George Robert Dimopoulos

Professor Hunt Allcott

ECON 177: Empirical Environmental Economics

15 December 2023

Analyzing Historic Superfund Cleanups and Infant Mortality:

Implications for Lake County, Indiana

Introduction

On September 6, 2023, the United States Environmental Protection Agency (EPA) announced the addition of three sites to the Superfund National Priorities List (NPL). One of these sites, Federated Metals Corp Whiting, is located in Hammond, Indiana. In 2017, the EPA concluded that lead and arsenic were "found at elevated concentrations" in the surrounding area of the facility—including but not limited to over seven-hundred residential properties (Pressley 2023). Long-term lead exposure, even in low levels, is linked to heightened "all-cause mortality, cardiovascular disease mortality, and ischemic heart disease mortality" levels; thus, the prospects of this superfund site cleanup have major implications for the wellbeing of future generations of Hammond residents (Lanphear et al. 2018).

The site's NPL placement therefore begs the question: do lead site Superfund cleanups lead to better long-term health outcomes for residents in the immediate surrounding area?

To measure the health impacts of lead site Superfund cleanups, we can conduct a staggered differences in differences (diff-in-diff) analysis on the county-wide changes of infant mortality due to the placement of a lead pollution site on the NPL. In particular, I used infant mortality data derived from the Center of Disease Control (CDC)'s CDC-WONDER datasets aggregated from 1999 to 2020. We are able to include Lake County, Indiana in the control group for the entire analysis: although Federated Metals Corp Whiting was just placed on the NPL in 2023, the site has not been in constant operation since 1983. Thus, the pollution in the site has not significantly increased during this time. Over forty-five other counties that have lead pollution sites already placed on the NPL comprise the rest of the data. Infant mortality rates from these counties are calculated in the control group for the years before their lead polluter was placed on the NPL, and are calculated in the treatment group for the years after the NPL placement.

I find that, following NPL placement, county-level infant mortality rates rose by .0167 percent relative to pre-NPL placement years. This shocking result may be somewhat explained by the percent change in both infant deaths and infant populations post-NPL placement, which both declined by .0192 percent and .0852 percent, respectively. These results may suggest other exogenous factors affecting the infant mortality rates post-NPL, such as the stigmatization of NPL placement, which could cause residents to move out of the affected counties. It is important to note, however, that none of the three results are statistically significant.

There are many limitations to this study. One prime example is the presence of other lead polluters near the Superfund sites that have not been placed on the NPL. In addition, the lack of granularity in the dataset may occlude the true effects of NPL placement on the immediate surrounding residents. There are also many other factors which contribute to infant mortality, such as inequal access to prenatal and postnatal care, which are not represented in this analysis. Finally, the (fortunate) rarity of child mortality inhibits statistical significance, as a single death constitutes a large percentage change from one year to the next.

Sections 1-7, respectively, present the literature review, overview of the purpose of NPL designation and Superfund site cleanups, dependent variable justification, data description, empirical strategy, results, discussion, and conclusion.

Literature Review

Superfund cleanups have broad implications on the socioeconomic vitality of the surrounding area. However, many of these economic benefits are not realized: Greenstone and Gallagher (2008) conclude that Superfund cleanups are "associated with economically small and statistically insignificant changes in residential property values, property rental rates, housing supply, total population, and types of individuals living near the sites". Although Greenstone and Gallagher (2008) show that housing-related factors do not change after the execution of a Superfund cleanup, they do not address their long-term health impacts: Klemick et al. (2020) find that Superfund cleanups "lowered the risk of elevated [blood lead level] for children living within 2 kilometers of lead-contaminated sites".

My approach closely follows Greenstone and Gallagher (2008). However, I implement their research process differently in four key ways. Firstly, instead of measuring property values, I measure infant mortality rates. By using as infant mortality rates an indicator for improvement in long-term health outcomes for communities affected by lead site Superfund cleanups, we are now able to capture benefits of Superfund cleanups that are not observed in housing markets. This also allows for the departure of using the Hedonic model to analyze the benefits of Superfund sites in favor of the staggered diff-in-diff analysis implemented in this paper. Furthermore, I focus my analysis on a different subset of Superfund cleanups that only include lead sites, as this allows for the clear observation of how such cleanups affect lead-related health complications. Finally, I use yearly data to add more granularity in measuring infant mortality before and after the treatment takes effect.

