

Article

## Lookout Below! The Downstream Infant Health Impacts of Mining Disasters

Benjamin A. Jones

Department of Economics, University of New Mexico, Albuquerque, NM 87106, USA; bajones@unm.edu; Tel.: +1-505-277-5304.

---

### ABSTRACT

Limited causal evidence exists on the acute infant health effects of environmental disasters, especially among mining accidents in developed countries. To address this gap, this paper exploits a natural experiment provided by the 2015 US EPA's accidental release of mine waste water into nearby river systems at the Gold King Mine in Colorado, USA to investigate the acute birth weight and gestation effects of a short-duration shock to water quality. Using both the synthetic control method and a difference-in-differences approach and information on the precise timing of the spill, I show that instances of low birth weight and premature births increased by 0.6 and 0.71 percentage points, or 7.7% and 9.1%, respectively, in downstream counties during the year of and year after the spill. These results highlight the policy importance of funding sustainable approaches for cleaning-up legacy metal mining operations in order to minimize their impacts on adjacent populations.

**KEYWORDS:** mine wastewater; river contamination; infant health; environmental disasters; sustainable coexistence

---

### ABBREVIATIONS

US EPA, US Environmental Protection Agency; SCM, synthetic control method; DID, difference-in-differences

### INTRODUCTION

There is a sizable literature linking water pollution to human health outcomes (e.g., [1–3]). However, much of the focus and attention of this literature has gravitated towards chronic health impacts associated with prolonged exposure to contaminated water supplies (e.g., [4,5]) and weather-related acute impacts (e.g., [6,7]), in predominately developing counties (e.g., [8–10]). While such works make novel contributions in their own right, concerns were recently raised by The *Lancet* Commission on Pollution and Health that insufficient attention is being given to the water quality health impacts associated with industrial and mining industries, including accidental releases [11]. Even in developed countries, it has recently been shown that mining water quality accidents can have meaningful health and economic impacts on local communities [12].

### Open Access

Received: 03 Feb 2026

Accepted: 10 Apr 2026

Published: 15 Apr 2026

Copyright © 2026 by the author. Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of Creative Commons Attribution 4.0 International License.

This paper contributes new causal evidence to the nascent literature on mining disaster related acute health impacts of water pollution by exploiting a natural experiment produced by the US Environmental Protection Agency's (US EPA) accidental release of approximately 3 million gallons of metal-rich mine waste water at the Gold King Mine in Colorado, USA in 2015. The deluge of metals and toxic compounds from the spill flowed into the nearby Animas River, which is a major source of municipal drinking water and recreation throughout southwestern Colorado and northwestern New Mexico. I investigate how releases of mine waste by-products such as arsenic, lead, and mercury into a river system affects infant health. Additionally, I investigate two states in the Western US where there is a historical legacy of abandoned mines that are in close proximity to major river systems (or their tributaries) such as the Colorado and Rio Grande rivers. This increases the generalizability of my study to other parts of the US where abandoned mines also exist (e.g., California, the mountain West, Appalachia, etc.).

There are several reasons why sustainability practitioners may be interested in this work. First, large population growth in former or current mining communities or along river systems that flow downstream from mining communities (e.g., along the Front Range of Colorado, in the Wasatch Range of Utah, etc.) means that there are increased opportunities for mine-related disasters to adversely affect individuals, families, and communities (Between 2000-2018, population growth rates in traditionally mining heavy states have outpaced the rest of the US: Utah (14.4%), Colorado (13.2%), Nevada (12.4%), Arizona, (12.2%), California (6.2%). Nevada, Arizona, and California are among the top five mineral producing states in 2016 [13]. Most of these states also rely on surface water supplies for municipal water, which mining waste water disasters can potentially contaminate.). Thus, there is a need to learn to sustainably co-exist with former and current mining operations since safe and reliable water supplies are necessary inputs into the economy [14]. Second, there are increasing concerns that early-in-life exposure to pollutants and contaminants is causally linked to long-term labor market and human capital outcomes—the “fetal origins hypothesis” [15]. Hence, *in utero* or neonatal exposure to contaminated water from an environmental disaster could therefore have adverse consequences on many later-in-life economic variables of interest, reducing society's ability to sustainably co-exist with abandoned mines. Finally, given the deteriorating physical support infrastructure in many abandoned mines across the US, there is a need to better understand the potential health impacts associated with water breaches at these sites. Knowing about the magnitude of potential health impacts would be helpful when setting mine clean-up priorities and in establishing appropriate financing of such ventures along sustainability dimensions.

My primary research question is: “To what extent did the temporary shock to river water quality caused by the release of mine waste water

from the Gold King Mine adversely impact infant health in downstream counties in Colorado and New Mexico during the months immediately following the spill?” Modern causal inference techniques are used to address this question, and I focus my attention on impacts to infant health outcomes. To increase the credibility of my empirical strategy, I employ the synthetic control method (SCM) of [16,17]. Importantly, synthetic control and treated counties are match-weighted to have not only similar pre-spill infant health outcomes, but are also matched on many socio-demographic and flexible weather characteristics that are known to influence infant health. In addition to the SCM model, a more traditional difference-in-differences (DID) model is also estimated, which allows me to also exploit the rich, individual birth data available to me.

Since I have data on the precise timing of the spill and subsequent movement of the pollutants downstream, I focus my analysis on a short period of time right after the spill (up to 16 months post-spill). For identification purposes, this increases the likelihood that other exogenous confounding factors affecting infant health are constant over the analysis period. However, several falsification and robustness checks are also performed in order to boost the causal interpretation of the results: (i) an internal validity test shows no significant changes in mother and infant characteristics, county socio-demographic indicators, or weather outcomes concurrent with the river water quality shock; (ii) the timing of impacts found in my empirical models are greatest during the first few months after the spill and then diminish to insignificance within one calendar year, consistent with an acute health impact environmental disaster story; (iii) falsification tests show no impact on infant health prior to the actual date of the spill, and no impact is observed on birth outcomes known to be unaffected by environmental exposures; (iv) results are robust to alternative specifications of the SCM donor pool, including alternative geographic and population-based specifications, and; (v) results are robust to alternative measures of birth outcomes and whether I look at singleton or multiple births.

This work contributes to the sustainability and environmental health literature in at least two ways. First, it provides new causal evidence that short-duration shocks to water quality can adversely affect infant health, which adds to the literature on water pollution (e.g., [18,19]) and water pollution and health (e.g., [8,20,21]). I improve upon the extant literature by showing adverse outcomes in a developed country where water quality concerns are often thought to have more marginal impacts on health (due to stringent Federal regulations, implementation of advanced water treatment technology, widespread availability of bottled water, etc.) and by looking at multiple measures of adverse health outcomes along both the intensive and extensive margins. Second, this work adds a new dimension to our understanding of the economic costs of mining-related environmental disasters, thereby complementing the small set of economic studies in this area (e.g., [12,22]). Importantly, this is the first

work to causally link *mine waste water* (as opposed to other types of water pollution) to downstream infant health outcomes.

### **Environmental Disasters, Water Quality, and Human Health**

There is an important distinction in the literature between the impacts of industrial environmental disasters on water quality and health and the general literature on water pollution and health outcomes. Disaster events result in a one-time shock to water quality that may cause a temporary shock to acute health outcomes. By contrast, continuous prolonged exposure to contaminated or polluted water supplies may result in different health outcomes, likely of a more chronic nature (e.g., [4,8]). The focus of this work is specifically on disaster impacts and there are surprisingly few published studies on the health impacts of shocks to water quality caused by industrial environmental disasters. Even in the epidemiology literature, few disaster-related papers have been published on water and health, and of these, an even smaller number are causally interpretable (see, for example, the discussions in [23,24]). This is troubling given that human-caused industrial environmental disasters can be significant sources of toxic releases into waterways and into the environment more generally. Recent industrial water disaster examples in the US, include: the 2014 Duke Energy spill of 27 million gallons of coal ash into the Dan River in North Carolina; the 2014 Freedom Industries spill of approximately 10,000 gallons of 4-methylcyclohexane methanol (MCHM) into the Elk River in West Virginia, and; the 2010 BP Deepwater Horizon blowout that resulted in a discharge of 210 million gallons of oil into the Gulf of Mexico.

Studies in this area have generally found that water quality shocks after environmental disasters result in acute physical and mental health impacts, from the relatively benign—headaches, scratchy throats, watery eyes, and worse birth outcomes [12,25]—to the more severe—cancer, mercury poisoning, PTSD, generalized anxiety disorder, and even mortality [26–28]. Exposure pathways include direct ingestion of contaminated water, indirect consumption through contaminated food products, or through recreational exposure in polluted waterways. However, most prior work in this area, including most of the studies referenced above, are from studies of oil spills, which is just one subset of water quality-related disaster events.

While health may be impacted by many types of environmental disasters, there is a particular concern about mining and industrial accidents, which tend to be major sources of environmental releases of compounds such as arsenic, lead, and mercury; compounds which extensive prior evidence has shown can adversely affect human health (e.g., [29,30]). The *Lancet* Commission on Pollution and Health recently echoed this concern by specifically noting the meaningful role that mining and industrial releases of toxins can have on health outcomes, while simultaneously highlighting a dearth of research in this area [11]. To

sustainability practitioners, there is therefore a concern that the health externalities of mining and industrial disasters are not fully understood due to a lack of credible empirical analyses in this area. One policy consequence of this knowledge gap is that damage assessments of mining and industrial disasters are likely to be downward biased due to an omission of health costs directly resulting from exposure to contaminated water supplies. This could negatively affect efforts to sustainably co-exist with current and former mining sites.

