
Accuracy of equity-related measures in 2020

An analysis of American Community Survey data

Introduction

As a leading source of socioeconomic data, the U.S. Census Bureau's American Community Survey (ACS) is used by community leaders, state and federal agency staff, business owners, philanthropists, and others who are planning and managing resources for future generations. These data are used to allocate funding and prioritize investments.

In 2020, the COVID-19 pandemic disrupted ways of reaching people, and new methods for protecting personal information were introduced by the Census Bureau. People who rely on ACS data have been concerned that the data gathered in 2020 may be less accurate and ultimately unreliable for decision-making.

We analyzed a suite of equity-related socioeconomic characteristics, including poverty, language barriers, and rental housing, to determine the changes in ACS data accuracy from 2019 and 2020. Some degree of inaccuracy is inevitable in surveyed data, and the Census Bureau reports estimated margins of error for every data value. We used a statistical formula, called a coefficient of variation, to compare accuracy between 10 ACS equity-related measures.

The 10 equity-related measures analyzed are commonly used for planning and resources management. They are used in Headwaters Economics' decision-support tools, including our [Economic Profile System](#), [Neighborhoods at Risk](#), [Wildfire Risk to Communities](#), and the [Rural Capacity Map](#), which are viewed more than 150,000 times annually. The measures we selected are also found in federal agency tools that influence allocations of resources, including the [Climate and Economic Justice Screening Tool](#), the [Environmental Justice Screening and Mapping Tool](#), the [Social Vulnerability Index](#), and the [Energy Justice Dashboard](#).

For each of the equity-related measures, we compared data accuracy overall, between types of counties (rural vs. metro), and among regions of the country. We found that:

- Equity-related measures from the American Community Survey were slightly less accurate in 2020 when compared to 2019. The median change in percent of counties with inaccurate data was 2.5%.
- Among the measures analyzed, declines in accuracy were most severe for data representing families in poverty, mobile homes, and households without a car.
- The lack of accurate data was most apparent in rural areas. In 2020, more than half of rural counties lacked accurate data for counts of people not proficient in English, households without a car, families in poverty, and mobile homes.

- Among U.S. regions, the South experienced the most widespread declines in ACS data accuracy.

We defined “inaccurate data” as having a coefficient of variation greater than 12%. Coefficients of variation (CVs), which are calculated from the margins of error reported in ACS, measure sampling error and are commonly used to assess the reliability of ACS estimates. A CV greater than 12% indicates that an ACS value should be interpreted with caution. At greater than 40%, the estimate is considered very unreliable. (See the Methods section below for more detail.)

The 2019 data analyzed includes survey responses from 2015-2019 while the 2020 data includes survey responses from 2016-2020. We selected these years because they have four years (2016-2019) in data congruency, and one year of data that separates them (2015 and 2020, respectively).

Why did American Community Survey data accuracy decline in 2020?

Two primary factors influence data accuracy in the [American Community Survey](#) (ACS):

- (1) sample size (the share of the population surveyed each year), and
- (2) the amount of “noise” introduced to protect the privacy of the people surveyed.

In 2020, both factors changed. The sample of surveyed individuals was smaller due to COVID-19, and the amount of “noise” introduced was greater than in previous years.

The ACS survey typically samples approximately 2.5% of U.S. households, or about 1% of the total population, annually to develop estimates of a host of socioeconomic indicators. Because the survey does not attempt to obtain information from every person, the data are estimates with inherent uncertainty. As the sample of households decreases, uncertainty increases, and the data become less reliable. COVID-19 disrupted ACS data collection, and as a result, the average county saw an [8.4% decrease](#) in data gathered in 2020. In some counties in Missouri, Alabama, and Alaska the decline in sampled households exceeded 20%.

Compounding the uncertainty in the 2020 ACS data, the Census Bureau began utilizing a different method for ensuring privacy after a 2019 case study found high exposure for personal identification. Since the 1980s the Census Bureau has been [introducing “noise”](#) into the publicly accessible data tables. Noise is a method for obscuring data to protect individuals who respond to the survey. Throughout the years, data scientists at the Census Bureau have developed new strategies for generating data noise to better protect privacy while attempting to ensure data are of reasonable quality for decision-making. New techniques were employed for ACS in 2020.

Findings

Finding 1: The share of counties with inaccurate data increased slightly in 2020.

Across the suite of equity-related measures we analyzed, the median increase in percent of counties with inaccurate data from 2019 to 2020 was 2.5%.

The number of counties with inaccurate data for families in poverty, mobile homes, and households without a car all increased by greater than 5 percentage points. For the other equity-related measures, the change in the number of counties with inaccurate data ranged from 0 to 4 percentage points.

Table 1. The number and share of counties with inaccurate data in 2019, 2020, and the change between those years.

