Chapter 7
Confronting Statistical Uncertainty in Rural America: Toward More Certain Data-Driven Policymaking Using American Community Survey (ACS) Data

Jason R. Jurjevich

Abstract Aging and lacking infrastructure are major impediments to economic development in rural America. To address these issues, civic leaders often look to state and federal infrastructure grant/loan funding, where eligibility is often based on income requirements established by the US Census Bureau's American Community Survey (ACS). The problem, especially for rural communities, is that ACS data contain a high degree of statistical uncertainty (i.e., margin of error) that is often disregarded for determining program eligibility. For rural communities with unreliable income estimates, the most common work-around involves hiring a consultant to conduct an income census or survey to formally challenge the US Census Bureau's ACS estimate. Many rural communities, however, elect not to formally challenge unreliable ACS estimates either because they are unaware that reimbursement for conducting an income survey is an allowable expense under some grant/loan programs or they are dissuaded by the necessary time and resources. First, I summarize whether federal infrastructure grant/loan programs incorporate MOE values when determining community eligibility. Second, I examine the degree to which ACS estimates are statistically reliable for communities across rural America. Finally, using an example from Oregon, I recommend guidelines for how states can assist rural communities with statistically unreliable ACS estimates. These findings can help rural communities secure infrastructure funding that advances economic development and quality of life, and potentially support reliable data-driven policy and decision-making more broadly.

Keywords American Community Survey (ACS) · Rural · Demographic data · Margin of error (MOE) · Data-driven policy

J. R. Jurjevich (✉)
Toulan School of Urban Studies and Planning, Population Research Center (PRC), Portland State University, Portland, OR, USA
e-mail: jjason@pdx.edu

R. S. Franklin (ed.), *Population, Place, and Spatial Interaction*, New Frontiers in Regional Science: Asian Perspectives 40,
https://doi.org/10.1007/978-981-13-9231-3_7

7.1 Introduction

Rural communities across the USA are facing similar developmental challenges. The lack of water, sewerage, and transportation infrastructure, combined with rapidly aging assets more broadly, are major impediments to advancing economic development. Rural civic leaders often look to secure funding from state and federal infrastructure grant/loan programs to address these issues. Qualifying for grants/loans typically requires that the community meet strict requirements around socioeconomic measures that are often determined using the US Census Bureau's (USCB) American Community Survey (ACS). In 2008, ACS data determined 29% ($416 billion) of government assistance program funding and 69% ($389 billion) of all federal grant funding (Carpenter and Reamer 2010).

Unlike the decennial census, an enumeration of an entire population, ACS data are drawn from a statistical sample and therefore contain sampling error. As with all survey data, sampling error represents the risk that data drawn from statistical samples may not accurately represent the broader population. The Bureau conveys this risk, referred to as statistical uncertainty, through a margin of error (MOE) statistic that accompanies ACS estimates.

Navigating the challenges of statistical uncertainty in ACS data and other sources of quantitative demographic data is a key challenge in today's data-driven world. Overcoming these challenges is even more difficult for rural communities, where MOE values are often disregarded for determining federal program eligibility (Nesse and Rahe 2015). Consider a town of 2000 people in rural Lake County, Oregon. The town's median household income (MHI) estimate, according to 5-year ACS data, is $35,900 with an accompanying MOE value of ±$3500. Given that most grant/loan programs disregard MOE values, this town would be considered ineligible because the MHI figure exceeds the $35,000 grant threshold. This approach is problematic because when the MOE is considered, the estimate range is $32,400 ($35,900 − $3500) to $39,400 ($35,900 + $3500): there is a roughly equal chance that the town is in fact eligible.

What does a rural community do in a situation like this? The typical response is to pursue alternative approaches that establish the "true" population value in order to qualify for grants/loans. The most common work-around involves hiring a consultant to conduct an income census or survey to formally challenge the USCB's ACS estimate. Many rural communities, however, elect not to pursue this approach. Often, either they are unaware that some grant/loan programs reimburse costs for conducting an income census or they are dissuaded by the necessary time and resources.

