URBAN INSTITUTE · ELEVATE · THE · DEBATE

Reenvisioning Rural America: How to Invest in the Strengths and Potential of Rural Communities

Amanda Gold, Eric Burnstein, Corianne Payton Scally, and Yipeng Su

September 2021

Appendixes

Contents

Appendix A. Study Methodology

Rural communities vary in their assets, capacities, and needs. However, few tools (e.g., datasets, web tools, online features) allow users to measure and categorize the diversity of rural strengths. Existing tools tend to fall short of adequately describing rural communities in three ways. First, they adopt a needs-based framework, pointing to problem areas more than to strengths. This is particularly challenging for communities with low incomes and communities of color that continue to be systematically identified by what they lack, rather than by the rich resources they have, and that are

frequently overlooked for public and private investment. Second, existing tools typically provide data only on the county level, which obscures local diversity and fails to provide meaningful information for smaller communities. Third, these tools provide large amounts of data and offer little support to users seeking to determine the relative importance of individual measures or to identify patterns across communities.

To address these gaps, we constructed a novel typology that categorizes rural census tracts into peer groups based on their assets using 50 distinct measures across seven types of community capital, provides data at the tract level, and leverages similarities in assets and challenges across places to help users identify areas that are similar. In addition, we assembled a set of companion measures that describe the institutional context in which rural communities exist, including their state, regional, county, and local government contexts. Understanding various social and economic realities in rural communities can provide crucial information about their capacities to absorb public- and private-sector investments and where institution building may need to precede larger-scale investments. The typology and the measures we used can be accessed in a data tool that uses a map to highlight the diversity of rural America, helping to dispel the negative stereotypes that are sometimes attributed to these communities.

Policymakers and practitioners can leverage this typology and the online tool to come up with the best approaches for building on the strengths of rural people and places. The primary audience for this work is stakeholders at the local and regional level, such as policymakers working in local or county government and practitioners working at the local or regional level. Our secondary audience includes federal policymakers and investors, such as employees of federal agencies working on rural policies or programs. It may also include investors in rural communities, such as community development financial institutions, place-based foundations, and private investors. Finally, researchers may leverage this detailed methodology, the downloadable data, and the typology for future studies.

These appendixes describe how we used 50 measures of rural assets and cluster analysis to group rural census tracts into peer groups and summarize additional capacity measures in our online tool. In addition to a detailed study methodology, we include a discussion of opportunities for future research and a landscape of complementary online tools.

Study Approach

COMMUNITY CAPITALS FRAMEWORK

Our study uses the Community Capitals Framework (CCF) as a starting point for describing rural assets. Developed by Cornelia Butler Flora and Jan Flora (Flora, Flora, and Fey 2003), the framework identifies seven types of capital available for communities to leverage: built, cultural, financial, human, natural, political, and social. Building on a growing body of literature that evaluates various social and economic capitals and institutional and community capacity, the CCF has been widely used in rural community and economic development. It can highlight the strengths that a community can leverage and identify unused assets that may deteriorate if they go unleveraged for an extended period. Thus, the framework provides a lens to help people plan investments based on their priorities and the strengths and potential that exist within each community (Flora et al., n.d.).

The CCF is traditionally used as an on-the-ground asset-identification tool to help with local asset mapping, so trying to use it as a national framework has some challenges:

- Because the framework is focused on assets, some community challenges, such as concentrated poverty or segregation, might not be captured well.
- Nonmaterial assets—such as cultural, social, and political capital—are difficult to quantify and are not mutually exclusive.
- Some assets may fit under multiple capitals. For example, schools are both human capital (education) and social capital (they can be a key community gathering place for strengthening community ties), and the buildings themselves could be thought of as built capital.
- Something that is considered an asset from the perspective of one capital type can be considered negative from the perspective of another. For example, extractive resources (such as mining) could have a negative impact on types of natural capital (e.g., air quality) or aspects of human capital (e.g., health) but also contribute to a community's financial capital, at least for those who work in the industry or own the resource.
- The framework does not distinguish between stocks of capital (the quantity of something that exists at a point in time) and flows of capital (measured over an interval of time).
- Data limitations and availability make measuring the quality of assets a challenge.

To address these challenges, our study builds on the CCF's theoretical definitions to construct the following working definitions of the seven community capitals.

- **Built:** "Infrastructure" casts a wide net, encompassing housing, transportation and communication systems, and anything else constructed by people to support society. To summarize a tract's built capital, we analyzed commute times, affordable-housing supply, housing quality, cell service coverage, and access to broadband internet, fire stations, highways, transportation, and emergency response centers.
- **Cultural:** Culture encompasses the fabric of community life, including the people who live in a region, the languages they speak, and the institutions they support. We measured a tract's cultural capital through its diversity of languages, religions, races, and ethnicities and its cultural institutions, including historic properties, cultural organizations and occupations, and convention centers.
- **Financial:** A region's financial stability depends on the wealth of its people, the availability of financial resources, and the strength of its financial organizations. To measure financial capital, we considered income, home value, banking capacity and access, and federal and community development investments.
- **Human:** People need access to institutions that promote their health, support their education and training, and provide them with a livelihood. We measured a tract's human capital through that lens, focusing on employment opportunities and access to and the capacity of health care and educational systems.
- **Natural:** Communities are often considered in relation to their built environment, but the natural environment also shapes a region's identity, industry, and culture. To measure an area's natural capital, we considered its air quality; its protections for natural resources and land; its land coverage in crops, parks, and water; and its access to extractive resources, such as oil, coal, and natural gas.
- **Political:** Small or large, every community has governments that shape its priorities and future. We measure a region's political capital based on the region's influence over policy: through government employment, political competition, and social welfare organizations and the degree to which its population participates in elections and the census.
- **Social:** If cultural capital represents the fabric of community life, social capital represents the stitches holding the fabric together. Social capital encourages economic growth that benefits an entire community, so we measure it through the access a region has to organizations that encourage interconnectedness: places of worship, public libraries, schools, and civic and social-minded organizations.

We also worked with project partners at the Robert Wood Johnson Foundation, the US Department of Agriculture (USDA), and the Federal Reserve Board of Governors to identify the most important types of assets to consider for the typology based on its anticipated uses. We solicited and received feedback and guidance from members of our two advisory groups: a technical working group of rural researchers, who helped inform measures and methodology and review findings, and a stakeholder advisory group of rural practitioners, who informed measures, discussed audiences and tool uses, and gave feedback on findings and group descriptions.

COMPANION MEASURES

By reviewing literature on location-specific indicators of capacity, we determined that institutional context, such as government and governance arrangements, was a key driver of variation in local capacity to leverage the assets identified in the CCF. These governance arrangements often include nongovernmental actors that provide core services, such as coordinating funding for direct services (e.g., United Way chapters) or acting as an organizing body and collective advocacy voice for the local business community (e.g., chambers of commerce). Because of this, we decided to include a set of companion measures to provide the user with information on local government bodies and private nonprofit and business organizing institutions. We determined these measures should not be included in the typology development for several reasons:

1. Specific government or governance structures are not a feature of a location but are largely shaped or even determined by supralocal influences. In the US's complex federal structure, each community exists under a multilayered state government and potentially county and municipal governments, as well as special districts such as school districts, water boards, wastewater districts, and regional commissions. Each of these operates under its own set of rules and restrictions (frequently circumscribed by state law) that determine, for example, what resources are available to local communities and which government bodies can levy taxes, collect fees, and receive external investments.
2. Any given government or governance arrangement has potential benefits and drawbacks and cannot be identified as an asset. Although some research (e.g., Moore, Severn, and Millar 2006) has identified these as institutional capital, other research has revealed potential conflicts when institutions and legal structures within a place fail to consistently work together or even communicate. A large body of work on the costs of inter- and intragovernmental coordination, for example, may not take into account the costs of breaking down silos in a resource-constrained environment (Deslatte and Stokan 2020). Moreover, a municipal government can benefit residents by increasing the services they can access, but it can also add development

pressures that can result in environmental degradation from increased intensity of land use and damage to farmland on the rural-urban frontier (Wu 2008).

3. The complexity of government and governance systems means that accessing high-quality data on local governments and related institutions for all rural communities in the United States at usable geographic levels is difficult.

