The Impact of the Flint Water Crisis on Fertility

Daniel S. Grossman
David J.G. Slusky

June 10, 2018

Abstract

Flint changed its public water source in April 2014, increasing lead and other contaminant exposure. We compare the change in the fertility rate and in health at birth in Flint before and after the water switch to the change in the other cities in Michigan. We find fertility rates decreased by 12%, and overall health at birth decreased. This effect on health at birth is a function of two countervailing mechanisms: negative selection of less healthy embryos and fetuses not surviving (raising the average health of survivors) and those that survived being scarred (decreasing average health). We untangle this and find a net of selection scarring effect of 5.4% decrease in birth weight. These effects are likely lower bounds on the overall effects of this exposure due to long-term effects of in utero exposure.

Keywords: Women's Health; Birth Rate; Fertility Rate; Birth Outcomes; Lead; Environmental Regulation

JEL Codes: H75, I12, I18, J13, Q53, Q58

Affiliations:

Grossman (corresponding author): Department of Economics, West Virginia University, 1601 University Ave., 411 College of Business and Economics, Morgantown, WV 26506-6025; daniel.grossman@mail.wvu.edu

Slusky: Department of Economics, University of Kansas, 1460 Jayhawk Blvd., 415 Snow Hall, Lawrence, KS, 66045, david.slusky@ku.edu

Acknowledgements:

We thank Vincent Francisco, Kate Lorenz, Matt Neidell, Dhaval Dave, Dietrich Earnhart, Josh Gottlieb, Ben Hansen, Shooshan Danagoulian, Scott Cunningham, Edson Severnini, David Keiser, Peter Christensen, Charles Pierce, Nicolas Ziebarth, Nigel Paneth, Tom Vogl, Nick Papageorge, Karen Clay, Bryce Steinberg, Ken Chay, Osea Giuntella, Werner Troesken, Emily Rauscher, Donna Ginther, Jarron Saint Onge, seminar participants at the University of Minnesota, University of North Carolina, Appalachian State University, Johns Hopkins University, University of Pittsburgh, Brown University, and the University of Kansas Medical Center, and other conference participants at the 2017 iHEA conference, the 2017 National Bureau of Economic Research Summer Institute, and the 2017 APPAM Conference for all of their suggestions and feedback. We also thank Glenn Copeland of the Michigan Department of Community Health, Vital Records and Health Statistics Division for providing vital statistics data, the University of Michigan-Flint GIS Center for sharing data, David Powell for sharing his imperfect synthetic control method code, the West Virginia University Center for Free Enterprise for financial support, the Big XII Faculty Research Fellow Program, and the staff at KU IT for managing our research server.

1. Introduction

Overwhelming evidence exists that lead in water contributes to higher rates of lead in the blood, and is related to eventual developmental problems in children (Edwards, Triantafyllidou, and Best 2009; Hanna-Attisha et al. 2016). Despite this, lead problems are well documented in many cities and EPA plans to cut back on testing for lead and other water pollutants and lead renovation, repair, and painting (Davis 2017).

Lead also may effect health through indirect channels by decreasing latent health of those infants carried to term. Latent health is difficult to measure and may not manifest until much later in life, as demonstrated interdisciplinarily in the literatures of epidemiology (Barker 1995), biology (Schultz 2010), and economics (Almond and Currie 2011). Exposure to lead in utero and in infancy may only represent a lower bound on the overall effect of lead on health and human capital development.

We study the effect of the higher lead content of water sourced from the Flint River on fertility and birth outcomes. Importantly, during the period in which water was sourced from the Flint River (beginning on April 25, 2014), local and state officials continually reassured residents that the water was safe. Officials did not issue a lead advisory until September 2015, just a few weeks before switching off Flint River water for good (Fonger 2015a). This reduced the scope of an avoidance behavioral response to the water crisis (see e.g. Neidell 2009).

High lead content in the blood affects nearly all organ systems and is associated with cardiovascular problems, high blood pressure, and developmental impairment affecting sexual maturity and the nervous system (Agency for Toxic Substances and Disease Registry; Zhu et al. 2010). Recent studies have linked maternal lead exposure to fetal death, prenatal growth abnormalities, reduced gestational period, and reduced birth weight (Edwards 2014; Zhu et al.

2010; Taylor, Golding, and Emond 2014); while historically lead is associated with increased fetal death and infant mortality rates (Troesken 2008; Clay, Troesken, and Haines 2014), and the poisoning of many adults as well (Troesken 2006a). Maternal lead crosses the placenta providing a potential direct link for lead poisoning of the fetus (Taylor, Golding, and Emond 2014; Lin et al. 1998). Lead exposure also decreases sperm count and male fecundity (Paul 1860; Vigeh et al. 2011).

The Flint water supply also contained higher rates of trihalomethanes, among other contaminants. Previous work has suggested that trihalomethanes may be detrimental to pregnant women (Gallagher et al. 1998; Nieuwenheijsen et al. 2000; Cao et al. 2016), although others dispute this association (e.g. Yang et al. 2007; Horton et al. 2011). The water change likely led to a Legionnaires disease outbreak in Flint that killed as many as 12 individuals (Rhoads et al. 2017). While we cannot separately identify the effects of these contaminants, we focus mostly on lead because of the large literature linking lead to poor pregnancy outcomes and the specific results from Flint showing elevated lead levels (Hanna-Attisha et al. 2016).

We leverage the fact that only the city of Flint switched its water source at this time, while the rest of Michigan did not. These areas provide a natural control group for Flint in that they are economically similar areas and otherwise followed similar trends in fertility and birth outcomes over this time period.

Using the universe of live births in Michigan from 2008 to 2015, we estimate the effect of a change in the water supply on fertility and health. Following the water change, the general fertility rate (GFR) in Flint decreased by 7.5 live births per 1,000 women aged 15-49 compared to control women of the same age group, a 12.0% decrease. Because the higher lead content of the new water supply was unknown at the time, this decrease in GFR is likely a reflection of an increase in fetal deaths and miscarriages and not a behavior change in sexual behavior related to conception like contraceptive use. We present suggestive evidence that behavioral changes are unlikely to explain our results. Additionally, the ratio of male to female live births decreases by 0.9 percentage points in Flint compared to surrounding areas.

Estimates of birth outcomes are less precise and at times contradictory. Because of the large decrease in fertility rates, selection into birth is a major concern for our birth outcome results. To account for this selected sample, we perform a bounding exercise which provides an upward bound on the birth weight effect caused by the water change (Bozzoli, Deaton, and Quintana-Domeque 2009). We find a nearly 5% decrease in birth weight. This estimate of the selection-scarring effect of in utero exposure to a contaminated water source is a contribution to both the fetal origin literature and the health effects of lead literature.

Futhermore, our paper contributes to a growing interdisciplinary literature on the consequences of the Flint water switch. Others show increased lead levels (Hanna-Attisha et al. 2016; Zahran et al. 2017), diminished test-scores (Sauve-Syed 2017), and some avoidance behavior through increased bottled water purchases (Christensen, Keiser, and Lade 2017). Abouk and Adams (2018) find worse health at birth for white mothers in Flint. We are the first to investigate the impact on fertility rates.

Because of the potential long term effects of lead on cognitive development (e.g., Aizer et al. 2018), we cannot make any definitive statement about whether babies born have a higher future health stock compared to control cohorts or if latent health for this group is actually worse. We can however estimate the selection effect by focusing on the birth rate, and investigate infant health of the surviving children to estimate the magnitute of the offsetting scarring effect on survivors.

2. Background on Flint[1]

Flint is an old manufacturing city and the birthplace of General Motors (GM) (Scorsone and Bateson 2011). The city has been shedding residents for many years; its contraction coinciding with GM closing several plants in and around Flint.[2]

In 1967, Flint switched its water source away from the Flint River because of concerns about serving a growing population (Carmody 2016), and signed a deal to receive Lake Huron water via pipeline from the Detroit Water and Sewerage Department (DWSD). In 2011 the Governor of Michigan installed an Emergency Manager in the city who would make all fiscal decisions and "rule local government," based on the city's precarious economic health (Longley 2011), changing the political economy of Flint. Citizens and elected officials would have little recourse to fight decisions made by the Emergency Manager. Concurrently, DWSD water rates were rising (Zahran, McElmurry, and Sadler 2017). To cut costs, the Emergency Manager together with other Genesee County officials began to build a pipeline directly to Lake Huron in March 2013 (City of Flint 2015; Walsh 2014). However, the project would take more than two years to complete. In the interim, Flint decided to use water from the Flint River to source its drinking water between April 2014 and the completion of the new pipeline, while Genesee County continued to receive water from DWSD (Carmody 2016).

Flint had to treat the new water source and while they used some of the same products as the DWSD, they did not use anti-corrosive inhibitors such as orthophosphate (Pieper et al. 2017; Olson et al. 2017). Flint citizens began complaining about the color and smell of their water but

[1] Appendix Figure B1 provides a timeline of events around the water change.

[2] The number of inhabitants employed by GM decreased from 80,000 in 1978 to 30,000 in the 1990s to under 10,000 today (Scorsone and Bateson 2011).

were continually assured that the water was safe to drink (City of Flint 2015a,b). The first sign of trouble came in August 2014 when a boil advisory was announced for part of the city due to a positive fecal coliform test, although the city minimized this adverse result claiming it was an "abnormal test" caused by a "sampling error" (Fonger 2014a; Adams 2014). Less than a month later a second boil adivisory was announced for a similar issue leading the city to increase chlorine levels in the water (Fonger 2014b). Then in October 2014, GM announced they would switch off of Flint River water in its Flint plant because the water was too corrosive for its engine parts (Fonger 2014c). The city confirmed the GM switch was best for engine parts but that the water was safe for human consumption. In December 2014, Flint received an EPA violation for excess trihalomethanes (TTHM) in the water, likely caused by the chemicals used to treat the water (Fonger 2015b).[3]

Throughout early 2015, Flint held public meetings to assure citizens the water was safe and that the TTHM violation would be fixed soon (City of Flint 2015a,b). Concurrently, the Emergency Manager commissioned a report on the safety of the water and rebuffed an offer from DWSD to return Flint to Lake Huron water. A team from Virginia Tech, led by Mark Edwards began independently testing Flint consumers' water and in August 2015 reported much higher levels of lead than previously reported, noting that Flint River water was many times more corrosive than the DWSD water (http://flintwaterstudy.org/wp-content/uploads/2015/10/Flint-Corrosion-Presentation-final.pdf). Mona Hanna-Attish, a Flint pediatrician and researcher, held a press conference September 24, 2015 to report a substantial

[3] Because of these additional contaminants found in Flint drinking water after the switch, we cannot attribute all health effects to lead. However, because of the well-established pathways through which lead affects fetal health, we focus on lead. To the extent that these other contaminants are present in the water, our estimates would be an upper bound on the effect of lead on fertility and birth outcomes.

increase in blood lead levels in children following the water switch (Fonger 2015c; Hanna-Attish et al. 2016). While the city initially attacked the results of this study, the resulting public outcry finally led the city to switch back to Lake Huron water on October 16, 2015 (Emery 2015). The crisis continues as those exposed to lead face potential life-long problems.[4]

3. Literature Review

3.1. Background on Lead

Lead is a naturally occuring heavy metal that is associated with health problems. Human activities, including burning fossil fuels and industrial chemical reactions, cause the majority of lead emission into the environment (Agency for Toxic Substances and Disease Registry (ATSDR) 2007). The US dramatically decreased the incidence of lead emissions and blood lead levels by banning lead paint in the 1970s and reducing leaded-gasoline throughout the 1980s before banning it in 1996 (CDC 2005, Zhu et al. 2010).

Previous work has investigated the effects of general exposure to lead, lead levels in the blood, lead exposure from a water source, and the mechanisms through which lead and other in utero exposure effects current and future health. We discuss each below.

3.2.Exposure to lead

Exposure to lead is associated with cardiovascular problems, high blood pressure, and developmental impairment affecting sexual maturity and the nervous system (ATSDR 2007; Zhu et al. 2010). Lead crosses the placenta (Amaral et al. 2010, Schell et al. 2003, Rudge et al. 2009, Lin et al. 1998) and is correlated with mental health issues, prenatal growth abnormalities, reduced gestational period, spontaneous abortion, and reduced birth weight (Borja-Aburto et al.

---

[4] https://www.washingtonpost.com/national/grant-to-create-registry-of-flint-residents-exposed-to-lead/2017/08/01/78c8fa66-7707-11e7-8c17-533c52b2f014_story.html?utm_term=.f89d3f64e273

1999; Hertz-Picciotto 2000; Joffe et al. 2003; Bellinger 2005; Hu et al. 2006; Cleveland et al. 2008; Vigeh et al. 2010; Zhu et al. 2010; Taylor, Golding, and Emond 2014). Clay, Portnykh, and Severnini (2018), using variation in lead exposure from the introduction of the Interstate Highway System and the Clean Air Act, find that exposure to lead in the air resulted in reductions in the birth rate and a worsening of birth outcomes. Additionally, men exposed to lead, including as workers in industrial settings, have lower fecundity (Paul 1860; Hamilton and Hardy 1983; Assennato et al. 1987; Coste et al. 1991; Winder 1993; Alexander et al. 1996; Lin et al. 1996; Bonde and Kolstad 1997; Apostoli, Porru, and Bisanti 1999; Apostoli et al. 2000; Hernberg 2000; Sallmén, Lindbohm, and Nurimnen 2000; Sallmén 2001; Shaiau, Wang, and Chen 2004; Wirth and Mijal 2010; Vigeh et al. 2011; Wu et al. 2012; Eibensteiner 2013).

3.3. Lead Effects in Water

High lead content in water leads to increases in lead content in the blood (Troesken and Beeson 2003; Edwards, Triantafyllidou, and Best 2009; Hanna-Attischa et al. 2016), increasing the risk of the negative health outcomes detailed above. Clay, Troesken, and Haines (2014) find historical evidence of higher rates of fetal deaths in cities with more lead service pipes and more acidic water. Areas with higher water lead levels have higher rates of preeclampsia (Troesken 2006b). Fetal death rates increased and birth rates decreased following the increase of lead in the water in Washington, DC from 2000 to 2003 (Edwards 2014). While our paper is similar to that of Edwards (2014), we use a substantially larger group of comparison cities to perform inference. That study uses only Washington, DC compared to overall US and Baltimore, MD. This comparison of just 2 areas makes proper inference difficult due to small clusters in both treatment and control cities (see e.g. Cameron, Gelbach, and Miller 2008).

While previous studies have used exact measures of lead in the blood (see e.g. Taylor, Golding, and Emond 2014; Zhu et al. 2010), these study designs do not include exogenous variation in lead supply and thus cannot rule out that these worse birth outcomes are actually associated with an omitted variable (or some other environmental factor that is associated with both birth outcomes and lead concentration).

Lead increased in the Flint water supply because of improper water treatment. Officials did not treat the Flint River water using corrosion inhibitors, while simultaneously using ferric chloride (to combat infectious bacteria in the water) which increased the likelihood of corrosion (Clark et al. 2015; Pieper, Tang, and Edwards 2017). Corrosion inhibitors aid in creating protective corrosion scales within pipes, reducing the amoung of lead leached from the pipes (Pieper, Tang, and Edwards 2017; Olson et al. 2017).

3.4. Other outcomes

Previous studies have found that increases in lead levels have a perverse effect on later life cognitive function (Hernberg 2000; Ferrie, Rolf, and Troesken 2012; Reuben et al. 2017), mental health and criminality (Reyes 2007; 2015), educational outcomes (Aizer et al. 2018), and school suspensions (Aizer and Currie 2017; Billings and Schnepel 2018). However, Billings and Schnepel (2018) and Gazze (2016) find that lead remediation can moderate the negative effects of those exposed to lead and reduce blood lead levels.

