# Energy Equity:
# Opportunities for Data Science, Machine Learning & AI Policy Impact in the US

An overview of the United States energy landscape, intersections with community social safety net needs, and opportunities for data science and AI to deliver policy impacts.

## Table of Contents

## The Problem Defined

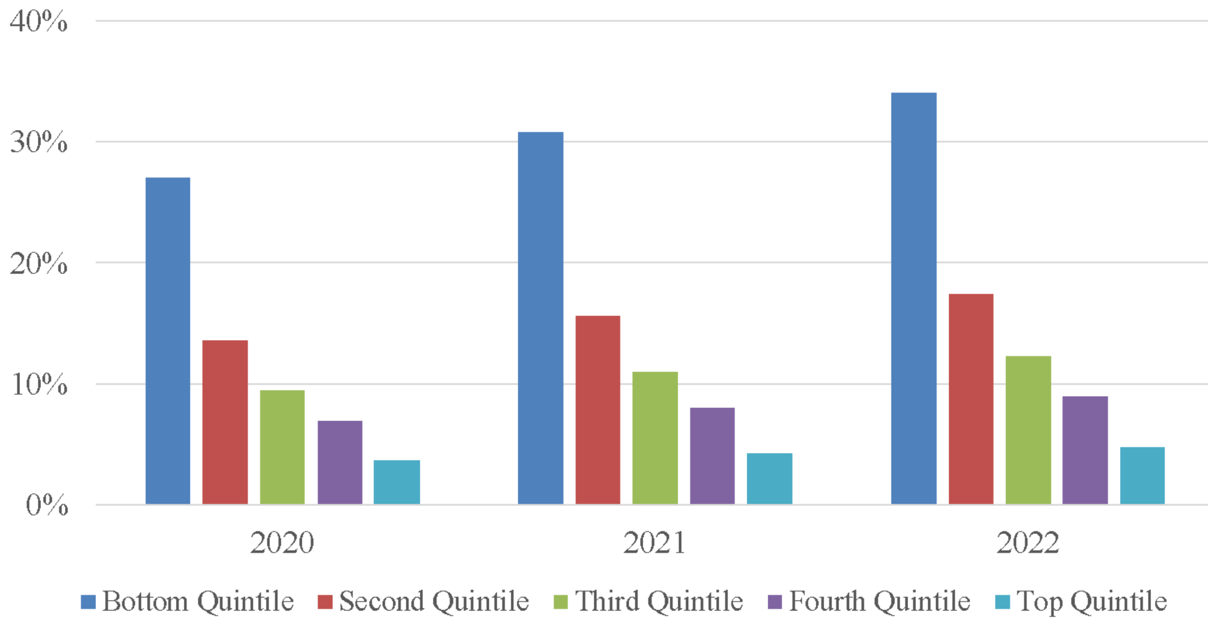
The Biden Administration has prioritized addressing Energy Equity issues as part of its broader push to improve Environmental Justice. Environmental Justice is generally defined as the fair treatment of all people regardless of race, color, national origin, or income regarding the implementation of environmental policies, and the recognition that disadvantaged communities bear a disproportionate share of environmental burdens.[1] Across many studies over the years, similar demographic variables are correlated with higher energy burdens, greater pollution burdens, and lower access to incentives such as federal tax benefits for solar installations or electric vehicle purchases. Income is a key variable, but across all income groups disparities exist with greater burdens in particular born by Black, Native, and Latinx households (as compared to white households).[2]

Energy Equity includes a variety of dimensions including reducing energy insecurity in low-income households as well as improving the delivery of benefits to (and participation in policymaking of) disadvantaged communities. Energy Equity issues in the United States are not necessarily driven by a physical lack of access (as one would expect in developing countries) but rather access to energy services at a reasonable percentage of income and a lack of fair access to clean energy and energy efficiency measures that could lower energy bills in disadvantaged communities, reduce pollution impacts, and increase overall health and well-being.

On average, households in the lowest quintile income bracket spend nearly a third of their income on energy (including gasoline). In comparison, the top quintile of households spends less than 5% of their income on energy.[3]

[1] https://www.epa.gov/environmentaljustice/learn-about-environmental-justice
[2] https://seas.umich.edu/sites/all/files/2022_EEP_Report.pdf?utm_source=pr&utm_campaign=eep&utm_id=eep+framework
[3] https://neada.org/wp-content/uploads/2022/06/Energy-Expenditures-and-Burdens-A-June-2022.pdf

*Figure 1: Share of Household Income Spent on Gasoline and Utilities by Income Quintile*[4]




