URBAN INSTITUTE · ELEVATE · THE · DEBATE

Where Are Children in Head Start Exposed to Environmental Hazards?

Technical Appendix

Christopher Davis, Anne N. Junod, and Amy Rogin

Last updated September 2023

This technical appendix describes the data sources and specifications, methodology, and related limitations for the "Where are Children in Head Start Exposed to Environmental Hazards?" feature. For more information about how to use this tool, see our Quick Start Guide.

Data Sources and Specifications

Head Start Data

We used administrative data about the number of participants in Head Start programs from the US Department of Health and Human Services (HHS). These data are updated on an ongoing basis; our research uses Head Start data as of April 11, 2023. The data represent the total number of reported funded slots for participants in Head Start programs by center location and include address, latitude, and longitude, which facilitated mapping the data to geographic regions. Home-based provider data were also provided. Geographic information below the block group level is not released due to privacy concerns.

Each observation included the grantee name, grant number, program, address, and reported number of funded slots for a Head Start center or centers where program activities occur. Data for the center locations represent where program participants receive the benefits of Head Start or family child

care programs, not where program participants live. The data were not considered confidential as they did not include participant information.

Environmental and Climate Hazard Data

We collaborated with staff from Head Start and the HHS Office of the Assistant Secretary for Planning and Evaluation (ASPE) and gathered information from key informant interviews to identify 10 environmental and climate hazard indicators to include in our analysis. We gave priority to indicators that are established in the climate change and environmental risk literature as posing risks to children, had data available at small enough geographies to use in a mapping feature, had relatively recent data available, and were of greatest interest or were most relevant to staff from Head Start and ASPE. All environmental data included in the analysis are publicly available to researchers.

EJ SCREEN 2.0 DATA

Five of the 10 indicators included in this feature were collected using data published in the US Environmental Protection Agency's (EPA) EJScreen 2.0, which is an environmental justice mapping and screening tool. The indicators include air quality, ozone concentration, diesel exhaust, airborne carcinogens, and air pollution. The tool uses publicly available and downloadable datasets released by the EPA and the Census Bureau. All indicators from EJScreen used in the feature represent environmental risks as of 2019; the tool was downloaded in March 2023 and included the most recent data available at that time.

Air Quality

Air quality is measured by the levels of particulate matter in the air that are smaller than 2.5 microns (2.5 μm) in diameter ("PM 2.5"). These microscopic particles in the air increase with pollution, wildfire smoke, and industrial and agricultural activity and can damage people's lungs and hearts when inhaled because they are small enough to enter the bloodstream. Repeated exposure to air pollution can lead to long-term health effects for children, including worsening asthma and respiratory illnesses and lung and heart disease later in life. This metric represents the annual average micrograms per cubic meter of PM 2.5 in the air and is sourced from the Office of Air and Radiation fusion of model and monitor data.

Ozone Concentration

Although ozone in the atmosphere helps shield the planet from harmful ultraviolet rays, pollution can increase ozone at ground levels. Higher concentrations of ozone can negatively affect lung function for all people but are especially harmful for children, as repeated exposure to high ozone levels can cause

permanent damage to their still-developing lungs. This metric represents the summer seasonal average of daily maximum 8-hour concentration in the air in parts per billion of ozone and is sourced from the Office of Air and Radiation fusion of model and monitor data.

Diesel Exhaust

Diesel exhaust refers to the soot produced by diesel engines and consists of more than 40 cancer-causing substances. Studies show that children who ride in diesel-fueled school buses or who live along major thoroughfares and other transportation and related industrial infrastructure may have increased risk of asthma and of developing cancer later in life. This metric represents the annual average micrograms per cubic meter of diesel particulate matter in the air and is sourced from EPA Hazardous Air Pollutants data.

Airborne Carcinogens

Air toxins include hazardous air pollutants that have been shown to increase lifetime cancer risk. They are produced by vehicular traffic, power plants, pesticides, wildfires, and other industrial, transportation, and pollution sources. Children are more sensitive to these toxins because their bodies are still developing and they breathe more air relative to their body weight. This metric represents lifetime cancer risk from inhalation of air toxics and is sourced from EPA Hazardous Air Pollutants data.