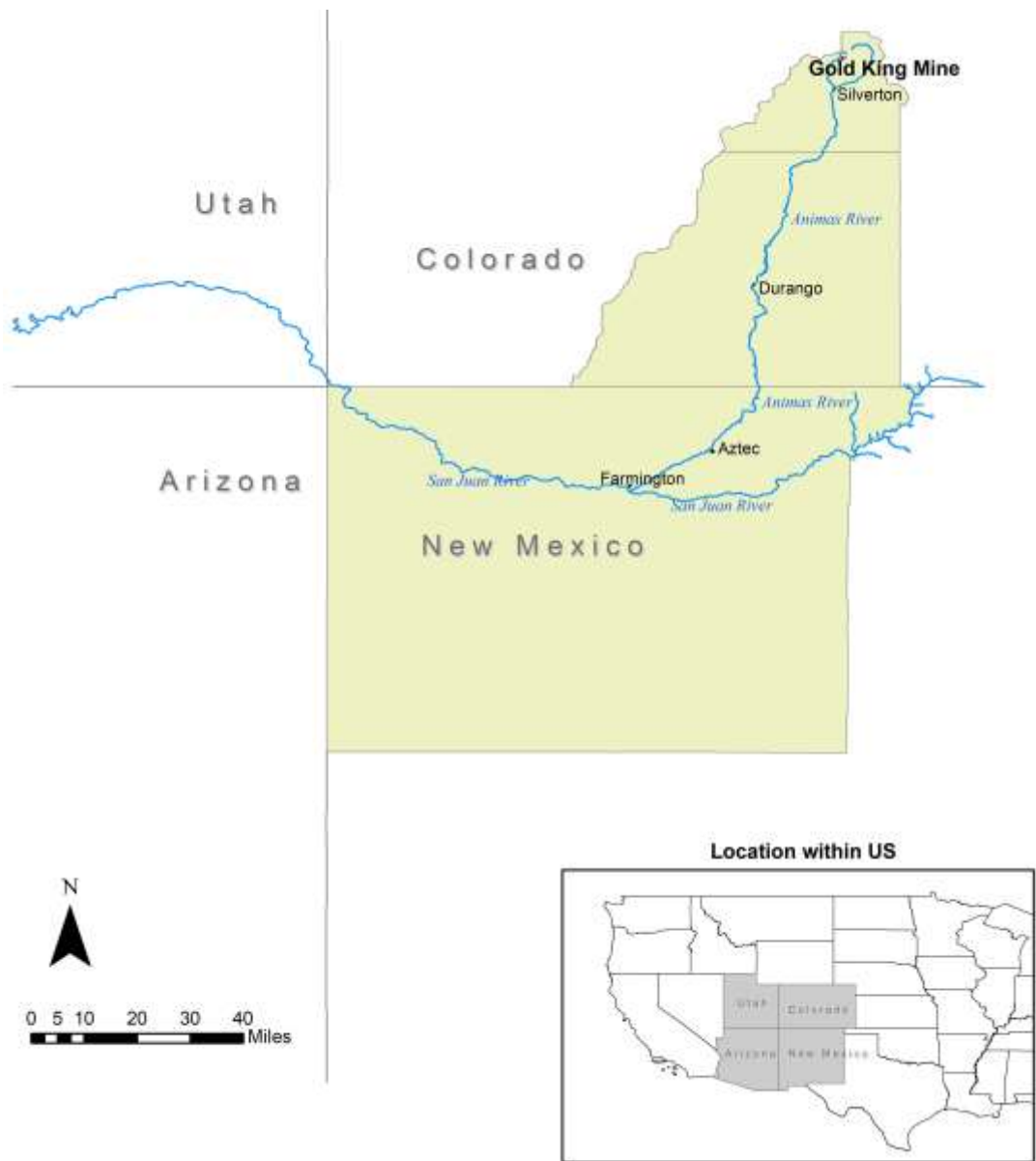
Of course, this is not the first paper to propose that the sustainability and environmental health literatures be extended to include investigations of mining water quality accidents. In [12], the synthetic control method was used to study the adverse health impacts and economic costs of a 2014 mining accident that resulted in the chemical MCHM being discharged into the Elk River, West Virginia. The authors found significant deteriorations to post-birthing infant outcomes as measured by 5-minute Apgar Scores after the spill, but did not observe significant changes in either birth weight or length of gestation. To the best of my knowledge, [12] are the first to rigorously apply modern causal inference techniques to study health outcomes in the aftermath of a mining water quality accident.

In this paper, I add to the literature by studying a mine waste water disaster, rather than a chemical spill, where large quantities of metals such as arsenic, lead, mercury, manganese, chromium, etc. were released into a nearby river system. Since there is strong prior evidence that these compounds can adversely impact health, especially among children and infants, I hypothesize that birth weight and gestation length will be affected, at least temporarily, by the mine water accident investigated here. Furthermore, my results are likely more generalizable than prior work because the source of our environmental disaster is an accidental release of waste water from an abandoned underground mine. Given that hundreds, if not thousands of similarly abandoned mines exist across the US, many that may have similar waste water build-ups as found at the Gold King Mine [31], the possibility of future accidents that are similar in size and scope cannot be ruled out (Federal investigators from the US Bureau of Reclamation found that conditions that led to the Gold King disaster “are not isolated or unique and in fact are surprisingly prevalent” [32]). In light of this, the results of this work are generally informative of the human health externalities associated with potential mine waste water disasters in places across the Western US (e.g., Colorado, California, Utah, Nevada) and in Appalachia (e.g., West Virginia, Virginia, North Carolina).

### **The Gold King Mine Spill**

On 2015 Aug 5, during routine US EPA reclamation work at the abandoned Gold King Mine near Silverton, CO, workers accidentally loosened some of the rocky material in one of the main mine tunnels, thereby allowing pressurized water that had built up to flow out of the

mine entrance and into nearby Cement Creek, which is a tributary of the Animas River. Over a several days period, approximately 3 million gallons of mine waste water were released into the Animas River, causing the river to turn an eerily bright orange color. The Animas River serves as a major source of drinking water for the downstream communities of Silverton, CO (2010 population of 645), Durango, CO (16,815 pop.), Aztec, NM (6765 pop.), and Farmington, NM (46,046 pop.)—see Figure 1. Near Farmington, NM, the Animas merges with the San Juan River, which is a major tributary of the Colorado River.



**Figure 1.** Map of the Gold King Mine Area and the Animas River. Shaded areas are the three downstream treated counties of (from north to south): San Juan, CO, La Plata, CO, and San Juan, NM.

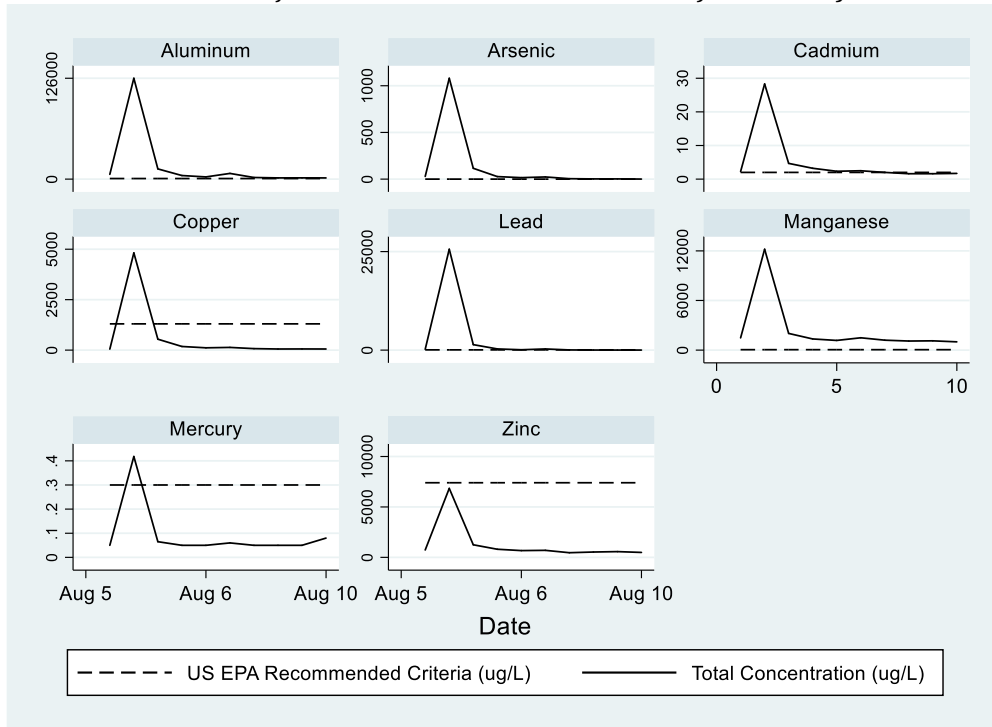
The waste water released from the Gold King Mine contained a diverse array of heavy metals and other compounds including: aluminum, arsenic, barium, beryllium, cadmium, chromium, copper, iron, lead, manganese,

mercury, silver, vanadium, and zinc (By weight, the top ten most prevalent metals in the mine waste water were, in order from highest to lowest: iron, aluminum, calcium, magnesium, potassium, lead, magnesium, zinc, copper, and sodium [33]). Using US EPA water quality data at two Animas River monitoring sites located downstream from the mine entrance: (i) the A72 site, located near Silverton, CO, and; (ii) the Animas-Rotary Park site, located in Durango, CO, I have plotted measured concentrations of various water pollutants in Figure 2, along with US EPA recommended water quality criteria for human health, or, if health criteria are unavailable, for aquatic life [34]. Concentrations are plotted immediately before and after the main plume of waste water flowed through each monitoring site: August 5–10 for site #1 (A72) and August 6–8 for site #2 (Animas-Rotary Park). As can be seen in Figure 2, pollutant concentrations during the main plume pass-by event exceeded US EPA recommendations for aluminum, arsenic, cadmium, lead, and manganese (at both sites), and copper and mercury at site #1. At peak impact, the US EPA reported that arsenic levels in the river water running through Durango were approximately 300 times the normal level, and lead levels were approximately 3500 times the normal level [35].