Over one quarter of U.S. households reported some type of energy insecurity in 2020. This number was down slightly from the previous survey (done in 2015), which may have been attributable to the fact that many states issued moratoria on utility shut-off notices in the latter part of 2020 in response to the COVID-19 pandemic.[5] A different survey in 2022 found increased rates of energy insecurity with 31% of households reporting they forwent basic necessities in order to pay an energy bill and 20% of all households kept their home at a temperature that felt unsafe or unhealthy at least one month of the previous year. Unsurprisingly, these percentages were much higher for low-income households with 52% reporting they forewent necessities to pay an energy bill at least one month, and 32% reported keeping their home at a temperature that felt unsafe or unhealthy.[6]

Households in low-income and disadvantaged communities have less ability to afford capital expenses to fund energy efficiency upgrades or clean energy installations, meaning not only are they paying a

---

[4] https://neada.org/wp-content/uploads/2022/06/Energy-Expenditures-and-Burdens-A-June-2022.pdf
[5]https://www.eia.gov/todayinenergy/detail.php?id=51979#:~:text=In%202020%2C%2027%25%20of%20U.S.,difficulty%20meeting%20their%20energy%20needs&text=In%202020%2C%2034%20million%20U.S.,because%20of%20energy%20cost%20concerns
[6] https://neada.org/wp-content/uploads/2022/06/Energy-Expenditures-and-Burdens-A-June-2022.pdf

percentage of their income for energy services, but they are also using more energy to achieve the same levels of comfort or energy service. Lower income households are also more likely to rent rather than own their residence, which may reduce incentives to make energy efficiency improvements. Property owners may be reluctant to spend money to improve energy efficiency when the money savings accrue to the renters; and the renters, even if they had resources to pay for energy efficiency improvements, are unwilling or unable to pay for improvements to property they do not own.

Federal and state programs aimed at facilitating energy access exist, but there are additional barriers in implementing these programs effectively for low-income and disadvantaged communities.[7] There is greater housing instability among lower income households making it more difficult to track the households receiving energy assistance. Lower income households have reduced access to basic financial services such as checking accounts, credit cards, or mobile banking, which makes it more difficult for them to pay bills. Similarly, even where there may be funds available to assist with capital expenses of energy efficiency upgrades, the lack of access to financial services can negatively impact the effectiveness of those programs. Where low-income or disadvantaged communities are in rural areas the geographic dispersion increases the challenges of addressing energy access.[8] Finally, digital platforms and smart-meter technology are being deployed as effective means to incentivize energy efficiency measures and the purchase of energy-efficient products, but low-income households often lack the internet access to benefit from these measures.[9]

An additional obstacle to developing federal and state policy is that there is not a single, federally recognized definition of energy poverty or energy equity or specific legislation to address it.[10] There are two, primary federal assistance programs that provide resources to improve energy equity, the Low Income Home Energy Assistance Program (LIHEAP) and the Weatherization Assistance Program (WAP),

---

[7] https://betterbuildingssolutioncenter.energy.gov/sites/default/files/IB_Using%20Data%20to%20Drive%20Low%20Income%20Program%20Success_final_4.pdf

[8] These barriers are outlined and discussed in greater depth in the DOE Better Buildings program Issue Brief on the Clean Energy for Low-Income Communities Accelerator (CELICA). https://betterbuildingssolutioncenter.energy.gov/sites/default/files/IB_Using%20Data%20to%20Drive%20Low%20Income%20Program%20Success_final_4.pdf

[9] At page 54 https://info.ornl.gov/sites/publications/Files/Pub124723.pdf

[10] https://www.nature.com/articles/s41560-020-0582-0#:~:text=Thus%2C%20we%20propose%20to%20define,factors4%2C40%2C41

administered by the Department of Health and Human Services and the Department of Energy. These programs were initiated in the 1970s, in response to the rising energy costs associated with the OPEC oil embargo and subsequent geopolitical instability. They have been renewed and expanded multiple times since and have provided assistance to millions of households. Their reach, however, is limited as annually only about 25% of LIHEAP eligible households receive benefits and approximately 17% of WAP-eligible households have been weatherized.[11]

The LIHEAP and WAP programs annually disburse approximately $3.7 billion and $300 million, respectively. State, local, and utility ratepayer funded assistance programs disburse an additional $4 -$5 billion annually. There have been increasingly organized efforts to gather and utilize data more systematically and more effectively to improve the delivery of services under these various programs. These efforts, particularly efforts to incorporate more sophisticated data-science tools, remain in their infancy.