Air Pollution

EPA's National Air Toxics Assessment (NATA) Respiratory Hazard Index compares an area's exposure and concentration of hazardous air pollutants with the levels of concentration that impair health. Children are more vulnerable to negative impacts from airborne pollution, which can create or exacerbate cardiopulmonary and other health problems. This metric represents the ratio of exposure concentration to health-based reference concentration and is sourced from EPA Hazardous Air Pollutants data.

AMERICAN COMMUNITY SURVEY DATA

Lead Exposure

The federal government banned the use of lead in consumer products in 1978 because of its harmful health effects, including nervous system damage, brain damage, and death. Children's exposure to lead can stunt learning and development and cause hearing and speech problems. We estimated lead exposure using the share of buildings in a community that were built before 1960, which are more likely

to contain lead-based paint. These data derive from the 2017–2021 American Community Survey administered by the Census Bureau. This metric was also used in EPA's EJScreen.

CLIMATE AND ECONOMIC JUSTICE SCREENING TOOL DATA

We drew 1 of the 10 indicators from data used in the White House Council on Environmental Quality's Climate and Economic Justice Screening Tool. The tool includes indicators of environmental and climate burdens across eight categories to help federal agencies identify disadvantaged communities that will benefit from programs and investments through the Justice 40 Initiative.

Paved Land Area

Impervious surfaces, or areas paved over with asphalt and other substances that do not allow water to pass through, are associated with many negative health effects. Studies show that exposure to paved surfaces may negatively affect children's development, while exposure to green spaces and vegetated areas improves it. This metric represents the share of a census tract's land area that is covered with artificial materials, such as concrete or pavement, excluding crop land used for agricultural purposes. Data from 2019 were collected in 2010 census tract geographies. We accessed these data from the Climate and Economic Justice Screening Tool download website. The data were originally developed by the Multi-Resolution Land Characteristics Consortium, with data analysis provided by the Trust for Public Lands and American Forests. These data were downloaded in March 2023 and were the most recent data available at that time.

FIRST STREET FOUNDATION DATA

First Street Foundation is a nonprofit research and technology group dedicated to quantifying and communicating climate risks. It provides flood, wildfire, wind, and extreme heat risk statistical data for noncommercial users. First Street Foundation data for flood risk, wildfire risk, and extreme heat metrics were downloaded in May 2023 and were the most recent data available at that time.

Flood Risk

Coastal, riverine, and severe storm flooding is increasing in severity, intensity, and frequency because of climate change. Flooding can cause power failures and contaminate groundwater, both of which pose health risks. Children are more vulnerable to toxins in the water because their bodies are still developing. This metric represents summed annual projected flood risk over the next 30 years at the census tract and ZIP code levels. Properties are given a risk score based on these data, and the metric derives from the risk scores as the proportion of properties that have severe or extreme risk in a given

geography. Data are provided in 2020 Census geographic terms. These data are generated by the First Street Foundation, released through the use of the Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license, and accessed through Amazon Web Services. This metric was also included in the Climate and Economic Justice Screening tool. See here for a more detailed methodology on these data.

Wildfire Risk

Wildfires release dozens of air pollutants and carcinogens into the atmosphere, all of which can cause long-term lung and heart damage. These risks are greater for children, as they tend to spend more time outdoors than adults and may inhale more smoke as a result. Wildfires can also cause emotional distress for children and other people who live near areas where they occur. This metric represents summed annual projected wildfire risk over the next 30 years at the census tract and ZIP code levels. Properties are given a risk score based on these data, and our metric derives from the risk scores as the proportion of properties that have severe or extreme risk in a given geography. Data are provided in 2020 Census geographic terms. These data are generated by the First Street Foundation, released through the use of the Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license, and accessed through Amazon Web Services. This metric was also included in the Climate and Economic Justice Screening tool. See here for a more detailed methodology on these data.