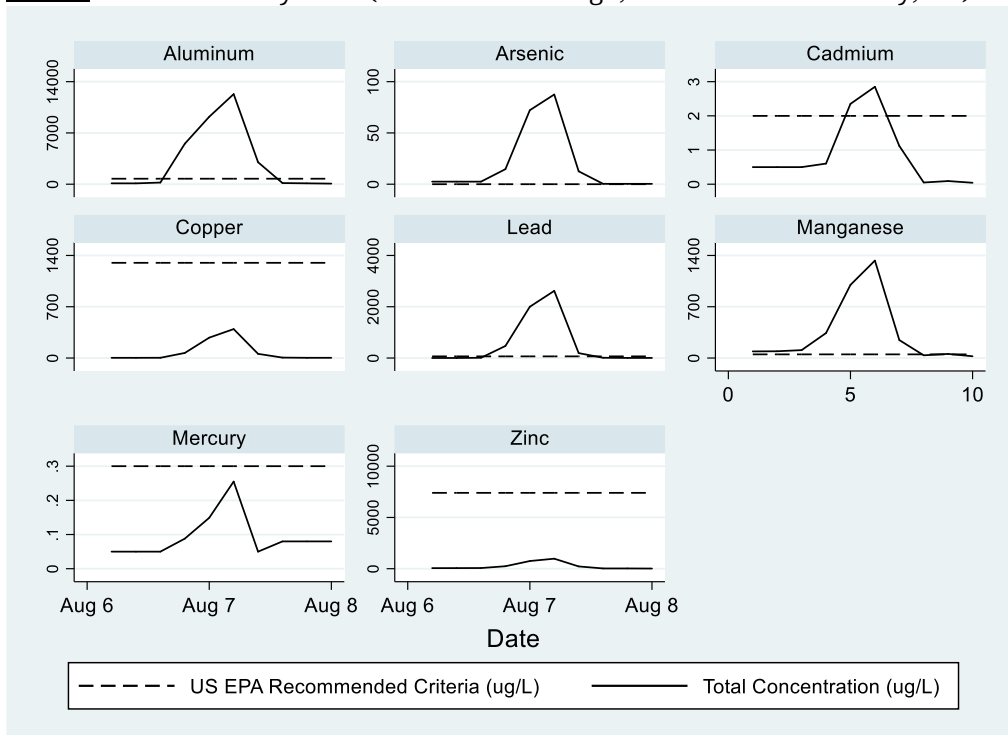
Coincidentally, the mine accident occurred during the peak summer recreation season, and the Animas is a large source of recreational and fishing opportunities in the region; over 58,000 user days were logged on the Animas River in 2005, making it the second most recreationally active river in all of Colorado [36]. While some local residents of southern Colorado and northern New Mexico would have been recreating in the Animas River during the spill period (this is a local destination hotspot in the summer, including by this author), many non-local residents would have also likely visited the area during the spill. The birth data I use only allows me to capture localized effects and not impacts from pregnant mothers who may have visited the animas and then returned home (e.g., to a different city, state, etc.). to the extent that these non-local mothers are in my counterfactual group, this would bias us against observing an impact of the spill on infant health (because some mothers in the counterfactual would have undergone degrees of treatment). Thus, my results may potentially be downward biased as a result. However, it is unknown how many non-local residents (or pregnant mothers, to be specific) were in the area during the spill event. A few days after the spill, public officials closed off access to the river in and around Durango and Silverton, CO (lasting until August 14) and residents with wells were advised to have their water tested before any use due to possible contamination [37]. However, photographs appearing in *The New York Times* showed rafters and kayakers still recreating in the bright orange-colored river even after the spill [37]. The City of Durango also preemptively stopped pumping water from the Animas River into their municipal supply, in order to prevent possible contamination, and people were urged to not catch or eat any fish from the river. However, it is

unknown how effective these public advisories were at preventing possible exposure to contaminants.

**Site #1: A72 (Located just south of Silverton, CO in San Juan county, CO)**



**Site #2: Animas-Rotary Park (located in Durango, CO in LA Plata County, CO)**



**Figure 2.** Concentrations of Water Pollution in the Animas River Before and After the Mine Spill Source: [34]. Notes: recommended criteria from US EPA National Recommended Water Quality Criteria—Human Health Criteria Table. Exceptions are for aluminum, lead, and cadmium where the Aquatic Life Criteria Table is used instead (due to the absence of a human health criterion for these pollutants).

During and immediately after the spill event, there were widespread concerns about adverse health impacts resulting from direct contact with contaminated water or from consuming crops that were irrigated by contaminated river water (e.g., [37,38]). Concerns were especially pronounced within the Navajo Nation, the largest Native American reservation in the US, who rely heavily on river water located downstream from the Gold King Mine as a source of drinking water, irrigation, and recreation [39]. Despite assurances from the Colorado Department of Public Health and Environment, among other government agencies, that any health impacts caused by the spill would be “low” [40], this is actually an open empirical question.

Of course not all residents in areas downstream from the Gold King Mine would have been necessarily exposed to contaminated water, however, based on the above discussion, there are several potential exposure pathways, including: (i) from recreation in the Animas River and its connected waterbodies; (ii) from the consumption of any marine life in the Animas or connected waterbodies, and; (iii) from possible contamination of municipal water supplies or domestic wells in the region.

There is particular concern about acute health effects immediately after the spill, when water pollution concentrations were elevated above baseline levels. Given that *in utero* and infant health are especially sensitive to small changes in exposure to environmental pollutants, it is possible that the spill may have resulted in temporary adverse impacts among this sensitive population. A 2018 joint Federal, state, and county agency report was released that found no long-term impacts of the Gold King Mine spill on water quality, and, moreover, the report states (without providing empirical justification) that the mine waste water poses “no risk to human health if common-sense precautions are followed”. However, this study provides no information regarding acute health impacts, which, on the basis of the prior environmental disaster literature [23], we would anticipate would be the primary source of any health impacts observed (rather than long-term or chronic health effects). The biological mechanisms linking heavy metals to fetal development are not precisely known, though there are several plausible hypotheses. For instance, arsenic in maternal blood can cross the placenta barrier and reach the fetus during the second and third trimesters. This can lead to decreased expression of adipose tissue-derived ENPP2, a phospholipase that regulates adipose tissue growth, thereby affecting fetal fat development and hence birth weight [41]. Furthermore, arsenic, even at low exposure levels, increases general inflammation in newborns exposed *in utero*, which can affect gestation length through cervical ripening and rupture of membranes [42]. Arsenic may also have direct toxic effects on DNA repair enzymes, which could affect fetal growth. Additionally, other heavy metals such as lead, cadmium, and mercury can reduce the thickness of the mother’s placenta and its overall functionality, with associated effects on birth weight and gestation length [43]. Heavy metals can also damage the

fetus through oxidative cell stress, neurological damage, DNA injury, and uptake of calcium and glucose [44]. Birth abnormalities and fetal development issues have been observed even at heavy metal exposure levels that are virtually non-toxic to the adult mother herself and with no overt clinical signs [45], thus reinforcing the possibility that even low exposure to contaminants from the Gold King Mine spill (such as through municipal water supplies, recreation, etc.) could have significant birth outcome impacts.

## MATERIALS AND METHODS

### Data

Restricted-use data on the near-universe of infant health outcomes and mother characteristics were obtained for Colorado and New Mexico, with permission, from the US CDC National Center for Health Statistics (NCHS) for the years 2011–2016. The analysis is limited to only those downstream counties in Colorado and New Mexico (and not areas of Utah or Arizona) since US EPA measurements showed small to negligible effects on water quality once the river reached the Utah state line (the metals had been heavily diluted by that point) [34]. These data contain individual birth records for >99% of all live births and include identification of each mother's county of residence. I focus on extensive and intensive margin measures of infant health: low birth weight (<2500 g at birth), prematurity (birth occurring before the 37th week of pregnancy), birth weight (in grams), and gestation length (in weeks). In addition, the data contain rich information on mother demographic characteristics (e.g., age, race, education) and behavioral actions taken during pregnancy (e.g., smoking status, prenatal care visits).

To account for potential confounders of infant health that are unrelated to the Gold King Mine spill, I combined the infant health data with annual data on household income and poverty rates (from the US Census Bureau Small Area Income and Poverty Estimates Program), unemployment rates (from the US Bureau of Labor Statistics), and demographic and educational attainment data (from the American Community Survey), all at the county-level. The specific variables used are presented in Table 1.

Since weather patterns are a known determinant of both health and pollution, I obtained detailed daily weather data on precipitation and temperature (maximum and minimum) from the NOAA National Centers for Environmental Information (NCEI). Given the presence of multiple weather monitoring sites in the counties in the study area, I follow [46] and take a population-weighted average of the weather stations within a 25-mile radius of the population centroid of the county. The weights are equal to the inverse of the square root of the station's distance to the population-weighted county centroid. Minimum and maximum temperature are averaged to obtain mean temperature in each county. I also use daily data on wind speed and direction from NCEI's North

American Regional Reanalysis. Data consist of 32 km × 32 km grid cells of East-West and North-South wind vectors. The inverse distance weighting approach was again used to construct wind conditions at the county-level. In the models that follow, I include flexible controls for both weather and wind. Five daily temperature bins (<25°, 25°–45°F, 45°–65°F, 65°–85°F, and >85°F), three daily wind direction bins (0–120°, 120–240°, and 240–360°) and three daily wind speed bins (0–2 m/s, 2–4 m/s, >4 m/s) are used. The results are robust to the exclusion of the wind controls. However, they are included in the main analysis to eliminate any confounding effects from odor from the river caused by the mine waste water spill that could have adversely affected human health during the treatment period. The bins capture the number of days during a given pregnancy that fall into each temperature or wind direction/speed range.

**Table 1.** Summary statistics. Sources: US CDC NCHS, US Census Bureau, US BLS, and NOAA NCEI.

Variable	A. Treatment Counties			B. Donor Pool Counties		
	Mean	Std. Dev.	Sample size	Mean	Std. Dev.	Sample size
Low birth weight	0.078	0.049	6	0.094	0.051	558
Prematurity	0.078	0.025	6	0.112	0.051	558
Birth weight (grams)	3185.89	120.63	6	3175.23	100.76	558
Gestation length (weeks)	38.99	0.23	6	38.72	0.39	558
Mother age	28.59	1.37	6	27.57	1.93	558
Mother education	4.19	0.20	6	4.12	0.609	558
Mother high school dropout (%)	0.171	0.051	6	0.155	0.084	558
Mother Hispanic (%)	0.244	0.091	6	0.353	0.238	558
Mother African American (%)	0.004	0.001	6	0.015	0.023	558
Mother smoked (%)	0.127	0.091	6	0.114	0.068	558
Teenage mother (%)	0.076	0.037	6	0.082	0.052	558
Birth order	1.92	0.109	6	2.16	0.270	558
Multiple birth (%)	0.015	0.006	6	0.029	0.031	558
Male infant (%)	0.467	0.058	6	0.506	0.074	558
Poverty rate (%)	0.156	0.009	6	0.174	0.075	558
Median household income	48,503.72	2633.66	6	47,661.81	15,834.09	558
Unemployment rate (%)	0.064	0.016	6	0.064	0.027	558
College graduation rate (%)	0.257	0.008	6	0.272	0.133	558
Demographics: African American (%)	0.004	0.000	6	0.016	0.022	558
Demographics: Hispanic (%)	0.155	0.006	6	0.293	0.201	558
Annual precipitation (in.)	19.11	1.24	6	16.22	1.09	558
Annual mean temperature (°F)	46.28	4.87	6	55.12	5.47	558
Annual wind speed (m/s)	2.95	0.474	6	3.14	0.589	558
Annual wind direction (°)	331.12	15.15	6	302.42	18.87	558

Mothers who reside in either San Juan County, CO, La Plata County, CO, or San Juan County, NM, are considered to be in the treated group—i.e., they live downstream from the Gold King Mine spill along or near the Animas River (see the shaded counties in Figure 1). These counties contain the population centers of Silverton, Durango, Aztec, and Farmington, which have a combined population of more than 70,000 people. While not all mothers in these counties live along the Animas River, there is a high

probability that many of them do since all the major population centers in this part of Colorado and New Mexico lie directly along the river's path.