There is an additional pressing need to improve the targeting and effectiveness of energy-equity programs rapidly because recent federal legislation is directing tens of billions of dollars over the remainder of this decade on energy equity. The 2021 Bipartisan Infrastructure Law (also referred to as the Infrastructure Investment and Job Act) allocated $3.5 billion to the WAP.[12] The 2022 Inflation Reduction Act[13] included billions of dollars to expand energy access, particularly access to energy efficiency and clean energy for low-income households.[14] These new efforts include $4.5 billion to provide rebates (up to 100%) for lower-income families to install energy-efficient appliances (including heat pumps); $4.3 billion for energy efficiency measures throughout houses, $1 billion for retrofits to increase energy and water efficiency in HUD-administered affordable housing,[15] and a $4,000 tax credit for used electric vehicles (which could help open the electric vehicle market to more households).[16]

---

[11] Ibid.
[12] https://www.energy.gov/sites/default/files/2022-03/wpn-bil-22-1.pdf
[13] Public Law 117-169, https://www.congress.gov/bill/117th-congress/house-bill/5376/text
[14] https://thehill.com/opinion/energy-environment/3598459-inflation-reduction-act-helping-low-income-families-adapt-to-rising-temperatures/
[15] https://neada.org/wp-content/uploads/2022/08/Inflation_Reduction_Act_LMI_Energy_Provisions.pdf
[16] https://thehill.com/opinion/energy-environment/3598459-inflation-reduction-act-helping-low-income-families-adapt-to-rising-temperatures/

These additional federal funds greatly expand the resources flowing from the Federal Government to address energy access. The requirements to distribute the IRA funds on short timelines (some funding blocks need to be allocated in two or three fiscal years) highlights the need to vastly improve the targeting of energy access improvements to make sure the most vulnerable and most in need receive their fair share. The next section describes the current and ongoing funding efforts in more detail and introduces the efforts that have been made to improve the gathering and use of data.

## Existing Efforts and Resources

The single largest source of funding aimed at improving energy equity (although it is not specifically identified as such) are the two federal programs: LIHEAP provides financial assistance to assist low-income households with high energy expenditures; and WAP subsidizes low-income families wanting to make whole-house energy efficiency upgrades. These programs focus on income eligibility (based on either percentage of federal policy level, a percentage of state median income, or participation in certain federal aid programs), energy burden (percent of income spent on energy) and demographic characteristics.[17]

The federal funds are distributed as block grants to the states, territories, and tribes, which then administer them at the state and local levels.[18] In recent years, the LIHEAP has been appropriated approximately $3.7 billion annually, with an additional $5.5 billion in supplemental appropriations in response to the COVID-19 pandemic.[19] The WAP has received approximately $300 million annually in recent years.[20] This current rate of annual funding generally reflects historic trends as the annual LIHEAP funding has on average been about 10 times greater than WAP funding, save for a spike in WAP funding from the 2009 American Recovery and Reinvestment Act (ARRA) and energy efficiency funding in the 2022 Inflation Reduction Act.

---

[17] https://www.nature.com/articles/s41560-020-0582-0#:~:text=Thus%2C%20we%20propose%20to%20define,factors4%2C40%2C41

[18] LIHEAP available at: https://www.acf.hhs.gov/ocs/fact-sheet/liheap-fact-sheet; WAP available at: https://www.energy.gov/eere/wap/how-apply-weatherization-assistance

[19] https://www.acf.hhs.gov/ocs/fact-sheet/liheap-fact-sheet;

[20] https://crsreports.congress.gov/product/pdf/R/R46418

These annual appropriations are not sufficient to cover the energy access needs nationally. Overall LIHEAP serves only approximately 25% of eligible households[21] and the percentage served varies greatly by state. A study in 2022 found that only a few states serve as many as 40% of eligible households, most states serve less than 20%, and the poorest performing states reach less than 10%.[22] Similarly, while DOE states that WAP has benefited 7.5 million homes in the program's history, nearly 40 million households are still eligible for assistance and in most years less than .2% of eligible homes receive assistance.[23]

The 2022 Inflation Reduction Act (IRA) will be transformational for the U.S. efforts to improve energy access and reduce energy poverty. It includes billions of dollars in funding for new programs focused on environmental justice, including:

- $4.5 billion to provide rebates (up to 100%) for lower income families to install energy efficient appliances (including heat pumps);
- $4.3 billion for additional energy efficiency measures;
- $27 billion for a Greenhouse Gas Reduction fund,[24] which will launch a national green bank, with 60% of those funds available to finance clean energy in low-income and disadvantaged communities;
- $1 billion for retrofits to increase energy and water efficiency in HUD-administered affordable housing, and
- $4,000 tax credit for used electric vehicles (which could help open the electric vehicle market to more households).[25]

The federal funds represent the largest chunks of funding aimed at improving energy equity, but a greater overall level of funding comes from state and local sources. In 2014 (the most recent year for which data is available) state and local sources provided an additional $3.4 billion in rate assistance and $1 billion devoted to energy efficiency improvements.[26] The bulk of these additional funds comes from states that mandate rate-payer funded programs, e.g. utilities in those states collect an additional fee