Equity-related measures	Counties with inaccurate data 2019	Counties with inaccurate data 2020	Change in counties with inaccurate data 2019 to 2020
Not proficient in English	2,478 (79%)	2,510 (80%)	+32 (1%)
Household without a car	2,007 (64%)	2,196 (70%)	+189 (6%)
Families in poverty	1,890 (60%)	2,064 (66%)	+174 (6%)
Mobile homes	1,321 (42%)	1,538 (49%)	+217 (7%)
Individuals in poverty	948 (30%)	1,073 (34%)	+125 (4%)
Housing units that are rentals	411 (13%)	545 (17%)	+134 (4%)
Population under 5	333 (11%)	376 (12%)	+43 (1%)
Population over 65	90 (3%)	116 (4%)	+26 (1%)
Median household income	85 (3%)	124 (4%)	+39 (1%)
Minority population	29 (1%)	37 (1%)	+8 (0%)

Values shown in parentheses are the share of total U.S. counties.

Although the change in the number of counties with inaccurate data was relatively small (less than or equal to 7%), the table shows that accuracy was poor to begin with for many equity-related measures. For example, in 2019, nearly 80% of counties (2,478 counties) lacked accurate counts of people who are not proficient in English. Also in 2019, more than two-thirds of counties lacked accurate counts of households without cars and families in poverty.

Among the equity-related measures we analyzed, estimates for mobile homes saw the largest degradation of accuracy. Beyond the 1,321 counties with inaccurate mobile home counts in 2019, an additional 217 counties had inaccurate mobile home counts in 2020.

As a result of poor data quality, many counties lack accurate data for multiple equity-related measures, such as households without a car and families in poverty. The likelihood that both demographics occur in a single household is high, compounding vulnerabilities and consequences if inaccurate data prevent resources from reaching those with the greatest need.

Finding 2: In 2020, ACS data quality was worse in rural counties.

Across the suite of equity-related measures we analyzed, the median percent of rural counties with inaccurate data in 2020 was 34%. By comparison, the median percent of urban counties with inaccurate data was 13%.

We found that more than half of rural counties lacked accurate data for counts of people not proficient in English, households without a car, families in poverty, and mobile homes. More than one-quarter of rural counties lacked accurate data for counts of rental units. More than 15% of rural counties lacked accurate counts of children under the age of five. Due to the lack of accurate socioeconomic data, federal programs and services may not be reaching rural residents or funds may be getting overallocated. Either situation interferes with the equitable distribution of benefits.

Table 2. The number and share of rural and urban counties with inaccurate data in 2020.

Equity-related measures	Rural counties with inaccurate data 2020	Urban counties with inaccurate data 2020
Not proficient in English	1,744 (88%)	766 (66%)
Household without a car	1,677 (85%)	519 (45%)
Families in poverty	1,584 (80%)	480 (41%)
Mobile homes	1,141 (58%)	397 (34%)
Individuals in poverty	848 (43%)	225 (19%)
Housing units that are rentals	467 (24%)	78 (7%)
Population under 5	337 (17%)	39 (3%)
Median household income	112 (6%)	12 (1%)
Population over 65	106 (5%)	10 (1%)
Minority population	36 (2%)	1 (0%)

Values shown in parentheses are the share of rural counties and share of urban counties.

Finding 3: Declines in 2020 data accuracy impacted all U.S. regions.

More than 20% of counties in every U.S. region saw at least one previously accurate equity-related measure become inaccurate in 2020.

Table 3. Counties where one or more equity-related measures became inaccurate in 2020.

Region	Counties where one or more equity-related measures became inaccurate in 2020
South	268 (38%)
Great Lakes	197 (38%)
Gulf Coast	189 (35%)
West	99 (35%)
Midwest	228 (33%)
Pacific Coast	39 (23%)
Northeast	55 (22%)

Values shown in parentheses are the share of counties per region.

The U.S. South experienced the most widespread declines in accuracy. More than 250 southern counties (38% of counties in the South) saw one or more equity-related measures change from accurate in 2019 to inaccurate in 2020. The Great Lakes, Gulf Coast, West, and Midwest also experienced widespread declines in accuracy. In these regions, more than one-third of counties saw one or more equity-related measures change from accurate in 2019 to inaccurate in 2020.

