Disrupting Food Insecurity:

Tapping Data for Strategies That Tackle the Root Causes

TECHNICAL APPENDIX

Caroline Ratcliffe, Elaine Waxman, Cary Lou, Hannah Hassani, and Victoria Tran

Creating County Peer Groups

We use cluster analysis to group the more than 3,100 US counties into ten "county peer groups" (or clusters). Our approach groups counties that are similar in terms of food insecurity, physical health, housing and transportation costs, income and employment, financial health, demographics, and geography.[1] To do this we use a hierarchical clustering model with Ward's method and a Euclidean distance measure.

We include 27 metrics in the cluster analysis grouped under seven broad categories: food insecurity, physical health, housing and transportation costs, income and employment, financial health, demographics, and geography. Only a subset of metrics used to create the county peer groups is displayed on the dashboard, but all of the metrics are available in a downloadable file containing dashboard data (county, peer group, state, national) and additional contextual information, along with an easy-to-use tool that allows users to compare metrics for custom groups of counties.[2]

These metrics are drawn from multiple sources:

- credit bureau data;[3]
- the US Census Bureau's decennial census and American Community Survey (ACS);[4]
- data from the County Health Rankings (CHR) drawn from various Center for Disease Control (CDC) and US Department of Agriculture (USDA) sources;
- data from Feeding America, the National Low Income Housing Coalition (NLIHC), the Center for Neighborhood Technology (CNT), and the Bureau of Labor Statistics (BLS).

The credit bureau data are from 2017. We use 2016 ACS data where possible, but for smaller counties with populations fewer than 65,000 for which one-year estimates are not available, we use the 2012–16 five-year ACS. The year(s) represented for data from other sources is listed below.

For the cluster analysis, all metrics were standardized to a comparable scale of 0–100, and some were normalized to help ensure the analysis was not sensitive to the scale and units of measure or outliers in skewed distributions.[5] Additionally, highly correlated metrics were analytically down-weighted to ensure they were not unintentionally influential, while key metrics in the analysis (food insecurity rates, share of people with limited access to food, share of households that are housing-cost burdened, and share of residents in rural areas) were up-weighted to increase their influence.[6] The metric weights were adjusted based on the results of stability analyses that measured the influence of each metric in the cluster analysis. We examined empirical measures of group separation and cohesion (silhouette scores) in order to arrive at 10 peer groups.

Metric Definitions

Food Insecurity

- *Food insecure, all people:* share of people living in households that experienced food insecurity in the past 12 months. Food-insecure households are those that at times lacked access to enough food for an active, healthy life or had limited or uncertain availability of nutritionally adequate foods for all household members (2016 and 2017 Feeding America Map the Meal Gap data).[7]

- *Food insecure, children:* share of children living in households that experienced food insecurity in the past 12 months (2016 and 2017 Feeding America Map the Meal Gap data).

- *Limited access to healthy food:* share of low-income people (less than or equal to 200 percent of the federal poverty threshold) that do not live within 10 miles of a grocery store in rural areas or within 1 mile in nonrural areas (2015 USDA Food Environment Atlas data, downloaded via CHR).

- *Households receiving SNAP:* percent of households that received benefits from the Supplemental Nutrition Assistance Program (SNAP) in the past 12 months (2016 or 2012–16 ACS data).[8,9]

- *Households receiving SNAP or cash assistance:* percent of households that received SNAP or public cash assistance, which includes cash benefits from the Temporary Assistance for Needy Families (TANF) and General Assistance (GA) programs, in the past 12 months (2016 or 2012–16 ACS data).

Physical Health

- *Low-birthweight births:* share of live births with birthweight below 2.5 kilograms (5.5 pounds) (2010–16 CDC National Vital Statistics System data, downloaded via CHR).

- *People with diabetes:* share of adults age 20 and older with diagnosed diabetes (2014 CDC Diabetes Interactive Atlas data, downloaded via CHR).

- *People with disability:* share of people with at least one disability (2016 or 2012–16 ACS data).

- *No health insurance:* share of people enrolled in neither a private nor public health insurance plan (2016 or 2012–16 ACS data).[10]

- *Premature deaths per 100,000 people:* number of deaths per 100,000 among people under age 75 (2014–16 CDC Wide-ranging Online Data for Epidemiologic Research [WONDER] mortality data, downloaded via CHR).

