Jenna Anders and Rachel Young
Aug 2025

Overview: This project is a large-scale study of the health effects of environmental contamination, and the benefits of remediation. The project has two parts. In the first part, we combine rich, geographically-coded data on groundwater quality with the location and timelines of remediation projects to assess the effectiveness of remediation at reducing groundwater contamination. In the second part of the project, we add the universe of cancer cases in several states and high-resolution data on migration away from contaminated areas (as a measure of exposure) to identify the role of environmental contamination in cancer incidence. We hypothesize that individuals with similar characteristics who are exposed to different pollutants across remediation sites face heightened risks of developing different cancers. The results of this study will improve understanding of the health impacts of specific pollutants and help quantify the benefits of targeted pollution cleanup efforts.

Part I. Remediation and water quality. Environmental remediation is infamously costly: in fiscal year 2025, the U.S. Environmental Protection Agency (EPA) spent over $1 billion on federal Superfund remediation activities, and states and private companies likely spent much more.[1] These expenditures are dedicated in the name of environmental quality; however, to our knowledge, no *systematic* description of the improvement to environmental quality exists.

We begin by quantifying changes in pollutant concentrations at remediation sites in Minnesota and California over several decades. To do this, we link water-sample-by-contaminant-level data from wells and monitors to data on remediation activities in cleanup projects overseen by the Minnesota Pollution Control Agency (MPCA) and the California Department of Toxic Substances Control (DTSC). We present visual and analytical descriptions of the reduction of groundwater contamination after cleanup activities. While informative on its own, this first stage of analysis also reveals important spatial and temporal variation in pollution levels across remediation sites that we will use in the second stage of the analysis.

Part II. Health effects of contamination. While economists and other social scientists have identified causal effects of environmental contamination on several health outcomes, including infant health (congenital anomalies, preterm birth) and children's blood lead levels, *large-scale* causal analyses of cancer incidence near environmental contamination have proven elusive. This is likely due to the long latency period cancers have; most quasi-experimental studies use site cleanup as an identification event, but only short-term health outcomes are amenable to this type of study.

We use a novel identification strategy: the fact that certain contaminants are more likely associated with certain cancers, regardless of the path of exposure (for example, benzene is associated with blood cancers, while arsenic is associated with skin cancers). We apply this causal analysis to test whether specific pollutant-cancer links that have been identified in case studies and theoretical models are observed at a broader scale with lower-level contamination exposure, and to estimate the increased risk associated with exposure to a specific pollutant. Finally, we test whether communities with lower outmigration (a measure of exposure) experience higher cancer risks.

---

[1] U.S. Environmental Protection Agency, FY 2026 Budget in Brief. https://www.epa.gov/system/files/documents/2025-05/fy-2026-epa-bib.pdf This number does not include expenditures on enforcement, research, operations, or other necessary aspects of the federal Superfund program.