Nongovernmental organizations can play important roles in organizing community capacity to leverage rural assets. A local United Way, for example, may be able to coordinate the nonprofit services sector in much the same way that a municipal government may be able to regulate land use or invest in local transportation systems. Likewise, a local chamber of commerce may be able to coordinate resources for the business community while connecting it to a national network of professionals focused on economic development. By considering these government and nongovernment measures together, we can gain insight into the actors that may choose to use locally available assets, as well as some of the opportunities and constraints those actors may experience in doing so.

DEFINING "RURAL" AND UNIT OF ANALYSIS

Before we could build our typology, we had to define which parts of the country are rural. Our primary goal was to select a definition of rural that was clear and inclusive enough to capture the many realities of rural life. We selected the rural-urban commuting area (RUCA) codes,[1] which classify US census tracts using measures of population density, urbanization, and daily commuting.[2] Created by the USDA Economic Research Service and used by the Federal Office of Rural Health Policy within the US Department of Health and Human Services,[3] RUCA codes build on the US Census Bureau's definitions of "urbanized areas" and "urban clusters."[4] The RUCA codes provide nuance beyond county-based classification systems (e.g., the US Office of Management and Budget's metropolitan and nonmetropolitan areas and USDA's rural-urban continuum codes),[5] allowing users to categorize places based on whether residents work within their census tract or travel to employment elsewhere. Using census tracts also allows us to identify rural places within counties that might be classified as "urban" by other classification systems while still aligning our typology with numerous rural definitions that exist for federal and other programs.

The RUCA classifications have 10 basic codes across four general categories: metropolitan (codes 1, 2, and 3), micropolitan (codes 4, 5, and 6), small town (codes 7, 8, and 9), and isolated rural (code 10). We define all micropolitan, small town, and isolated rural areas as rural (codes 4 through 10).[6] For this study, we focused only on the 50 US states. Although many US territories have large rural populations, data availability and reliability for these areas are a significant challenge.[7] This definition of rural yielded

13,674 census tracts. From this group, we excluded tracts that had zero housing units (327) or zero population (299) using 2014–18 estimates from the US Census Bureau's American Community Survey. Most of the excluded tracts were covered by water or parkland. This left a final study universe of 13,048 rural census tracts.

Typology Construction

DATA SELECTION AND MEASURE IDENTIFICATION

USDA Rural Development interns in 2019 conducted a literature scan of datasets that could be used to measure characteristics of each of the seven community capitals. For this study, the Urban research team assessed and added datasets to fill in any gaps. In total, we reviewed 209 datasets and assessed each using the following four criteria:

1. The data must be available at no cost.
2. The data should be from 2010 or later (some exceptions were made).[8]
3. The data must be at least at the county level, with the census tract preferred.
4. The data should be as reliable and accurate for rural areas as possible.[9]

For each community capital, our goal was to identify no more than 10 datasets and to prioritize those that performed the best based on the four criteria. Ultimately, we selected a subset of 25 sources that were all at the census tract, zip code, or county level. The most common reasons for excluding a dataset were that it did not fit well within the CCF or was not asset-based, that we had another source that better measured the same concept, that the data were not at the county or census tract level, and that the dataset potentially had an urban bias or did not have good coverage for rural tracts.

After finalizing our data sources, we identified 50 indicators across all seven community capitals. The process of finalizing this list was iterative. Indicators were selected based on measures that have been used by researchers elsewhere (e.g., median home value or labor force participation rate) and recommendations compiled from our two advisory groups. Using analytic methods described later, we prioritized indicators that were continuous, because binary or categorical data are not useful for cluster analysis. Our final list of measures is in table 1, which includes details about their definitions and their sources. Although these measures are framed as assets in this study, some may have both positive and negative impacts on a community.

TABLE 1

Summary of Typology Measures and Units, by Community Capital

Capital type	Measure	Source	Indicator	Year	Level of geography
Built	Access to broadband internet	Census Bureau	Share of households with a broadband subscription of any type or a cellular data plan of any type	2018	Tract
Built	Access to domestic emergency operations centers[a]	Homeland Infrastructure Foundation-Level Data	For tracts with one or more emergency operations centers within their boundaries: average distance from tract emergency operations centers to tract population-weighted centroid. For tracts without emergency operations centers within their boundaries: distance from nearest emergency operation center to tract population-weighted centroid.	2009	Tract
Built	Access to fire stations or emergency medical services (EMS)	Homeland Infrastructure Foundation-Level Data	For tracts with one or more fire or EMS centers within their boundaries: average distance from tract fire/EMS centers to tract population-weighted centroid. For tracts without fire/EMS centers within their boundaries: distance from nearest fire/EMS center to tract population-weighted centroid.	2010	Tract
Built	Access to highways	ESRI maps and data	For tracts with one or more highway exits within their boundaries: average distance from highway exits to tract population-weighted centroid. For tracts without highway exits within their boundaries: distance from nearest highway exit to tract population-weighted centroid.	2005	Tract
Built	Access to transportation stations (train, air, bus, ferry, light rail, and bike-share)	Homeland Infrastructure Foundation-Level Data	For tracts with one or more transportation stations within their boundaries: average distance from all transportation stations to tract population-weighted centroid. For tracts without transportation stations within their boundaries: distance from nearest transportation stations to tract population-weighted centroid.	2018	Tract

Capital type	Measure	Source	Indicator	Year	Level of geography
Built	Affordable-housing supply	National Low Income Housing Coalition	Housing units with active federal subsidies as a share of the total rental units in a tract. Includes the following programs: HOME rental assistance, US Department of Housing and Urban Development (HUD) project-based rental assistance, Section 202 direct loans program, HUD insurance programs, state housing finance agency–funded Section 236, Low-Income Housing Tax Credit program, Section 515 rural rental housing loans, Section 538 US Department of Agriculture Guaranteed Rural Rental Housing Program, and public housing.	2020	Tract
Built	Cell service coverage	Homeland Infrastructure Foundation-Level Data	Share of tract covered by cell service providers	2018	Tract
Built	Housing quality	Census Bureau	Share of occupied housing units with complete kitchen and plumbing facilities	2018	Tract
Built	Workers with a short commute	Census Bureau	Share of workers who are 16 years or older, do not work at home, and travel less than 30 minutes to their jobs	2018	Tract
Cultural	Access to convention centers	Homeland Infrastructure Foundation-Level Data	For tracts with one or more convention centers within their boundaries: average distance from convention centers to tract population-weighted centroid. For tracts without convention centers within their boundaries: distance from nearest convention center to tract population-weighted centroid.	2018–19	Tract
Cultural	Capacity of cultural organizations	National Center for Charitable Statistics Core Files; Census Bureau	Average annual revenue per 1,000 people of all organizations categorized under the National Taxonomy of Exempt Entities groups A (arts, culture, and humanities) and X (religion related, spiritual development) in the county where the tract is located. The National Taxonomy of Exempt Entities is a classification system for	2017[b]	County

Capital type	Measure	Source	Indicator	Year	Level of geography
			nonprofit organizations developed by the National Center for Charitable Statistics and used by the Internal Revenue Service.		
Cultural	Employment in cultural occupations	Census Bureau	Share of the county civilian population aged 16 or older employed in the arts, entertainment, and recreation industry	2018	County
Cultural	Language diversity	Census Bureau	Simpson's Diversity Index for languages, using five spoken language categories: (1) English only; (2) Spanish; (3) French, Haitian, or Cajun; (4) German or other West Germanic languages; and (5) all remaining languages from the American Community Survey.[c]	2018	Tract
Cultural	Presence of historic properties	National Park Service; Census Bureau	Number of historic properties in a tract per 1,000 people, as identified by the National Register of Historic Places, which was authorized by the National Historic Preservation Act of 1966 and is maintained by the National Park Service	2017, National Park Service; 2018, population, American Community Survey	Tract
Cultural	Racial and ethnic diversity	Census Bureau	Simpson's Diversity Index for race and ethnicity, using the following groups: (1) white alone, not Hispanic or Latino; (2) Black or African American alone, not Hispanic or Latino; (3) American Indian and Alaska Native alone, not Hispanic or Latino; (4) Asian alone, not Hispanic or Latino; (5) Native Hawaiian and other Pacific Islander alone, not Hispanic or Latino; (6) some other race alone, not Hispanic or Latino; (7) two or more races, not Hispanic or Latino; (8) Hispanic or Latino.	2018	Tract
Cultural	Religious diversity	US Religion Census: Religious Congregations and Membership Study (county file),	Simpson's Diversity Index for adherents of religion using all 152 unique groups counted by the Census Bureau individually[d]	2010	County