The donor pool of counties from which the synthetic control will be constructed consists of all other counties in Colorado and New Mexico, not including the three treated counties listed above. Specifically, the donor pool consists of 94 counties with a combined 2010 US Census population of 6,906,298. By comparison, the 2010 combined population of the three treated counties is 182,077. Health impacts of the mine accident in these counties are expected to be negligible given that the Animas River does not extend to these counties, though, intrastate travel effects cannot be completely ruled out. However, as part of my specification checks later in the paper, alternative definitions of the donor pool are explored, including more limited geographic and population-based specifications.

My identifying assumption is that water contamination from the Gold King Mine spill only impacted mothers living downstream from the mine and not mothers in other counties in Colorado and New Mexico. Additionally, I also assume that any observed changes to infant health immediately after the water quality shock can be directly attributable to the mine disaster and are not due to changes in other factors or circumstances, unrelated to the disaster event. In the results section, I provide evidence supporting this assumption. Furthermore, as previously mentioned, given the short time period that is investigated in this work, it seems unlikely that any other predictors of infant health would have significantly changed at the same time that the spill occurred. I'm looking no more than 16 months after the spill event, and, as will be shown later, peak impacts occur within the first three months after the spill, thus virtually eliminating any non-spill-related environmental confounding concerns.

The analysis period is 2011–2016, inclusive, and data are averaged at the county-year level for use with the synthetic control method. Table 1 presents the summary statistics. 7.8% of births are of low weight or are premature in the treated counties, while the comparable figures for low birth weight and prematurity in the donor pool counties are 9.4% of 11.2%, respectively. Birth weight and gestation length are very similar between treated and donor pool counties, but there are notable differences between some of the socio-demographic characteristics.

While Table 1 only presents county-level averaged data, I note that later in the paper I also take advantage of the individual birth data available to me. Additionally, in the extended analysis, I also explore models using a monthly, rather than annual, timestep. In total, three different units of analysis are investigated in the paper: county-year (baseline), county-month, and individual birth-year. Given the rich data available to me, in addition to the short-term nature of the health impacts I'm seeking to capture, it is necessary to use more than one data aggregation type.

### Synthetic Control Method

To estimate the impact of the spill on infant health, the counterfactual path of birth outcomes in the downstream counties needs to be credibly identified. An increasingly preferred approach for constructing counterfactuals in causal inference settings is the synthetic control method or SCM of [16,17]. There are several advantages of the SCM over more traditional difference-in-differences (DID) approaches: (i) SCM is data driven, meaning that the subjectivity inherent in arbitrarily selecting a counterfactual is removed from the hands of the researcher; (ii) SCM moves away from using a single counterfactual unit or a simple average of counterfactual units, and instead uses a weighted average of units, and; (iii) SCM controls for time-varying unobserved confounders, thus helping to eliminate endogeneity resulting from omitted variables [47].

In this particular case, the SCM will be used to develop a synthetic control for the three downstream treated counties where water quality shocks occurred due to the Gold King Mine spill. The synthetic control will consist of a weighted combination of the 94 donor pool counties in Colorado and New Mexico that did not concurrently experience water quality shocks. Importantly, weights will be selected by the SCM such that the weighted synthetic control resembles the treated counties in terms of pre-spill observable infant health outcomes and relevant socio-demographic and weather characteristics.

Let  $J$  be the number of available donor pool counties and define the  $J \times 1$  weighting vector  $W = (w_1, \dots, w_J)'$  such that  $\sum_{j=1}^J w_j = 1$  and  $w_j \geq 0$  for  $j = (1, \dots, J)$ . Each scalar  $w_j$  represents the nonnegative weight placed on the  $j^{th}$  county in the synthetic control. Let  $H_0$  be a  $K \times 1$  vector of  $K$  birth outcomes and observable covariates for the downstream treated counties prior to the mine waste water spill. Let  $H_1$  be a  $K \times J$  matrix of comparable data vectors for each of the  $J$  counties in the donor pool. Following [16], the vector of weights  $W^*$  is chosen to minimize,

$$W^* = \arg \min_W (H_0 - H_1 W)' V (H_0 - H_1 W)$$

s. t.

$$w_j \geq 0, \quad \sum_{j=1}^J w_j = 1 \quad \text{for } j = (1, \dots, J) \quad (1)$$

where  $V$  is a positive-definite matrix whose diagonal elements reflect the relative importance of the variables in  $H_0$  and  $H_1$  ([48] also suggest an alternative approach where the synthetic control is constructed based solely on pre-treatment trends in the dependent variable of interest, thus ignoring all observable covariates. The results from this alternative approach are very similar to the results when the covariates are included, consistent with the findings in [48].). Variables in  $H_0$  and  $H_1$  include pre-disaster infant health outcomes, mother's age, Hispanic mother, African American mother, mother's educational attainment, mother is a high school dropout, mother smoked during pregnancy, teenage mother, birth

order, whether the birth was a multiple birth, sex of the child, and several county-level socio-demographic and weather indicators: poverty rate, median household income, unemployment rate, college graduation rate, percent African American, percent Hispanic, annual average precipitation, five daily mean temperature bins, three daily wind direction bins, and three daily wind speed bins. Minimizations are performed separately for each birth outcome (low birth weight, prematurity, birth weight, and gestation length) and the resulting weighting vectors are provided in Table S1 (Weights are estimated separately for each birth outcome since the common trends assumption must be met for each dependent variable investigated.).

To compare infant health outcomes between the downstream treated counties and the synthetic control, I follow [48] and use the following estimator,

$$SCM_{effect} = (Health_{post}^{treated} - Health_{post}^{synthetic}) - (Health_{pre}^{treated} - Health_{pre}^{synthetic}) \quad (2)$$

where  $Health_{post}$  ( $Health_{pre}$ ) is the average infant health outcome after (before) the mine waste water accident.

Characterization of uncertainty around  $SCM_{effect}$  is based on the results from placebo or falsification tests, wherein each donor pool county is (wrongly) assumed to have actually experienced treatment and the SCM is used to determine the magnitude of the placebo treatment effect [17]. This process is repeated for every donor pool county, thereby creating a distribution of placebo effect estimates:  $SCM_{placebo}$ . A two-sided p-value can be obtained from this process, following [49],

$$p - value = \Pr(|SCM_{placebo}| \geq |SCM_{effect}|) = \frac{\sum 1(|SCM_{placebo,j}| \geq |SCM_{effect}|)}{J} \quad (3)$$

where  $j$  is the  $j^{th}$  placebo county SCM estimate for  $j = (1, \dots, J)$ . Since placebo effect estimates in cases where the pre-treatment fit is poor (as defined by the root mean squared predictor error, RMSPE) do not provide credible information to measure the relative rarity of estimating a large post-treatment effect for a unit that was well-fitted prior to treatment [17], I weight each  $SCM_{placebo,j}$  by its pre-treatment RMSPE. This will give more (less) weight to placebo units with good (poor) pre-treatment fit with the downstream treated counties (Effectively, I'm downweighting SCM placebo estimates that have large pre-treatment RMSPE (meaning the SCM method didn't find a good synthetic control for these units during the pre-treatment period). This means that the p-value I calculate is based on placebo results from units with good pre-treatment fit (as defined by a low RMSPE), which allows for a more credible calculation of uncertainty since one is comparing good pre-treatment fit units from the placebo to a good pre-treatment fit from the treated unit.). In the robustness checks later in the paper, unweighted results are also reported.

### Traditional Difference-In-Differences Approach

The SCM uses county-level data to determine the infant health effect estimates. However, given the individual birth data available in the NCHS dataset, I would also like to employ a secondary approach that allows me to take advantage of this rich data. To do so, a more traditional DID approach is used, but with the caveat that I still use the previously obtained synthetic control weights. Specifically, the county-level weights generated by the SCM are used to re-weight the contribution of each individual NCHS observation in a given county by the SCM-derived county weight. In other words, individual observations are weighted by their county-level SCM weight in an attempt to construct a credible control group. This follows similar approaches in [48,50] and means that I can combine some of the advantages of the SCM (in terms of obtaining similar pre-treatment trends) with the advantages afforded by the traditional DID regression approach using individual birth observations.

The empirical model takes the following form,

$$Health_{ict} = \beta_0 + \beta_1 DownstreamAfter_{ct} + X'_{it}\beta_4 + M'_{ct}\beta_5 + W'_{ct}\beta_6 + Year_t + County_c + \varepsilon_{ict} \quad (4)$$

where  $Health_{ict}$  is the infant health outcome for mother  $i$  living in county  $c$  at year  $t$ ,  $DownstreamAfter_{ct}$  is an indicator variable for downstream county treatment status,  $X'_{it}$  is a vector of mother and birth characteristics,  $M'_{ct}$  is a vector of county socio-demographic characteristics,  $W'_{ct}$  is a vector of county weather characteristics,  $Year_t$  is the year fixed effect,  $County_c$  is the county fixed effect, and  $\varepsilon_{ict}$  is the idiosyncratic error term. The same set of mother, socio-demographic, and weather variables are used here as in the SCM (Later in the paper, I also explore a monthly version of Equation (4), where month, in addition to year, fixed effects are included.). The DID estimator is  $\beta_1$ , which provides an estimate of the changes in infant health outcomes in downstream treated counties relative to similar nearby counties that did not experience a concurrent change in river water quality (I also explored a version of the model using state fixed effects (Colorado and New Mexico), but the results were little changed as a result.). Standard errors are clustered by county.

## RESULTS

### Internal Validity Test

I begin by presenting evidence showing no significant changes in mother and infant characteristics, socio-demographics, or weather outcomes concurrent with the mine waste water disaster. This internal validity test was performed by using the previously obtained synthetic control weights for low birth weight to calculate  $SCM_{effect}$ , separately, for each observable covariate (Findings are similar to those reported here if the synthetic weights for the other infant health outcomes were used instead of the weights for low birth weight (e.g., prematurity, gestation, etc.)). In other words, I'm using the same synthetic control to calculate

how observable covariates across treatment and control groups changed from before and after the Gold King Mine spill. As shown in Table 2, none of the  $SCM_{effect}$  estimates are statistically significant, indicating that changes in the composition of mothers or changes in county socio-demographic or weather characteristics are not driving any mine-related infant health outcomes that might be found below.