Challenges that arise from statistical uncertainty introduce state-level policy implications and yield real, on-the-ground effects. Currently, states must prioritize financial resources to ensure that eligible communities receive their share of federal funding for community development. This challenge is particularly difficult for states with limited financial resources and/or a large rural population. Without sufficient funds to file a formal challenge, towns often delay or abandon capital projects until they can secure resources to hire a consultant. This situation creates negative impacts on citizens, particularly among the most vulnerable—including racial/ethnic minorities and low-income populations—by forcing individuals to commute long distances for affordable housing, food, and work.

Best practices for working with ACS data and practical examples and alternative approaches for reducing MOE values are well-documented in the existing literature. Aggregating geographies to reduce MOE, for example, is generally not a feasible approach because many infrastructure grant/loan programs require rural communities use place-level income estimates from the ACS. To highlight the challenge of statistical uncertainty in rural America and flesh out the unique challenge for rural communities—there are few, if any, alternative approaches to improving data quality—I address two interrelated questions in this chapter: (1) To what extent are ACS income estimates unreliable for rural communities across the USA? and (2) What policies can states implement to assist rural communities with statistically unreliable ACS estimates and qualify them for infrastructure grant/loan programs?

This chapter proceeds as follows: first, I summarize whether federal infrastructure grant/loan programs incorporate MOE values when determining community eligibility. Second, I examine the degree to which ACS estimates are statistically reliable for communities across rural America. Finally, using an example from Oregon, I recommend approaches and policies that states can implement to help rural communities with statistically unreliable ACS estimates. Presently, MOE values are often disregarded for determining program eligibility, and state support is either insufficient or nonexistent to help rural communities formally challenge unreliable ACS estimates. Given this scenario, I offer recommendations and brightlines for prioritizing state resources to account for MOE values when determining infrastructure grant/loan eligibility. These results are critical steps toward more certain data-driven policy and decision-making in rural America.

7.2 American Community Survey (ACS)

Between 1970 and 2000, the USCB administered two different forms to collect decennial census data: the short and long form (USCB 2019). In Census 2000, five out of six households received the short form, which contained approximately ten questions gathering data on age, sex, race/ethnicity, household relationship, and housing tenure. The long form surveyed one of six households and asked more detailed social, economic, and housing-specific questions. Long-form data were obsolete soon after each decennial census, however, a significant limitation for rapidly changing communities (Citro and Kalton 2007).

Rising costs of administering the long form, demand for timelier sociodemographic data, and concerns around confidentiality led the USCB to implement the ACS in 2005, replacing the long form (Torrieri 2007). ACS data, like

$$CV = \frac{SE}{\hat{X}} \times 100$$

Fig. 7.1 Coefficient of variation CV

decennial long-form data, are drawn from a statistical sample and are an approximated quantity rather than actual, true counts. This means that the degree of statistical uncertainty, expressed in the MOE statistic, represents a range of values expected to contain the true value of the quantity being estimated. For example, when the MHI estimate and corresponding MOE ($35,900, ±$3500) for a town in Lake County, Oregon, are considered, the range of values containing the true estimate is $32,400 ($35,900 – $3500) to $39,400 ($35,900 + $3500). The USCB reports MOE values with 90% statistical confidence, meaning that users desiring greater statistical confidence (e.g., 99%) will encounter greater statistical uncertainty.

Compared to long-form data, ACS data are drawn from a greatly reduced sample size. To put this into context, consider that the long form was administered to 1 in 6 households, while ACS estimates are derived from a sample of roughly 1 in 40 households.[1] The differences in sample size are meaningful for three key reasons. First, because long-form data were drawn from such a large statistical sample, the USCB did not report MOE values for long-form estimates (Spielman et al. 2014). This has, in part, created a situation where some data users do not understand statistical uncertainty in ACS data and avoid engaging with or reporting MOE values altogether (Jurjevich et al. 2018). Second, ACS data contain a greater degree of statistical uncertainty than decennial long-form data. USCB officials initially estimated that ACS estimates would have a 33% higher sampling error than long-form estimates. In the end, however, the sampling error ended up being roughly 75% higher than decennial long-form estimates (Spielman et al. 2014; Navarro 2012). Third, MOE values are higher for cross-tabulated data (e.g., child poverty) and for small-area and rural geographies.