Housing and Transportation Costs

- *Housing-cost burdened:* share of households (renters and owners) that spend 30 percent or more of their incomes on housing and utility costs (2016 or 2012–16 ACS data).

- *Severely housing-cost burdened:* share of households (renters and owners) that spend 50 percent or more of their incomes on housing and utility costs (2016 or 2012–16 ACS data).

- *Wage to afford fair market rent:* hourly wage a person working full time, full year needs to make so he or she spends no more than 30 percent of income on rent at HUD's prevailing fair market rent for a two-bedroom apartment (2018 NLIHC Out of Reach data).

- *Transportation costs as percent of income:* median transportation cost as percent of income for regional typical households (2017 CNT Housing and Transportation Affordability Index data).[11]

Income and Employment

- *Median household income:* median household income (2016 or 2012–16 ACS data).

- *Below 200% of federal poverty level:* share of people with incomes below 200 percent of the federal poverty threshold (2016 or 2012–16 ACS data).

- *Unemployment rate:* share of people in the labor force who are not employed (2016 BLS Local Area Unemployment Statistics annual-average estimates for county and state figures and 2016 BLS Current Population Survey annual-average estimates for the national figure).

- *Income inequality:* Gini coefficient multiplied by 100 (higher value indicates greater inequality) (2016 or 2012–16 ACS data). The Gini index is measured on a scale of 0 to 1, with higher numbers indicating more inequality. A Gini of 0 reflects perfect equality (i.e., all households have the same income), and a Gini of 1 reflects perfect inequality (i.e., one household has all the income).

Financial Health

- *Median credit score:* median VantageScore credit score ranging from 300 to 850 among those with a credit record. Subprime scores range from 300 to 600, near-prime from 601 to 660, and prime or more from 661 to 850 (2017 credit bureau data).[12]

- *Debt in collections:* share of people with a credit record who have debt in collections (i.e., severely delinquent debt) (2017 credit bureau data).

- *Medical debt in collections:* share of people with a credit record who have medical debt in collections (2017 credit bureau data).

- *Average unsecured debt*: average amount of unsecured debt (e.g., credit card debt, student loan debt, auto debt) among people with a credit record (2017 credit bureau data).

Demographics

- *Households with children:* share of households with any children younger than age 18 (2016 or 2012–16 ACS data).

- *Households with seniors (65+):* share of households with a member age 65 or older (2016 or 2012–16 ACS data).

- *People of color:* share of people who report being Hispanic or non-Hispanic and one of the following races: Black, Asian or Pacific Islander, American Indian or Alaska Native, another race, or multiracial (2016 or 2012–16 ACS data).

- *Black, non-Hispanic:* share of people who reported being non-Hispanic and black or African American alone (2016 or 2012–16 ACS data).

- *Hispanic:* share of people who reported being Hispanic and of any one or multiple races (2016 or 2012–16 ACS data).

- *Asian or Pacific Islander:* share of people who reported being non-Hispanic and Asian or Pacific Islander alone (2016 or 2012–16 ACS data).

- *Native American:* share of people who reported being non-Hispanic and American Indian or Alaska Native alone (2016 or 2012–16 ACS data).

- *Another race or multiracial, non-Hispanic:* share of people who reported being non-Hispanic and a race other than those listed above or belonging to more than one racial category (2016 or 2012–16 ACS data).

- *White, non-Hispanic:* share of people who reported being non-Hispanic and white alone (2016 or 2012–16 ACS data).

- *No college degree:* share of people ages 25 or older with no college degree (includes associate's degrees, bachelor's degrees, and graduate degrees) (2016 or 2012–16 ACS data).

- *Immigrants:* share of people who are foreign born (2016 or 2012–16 ACS data).

- *Total population:* total number of people residing in the area (2016 or 2012–16 ACS data).

Geography

- *Population in rural area:* share of the population living in rural areas (2010 decennial census data).

Notes

1. Other analyses have used cluster analysis to form county peer groups that classify counties in terms of their similarity or difference across alternative arrays of metrics. For example, the America at Work initiative groups counties based on their workforce and economic profiles: "Automation is Reshaping Work Across America. A New Report Explores the Impact—and How Communities Might Respond," Walmart, accessed May 1, 2019, https://corporate.walmart.com/2019/02/13/automation-is-reshaping-work-across-america-a-new-report-explores-the-impact-and-how-communities-might-respond.