Extreme Heat

Heat waves are becoming longer, more frequent, and more intense as a result of climate change. Because children have less body mass, they are more susceptible to dehydration and death during extreme heat events than are healthy adults. This metric represents an average of current and projected heat risk over the next 30 years at the census tract and ZIP code levels. Properties are given a risk score based on these data, and our metric derives from the risk scores as the proportion of properties that have severe or extreme risk in a given geography. These data are generated by the First Street Foundation and released through the use of the Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license. The data are provided in 2020 Census geographic terms. See here for a more detailed methodology on this data. Data were accessed through Amazon Web Services.

2020 DECENNIAL CENSUS RACE AND ETHNICITY DATA

We collected 2020 Decennial Census race and ethnicity data at the census block group, census tract, and ZIP code tabulation area (ZCTA) levels to show the demographic makeup of each geography.

Methodology

Map Geographies

Data were aggregated to or collected at three geographic levels for this feature: the census block group level, the census tract level, and the census ZCTA level. **Census block groups** are geographies delineated by the Census Bureau and are clusters of blocks, which are statistical areas bounded by visible and nonvisible features, such as roads, streets, and property or city lines. They generally contain 600 to 3,000 people. **Census tracts** are also geographies delineated by the Census Bureau and were created to provide a stable set of geographic units for the presentation of statistical data. They generally contain 1,200 to 8,000 people. **ZCTAs** are approximate area representations of US Postal Service ZIP code service routes that are approximated from census blocks. We use ZCTAs instead of ZIP codes because ZIP codes are not areal features but rather a collection of mail delivery routes. Geographies depicted in the feature are 2020 spatial files.

Head Start–reported funded slot data were available at the program address level. We conducted a spatial join between the address level and each of the three geographies (census block group, census tract, and ZCTA) and aggregated funded slots by those geographies to determine the total number of reported funded slots in each geography.

Data for metrics from EJScreen were retrieved at the census block group level. All indicators from EJScreen used in the feature were supplied at the census tract level to EJScreen, which were then attributed to block groups. We used a population-weighted block group to ZCTA crosswalk collected from Missouri Census Data Center's Geocorr to approximate EJ Screen indicators to the ZCTA level. The crosswalk was merged to the data at the block group level. Then, we multiplied each indicator by the total population of the census tract from the 2020 Decennial Census and multiplied by the weight supplied by the crosswalk. We also multiplied the total population by the weight in the crosswalk. We next aggregated the indicators and total population at the ZCTA level. Finally, we divided each indicator by the resultant total population to develop the approximated ZCTA metric.

The lead indicator was collected at the block group level and the ZCTA level from 2017–2021 American Community Survey data.

Data for metrics from the First Street Foundation were retrieved at the census tract level and the ZCTA level in vintage 2020.

Data for the paved land area indicator were provided at the 2010 census tract geographic level. Data were adjusted from 2010 census tracts to 2020 census tracts using a 2010 to 2020 census tract crosswalk derived from Brown's Longitudinal Tract Database. Then, we transformed these data to the 2020 ZCTA level using a census tract to ZCTA population-weighted crosswalk, similar to the method we used to transform block groups to ZCTAs for the EJScreen metrics.

We include the county and state in which each selected geography is located. Because ZCTAs do not nest perfectly within counties, we attribute the county with the highest population intersection to the ZCTA using a ZCTA-to-county crosswalk from Missouri Census Data Center's Geocorr.

Head Start Map View and Index

Using administrative and program data from HHS, this map view shows both the number of reported funded slots in each selected geography and the number of reported funded slots as an index, normalized on a 0 to 100 scale across each geography. Two geographic levels are represented in this view: the census block group and the ZCTA. Users can zoom in to see block groups and zoom out to see ZCTAs. Areas that do not have Head Start–funded slots are not included for purposes of the index.

On the map legend, an index score between 0 and 50 represents a "low" number of reported funded slots; an index score between 50 and 75 is "medium"; a score between 75 and 90 is "high"; and a score between 90 and 100 is considered "very high." These are interpreted as a percentile between 0 and 1; for example, the "very high" bucket represents the top 10 percent of census block groups or ZCTAs in terms of the number of reported funded slots.

Exposure Map View and Index

This map view represents an index for each of the 10 environmental indicators normalized on a 0 to 100 scale across each geographic level. A score of 100 means the selected geography has the highest level of relative exposure for the selected hazard across all other geographic regions at the same geographic level.