**Table 2.** Internal Validity Test (Synthetic Control Results).

Outcome	$SCM_{effect}$	p-Value	Sample Size
Mother age	1.12	0.91	564
Mother education	0.191	0.39	564
Mother HS dropout	-0.123	0.17	564
Mother Hispanic	0.117	0.85	564
Mother African American	-0.005	0.90	564
Mother smoked	0.127	0.25	564
Teenage mother	0.017	0.21	564
Birth order	0.070	0.19	564
Multiple birth	-0.007	0.56	564
Male infant	-0.059	0.43	564
Poverty rate	-0.807	0.40	564
Median household income	1105.66	0.77	564
Unemployment rate	0.311	0.64	564
College graduation rate	-0.051	0.75	564
Demographics: African American	0.0001	0.33	564
Demographics: Hispanic	0.0002	0.94	564
Annual precipitation	0.0001	0.89	564
Mean temperature bins:			
<25°	0.001	0.89	564
25–45°	0.002	0.75	564
45–65°	0.001	0.95	564
65–85°	-0.001	0.96	564
>85°	0.004	0.90	564
Wind direction bins:			
0–120°	0.015	0.89	564
120–240°	0.017	0.87	564
240–360°	0.018	0.84	564
Wind speed bins:			
0–2 m/s	0.001	0.96	564
2–4 m/s	0.002	0.90	564
>4 m/s	0.005	0.91	564

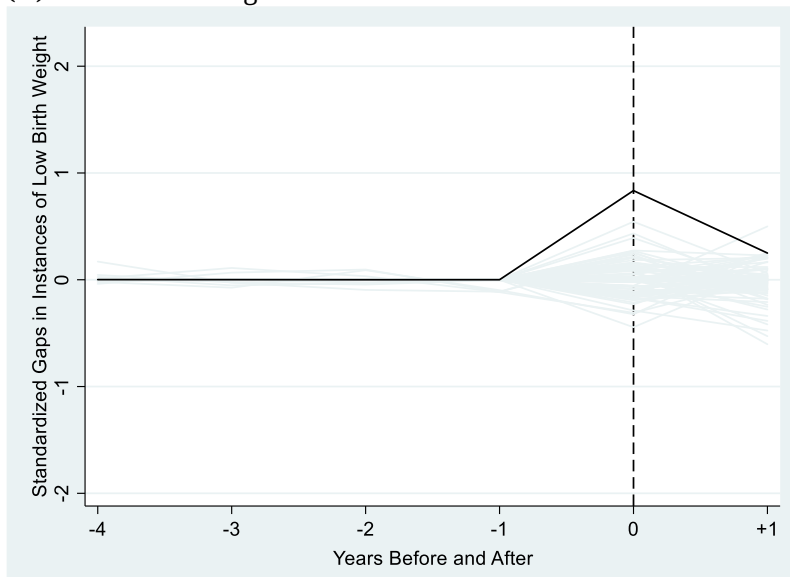
Notes: donor pool consists of all non-treated counties in NM and CO. For comparability purposes, the same synthetic control weights were used across all outcomes in the table (the low-birth-weight weights).  $SCM_{effect}$  estimates provide information on how a given outcome changed before and after the mine spill. Temperature and wind variables should be interpreted as the number of days that fall into each category.

### SCM Results

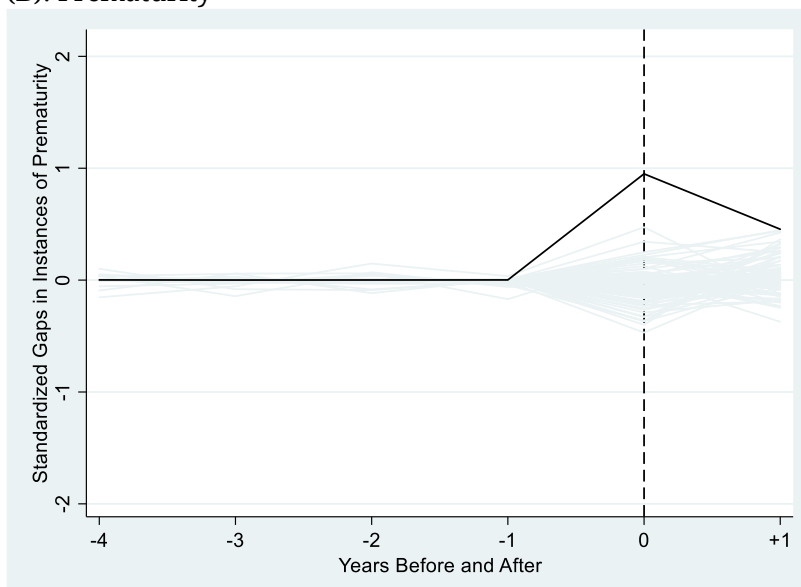
Results of using the SCM to estimate changes in low birth weight and prematurity are presented in Figure 3. Shown are the standardized gaps in annual outcomes between the treated and synthetic control groups (black lines) and similar gaps for each placebo effect (gray lines). The near zero gaps during the years prior to the mine spill indicate nearly identical pre-treatment levels and trends in low birth weight and prematurity—the

necessary common trends assumption. In the year of the spill, gaps in the instances of low birth weight and prematurity spike, by 0.8 and 0.9 percentage points, respectively (each from a base of 7.8%), suggesting that there was an increase in adverse infant health outcomes in the downstream counties that were not simultaneously observed in the synthetic control. Low birth weight and prematurity effects linger into the year after the spill, though are diminished in magnitude.

(A): Low Birth Weight



(B): Prematurity

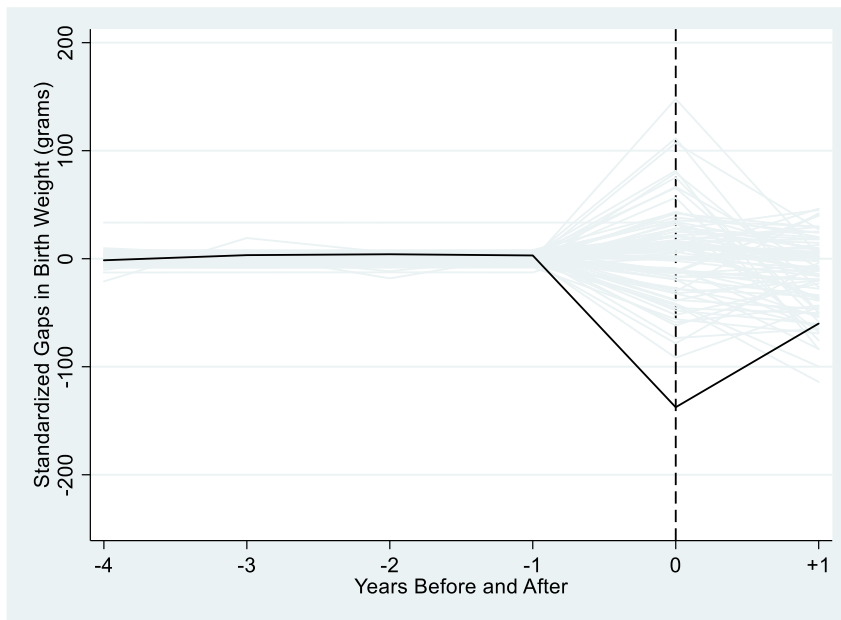
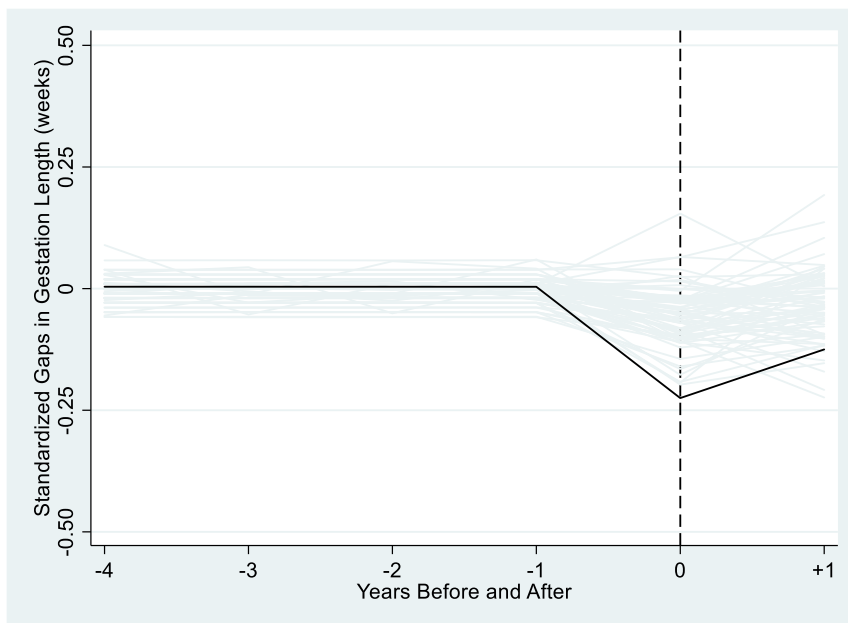


**Figure 3.** Standardized Gaps in Infant Health Outcomes between Downstream Counties and the Synthetic Control. Notes: plotted are the standardized gaps in infant health outcomes between downstream counties and the synthetic control (black line) and the standardized gaps between each county in the donor pool and its unique synthetic control for 94 placebo permutations (gray lines) for low birth weight (A) and prematurity (B). Gaps standardized by pre-treatment RMSPE. Different synthetic controls used in each panel. Gaps are defined as treated minus synthetic control.

Visually, the treatment gap is a relative outlier compared to the placebo effect gaps. For the year of the mine spill, this is clearly the case for both low birth weight and prematurity; both black lines lie above all the placebo gaps. Additionally, in the year after the spill, the downstream gap is still mostly an outlier for prematurity, but begins to converge towards the placebo gaps for low birth weight. These results should be interpreted as providing evidence that the magnitude of increases in instances of low birth weight and prematurity cannot be replicated by applying the SCM to counties that are not located downstream from the Gold King Mine spill.

The same analysis was conducted for the intensive margin infant health outcomes birth weight and gestation length, with results provided in Figure 4. During the year of the spill, birth weight is lower by 130 g (from a base of 3185 g) and gestation length is lower by 0.22 weeks (from a base of 38.99 weeks). The evidence is mixed for the year after the spill since the gaps begin reverting back to their pre-disaster levels. This suggests that the mine disaster, at least temporarily, resulted in lower weight infants and in infants born earlier than those in the synthetic control group that otherwise had nearly identical pre-treatment levels and trends along the intensive margin.

To numerically present the prior graphical results, I calculated  $SCM_{effect}$  using Equation (2) and the corresponding p-values for each infant outcome (Table 3). Panel A (panel B) reports the extensive margin (intensive margin) results. Along the extensive margin, the Gold King Mine spill is associated with a 0.6 percentage point average two-year increase in instances of low birth weight ( $p = 0.01$ ) and a 0.71 percentage point average two-year increase in instances of prematurity ( $p = 0.00$ ) in the downstream Animas River counties relative to the synthetic control. These represent percentage increases of 7.7% (low birth weight) and 9.1% (prematurity), relative to the mean. Intensive margin outcomes birth weight and gestation length are, on average, lower by 98.31 g ( $p = 0.05$ ) and 0.25 weeks ( $p = 0.09$ ) in downstream counties in the year of and year after the mine disaster. That is, relative to the mean, birth weight and gestation length are lower by 3.1% and 0.64%, respectively, over the two-year post-spill period (As a specification check, I additionally estimated a version of Table 3 that used a single synthetic control (with one set of weights) for all four infant health outcomes, rather than separate controls for each outcome. While doing so led to some violations of the common trends assumption, especially for the intensive margin outcomes, the results continue to show significant adverse infant health outcomes by using this alternative specification, though the magnitudes of effects were different from those presented here.).