3.5. Mechanisms

This study contributes to the large literature on fetal origins hypothesis, where utero shocks may affect health. The sign of the effect of these shocks is ambiguous due to two countervailing mechanisms (Almond 2006).

As discussed above, fetal insults may lead to "selective attrition," or the culling of weaker fetuses through miscarriage or fetal death (Edwards 2014; Clay, Troesken, and Haines 2014; Almond 2006). Thus, the less healthy fetuses would not be born, leaving only the healthier fetuses, or a potentially positive effect on population health. Additionally, higher rates of lead may shift the overall health distribution of infants affected in utero towards being more unhealthy, leading to worse health outcomes. The two effects (selection and scarring) could even approximately cancel each other out for surivors (Bozzoli, Deaton, and Quintana-Domeque 2009).[5] Behavioral selection into pregnancy may occur if women decide not to get pregnant because of worries about their future child's health. Dehejia and Lleras-Muney (2004) document non-random selection into pregnancy in response to changing labor market conditions while Clay, Portnykh, and Severnini (2018) provide evidence of more educated women reducing fertility in response to lead exposure. However, women would need to be aware of the water crisis in advance for this explanation to affect our analysis. While women were aware of several issues with Flint water following the change, they had no way of knowing about the lead content in the water until nearly the end of the Flint River water regime.[6]

4. Data

We use vital statistics data for the state of Michigan from 2008 to 2015. These data contain detailed information on every birth in the state including health at birth and background information on the mother and father which includes race, ethnicity, education, marital status, as well as prenatal care during her pregnancy. We calculate the date of conception for a woman

[5] For example, in the Great Chinese Famine, taller children were more likely to survive but then were stunted, resulting in a minimal change in height for the affected cohort but their unscarred children being taller (Gørgens, Meng, and Vaithianathan 2012).

[6] See Appendix Figure B2 for support of this.

from the clinical gestational estimate and exact date of birth. We define Flint per the census tract-level (University of Michigan-Flint GIS Center 2017) data on lead pipes, and then use HUD census tract to ZIP code matching[7] and SAS ZIP code to city matching[8] for the 15 largest non-Flint cities (i.e., Ann Arbor, Dearborn, Detroit, Farmington Hills, Grand Rapids, Kalamazoo, Lansing, Livonia, Rochester Hills, Southfield, Sterling Heights, Troy, Warren, Westland, and Wyoming).[9]

Using population data from the American Community Survey[10], we calculate general fertility rate (GFR) as:

$$GFR_{ct} = 12 * 1000 * \frac{Total\ Births_{ct}}{Population\ Aged\ 15{-}49_{ct}} \tag{1}$$

where c indexes the city, and t the month and year. Total births are the exact number of births occurring in the area for a given conception month, while population is a measure of the female population of childbearing age.[11] We multiply by 12 to make this an annual measure.

5. Methods

To assess the relationship between water source and fertility outcomes, we use a difference-in-differences model to compare areas that received the new source to areas that did not change their water source but were trending similarly in the pre-period. The model takes the

[7] https://www.huduser.gov/portal/datasets/usps_crosswalk.html#data

[8] https://support.sas.com/downloads/download.htm?did=104285#

[9] In Appendix C, we instead compare Flint to counties in Michigan.

[10] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_1YR_S0101&prodType=table – “State 040,” “Place 160”

[11] Our analysis sample covers 95 months from May 2007 through March 2015. Because we use conception date, our 2008-2015 data contains complete date of conception data from approximately May 2007 through March 2015.

following form:

$$Outcome_{ct} = a + \beta_1 Water_{ct} + \alpha_c + \delta_t + \varepsilon_{ct} \qquad (2)$$

Where c indexes the city, and t the month and year. We chose the 15 most populous cities in Michigan (excluding Flint) as comparison cities, as shown in the map in Figure 1. *Outcome* includes measures of *GFR* and male to female sex ratio (sex ratio).[12] *Water* is a binary variable indicating whether the date of conception of the child occurred after the water supply changed and whether the mother lived in Flint. We define being in utero during the new water regime as a birth conceived in November 2013 or later, which would mean that at least one trimester of the pregnancy was affected by the water change.[13] We include city fixed effects, α_c, to control for time-invariant characteristics of the city. δ_t is a vector of month and year fixed effects. City and time fixed effects subsume the main effects of living in Flint and being in utero during the new water regime, respectively. We also investigate these models using subsets of the data based on race, educational attainment, and age of mothers.

For birth outcomes, we estimate the following model:

$$Birthoutcome_{ict} = a + \beta_1 Water_{ct} + \beta_2 X_{ict} + \gamma_{cen} + \delta_t + \varepsilon_{ict} \qquad (3)$$

where i indexes the individual, c the city, and t the month. *Birthoutcome* includes a continuous variable for birth weight in grams, a binary variable for low birth weight, estimated time of gestation in weeks, or fetal growth rate, defined as the birth weight divided by weeks in gestation. *Water* is defined as above. X_{ict} is a vector of variables capturing individual level

---

[12] Our results are robust to using alternative specifications, including the natural log of the count of births and a nonlinear Poisson specification of the count of births. See Appendix Tables B3-B4, and note that the coefficients are in log points, which for this range are approximately numerically the same as percentage points.

[13] We show in Appendix Figure B3 that our results are robust to varying the date of treatment.

socioeconomic characteristics of the mother and child including gender of the child, race, ethnicity, marital status, and educational attainment of the mother, which come from birth records. We include census tract fixed effects, γ_{cen}, to control for time-invariant characteristics of the direct neighborhood of the mother. δ_t is a vector of month and year fixed effects, which control for seasonality of births and a general trend in birth outcomes across Michigan over time. ε is an error term clustered at the city level.

A strength of our study is that it exploits a natural experiment in the exposure of women to contaminants, including lead, caused by an exogenous change in the water supply. Any time a policy shift occurs that potentially causes an exogenous change, one worries about policy endogeneity: the idea that this policy change occurred in response to conditions that were already changing; in response to public pressure; or other additional factors unobservable to the econometrician were present. However, an EPA memo citing lead concerns was leaked to the public only in July, 2015, and not confirmed by other researchers until September, 2015 (Robbins 2016).

6. Results

Table 1 presents summary statistics. Columns (1) and (2) present means of births to individuals who did not reside in Flint before and after the water change, respectively. Descriptive statistics for mothers who lived in Flint at the time of birth before the water change are presented in Column (3) while results for Flint mothers who gave birth after the water change are presented in Column (4). In general, we consider a birth as occurring after the water change if the mother conceived in November 2013 or later.[14]

Mothers who gave birth outside of Flint were older in the pre-period. However, we find no differential change in age between the periods. Women in Flint also had lower educational attainment. They were much more likely not to have a high school diploma and less likely to have obtained a college degree. While the proportion of mothers who did not receive a high school diploma decreased by approximately 2.5 percentage points for both Flint and non-Flint mothers following the water change, Flint mothers were more likely to receive a high school diploma and non-Flint mothers were more likely to complete some college or a college degree.

The GFR in Flint was nearly 8.5 births per 1000 women aged 15-49 lower in Flint following the water change compared to control areas. The sex ratio of babies born in Flint skewed more female following the water change, a decrease of 0.74 percentage points. Babies born in Flint were nearly 150 grams lighter than in other areas, were born ½ a week earlier and gained 5 grams per week less than babies in other areas in the pre-period. The unadjusted difference-in-differences for these variables was a decrease of 15 grams, 0.12 weeks of gestational age and 0.27 grams per week in growth rate.

---

[14] This allows for a mother to be considered "treated" if she lived under the new water regime for at least one trimester of her pregnancy.

6.1. Fertility Results

In Figure 2 we present trends in GFR for Flint and the rest of Michigan separately. We present unadjusted fertility rates.[15] While births in Flint are more volatile due to the smaller base sample in the area, the graph demonstrates a substantial decrease in fertility rates in Flint for births conceived around November 2013, which persisted through the beginning of 2015. Flint switched its water source in April 2014, meaning these births would have been exposed to this new water for a substantial period in utero (i.e., at least one trimester). Other cities in Michigan had similarly seasonally volatile GFR trends, but did not display large decreases in GFR following the Flint water switch.

Table 2 presents regression results for GFR by city. The main coefficient of interest is β_1, the parameter of $Water_{ct}$ calculated using equation (2) above. The unit of observation is city-month. We estimate that women living in Flint following the water change gave birth to 7.5 fewer infants per 1,000 women aged 15-49 compared to control counties. These results are statistically significant at the 0.001 (0.1%) level. This is on a base of 62 births per 1000 women aged 15-49, or a 12.0% decrease in births in Flint. In column 2 we include a more flexible measure of time fixed effects by interacting month into year. Estimates are nearly identical in these more saturated models. We adjust our standard errors using the wild bootstrap method (Cameron, Gelbach, and Miller 2008) because we only have one treated cluster. We include city specific linear time trends in column 3, and the results are statistically significant, but we do not consider them our main results because of concerns about potentially biased treatment group-

[15] We calculated a 13 month moving average (+/- 6 months) to remove both seasonality and idiosyncratic noise in Appendix Figure B4.

specific time trends (Wolfers 2006, Lindo and Packham 2017).[16] We examine how the sex ratio of live births changed in Flint, given the medical literature that male fetuses are more susceptible to fetal insults (Trivers and Willard 1973, Sanders and Stoecker 2015). We find that sex ratios decrease by 0.9 percentage points (1.8 percent) in Flint, compared to other Michigan cities. Sanders and Stoecker (2015), investigating the health effects of the Clean Air Act, find that birth ratios skew more male following the implementation of the act. They argue this is consistent with an increase in health. While this increase in the proportion of births that are female likely represents a level of selection consistent with an increase in fetal deaths, it is also consistent with a decrease in health at the time of birth. Given the concerns of biased treatment group-specific time trends (Wolfers 2006, Lindo and Packham 2017), we are not concern by the results in column 6 being far smaller in magnitude.

Given the large decreases in GFR and the shift towards more female births we find for Flint overall, we investigate whether these results hold for all individuals in Flint. In Table 3, we limit our sample to demographic subgroups. We use demographic characteristics of mothers because this is more accurately measured and available for nearly all births. First, for the sample of mothers 18-44 (the closest group for annual city population estimates by educational attainment),[17] as a whole, we find a decrease of 8.188 which can be interpreted as a 10.6% decrease in births in Flint following the water change. This is slightly smaller than the 12% decrease we find in Table 2 using GFR. Results by education suggest that the decrease in births

---

[16] We sought administrative records on fetal deaths from the state of Michigan. Unfortunately, the state sent us two different data files which show different results, and so we are unable to report whether the fetal death rate increased or decreased. However, using the data set that gives the largest increase in the fetal death rate only explains 2% of the decrease in the fertility rate.

[17] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_1YR_B15001&prodType=table – "State 040," "Place 160"

is concentrated among the more educated mothers, with those with some college decreasing births by 30% and those with a post-secondary degree decreasing births by 20%.[18] However, those with some college behave differently from those with a college degree, suggesting they are not necessarily a convex combination of college graduates and only high school graduates. For example, many of them had to drop out of college due to a financial or health shock or other family emergency, making them more economically disadvantaged than those with no college even though they have more education (Pollak and Lundberg 2014).[19] Furthermore, the college graduate group in Flint is particularly small, less than 200 total births post water change (even including those with associate's degrees) and thus readers should be careful not to ascribe too much weight to these results as small variations may substantially affect our results.

In Table 3 Panel B, we perform a subsample analysis of the effect of the water change on sex ratios. The results by education are consistent with the fertility results above, with large decreases in the ratio of female births for mothers with some college and a college degree. These results suggest that if this is an avoidance story for these mothers, the mothers of this education level who do give birth are a selected sample that differs from similarly educated mothers who gave birth in Flint before the water switch in unobservable ways.

6.2.Birth Outcomes

The results in the section above provide direct support for the Flint water change causing a culling of the weakest fetuses. Next, we turn our focus to birth outcomes. If increased lead in

[18] We use the following categories that have a substantial share of mothers in Flint in the post period: no high school degree (25%), only a high school degree (35%), some post-secondary education but no degree (30%), and any post-secondary degree, including associate (10%).

[19] Unlike with education, we find comparable results for the fertility rate and sex ratio when stratifying by mother's race or age.

the water only has a selective attrition effect then we would expect an increase in health among the births in Flint as the selection would remove only the weakest and leave the healthier fetuses to come to term. If, alternatively, a scarring effect also is present, then we would expect a decrease in health for those births that actually occurred. Lastly, if more educated women were exhibiting avoidance behavior we would expect overall health to decrease due to this selective sample.

We first investigate whether the change in water supply caused a change in birth weight in Table 4.[20] We cluster standard errors in these regressions conservatively at the census tract level. While we find negative results for birth weight, they are imprecisely measured. Adding census tract, month and year of conception fixed effects and additional covariates in columns 2-5 does not substantially change the coefficient.

Results for low birth weight, gestational age and gestational growth rate are all in the direction of worse health, but not statistically significant. The magnitudes on the coefficients are all rather small, with the exception of low birth weight, suggesting non-economically meaningful effect sizes.[21]

6.3. Behavioral Changes

Behavioral changes (i.e., less sex) and not the physiological impacts of lead could be driving our results. Following Barreca, Deschenes, and Guldi (Forthcoming), in Table 5 we use

[20] We also estimate models using abnormal conditions as the dependent variable, but this variable has a substantial amount of missing values for Flint compared to other cities for 2007-2009. We estimate models only for 2010 onwards with abnormal conditions as the dependent variable and find no evidence of changes in abnormal conditions caused by the Flint water change.

[21] Clustering at the city level, rather than the census tract level provides mostly statistically significant estimates. Still, scarring and selection may be negating each other, and so the disentangling below applies.

the American Time Use Survey to investigate time spent engaged in sexual relations, proxied by any time spent in "personal or private activities".[22] Note that these analyses are at the county or CBSA-level and are thus not directly comparable to our main results as Flint comprises approximately ¼ of the population of Genesee County. We find that sexual activity *increased* in the post period, which would bias our main result of a decrease in the fertility rate toward zero.[23] While this is only suggestive evidence, it supports our conclusion that the reduction in the conception rate is not driven by a reduction in sexual activity.[24]

6.4. Synthetic Control Methods

Lastly, we perform an analysis of fertility rates using a synthetic control methods approach (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010).[25] This method creates a weighted control group that more closely resembles the characteristics of Flint in the period before the water change on both level and trend of fertility rates. It also controls for demographic characteristics of mothers in the pre-period, including race/ethnicity, educational attainment, and gender of the child. Figure 3, Panel A displays GFR trends in Flint and its synthetic control group before and after the water switch, which is visualized as the vertical line at November, 2013, the last conception date for which women would have been exposed for at

[22] I.e., "having sex, private activity (unspecified), making out, personal activity (unspecified), cuddling partner in bed, spouse gave me a massage."

[23] This is analogous to Barreca, Deschenes, and Guldi (Forthcoming) which also finds a statistically significant *increase* in time in the probability that individuals spend time on sex during environmental conditions that reduce fertility.

[24] The ATUS only has county/CBSA identifiers. In Appendix Table B1, we repeat our results are the county level and show that while the inclusion of the rest of Genesee County (where Flint is located) as treated reduces the magnitude of our results, they are still directionally consistent and statistically significant in some specifications.