---

[21] https://www.nature.com/articles/s41560-020-0582-0#:~:text=Thus%2C%20we%20propose%20to%20define,factors4%2C40%2C41
[22] https://rmi.org/by-the-numbers-low-income-energy-assistance/
[23] https://www.aceee.org/blog-post/2020/07/weatherization-cuts-bills-and-creates-jobs-serves-only-tiny-share-low-income
[24] https://fortune.com/2022/08/22/climate-inflation-reduction-act-national-green-bank-environment-biden-andrei-cherny/
[25] https://neada.org/wp-content/uploads/2022/08/Inflation_Reduction_Act_LMI_Energy_Provisions.pdf
[26] https://liheapch.acf.hhs.gov/Supplements/2014/supplement14.htm

from all rate-payers to support the assistance programs. The last year further-disaggregated data is available is from 2010, and that indicates that state and local taxes provided $250 million in assistance and charitable and non-profit organizations provided approximately $130 million.[27]

In addition to the ongoing assistance programs and the specific measures in the Bipartisan Infrastructure Law and the Inflation Reduction Act, the current Administration has announced a Justice40 initiative, which aims to provide 40% of overall benefits of federal investments in seven key areas to disadvantaged communities.[28] Those areas are:

- Climate change
- Clean energy
- Energy efficiency
- Clean transit
- Affordable and sustainable housing
- Training and workforce development
- Remediation and reduction of legacy pollution
- Development of critical clean-water infrastructure

The next section outlines the fact that although increased efforts have been made in the past decade to improve data collection, establish metrics, and better utilize data for policy planning and program delivery, more advanced data-science techniques are required to successfully meet (and verify) these various funding goals.

## Data and Metrics

Most of the datasets currently being utilized in efforts to improve energy equity are derived from more general surveys conducted by federal agencies.

The Federal Government collects information on the burden of energy expenditures on households through several channels. The Energy Information Administration (EIA) conducts a regular Residential

[27] https://liheapch.acf.hhs.gov/Supplements/2010/supplement10.htm
[28] https://screeningtool.geoplatform.gov/en/about

Energy Consumption Survey (RECS) that includes information on energy insecurity and energy Poverty.[29] The RECS survey includes questions about the characteristics of households (including physical characteristics of homes as well as demographic and income information) and questions about energy-use behaviors combined with information from energy suppliers on household fuel use and expenditures.[30] The Bureau of Labor Statistics conducts annual Consumer Expenditure Surveys that include information on energy expenditures and income that can be used to calculate household energy cost burden.[31]

The Census Bureau has been conducting a frequent, experimental Household Pulse Survey to collect data on how the Covid-19 pandemic has impacted people, which has included questions about ability to pay energy bills, whether households forwent other necessities to pay energy bills, and whether the households kept their homes at a temperature that felt unsafe or unhealthy (these are similar to questions asked on the less-frequent RECS survey).[32]

These surveys suffer from similar shortcomings in that, other than the Census Pulse Surveys, there are multi-year delays between gathering and disseminating data. The Pulse Surveys are designed to be rapid turnaround survey products that reduce the dissemination time of data to a matter of weeks. It is unclear how long the Census Bureau will continue the Pulse Surveys and they are an experimental product which may not meet the same data standards as the Census Bureau's usual products. [33]

The Environmental Protection Agency (EPA) built an Environmental Justice Screening tool – EJScreen – from the above data sets as well as sources of information on environmental burdens. EJScreen is the Federal Government's most comprehensive collection of data related to environmental burdens to disadvantaged communities.

---

[29] https://www.eia.gov/todayinenergy/detail.php?id=51979#:~:text=In%202020%2C%2027%25%20of%20U.S.,difficulty%20meeting%20their%20energy%20needs&text=In%202020%2C%2034%20million%20U.S.,because%20of%20energy%20cost%20concern

[30] https://www.eia.gov/consumption/residential/background-technical.php

[31] https://neada.org/wp-content/uploads/2022/06/Energy-Expenditures-and-Burdens-A-June-2022.pdf; https://www.bls.gov/cex/

[32] https://neada.org/wp-content/uploads/2022/06/Energy-Expenditures-and-Burdens-A-June-2022.pdf; https://www.census.gov/programs-surveys/household-pulse-survey/data.html

[33] https://www.census.gov/data/experimental-data-products/household-pulse-survey.html

EJScreen includes demographic information along with data on indicators of 11 environmental burdens such as air pollution, toxic substances, lead paint, chemical plants, and waste management facilities.[34] The technical documentation for EJScreen also notes that it does not include data on overall health indicators, drinking water or surface water quality, radon gas exposure, criteria air pollutants other than $PM_{2.5}$ and ozone, exposure to undocumented emissions caused by lease, exposures related to oil and gas extraction, combines animal feeding operations, pesticide exposures, noise pollution. In some instances, indicators for these other environmental burdens were not included because they substantially overlap in coverage with existing indicators – but EPA also noted that the most common limitation on including these other indicators was the lack of data, particularly high-resolution data (at the census tract level, census-block level or finer). [35]