2. The clustering metrics displayed on the dashboard are food insecure, all people (2016); food insecure, children (2016); low-birthweight births; people with diabetes; people with disability; no health insurance; housing-cost burdened; severely housing-cost burdened; wage to afford fair market rent; median household income; below 200% of federal poverty level; unemployment rate; median credit score; debt in collections; households with children; households with seniors (65+); people of color; no college degree; and population in a rural area. The clustering metrics not presented on the dashboard but available in the data download and county comparison file along with those on the dashboard are limited access to healthy food; households receiving SNAP or cash assistance; premature deaths per 100,000 people; transportation costs as percent of income; income inequality; medical debt in collections; average unsecured debt; and immigrants. Additional metrics not included in the cluster analysis or on the dashboard but available in the data download and county comparison file as contextual information along with the metrics listed above are: food insecure, all people (2017); food insecure, children (2017); households receiving SNAP; Black, non-Hispanic; Hispanic; Asian or Pacific Islander; Native American; another race or multiracial, non-Hispanic; white, non-Hispanic; and total population.

3. Individual, de-identified credit record data from a major credit bureau.

4. Downloaded via the National Historical Geographic Information System (NHGIS), IPUMS NHGIS, University of Minnesota, accessed March 1, 2019, www.nhgis.org.

5. The following variables were transformed to a more normal distribution in order to minimize the influence of outliers: limited access to healthy food, wage to afford fair market rent, unemployment rate, share of people of color, and share of immigrants.

6. The following variables were up-weighted by the number of times indicated to increase their influence in line with their analytic importance: food insecure, all people, 6; food insecure, children, 3; limited access to healthy food, 2; housing-cost burdened, 2; severely housing-cost burdened, 2; and population in rural area, 4.

7. Households that are food insecure reported three or more food-insecure conditions or behaviors in the Current Population Survey Food Security Supplement. County and state rates are estimated based on a model developed for Feeding America (see "How We Got the Map Data," Feeding America, accessed July 1, 2019, https://www.feedingamerica.org/research/map-the-meal-gap/how-we-got-the-map-data for additional information).

8. We use 2016 ACS data where possible, but for smaller counties with populations fewer than 65,000 for which one-year estimates are not available, we use the 2012–16 five-year ACS.

9. SNAP and cash assistance data has not been corrected for benefits underreporting, which makes the reported number and share of people in households that receive these benefits lower than they would be otherwise (see "Data & Methods: The Transfer Income Model TRIM," Urban Institute, accessed July 1, 2019, https://www.urban.org/research/data-methods/data-analysis/quantitative-data-analysis/microsimulation/transfer-income-model-trim for additional information).

10. This health insurance data has not been corrected for coverage underreporting, which makes the reported number and share of people who do not have health insurance higher than they would be otherwise (see Lynch, Victoria, Genevieve M. Kenney, Jennifer Haley, and Dean M. Resnick, *Improving the Validity of the Medicaid/CHIP Estimates on the American Community Survey: The Role of Logical Coverage Edits* (US Census Bureau, September 27, 2011), https://www.census.gov/content/dam/Census/library/working-papers/2011/demo/improving-the-validity-of-the-medicaid-chip-estimates-on-the-acs.pdf for additional information).

[11] CNT's 2017-release data estimates transportation costs at various geographic levels based on a model that uses 2015 and 2014 data from various sources. A regional typical household is defined as one that has the median income, average number of members, and average number of commuters among households in the region. The region is defined as the core-based statistical area (CBSA) for counties that are part of a CBSA and the individual county for those that are not part of a CBSA. State and national transportation costs as percent of income were computed by the Urban Institute based on guidance from Dr. Peter Haas of CNT and following the method CNT uses to calculate this metric for congressional districts. The state and national figures are calculated as an average of county-level data (weighted by number of households in each county) of transportation costs as a share of household income for the regional typical household (see "Methodology," CNT, accessed February 1, 2019, https://htaindex.cnt.org/about/#methodology for additional information).

[12] Individuals for whom a credit record exists but for whom the record contains insufficient information to produce a score are counted as having a credit score of zero when calculating the median.

Acknowledgments

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of our experts. More information on our funding principles is available here. Read our terms of service here.

For more information on this project, see "Disrupting Food Insecurity: Tapping Data for Strategies that Tackle the Root Causes."

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org