These four ways significantly differ from the Greenstone and Gallagher (2008) paper for a variety of reasons. Apart from choosing a different dependent variable, the primary difference between my analysis and the analysis conducted by Greenstone and Gallagher (2008) is their decision to use census data, which is measured once every ten years. Because they are restricted to choosing a greater timespan in order to observe a change before and after their 1982 treatment, they were required to install gradients for datapoints that were placed on the NPL after that time. My analysis avoids this issue, as a county enters the treatment group in the same year that their lead polluter is placed on the NPL.

My work largely looks to echo the conclusions provided by Klemick et al. (2020). However, it differs from their study by incorporating prenatal lead exposure in addition to postnatal lead exposure in the analysis. Furthermore, the data used by Klemick et al. (2020) relies on reports from blood lead level testing, which may be inconsistent in different states. Conversely, infant mortality data is provided by the fifty-seven vital statistics jurisdictions through the Vital Statistics Cooperative, which collects and publishes “official national statistics on… deaths… based on U.S. standard certificates” (National Center for Health Statistics 2022). This singular body for data collection prevents inconsistencies in reporting across states.

The Purpose of NPL Designation

Federated Metals Corp Whiting, a former metal smelting, refining, recovery, and recycling facility in operation from 1937 to 1983, smelted metals such as brass, lead, and zinc; the raw materials necessary for such processes “included scrap, non-ferrous metals and…

drosses and skimmings from other smelters" (Banik 2017). Although it is known that facilities such as this often "[emit] lead and other heavy metals from the facilities' stack, baghouse, and waste piles", it was not until 2017 that the EPA had conducted a soil contamination study on the land of the former facility (Environmental Protection Agency 2017).

Placement on the NPL makes a site eligible for remediation financed under the federal Superfund program. EPA Regional Administrator Debra Shore emphasized these benefits in a press release following the site's designation on the NPL: "Adding the Federated Metals site to the [NPL] showcases EPA's commitment to protecting human health… By cleaning up this site, EPA can support residents to help create healthy, thriving communities" (Pressley 2023).

The actions taken to clean up lead pollution sites after their designation on the NPL are almost immediate. The EPA conducts preliminary studies of a site to assess the severity of the pollution and to determine whether it should be placed on the NPL. As in the case of Federated Metals Corp Whiting, these studies may take multiple years to complete. During this time, the EPA does not only conduct an assessment of the current state of the site: it also develops plans for how to begin the cleanup process. The site's placement on the NPL greenlights the EPA to take remedial action in the site, which often begins mere months after. Thus, we are able to use NPL placement as a standardized proxy for the beginning of a Superfund site cleanup, as many of the sites are placed on the NPL and begin their cleanups in the same calendar year.

Dependent Variable Justification

A key to determining long-term health outcomes after a lead site Superfund cleanup is to decide how to quantify lead-related health concerns. Complications associated with lead exposure do not uniformly appear across humans' lifespans; therefore, it is difficult to measure the exact point in which a decrease in lead exposure leads to better health outcomes. However, one tragic complication—infant mortality—is significantly influenced by maternal lead exposure during pregnancy, as it is linked to "higher incidence of miscarriages and fetal death, even at blood lead elevations (≈5 μg/dL) once considered relatively low" (Edwards 2013). Furthermore, infant deaths resulting from low birthweight, sudden unexpected infant death (SUID), sudden infant death syndrome (SIDS), and other complications "are strongly related to air lead concentration at one year and at one month", indicating that postnatal exogenous lead exposure also contributes to infant mortality (Clay et al. 2022).

Data Description

The data used to conduct the analysis in this paper is found through the CDC WONDER Underlying Cause of Death, 1999-2020 dataset. In order to create a comprehensive dataset, I included all forty-five counties which had a lead polluter placed on the NPL between the years of 2000-2019 according to the county's Superfund Site Progress Profile. To denote when a lead polluter was placed on the NPL, I created a dataset consisting of the aforementioned counties and their respective placement year.

To generate the infant mortality data, I first grouped the results by county, and only included Infant Age Groups in the data compilation. Infant Age Groups are separated by "< 1 day", "1-6 days", "7-27 days", and "28-364 days"; I opted to include all groups, as infant lead exposure can cause mortality in all four cases. In addition, lead exposure is associated with increased deaths "from causes that science suggest may be related to lead, such as low birthweight, sudden unexplained infant death, and respiratory causes", among many others (Clay et al. 2023). Therefore, all infant deaths are taken into account. To ensure the dataset is a comprehensive representation of how lead affects infant mortality across an entire county, I included all races (including all "Hispanic Origins"), and all genders. I merged this dataset with the placement year dataset.

