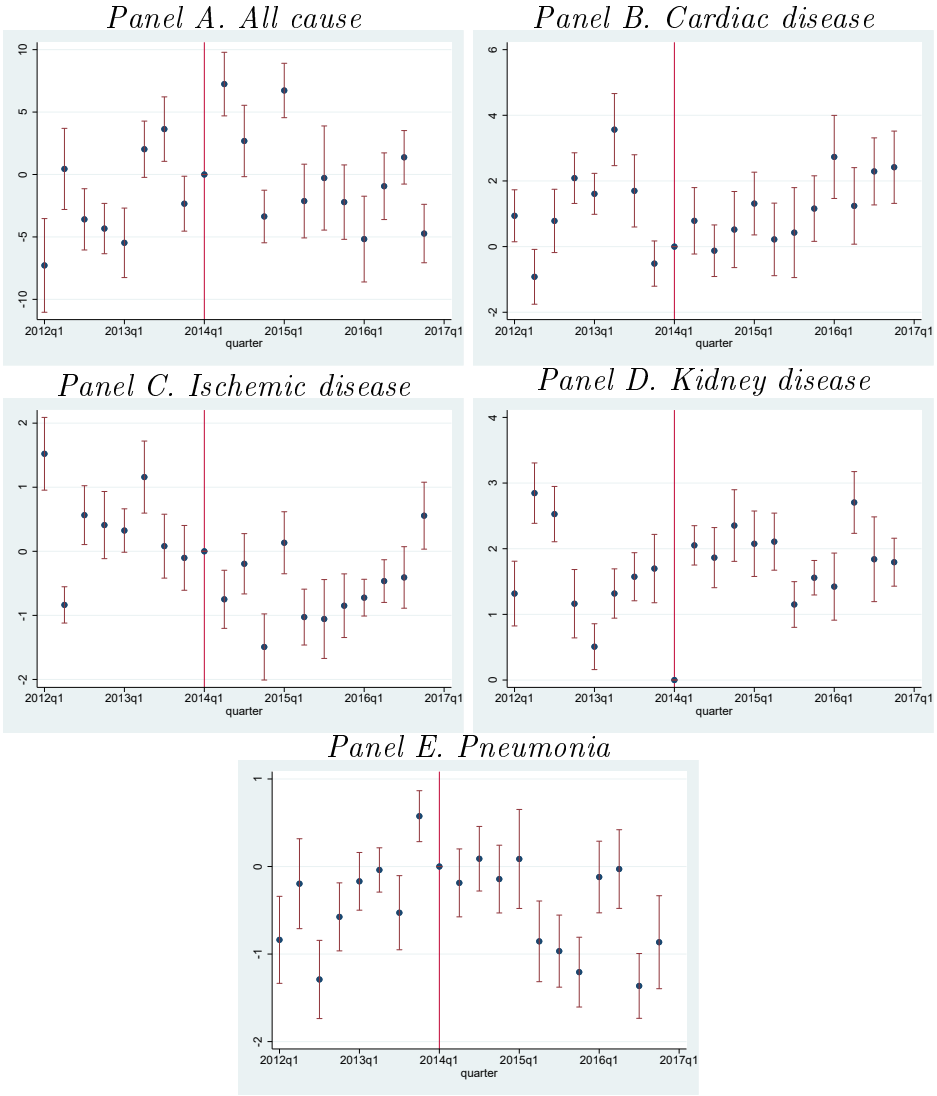

*Notes:* Sample includes all adults ages 65 and older. Black and white dots show raw quarterly trends for Flint and Other cities, respectively. Dashed and solid lines are quadratic fits derived after running separate local quadratic regressions on both sides of the event for Flint and Other cities, respectively. The red vertical line indicates the second quarter of 2014 when the crisis started, around which the x-axis is centered. Other cities include Ann Arbor, Dearborn, Detroit, Farmington Heights, Grand Rapids, Kalamazoo, Lansing, Livonia, Rochester Hills, Southfield, Sterling Heights, Troy, Warren, Westland, and Wyoming.

Figure 2.4: Adjusted mortality rates in Flint vs other cities in Michigan, quarterly trends