To reduce statistical uncertainty and corresponding MOE values in ACS data, scholars recommend various alternative approaches, including collapsing data detail or aggregating census geographies (Citro and Kalton 2007; Heuvelink and Burrough 2002; National Academy of Sciences 2015; Spielman and Folch 2015; Spielman et al. 2014; USCB 2009). However, these recommendations are not feasible strategies for rural communities trying to become eligible for state and federal infrastructure grant/loan programs because many programs require rural communities use place-level income estimates from the ACS.

[1]Sampling odds are an effective way to compare the sampling frame of the ACS to the decennial long form. However, to be clear, the ACS sample is drawn from housing units and the group quarters population and is not based on a sampling rate. The 2017 ACS sample, for example, was drawn from roughly 2.1 million housing units and almost 300,000 individuals rising in group quarters, which together include more than 5,000,000 individuals.

A common way to express survey data reliability is through a statistic called the coefficient of variation (CV) (Fig. 7.1). Recommended by the USCB to evaluate data reliability, the CV expresses sampling error relative to the estimate and is calculated by dividing the standard error (SE) by the statistical estimate and multiplying the result by 100 (USCB 2009). Larger CV values indicate lower reliability. The Environmental Systems Research Institute's (ESRI 2014) "red-yellow-green" schematic is most often cited because the color schematic, combined with the simplicity of the CV values, is useful for conveying the reliability of ACS data to users unfamiliar with the concepts of data reliability (Jurjevich et al. 2018). Here, CV values less than 12% indicate a high degree (i.e., green) of reliability; CVs between 12 and 40 are somewhat (i.e., yellow) reliable; and CVs greater than 40 indicate little, if any (i.e., red), reliability.

Consider, for example, the estimate of children living at or below the poverty level is 20% (±3%) for Tract A and 25% (±15%) for Tract B. After deriving the standard errors from the MOE values (1.8% for Tract A and 9.1% for Tract B), the CVs are 9.0% and 36.4% for the two tracts, respectively.[2] This means that the poverty estimate for Tract 1 is reliable while the estimate for Tract 2 is somewhat reliable.

One key limitation of CV values is they are somewhat subjective; there are no hard-and-fast rules on cutoff values. To illustrate this point, consider these two different standards: (1) according to the National Academy of Sciences (2015), the USCB judges an estimate to be statistically reliable when the tract-level estimate for a "key variable" is less than or equal to 30%, and (2) the National Research Council (see Citro and Kalton 2007), on the other hand, recommends CV values be less than or equal to 10% as a standard level of precision. Together, these differences underscore a critically important point: achieving consensus on generally accepted CV threshold values for statistical reliability (or for specific use-case scenarios) is essential for professionals working in government, academia, and applied practice, and can also make it easier for novice data users to navigate and account for statistical uncertainty in ACS data.

7.2.1 *Sampling Methodology*

To ensure that ACS data are representative across age, sex, race/ethnicity, educational attainment, and urban/rural communities, the USCB implements a highly complex sampling methodology. What is particularly salient is the way the Bureau develops and implements its sampling methodology for targeting households. The likelihood of being selected for the ACS sample varies according to the geography

[2]The standard error is calculated by dividing the MOE values (3 and 15%) by 1.645 (z-score at 90% statistical confidence, which is the statistical level of confidence at which the USCB reports the MOE value).

Table 7.1 Corresponding CV values of ACS sampling methodology, 2005–2010

Tract size category	Average tract size	Coefficient of variation (CV)
0–400	291	66%
401–1000	766	41%
1001–2000	1485	29%
2001–4000	2636	26%
4001–6000	4684	19%
6000 +	8337	15%

Source: Asiala (2012)

Table 7.2 Corresponding CV values of ACS sampling methodology, 2011–present

Tract size category	Average tract size	Coefficient of variation (CV)
0–400	291	41%
401–1000	766	30%
1001–2000	1485	29%
2001–4000	2636	29%
4001–6000	4684	29%
6000 +	8337	28%

Source: Asiala (2012)

(i.e., urban or rural area) and the size of the census tract (i.e., the deviation from the average tract size of 6000 residents). For each ACS executed between 2005 and 2010, the sample was largely drawn from larger census tracts in order to yield more reliable estimates (Asiala 2012). As such, Table 7.1 shows that CV values were lowest and most reliable for larger (more populated) census tracts, which more often than not are in urban areas. On the other hand, smaller census tracts (largely in rural areas) contained higher CV values, indicating lower levels of statistical reliability.