From the choices I made above, the output consists of the populations of each county's Infant Age Groups, along with their respective deaths, and the year in which each county's lead polluter was placed on the NPL. I then aggregated the deaths and population of every age group into singular respective "total death" and "total population" variables over the course of a single year. In order to account for differences in population between counties, I calculated the infant mortality percent of the population as opposed to using the total deaths. To do this, I summed each county's populations and deaths across their reported Infant Age Groups, and divided the summed deaths by the summed population to generate an infant mortality rate. To create the staggering required of the diff-in-diff analysis I wished to implement, I created a dummy variable, "treat", which equaled zero for the years before a county's lead polluter was placed on the NPL, and one for the years during and after NPL placement. A descriptive statistics table of the dataset can be viewed in Figure 1.

county	year	treat	deaths	population	mortality_rate
Length:388	Min. :1999	Min. :0.0000	Min. : 10.0	Min. : 1529	Min. :0.0830
Class :character	1st Qu.:2004	1st Qu.:0.0000	1st Qu.: 19.0	1st Qu.: 6679	1st Qu.:0.1788
Mode :character	Median :2009	Median :0.0000	Median : 36.0	Median : 14196	Median :0.2501
	Mean :2009	Mean :0.4485	Mean :105.1	Mean : 65909	Mean :0.2662
	3rd Qu.:2014	3rd Qu.:1.0000	3rd Qu.: 66.0	3rd Qu.: 34791	3rd Qu.:0.3244
	Max. :2020	Max. :1.0000	Max. :900.0	Max. :630032	Max. :0.9048

Figure 1: Descriptive statistics table.

There exist multiple limitations with the data used in this analysis. Firstly, not all counties have complete data for every Infant Age Group. The varying health complications from infant lead exposure do not significantly change across infant age groups: lead exposure "causes higher infant mortality in the first month and in the first year" (Clay et al. 2023). With this information, I decided that missing data for certain age groups would not be an issue, and went ahead with aggregating populations and deaths across Infant Age Groups in each county.

Another limitation is the timeframe of the data. There have been many sites that have been placed on the NPL since the inception of the Superfund program in 1980. However, the data is restricted to the years between 1999 and 2020; thus, I am unable to include counties which have sites placed on the NPL before or after these upper or lower bounds. This is to avoid the inclusion of sites that are already in the treatment of the start of the analysis.

Empirical Strategy

In order to analyze how the placement of Federated Metals Corp Whiting on the NPL will affect infant mortality rates in Lake County, Indiana, I conducted a diff-in-diff analysis using Lake County as the control group, and all other counties which have had a lead polluter placed on the NPL beforehand as originally part of the control group, as well. The other counties then became part of the treatment group after their site was placed on the NPL. It is important to note that the counties in the treatment group had lead polluters placed on the NPL at different times; thus, it will be best to utilize a staggered diff-in-diff analysis. To do so, the regression equation is as follows:

$$\mathbf{Mortality}_{it} = \beta_0 + \beta_1 \mathbf{Treat}_{it} + \delta_i + \gamma_t + \varepsilon_{it}$$

This regression shows how the infant mortality rate for county i in year t ($\mathbf{Mortality}_{it}$) is dependent on a dummy variable of whether a county is from the pre-condition cohort or the post-condition cohort in a given year ($\mathbf{Treat}_{it}$), the county fixed effects (δ_i), the time fixed effects (γ_t), and an error term (ε_{it}). If this coefficient is negative, a lead polluter placement on the NPL is associated with a decrease in infant mortality rate in comparison to the control group.

Conversely, a positive coefficient indicates an association between a lead polluter placement on the NPL and an increase in infant mortality rate in comparison to the control group.

Infant mortality rate is based on both the number of infant deaths and the total population of infant groups. Thus, in order to develop a complete picture regarding changes to the infant mortality rate, it was necessary to conduct two additional analyses:

$$\mathbf{\log(Death_{\mathit{it}}) = \beta_0 + \beta_1 Treat_{\mathit{it}} + \delta_{\mathit{i}} + \gamma_{\mathit{t}} + \varepsilon_{\mathit{it}}} \, ,$$

which is the same regression but accounts for the percent change in infant death, **log(Death$_{it}$)**, due to NPL placement, and:

$$\mathbf{\log(Population_{\mathit{it}}) = \beta_0 + \beta_1 Treat_{\mathit{it}} + \delta_{\mathit{i}} + \gamma_{\mathit{t}} + \varepsilon_{\mathit{it}}} \, ,$$

which is the same regression but accounts for the percent change in infant group population, **log(Population$_{it}$)**, due to NPL placement.