Capital type	Measure	Source	Indicator	Year	Level of geography
		Association of Religion Data Archives			
Financial	Access to banks and credit unions	Federal Deposit Insurance Corporation; National Credit Union Administration	For tracts with one or more full service[e] banking locations within their boundaries: average distance from banking locations to tract population-weighted centroid. For tracts without banking locations within their boundaries: distance from nearest banking location to tract population-weighted centroid.	2020	Tract
Financial	Banking capacity	Community Reinvestment Act; Federal Deposit Insurance Corporation (FDIC); Census Bureau	Inflation-adjusted total amount of bank deposits among all FDIC-insured bank branch offices at the time of data collection per 1,000 people	2020	Tract
Financial	Community development financial institution investments	Community Development Financial Institutions Fund; Census Bureau	Inflation-adjusted total annual amount of community development financial institution investments per 1,000 people	2017	Tract
Financial	Federal investment	USAspending.gov; Census Bureau	The sum of all federal block grants, formula grants, project grants, cooperative agreements, direct loans, and guaranteed/insured loans invested in the county where a tract is located per 1,000 people (based on place of performance). For investments to be included, either "federal action obligation" or "face value of loan" must be positive; the other variable could be missing, but not negative. Investments also must have an action date between federal fiscal years 2018 and 2019 and have a primary place of performance at the county or zip code level. Investments across all federal agencies and programs were included.	Federal fiscal years 2018–19	County
Financial	Home value	Census Bureau	Median home value of owner-occupied units	2018	Tract

Capital type	Measure	Source	Indicator	Year	Level of geography
Financial	Individual income	Census Bureau	Inflation-adjusted median individual income in the previous 12 months	2018	Tract
Human	Access to child care centers	Homeland Infrastructure Foundation-Level Data; American Community Survey	Number of child care centers per 1,000 children younger than 5	2018–19, Homeland Infrastructure Foundation-Level Data; 2018, population, American Community Survey	Tract
Human	Access to college and university campuses	Homeland Infrastructure Foundation-Level Data	For tracts with one or more colleges/universities within their boundaries: average distance from colleges/universities to tract population-weighted centroid. For tracts without colleges/universities within their boundaries: distance from nearest college/university to tract population-weighted centroid.	2017–19	Tract
Human	Access to farmers' markets	US Department of Agriculture	For tracts with one or more farmers' markets within their boundaries: average distance from farmers' markets to tract population-weighted centroid. For tracts without farmers' markets within their boundaries: distance from nearest farmers' market to tract population-weighted centroid.	2020	Tract
Human	Access to health care facilities	Health Resources and Services Administration; Homeland Infrastructure Foundation-Level Data	For tracts with one or more federally qualified health centers, rural health clinics, and hospitals within their boundaries: average distance from facilities to tract population-weighted centroid. For tracts without federally qualified health centers, rural health clinics, and hospitals within their boundaries: distance from nearest facility to tract population-weighted centroid.	2019	Tract
Human	Access to health care professionals	Health Resources and	Number of health care professionals (primary care, mental health, and dental care)	2019	County

Capital type	Measure	Source	Indicator	Year	Level of geography
		Services Administration	per 1,000 people in the county where the tract is located		
Human	Educational attainment	Census Bureau	Share of people 25 years or older with at least a high school degree or its equivalent	2018	Tract
Human	Employment to population ratio	Census Bureau	Employment-to-tract-population ratio, calculated by dividing the number of people employed by the total number of people 16 or older	2018	Tract
Human	Health insurance coverage	Census Bureau	Share of total civilian noninstitutionalized population with health insurance coverage	2017	Tract
Human	Hospital capacity	Homeland Infrastructure Foundation-Level Data; American Community Survey	General acute care and critical access hospital beds per 1,000 people	2012–19, Homeland Infrastructure Foundation-Level Data; 2018, population, American Community Survey	Tract
Human	Labor force participation	Census Bureau	Share of 18- to 64-year-olds in the labor force, calculated as the sum of all workers ages 18 to 64 who are employed or actively seeking employment divided by the total noninstitutionalized, civilian population of 18- to 64-year-olds	2018	Tract
Human	Life expectancy	National Association for Public Health Statistics and Information Systems; County Health Rankings and Roadmaps	Estimated life expectancy for total population of tract. If data are missing, the measure is backfilled with county-level data.	2010–15, US Small-Area Life Expectancy Estimates Project; 2016–18, County Health Rankings and Roadmaps	Tract; county
Human	Nursing home capacity	Homeland Infrastructure Foundation-Level Data; American Community Survey	Nursing home beds per 1,000 people 65 or older	2019, nursing home beds; 2018, population	Tract

Capital type	Measure	Source	Indicator	Year	Level of geography
Natural	Access to coal mines	US Labor Department, Mine Safety and Health Administration Form 7000-2: "Quarterly Mine Employment and Coal Production Report"	For tracts with one or more operating coal mines within their boundaries: average distance from coal mines to tract population-weighted centroid. For tracts without coal mines within their boundaries: distance from nearest coal mine to tract population-weighted centroid.	2018	Tract
Natural	Access to oil and natural gas	Homeland Infrastructure Foundation-Level Data and Oak Ridge National Laboratory	For tracts with one or more oil and gas wells within their boundaries: average distance from oil and gas wells to tract population-weighted centroid. For tracts without oil and gas wells within their boundaries: distance from nearest oil and gas well to tract population-weighted centroid.	2014–17	Tract
Natural	Air quality	Environmental Protection Agency	Average particulate matter (PM2.5) for the county	2020	County
Natural	Cropland coverage	National Land Cover Database, US Geological Survey	Share of tract covered by cropland (areas used to produce annual crops such as vegetables, tobacco, and cotton and perennial woody crops such as orchards and vineyards). This also includes all land actively tilled.	2011	Tract
Natural	Lake or river coverage	Homeland Infrastructure Foundation-Level Data; US Army Corps of Engineers Inland Electronic Navigational Chart	Share of tract covered by lakes and rivers	1995–2020	Tract
Natural	Land and natural resource protections	National Conservation Easement Database; Natural Resources	Share of tract covered by conservation easements (land that is owned by private land trusts or public entities and is protected to benefit agriculture and/or the environment)	2020	Tract

Capital type	Measure	Source	Indicator	Year	Level of geography
		Conservation Service			
Natural	Park coverage	ArcGIS/ESRI	Share of tract covered by parks	2010–16	Tract
Political	Census participation	Census Bureau	Census response rate, calculated as the share of questionnaires mailed back by households that received them. This excludes households whose forms were returned by the US Postal Service as "undeliverable."	2010	Tract
Political	Financial capacity of social welfare organizations	National Center for Charitable Statistics; Census Bureau	Inflation-adjusted average annual revenue per 1,000 people of 501(c)(4) organizations within the county where a tract is located	2017[b]	County
Political	Government employment	Census Bureau	Share of the county civilian population 16 or older who is employed in public administration	2018	County
Political	Political competition	Massachusetts Institute of Technology Election Data and Science Lab	Absolute value of the percentage point difference between the share of county votes for the Republican candidate and the share for the Democratic candidate in the 2016 presidential election. Lower values (e.g., 1) mean more competition and higher values (e.g., 50) mean less competition.	2016	County
Political	Voter participation	Massachusetts Institute of Technology Election Data and Science Lab	Participation in the 2016 presidential election in the county where a tract is located, calculated as the number of people 18 or older who voted divided by all people 18 or older	2016	County
Social	Access to civic and social organizations	Northeast Regional Center for Rural Development, using data from the Census Bureau	Sum of all civic and social organizations within the county per 1,000 people, including civic (North American Industry Classification System 813410), business (NAICS 813910), political (NAICS 813940), professional (NAICS 813920), labor (NAICS 813930), bowling (NAICS 713950), recreational (NAICS 713940), golf (NAICS 713910), and sports (NAICS 711211) organizations	2014	County