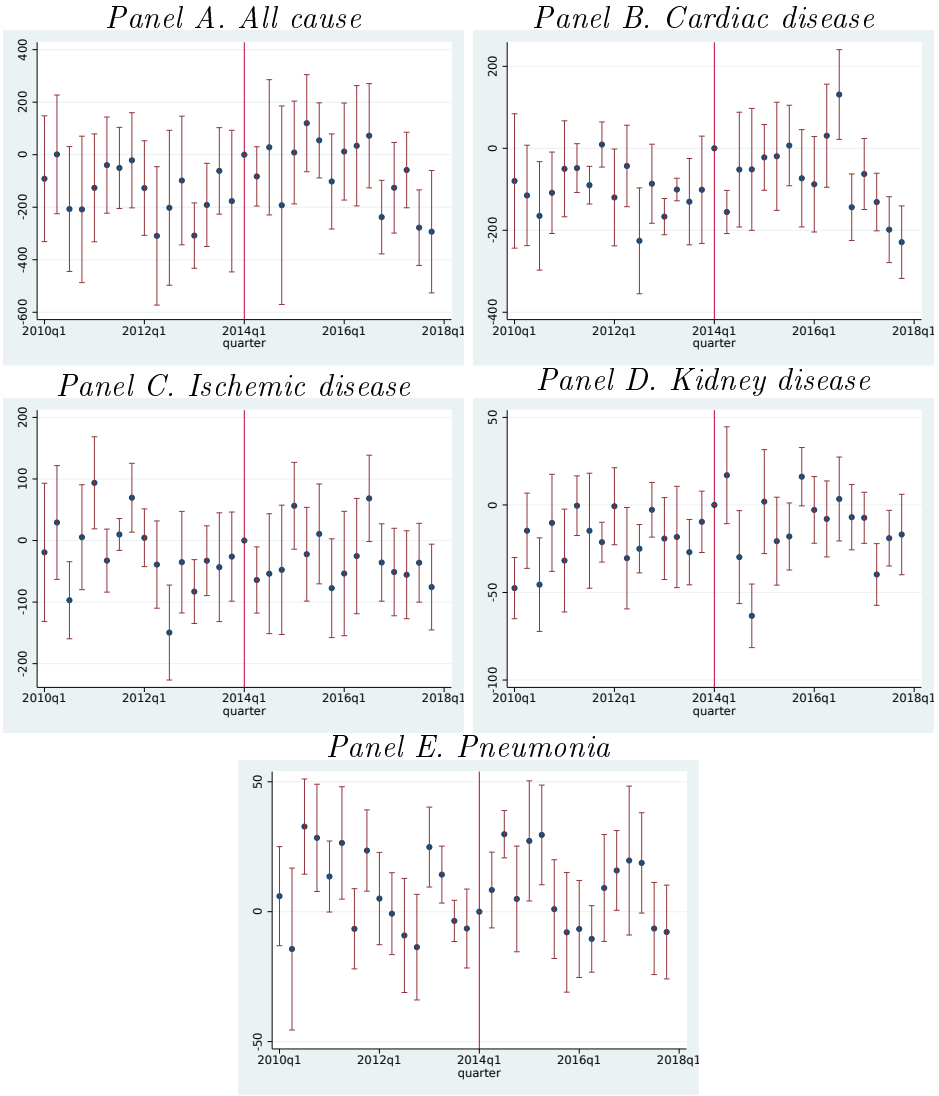

*Notes:* Sample includes all adults ages 65 and older. Black and white dots show raw quarterly trends for Flint and Other cities, respectively. Dashed and solid lines are quadratic fits derived after running separate local quadratic regressions on both sides of the event for Flint and Other cities, respectively. The red vertical line indicates the second quarter of 2014 when the crisis started, around which the x-axis is centered. Other cities include Ann Arbor, Detroit, Grand Rapids, Lansing, Sterling Heights, and Warren.

Figure 2.5: Graphic analysis parallel trends - Hospitalization rates



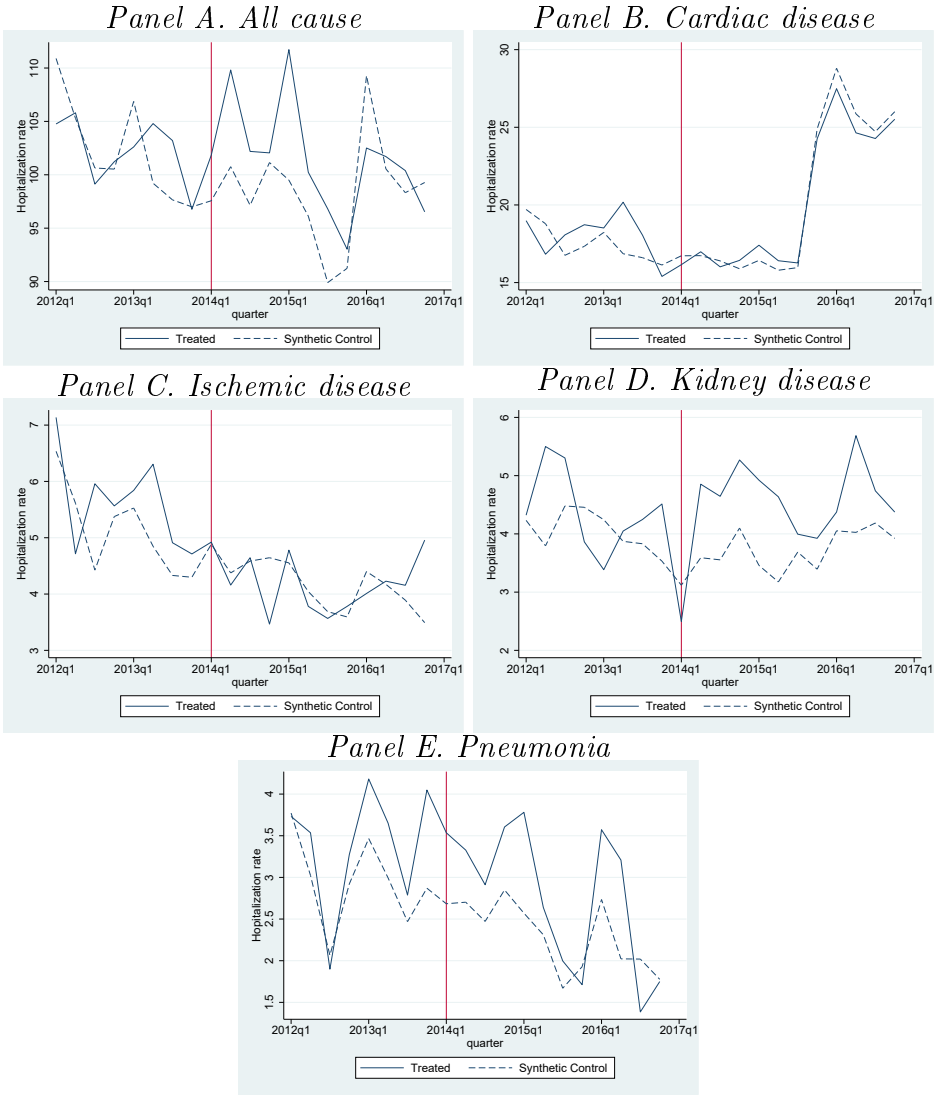

*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the first quarter of 2014 with respect to which the coefficients were estimated.