[25] We describe this method in detail in Appendix A.

least one trimester to the new water supply. Panel B shows the difference between each city systematically assigned to treatment and the synthetic version of the city for each month. Flint is denoted by the solid line. The average treatment effect in Flint compared to the synthetic control is a decrease of 11.6 births, presented in Panel C by the horizontal blue bar. This effect size is slightly larger than that found above in Table 2. The average treatment effect in Flint is substantially larger than the average treatment effect for all other cities.[26]

As an additional robustness check, we perform a synthetic control model matching on all GFR for the month of March in each year before the water change (2008, 2009, 2010, 2011, 2012, 2013). The strength of this analysis is that it creates a better pre trend match on GFR, but the weakness is that it may over-fit on GFR and ignore other covariates (see Kaul et al. 2017). [27]

Finally, because our pre-trend match may be imperfect which would affect our inference, we use an *imperfect* synthetic control method (Powell 2017). This method solves two issues in the Abadie et al. (2010) method: it improves inconsistent pre-period match due to transitory shocks by using pre-period outcomes predicted from city-specific flexible time trends, instead of the actual per-period outcomes; and it allows for treated group to be an outlier by using treated group's presence in other group's synthetic controls. More specifically, an outlier treated group cannot match the outcomes of a convex combination (e.g., with nonnegative weights) of the control units. But a convex combination of the treated group and the rest of the control group

[26] We find similar effect sizes and inference interpretations using quarter of birth rather than month of birth, dropping outlier cities (Appendix Figure B5), and dropping Flint from the inference analysis so that when we assign treatment to each control group, Flint cannot be used as a synthetic control.

[27] Our estimates are robust to this alternative specification and we present these results in Appendix Figure B6. We also find similar results matching on the 4th quarter GFR for each year before the water change (2007, 2008, 2009, 2010, 2011, 2012) and using a 13 month moving average for GFR.

could match the outcome of control group that has been assigned a falsification treatment. If the treated group has a positive weight in this situation, that weight can be inverted to describe the mapping from the control group to the treated group.

These two fixes reduce bias in the estimates and out performs the synthetic control method in simulations (Powell 2017). Figure 4, Panel A shows that the *imperfect* synthetic control group is a better match for Flint in the pre-period. Panel B demonstrates that the decrease in GFR in Flint is larger than in any other area following the water change which provides additional evidence of the statistical significance of our estimates.

7. Robustness Checks

We perform a number of robustness checks. First, we perform a randomization inference permutation test (e.g., Cunningham and Shah 2018; Fischer 1935). This test essentially systematically assigns treatment status to each control area and compares the size of the implied treatment effect for each control area to that of the actual treated area. As shown in Appendix Figure B7, we find our effect size in Flint is larger than all control area "treatment effects" which suggests our inference above is indeed correct.

We compare county level GFR rates in Appendix Table B1. The treatment in this table includes all of Genesee County, of which Flint comprises approximately ¼ of the population. The results are greatly reduced in this table, which is to be expected given that the treatment sample is contaminated with non-affected areas. However, GFR still decreases in a statistically significant way in Genesee County compared to other counties in Michigan following the Flint water change.[28]

In Appendix Table B2, we show our results are robust to limiting our sample to GFR of

---

[28] With the exception of the county-specific linear time trend, which biases the results for the reasons described above.

births conceived before September 2014 and dropping the cities with the highest and lowest GFRs. We investigate births conceived before September 2014 because of potential avoidance given boil advisories due to fecal coliform in the Flint water source reported around this time. The decrease in fertility rates was actually larger in this early period, with Flint's GFR nearly 9 births per 1000 women aged 15-49 lower than other cities' GFR following the water change. One potential explanation for this change is that avoidance behaviors led Flint residents to begin buying bottled water at higher rates after September 2014 (Christensen, Keiser, and Lade 2017).[29]

In Appendix Table B3 we estimate the effect of the water change on log births. We find a 15%-18% decrease in Flint following the water change, which is comparable to our 12%-14% result in Table 2. In Appendix Table B4 we estimate a Poisson model and find a decrease in births of 0.15, which can be interpreted as similar to a 15% decrease in births in Flint. These results assume a constant population in Flint over the study period. Estimates of population in Flint decrease over the study period which may partially explain the larger magnitude of the effect in log and Poisson models.

We find consistent results comparing only Flint and the rest of Genesee County (Appendix Table B5). As a falsification analysis, we compare Genesee County, excluding Flint, to the rest of Michigan in Appendix Table B6 and find no change in GFR or sex ratios providing strong support for a change within Flint at the time of the water switch driving our results.

In Appendix C, we focus on Flint compared to counties in Michigan rather than cities. The results are largely robust to this alternative definition of control areas. The main difference

[29] Additionally, our final time period, March 2015 shows a large increase in GFR for Flint, from which we are unable to determine if that month GFR is an outlier or a general trend towards higher GFRs.

between these results and our main city comparison results are that the effect of the water switch on fertility rates in Flint is slightly smaller than our main results, but birth outcome results are slightly larger.

8. Discussion

Our results for the decrease in the fertility rate are plausible given the broader scientific literature on this topic. Specifically, Edwards (2014) studies an increase in lead in drinking water in Washington, D.C. in the early 2000s, and using somewhat different methods finds a 12% decrease in the fertility rate, similar to the magnitude of the fertility rate decrease that we find.

We attempt to extrapolate the consequences of our results. The population of women aged 15-49 in Flint during our study period is approximately 26,000. The GFR dropped from 62 to 57, suggesting that over our study window of 17 months (births conceived from November 2013 through March 2015) between 198 and 276 more children would have been born had Flint not enacted the switch in water.[30] We consider this strong empirical support for the existence of a culling effect caused by increased lead in the water. Our results on sex ratios suggest that among the live births that occurred in Flint following the change in water supply, an additional 18 female infants were born than expected.[31] While birth outcome results are not as definitive as our fertility results, they provide evidence that the effect we find is likely a combination of a selection and a scarring effect. In fact, even an effect size of zero for these birth outcomes

[30] We calculate this as the either the change in GFR in Flint only (62-57) or the difference-in-differences estimate (7.5) * population aged 15-49 in Flint (26,000) * the number of years affected (17/12) which gives us a range of 198 to 276.

[31] We multiply the change in sex ratio (0.009) * the number of post water change births (2,010) to get 18.

provides evidence of scarring because had there only been a selection effect, we would expect the health effects to be positive.

We perform an analysis in the spirit of Bozzoli, Deaton, and Quintana-Domeque (2009) to disentangle scarring and selection effects. We assume that the pre-water change birth weight distribution in Flint is normally distributed (see Figure 5) with mean (3082 g) and standard deviation (632 g) as in column 3 of Table 1. Using the 12% reduction in the live birth rate, we assume that this reduction all came from the left tail of the birth weight distribution, as birth weight is often thought of as a proxy for infant health. Another way to think of this is that there is some minimal birth weight cutoff for live birth, and the selection shock of adding lead to the water shifted the entire distribution left such that the bottom 12% of birth weight did not survive.

Using the standard formula for the mean of a truncated normal[32] we calculate that mean birth weight of the surviving newborns, without any scarring, would have been 3242 g. The observed Flint mean birth weight in the post period is 3042 g, a decrease of 200 g from 3242 g. Removing the pre-post difference in the rest of Michigan (from Columns 1 and 2 of Table 1) reduces this by 25 g (to 3217 g) and a scarring effect of 175 g, which is a 5.4% decrease compared to 3217 g. This is much larger than the scarring effect found from ignoring how scarring and selection cancel each other out (as in Gørgens, Meng, and Vaithianathan 2012) and naïvely using the coefficient in Table 4. We consider this a bounding exercise for the full effect of scarring had no selective attrition occurred. As Figure 5 makes clear, despite the large amount of selective attrition we document in Table 2, our pdf for Flint show that the health distribution shifted to the left in Flint following the water change and did not shift in comparison cities.

[32] I.e., $E\left(X \middle| X > \mu + \sigma\Phi^{-1}(p)\right) = \mu + \frac{\sigma\phi\left(\Phi^{-1}(p)\right)}{1-p}$, where μ is the mean, σ the standard deviation, Φ the standard normal CDF, φ the standard normal PDF, and p the truncation cutoff probability.

Additionally, while our sex ratio results are not definitive, they support our main result that fertility rates decrease because of both selective attrition and scarring from a biological effect of an increase in contaminants including lead in the water. The 0.9 percentage point increase (1.8%) in female births following the water change, is consistent with worse health at birth (Trivers and Willard 1973; Sanders and Stoecker 2015). Additionally, in Table 5, we find no evidence to support a decrease in sexual relations among individuals living in Flint during this time period. For our results to be explained by behavioral changes, we would have to postulate a theory that at the same time Flint changed its water source, parents changed their preference for male children and began performing sex-selective abortions showing a preference for *female* children. This result would run counter to the prevailing evidence of lower female births than expected, especially in Asian countries (e.g. Sen 1990; Das Gupta 2005), but also in the US (Abrevaya 2009).

Finally, we stress that our measure of health may not capture the full health effects of this water change. Firstly, infants born during this time period would have been exposed to water both in utero and for a period post-birth. Hanna-Attischa et al. (2016) show that children exposed to the new water regime had higher levels of blood lead. Secondly, the Barker hypothesis posits that measured health at birth only partially describes later life health. An additional component, latent health, may be exhibited later as poor health in adulthood, decreased educational attainment, increased behavioral problems and criminal behavior, and worse labor market outcomes (e.g., Almond and Currie 2011, Aizer and Currie 2017, Aizer et al. 2018, Reyes 2007, 2015, Billings and Schnepel 2018).

9. **Conclusion**

Failure to provide safe drinking water has large health implications. We provide the first estimates of the in utero effect of increased amounts of lead and other contaminants in drinking water in Flint. General fertility rates in Flint decreased substantially following the water change while health outcomes displayed mixed results, with suggestive evidence of an overall decrease in birth weight and an increase in low birth weight.

An overall decrease in fertility rates can have lasting effects on a community, including decreasing school funding due to a decrease in the number of students. Alternatively, if the decrease in births truly decreased the number of less healthy babies, it may reduce the health expenditures of the community. However, given the research demonstrating a substantial increase in blood lead levels among children in the community, an overall decrease in health expenditures in both the short and long-term seem highly unlikely (Hanna-Attischa et al. 2016; Edwards, Triantafyllidou, and Best 2009). Furthermore, children that were seemingly born healthy may still have worse latent health at birth, which could manifest itself later in life (Barker 1992; Barker 1995). While many on the ground programs have or are currently being implemented to counteract the water crisis (see e.g. Hanna-Attisha 2017), these latent health effects remain a concern.

This study has several limitations. First, previous work has demonstrated that lead builds up in the body over time, so that focusing on neonatal outcomes may underestimate the overall effects of lead on health and human development. Other contaminants may be present in the water that also effect health making our results an estimate of the *overall* effect of the water change on these outcomes. Additionally, the health effects of a change in water supply are not limited to pregnant women and neonates. This is just one piece of the health effects of this switch in water supply; however, given the litany of evidence linking fetal and birth outcomes to later life health, education, and labor outcomes, this study is an important step in investigating this public health issue. Despite these limitations, the culling of births in Flint provides robust evidence of the effect of lead on the health of not just infants, but on the health of potential newborns in utero.

This paper presents the first study of the Flint water change on fertility and birth outcomes. This is a natural experiment from which to study the effect of high concentrations of lead in water on birth outcomes. Lead problems in many municipalities have recently been reported, making these estimates important in informing public policy (see Wines and Schwartz 2016).

This study is of great importance as the current legislative environment includes calls for a substantial decrease in funding for the EPA which is charged with ensuring localities maintain minimum water standards. Our results suggest that a more lax regulatory environment in the context of drinking water may have substantial unforeseen effects on maternal and infant health, including large reductions in the number of births.

References:

Abadie, A and J Gardeazabal. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1): 113-132.

Abadie, A, A Diamond, and J Hainmueller. (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.

Abouk, R and S Adams. (2018). Birth Outcomes in Flint in the Early Stages of the Water Crisis. *Journal of Public Health Policy*, 39: 68-85.

Abreyava, J. (2009). Are There Missing Girls in the United States? Evidence from Birth Data. *American Economic Journal: Applied Economics*, 1(2): 1-34.

Adams, D. (2014). Flint officials say 'abnormal' test to blame in E. coli scare, water boil advisory remains. Mlive.com, August 18, 2014. Available at: http://www.mlive.com/news/flint/index.ssf/2014/08/flint_officials_say_abnormal_t.html.

Aizer, A, and J Currie. (2017). Lead and Juvenile Delinquency: New Evidence from Linked Birth, School and Juvenile Detention Records. *NBER Working Paper Series* No. 23392.

Aizer, A, J Currie, P Simon, and P Vivier. (2018). Do Low Levels of Blood Lead Reduce Children's Future Test Scores? *American Economic Journal: Applied Economics*, 10(1): 307-41.

Alexander, BH, et al. (1996.) Semen quality of men employed at a lead smelter. *Occupational and Environmental Medicine*, 53:411-416.

Almond, D. (2006). Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population. *Journal of Political Economy*, 114(4): 672-712.

Almond D and J Currie. (2011). Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, 25(3): 153-72.

Amaral, JH, VB Rezende, SM Quintana, RF Gerlach, F Barbosa Jr, and JE Tanus-Santos. (2010). The relationship between blood and serum lead levels in peripartum women and their respective umbilical cords. *Basic Clini Pharmacol Toxicol*, 107:971–5.

Apostoli, P, S Porru, and L Bisanti. (1999). Critical aspects of male fertility in the assessment of exposure to lead. *Scandinavian Journal of Work, Environment & Health*, 25(1): 40-43.

Apostoli, P, A Bellini, S Porru, and L Bistani. (2000). The Effect of Lead on Male Fertility: A Time To Pregnancy (TTP) Study. *American Journal of Industrial Medicine*, 38:310-315.

Assennato et al. (1987). Sperm Count Suppression without Endocrine Dysfunction in Lead-Exposed Men. *Archives of Environmental Health: An International Journal*, 42: 124-127.

ATSDR (Agency for Toxic Substances and Disease Registry). (2007). Toxicological Profile for Lead. Case No. 7439-92-1. Atlanta, Georgia:ATSDR. Available at: http://www.atsdr.cdc.gov/toxprofiles/tp13.pdf

Barker, DJ. (1992). The fetal origins of adult hyptertension. *Journal of Hypertension*, 10(suppl 7):S39-S44.

Barker, DJ. (1995). Fetal origins of coronary heart disease. *BMJ*, 311(6998):171-4.

Barreca, A, O Deschenes, and M Guldi. (2016). Maybe Next Month? The Dynamic Effects of Ambient Temperature on Fertility. Forthcoming *Demography* Also *NBER Working Paper* No. w21681.