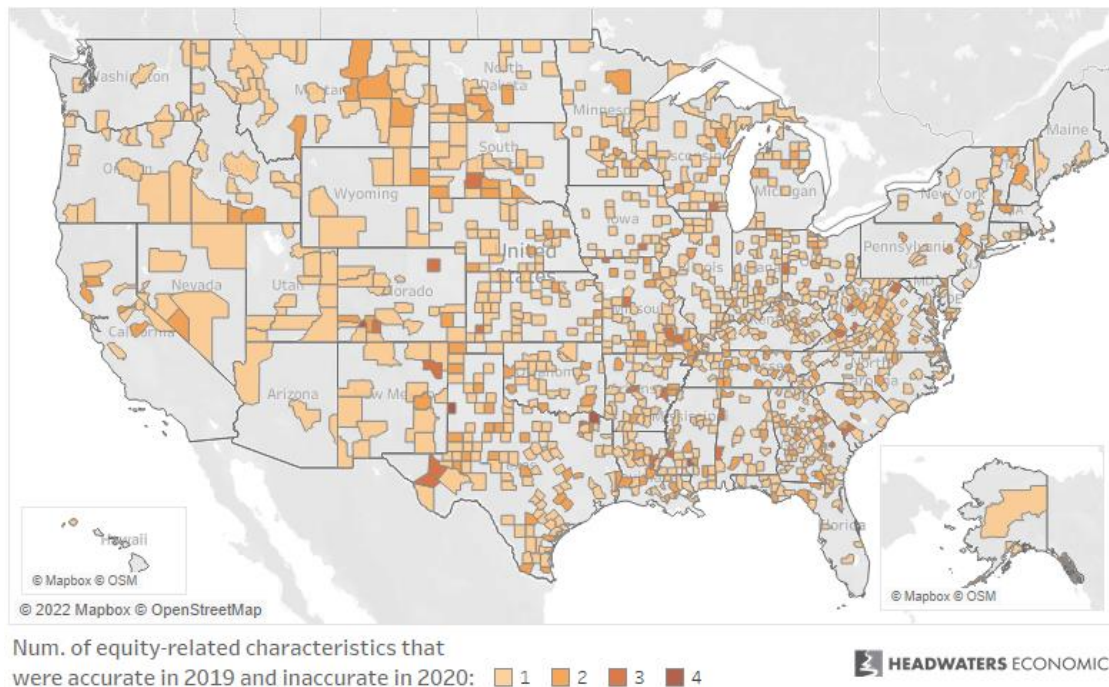


Figure 1. The number of equity-related measures that became inaccurate in 2020 by county.

Summary

Overall, equity-related data from the American Community Survey were slightly less accurate in 2020. Although the change in counties with inaccurate data was relatively small, the impact was geographically widespread – affecting counties in every U.S. region. Across the country, rural areas lacked reliable data far more often than their urban counterparts.

Many equity-related measures had poor accuracy prior to 2020. Because of this, any degradation of accuracy is concerning. This is particularly true for rural areas and for areas with high poverty where the lack of accurate socioeconomic data may impact eligibility for programs and services.

With dedication and resources, the Census Bureau can improve data accuracy. The agency has previously improved data for geographies and populations where accurate data was historically lacking. For example, from 2010 to 2020, the share of tribal areas and reservations with accurate or moderately accurate data increased from 66% to 71%.

To achieve equity goals, more attention and resources are needed to improve data accuracy for rural communities and for specific equity-related measures. Users of ACS data can advocate for that change and seek out tools that clearly communicate ACS data accuracy. Agencies using ACS to allocate resources should allow flexibility and supplemental data to ensure help reaches communities most in need.

Methods

We define “inaccurate data” as having a coefficient of variation (at a 90% confidence level) greater than 12%. Coefficients of variation (CVs), which are calculated from the margins of error reported in ACS, measure sampling error and are commonly used to assess the reliability of ACS estimates. A CV greater than 12% indicates that an ACS value should be interpreted with caution. At greater than 40%, the estimate is considered very unreliable. The following equation is used to calculate a CV from the ACS-reported margin of error (MOE) for any reported estimate: $CV = ((MOE/1.645)/Estimate)*100$.

This analysis compares ACS data accuracy in 2019 versus 2020. The 2019 data include survey responses from 2015-2019 while the 2020 data include survey responses from 2016-2020. We selected these years because they have four years (2016-2019) in data congruency, and one year of data that separates them (2015 and 2020, respectively).

In this report, we refer to metropolitan or micropolitan counties as “urban” and nonmetropolitan counties as “rural.” The delineation is based on Census Bureau data. Metropolitan Statistical Areas are counties that have at least one urbanized area of 50,000 or more people, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. Micropolitan Statistical Areas are counties that have at least one urbanized area of 10,000 to 50,000 people, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. We refer to a county as “rural” if it is not designated as either metropolitan or micropolitan.

Contact

Patricia Hernandez, Executive Director | 406.599.7425 | patty@headwaterseconomics.org
Kelly Pohl, Associate Director | 406.599.7841 | kelly@headwaterseconomics.org

About Headwaters Economics

Headwaters Economics is an independent, nonprofit research group whose mission is to improve community development and land management decisions. <https://headwaterseconomics.org/>