Capital type	Measure	Source	Indicator	Year	Level of geography
Social	Access to places of worship	Homeland Infrastructure Foundation-Level Data; American Community Survey	Number of 501(c)(3) organizations that are places of worship—including churches, temples, mosques, and synagogues—per 1,000 people	No year specified. Data were downloaded on February 21, 2020.	Tract
Social	Access to public libraries	Institute of Museum and Library Services	For tracts with one or more public libraries within their boundaries: average distance from public libraries to tract population-weighted centroid. For tracts without public libraries within their boundaries: distance from nearest public library to tract population-weighted centroid.	2018[f]	Tract
Social	Access to schools	Homeland Infrastructure Foundation-Level Data	For tracts with one or more schools within their boundaries: average distance from schools to tract population-weighted centroid. For tracts without schools within their boundaries: distance from nearest school to tract population-weighted centroid.	2009–18	Tract

[a] Emergency operations centers coordinate national, regional, state, county, and tribal emergency responses. These facilities are designed to respond quickly to a major incident, such as a weather disaster, and help coordinate efforts across multiple organizations, agencies, and departments. Centers may be operated by the county or tribal government. Some centers may operate on an ongoing basis or be activated in response to a specific emergency.

[b] The 2017 Core file contains the most recent filing available for a given organization within two years of the file date. Thus, each Core file contains multiple fiscal years' worth of information but only one record for each Employer Identification Number. More information is available at https://nccs.urban.org/sites/default/files/2018-12/Guide%20to%20Using%20NCCS%20Data.pdf.

[c] This measure could more appropriately be called an integration index, but we use the common name of the measure.

[d] For a full list of the religious groups, see "Religious Groups," Association of Religion Data Archives, accessed August 13, 2021, https://www.thearda.com/denoms/families/groups.asp.

[e] This measure includes only full service bank locations. It does not include locations for ATMs unless they are part of a full service bank.

[f] The 2018 public library file includes the most recent data for states according to their fiscal years, although the reporting period for each state varies. Most state fiscal years are a calendar year or run from July to June. In some states, the fiscal year reporting period varies among local jurisdictions. More information is available at https://www.imls.gov/sites/default/files/2018_pls_data_file_documentation.pdf.

VARIABLE CONSTRUCTION

Many of our sources already provided tract-level data (e.g., the American Community Survey from the US Census Bureau), and variable construction for these measures was relatively straightforward. However, building other measures, including those for which the data were at the county or zip code

level, was more complicated. For county-level measures, we assigned the county value to all the rural tracts within a county's boundary. Our measure of federal investment uses a mix of county- and zip-code-level data from USAspending.gov.[10] To assign a value to each rural census tract, we used a zip code-to-county crosswalk to apportion zip code data to counties (using population weights) and then assigned the combined county-level data to all rural tracts within each county (box 1).

BOX 1

Analyzing Data from USAspending.gov

The US Treasury Department shares data on all federal spending on USAspending.gov. We used this source to estimate the amount of federal investment flowing into rural census tracts. For this study, we define federal investment as the sum of all federal block grants, formula grants, project grants, cooperative agreements, direct loans, and guaranteed/insured loans invested in a tract per 1,000 people (based on place of performance). For investments to be included, either "federal action obligation" or "face value of loan" had to be positive; the other variable could be missing, but not negative. Investments also had to have had an action date between federal fiscal years 2018 and 2019 and a primary place of performance at the county or zip code level (i.e., investments with a place of performance at the state, multicounty, or multiple zip code level were excluded, as were investments made to a specific city). Investments across all federal agencies and programs were included.

As a newer data source that collects information across agencies and programs, USAspending.gov has faced data quality challenges. The US Government Accountability Office documented these in several recent reports (GAO 2017, 2018, and 2019). The 2019 report noted that even though data quality has improved, issues persist. The most concerning one for our study is related to the measure "primary place of performance," which we use to assign spending to tracts. The agency estimated that this measure is accurate at the county level for only 86 to 93 percent of transactions. Accuracy for zip code–level data is slightly worse (76 to 86 percent of transactions).

No other public data source captures the full scope of federal investment in US communities. We determined that including this measure despite its limitations was preferable to excluding it. We hope Congress will continue to put resources into improving USAspending.gov so it can be used with greater precision in future studies and can be leveraged to improve transparency and public trust in government.

Our study also uses spatial data, which had to be converted to tract-level measures of distance and coverage. We constructed distance-based measures by calculating the Euclidean distance from the population-weighted centroid to the nearest asset (all measured by point data). For tracts with one or more assets within their boundaries, we took the average distance from all assets within the tract to the

population-weighted centroid. For tracts without any assets within their boundaries, we calculated the distance to the nearest asset outside the tract from the population-weighted centroid. Because continuous measures are better for cluster analysis, we do not cut off any of our distance measures (i.e., replace them with a 0 if the asset is more than a certain distance away). Instead, we include the distance to the nearest asset even if it is in a different tract, county, or state. We also used polygon data to calculate coverage measures (i.e., the share of a tract's land area that is covered by an asset). In some instances, an area might be covered by more than one asset of the same type (e.g., a community that has access to cell service from more than one provider). When this happened, we did not double-count areas of overlap.

Finally, we used Simpson's Diversity Index to construct three measures of diversity: racial and ethnic, linguistic, and religious. The index values range from 0 to 1, with higher values meaning greater diversity. The resulting score measures both the richness (i.e., the number) and evenness (i.e., the distribution) of the groups. For example, given two communities with three groups, the Simpson's Diversity Index score would tell us that community A (80 percent, 10 percent, 10 percent) is less diverse than community B (33 percent, 33 percent, 33 percent). For each tract, the score is calculated using the formula below, where n_i is the population of group i, and N is the total tract population or $N=\sum n_i$.

$$\text{Diversity Index Score} = \sum_{i=1}^{k} \frac{n_i(n_i-1)}{N(N-1)}$$

One limitation of this measure is that it defines any tract that is heavily skewed toward a single group as having low diversity, regardless of the group. For example, a community that is 100 percent white non-Hispanic would receive the same low diversity score as a group that is 100 percent Black non-Hispanic (a score of 0 diversity).

DATA CLEANING

To prepare the data for cluster analysis, we had to ensure that all 13,048 rural tracts had values across all 50 measures. This meant we had to impute (estimate) values when data were missing. Overall, we tried to take the simplest approach possible to facilitate reliable updating of the tool in the future (the trade-off being that sometimes a more complex solution might yield more accurate data). For most measures, the number of missing values was low (less than 1 percent of the data for each measure). In these instances, we assigned tracts with missing data the average for their county. In a small number of cases when these county-level data had missing values, we assigned those counties the average for their state.

We experienced additional challenges with missing data for some states. All tracts in Hawaii and Alaska were missing data on air quality and cropland coverage, so we used a national average instead. In

Alaska, measures that relied on voting data for national elections (voter participation and political competition) had to be imputed based on the state average because Alaska reports voting data by borough, whose boundaries do not align with those of tracts or counties.[11] The data behind our tract-level measure for life expectancy, which comes from the US Small-Area Life Expectancy Estimates Project, had missing values for all the rural tracts in Wisconsin and Maine and 550 other tracts across the country. We supplemented this data source with county-level estimates from the County Health Rankings and Roadmaps website, which is published by the University of Wisconsin Population Health Institute. For a small number of tracts (26), both sources were missing data. In those cases, we used state averages.

After correcting for missing data, we put all data on the same scale. This was important because the cluster analysis method we used (k-means) is sensitive to the scale of the measures used (i.e., measures without the same units or magnitude). Our measures have various units (e.g., percentages, dollars, and distances in meters) and a range of magnitudes, both within and across units. A measure with a larger magnitude often has a larger variance, and without standardizing our measures, our analysis would have been heavily weighted toward higher-variance measures vis-à-vis lower-variance measures. Thus, aligning the scales of our measures was an important data-cleaning step before running the cluster analysis. This step allowed us to eliminate the influence of scale in our results and more easily make comparisons across measures.