Figure 2.6: Graphic analysis parallel trends - Mortality rates



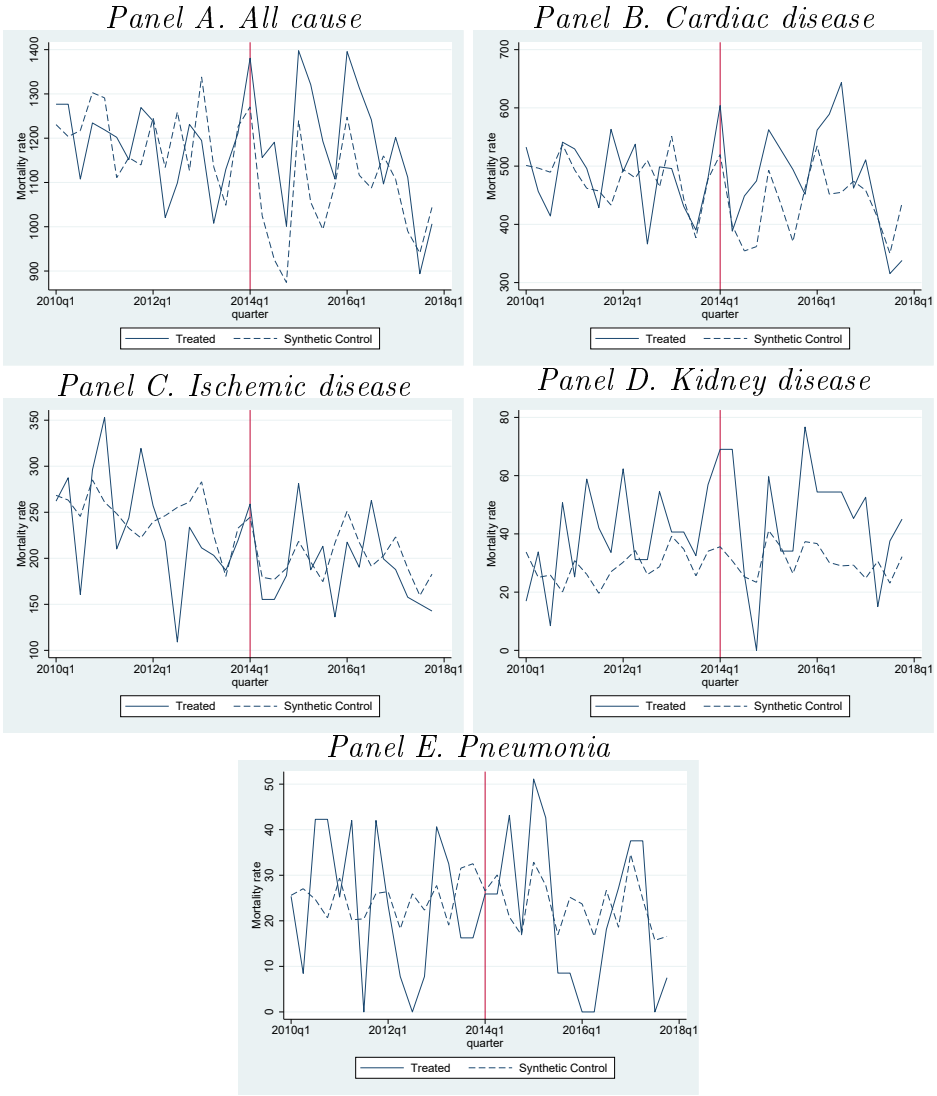

*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the first quarter of 2014 with respect to which the coefficients were estimated.