Bellinger, DC. (2005). Teratogen Update: Lead and Pregnancy. *Birth Defects Research*, 73:409–420.
Billings, SB., and KT Schnepel. (2018). Life After Lead: Effects of Early Interventions for Children Exposed to Lead. Forthcoming *American Economic Journal: Applied Economics*.
Bonde, JPE and H Kolstad. 1997. Fertility of Danish Battery Workers Exposed to Lead. *International Journal of Epidemiology*, 26(6): 1281-1288.
Borje-Aburto, VH, I Hertz-Picciotto, M Rojas Lopez, P Farias, C Rios, J Blanco. 1999. Blood lead levels measured prospectively and risk of spontaneous abortion. *Am J Epidemiol*, 150(6):590-7.
Bozzoli, C, A Deaton, and C Quintana-Domeque. (2009). Adult Height and Childhood Disease. *Demography*, 46(4): 647-669.
Cameron, AC, JB Gelbach, and DL Miller. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3): 414-427.
Carmody, T. (2016). How the Flint River got so toxic: Factories and people have been dumping sewage, chemicals, and road salt in the Flint River for more than a century. The Verge. https://www.theverge.com/2016/2/26/11117022/flint-michigan-water-crisis-lead-pollution-history
Cao, WC, Q Zeng, Y Luo, et al. (2016). Blood Biomarkers of Late Pregnancy Exposure to Trihalomethanes in Drinking Water and Fetal Growth Measures and Gestational Age in a Chinese Cohort. Environmental Health Perspectives 124: 536-541.
CDC (Centers for Disease Control and Prevention) Third National Report on Human Exposure to Environmental Chemicals. Atlanta, GA: CDC, National Center for Environmental Health; 2005. Pub. No. 05-0570.
Christensen, P, DA Keiser, and GE Lade. (2017). Economic Effects of Environmental Crises: Evidence from Flint, Michigan. Unpublished Manuscript.
City of Flint: Water System Questions & Answers. (2015). January 13, 2015 Available at: https://www.cityofflint.com/wp-content/uploads/CoF-Water-System-QA.pdf
City of Flint: Water System Update with Questions & Answers. (2015). February 16, 2015 Available at: https://www.cityofflint.com/wp-content/uploads/Water-Sysytem-FAQ-Update-2-16-151.pdf
Clark, BN, SV Masters, and MA Edwards. (2015). Lead Release to Drinking Water from Galvanized Steel Pipe Coatings. *Environmental Engineering Science*, 32(8): 713-21.
Clay K, W Troesken, and M Haines. (2014). Lead, Mortality, and Productivity. *Review of Economics and Statistics*, 96(3): 458-70.
Clay, K, M Portnykh, and E Severini. (2018). Toxic Truth: Lead Exposure and Fertility Choices. *NBER Working Paper Series* No. w24607.
Cleveland, LM, ML Minter, KA Cobb, AA Scott, and VF German. (2008). Lead Hazards for Pregnant Women and Children. *American Journal of Nursing*, 108(10): 40–49.
Coste, J, et al. 1991. Lead-Exposed Workmen and Fertility: A Cohort Study on 354 Subjects. *European Journal of Epidemiology*, 7(2): 154-158.
Cunningham, S and M Shah. (2018). Decriminalizing Indoor Prostitution: Implications for Sexual Violence and Public Health. Forthcoming at *Review of Economic Studies*.
Das Gupta, M. (2005). Explaining Asia's 'missing women': a new look at the data. *Population and Development Review*, 31(3), 529–535.

Davis 2017
http://www.oregonlive.com/environment/index.ssf/2017/03/here_are_42_of_president_donal.html
Dehejia, R, and A Lleras-Muney. (2004). "Booms, Busts, and Babies' Health," *Quarterly Journal of Economics*, 119(3): 1091-1130.
Edwards, M. (2014). Fetal Death and Reduced Birth Rates Associated with Exposure to Lead-Contaminated Drinking Water. *Environ. Sci. Technol.*, 48 (1): 739–746.
Edwards, M, S Triantafyllidou, and D Best. (2009). Elevated Blood Lead in Young Children Due to Lead-Contaminated Drinking Water: Washington, DC, 2001-2004. Environ. Sci. Technol,. 43:1618–1623.
Egan, P. (2017). These are the 15 people criminaaly charged in the Flint water crisis. Detroit Free Press, June 14, 2017. Available at: http://www.freep.com/story/news/local/michigan/flint-water-crisis/2017/06/14/flint-water-crisis-charges/397425001/
Eibensteiner, L, ADC Sanz, H Frumkin, C Gonzales, and GF Gonzales. (2013). Lead Exposure and Semen Quality among Traffic Police in Arequipa, Peru. *International Journal of Occupational and Environmental Health* 11: 161-166.
Emery, A. (2015). Flint Reconnects to Detroit Water, may take 3 weeks to clear all pipes. Mlive.com, October 16, 2015. Available at: http://www.mlive.com/news/flint/index.ssf/2015/10/flint_reconnecting_to_detroit.html.
Ferrie, JP, K Rolf, and W Troesken. (2012). Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence test scores among World War Two U.S. Army enlistees. *Economics and Human Biology*, 10: 98–111.
Fisher, R.A. (1935). The Design of Experiments (Edinburgh: Oliver and Boyd).
Fonger R. (2014a.) Tests positive for total coliform again in water-boil area on Flint's west side. Mlive.com**,** August 19, 2014. Available at: http://www.mlive.com/news/flint/index.ssf/2014/08/water_boil_area_in_flint_gets.html.
Fonger, R. (2014b). Flint issues boil water advisory for section of the city after positive test for total coliform bacteria. Mlive.com**,** September 5, 2014. Available at: http://www.mlive.com/news/flint/index.ssf/2014/09/flint_issues_boil_water_adviso.html.
Fonger, R. (2014c). General Motors shutting off Flint River water at engine plant over corrosion worries. Mlive.com**,** October 13, 2014. Available at: http://www.mlive.com/news/flint/index.ssf/2014/10/general_motors_wont_use_flint.html
Fonger R. (2014d). Emergency manager accepts $3.9 million Genesee County offer to buy Flint-owned pipeline. Mlive.com**,** June 12, 2014. Available at http://www.mlive.com/news/flint/index.ssf/2014/06/emergency_manager_accepts_39_m.html.
Fonger, R. (2015a). Flint to issue lead in water warning after push from doctors, health officials. Mlive.com, September 24, 2015. Fonger, R. 2015b. Elevated lead found in more Flint kids after water switch, study finds. Mlive.com, September 24, 2015. http://www.mlive.com/news/flint/index.ssf/2015/09/study_shows_twice_as_many_flin.html.
Fonger, R. (2015b). City warns of potential health risks after Flint water tests revealed too much disinfection byproduct. Mlive.com**,** January 2, 2015. Available at: http://www.mlive.com/news/flint/index.ssf/2015/01/flint_water_has_high_disinfect.html

Fonger, R. (2015c). Elevated lead found in more Flint kids after water switch, study finds. Mlive.com, September 24, 2015. http://www.mlive.com/news/flint/index.ssf/2015/09/study_shows_twice_as_many_flin.html.

Gallagher, MD, JR Nuckols, L Stallones, and DA Savitz. (1998). Exposure to Trihalomethanes and Adverse Pregnancy Outcomes. *Epidemiology*, 9(5): 484-489.

Gazze, L. (2016). Lead Policies, Lead Poisoning, and Government Spending. Unpublished Manuscript.

Gørgens, T, X Meng, and R Vaithianathan. (2012). Stunting and selection effects of famine: A case study of the Great Chinese Famine. *Journal of Development Economics*, 97(1): 99-111.

Hamilton and Hardy's Industrial Toxicology, Revised by Asher J. Finkel. 1983. Boston, Bristol, London: John Wright PSG

Hanna-Attisha M, (2017). Flint Kids: Tragic, Resilient, and Exemplary. *American Journal of Public Health*, 107(5): 651-652.

Hanna-Attisha M, J LaChance, RC Sadler, and A Champney Schnepp. (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: 283-290.

Hernberg, S. (2000). Lead Poisoning in a Historical Perspective. *American Journal of Industrial Medicine*, 38:244-254.

Hertz-Picciotto, I. (2003). The Evidence that Lead Increases the Risk for Spontaneous Abortion. *American Journal of Industrial Medicine*, 38:300-309.

Horton, BJ, TJ Luben, AH Herring, DA Savitz, PC Singer, HS Weinberg, and KE Hartmann. (2011). The Effect of Water Disinfection By-products on Pregnancy Outcomes in Two Southeastern U.S. Communities. *Journal of Occupational and Environmental Medicine,* 53(10): 1172-1178.

Hu, H, MM Téllez-Rojo, D Bellinger, D Smith, AS Ettinger, H Lamadrid-Figueroa, J Schwartz, L Schnaas, A Mercado-García, and M Hernández-Avila. (2006). Fetal Lead Exposure at Each Stage of Pregnancy as a Predictor of Infant Mental Development. *Environ Health Perspect*, 114:1730–1735.

Joffe, M. et al. (2003). Time To Pregnancy and occupational lead exposure. *Occup Environ Med*, 60:752–758.

Kaul, A, S Klößner, G Pfeifer, and M Schieler. (2017). Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates. *Mimeo* University of Hohenheim.

Lin, S, S Hwang, EG Marshall, R Stone, and J Chen. (1996). Fertility Rates among Lead Workers and Professional Bus Drivers: A Comparative Study. *AEP*, 6(3): 201-208.

Lin S, S Hwang, EG Marshall, and DMarion. 1998. Does paternal occupational lead exposure increase the risks of low birth weight or prematurity? *Am J Epidemiol*, 148:173–181.

Lindo, Jason M., and Analisa Packham. (2017). "How Much Can Expanding Access to Long-Acting Reversible Contraceptives Reduce Teen Birth Rates?" *American Economic Journal: Economic Policy*, 9(3): 348-76.

Longley, K. (2011). Emergency Manager Michael Brown appointed to lead Flint through second state takeover. Mlive.com, November 29, 2011. Available at:

http://www.mlive.com/news/flint/index.ssf/2011/11/emergency_manager_michael_brow.html.

Masten, SJ, SH Davies, and SP Mcelmurry. (2016). Flint Water Crisis: What Happened and Why? J Am Water Works Assoc, 108(12):22-34.Neidell, M. (2009). Information, Avoidance Behavior, and Health The Effect of Ozone on Asthma Hospitalizations. *Journal of Human Resources*, 44(2): 450-478.

Niewenhuijsen, MJ, MB Toledano, NE Eaton, J Fawell, and P Elliott. (2000). Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review. *Occup Environ Med*, 57: 73-85.

Olson, TM, M Wax, J Yonts, K Heidecorn, SJ Haig, D Yeoman, Z Hayes, L Raskin, and BR Ellis. (2017). Forensic Estimates of Lead Release from Lead Service Lines during the Water Crisis in Flint, Michigan. *Environmental Science & Technology Letters*, 4(9): 356-361.

Paul, C. (1860). Étude sur l'intoxication lente par les préparations de plomb et son influence sur le production de la conception. (Studies on the chronic poisoning by lead compounds and its influence on the fecundity.) *Arch Gén de Med*, 15:344-360.

Pieper, KJ, M Tang, and MA Edwards. (2017). Flint Water Crisis Caused by Interrupted Corrosion Control: Investigating "Ground Zero" Home. *Environmental Science & Technology*, 51(4): 2007−2014.

Powell, D. (2017). Imperfect Synthetic Controls: Did the Massachusetts Health Care Reform Save Lives? Unpublished manuscript.

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

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

Rhoads, WJ, E Garner, P Ji, N Zhu, J Parks, DO Schwake, A Pruden, and MA Edwards. Distribution System Operational Deficiencies Coincide with Reported Legionnaires' Disease Clusters in Flint, Michigan. Environ. *Sci*. *Technol*., 51(20): 11986-11995.

Robbins, D. (2016). ANALYSIS: How Michigan And National Reporters Covered The Flint Water Crisis. https://www.mediamatters.org/research/2016/02/02/analysis-how-michigan-and-national-reporters-co/208290

Rudge, CV, HB Rollin, CM Nogueira, Y Thomassen, MC Rudge, and JO Odland. (2009). The placenta as a barrier for toxic and essential elements in paired maternal and cord blood samples of South African delivering women. *J Environ Monitor*, 11:1322–30. .

Sallmén, M, M Lindbhm, and M Nirumnem. (2000). Paternal Exposure to Lead and Infertility. *Epidemiology*, 11(2): 148-152.

Sallmén, M. (2001). Exposure to lead and male fertility. *Int J Occup Med Environ Health*, 14(3):219-22.

Sanders, NJ, and C Stoecker. (2015). Where have all the young men gone? Using sex ratios to measure fetal death rates, *Journal of Health Economics*, 41: 30-45.

Sauve-Syed, K. (2017). Lead Exposure and Student Performance: A Study of Flint Schools. Unpublished Manuscript.

Schell, LM, M Denham, AD Stark, M Gomez, J Ravenscroft, PJ Parsons, et al. (2003). Maternal blood lead concentration, diet during pregnancy, and anthropometry predict neonatal blood lead in a socioeconomically disadvantaged population. *Environ Health Persp*, 111:195–200.

Schultz, LC. (2010). The Dutch Hunger Winter and the Developmental Origins of Health and Disease. *Proceedings of the National Academy of Sciences*, 107(39): 16757–58

Scorsone, E, and N Bateson. (2011). Long-Term Crisis and Systemic Failure: Taking the Fiscal Stress of America's Older Cities Seriously. Case Study: City of Flint, Michigan. Michigan State University Extension. Available at: https://www.cityofflint.com/wp-content/uploads/Reports/MSUE_FlintStudy2011.pdf

Sen, A., 1990. More than 100 million women are missing. The New York Review of Books.

Shaiau, C, J Wang, and P Chen. (2004). Decreased fecundity among male lead workers. *Occup Environ Med*, 61:915–923.

Taylor, CM, J Golding, and AM Emond. (2015). Adverse Effects of Maternal Lead Levels on Birth Outcomes in the ALSPAC Study: A Prospective Birth Cohort Study. *BJOG*, 122(3): 322-328.

Trivers, RL, and DE Willard. (1973). Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring. *Science*, 179 (4068):90-92.

Troesken, W and PE Beeson. (2003). The Significance of Lead Water Mains in American Cities: Some Historical Evidence. "Health and Labor Force Participation over the Life Cycle: Evidence from the Past," ed. DL Costa. University of Chicago Press.

Troesken, W. (2006a). "The Great Lead Water Pipe Disaster." MIT Press.

Troesken, W. (2006b). Lead exposure and eclampsia in Britain, 1883–1934. *Environmental Research*, 101: 395–400.

Troesken, W. 2008 Lead Water Pipes and Infant Mortality at the Turn of the Twentieth Century. *Journal of Human Resources*, 43(3): 553-575.

University of Michigan-Flint GIS Center. Map of Flint's Lead Water Pipes. Available at: https://www.umflint.edu/gis.

Vigeh, M, K Yokoyama, F Kitamura, M Afshinrokh, A Beygi, and S Niroomanesh. (2010). Early Pregnancy Blood Lead and Spontaneous Abortion. *Women & Health*, 50:756–766.

Vigeh, M, DR Smith, and P Hsu. (2011). How does lead induce male infertility? *Iranian Journal of Reproductive Medicine*, 9(1): 1-8.

Walsh, MW. (2014). Detroit's Plan to Profit on Its Water, by Selling to Its Neighbors, Looks Half Empty. The New York Times, May 26, 2014. Available at: http://www.nytimes.com/2014/05/26/business/detroit-plan-to-profit-on-water-looks-half-empty.html

Winder, C. (1993). Lead, reproduction and development. *Neurotoxicology*, 14(2-3):303-317.

Wines, M and J Schwartz. (2016). Unsafe lead levels in tap water not limited to Flint. The New York Times, Feb. 8, 2016. Available at: http://www.nytimes.com/2016/02/09/us/regulatory-gaps-leave-unsafe-lead-levels-in-water-nationwide.html?_r=0.

Wolfers, Justin. (2006). "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review*, 96(5): 1802-20.

Worth, JJ and RS Mijal. (2010). Special Issue: SBiRM: Focus on Impact of Environmental Toxicants on Reproductive Function. *Systems Biology in Reproductive Medicine*, 56:147–167.