We used the min-max normalization method of scaling. This approach transforms data so they are on a 0 to 1 scale, where the lowest (min) value after the transformation is set to 0.0 and the highest (max) value is set to 1.0. Every other value is transformed into a decimal between 0.0 and 1.0 using the following formula:

$$\frac{\text{value} - \min \text{value}}{\max \text{value} - \min \text{value}}$$

After applying this method to our data, the variances for all measures fell between 0 and 1, but variances across measures still varied widely. For example, the variance for cropland coverage (0.07) is higher than the variance for access to public libraries (0.0005). When we conducted our cluster analysis, these higher-variance measures factored more strongly into our results than did measures with lower variances. This approach produced cluster results that were mostly driven by the capital types with higher-variance measures (e.g., natural capital, cultural capital, and political capital), as discussed later.

ANALYSIS

Because of the large number of measures selected, we used principal component analysis (PCA) before the cluster analysis. PCA is a common variable-reduction technique that creates a linear combination of

a set of variables to produce one or more index variables called components. This approach served two purposes. First, it eliminated the issue of correlation among measures. The second benefit of PCA was that it made our results easier to interpret. That is, understanding how groups differed from one another was simpler when comparing them across a smaller set of synthetic measures, rather than across all 50.

We ran PCA on all 50 measures and selected the first seven components to use in the cluster analysis. Together, they describe about 80 percent of the variation in our data. We settled on seven components through a process of weighing the marginal value of adding a component against the diminishing amount of variance described by each component. The value of adding a seventh component, when we had six, was significant, but the value of adding an eighth dropped considerably, suggesting that the return on adding components beyond the first seven was diminishing. Appendix B details how each of the original 50 measures load onto the first seven principal components used in our analysis.[12]

Using the seven components generated by PCA, we used cluster analysis to construct groups of rural census tracts based on their similarities. We selected the k-means clustering method because it is one of the most popular approaches to clustering, is easy to interpret and understand, and is relatively computationally efficient. K-means clustering works by minimizing the distance within groups (i.e., minimizing within cluster sum of squares). It does this by dropping a predefined number of random points, assigning each of the tracts to the nearest point, and then moving the point to the center of that cluster and recalculating everything again. It does this a predefined number of times until it finds the result that minimizes all the distances between points within a cluster.[13]

Our analysis yielded seven types of rural census tracts. We used the "elbow method" to help determine the optimal number of groups. We did this by graphing the sum of squares for each number of clusters and looking for the number of groups where the marginal benefit of adding one group changed the angle of the line from one that was steep to one that appeared to be flatter (i.e., forming an elbow). The result of this test suggested that six or seven groups would be best. We balanced this information against the relative group sizes that emerged from a six-group cluster and a seven-group cluster (avoiding groups with an extremely small or large number of tracts) and determined that seven groups would be optimal.

RESULTS

We worked with our research partners to develop names and descriptions for the seven types of rural census tracts. In describing the groups, we relied on the original, unscaled measures (i.e., not the

principal components) and highlighted measures where a group's average was best or worst compared with the other groups. We named each group by identifying themes of strengths and weaknesses that cut across multiple measures or capital types. We also shared the group names and descriptions with members of our stakeholder advisory group and further refined the language and analysis based on their feedback. Key adjustments included:

- We moved a measure that gauged how much of a tract was tribal land from the typology to the companion measures. The high variance of this measure (mostly bimodal) caused tribal tracts to cluster together into a single group. Advisers suggested this was not useful for tribal communities searching for rural peers. Moreover, when we removed this measure, tribal tracts were distributed more evenly across groups (suggesting their underlying asset structure across the remaining 50 measures was more similar to nontribal rural areas than to one another).
- We refined group descriptions to highlight challenges in addition to strengths. We made this change after an advisory group member noted that challenges (represented by lower relative levels of assets in our analysis) were still important for describing group realities and providing information on how policymakers, practitioners, and investors can work to fill gaps in assets.

The seven groups are Accessible, Energy-Rich Hubs; High-Employment Agricultural Areas; Centers of Wealth and Health; Diverse, Institution-Rich Hubs; Remote, Energy-Rich Tracts; Diverse, Outlying Tracts; and Remote Recreational and Cultural Areas. The descriptions of each group and where the tracts in each group are located can be viewed in the data tool, at https://reenvisioning-rural-america.urban.org/. The full table of group averages across all 50 measures can be accessed via the data download on the tool's page.

ROBUSTNESS CHECKS AND DATA VALIDATION

We conducted robustness checks on our analysis before finalizing our model. They included the following tests:

- **Model.** We tested hierarchical clustering and k-means clustering without PCA and got similar results. We did not make any changes.
- **Scaling.** We tested normalizing our measures using the standard scaling approach (also known as a z-score), setting all variances to 1.[14] Land use coverage (e.g., park coverage, cropland coverage) became a less important factor driving how tracts clustered, distance measures gained more prominence, and the resulting clusters were uneven in size (the smallest group was 19 tracts, while the largest was nearly 4,000). We did not make any changes.

- **Outliers.** We replaced any observations that were above the 99th percentile with the value that was at the 99th percentile and any observation that was below the 1st percentile with the value that was at the 1st percentile (also known as top and bottom coding). We did not observe significant differences. We did not make any changes.
- **Measures.** We originally included a different measure of coal resources, the share of a tract covered with coal reserves. However, this measure was mostly bimodal (with tracts having either 0 or 100 percent coverage), and it was the most influential measure driving the clustering. We also discovered a conflict between this measure and our measure for cropland coverage. Some tracts that had high levels of coal reserves also had high levels of cropland coverage, suggesting that even though these areas may have coal reserves, the land was already being used for farming. Because our goal was to capture the predominant land use contributing to the economic activity in a tract, we replaced the original measure of coal reserves with one that measured the distance to the nearest active coal mine. This new measure was also more similar conceptually to our measure related to oil and gas wells (also a distance measure).
- **Imputing.** Our measures for life expectancy, political competition, and voter participation required the most imputing. To test our methods for estimating missing data, we reran our cluster analysis without each of these measures, one at a time. When we removed life expectancy, we saw no major changes in the number of tracts in each group or their key characteristics. When we dropped political competition and voter participation, there were some subtle changes in how measures loaded onto the first seven principal components, but they were not significant. We did not make any changes.

Companion Measures

The companion measures we included in our tool are government and governance features of a locality that may be important for tool users to understand but cannot be clearly labeled as assets and therefore were not included in the typology construction. We identified potential companion measures and data sources inductively by speaking with our advisers and stakeholder advisory committee and compiled a list of 36 measures. They include the type and number of local government bodies, whether an area has previously received federal funding sources, and whether an area has local organizing bodies for the nonprofit/direct service sector and the local business sector. We also spoke to experts at several national governance coordinating nonprofits—including the National Association of Counties, the

National Association of Development Organizations, Indiana University, and the United Way—about structural features they believed would be relevant.

Through this process, we confirmed that national datasets for many of the measures we identified had gaps because they were collected through particular networks (such as the National Association of Counties or the United Way) or through surveys with uneven coverage across rural places. As such, we used a different set of criteria from those we used to collect data for the measures used in the typology, to ensure the data we included would be of value:

1. **Data accessibility.** We did not restrict our search to publicly available datasets but were limited by data that were accessible within the project's budget and time frame.
2. **Reliability.** Although some datasets we used do not include data for all 50 states or may not include some institutions that provide a particular service, we prioritized data from reliable sources.
3. **Clarity of purpose and relevance to potential users.** We assessed potential measures based on their ability to communicate an institutional factor without a need for further information and whether our partners and stakeholder advisory group found the measures relevant to their interests.

In total, we identified 12 companion measures to include in the online tool. The final list is in table 2.