Figure 2.7: Trends in hospitalization rates - Flint and synthetic Flint



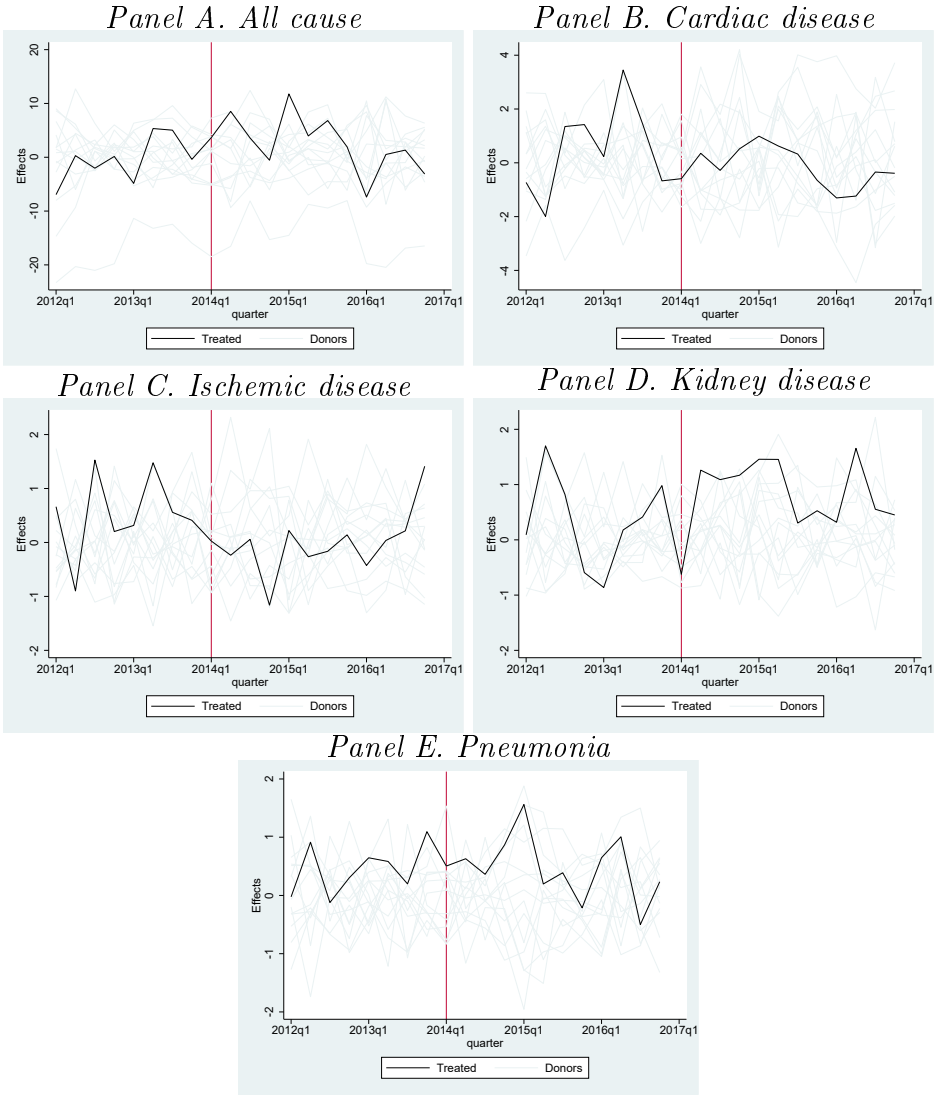

*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the first quarter of 2014.

Figure 2.8: Trends in mortality rates - Flint and synthetic Flint



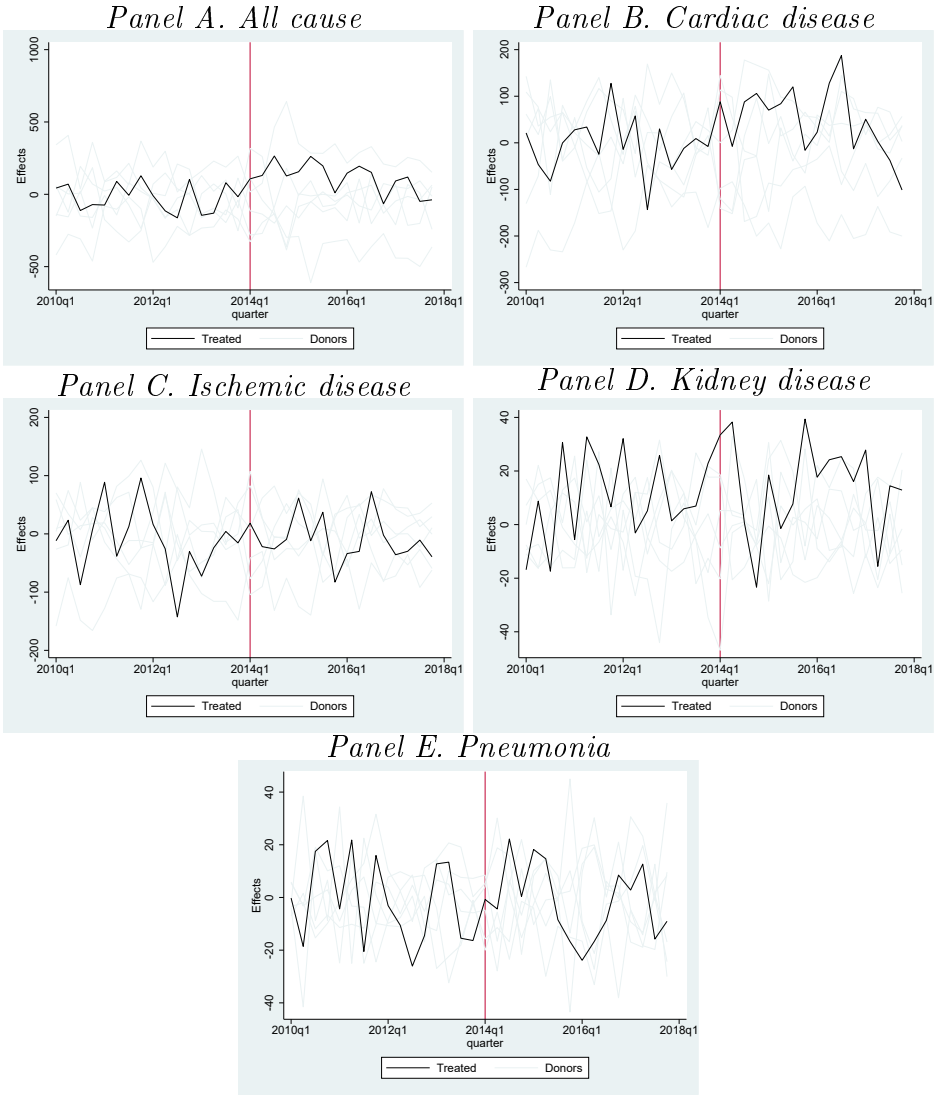

*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the first quarter of 2014.

Figure 2.9: Difference in hospitalization rates relative to synthetic control group, all cities




*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the first quarter of 2014. Flint is displayed with the black line while cities in the donor pool appear in gray.

Figure 2.10: Difference in mortality rates relative to synthetic control group, all cities




*Notes:* Sample includes all adults ages 65 and older. The red vertical line indicates the first quarter of 2014. Flint is displayed with the black line while cities in the donor pool appear in gray.