Wu, H. et al. (2012). Lead level in seminal plasma may affect semen quality for men without occupational exposure to lead. *Reproductive Biology and Endocrinology*, 10:91.

Yang, CY, ZP Xiao, SC Ho, TN Wu, and SS Tsai. (2007). Association between trihalomethane concentrations in drinking water and adverse pregnancy outcome in Taiwan. *Environmental Research*, 104: 390-395.

Zahran, S, SP McElmurry, and RC Sadler. (2017). Four phases of the Flint Water Crisis: Evidence from blood lead levels in children. *Environmental Research*, 157: 160-172.

Zhu M, EF Fitzgerald, KH Gelberg, S Lin, and CM Druschel. (2010). Maternal low-level lead exposure and fetal growth. *Environmental Health Perspectives*, 118: 1471-1475.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)
	Non-Flint Births		Flint Births		
	Pre-Water Change (N=238,733)	Post-Water Change (N=52,311)	Pre-Water Change (N=10,620)	Post-Water Change (N=2,010)	Difference in Differences
Demographic variables:					
Mother's age (years)	27.49 (6.07)	28.10 (5.78)	24.66 (5.59)	25.17 (5.37)	-0.105
Mother no high school	0.190	0.155	0.294	0.271	0.011
Mother high school grad	0.268	0.266	0.317	0.343	0.028*
Mother some college	0.275	0.292	0.337	0.337	-0.0168
Mother college grad	0.258	0.276	0.050	0.047	-0.021**
Outcome variables:					
General fertility rate	67.14 (33.38)	69.18 (31.83)	62.28 (6.81)	56.87 (6.76)	-7.45**
Male-Female Sex Ratio (percent male)	51.11 (4.44)	51.20 (4.95)	51.05 (4.59)	50.20 (3.06)	-0.92
Birth weight (grams)	3,225 (631)	3,200 (645)	3,082 (632)	3,042 (651)	-15
Low Birth Weight	0.097 (0.30)	0.108 (0.31)	0.135 (0.34)	0.158 (0.37)	
Estimated gestational age (weeks)	38.50 (2.99)	38.38 (2.58)	38.08 (2.97)	37.89 (2.69)	-0.078
Gestational Growth (grams/week)	83.32 (14.63)	82.79 (14.61)	80.38 (14.33)	79.58 (14.48)	-0.27

Notes: For Columns (1)-(4), standard deviation for non-dummy variables in parenthesis. For Column (5), robust standard errors are in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001

Table 2: Lead in Water on General Fertility Rate and Sex Ratios

	General Fertility Rates			Sex Ratios		
	(1)	(2)	(3)	(4)	(5)	(6)
Water (β_1)	-7.451***	-7.451***	-5.682***	-0.0092***	-0.0092***	-0.00121
	(0.791)	(0.811)	(0.603)	(0.00262)	(0.00268)	(0.00411)
		[0.004]			[0.004]	
Conception Month Fixed Effects (FE)	X	X	X	X	X	X
Conception Year FE	X	X	X	X	X	X
City FE	X	X	X	X	X	X
Conception Month into Year FE		X	X		X	X
City Linear Time Trends			X			X
Observations	1,520	1,520	1,520	1,520	1,520	1,520
Cities	16	16	16	16	16	16
R-squared	0.235	0.269	0.303	0.235	0.269	0.303
Mean	62.28	62.28	62.28	0.510	0.510	0.510

Notes: Robust standard errors clustered at the city level in parentheses. Brackets contain wild bootstrapped p-values for the most saturated models. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001

Table 3: Lead in Water on General Fertility Rate and Sex Ratio, Subsample Analyses

	(1)	(2)	(3)	(4)	(5)
	All	No High School Degree	Only High School Degree	Some College (no degree)	Post-secondary degree
Panel A. GFR					
	-8.188***	26.81*	9.124†	-20.27***	-9.876***
	(1.112)	(9.224)	(4.836)	(3.170)	(2.079)
R-squared	0.247	0.091	0.101	0.149	0.204
Mean	76.92	109.2	85.75	68.01	49.53
Panel B. Sex Ratio					
	-0.0169***	0.00944	0.00816	-0.0162†	-0.0481***
	(0.00403)	(0.0176)	(0.00755)	(0.00882)	(0.00705)
R-squared	0.008	0.055	0.077	0.054	0.060
Mean	0.493	0.505	0.501	0.523	0.487

Notes: N=1520 observations across 16 cities for all regressions. Only considering mothers 18-44 to be consistent with the annual city population counts by educational attainment. All contain conception month into year fixed effects and city fixed effects. Robust standard errors clustered at the city level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001 .

Table 4: Lead in Water on Other Birth Outcomes

	(1)	(2)	(3)	(4)	(5)
Birth Weight (grams)	-14.69	-20.82	-19.47	-18.42	-10.78
	(13.65)	(14.65)	(14.48)	(14.77)	(14.64)
Low Birth Weight	0.0124	0.0142	0.0136	0.0135	0.0114
	(0.00873)	(0.00884)	(0.00880)	(0.00887)	(0.00900)
Gestational Age (weeks)	-0.0668	-0.0878	-0.0830	-0.0835	-0.0701
	(0.0586)	(0.0612)	(0.0603)	(0.0602)	(0.0603)
Gestational Growth	-0.265	-0.392	-0.364	-0.336	-0.159
(grams/week)	(0.307)	(0.328)	(0.325)	(0.334)	(0.332)
Census Tract Fixed Effects		X	X	X	X
Conception Month Fixed Effects			X	X	X
Conception Year Fixed Effects			X	X	X
Child Sex Control				X	X
Mom Controls					X
N	303,674	303,674	303,674	303,674	303,674

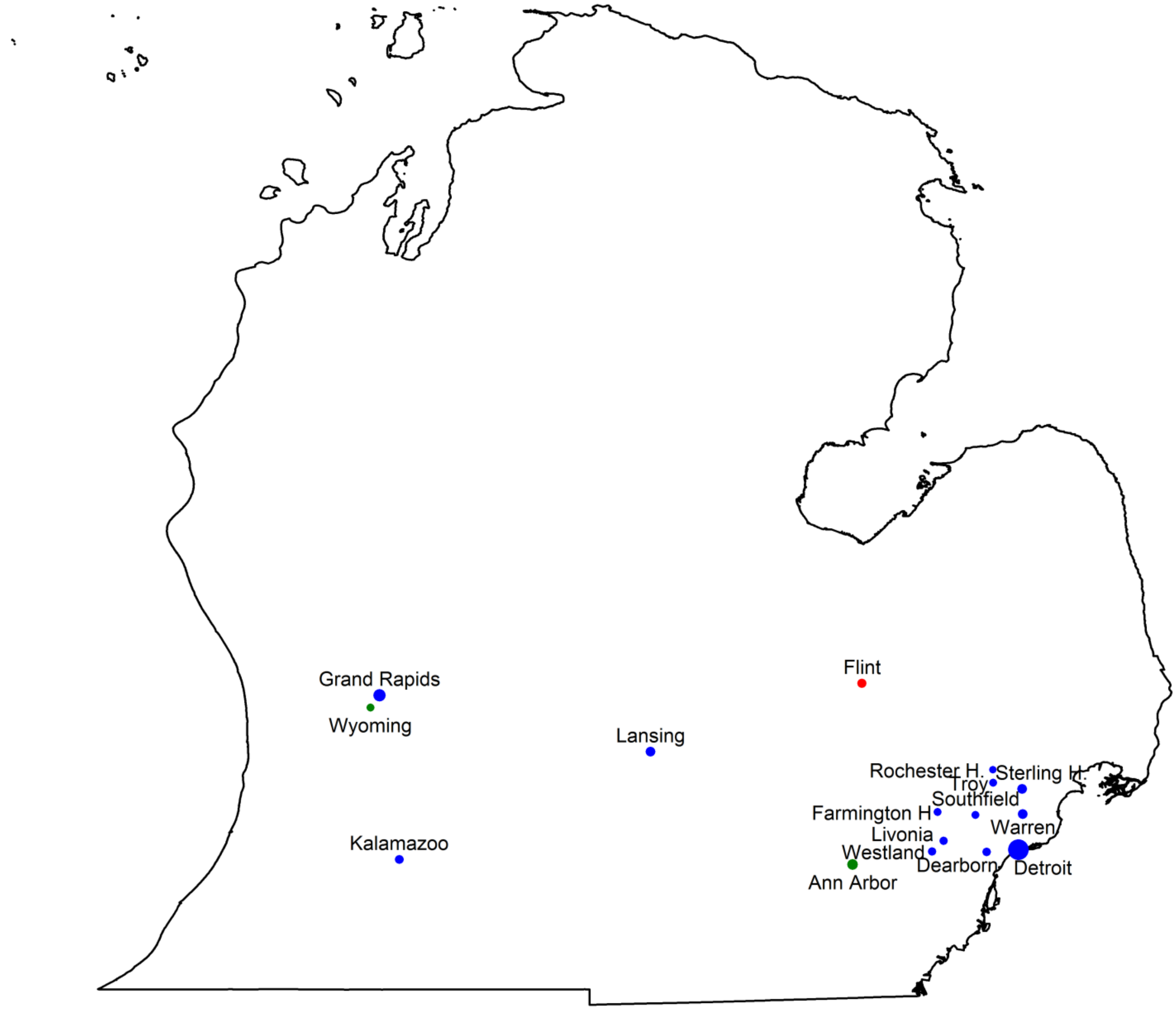
Notes: Robust standard errors clustered at the census tract level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001

Table 5: Time Use Data on Sex

	(1)	(2)	(3)	(4)	(5)	(6)
	County-level			CBSA-level		
Water (β_1)	0.0148***	0.0158***	0.0157***	0.0186***	0.0206***	0.0205***
	(0.00203)	(0.00133)	(0.00131)	(0.00229)	(0.00319)	(0.00310)
Conception Month Fixed Effects		X	X		X	
Conception Year Fixed Effects		X	X		X	X
County Fixed Effects			X			
CBSA Fixed Effects						X
Observations	861	861	861	745	745	745
Counties/CBSAs	16	16	16	13	13	13
R-squared	0.011	0.037	0.036	0.003	0.028	0.027

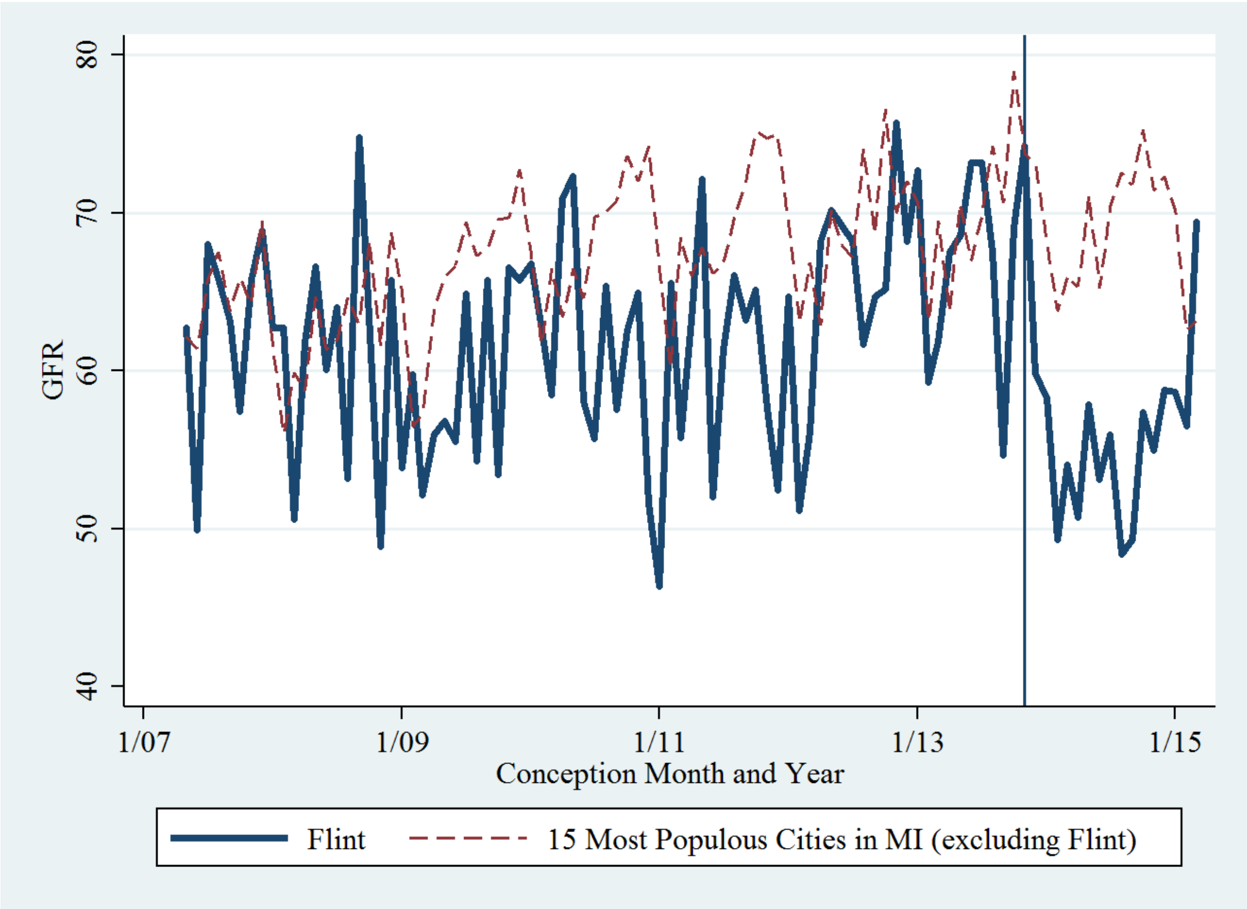
Notes: Robust standard errors clustered at the county or CBSA level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001

Figure 1: Comparison Cities



Note: Comparison cities are in blue, Flint in red, and cities with outlier GFR in green. Point size is proportional to the population of women age 15-49 in that city in 2014.

Figure 2: Unadjusted Monthly GFR[33]

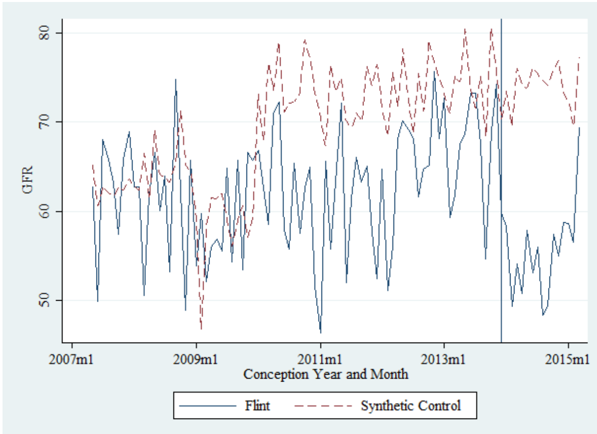


Note: The vertical blue line is at November 2013, the first month in which all conceived births would have been affected by the water supply for at least one trimester.