TABLE 2
Companion Measures

Measure	Source	Indicator	Year
Population	2014–18 American Community Survey five-year estimates	Total population of census tract	2018
Land area	2014–18 American Community Survey five-year estimates	Square miles in tract	2018
Number of municipalities	Census of Governments	Number of incorporated municipal governments that have boundaries that fall within a tract	2017
Degree of county independence from state	National Association of Counties' adaptation of Census of Governments data	The type of authorizing legislation in effect for the county in which a tract falls. Home-rule counties are allowed to take any governmental actions not specifically prohibited by state government. Dillon's Rule counties are allowed only those actions specifically allowed by state government. Hutchinson's rule is a complex blend that applies exclusively in Utah.	2017

Measure	Source	Indicator	Year
Prior experience with federal investment	USAspending.gov	Whether the county where the tract is located has previously received federal investment	Federal fiscal years 2018–19
Elected executive	National Association of Counties' adaptation of Census of Governments data	Whether the county where the tract is located has an elected executive	2017
Regional government network	University of Pittsburgh Regional Intergovernmental Organizations Project	Whether the county where the tract is located is served by a regional intergovernmental organization, which is made up of the local governments within a region and supported by active engagement of that region's civic sector	2018
Regional commission service area	Appalachian Regional Commission; Delta Regional Authority; Northern Border Regional Commission	Whether the county where a tract is located qualifies for funds from the Appalachian Regional Commission (ARC), the Delta Regional Authority (DRA), or the Northern Border Regional Commission (NBRC). All three are federal agencies devoted to funding economic development. The ARC serves 420 counties in the Appalachian region. The DRA serves 252 counties in the greater Mississippi Delta region. The NBRC serves counties in states in the Northeast that border Canada.	2021
Tribal lands	Census Bureau	Whether the county where the tract is located overlaps with American Indian/Alaska Native/Native Hawaiian areas, including American Indian reservations and trust lands, tribal jurisdiction statistical areas, Alaska Native Regional Corporations, Alaska Native village statistical areas, and tribal designated statistical areas	2017
United Way	United Way	Whether the county where the tract is located falls within the service area of a United Way	2021
Chamber of Commerce	US Chamber of Commerce	Whether the county where the tract is located falls within the service area of one or more chambers of commerce	2021
Number of chambers of commerce	US Chamber of Commerce	The number of chambers of commerce serving the county where the tract is located	2021

Study Limitations

TYPOLOGY

In addition to the missing data challenges discussed earlier, our study has limitations related to how we operationalized the CCF and the availability of public data for rural census tracts. Broadly, our focus on assets resulted in the exclusion of many common, deficit-based measures that shape rural health and

well-being. For example, we did not include commonly used metrics for poverty (poverty rates), economic insecurity (unemployment), health problems (low infant birth weight, toxic exposures, traumas), safety (crime rates), or housing insecurity (overcrowding, cost-burdened households). However, other typologies do; among them are the University of Michigan's Understanding Communities of Deep Disadvantage Index and the Economic Innovation Group's Distressed Communities Index (although the latter is presented by zip code and the map does not distinguish between urban and rural areas). Those and other, similar projects are listed in appendix C.

Our decision to use the RUCA codes meant that we prioritized tract-level data sources over more commonly used measures that are available only at the county level. However, finding tract-level data with national coverage for rural places was a challenge, and it was not possible for every measure. For some crucial measures, we were forced to use county data and in rare cases national or state averages.

Our commitment to using publicly available data for the typology, to allow for replicability of this project, also limited our potential data sources. Proprietary data might have yielded potentially useful measures. However, many private data sources do not have good coverage of rural communities (many measures may not exist publicly or privately for rural places).[15] In particular, we could not capture many important measures of natural capital, such as the potential for carbon capture, climate resilience, and renewable energy resources like solar and wind. Similarly, many sources provide data for only certain regions, states, or localities.

Some types of the measures we constructed had additional limitations. All our distance measures calculated the Euclidean distance (i.e., a straight line between two points) from the population-weighted centroid to the asset. These measures did not take into account differences in topography, roads, and transportation. For very mountainous communities (e.g., Appalachia or the Rocky Mountains), we may have overstated residents' access to some physical assets by not accounting for actual travel times. Finally, our measure for cropland coverage only considers the production of annual crops such as corn, soybeans, vegetables, tobacco, and cotton and perennial woody crops such as orchards and vineyards. This measure does not include other types of agricultural activities, particularly livestock ranching, which is a dominant industry in some rural places. Unfortunately, we could not find data to capture this asset.

Data limitations also led us to focus on measures of access and capacity rather than quality. For example, our measure for broadband connectivity is a measure of whether households have access via a subscription or a cellular data plan, rather than the speed of the connection, even though we know from other studies and anecdotally that slow connection speeds continue to be a major challenge in rural

places.[16] Similarly, we have no standardized measure of the quality of many of the services we measure tract-level access to, including banks, child care centers, and emergency services. Some of these data exist but cannot be disaggregated for a tract-level analysis or are available only at the local or state level and cannot be used for national studies. For example, Child Care Aware of America publishes data on the quality of child care, preschool, and school-age care programs, but the methodology varies by state and thus the data cannot be used for comparisons across states.[17] And J.D. Power publishes the results of its US Retail Banking Satisfaction Study on its website, but the public data are not detailed enough to allow for a tract- or county-level analysis.[18] One notable exception is for health services—the Centers for Medicare and Medicaid Services publish data on hospital quality via their Hospital Compare tool.[19]

COMPANION MEASURES

The companion measures had two key limitations: the absence of national data on many nongovernmental institutions and the reliability of available national data on local governments.

We wanted to capture in our measures whether an area had nongovernmental bodies with key roles in "governance arrangements" (convening, coordinating, directing, and centralized funding services for private for-profit and nonprofit institutions that serve a community).[20] Such organizations are highly varied, and although some may belong to a national association, others are independent or unique to a community. Some researchers have tried to compile national lists of civil society organizations, but our searches uncovered none with sufficient documentation to allow us to determine the completeness or accuracy of the data or whether the included organizations play the crucial governance roles we sought to capture. Although the data we used from United Way and US Chamber of Commerce sources do not provide a complete picture of the convening and coordinating institutions in a given rural community, they indicate the existence of an important governance arrangement beyond government and represent a potential partner or stakeholder for external engagement.

The limitations of our government measures are related to the recentness and reliability of the data collected. Our base source for government measures—the number of municipalities in an area, the type of local government, and whether the county a tract is located in has an elected executive—is the national Census of Governments conducted by the US Census Bureau. This dataset is the result of a series of five-year surveys and depends on self-reporting by local governments. Although response rates are believed to be high, knowing how many local government bodies are not counted because of a lack of capacity to respond is impossible. Moreover, the Census Bureau will at times impute data based on national trends for very small governments if it has received no unique response. In these cases, a government that previously responded may have experienced an unmeasured change in function or capacity. While we cannot know the magnitude of these inaccuracies, noting them here is important

because they are more likely to apply to very small governments of the type that can be found in rural areas.

Opportunities for Future Research

The results from this research suggest that there is no single rural America. Rural places have different strengths and capacities and face diverse challenges. Yet a national, quantitative analysis can only reveal so much. Below, we highlight five key opportunities for future research.

- **Explore, validate, and build upon data locally.** National data sources can provide only a limited amount of detail. Rural communities should test the results of this analysis and validate them against local data. They should also work with community members to identify assets that are not within these datasets to expand on what is available in the tool. Case studies can highlight more nuanced similarities and differences across groups, including through explorations of local histories and communities' successes and challenges in leveraging assets.
- **Evaluate utility and impact.** More evidence is needed to understand the causal connections between the measures used in the typology (both within and across capitals) so we know how they interact to enhance or hinder capacity in rural areas.
- **Contextualize nationally.** Future research could also look for ways to contextualize rural averages by comparing them to national statistics or absolute thresholds (e.g., the ideal ratio of hospital beds per capita), with the goal of highlighting rural strengths in more absolute terms, rather than relative to other types of rural areas.
- **Understand change over time.** This study categorizes rural tracts based on their assets and capacities. Yet questions remain about how rural communities evolve, including how resilient the seven peer groups will be after health and economic shocks such as the COVID-19 pandemic.
- **Connect indicators to outcomes**. In conjunction with case studies, future efforts could also explore how assets and strengths compare with outcomes that rural areas are trying to achieve, such as equitable access to high-quality resources, reductions in poverty, and improvements in people's health.