## 2.8 Tables

Table 2.1: Summary statistics

| | Flint | | Other cities | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| Panel A: Socioeconomic characteristics, city-by-year, 2010-2014 | | | | |
| Median HH income | 24592.2 | 1965.857 | 50629.97 | 17169.84 |
| Share nonwhite | 0.596 | 0.012 | 0.323 | 0.219 |
| Share high school diploma or more | 0.824 | 0.006 | 0.888 | 0.056 |
| Unemployment rate | 0.261 | 0.041 | 0.115 | 0.057 |
| Observations | 5 | | 70 | |
| Panel B: Outcomes, city-by-quarter, 2010q1-2014q1 | | | | |
| *Hospitalization rates* | | | | |
| All-cause | 102.238 | 2.904 | 84.078 | 15.509 |
| Cardiovascular | 17.878 | 1.497 | 12.216 | 3.117 |
| Ischemic | 5.562 | 0.832 | 3.803 | 1.17 |
| Kidney | 4.186 | 0.916 | 2.422 | 0.938 |
| Pneumonia | 3.404 | 0.696 | 2.74 | 0.731 |
| Observations | 9 | | 126 | |
| *Mortality rates* | | | | |
| All-cause | 1191.248 | 95.917 | 1247.938 | 231.602 |
| Cardiovascualr | 485.644 | 64.211 | 469.593 | 108.638 |
| Ischemic | 237.09 | 59.015 | 232.693 | 73.589 |
| Kidney | 40.504 | 16.484 | 25.612 | 11.835 |
| Pneumonia | 23.421 | 15.223 | 26.808 | 12.46 |
| Observations | 17 | | 102 | |

*Notes:* This table reports summary statistics on socioeconomic characteristics used in the synthetic control analysis (Panel A) and outcome variables (Panel B). Variables are measured during the pre-intervention period at year level for city characteristics (2010-2014), and quarter level for outcomes (2010q1-2014q1/2012q1-2014q1). As hospitalization rates are only available from 2012 but the data allow me to identify more cities, there is a difference in the number of observations between hospitalization and mortality rates. Other cities include Ann Arbor, Dearborn, Detroit, Farmington Heights, Grand Rapids, Kalamazoo, Lansing, Livonia, Rochester Hills, Southfield, Sterling Heights, Troy, Warren, Westland, and Wyoming (Ann Arbor, Detroit, Grand Rapids, Lansing, Sterling Heights, and Warren for mortality outcomes).

Table 2.2: Effect of Flint crisis on hospitalizations - Other Michigan cities

| | All-cause | Cardiovascular | Ischemic | Kidney | Pneumonia |
|---|---|---|---|---|---|
| *Panel A: Hospitalization rates* | | | | | |
| Flint x post 2014q2 | 1.806*** | 0.153 | -0.917*** | 0.463*** | -0.165 |
| | (0.550) | (0.143) | (0.114) | (0.0680) | (0.101) |
| Flint x (2014q2-2015q4) | 3.267*** | -0.765** | -0.908*** | 0.287*** | -0.0364 |
| | (0.626) | (0.281) | (0.108) | (0.0786) | (0.0840) |
| Mean of dependent variable | 102.2 | 17.8 | 5.5 | 4.1 | 3.4 |
| Observations | 300 | 300 | 300 | 300 | 300 |
| *Panel B: Mortality rates* | | | | | |
| Flint x post 2014q2 | 61.05* | 24.84 | -10.52 | 5.830** | 1.251 |
| | (30.96) | (13.84) | (9.801) | (2.021) | (1.482) |
| Mean of dependent variable | 1191 | 485.6 | 237.1 | 40.5 | 23.4 |
| Observations | 224 | 224 | 224 | 224 | 224 |

*Notes:* This table presents the difference-in-differences results using Medicare (2012-2016) and Mortality files data (2010-2017), where the dependent variables are hospitalization and mortality rates by different causes. Regressions include city and quarter-year fixed effects. Robust standard errors, clustered at the city level, in parenthesis. * p<0.10, ** p<0.05, *** p<0.01.

Table 2.3: Cities receiving positive weights for the synthetic control groups

| | All-cause | Cardiovascular | Ischemic | Kidney | Pneumonia |
|---|---|---|---|---|---|
| *Panel A: Hospitalization rates* | | | | | |
| | Lansing | Warren | Detroit | Detroit | Detroit |
| | Detroit | Detroit | Dearborn | Dearborn | Dearborn |
| | Dearborn | | | | |
| | Livonia | | | | |
| *Panel B: Mortality rates* | | | | | |
| | Ann Arbor | Lansing | Lansing | Sterling Heights | Lansing |
| | Detroit | Ann Arbor | Ann Arbor | Detroit | Ann Arbor |
| | | Detroit | Detroit | | Detroit |

Table 2.4: Effect of Flint crisis on hospitalization rates, 2014q2-2016q4

| | 2014q2 | 2014q3 | 2014q4 | 2015q1 | 2015q2 | 2015q3 | 2015q4 | 2016q1 | 2016q2 | 2016q3 | 2016q4 | Average effect | Ratio RMSPE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| All-cause | 8.528 | 3.575 | -0.543 | 11.768*** | 3.962 | 6.822 | 1.934 | -7.391 | 0.519 | 1.330 | -3.119 | 2.489 | |
| *p*-value | 0.214 | 0.428 | 0.928 | 0.000 | 0.357 | 0.214 | 0.500 | 0.357 | 0.928 | 0.928 | 0.571 | | 0.285 |
| Cardiovascular | 0.353 | -0.281 | 0.521 | 0.989 | 0.624 | 0.325 | -0.649 | -1.305 | -1.236 | -0.343 | -0.388 | -0.126 | |
| *p*-value | 0.928 | 0.857 | 0.928 | 0.571 | 0.785 | 0.928 | 0.928 | 0.571 | 0.714 | 0.857 | 0.928 | | 1.000 |
| Ischemic | -0.236 | 0.057 | -1.161 | 0.220 | -0.263 | -0.163 | 0.141 | -0.426 | 0.037 | 0.215 | 1.412 | -0.015 | |
| *p*-value | 0.571 | 0.928 | 0.428 | 0.714 | 0.714 | 0.785 | 0.714 | 0.571 | 1.000 | 1.000 | 0.071 | | 0.857 |
| Kidney | 1.261 | 1.089 | 1.169 | 1.458 | 1.455 | 0.304 | 0.526 | 0.318 | 1.656*** | 0.553 | 0.450 | 0.931 | |
| *p*-value | 0.142 | 0.357 | 0.285 | 0.214 | 0.214 | 0.714 | 0.642 | 0.785 | 0.000 | 0.642 | 0.714 | | 0.428 |
| Pneumonia | 0.630 | 0.362 | 0.867 | 1.562 | 0.197 | 0.389 | -0.213 | 0.647 | 1.007 | -0.500 | 0.235 | 0.471 | |
| *p*-value | 0.142 | 0.500 | 0.285 | 0.285 | 0.714 | 0.285 | 0.785 | 0.428 | 0.142 | 0.500 | 0.714 | | 0.500 |