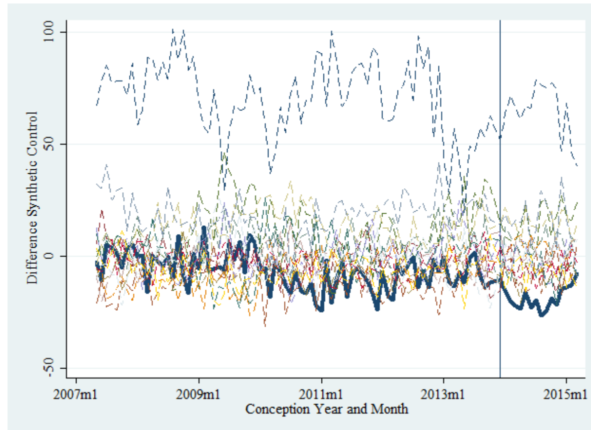
[33] This figure does appear to show city specific seasonality. Given the differences in the demographic composition of Flint and some other cities and the correlation between socioeconomic status and birth seasonality (Trivers and Willard 1973) we run a specification using city specific month fixed effects and find virtually identical results.

Figure 3: Synthetic Control Results for General Fertility Rates

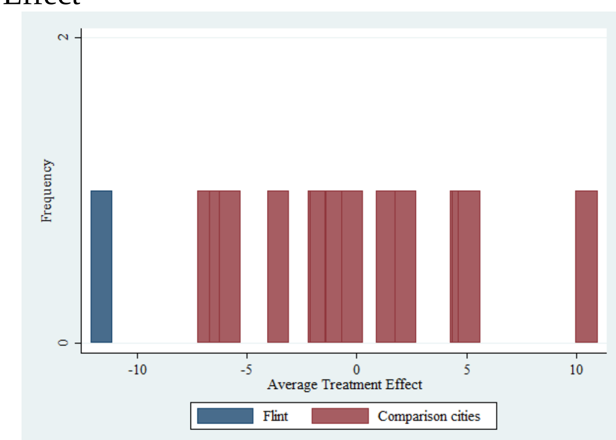
Panel A. Flint GFR compared to Synthetic Flint GFR



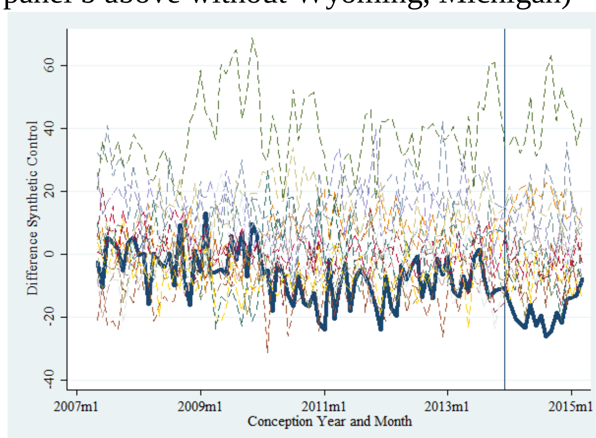
Panel B. Difference Between Each City and its Synthetic Counterpart



Panel C. Inference using Average Treatment Effect



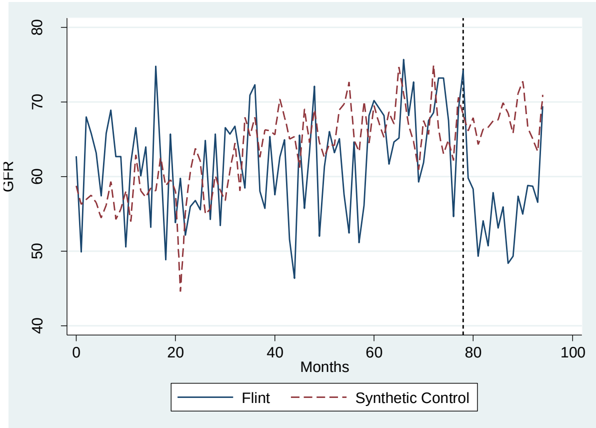
Panel D. Dropping outlier (same as panel b above without Wyoming, Michigan)



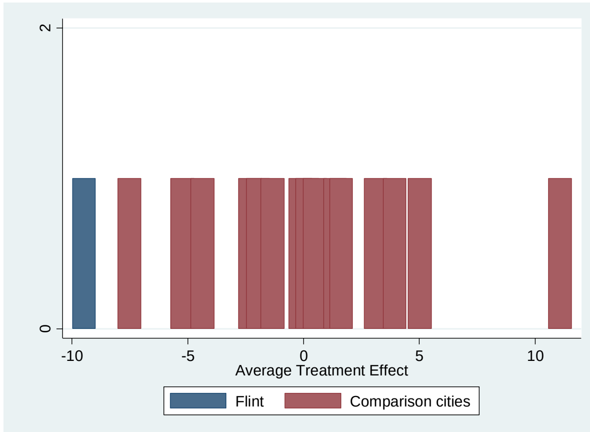
Note: The blue vertical line in Panel A is at November 2013, which is the last conception date for which women would have been exposed for at least one trimester to the new water supply. The blue solid line in Panel B represents the difference between GFR in Flint and "synthetic Flint." The blue bar in Panel C displays the average treatment effect (ATE) for Flint, while red bars show comparison city ATEs. It is the most negative ATE compared to assigning all areas to treatment, suggesting statistical significance.

Figure 4: Imperfect Synthetic Controls

Panel A. Flint GFR compared to Imperfect Synthetic Flint GFR

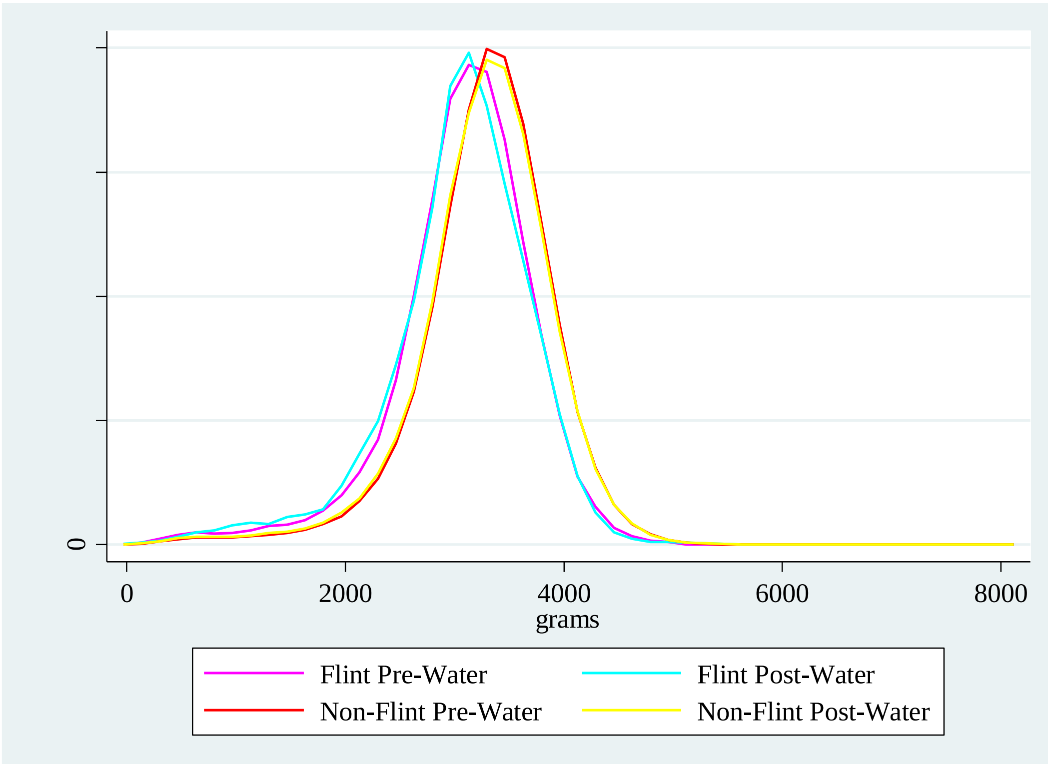


Panel B. Inference using Average Treatment Effect



Note: The blue vertical line in Panel A is at November 2013, which is the last conception date for which women would have been exposed for at least one trimester to the new water supply. The blue bar in Panel B displays the average treatment effect (ATE) for Flint, while red bars show comparison city ATEs. It is the most negative ATE compared to assigning all areas to treatment, suggesting statistical significance.

Figure 5: Probability Density Function of Birth Weight Before and After the Water Switch.



The Impact of the Flint Water Crisis on Fertility: Online Appendix

Appendix A: Synthetic Control Methods

The synthetic control method creates a weighted control group matched on pre-water supply trends, including the outcome of interest fertility rates and birth outcomes, such that the vector of weights (W) minimizes:

$$\|X_1 - X_0W\| = \sqrt{(X_1 - X_0W)'V(X_1 - X_0W)}$$

where X_1 is an unweighted vector of pre-intervention characteristics of the treatment counties and X_0 denotes a similar vector for control counties. The pool of control counties consists of the largest 15 cities in Michigan that did not change their water supply over this time period.[34] One strength of a synthetic control analysis is if a control county is trending differently from the treatment, it can receive zero weight. This method creates a weighted comparison group that minimizes the root mean squared error of the outcome variables in the pre-treatment period, which is the standard deviation in the difference between the actual outcome value of the treatment group and the predicted outcome value of the synthetic control group (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010).

The basic specification adjusts for the average pre-period general fertility rate of interest in each and the average of the following variables over the same pre-period: mother's educational attainment including less than high school, high school graduate, some college, and college graduate, race, age of mother, and gender of the child.

[34] Cities included are Ann Arbor, Dearborn, Detroit, Farmington Hills, Flint, Grand Rapids, Kalamazoo, Lansing, Livonia, Rochester Hills, Southfield, Sterling Heights, Troy, Warren, Westland, and Wyoming.

The main strengths of this method are it creates a matched control group that follows similar pre-trends in terms of the outcome of interest, and it allows for rigorous inference testing. Because the control areas follow similar pre-trends and are matched on level as well, they are plausibly a better counterfactual representation of what one would expect to have happened to pregnancy and birth outcomes in Flint had the city never switched its water source.

Inference testing consists of systematically assigning treatment to each control zone, creating a synthetic control group using the city of Flint (the treatment zone) as a control as well as the full pool of control zones, minus the city assigned to treatment. We separately calculate the average treatment effect in the post-period of assigning treatment to each control zone. This creates a distribution of average treatment effects by which to evaluate the average treatment effect of the actual water supply change in Flint. So if there are 16 average treatment effects and the Flint effect is larger than the other 15 control area average treatment effects, the estimate is statistically significant at the 6.25 percent level.[35]

[35] 1/16=0.06

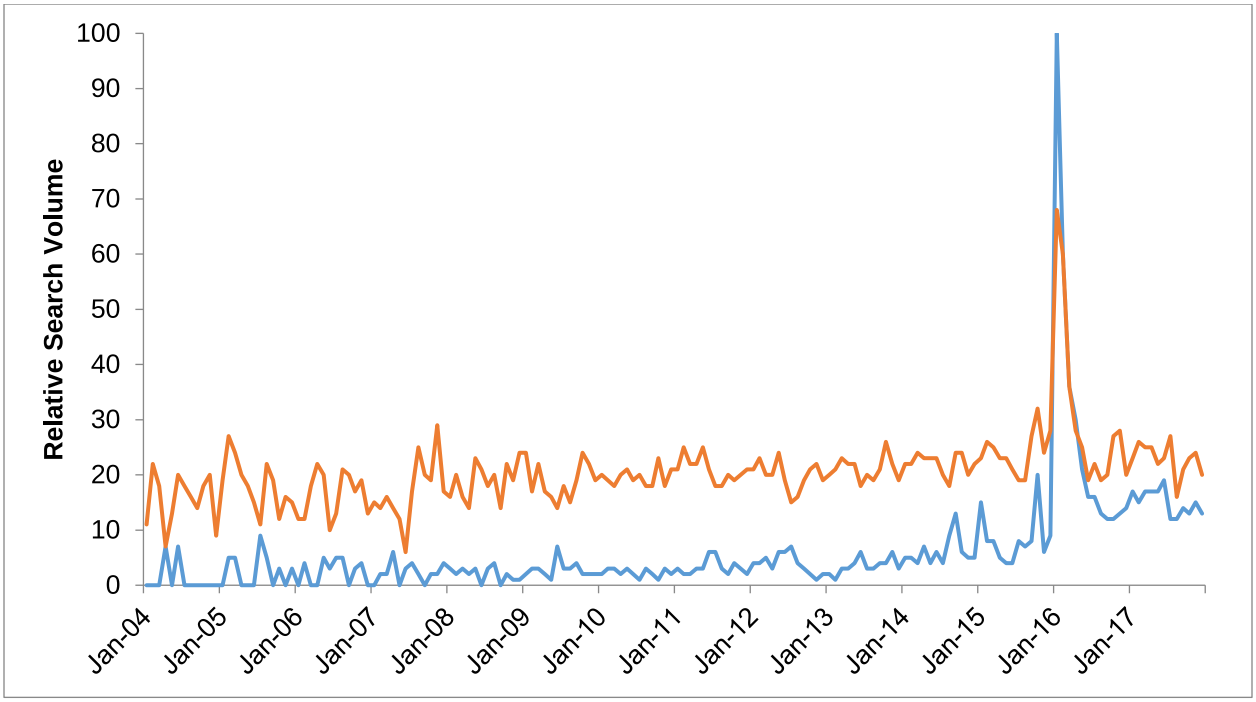
Appendix B: Additional Tables and Figures:

Appendix Figure B1: Timeline of Important Events in Flint

1897: Flint passes ordinance that all connections with any water main be made with lead pipes (Masten et al. 2016)	1967-2014: Flint receives water from Detroit Water and Sewerage Department (DWSD)	2011: Governor appoints Emergency Manager	2009-2013: Water rates (prices) consistently increase	March 2014: Flint and Genesee County plan own pipeline to Lake Huron	April 2014: Flint changes water source to Flint River, Genesee County stays with DWSD	Aug – Sept 2014: Positive test for fecal coliform, first boil advisory

Oct 2014: Flint GM plant switches off Flint water supply because of engine corrosion.	Dec 2014: EPA violation for too much trihalomethane concentration in the Flint water.	Jan – Mar 2015: Emergency manager stresses water is safe, refuses to return to DWSD	Jun – Jul 2015: Dr. Edwards independently tests Flint water lead levels, 19 times more corrosive than DWSD.	Sept 2015: Dr. Hanna-Attisha holds press conference announcing increased rates of child blood lead levels.	Oct 2015: Flint stops receiving water from Flint River.