EPA also emphasizes that the EJScreen should be used only for screening purposes because of the uncertainty of data at the finer scales (i.e., census block level). The demographic indicators for those smaller groups are based on surveys rather than a full census. Also, the environmental indicators are proxies for actual health impacts and uncertainty is especially acute for the "proximity" indicators (i.e., distance from a chemical plant or waste facility).[36]

In the late 2010s the Department of Energy (DOE) operated a voluntary partnership with key states focused on energy equity called the Clean Energy for Low-Income Communities Accelerator (CELICA). A key output of the CELICA effort was the development of the first federal Low Income Energy Affordability Data (LEAD) tool by the National Renewable Energy Laboratory (NREL).[37] The LEAD Tool combines census and energy data (including costs) to allow users to examine the energy burden (energy costs as a percentage of income) across multiple factors. The tool can provide results:

- From the national down to the census tract level (although energy consumption data is not available down to the census tract level);
- Based on different income levels (either as percentages of median income or as percentages of federal poverty level);

[34] https://www.epa.gov/sites/default/files/2021-04/documents/ejscreen_technical_document.pdf
[35] Ibid.
[36] https://www.epa.gov/ejscreen/limitations-and-caveats-using-ejscreen
[37] https://www.energy.gov/eere/slsc/maps/lead-tool

- Based on housing ownership characteristics (own or rent, number of units in building, age of building, primary heating fuel type);
- Cost of energy either as average monthly expenditures or energy burden as a percentage of income spent on household energy (excluding fuel for transportation).[38]

According to the Department of Energy, stakeholders have used the LEAD Tool to understand income characteristics in their community, improve outreach, and start new programs.

> The LEAD Tool has been used by stakeholders to improve understanding of low-income and moderate-income characteristics in their locality, identify target areas, start new low-income programs, and use the information for outreach or educational purposes. The tool can also be used to inform strategic planning or to support independent research.[39]

To effectively implement and track progress on the Justice40 initiative, agencies have increased their efforts to improve the availability and use of data. The Council on Environmental Quality (CEQ) is a Climate and Economic Justice Screening tool to help Federal agencies identify disadvantaged communities that are overburdened by environmental degradation.[40] The tool identifies "disadvantaged communities" at the census tract scale if the tract is at or above the 65th percentile for low-income and at the 80th percentile or above for the percentage of residents above 15 years old not enrolled in college and has a disproportionate burden related to one of the seven focus areas. The data sources it draws on are from the Census Department, FEMA's National Risk Index, LEAD Tool developed by NREL, EPA's EJ Screen, Department of Housing and Urban Development, and Centers for Disease Control PLACES data.[41]

In conjunction with the Justice40 initiative, the DOE has launched an Energy Justice Dashboard (EJD) to better understand and track how DOE's funding and investments benefit overburdened and disadvantaged communities.[42] The EJD combines data from a Low Income Energy Affordability (LEAD) tool (described below) developed previously by DOE with extensive data on overall environmental

[38] https://lead.openei.org/assets/docs/LEAD-Tool-Methodology.pdf
[39] https://lead.openei.org/assets/docs/LEAD-Factsheet.pdf
[40] https://screeningtool.geoplatform.gov/en/about
[41] https://screeningtool.geoplatform.gov/en/methodology
[42] https://www.energy.gov/diversity/energy-justice-dashboard-beta

burdens experienced by disadvantaged communities that is maintained by the EPA in an Environmental Justice Screening and Mapping Tool (EJScreen).[43]

DOE National Renewable Energy Laboratory (NREL) has pulled together many of these above-described data sets into an Equitable Energy Investment Prioritization Data Set, including from SLOPE, EJScreen, LEAD, REopt, and USDA Rural Atlas.[44] The Department of Energy also maintains resources such as a State & Local Energy Efficiency Action Network to provide tools to local regulators[45] and the State and Local Planning for Energy (SLOPE).[46] SLOPE is an online platform that allows local officials to utilize DOE data to inform their energy planning, including energy modeling tools and a data viewer providing information on energy costs, emissions, renewables deployment, efficiency, and transportation at a city, county, and state level. Finally, DOE National Renewable Energy Laboratory maintains a variety of geospatial data science applications and visualizations such as the National Solar Radiation Database.[47]

As mentioned, much of the funding for these various federal programs is done through state-block grants. In addition, some states have adopted their own definitions and metrics to define disadvantaged communities and improve energy equity. Over half the states have either enacted legislation or implemented regulations formally identifying Environmental Justice communities to improve the targeting and delivery of services and investments.[48] New York State has gone as far as adopting a requirement similar to the Justice40 initiative for disadvantaged communities to receive at least 40% of the overall benefits of funding for spending on clean energy and energy efficiency programs.[49]