**(A): Birth Weight****(B): Gestation Length**

**Figure 4.** Standardized Gaps in Birth Weight and Gestation Outcomes between Downstream Counties and the Synthetic Control (Intensive Margin Infant Health Outcomes). Notes: plotted are the standardized gaps in intensive margin birth outcomes between downstream counties and the synthetic control (black line) and the standardized gaps between each county in the donor pool and its unique synthetic control (gray lines) for birth weight **(A)** and gestation length **(B)**. Gaps standardized by pre-treatment RMSPE.

**Table 3.** Impact of the Gold King Mine Spill on Infant Health (Synthetic Control Results).

Outcome	Average Pre-treatment Difference	Average Post-treatment Difference	SCM <sub>effect</sub>	p-value
<i>Panel A: Extensive margin outcomes</i>				
Low birth weight	0.0001	0.0061	0.0060	0.01
Prematurity	0.0002	0.0073	0.0071	0.00
Mother characteristic controls	Yes	Yes	Yes	
Socio-demographic controls	Yes	Yes	Yes	
Weather controls	Yes	Yes	Yes	
Sample size	564	564	564	
<i>Panel B: Intensive margin outcomes</i>				
Birth weight	2.87	-95.44	-98.31	0.05
Gestation length	0.087	-0.158	-0.245	0.09
Mother characteristic controls	Yes	Yes	Yes	
Socio-demographic controls	Yes	Yes	Yes	
Weather controls	Yes	Yes	Yes	
Sample size	564	564	564	

Notes: average differences reported are synthetic control estimates of the difference in extensive margin infant health outcomes (panel A) and intensive margin outcomes (panel B) between downstream counties and the matched synthetic control group. Each outcome has a unique synthetic control. P-values obtained from placebo tests. Controls are the same as those used to produce Figures 3 and 4 and are used in the SCM weighting process to construct the synthetic control.  $SCM_{effect}$  calculated as  $(Health_{post}^{treated} - Health_{post}^{synth}) - (Health_{pre}^{treated} - Health_{pre}^{synth})$ .

### DID Results

The results up until now have used county-level averaged data. Here, I use individual birth data to estimate the traditional DID model in Equation (4) using the synthetic control weights, as previously described. Results of these regressions are presented in Table 4. Note first that the sample sizes have increased tremendously; from 564 in the SCM results to over 490,000 now. Consistent with the earlier story, I continue to observe adverse impacts of the mine water disaster on both extensive and intensive margin infant health outcomes. DID coefficients suggest that instances of low birth weight are higher by 0.64 percentage points ( $p < 0.01$ ) and prematurity cases are higher by 0.69 percentage points ( $p < 0.01$ ). Birth weight is lower by 101.49 g ( $p < 0.05$ ) and gestation length is lower by 0.251 weeks ( $p < 0.10$ ), though with only marginal significance. These effect estimates are very similar in magnitude and significance to what was observed in the SCM results, suggesting that results are robust to: (i) different empirical models (DID vs. SCM), and; (ii) using either county-level averaged data or individual birth data.

**Table 4.** Traditional Regression Difference-in-Differences Results.

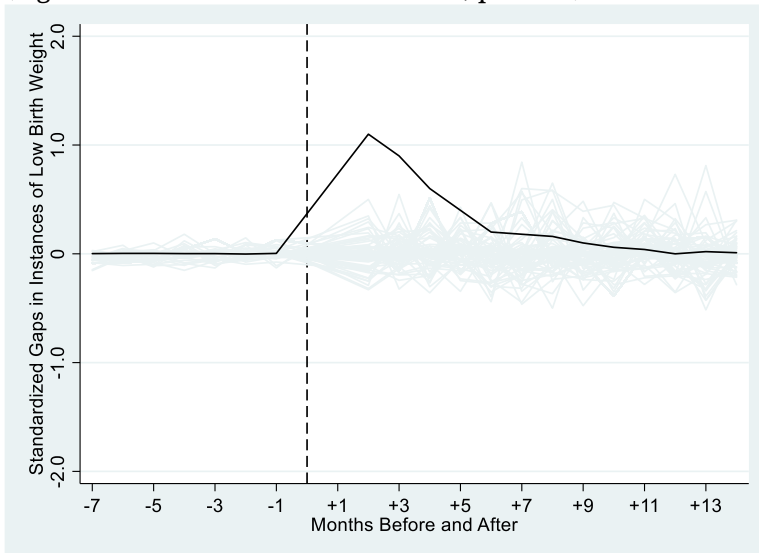
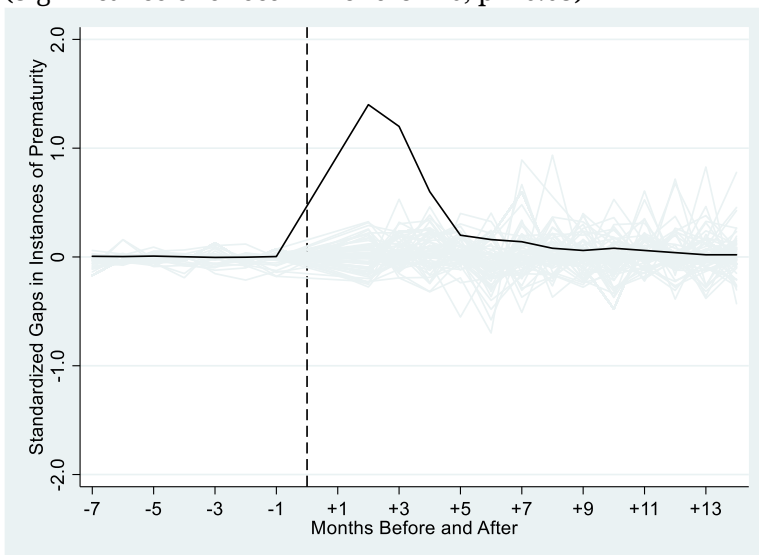
	(1)	(2)	(3)	(4)
	Low Birth Weight	Prematurity	Birth Weight	Gestation Length
DID coefficient	0.0064*** (0.002)	0.0069*** (0.002)	-101.49** (51.47)	-0.251* (0.152)
Mother characteristics	Yes	Yes	Yes	Yes
County socio-demographic controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
R <sup>2</sup> (within)	0.18	0.17	0.20	0.21
Sample size	494,050	493,810	493,358	494,602

Notes: each column is a separate regression of Equation (4) in the text. individual birth data used here. weights generated by the synthetic control method are used to select control counties and re-weight the contribution of each individual observation such that the cumulative weight associated with the observations from a county matches the synthetic weights. controls are the same as those used above. county clustered standard errors in parentheses. \*\*\*p < 0.01; \*\*p < 0.05; \*p < 0.10.

### Monthly Analysis Results

Perhaps the most convincing evidence to support an acute infant health impact story can be provided by using higher frequency monthly data, rather than the annual data used up until now. I again employ the SCM to construct synthetic counterfactuals for the downstream counties for low birth weight and prematurity, using the same donor pool and the same set of observable characteristics as before. However, a monthly calendar timestep is now employed. Figure 5 presents the results and associated p-values of this analysis. Gaps between the treated and synthetic control counties spike in the first three months immediately following the mine spill. After the immediate post-spill spikes, there is a gradual reversion to pre-spill outcomes, and by 12 calendar months after the disaster, both the low birth weight and prematurity gaps are near zero (and insignificant), indicating no difference between the downstream and synthetic control counties. At peak impact, instances of low birth weight are elevated by 1.1 percentage points, or 14.1% above the mean (two months after the spill) and instances of prematurity are elevated by 1.2 percentage points, or 15.4% above the mean (also two months after the spill). Placebo test p-values indicate that the observed post-spill outcomes are significantly different from zero.

The overall shape and timing of the effect curves are strongly suggestive of an acute health impact story consistent with a short-term change in environmental exposures. Clearly, an abrupt change in infant health outcomes occurred at the same time as the Gold King Mine spill that did not also occur in the donor pool counties. These results provide perhaps the strongest evidence yet to support a causal story.

**(A): Low Birth Weight**(Significance of effect in months 1–6;  $p = 0.01$ )**(B): Prematurity**(Significance of effect in months 1–6;  $p = 0.03$ )

**Figure 5.** Monthly Impacts of the Mine Water Disaster on Extensive Margin Infant Health Outcomes (Synthetic Control Results). Notes: monthly impacts of the mine waste water disaster on gaps in extensive margin infant health outcomes shown (black lines) along with placebo test gaps (gray lines).  $p$ -value from donor pool placebo tests. Synthetic control constructed by matching treated and donor pool counties on pre-treatment infant health outcomes, mother and infant characteristics, county socio-demographic characteristics, and county weather characteristics, as before.

### Alternative Donor Pool Specification Results

I also performed many robustness checks to determine if the main specifications are robust to different constructions of the donor pool. Results from these alternative specifications are provided in Table 5. In panel A, the donor pool has been restricted to those counties immediately adjacent to the three downstream counties. Panel B restricts the donor

pool to counties with a population centroid within 100 miles (161 km) of the population centroids of the downstream counties. In panels C and D, rather than constructing the donor pool based on geographic distance, I instead use a population-based metric. In panel C, the top 10 highest population counties in Colorado and New Mexico are included in the donor pool, whereas in panel D, the top 10 lowest population counties in these states constitute the donor pool. Finally, panel E drops from the donor pool all counties adjacent to the downstream counties (in case there were spillover effects to these counties).