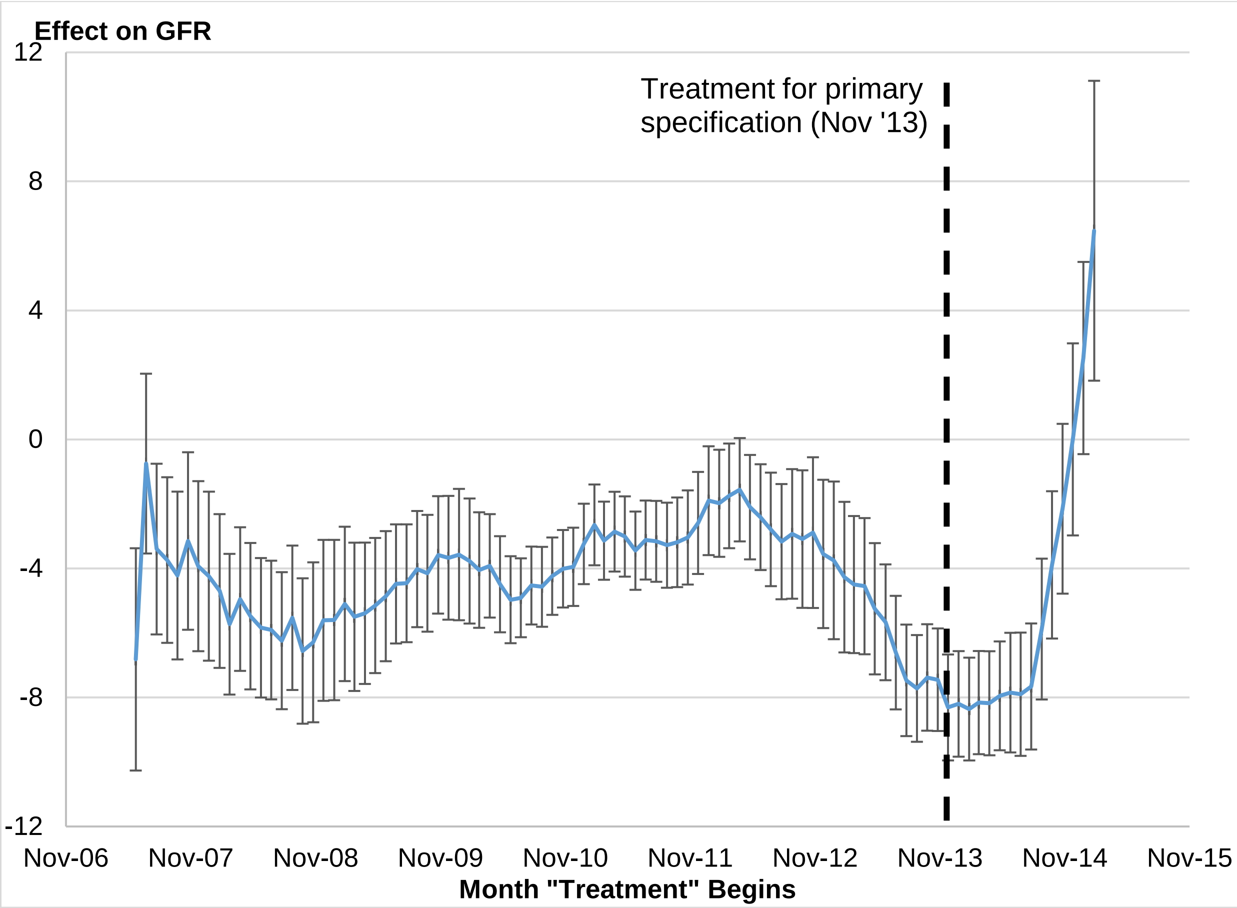
Appendix Figure B2: Google Trend Data on Searches for Water and Lead in Flint



Source: Google Trends

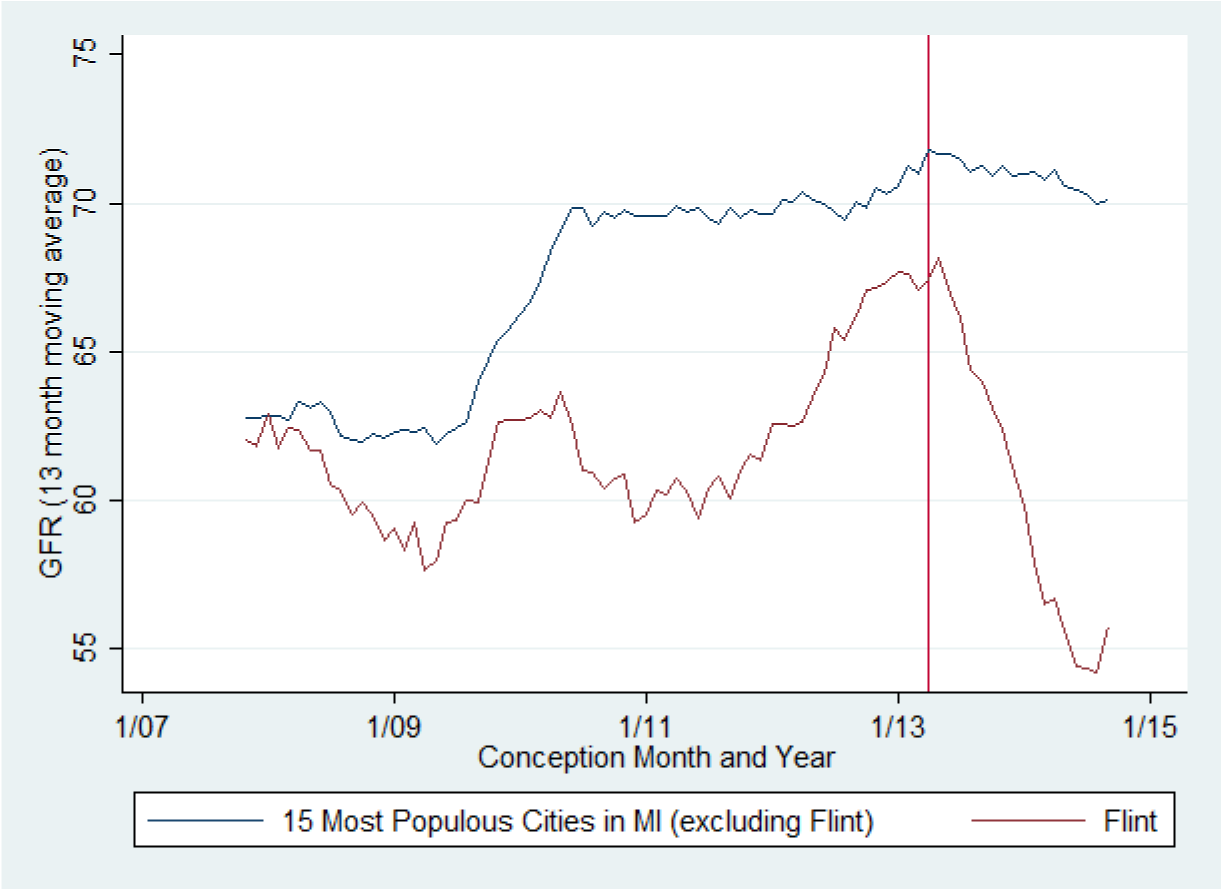
Notes: Searches for "flint water" in blue and "lead" in orange.

Appendix Figure B3: Results from Regressions with Alternate Treatment Dates



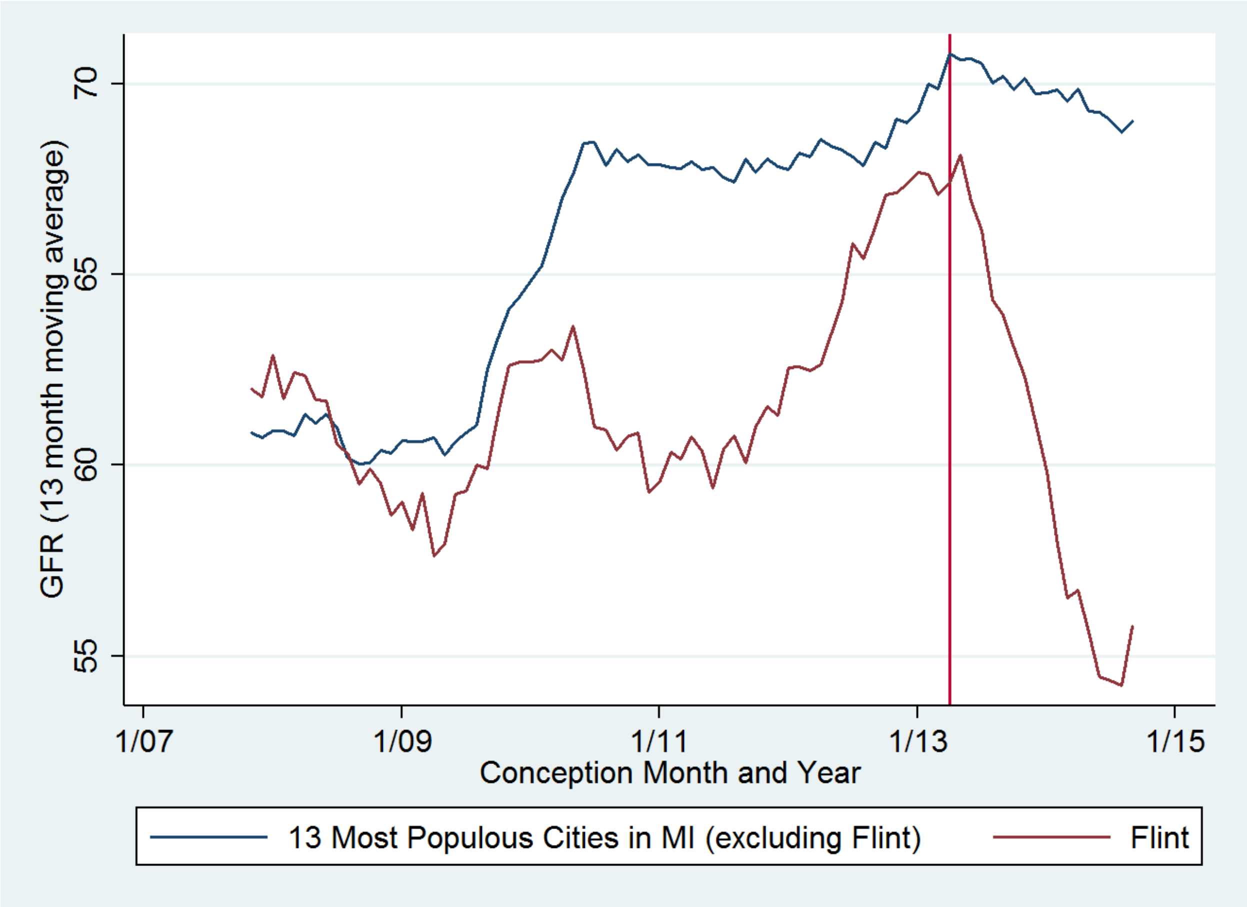
Note: Each point is the coefficient from a different regression. 95% confidence interval around each regression coefficient. All regressions include city and conception month into year fixed effects.

Appendix Figure B4: Moving Average Fertility Rate in Flint and Comparison Cities



Note: The red vertical line is at April 2013, which is the last conception date for which no affected birth rates are included in the moving average.

Appendix Figure B5: Moving Average Fertility Rate Over Time in Flint and Comparison Cities – Dropping Outlier Cities



Note: The red vertical line is at April 2013, which is the last conception date for which no affected birth rates are included in the moving average.

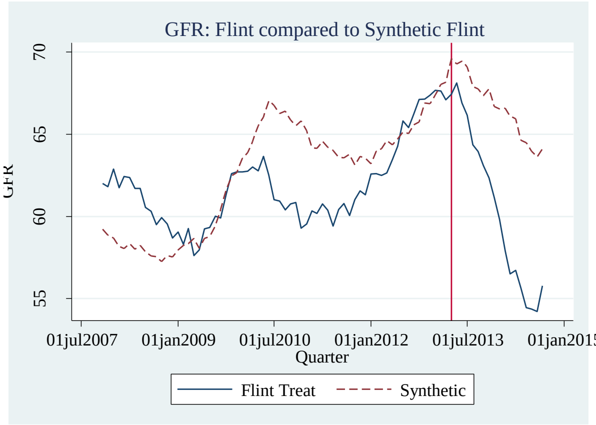
Appendix Table B1: Lead in Water on General Fertility Rate at the County Level

	(1)	(2)	(3)	(4)	(5)
Water (β_1)	-1.360***	-1.360***	-1.360***	-1.360***	-0.382
	(0.341)	(0.342)	(0.342)	(0.347)	(0.366)
Conception Month Fixed Effects		X	X	X	X
Conception Year Fixed Effects		X	X	X	X
City Fixed Effects			X	X	X
Conception Month into Year Fixed Effects				X	X
County Linear Time Trends					X
Observations	2,755	2,755	2,755	2,755	2,755
Counties	29	29	29	29	29
R-squared	0.009	0.122	0.257	0.296	0.315
Mean	51.77	51.77	51.77	51.77	51.77

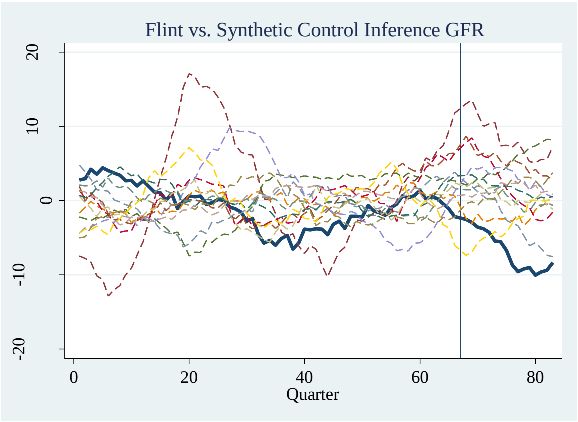
Notes: Robust standard errors clustered at the county level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001 . This table defines treatment as all of Genesee County and uses the 28 largest counties in Michigan as the comparison group.

Appendix Figure B6: Synthetic Control Results for General Fertility Rates, Adjusting for March 2008-2013 GFR

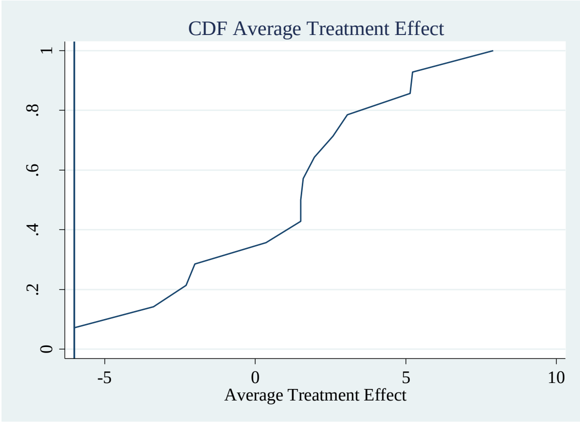
Panel A. Flint GFR compared to Synthetic Flint GFR



Panel B. Difference Between Each City and its Synthetic Counterpart



Panel C. Inference using Average Treatment Effect



Note: We include GFR for March 2008, March 2009, March 2010, March 2011, March 2012, and March 2013 in the Synthetic Control Model to create a better pre-treatment control group for Flint. The red vertical line in Panel A is at April 2013, which is the last conception date for which no affected birth rates are included in the moving average. The blue solid line in Panel B represents the difference between GFR in Flint and "synthetic Flint." The horizontal blue line in Panel C displays the average treatment effect. It is the largest average treatment effect compared to assigning all areas to treatment, suggesting statistical significance.

Appendix Figure B7: Randomization Inference Permutation Test

Appendix Table B2: Lead in Water on General Fertility Rate and Sex Ratios, Sample Changes

	General Fertility Rates			Sex Ratios		
	(1)	(2)	(3)	(4)	(5)	(6)
Main Results (N=1520)	-7.451*** (0.791)	-7.451*** (0.811) [0.004]	-5.682*** (0.603)	-0.0092*** (0.00262)	-0.0092*** (0.00268) [0.004]	-0.00121 (0.00411)
Before 9/2014 (N=1424)	-8.797*** (0.694)	-8.797*** (0.712) [0.004]	-6.900*** (0.585)	-0.00231 (0.00292)	-0.00231 (0.00300) [0.468]	0.00447 (0.00445)
Drop Outlier Cities (cities=14, N=1330)	-8.173*** (0.697)	-8.173*** (0.718) [0.004]	-5.549*** (0.678)	-0.0090** (0.00301)	-0.0090** (0.00310) [0.012]	-0.00352 (0.00409)
Conception Month Fixed Effects (FE)	X	X	X	X	X	X
Conception Year FE	X	X	X	X	X	X
City FE	X	X	X	X	X	X
Conception Month into Year FE		X	X		X	X
City Linear Time Trends			X			X

Notes: Robust standard errors clustered at the city level in parentheses. Brackets contain wild bootstrapped p-values for the most saturated models. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001

Appendix Table B3: Lead in Water on General Fertility Rate - ln(births) – All Cities

	(1)	(2)	(3)	(4)	(5)
Water (β_1)	-0.175***	-0.175***	-0.175***	-0.175***	-0.175***
	(0.0123)	(0.0124)	(0.0124)	(0.0128)	(0.0128)
Conception Month Fixed Effects		X	X	X	X
Conception Year Fixed Effects		X	X	X	X
City Fixed Effects			X	X	X
Conception Month into Year Fixed Effects				X	X
City Linear Time Trends					X
Observations	1,520	1,520	1,520	1,520	1,520
Counties & Flint	16	16	16	16	16
R-squared	0.001	0.007	0.980	0.981	0.981

Notes: Robust standard errors clustered at the city level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001 . Note that coefficients are in log points.