Combined with the fact that the federal funds are distributed to states as block grants to implement assistance programs at the local level, the volume of these state and local resources highlights the importance of providing data-driven solutions that are adaptable and deployable by smaller state and

---

[43] https://www.epa.gov/ejscreen
[44] https://data.nrel.gov/submissions/175
[45] https://www.energy.gov/eere/state-and-local-energy-efficiency-action-network-see-action. The Environmental Protection Agency also keeps a list of resources for state and local energy efficiency program sponsors. https://www.epa.gov/statelocalenergy/local-utilities-and-other-energy-efficiency-program-sponsors
[46] https://maps.nrel.gov/slope
[47] https://maps.nrel.gov/solar-for-all/
[48] https://docs.google.com/spreadsheets/d/1TgwZAAlmLlBYJciXSpUfOOZifuk8MJUU4_NagVnulU8/edit#gid=0
[49] https://www.nyserda.ny.gov/ny/disadvantaged-communities

local agencies that may be relatively under-resourced even in areas that do not serve primarily disadvantaged communities.

The University of Michigan has launched an "Energy Equity Project" at the School of Environment and Sustainability. The goal of the project was to create a national framework for advancing equity in clean energy programs. It identified a variety of data that would be necessary to support efforts to improve energy equity and examined nearly 150 existing datasets.[50] The project made recommendations on whether the data sets should be included in quantitative metrics, qualitative best practices, or eliminated from consideration and assessed where additional data might be needed.[51] A key output of the project was the identification of 26 national, census tract level datasets that are being used to create an interactive tool. These datasets primarily come from U.S. government sources such as the census surveys referenced above, federal emergency management databases, or Department of Energy EIA datasets. The project also produced an extensive report providing transparency on how the metrics and datasets were identified and assessed, as well as explaining important historical background on energy and environmental justice and the importance of taking an intersectional approach to assessing energy equity to help ensure clean energy and climate interventions benefit all communities.

The project also identified priority data gaps. One set of priority data gaps were in the area of affordability, household benefits (such as energy savings by race and health benefits of clean energy programs by demographic segment) and broader community benefits such as job creation and quality.[52] Another key data gap is the rate of utility shut offs by utility and demographic variables. As described below, the LIHEAP program has only in the past few years started to attempt to gather this data and report at the state level, but no national data set yet exists, the resolution on the data that is being gathered is at a high level, and the data does not identify disconnection rate differences across the relevant demographic variables.

---

[50] Energy Equity Project, 2022. "Energy Equity Framework: Combining data and qualitative approaches to ensure equity in the energy transition." University of Michigan – School for Environment and Sustainability (SEAS)

[51] https://docs.google.com/spreadsheets/u/0/d/1Pfh9VlOYJvCwqe_hkiMT3wLRz_aiHxHroS-CSY-npLk/htmlview?urp=gmail_link#

[52] https://seas.umich.edu/sites/all/files/2022_EEP_Report.pdf?utm_source=pr&utm_campaign=eep&utm_id=eep+framework

The implementation and review of the LIHEAP program provide an illustrative example of the data challenges associated with implementing and measuring the effectiveness of energy-equity programs. It was only in 2016 that LIHEAP identified four performance measures: [53]

- **Benefit Targeting Index** – are the households with the highest energy needs receiving the highest benefits;
- **Burden Reduction Targeting Index** – are the households with the highest energy needs the ones with the largest share of reductions on their bills;
- **Restoration of Home Energy Service** – number of times households lost energy service and had it restored by LIHEAP; and
- **Prevention of Loss of Home Energy Service** – number of times households would have lost energy services without LIHEAP assistance.

The reporting on these measures relies on tracking in each individual state and is largely based on survey methods. The most recent annual report from FY 2017[54] noted at that time only thirty states had moderate or better data reliability for the four measures. LIHEAP maintains a database, but it is not publicly accessible. As it stands, the data on performance measures do not provide sufficient information to allow policy-makers to actually judge the extent that LIHEAP improves the health and wellbeing of recipients (and eligible recipients) and is successful in targeting the most resources to those most in need.[55] LIHEAP also has not been subjected to a comprehensive, national program evaluation.[56] WAP has been independently evaluated and studied monetizing both energy savings and broader benefits, but has not been assessed by its effectiveness at reducing energy poverty.[57]

## Direction of Data Use and Areas for Engagement

While these traditional surveys and data-collection efforts have been underway for decades, there are limited examples of entities (the government or otherwise) beginning to tap into the power of data analytics to address energy equity. Most of the data efforts of the past twenty years have focused on organizing publicly available datasets, identifying better quantitative metrics for assessing ongoing

---

[53] https://liheappm.acf.hhs.gov/what-are-pm/
[54] https://liheappm.acf.hhs.gov/sites/default/files/private/notebooks/2017/FY17_HEN_Part3_Perf_Meas.pdf
[55] https://www.nature.com/articles/s41560-020-0582-0#:~:text=Thus%2C%20we%20propose%20to%20define,factors4%2C40%2C41
[56] Ibid.
[57] Ibid.

energy equity policies, and better defining what is meant by energy equity or energy justice. There remains, however, substantial room for improvement in utilization of that data and more sophisticated data science methods will play a key role in driving that improvement. This section identifies key areas where not only can data analytics improve efforts to improve energy equity, but it is also essential to making those improvements.