*Notes:* This table shows in each column the magnitude of the difference between Flint and synthetic Flint for different hospitalization rates across all quarters of the post-intervention period, with their corresponding standardized *p*-value derived from equation (2.7). Additionally, the last two columns show the average affect across post-intervention quarters and the *p*-value of the ratio of post/pre-treatment RMSPE respectively. * p<0.10, ** p<0.05, *** p<0.01.

Table 2.5: Effect of Flint crisis on mortality rates, 2014q2-2016q4

| | 2014q2 | 2014q3 | 2014q4 | 2015q1 | 2015q2 | 2015q3 | 2015q4 |
|---|---|---|---|---|---|---|---|
| All-cause | 130.355 | 264.590*** | 126.718 | 155.872*** | 261.908*** | 196.483*** | 10.308 |
| *p*-value | 0.333 | 0.000 | 0.500 | 0.000 | 0.000 | 0.000 | 0.833 |
| Cardiovascular | -7.550 | 87.829 | 105.980 | 70.151 | 83.953 | 120.074*** | -16.030 |
| *p*-value | 1.000 | 0.333 | 0.166 | 0.500 | 0.333 | 0.000 | 0.833 |
| Ischemic | -21.886 | -25.808 | -9.444 | 61.429 | -11.790 | 37.168 | -82.803*** |
| *p*-value | 0.666 | 0.500 | 0.833 | 0.500 | 0.500 | 0.666 | 0.000 |
| Kidney | 38.255 | 0.638 | -23.360 | 18.497 | -1.455 | 7.726 | 39.372 |
| *p*-value | 0.166 | 0.833 | 0.166 | 0.666 | 1.000 | 0.666 | 0.166 |
| Pneumonia | -4.349 | 22.208 | 0.354 | 18.224 | 14.685 | -8.447 | -16.715 |
| *p*-value | 1.000 | 0.166 | 1.000 | 0.166 | 0.333 | 0.333 | 0.666 |

*Notes:* Continues in next page.

Effect of Flint crisis on mortality rates, 2014q2-2016q4 (cont.)

| | 2016q1 | 2016q2 | 2016q3 | 2016q4 | 2017q1 | 2017q2 | 2017q3 | 2017q4 | Average Effect | Ratio RMSPE |
|---|---|---|---|---|---|---|---|---|---|---|
| All-cause | 146.095*** | 194.414*** | 151.831 | -65.433 | 92.301 | 119.440 | -48.906 | -38.088 | 113.192 | |
| *p*-value | 0.000 | 0.000 | 0.166 | 0.500 | 0.333 | 0.333 | 0.666 | 0.666 | | 0.000 |
| Cardiovascular | 22.796 | 128.010*** | 187.357*** | -12.899 | 50.619 | 3.194 | -37.627 | -101.225*** | 45.642 | |
| *p*-value | 0.666 | 0.000 | 0.000 | 1.000 | 0.333 | 0.833 | 0.500 | 0.000 | | 0.166 |
| Ischemic | -33.799 | -30.095 | 72.466*** | -2.177 | -35.824 | -29.852 | -10.657 | -39.347 | -10.828 | |
| *p*-value | 0.666 | 0.666 | 0.000 | 1.000 | 0.666 | 0.500 | 1.000 | 0.333 | | 0.666 |
| Kidney | 17.693 | 24.190*** | 25.363 | 16.055 | 27.824*** | -15.587 | 14.474 | 12.885 | 13.504 | |
| *p*-value | 0.166 | 0.000 | 0.166 | 0.500 | 0.000 | 0.666 | 0.500 | 0.833 | | 0.500 |
| Pneumonia | -23.839 | -16.695 | -8.699 | 8.489 | 2.831 | 12.663 | -15.785 | -9.035 | -1.607 | |
| *p*-value | 0.166 | 1.000 | 0.166 | 0.333 | 1.000 | 1.000 | 0.333 | 0.833 | | 0.666 |

*Notes:* This table shows in each column the magnitude of the difference between Flint and synthetic Flint for different hospitalization rates across all quarters of the post-intervention period, with their corresponding standardized *p*-value derived from equation (2.7). Additionally, the last two columns show the average affect across post-intervention quarters and the *p*-value of the ratio of post/pre-treatment RMSPE respectively. * p<0.10, ** p<0.05, *** p<0.01.

## BIBLIOGRAPHY


AARP (2016). City of flint senior needs assessment: How does the flint water crisis impact its older citizens? July, 2016.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, 105(490), 493–505.

Abadie, A. & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, 93(1), 113–132.

Abouk, R. & Adams, S. (2018). Birth outcomes in flint in the early stages of the water crisis. *Journal of public health policy*, 39(1), 68–85.

Aizer, A. & Currie, J. (2019). Lead and juvenile delinquency: New evidence from linked birth, school and juvenile detention records. *Review of Economics and Statistics*, 101(4), 645–657.

Aizer, A., Currie, J., Simon, P., & Vivier, P. (2018). Do low levels of blood lead reduce children's future test scores? *American Economic Journal: Applied Economics*, 10(1), 307–41.

Alexander, D. (2015). Does physician pay affect procedure choice and patient health? evidence from medicaid c-section use.

Baicker, K. & Chandra, A. (2004). Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Affairs*, (pp. W184).

Bellware, K. (2019). There may have been dozens more deaths linked to the flint water crisis than previously known. *The Washington Post*, September 12, 2019.