In 2011, the Bureau made two significant changes in the ACS sampling methodology: (1) increasing the sample size from 2.9 to 3.5 million housing units and (2) changing the sampling frame to select more cases from smaller tracts, largely in rural areas (see National Academy of Sciences 2015, Alvarez and Salvo 2014, and USCB 2011). Table 7.2 illustrates how the combination of a larger ACS sample size, along with the shift in methodology, resulted in more equitable reliability across different sized census tracts. The main takeaway for rural areas is that data from the 2011 ACS (and subsequent years) should yield more reliable estimates than earlier ACS periods (i.e., 2010 and prior).

7.3 Federal Grant and Loan Infrastructure Programs

Many federal grant and loan programs rely on ACS data to determine program eligibility, to allocate funds, and for program assessment. Federal agency programs using ACS income data include, but are not limited to, the following: the US Department of Housing and Urban Development's (HUD) assisted housing programs; identifying distressed or underserved nonmetropolitan areas as part of the Community Reinvestment Act (CRA); determining the grant amount and fixed interest rate for community loans related to the US Department of Agriculture (USDA) Rural Development Wastewater Infrastructure Fund; setting eligibility parameters for the US Department of Agriculture Supplemental Nutrition Assistance Program (SNAP); and determining which communities are eligible for Technical Assistance and Public Works grants from the US Economic Development Administration.

Despite their widespread use, many agencies use ACS income data to determine program eligibility without accounting for accompanying MOE values. Nesse and Rahe (2015) explored this issue in more depth. They specifically looked at how the US Department of Housing and Urban Development (HUD), Department of Agriculture, and the Department of Transportation[3] use ACS data for administering their programs. Of particular interest was how these agencies incorporated MOE, if at all, in policy governance. Nesse and Rahe (2015) found that although most federal agencies recognize that ACS estimates are subject to accompanying statistical error, they generally do not consider MOE because the added complexity provides a marginal program benefit. Along these lines, the USDA Rural Development issued an administrative notice in 2012 instructing state and local officials to administer loan guarantee and grant programs using income and poverty data from the ACS. The notice, however, does not mention or establish policies for incorporating MOE (USDA 2012).

One of the only federal agencies to explicitly consider statistical uncertainty of ACS data in their programs, particularly for small areas, is HUD.[4] According to Usowski et al. (2008, p. 206), HUD aims to "minimize the possibility of publishing income estimates in which the annual change is more a reflection of the variation in estimation errors than a reflection of changes in underlying economic conditions."

[3]The specific programs examined by Nesse and Rahe (2015) include the Housing Choice Voucher Program, Supplemental Nutrition Assistance Program (SNAP), and the Urbanized Area Formula Program.

[4]Although HUD considers MOE for developing annual income limits, the agency does not consistently report MOE across all data programs. For example, see the Comprehensive Housing Affordability Strategy (CHAS) dataset at: http://www.huduser.org/portal/datasets/cp.html

To this end, HUD develops annual income limit estimates[5] based on a calculation that gives less emphasis to area estimates with high MOE values.

Why, given the potentially serious limitations, do some grant/loan programs disregard statistical uncertainty in their eligibility criteria? In addition to Nesse and Rahe's (2015) conclusion that agencies generally do not consider MOE, another possible factor is there may be little incentive to address this problem. Underscoring this point, in 2017 the Appropriations Committee in the US House of Representatives approved a measure mandating that HUD report areas in the USA where income data from the ACS, used to determine program eligibility, had an accompanying MOE of 20% or higher (Caster 2017). The provision, part of the FY 2018 omnibus spending package, was approved in March 2018.[6]

7.4 Geographic Delineation of "Rural America"

The geographic delineation of areas considered "rural" has long been of interest to researchers. As a result, there is a well-established body of research that defines, quantifies, and operationalizes varying degrees of rurality (e.g., Isserman 2005, Cromartie and Swanson 1996, McGranahan et al. 1986). A common starting point is the USCB's urban-rural classification. With census blocks as the primary geographic unit, the Bureau uses population density, land use, and distance measures to determine whether a census block qualifies as urban (Ratcliffe et al. 2016). Urban areas include communities that meet one of two following subclassifications: (1) an urbanized area (UA), which includes 50,000 or more people, and (2) an urban cluster (UC), which includes at least 2500 and fewer than 50,000 people.[7] Population, housing, and areas outside of urbanized areas and urban clusters are considered rural (Ratcliffe et al. 2016).