- National Grid utilized outputs from the LEAD Tool on the number and geographic distribution of households eligible for a weatherization program then compared them with internal enrollment data to better target marketing efforts.
- Kentucky Office of Energy Policy used the tool to identify counties with the highest energy burden to prioritize issuing grants to local nonprofits implementing energy efficiency programs.
- Rochester, New York and New Haven, Connecticut utilized the tool and other data sets to better understand the distribution of extremely low-income households in the city for targeted energy efficiency upgrades; they also gained insights on relationships between fuel-type and housing unit size for extremely low-income citizens.
- Carrboro, North Carolina was able to identify that targeting energy efficiency measures of replacing bottled-gas for heating with heat-pumps would best target residents with the highest energy burden.
- Connecticut has a goal to weatherize 80% of homes by 2030. The state used the LEAD tool to target their program on single-family owner-occupied homes and large multifamily buildings, which account for nearly 2/3 of low-income households in the state. Officials also then utilized a Nielson customer market segmentation analysis to gain additional insight into customer behavior and biases they were likely to encounter in designing their outreach efforts to the target households.

Producing and Utilizing Finer-Scale Data

The publicly available datasets typically only resolve to the census tract (or sometimes to the census-block) level. Although less-readily available data (including proprietary or non-public data) may be available at finer scales, that will not always be the case.

This finer-scale data, however, could be key to better understand the problem of energy equity, and to better target the limited resources currently available to improve equity. For example, one study utilized finer-scale data (down to the household level) to examine differences in energy use intensity (EUI) between the lowest and highest income groups within two individual cities.[58] Previous studies, which

[58] https://www.pnas.org/doi/10.1073/pnas.2023554118

did not use fine-scale data, had found the poorest households had a 25% greater EUI than the richest households. The fine-scale data study found the disparity actually ranged from 27% to 167% for income and a disparity of 40% to 156% for race. The study also showed that conducting their analysis using data averaged on a larger spatial scale masked some of this disparity.

The narrative in the EUI study emphasized some of the further data limitations, including the fact that EUI, even at a fine scale, can obscure the true level of inequity because it misses changes in behavior. Energy surveys indicate low-income households are more likely to tolerate uncomfortable or unsafe temperature levels because of the reduced ability to pay energy bills. Just measuring EUI or energy burden (energy costs divided by income) will fail to capture those behavioral changes and will understate the inequity.

A study was able to move towards quantifying these behavioral differences between low income and high-income households. Using actual energy use data (from a local utility) cross-referenced with daily temperature changes and household demographic information, researchers were able to quantify the difference in outdoor temperatures households of different income levels tolerated before increasing air-conditioner use.[59] This type of quantitative measure of behavioral differences could be an important complement to the qualitative data derived solely from survey results (i.e., self-reported instances of keeping households at an "unsafe" or "unhealthy" temperature).

Filling Data Gaps

Even at the census tract level, there remain some crucial gaps in the publicly available. For example, the EIA's RECS surveys provide one of the key datasets the Federal Government has on household energy consumption. But RECs only survey several thousand homes nationwide, far too few to provide data at census tract level. Thus, tools like LEAD build on the RECs data to estimate average household energy consumption at the census tract level from a top-down approach.

Machine learning approaches could be used to improve those estimates. One study used the RECS survey results, information from the Census Bureau's American Community Survey (ACS), and the

[59] https://www.nature.com/articles/s41467-022-30146-5#Sec8

Census Bureau's Public Use Microdata Samples (PUMS) to build a bottom-up model of household energy use at the neighborhood scale. The RECS survey includes energy consumption data but is not geolocated. The PUMS data is more fine-scale but does not contain energy consumption data. The study effort statistically matched the RECS and PUMS data (using variables consistent to each) then brought in additional energy use information for the geographic areas and built out the model that creates a complete synthetic population of households with energy consumption data.

In another example of gap-filling and utilizing finer-scale data, Oak Ridge National Laboratory is developing a tool that can analyze the impacts of climate change in urban environments across different socio-economic groups down to street level.[60] Scientists at the lab utilized micro-data from the Census on income, age, gender, ethnic groups and housing status. They then used machine learning techniques to generate building characteristics (based on satellite imagery) to classify different buildings. The result (the work is in progress) would be able to apply down-scaled impacts from climate models on specific urban areas to assess climate impacts at a granular scale in an urban area.