Results

	mortality_model		death_model		population_model
Dependent Var.:	mortality_rate	Dependent Var.:	log(deaths)	Dependent Var.:	log(population)
treat	0.0167 (0.0124)	treat	-0.0192 (0.0558)	treat	-0.0582 (0.0721)
Fixed-Effects:	---------------	Fixed-Effects:	----------------	Fixed-Effects:	----------------
county	Yes	county	Yes	county	Yes
year	Yes	year	Yes	year	Yes
S.E.: Clustered	by: county	S.E.: Clustered	by: county	S.E.: Clustered	by: county
Observations	388	Observations	388	Observations	388
R2	0.82243	R2	0.93947	R2	0.96014
Within R2	0.00801	Within R2	0.00036	Within R2	0.00311
---		---		---	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 2: Results of the staggered diff-in-diff analysis, in order as presented in the prior section.

Surprisingly, the placement of a lead polluter on the NPL is associated with a .0167% increase in child mortality rate, as seen in Figure 2. However, this result alone may not provide an all-encompassing explanation: the second and third regression results show that the placement of a lead polluter on the NPL is associated with both a .0192% decrease in infant deaths and a .0582% decrease in infant population, respectively.

Discussion

Above all else, it is important to note that none of these values are statistically significant. Therefore, the results do not provide enough evidence to reject the null hypothesis. This could be due to a variety of factors, such as incredibly low infant mortality rates present within these counties both before and after NPL placement, or a lack of higher granularity within the data, among other concerns. However, it is nonetheless important to explain the relationships between infant mortality rate, percent change in infant death, and percent change in infant population with NPL placement that are outlined in this analysis.

The results of the first regression show a shocking departure from the expected impact of NPL placement, as cleaning these sites decreases lead pollution in a county. However, perhaps the subsequent two regressions are able to elaborate on why it may seem as though the placement of a lead polluter on the NPL would lead to an increase in infant mortality rate. The placement of a lead polluter on the NPL lead to decreases in both the percent change of infant deaths and the percent change of population; however, the rate of change for infant population was more than that of infant deaths, which explains the increase in infant mortality rate.

There are several possible explanations for this. One possible explanation is that, although NPL placement marks the beginning of cleaning a lead pollution site, the negative stigma a neighborhood receives by containing a Superfund site is large. Upon learning that their neighborhood contains a Superfund site, some current homeowners “may simply be unwilling to continue living in their homes, and likewise, potential buyers will be unwilling to consider buying a home in that community" (Messer et al. 2006). This has led to declining home property values due to "the notion of progressive stigmatization of the site": relative property values of homes near the Superfund sites sampled in the study have "just 60.5%, … 76.5%, … [and] 86.0% of their original value", respectively (Messer et al. 2006). This stigmatization is not only realized in academic studies: officials of Silverton, Colorado, pushed against the EPA's placement of Gold King mine on the NPL in 2015, as they believed such placement "would kill a hoped-for revival in the area's mining industry" (Moreno and Knickmeyer 2015). Thus, stigmatization of surrounding neighborhoods may account for the decrease in relative infant population size after NPL placement, as residents who were originally unaware of the state of their environment now have a desire to leave.

Stigmatization of counties with polluters placed on the NPL may play a large role in not only the change in infant population, but the mortality rate, as well. Receiving placement on the NPL does not signify immediate erasure of pollution: although cleanups often begin very shortly after NPL designation, the "cleanup processes have taken over 10 years" on average (US Cong. House of Rep. Comm. Government Reform and Oversite). Thus, the health impacts of the pollution may still be present years after NPL placement. In addition, residents within close proximity of a Superfund site are "more low-income, [are] more indigenous and people of color" than other residents within their county; it is therefore more difficult for these individuals to move out in comparison to other residents (Environmental Protection Agency 2023).

Therefore, one possible explanation of the increase in mortality rate is that higher-income residents who do not experience as much lead pollution as people closer to the Superfund site elect to move out of their counties at a higher rate than lower-income residents closer to the

Superfund sites. This leaves a lower-income population—which is more susceptible to inheriting negative externalities such as infant mortality due to lead pollution—behind, thus inflating the infant mortality rate.