Data for each hazard indicator are available at two geographic levels, but the levels vary by indicator. For data from EJScreen and the American Community Survey, which include air quality, ozone concentration, diesel exhaust, airborne carcinogens, air pollution, and lead exposure, the geographies available include the block group level and the ZCTA level. For all other indicators, which include paved land area, heat risk, flood risk, and wildfire risk, the geographies available include the census tract level

and the ZCTA level. These views are provided to allow users to examine data at both the ZIP code–approximate level and the smallest geography possible.

On the map legend, an index score between 0 and 50 represents a "low" level of environmental risk or exposure; "medium" is represented by scores between 50 and 75; "high" is represented by scores between 75 and 90; and "very high" by scores between 9 and 100. These scores represent a percentile between 0 and 1; for example, the "very high" bucket represents the top 10 percent of census block groups, census tracts, or ZCTAs in terms of their environmental risk or exposure.

Exposure and Head Start Map View and Index

This map view represents a combined index of selected environmental or climate hazards and Head Start participation for each geography, wherein higher numbers indicate higher levels of relative exposure by environmental indicator and higher numbers of reported Head Start slots.

First, we modify the environmental and climate hazard data used in the exposure map to omit areas that do not have Head Start–reported funded slots, and then re-index Head Start and environmental hazard data on a scale from 0 to 100. Next, data at the ZCTA level are recalibrated for the EJScreen indicators, the lead exposure indicator, and paved land area indicator to weight by the number of funded slots when conducting the geographic transformation rather than total population (see the Geographic Levels section for a full methodology of the geographic transformation) to better represent how the hazard metrics may pose a risk to Head Start recipients based on Head Start–reported funded slot numbers.

Next, the Head Start–reported slot index and each environmental or climate hazard index is averaged on a weighted basis, with the environmental hazard index weighted on a scale of 3 to 1, relative to the Head Start–reported slot index. We weight the environmental hazard index more than the Head Start index to ensure that areas that are low in environmental risk or exposure do not inaccurately populate as "high" or "very high" in the final composite score, indicating substantial environmental or climate hazard exposure, as they would for geographies with very high Head Start–reported funded slots absent weighting.

As in the exposure map view, each composite indicator has two geographic levels available, but the available geographies vary by indicator based on the geographies available in the original hazard data.

The recipient exposure box score represents a percentile where higher numbers indicate higher levels of relative exposure by environmental indicator combined with higher numbers of reported

funded slots relative to other geographies at the same geographic level. For example, a recipient exposure score of 83.7 indicates that 83.7 percent of geographies at the same geographic level have a lower relative environmental risk or lower relative exposure to Head Start participants than the selected geography does.

On the map legend, an index score between 0 and 50 represents a "low" level of environmental risk or exposure to Head Start participants; scores between 50 and 75 are "medium"; scores between 75 and 90 are "high"; and scores between 90 and 100 are "very high." These scores represent a percentile between 0 and 1; for example, the "very high" bucket represents the top 10 percent of census block groups, census tracts, or ZCTAs in terms of their relative hazard risk or exposure to Head Start participants.

Limitations

Although this tool seeks to provide an overview of potential environmental hazards and potential risks or exposures to Head Start participants, gaps in data availability and scale limit its scope.

First, race and ethnicity data are not available for Head Start participants because the program tracks the number of slots rather than participant characteristics. As a rough proxy, race and ethnicity data derived from the American Community Survey are included for selected geographies as a whole.

Second, the number of houses built before 1960 is used as a proxy measure for lead paint exposure. While many homes built before 1960 used lead paint, this measure obscures the true location and number of households affected, as it is not a direct measure of lead exposure or risk.

Third, EJScreen attributes tract-level observations to block groups contained within them. Although not inaccurate, this is less precise than deriving block-group level data directly.

Finally, this feature cannot show the cumulative risk of hazard exposures or compounding risks from exposure to multiple hazards, faced by many locales—that is, the visualization does not depict the cumulative exposure to multiple environmental and climate hazards at once. The research behind this feature can establish a baseline for individual indicators, but examining cumulative and compounding risks is an important avenue for future analyses to better understand the total impact of climate-related harms and inform investments in effective policies and programs.