Appendix B. Variable Loadings for Principal Component Analysis

Capital type	Measure	Principal Components 1	2	3	4	5	6	7
Built	Access to broadband internet	-0.0859	0.020889	-0.21333	-0.27096	0.158298	-0.02953	-0.02209
Built	Access to domestic emergency operations centers	0.006863	0.008196	-0.00832	-0.00522	-0.01597	0.008272	0.017671
Built	Access to fire stations or emergency medical services	-0.00029	-0.00122	0.001759	-0.01654	-0.02017	0.013166	0.009527
Built	Access to highways	0.011615	0.017037	-0.01068	-0.03785	-0.00594	-0.00876	-0.03116
Built	Access to transportation stations	-0.00726	-0.04491	0.07941	0.00081	-0.05345	0.034712	0.002671
Built	Affordable-housing supply	-0.01512	0.004377	0.009058	0.016325	-0.01326	0.000398	-0.06033
Built	Cell service coverage	-0.01203	-0.01146	0.008676	0.021452	0.013261	-0.00213	0.016862
Built	Housing quality	-0.00381	-0.00415	-0.00329	0.000306	0.001509	0.001124	-0.01454
Built	Workers with a short commute	-0.00784	0.145066	-0.04931	-0.19625	0.215659	-0.24873	-0.69593
Cultural	Access to convention centers	0.009816	0.005838	-0.00555	-0.03746	-0.00256	0.004381	-0.02388
Cultural	Capacity of cultural organizations	0.000678	0.000289	-0.00218	-0.0016	-0.00077	-0.00206	-0.0013
Cultural	Employment in cultural occupations	0.039104	0.012911	-0.07146	-0.05052	-0.0227	0.018689	-0.0298
Cultural	Language diversity	0.335904	0.578017	0.151254	-0.38135	-0.05864	0.009446	0.417761
Cultural	Presence of historic properties	0.001111	-0.00332	-0.007	-0.00425	-0.00873	-0.00148	-0.00943
Cultural	Racial and ethnic diversity	0.390511	0.443869	0.174823	0.156771	0.026294	-0.20522	-0.34349
Cultural	Religious diversity	-0.13251	0.020985	-0.12279	0.114788	0.296656	-0.81807	0.377714
Financial	Access to banks and credit unions	0.001762	0.000425	-9.60E-05	-0.00353	-0.00912	0.006648	0.007151
Financial	Banking capacity	0.001667	0.004208	-0.00054	-0.01141	0.014719	-0.01384	-0.03019
Financial	Community development financial institution investments	-0.00064	0.001973	0.000358	0.011124	-0.00493	0.000782	-0.0127
Financial	Federal investment	0.000234	0.001319	-0.00162	-0.00217	-5.61E-05	0.001472	-0.00355
Financial	Home value	0.014611	0.02017	-0.09446	-0.10222	0.011359	0.040132	0.044075
Financial	Individual income	-0.09602	0.019798	-0.13172	-0.18522	0.063311	0.066835	0.07664
Human	Access to child care centers	0.001786	-0.00026	-0.0023	-0.00031	0.000597	-0.0016	-0.00605
Human	Access to college and university campuses	0.002867	0.004953	0.002756	-0.02379	-0.01396	0.012106	-0.00115
Human	Access to farmers' markets	0.014407	0.014456	0.023778	-0.02147	-0.02039	0.021744	0.006214
Human	Access to health care facilities	-0.00173	-0.00307	-0.00303	-0.00773	-0.01728	0.014182	0.020234
Human	Access to health care professionals	0.009852	0.010942	-0.11699	-0.0551	0.036728	-0.01232	-0.01237
Human	Educational attainment	-0.10358	-0.06316	-0.19849	-0.12925	0.105677	-0.0274	-0.10274
Human	Employment to population ratio	-0.14417	0.079523	-0.10771	-0.24482	0.13638	0.057255	0.004165
Human	Health insurance coverage	-0.09092	-0.0519	-0.112	-0.01551	0.036736	0.008002	-0.04664

Capital type	Measure	Principal Components						
		1	2	3	4	5	6	7
Human	Hospital capacity	0.005242	0.000996	0.004831	0.003022	0.021084	-0.01943	-0.04833
Human	Labor force participation	-0.13796	0.075528	-0.11517	-0.21876	0.123529	0.02448	-0.02327
Human	Life expectancy	-0.06852	0.051171	-0.15025	-0.21835	-0.02029	0.098988	0.125546
Human	Nursing home capacity	-0.03005	0.024234	0.002161	-0.01942	0.041292	-0.02742	-0.09099
Natural	Access to coal mines	0.029052	0.104892	-0.09588	-0.08115	-0.01713	-0.03956	0.007856
Natural	Access to oil and natural gas	0.016417	0.056105	-0.062	-0.03879	-0.00723	-0.02106	-0.02725
Natural	Air quality	-0.04403	-0.03579	0.247851	0.40923	0.007783	-0.08867	0.057634
Natural	Cropland coverage	-0.75633	0.496391	0.193135	0.085474	-0.31407	-0.00142	-0.06186
Natural	Lake or river coverage	-0.00312	-0.00045	-0.00088	0.005715	0.001349	-0.00176	0.002293
Natural	Land and natural resource protections	0.003992	-0.00141	-0.04126	-0.0094	-0.03655	0.036244	0.013179
Natural	Park coverage	0.09615	-0.21661	-0.14149	-0.24439	-0.79885	-0.41002	-0.09095
Political	Census participation	-0.13284	-0.00458	-0.04456	0.035074	0.117786	-0.04332	0.006192
Political	Financial capacity of social welfare organizations	5.84E-06	-0.00034	-0.00087	-0.00098	-0.00048	-0.00056	0.001447
Political	Government employment	0.087214	0.01857	-0.00638	0.022891	-0.04403	0.01085	-0.07485
Political	Political competition	-0.10823	-0.32887	0.754293	-0.47109	0.138775	-0.12696	-0.01859
Political	Voter participation	-0.09153	-0.06103	-0.15201	-0.06301	0.018253	0.029582	-0.01993
Social	Access to civic and social organizations	-0.05926	-0.00741	-0.0643	-0.07663	0.036027	-0.04802	-0.01715
Social	Access to places of worship	0.001449	-0.00067	0.001136	0.005533	0.000701	-0.0027	-0.00492
Social	Access to public libraries	0.003077	-0.00537	0.002658	-0.00737	-0.01993	0.01438	0.019781
Social	Access to schools	-0.00107	-0.00331	0.00054	-0.01337	-0.02218	0.013889	0.013674

Appendix C. Landscape of Complementary Tools

American Communities Project

https://www.americancommunities.org/

The American Communities Project groups all US counties into 15 categories, highlighting the diversity of the US in terms of health, education, infrastructure, and community demographics. The project defines rural communities based on population density. A preliminary cluster analysis yielded 13 groupings, and two groupings were created later. First, the Urban Suburbs group was split into two groups depending on whether counties contained any of the nation's 50 largest cities and had population densities above the US average. Second, researchers split a very large category into two groupings (Middle Suburbs and Rural Middle America) based on population density.

Distressed Communities Index

https://eig.org/dci

The Distressed Communities Index focuses on business activities and levels of community prosperity or distress. The index combines metrics of economic and social well-being into a single index to show the economic state of communities relative to others across the country. The index is presented by zip code, and the map does not distinguish between rural and urban.

Economic Resilience and Inclusion Navigator

https://www.stlouisfed.org/publications/bridges/winter-2018-2019/hows-your-county-doing

The Economic Resilience and Inclusion Navigator curates data across a range of topics, including demographic measures (e.g., population size, median age, education, and household structure), economic measures (e.g., household income), health measures (e.g., preventable hospital admissions and premature deaths), and community measures (e.g., racial segregation, crime, and housing affordability). Data are at the county level. This resource is useful for exploring national patterns across a single indicator or making fine-grained comparisons between one and three counties.

Opportunity360

https://www.enterprisecommunity.org/opportunity360

The Opportunity360 Community Dashboard uses indicators organized around the five categories of the Opportunity Framework: housing stability, education, health and well-being, economic security, and mobility. It is an outcome-oriented tool designed for investors to track and evaluate investment

outcomes across the five categories. The measures are provided at the census tract level and cover all tracts in the US.