Bernstein, L. & Dennis, B. (2016). Did flint's contaminated water cause deadly legionnaires' outbreaks? *The Washington Post*, February 27, 2016.

Canfield, R. L., Henderson Jr, C. R., Cory-Slechta, D. A., Cox, C., Jusko, T. A., & Lanphear, B. P. (2003). Intellectual impairment in children with blood lead concentrations below 10 $\mu$g per deciliter. *New England journal of medicine*, 348(16), 1517–1526.

Chandramouli, K., Steer, C. D., Ellis, M., & Emond, A. M. (2009). Effects of early childhood lead exposure on academic performance and behaviour of school age children. *Archives of Disease in Childhood*, 94(11), 844–848.

Chang, C.-H., Stukel, T. A., Flood, A. B., & Goodman, D. C. (2011). Primary care physician workforce and medicare beneficiaries' health outcomes. *Jama*, 305(20), 2096–2104.

Childress, S. (2019). We found dozens of uncounted deaths during the flint water crisis. here's how. *PBS Frontline*, September 10, 2019.

Christensen, P., Keiser, D., & Lade, G. (2019). Economic effects of environmental crises: Evidence from flint, michigan. Unpublished Manuscript.

Clemens, J. & Gottlieb, J. D. (2014). Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, 104(4), 1320–49.

Coey, D. et al. (2013). *Physician incentives and treatment choices in heart attack management*. Technical report.

Corasaniti, N., Kilgannon, C., & Schwartz, J. (2019). Lead crisis in newark grows, as bottled water distribution is bungled. *The New York Times*, August 14, 2019.

Dranove, D. & Wehner, P. (1994). Physician-induced demand for childbirths. *Journal of Health Economics*, 13(1), 61–73.

Dummit, L. A. (2009). Relative value units (rvus). In *The Basics Natl Health Policy Forum*.

Ekong, E., Jaar, B., & Weaver, V. M. (2006). Lead-related nephrotoxicity: a review of the epidemiologic evidence. *Kidney international*, 70(12), 2074–2084.

Ferrie, J. P., Rolf, K., & Troesken, W. (2012). Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence test scores among world war two us army enlistees. *Economics & Human Biology*, 10(1), 98–111.

Fleming, L. N. (2018). Expert: Flint plant not ready before water switch. *The Detroit News*, February 5, 2018.

Fonger, R. (2014a). State says flint river water meets all standards but more than twice the hardness of lake water. *MLive Michigan*, May 23, 2014.

Fonger, R. (2014b). City adding more lime to flint river water as resident complaints pour in. *MLive Michigan*, June 12, 2014.

Fonger, R. (2014c). Flint flushes out latest water contamination, but repeat boil advisories show system is vulnerable. *MLive Michigan*, September 14, 2014.

Fonger, R. (2014d). General motors shutting off flint river water at engine plant over corrosion worries. *MLive Michigan*, October 13, 2014.

Ford, M. (2016). A legionnaires' disease outbreak in flint. *The Atlantic*, January 13, 2016.

Galiani, S. & Quistorff, B. (2017). The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal*, 17(4), 834–849.

Glenza, J. (2018). Flint crisis, four years on: what little trust is left continues to wash away. *The Guardian*, April 25, 2018.

Grant, D. (2009). Physician financial incentives and cesarean delivery: new conclusions from the healthcare cost and utilization project. *Journal of health economics*, 28(1), 244–250.

Grossman, D. S. & Slusky, D. J. (2019). The impact of the flint water crisis on fertility. *Demography*, Forthcoming.

Gruber, J., Kim, J., & Mayzlin, D. (1999). Physician fees and procedure intensity: the case of cesarean delivery. *Journal of health economics*, 18(4), 473–490.

Hadley, J., Mitchell, J. M., & Mandelblatt, J. (2001). Medicare fees and small area variations in breast-conserving surgery among elderly women. *Medical Care Research and Review*, 58(3), 334–360.

Hadley, J., Reschovsky, J., Corey, C., & Zuckerman, S. (2009). Medicare fees and the volume of physicians' services. *Inquiry*, (pp. 372–390).

Hanna-Attisha, M., LaChance, J., Sadler, R. C., & Champney Schnepp, A. (2016). Elevated blood lead levels in children associated with the flint drinking water crisis: a spatial analysis of risk and public health response. *American journal of public health*, 106(2), 283–290.

Hulett, S. (2015). High lead levels in michigan kids after city switches water source. *NPR*, September 29, 2015.

Jain, N. B., Potula, V., Schwartz, J., Vokonas, P. S., Sparrow, D., Wright, R. O., Nie, H., & Hu, H. (2007). Lead levels and ischemic heart disease in a prospective study of middle-aged and elderly men: the va normative aging study. *Environmental health perspectives*, 115(6), 871–875.

Jenkins, D. & Danagoulian, S. (2019). Maternal health and pregnancy exposure to lead: A case study of flint, michigan. Unpublished Manuscript.

Keeler, E. B. & Fok, T. (1996). Equalizing physician fees had little effect on cesarean rates. *Medical Care Research and Review*, 53(4), 465–471.

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., Canfield, R. L., Dietrich, K. N., Bornschein, R., Greene, T., et al. (2005). Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. *Environmental health perspectives*, 113(7), 894–899.

Lanphear, B. P., Rauch, S., Auinger, P., Allen, R. W., & Hornung, R. W. (2018). Low-level lead exposure and mortality in us adults: a population-based cohort study. *The Lancet Public Health*, 3(4), e177–e184.

Lustberg, M. & Silbergeld, E. (2002). Blood lead levels and mortality. *Archives of internal medicine*, 162(21), 2443–2449.

Maher, K. (2016). Flint's water woes make residents feel like the walking dead. *The Wall Street Journal*, January 21, 2016.