Appendix Table B4: Lead in Water on General Fertility Rate – Poisson (All Cities)

	(1)	(2)	(3)	(4)	(5)
Water (β_1)	-0.151***	-0.151***	-0.151***	-0.151***	-0.151***
	(0.0166)	(0.0166)	(0.0166)	(0.0166)	(0.0166)
Conception Month Fixed Effects		X	X	X	X
Conception Year Fixed Effects		X	X	X	X
City Fixed Effects			X	X	X
Conception Month into Year Fixed Effects				X	X
City Linear Time Trends					X
Observations	1,520	1,520	1,520	1,520	1,520
Counties & Flint	16	16	16	16	16
Pseudo R-squared	0.0092	0.0113	0.9553	0.9558	0.9558

Notes: Robust standard errors clustered at the city level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001 . Note that coefficients are in log points.[36]

[36] We do not include wild bootstrap standard errors in these analyses because these are Poisson regressions.

Appendix Table B5: Flint Compared Only to Genesee County GFR and Sex Ratio

	(1)	(2)	(3)	(4)	(5)	(6)
	GFR	GFR	GFR	Sex Ratio	Sex Ratio	Sex Ratio
Water (β_1)	-6.568**	-6.568***	-6.568***	-0.00711	-0.00711	-0.00711
	(2.071)	(1.918)	(1.918)	(0.0137)	(0.0136)	(0.0136)
Conception Month Fixed Effects		X	X		X	X
Conception Year Fixed Effects		X	X		X	X
County Fixed Effects			X			X
Observations	190	190	190	190	190	190
Counties & Flint	2	2	2	2	2	2
R-squared	0.604	0.695	0.285	0.015	0.123	0.114
Mean	62.28	62.28	62.28	0.510	0.510	0.510

Notes: †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001. This table defines treatment as Flint and uses the rest of Genesse County as the comparison group.

Appendix Table B6: Genesee County Except Flint as Treatment GFR and Sex Ratio

	(1)	(2)	(3)	(4)	(5)	(6)
	GFR	GFR	GFR	Sex Ratio	Sex Ratio	Sex Ratio
Water (β_1)	0.366	0.366	0.366	0.00476†	0.00476†	0.000387
	(0.341)	(0.342)	(0.342)	(0.00260)	(0.00261)	(0.00296)
Conception Month Fixed Effects		X	X		X	X
Conception Year Fixed Effects		X	X		X	X
County Fixed Effects			X			X
Observations	2,755	2,755	2,755	2,755	2,755	2,755
Counties & Flint	29	29	29	29	29	29
R-squared	0.002	0.116	0.257	0.000	0.004	0.004
Mean	48.08	48.08	48.08	0.510	0.510	0.510

Notes: Robust standard errors clustered at the county level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001 . This table defines treatment as the rest (i.e. parts that are not in Flint) of Genesee county and uses the 28 largest counties in Michigan as the comparison group.

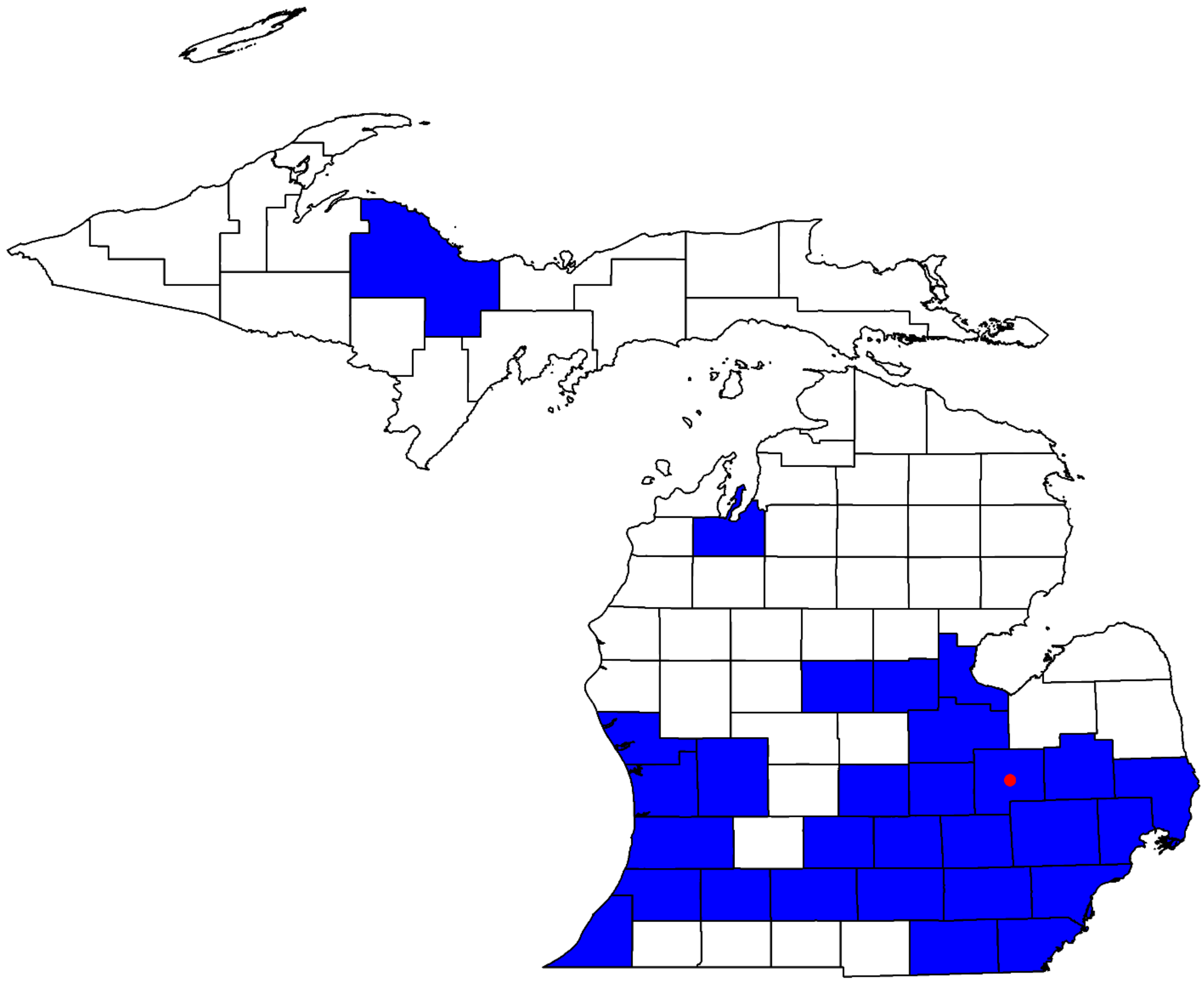
Appendix C: County Level Analysis

For the county level analysis, we consider Flint as the treatment unit, and then assign the rest of Genesee County as a rump control Genesee County with the remainder of the county's population.[37] Annual population data at the county level is only available from Census for high population counties, and so our main specification only uses those counties.[38]

[37] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_1YR_S0101&prodType=table

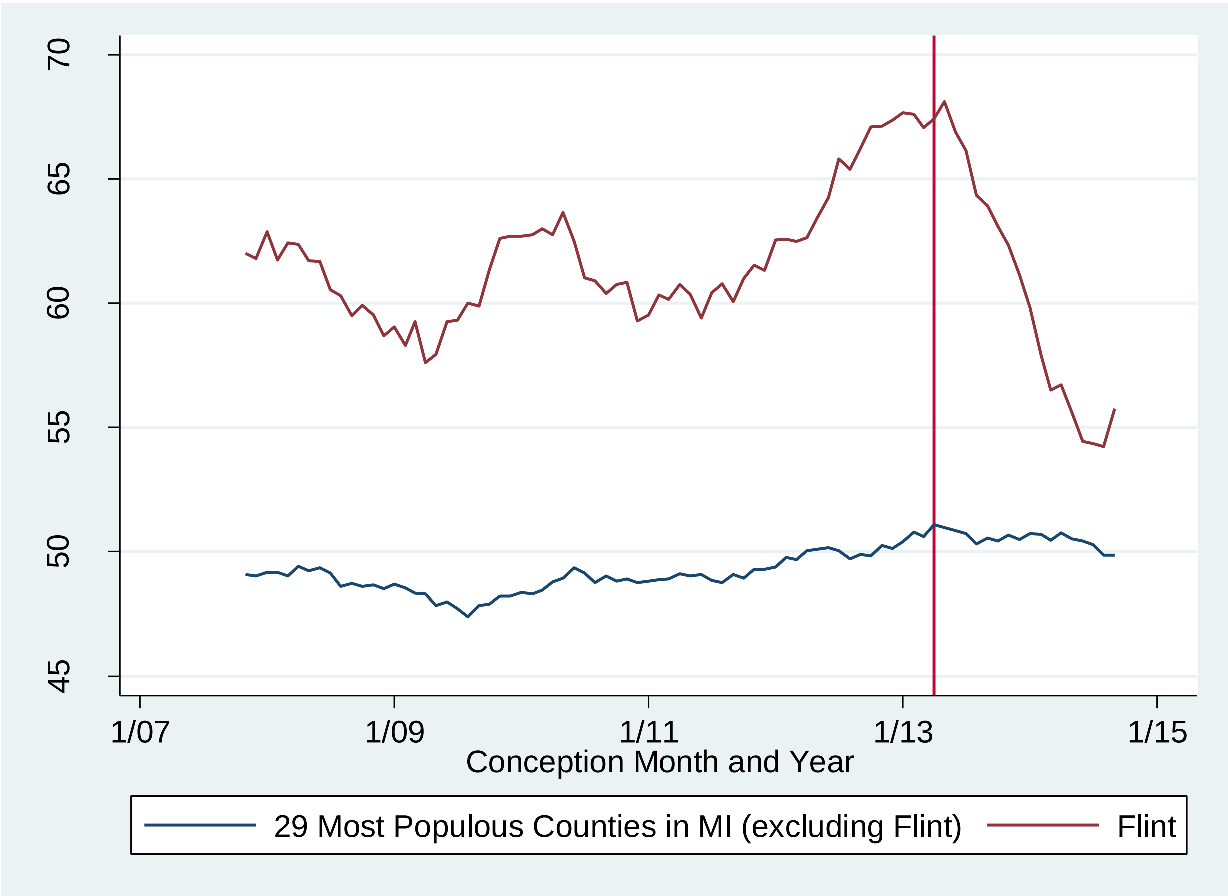
[38] I.e., Allegan County, Bay County, Berrien County, Calhoun County, Clinton County, Eaton County, Genesee County, Grand Traverse County, Ingham County, Isabella County, Jackson County, Kalamazoo County, Kent County, Lapeer County, Lenawee County, Livingston County, Macomb County, Marquette County, Midland County, Monroe County, Muskegon County, Oakland County, Ottawa County, Saginaw County, St. Clair County, Shiawassee County, Van Buren County, Washtenaw County, and Wayne County.

Appendix Figure C1: Comparison Counties



Notes: Blue counties are comparison counties. Flint is shown in red.

Appendix Figure C2: Moving Average Fertility Rate Over Time in Flint and Comparison Cities



Note: The red vertical line is at April 2013, which is the last conception date for which no affected birth rates are included in the moving average.

Appendix Table C1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)
	Non-Flint Births		Flint Births		
	Pre-Water Change (N=643,955)	Post-Water Change (N=137,808)	Pre-Water Change (N=10,620)	Post-Water Change (N=2,010)	Difference in Differences
Demographic variables:					
Mother's age (years)	27.78 (5.90)	28.32 (5.63)	24.66 (5.60)	25.17 (5.37)	-0.024
Mother no high school	0.141	0.115	0.294	0.271	0.003
Mother high school grad	0.249	0.240	0.317	0.343	0.035**
Mother some college	0.315	0.329	0.337	0.337	-0.014
Mother college grad	0.289	0.308	0.050	0.047	-0.023***
Outcome variables:					
General fertility rate	47.59 (7.96)	48.39 (8.27)	62.28 (6.81)	56.87 (6.76)	-6.22**
Male-Female Sex Ratio (percent male)	51.21 (0.50)	51.19 (0.63)	51.05 (4.59)	50.20 (3.06)	-0.82
Birth Weight (grams)	3,279 (616)	3,262 (627)	3,082 (632)	3,042 (651)	-23.7
Low Birth Weight	0.085 (0.28)	0.092 (0.29)	0.135 (0.34)	0.158 (0.37)	0.017†
Estimated gestational age (weeks)	38.56 (2.77)	38.48 (2.41)	38.08 (2.97)	37.89 (2.69)	-0.108
Gestational Growth (grams/week)	84.65 (14.44)	84.29 (14.27)	80.38 (14.33)	79.58 (14.48)	-0.437

Notes: For Columns (1)-(4), standard deviation for non-dummy variables in parenthesis. For Column (5), we present robust standard errors. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001

Appendix Table C2: Lead in Water on General Fertility Rate at the County Level

	(1)	(2)	(3)	(4)	(5)
Water (β_1)	-6.215*** (0.329)	-6.215*** (0.330)	-6.215*** (0.330)	-6.215*** (0.335)	-8.711*** (0.363)
Conception Month Fixed Effects		X	X	X	X
Conception Year Fixed Effects		X	X	X	X
City Fixed Effects			X	X	X
Conception Month into Year Fixed Effects				X	X
County Linear Time Trends					X

Notes: Robust standard errors clustered at the county level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001 .

Appendix Table C3: Lead in Water on Other Birth Outcomes by County

	(1)	(2)	(3)	(4)	(5)
Birth weight (grams)	-23.72†	-29.70*	-27.95†	-26.74†	-18.77
	(13.38)	(14.41)	(14.26)	(14.57)	(14.47)
Low Birth Weight	0.017†	0.019*	0.018*	0.018*	0.016†
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Gestational Age (weeks)	-0.108†	-0.126*	-0.119*	-0.119*	-0.104†
	(0.0574)	(0.0601)	(0.0593)	(0.0591)	(0.0594)
Gestational Growth (grams/week)	-0.437	-0.567†	-0.532†	-0.499	-0.316
	(0.301)	(0.322)	(0.320)	(0.329)	(0.328)
Census Tract Fixed Effects		X	X	X	X
Conception Month Fixed Effects			X	X	X
Conception Year Fixed Effects			X	X	X
Child Sex Control				X	X
Mom Controls					X

Notes: Robust standard errors clustered at the census tract level in parentheses. †p ‹ .10; *p ‹ .05; **p ‹ .01; ***p ‹ .001