Existing data gaps present a particular challenge in the implementation of the current weatherization (WAP) programs, particularly in the way the deferral program works. The federal rules specify that homes should be deferred from receiving weatherization funds if inspectors find significant issues that would make the weatherization work ineffective, such as other repair needs, health or safety concerns, or exceedingly high repair costs.[61] There may be other subsidy programs that can assist with non-energy repairs, but these often lack sufficient funds. Trying to piece together funding from a variety of different programs creates additional complexity and barriers to implementation. Also, it almost goes without saying that many of the disadvantaged households do not have their own funds for these other repairs. The recent legislation (and other legislative proposals) has devoted some additional resources to help with these non-energy repairs. There remains, however, a large data gap around this problem because agencies typically do not track households that have been deferred to see who may return after fixing those problems, and generally these deferrals are not tracked in any of the state or federal statistics.

---

[60] University of Michigan, Energy Equity Project Summary of EEP Metrics available at: https://docs.google.com/spreadsheets/u/0/d/1Pfh9VlOYJvCwqe_hkiMT3wLRz_aiHxHroS-CSY-npLk/htmlview?urp=gmail_link

[61] https://www.npr.org/2022/05/13/1096114029/low-income-energy-efficient-weatherization-program-3-5b-needy

Improving Efforts to Overcome Barriers

Another key avenue where data analytics could be deployed is to help overcome barriers to the uptake of existing energy equity efforts with the most marginalized and most in need individuals. As noted, there are a variety of barriers to implementing energy equity programs related to housing instability, inadequate access to financial resources, lack of information about available energy assistance programs, and unequal access to internet and broadband resources.

A first step for overcoming some of these barriers is increasing awareness among low-income communities of the resources that are available and marketing those resources as desirable and attainable. Retail businesses increasingly utilize sophisticated data analysis to segment their target markets more effectively and customize marketing efforts to the specific segments. One non-data-dependent approach to overcome these barriers has been to develop community-based programs to assist harder-to reach individuals.[62] Such community-based programs could improve their outreach by utilizing data science to identify the individuals to target in an outreach program and to better tailor the outreach strategies.

Data analytics can identify unknown barriers and help design the most effective “nudges'' to incentivize individuals to utilize available programs or adopt more energy-efficient practices. Assessing the effectiveness of different policy interventions can be improved with data analytics. One study used agent-based modeling to assess how different policy interventions would impact the uptake of rooftop solar programs among different groups.[63] Another mapped the relationship between energy burden and household evictions in Virginia’s counties.[64] Yet another has utilized high-resolution data to improve investment in demand-side management to help keep down utility rates by avoiding additional higher-cost generation and help track if lower-income households are adequately benefiting.[65]

---

[62]https://www.tandfonline.com/doi/abs/10.1080/13549839.2015.1136995

[63] Using agent-based models to examine policy effectiveness. https://link.springer.com/article/10.1007/s10458-016-9326-8

[64] https://info.ornl.gov/sites/publications/Files/Pub124723.pdf

[65] https://info.ornl.gov/sites/publications/Files/Pub124723.pdf

Producing Dynamic and Predictive Models

A key limitation of existing datasets is that they are not done annually, or even where they are annual may not be updated for several years after their completion. Energy burden, on the other hand, might change significantly on a month-to-month basis based on fluctuations in energy prices or changes in income streams or expenses in vulnerable households. And on top of this the fact that the energy burden will be changing going forward we experience the impacts of climate change. As one example, the formulas for distributing LIHEAP and WAP funds to the states were developed based on a history of the greatest energy assistance needs being for cold-weather states in the winter. Increasingly, if there is a need for energy assistance to provide more cooling as communities deal with hotter summers.[66]

One data-driven tool to help address this are urban heat maps.[67] HEAT.gov has heat maps produced through innovative public-private cooperation, combined with community science, and advanced data analytics. CAPA Strategies has engaged local volunteers to gather highly localized heat and humidity data in urban areas.[68] This direct-measured heat and humidity data is then combined with satellite imagery and databases on land cover and applies machine learning to create detailed map of heat variation.[69] These maps are then further cross-referenced with a variety of databases maintained by federal agencies such as FEMA, EPA, and Centers for Disease Control. Among the outputs is a heat equity map that combines the localized heat data with data from the White House Climate and Environmental Justice screening tool to identify and examine how disadvantaged communities are exposed to higher temps. They have also produced a map projecting the expected increase of 90 degrees and higher days under a given emissions scenario.

---

[66] https://www.washingtonpost.com/climate-environment/2022/09/17/us-safety-net-was-built-cold-winters-hot-summers-threaten-it/
[67] http://Heat.gov
[68] https://www.capastrategies.com/heat-watch
[69] https://www.mdpi.com/2225-1154/7/1/5