**Table 5.** Alternative Donor Pool Specifications (Synthetic Control Results).

Outcome	Average Pre-treatment Difference	Average Post-treatment Difference	$SCM_{effect}$	P-value	Sample size
<i>Panel A: Adjacent Counties Only</i>					
Low birth weight	-0.0004	0.0057	0.0061	0.02	60
Prematurity	0.0001	0.0075	0.0074	0.01	60
Birth weight	2.99	-97.51	-100.50	0.04	60
Gestation length	0.041	-0.156	-0.197	0.10	60
<i>Panel B: 100-mile Population Centroid</i>					
Low birth weight	-0.0003	0.0059	0.0062	0.01	84
Prematurity	-0.0001	0.0072	0.0073	0.01	84
Birth weight	3.07	-99.14	-102.21	0.05	84
Gestation length	0.057	-0.161	-0.218	0.09	84
<i>Panel C: Highest Population Counties</i>					
Low birth weight	-0.0014	0.0052	0.0066	0.01	60
Prematurity	-0.0019	0.0069	0.0088	0.02	60
Birth weight	4.11	-94.97	-99.08	0.05	60
Gestation length	0.049	-0.159	-0.208	0.12	60
<i>Panel D: Lowest Population Counties</i>					
Low birth weight	0.0006	0.0056	0.0050	0.03	60
Prematurity	0.0011	0.0044	0.0033	0.05	60
Birth weight	4.74	-91.46	-95.93	0.07	60
Gestation length	0.055	-0.150	-0.205	0.14	60
<i>Panel E: Dropping Adjacent Counties</i>					
Low birth weight	0.0001	0.0058	0.0057	0.00	510
Prematurity	0.0002	0.0076	0.0074	0.01	510
Birth weight	3.08	-99.67	-102.75	0.03	510
Gestation length	0.027	-0.164	-0.191	0.08	510

Notes: synthetic control results shown across various alternative donor pool specifications. In panel A, only the counties adjacent to the treated counties are included in the donor pool. In panel B, the donor pool consists of those counties with a population centroid within 100 miles of the population centroid of the treated counties. In panel C (panel D), the donor pool is restricted to the 10 highest (lowest) population counties in NM and CO. In panel E, counties adjacent to the treated counties are dropped from the donor pool.

With very few exceptions, the  $SCM_{effect}$  estimates remain statistically significant across all five specifications and also have similar magnitudes both to each other and to the baseline results previously presented. Two notable exceptions are for gestation length in panel C (the highest population specification) and again in panel D (the lowest population specification). In these cases, the SCM results are no longer statistically significant, though it is important to note that the gestation length results

were only marginally significant before, thus, the change in significance observed here is not especially meaningful. Despite these two isolated cases, I conclude from this specification check that the main findings are generalizable across alternative constructions of the donor pool, suggesting that the results are not being driven by the original donor pool choice.

### **Falsification Test Results**

Two falsification tests are performed on the baseline results. For the first test, I show no effect of the spill on outcomes measured before the spill actually occurred. Specifically, the SCM was again employed (using the same set of matching covariates as before), but now using the years 2011 and 2012 as the pre-treatment period and the years 2013 and 2014 as the treatment period. Recall that the actual mine disaster occurred in August 2015. The results from this test are presented in Figure S1. In short, none of the infant health outcome gaps are significantly different from zero at conventional levels and all the gaps are fairly close to zero. Hence, as expected, there is no impact of the mine spill prior to when it actually occurred, consistent with a causal story.

For the second falsification test, I show no impact of the mine disaster on birth outcomes that are known to be unaffected by temporary environmental shocks: instances of large for gestational age (LGA) births and instances of postterm pregnancy. LGA describes infants who weigh more than the 90th percentile of babies of the same gestational age. In the US, this commonly translates into weighing  $\geq 4000$  g at birth [51]. A postterm pregnancy is a pregnancy that extends to 42 weeks of gestation or beyond. Prior literature has suggested that genetic factors and maternal diabetes are the most likely causes of both of these birth outcomes [51,52]. As expected, I show in Figure S2 that there is indeed no impact of the Gold King Mine spill on these birth outcomes, thereby further boosting the causal interpretation of the earlier findings.

### **Additional Robustness Check Results**

Several additional robustness checks were performed on the types of births included in the analysis and the variables used in the SCM weighting process. These results are reported in Table S2 and are briefly summarized here. First, I limit the sample to singleton births (i.e., the birth of only one child during a single delivery) rather than both multiple and singleton births as was previously done. This check is performed because multiple births are more likely to have poor health outcomes independent of exposures to environmental pollution [18]. Second, I add controls for health insurance status (Medicaid, private insurance, and self-pay) and an indicator for receipt of WIC (Women, Infants, and Children) to the set of variables included in the SCM weighting process. Third, I do not weight the placebo effect estimates by their pre-treatment RMSPE as was done in the

main specifications. Overall, as reported in Table S2, I find that the main results are largely robust to these checks.

## DISCUSSION

This work provides new causal evidence that temporary shocks to water quality caused by a major mining disaster can adversely affect infant health outcomes in counties located immediately downstream from the spill site. I provide evidence that concentrations of heavy metals and toxic compounds such as arsenic, lead, and mercury were above US EPA water quality guidelines after the spill in the downstream counties investigated, making them the most probable exposure pathways for the effects observed. However, additional data and analysis would be needed to fully investigate the actual exposure mechanisms, which is outside the scope of this paper. Several possible mechanisms could stem from the fact that the Animas River is a major source of drinking water in the counties investigated, in addition to being a significant source of recreational opportunities and economic livelihoods in Colorado and New Mexico.

The main finding in this work is that instances of low birth weight and prematurity were elevated by an average of 0.6 and 0.71 percentage points, respectively, in the two years after the August 2015 Gold King Mine spill, or, by 7.7% and 9.1% relative to the mean. Based on the empirical relationship reported in [53], the observed percentage increase in low birth weight is approximately equivalent to what would be expected from a mother smoking 4.7 cigarettes per day for the duration of her pregnancy, or, alternatively, being exposed to an additional 0.14 parts per million (ppm) of carbon monoxide (CO) during the last trimester of pregnancy. For comparison, US EPA AirData reports that average CO concentrations near Durango, CO (located in our treatment area) were 0.3 ppm in 2015. Thus, the 0.14 ppm increase required to obtain similar low birth weight outcomes as caused by the Gold King Mine spill is significant by comparison. Hence, this reinforces the significance of the low birth weight finding, as well as the importance of accounting for infant health externalities when assessing the damages of mine waste water disasters.

However, while the results are significant, they are not unreasonably large when compared with other investigations of infant health impacts associated with pollution exposures. For instance, [54] found that the closure of a coal-fired power plant in New Jersey led to a 15% change in low birth weight and a 28% change in prematurity among downwind mothers. Similarly, [55] found that reduced vehicle emissions surrounding toll plazas in the northeastern US improved low birth weight by 10.8% and prematurity by 11.8% among nearby mothers. By contrast, the 7.7% (low birth weight) and 9.1% (prematurity) average two-year findings, and the peak impact findings of 14.1% (low birth weight) and 15.1% (prematurity) within the first two months after the spill are slightly smaller by comparison, yet are still meaningful.

## CONCLUSIONS

Given the legacy of abandoned underground mines both in the US and globally, the potential for future mine waste water disasters is likely. The US House Committee on Natural Resources estimates that there are approximately 500,000 abandoned hard rock mines in the US as of 2016, and in Colorado alone, the site of the Gold King Mine disaster, there are an estimated 23,000 abandoned mines. Since the 1970s, there have been a total of 32 uncontrolled releases of mine waste water in the US, including in the states of California, Colorado, Kentucky, Montana, New Jersey, Pennsylvania, Virginia, and West Virginia [32]. As noted by the US Bureau of Reclamation, in their incident report on the Gold King Mine disaster, there are significant risks that future mine waste water disasters similar to the Gold King Mine incident will occur in the near future due to the sheer number of similar abandoned mines in the US, the lack of sufficient resources for proper long-term maintenance and study, and due to inconsistencies in abandoned mine management from agency to agency, which shows “little appreciation for the engineering complexity” of abandoned mine reclamation projects [32]. Thus, the results of this study are broadly informative of the potential short-term infant health risks associated with abandoned mine water releases across the US. Understanding the acute health impacts of these environmental disaster events is therefore necessary and can aid in the management of scarce reclamation budgets along sustainability dimensions.

Several potential policy recommendations emerge from this work. First, urgent Federal and state funding is needed to clean-up and sustainably manage abandoned and legacy metal mines in the US. Second, considerations of abandoned mines’ impacts (i.e., in clean-up benefit-cost analyses) needs to extend to downstream communities in instances of close proximity of mines to waterbodies. As this work shows, downstream communities far from abandoned mines can be affected by mine waste water disasters. Finally, issues of liability and responsibility in instances of mine waste water disasters needs to be formally established in law in order to properly compensate affected individuals and communities, such as for negative health impacts.

More broadly, this work illustrates that temporary water quality shocks, even in a developed country, can have meaningful impacts on human health. To the extent that people continue to migrate to abandoned industrial and mining communities, or toward cities located downstream from former industrial mining operations, then this work, and others like it, will continue to rise in importance and significance. This is especially true in cases where infant health outcomes are impacted, given their important role as determinants of long-term human capital and labor market outcomes.

## SUPPLEMENTARY MATERIALS

The following supplementary materials are available online, Figure S1: Pre-Mine Disaster Falsification Test (Synthetic Control Results), Figure S2: Falsification Tests on Alternative Birth Outcomes Not Known to Be Impacted by Acute Environmental Exposures (Synthetic Control Results), Table S1: Synthetic Control Weights by Birth Outcome, Table S2: Additional Robustness Checks (Synthetic Control Results).

## DATA AVAILABILITY

The dataset from the study is not available because it is restricted use under a data use agreement.

## CONFLICTS OF INTEREST

The author declares that they have no conflicts of interest.

## FUNDING

No funding was received for this work.

## REFERENCES

1. Sibanda T, Selvarajan R, Tekere M. Urban effluent discharges as causes of public and environmental health concerns in South Africa's aquatic milieu. *Environ Sci Pollut Res*. 2015;22(23):18301-17.
2. Soller JA, Schoen ME, Bartrand T, Ravenscroft JE, Ashbolt NJ. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Res*. 2010;44(16):4674-91.
3. Smith AH, Lopipero PA, Bates MN, Steinmaus CM. Arsenic epidemiology and drinking water standards. *Science*. 2002;296(5576):2145-7.
4. Kapaj S, Peterson H, Liber K, Bhattacharya P. Human health effects from chronic arsenic poisoning—a review. *J Environ Sci Health Part A*. 2006;41(10):2399-428.
5. Ward MH, DeKok TM, Levallois P, Brender J, Gulis G, Nolan BT, et al. Workgroup report: drinking-water nitrate and health—recent findings and research needs. *Environ Health Perspect*. 2005;113(11):1607-14.
6. Jagai JS, Li Q, Wang S, Messier KP, Wade TJ, Hilborn ED. Extreme precipitation and emergency room visits for gastrointestinal illness in areas with and without combined sewer systems: an analysis of Massachusetts data, 2003–2007. *Environ Health Perspect*. 2015;123(9):873.
7. Alderman K, Turner LR, Tong S. Floods and human health: a systematic review. *Environ Int*. 2012;47:37-47.
8. Lai W. Pesticide use and health outcomes: evidence from agricultural water pollution in China. *J Environ Econ Manage*. 2017;86:93-120.