Medicare Payment Advisory Commission (MEDPAC) (2017). Physician and other health professional services.

Menke, A., Muntner, P., Batuman, V., Silbergeld, E. K., & Guallar, E. (2006). Blood lead below 0.48 mmol/l (10 mg/dl) and mortality among us adults. *Circulation*, 114(13), 1388–1394.

Muntner, P., He, J., Vupputuri, S., Coresh, J., & Batuman, V. (2003). Blood lead and chronic kidney disease in the general united states population: results from nhanes iii. *Kidney international*, 63(3), 1044–1050.

Navas-Acien, A., Guallar, E., Silbergeld, E. K., & Rothenberg, S. J. (2006). Lead exposure and cardiovascular disease—a systematic review. *Environmental health perspectives*, 115(3), 472–482.

Nigg, J. T., Nikolas, M., Mark Knottnerus, G., Cavanagh, K., & Friderici, K. (2010). Confirmation and extension of association of blood lead with attention-deficit/hyperactivity disorder (adhd) and adhd symptom domains at population-typical exposure levels. *Journal of Child Psychology and Psychiatry*, 51(1), 58–65.

Rabinowitz, H. K., Diamond, J. J., Markham, F. W., & Paynter, N. P. (2001). Critical factors for designing programs to increase the supply and retention of rural primary care physicians. *Jama*, 286(9), 1041–1048.

Rau, T., Urzúa, S., & Reyes, L. (2015). Early exposure to hazardous waste and academic achievement: evidence from a case of environmental negligence. *Journal of the Association of Environmental and Resource Economists*, 2(4), 527–563.

Reyes, J. W. (2007). Environmental policy as social policy? the impact of childhood lead exposure on crime. *The BE Journal of Economic Analysis & Policy*, 7(1).

Reyes, J. W. (2015). Lead exposure and behavior: effects on antisocial and risky behavior among children and adolescents. *Economic Inquiry*, 53(3), 1580–1605.

Schober, S. E., Mirel, L. B., Graubard, B. I., Brody, D. J., & Flegal, K. M. (2006). Blood lead levels and death from all causes, cardiovascular disease, and cancer: results from the nhanes iii mortality study. *Environmental health perspectives*, 114(10), 1538–1541.

Schwake, D. O., Garner, E., Strom, O. R., Pruden, A., & Edwards, M. A. (2016). Legionella dna markers in tap water coincident with a spike in legionnaires' disease in flint, mi. *Environmental Science & Technology Letters*, 3(9), 311–315.

Schwartz, J. & Wines, M. (2016). Unsafe lead levels in tap water not limited to flint. *The New York Times*, February 8, 2016.

Semuels, A. (2015). Aging pipes are poisoning america's tap water. *The Atlantic*, July 29, 2015.

Shugarman, L. R. & Farley, D. O. (2003). Shortcomings in medicare bonus payments for physicians in underserved areas. *Health Affairs*, 22(4), 173–178.

Shugarman, L. R., Farley, D. O., Taylor, P., & Ashwood, J. S. (2001). *Trends in Bonus Payments for Physician Services to Rural Medicare Beneficiaries*. Technical report, RAND CORP SANTA MONICA CA.

Starfield, B. (1998). *Primary Care: Balancing Health Needs, Services, and Technology*. Oxford University Press, USA.

Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *The milbank quarterly*, 83(3), 457–502.

Taddonio, P. (2019). Flint's deadly water: 8 key takeaways from frontline's investigation of the flint water crisis. *PBS*, September 11, 2019.

U.S. Department of Health and Human Services, Health Resources and Services Administration, National Center for Health Workforce Analysis (2013). National and regional projections of supply and demand for primary care practitioners: 2013-2025.

Vaziri, N. D. (2008). Mechanisms of lead-induced hypertension and cardiovascular disease. *American Journal of Physiology-Heart and Circulatory Physiology*, 295(2), H454–H465.

Vig, E. K. & Hu, H. (2000). Lead toxicity in older adults. *Journal of the American Geriatrics Society*, 48(11), 1501–1506.

Wasserman, G. A., Liu, X., Lolacono, N. J., Factor-Litvak, P., Kline, J. K., Popovac, D., Morina, N., Musabegovic, A., Vrenezi, N., Capuni-Paracka, S., et al. (1997). Lead exposure and intelligence in 7-year-old children: the yugoslavia prospective study. *Environmental health perspectives*, 105(9), 956–962.

Weisskopf, M. G., Proctor, S. P., Wright, R. O., Schwartz, J., Spiro III, A., Sparrow, D., Nie, H., & Hu, H. (2007). Cumulative lead exposure and cognitive performance among elderly men. *Epidemiology*, 18(1), 59–66.

Weisskopf, M. G., Wright, R. O., Schwartz, J., Spiro III, A., Sparrow, D., Aro, A., & Hu, H. (2004). Cumulative lead exposure and prospective change in cognition among elderly men: the va normative aging study. *American journal of epidemiology*, 160(12), 1184–1193.

Yip, W. C. (1998). Physician response to medicare fee reductions: changes in the volume of coronary artery bypass graft (cabg) surgeries in the medicare and private sectors. *Journal of health economics*, 17(6), 675–699.

Zahran, S., McElmurry, S. P., Kilgore, P. E., Mushinski, D., Press, J., Love, N. G., Sadler, R. C., & Swanson, M. S. (2018). Assessment of the legionnaires' disease outbreak in flint, michigan. *Proceedings of the National Academy of Sciences*, 115(8), E1730–E1739.

Zahran, S., McElmurry, S. P., & Sadler, R. C. (2017). Four phases of the flint water crisis: Evidence from blood lead levels in children. *Environmental research*, 157, 160–172.

Zerehi, M. R. (2008). *How is a Shortage of Primary Care Physicians Affecting the Quality and Cost of Medical Care?: A Comprehensive Evidence Review*. American College of Physicians.