Rural Opportunity Map

https://ruralopportunitymap.us/

The Rural Opportunity Map visualizes descriptive metrics on business activity and infrastructure, Opportunity Zone designation, and a range of community assets, such as museums and libraries. The data in this tool are either at the point, block, tract, or county level. The user can choose from different rural/urban definitions, including metropolitan and nonmetropolitan counties (Office of Management and Budget), urbanized areas/urban clusters (US Census Bureau), rural areas (US Census Bureau), core-based statistical areas (Office of Management and Budget), rural-urban continuum codes (USDA), frontier and remote areas (USDA), urban influence codes (USDA), rural counties (Economic Development Administration), rural health clinics program eligibility (Centers for Medicare and Medicaid Services), grant program eligibility (Federal Office of Rural Health Policy), and RUCA codes.

Understanding Communities of Deep Disadvantage

https://poverty.umich.edu/projects/understanding-communities-of-deep-disadvantage/

The purpose of this project is to broaden the poverty lens from more traditional, income-based measures to other indicators of disadvantage, including life expectancy, birth weight, and social mobility. The result is an index of deep disadvantage for all counties and the 500 largest cities (constructed using principal component analysis), as well as five case studies. The online tool does not allow users to differentiate between rural or urban areas, but the accompanying report does note that most communities that rank high on the index are rural. The final, cleaned data are available for download from the website.

Notes

[1] "RUCA Data," Rural Health Research Center, accessed March 10, 2020, https://depts.washington.edu/uwruca/ruca-codes.php.

[2] A brief overview of RUCA codes and the codes themselves are available at https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/.

[3] "Defining Rural Population," Health Resources and Services Administration, accessed March 10, 2020, https://www.hrsa.gov/rural-health/about-us/definition/index.html.

[4] A brief overview of the Census Bureau's urban-rural classification is available at https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html.

[5] More information on the Office of Management and Budget's classification is available at https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural/. More information on rural-urban continuum codes is available at https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx.

[6] We also explored aggregating adjacent tracts that have the same RUCA code and are in the same county. Ultimately, we maintained a tract-level analysis after identifying three challenges with tract aggregation: (1) greater variation across units of analysis if some units are entire counties and others are single tracts, (2) the potential for oddly shaped areas (e.g., creating a "doughnut," with one RUCA type surrounding another in the center), which would complicate distance calculations; and (3) the potential for interpretation confusion because people are likely more familiar with census tracts than with RUCA codes.

[7] Washington, DC, is not included because it does not have any rural tracts according to the definition we used.

[8] For example, data we used on lakes and rivers from the US Army Corps of Engineers and data we used on public and private schools include some data points from before 2010.

[9] For example, we excluded some measures that would not have good coverage for rural areas, such as state fairgrounds, which are commonly located outside of rural areas, or major sports venues. We also excluded measures that had the potential for urban bias, such as the US Environmental Protection Agency's National Walkability Index.

[10] We excluded any spending for other geographic levels, including state, multicounty, multiple zip codes, and city.

[11] No crosswalk exists to convert Alaska's voting districts to counties.

[12] We also tested running PCA by capital type, with the goal of having each type of capital described by one or two components. However, this method was much less efficient and would have required using about half of all 50 principal components to reach a similar amount of variance as the previously described method.

[13] Results can vary slightly based on the starting point. We used a random number generator online and selected 888 as the random number seed.

[14] Standard scale ("z-score") transforms data so all measures have a standard score, meaning the mean is 0, the standard deviation for the measure is 1, and thus the variance is also 1. Values that are greater than the mean will be positive, and values that are below the mean will be negative. The relative distance of the negative and positive values is determined by the standard deviation of the original measure. See Juhi Ramzai, "Clearly Explained: What, Why and How of Feature Scaling—Normalization and Standardization," Towards Data Science, May 17, 2020, https://towardsdatascience.com/clearly-explained-what-why-and-how-of-feature-scaling-normalization-standardization-e9207042d971.

[15] For more information on rural data sources and challenges, see Scally, Burnstein, and Gerken (2020).

[16] For a study on how rural broadband speeds compare with those in urban areas, see Prieger (2013). For studies that explore the relationship between broadband speed and employment, see Lobo, Alama, and Whitacre (2020) and Whitacre, Gallardo, and Strover (2014).

[17] See "Child Care Quality Ratings," Child Care Aware of America, https://www.childcareaware.org/families/child-care-quality-ratings/.

[18] See J.D. Power, "Retail Banks Face Major Customer Satisfaction Challenge as World Shifts to Digital-Only Engagement," news release, April 30, 2020, https://www.jdpower.com/business/press-releases/2020-us-retail-banking-satisfaction-study.

[19] See "Hospital Compare," Centers for Medicare and Medicaid Services, last modified October 1, 2020, https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare.

[20] We excluded places of worship because the social capital component of the main framework includes a measure for these institutions.

References

Deslatte, Aaron, and Eric Stokan. 2020. "Sustainability Synergies or Silos? The Opportunity Costs of Local Government Organizational Capabilities." *Public Administration Review* 80 (6): 1024–34. https://doi.org/10.1111/puar.13237.

Flora, Cornelia Butler, Mary Emery, Susan Fey, and Corry Bregendahl. n.d. "Community Capitals: A Tool for Evaluating Strategic Interventions and Projects." Ames, IA: Iowa State University, North Central Regional Center for Rural Development.

Flora, Cornelia Butler, Jan L. Flora, and Susan Fey. 2003. *Rural Communities: Legacy and Change*. 2nd ed. Boulder, CO: Westview Press.

GAO (US Government Accountability Office). 2017. *DATA Act: OMB, Treasury, and Agencies Need to Improve Completeness and Accuracy of Spending Data and Disclose Limitations*. Washington, DC: GAO.

———. 2018. *Open Data: Treasury Could Better Align USAspending.gov with Key Practices and Search Requirements*. Washington, DC: GAO.

———. 2019. *DATA Act: Quality of Data Submissions Has Improved but Further Action Is Needed to Disclose Known Data Limitations*. Washington, DC: GAO.

Lobo, Bento J., Md Rafayet Alama, and Brian E. Whitacre. 2020. "Broadband Speed and Unemployment Rates: Data and Measurement Issues." *Telecommunications Policy* 44 (1), 101829.

Moore, Susan A., Roger C. Severn, and Rebecca Millar. 2006. "A Conceptual Model of Community Capacity for Biodiversity Conservation Outcomes." *Geographical Research* 44 (4): 361–71. https://doi.org/10.1111/j.1745-5871.2006.00407.x.

Prieger, James E. 2013. "The Broadband Digital Divide and the Economic Benefits of Mobile Broadband for Rural Areas." *Telecommunications Policy* 37 (6–7): 483–502.

Scally, Corianne Payton, Eric Burnstein, and Matthew Gerken. 2020. *In Search of "Good" Rural Data: Measuring Rural Prosperity*. Washington, DC: Urban Institute.

Whitacre, Brian, Roberto Gallardo, and Sharon Strover. 2014. "Broadband's Contribution to Economic Growth in Rural Areas: Moving towards a Causal Relationship." *Telecommunications Policy* 38 (11): 1011–23.

Wu, JunJie. 2008. "Land Use Changes: Economic, Social, and Environmental Impacts." *Choices: The Magazine of Food, Farm, and Resource Issues* 23 (4): 6–10.

Acknowledgments

This research was funded by the Robert Wood Johnson Foundation. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and do not necessarily reflect the views of the Robert Wood Johnson Foundation or the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

Special thanks to our partners at US Department of Agriculture Rural Development and the Federal Reserve Board of Governors and to the members of our stakeholder advisory group and technical working group for their expert guidance and feedback on data, methodology, and tool development.

We would also like to thank our Urban Institute colleagues Jorge Morales-Burnett and Patrick Spauster for their assistance in cleaning and analyzing data, Aaron Williams for his guidance on principal component analysis and cluster analysis, and Christina Stacy for her methodological guidance during the design of this study and her detailed review and feedback on the technical appendixes.

Finally, we appreciate all the organizations that collect, maintain, and share data on rural places, including the Homeland Infrastructure Foundation-Level Data, the National Center for Charitable Statistics, United Way, and the National Association of Counties. Without their careful stewardship and commitment to data sharing, this study would not have been possible.

For more information on this project, see https://reenvisioning-rural-america.urban.org/.

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.