9. Chen Y, Parvez F, Gamble M, Islam T, Ahmed A, Argos M, et al. Arsenic exposure at low-to-moderate levels and skin lesions, arsenic metabolism, neurological functions, and biomarkers for respiratory and cardiovascular diseases: review of recent findings from the Health Effects of Arsenic Longitudinal Study (HEALS) in Bangladesh. *Toxicol Appl Pharmacol*. 2009;239(2):184-92.
10. Hunter PR, Zmirou-Navier D, Hartemann P. Estimating the impact on health of poor reliability of drinking water interventions in developing countries. *Sci Total Environ*. 2009;407(8):2621-4.
11. Landrigan PJ, Fuller R, Acosta NJ, Adeyi O, Arnold R, Baldé AB, et al. The Lancet Commission on pollution and health. *Lancet*. 2018;391(10119):462-512.
12. Guilfoos T, Kell D, Boslett A, Hill EL. The economic and health effects of the 2014 chemical spill in the Elk River, West Virginia. *Am J Agric Econ*. 2017;100(2):609-24.
13. USGS. The top 5 mineral-producing states [Internet]. Available from: <https://www.usgs.gov/news/science-snippet/top-5-mineral-producing-states#targetText=California%20ranks%20number%204%20overall,production%2C%20California's%20contribution%20is%20significant>. Accessed on 2019 Oct 18.
14. Olmstead SM. The economics of managing scarce water resources. *Rev Environ Econ Policy*. 2010;4(2):179-98.
15. Almond D, Currie J. Killing me softly: the fetal origins hypothesis. *J Econ Perspect*. 2011;25(3):153-72.
16. Abadie A, Gardeazabal J. The economic costs of conflict: a case study of the Basque Country. *Am Econ Rev*. 2003;93(1):113-32.
17. Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *J Am Stat Assoc*. 2010;105(490):493-505.
18. Hill E, Ma L. Shale gas development and drinking water quality. *Am Econ Rev*. 2017;107(5):522-5.
19. Olmstead SM. The economics of water quality. *Rev Environ Econ Policy*. 2009;4(1):44-62.
20. Zhang J. The impact of water quality on health: evidence from the drinking water infrastructure program in rural China. *J Health Econ*. 2012;31(1):122-34.
21. Srinivasan JT, Reddy VR. Impact of irrigation water quality on human health: a case study in India. *Ecol Econ*. 2009;68(11):2800-7.
22. Marshall University Center for Business and Economic Research (CBER). CBER calculates impact from chemical spill into Elk River. Available from: [http://mediad.publicbroadcasting.net/p/wvvpn/files/201402/CBER\\_PR\\_WaterUsageBan\\_02.04.14.pdf](http://mediad.publicbroadcasting.net/p/wvvpn/files/201402/CBER_PR_WaterUsageBan_02.04.14.pdf). Accessed on 2026 Apr 10.
23. Knap AH, Rusyn I. Environmental exposures due to natural disasters. *Rev Environ Health*. 2016;31(1):89-92.
24. Aguilera F, Méndez J, Pásaro E, Laffon B. Review on the effects of exposure to spilled oils on human health. *J Appl Toxicol*. 2010;30(4):291-301.

25. Campbell D, Cox D, Crum J, Foster K, Riley A. Later effects of grounding of tanker Braer on health in Shetland. *BMJ*. 1994;309(6957):773-4.
26. Harada M. Minamata disease: methylmercury poisoning in Japan caused by environmental pollution. *Crit Rev Toxicol*. 1995;25(1):1-24.
27. Palinkas LA, Petterson JS, Russell J, Downs MA. Community patterns of psychiatric disorders after the Exxon Valdez oil spill. *Am J Psychiatry*. 1993;150(10):1517.
28. Griffith J, Duncan RC, Riggan WB, Pellom AC. Cancer mortality in US counties with hazardous waste sites and ground water pollution. *Arch Environ Health*. 1989;44(2):69-74.
29. Naujokas MF, Anderson B, Ahsan H, Aposhian HV, Graziano JH, Thompson C, Suk WA. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect*. 2013;121(3):295.
30. de Vries W, Römkens PF, Schütze G. Critical soil concentrations of cadmium, lead, and mercury in view of health effects on humans and animals. In: *Reviews of environmental contamination and toxicology*. New York (NY): Springer; 2007. p. 91-130.
31. Associated Press. Robots could climb into old mines to prevent toxic spills. Available from: <https://nypost.com/2018/01/31/robots-could-climb-into-old-mines-to-prevent-toxic-spills/>. Accessed on 2026 Apr 10.
32. US Bureau of Reclamation. Technical evaluation of the Gold King Mine incident. Denver (CO, US): US Department of Interior, Bureau of Reclamation, Technical Service Center; 2015. Available from: <https://assets.documentcloud.org/documents/2475483/technical-evaluation-of-the-gold-king-mine.pdf>. Accessed on 2018 Oct 16.
33. Chief K, Artiola JF, Wilkinson ST, Beamer P, Maier RM. Understanding the Gold King Mine spill. Tucson (AZ, US): University of Arizona, Superfund Research Program; 2015. Available from: [https://superfund.arizona.edu/sites/default/files/gold\\_king\\_mine\\_spill.pdf](https://superfund.arizona.edu/sites/default/files/gold_king_mine_spill.pdf). Accessed on 2018 Oct 16.
34. US EPA Region 8. Emergency response monitoring data from the Gold King Mine incident. Available from: <https://www.epa.gov/goldkingmine/emergency-response-monitoring-data-gold-king-mine-incident>. Accessed on 2018 Oct 16.
35. Garrison S, Kellogg J. EPA: pollution from mine spill much worse than feared [Internet]. *USA Today*. 2015 Aug 9. Available from: <https://www.usatoday.com/story/news/nation/2015/08/09/navajo-nation-epa-spill/31384515/>. Accessed on 2018 Oct 16.
36. RPI Consulting Inc. City of Durango, Colorado: economic impacts of whitewater recreation. Durango (CO, US): RPI Consulting; 2006. Available from: <https://www.durangogov.org/DocumentCenter/View/57>. Accessed on 2018 Oct 16.

37. Turkewitz J. Environmental agency uncorks its own toxic water spill at Colorado mine. Available from: <https://www.nytimes.com/2015/08/11/us/durango-colorado-mine-spill-environmental-protection-agency.html>. Accessed on 2026 Apr 10.
38. US EPA. Frequent questions related to Gold King Mine response. Available from: <https://www.epa.gov/goldkingmine/frequent-questions-related-gold-king-mine-response>. Accessed on 2018 Oct 17.
39. Chief K, Beamer P, Teufel-Shone N, Billheimer D, Ingram J, Begay M, Begay M. Navajo Gold King Mine exposure project [Internet]. Fall 2016, Issue 1. Tucson (AZ, US): University of Arizona; 2016. Available from: [https://superfund.arizona.edu/sites/default/files/u68/gkm\\_navajo\\_exposure\\_project\\_update\\_oct\\_2016\\_rdx\\_0.pdf](https://superfund.arizona.edu/sites/default/files/u68/gkm_navajo_exposure_project_update_oct_2016_rdx_0.pdf). Accessed on 2026 Apr 10.
40. Joint Information Center. Animas River spill incident. Available from: [https://drive.google.com/file/d/1I\\_P\\_cT4AaRbdnJbSNm1VSXMSvW\\_nL3M/view](https://drive.google.com/file/d/1I_P_cT4AaRbdnJbSNm1VSXMSvW_nL3M/view). Accessed on 2026 Feb 17.
41. Liu H, Lu S, Zhang B, Xia W, Liu W, Peng Y, et al. Maternal arsenic exposure and birth outcomes: a birth cohort study in Wuhan, China. *Environ Pollut*. 2018;236:817-23.
42. Almberg KS, Turyk ME, Jones RM, Rankin K, Freels S, Graber JM, Stayner LT. Arsenic in drinking water and adverse birth outcomes in Ohio. *Environ Res*. 2017;157:52-9.
43. Al-Saleh I, Shinwari N, Mashhour A, Rabah A. Birth outcome measures and maternal exposure to heavy metals (lead, cadmium and mercury) in Saudi Arabian population. *Int J Hyg Environ Health*. 2014;217(2-3):205-18.
44. Caserta D, Graziano A, Monte GL, Bordi G, Moscarini M. Heavy metals and placental fetal-maternal barrier: a mini-review on the major concerns. *Eur Rev Med Pharmacol Sci*. 2013;17(16):2198-206.
45. Shirai S, Suzuki Y, Yoshinaga J, Mizumoto Y. Maternal exposure to low-level heavy metals during pregnancy and birth size. *J Environ Sci Health Part A*. 2010;45(11):1468-74.
46. Levinson A. Valuing public goods using happiness data: the case of air quality. *J Public Econ*. 2012;96(9-10):869-80.
47. Adhikari B, Alm J. Evaluating the economic effects of flat tax reforms using synthetic control methods. *South Econ J*. 2016;83(2):437-63.
48. Bohn S, Lofstrom M, Raphael S. Did the 2007 Legal Arizona Workers Act reduce the state's unauthorized immigrant population? *Rev Econ Stat*. 2014;96(2):258-69.
49. Galiani S, Quistorff B. The synth\_runner package: utilities to automate synthetic control estimation using synth. Working paper. College Park (MD): University of Maryland; 2016.
50. Jones BA. Spillover health effects of energy efficiency investments: quasi-experimental evidence from the Los Angeles LED streetlight program. *J Environ Econ Manage*. 2018;88:283-99.

51. University of Rochester Medical Center. Large for gestational age. Available from:  
<https://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=90&ContentID=P02383>. Accessed on 2018 Oct 23.
52. Galal M, Symonds I, Murray H, Petraglia F, Smith R. Postterm pregnancy. *Facts Views Vis ObGyn*. 2012;4(3):175-87.
53. Currie J, Neidell M, Schmieder JF. Air pollution and infant health: lessons from New Jersey. *J Health Econ*. 2009;28(3):688-703.
54. Yang M, Chou SY. The impact of environmental regulation on fetal health: evidence from the shutdown of a coal-fired power plant located upwind of New Jersey. *J Environ Econ Manage*. 2018;89:94-115.
55. Currie J, Walker R. Traffic congestion and infant health: evidence from E-ZPass. *Am Econ J Appl Econ*. 2011;3(1):65-90.

How to cite this article:

Jones BA. Lookout Below! The Downstream Infant Health Impacts of Mining Disasters. *J Sustain Res*. 2026;8(2):e260035. <https://doi.org/10.20900/jsr20260035>.