Conclusion

In an attempt to measure the benefits and detriments of the placement of Federated Metals Corp Whiting on the NPL, this analysis aimed to investigate the impact of historic NPL placement of lead polluting sites on county-level infant mortality rates. Although not statistically significant, the results revealed an unexpected uptick in the infant mortality rate post-NPL placement. This counterintuitive outcome was explained by the discovery of a decrease in percent change in infant deaths coupled with a more substantial decrease in percent change in infant population, which resulted in an overall increase in the infant mortality rate. One possible explanation for this is the stigmatization of a county post-NPL placement: residents may be more driven to leave their county, as they are now aware of the pollutants present in their community.

This study gives a nod to the intricate interplay of environmental policies and psychology, both of which shape health outcomes in a county. Although there are many limitations to this study, such as lack of granular data and possible omitted variable bias, this analysis nonetheless provides valuable insights into the dynamics surrounding unintended consequences of environmental remediation efforts. Future papers should address the lack of granularity in this paper by perhaps evaluating the effects of NPL placement on census tract-level infant mortality rate or on another indicator of human health. Furthermore, in order to examine the role of stigmatization and income inequality suggested but not conclusively determined in this study, future papers may look to analyze how stigmatization affects residents of different income groups and proximities to the Superfund site.

Works Cited

Banik, Jerry. “Tragedy at Federated Metals in 1949.” *Whiting-Robertsdale Historical Society*, 2022, www.wrhistoricalsociety.com/tragedy-at-federated-metals-in-1949.

Clay, Karen, et al. “Airborne Lead Pollution and Infant Mortality.” *IZA Discussion Paper Series*, June 2022.

Clay, Karen, et al. “The Impact of Lead Exposure on Fertility, Infant Mortality, and Infant Birth Outcomes”, June 2023, https://doi.org/10.3386/w31379.

Edwards, Marc. “Fetal death and reduced birth rates associated with exposure to lead-contaminated drinking water.” *Environmental Science & Technology*, vol. 48, no. 1, 2013, pp. 739–746, https://doi.org/10.1021/es4034952.

Environmental Protection Agency. “Federated Metals Corp Whiting Site Profile.” *EPA*, Environmental Protection Agency, 20 Oct. 2017, cumulis.epa.gov/supercpad/SiteProfiles/index.cfm?fuseaction=second.Cleanup&id=0501275#Status.

Environmental Protection Agency. “Supporting Environmental Justice at Superfund Sites.” *EPA*, 30 Oct. 2023, www.epa.gov/superfund/supporting-environmental-justice-superfund-sites.

Greenstone, Michael, and Justin Gallagher. “Does hazardous waste matter? evidence from the Housing Market and the Superfund program.” *SSRN Electronic Journal*, 2008, https://doi.org/10.2139/ssrn.840207.

Klemick, Heather, et al. “Superfund cleanups and children’s lead exposure.” *Journal of Environmental Economics and Management*, vol. 100, 2020, p. 102289, https://doi.org/10.1016/j.jeem.2019.102289.

Lanphear, Bruce P, et al. “Low-level lead exposure and mortality in US adults: A population-based Cohort Study.” *The Lancet Public Health*, vol. 3, no. 4, 2018, https://doi.org/10.1016/s2468-2667(18)30025-2.

Messer, Kent D., et al. “Can stigma explain large property value losses? The Psychology and Economics of Superfund.” *Environmental & Resource Economics*, vol. 33, no. 3, 2006, pp. 299–324, https://doi.org/10.1007/s10640-005-3609-x.

Moreno, Ivan, and Ellen Knickmeyer. “Fearing Stigma, Colorado Contested Superfund Status for Mine.” *CBS News*, CBS Interactive, 12 Aug. 2015, www.cbsnews.com/colorado/news/fearing-stigma-colorado-contested-superfund-status-for-mine/.

National Center for Health Statistics. “National Vital Statistics System (NVSS) - Health, United States.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 12 Aug. 2022, www.cdc.gov/nchs/hus/sources-definitions/nvss.htm#:~:text=NVSS%20collects%20and%20publishes%20official,Mortality%E2%80%94are%20detailed%20as%20follows.

Pressley, Mary. “EPA Adds Federated Metals Corp Whiting Site in Hammond, Indiana To ...” *EPA*, Environmental Protection Agency, 6 Sept. 2023, www.epa.gov/newsreleases/epa-adds-federated-metals-corp-whiting-site-hammond-indiana-superfund-national.

United States, Congress, House of Representatives, Subcommittee on National Economic Growth, Natural Resources, and Regulatory Affairs, Committee on Government Reform and Oversight. *Times to Assess and Clean Up Hazardous Waste Sites Exceed Program Goals*. Federal Testimony February 1997, General Accounting Office, 1997. 105th Congress, 1st session.