A limitation of the USCB's urban-rural continuum is that the binary nature of the definition is not particularly well-suited for classifying small towns. Small towns of 3000 to 5000, for example, are often classified as urban.[8] Consider the village of Lancaster, Wisconsin, a community of almost 4000 located in a remote corner of

[5]As Usowski et al. (2008) point out, annual income limits determine eligibility for the Public Housing program, Section 8 Housing Assistance Payments program, Section 2020 Supportive Housing for the Elderly, and Section 811 Supportive Housing for Persons with Disabilities.

[6]In response to the directive, HUD now publishes the MOE data for all block groups and all places. See: https://www.hudexchange.info/programs/acs-low-mod-summary-data/

[7]Ratcliffe et al. (2016) note that a minimum of 1500 people must reside outside of group quarters in order for an area to be classified as urban.

[8]Scholars have developed innovative approaches, including the ERC Rural-Urban continuum codes, that contextualize the degree of rurality.

Southwest Wisconsin. In 2010, the USCB classified Lancaster as urban.[9] Arguably, communities like Lancaster are more closely aligned with small towns of a few hundred people compared to more urbanized places. Given this situation, I define rural America by including rural places and small towns (possibly considered urban by the USCB) with fewer than 20,000 people.[10]

More populated small towns (e.g., a community of 16,500) might challenge traditional notions of what constitutes "rural." However, these communities are included for a practical reason. All communities with populations less than 20,000—rural places *and* small towns alike—do not have access to single-year annual ACS data (e.g., 2017) for data-driven decision-making. Instead, they must rely on a 5-year combined ACS data (e.g., 2013–2017). In the end, the 20,000 population threshold is an appropriate cutoff for revealing the challenges that both rural and small-town communities face: navigating the statistical uncertainty of ACS data for data-driven policymaking.

To assess the statistical reliability of ACS estimates across rural America, I selected a state from each of the USCB-defined regions (i.e., Northeast, South, Midwest, and West) (see Appendix). An upside to this approach is that it also makes it possible to assess the degree to which, if any, statistical reliability of ACS data varies across rural America by census region.

More than one in four Americans (28.8%) lived in a rural area or urban cluster in 2010 (Table 7.3). In the Midwest and South, more than one-third of residents, 37.2% and 33.8%, respectively, lived in a rural area or urban cluster in 2010. Across the West and Northeast, just 19.4% and 20.3% of residents, respectively, lived in rural areas or urban clusters. After examining data for all 50 states, I selected New Hampshire, North Carolina, Oregon, and Wisconsin as representative states for their respective census regions (Table 7.3). The population residing in rural and urban clusters in these states was slightly above each region's median.

In this analysis I rely on median household income (MHI) to assess the degree to which ACS estimates are statistically reliable. The decision is based on two key factors. First, MHI is used by most state and federal infrastructure grant/loan managers more for determining program eligibility. Second, MHI applies to all households and is not specific to a particular subpopulation. This is important because cross-tabulated data (e.g., child poverty) are drawn from a smaller population (i.e., children), resulting in higher MOE values. Therefore, although the challenges of statistical uncertainty of ACS data are more severe for cross-tabulated data, using MHI most closely illustrates the statistical challenges that rural and small-town communities must confront to secure infrastructure grant/loan funding.

---

[9]According to Census 2010 data, the population of Lancaster, WI was 3868. Roughly 94% (3642 persons) of the population was classified as urban and the remaining 6% (226 persons) was classified as rural (US Census 2010a).

[10]The mean and median CV values calculated for places with fewer than 20,000 residents include both incorporated towns and cities (e.g., Drain, OR), as well as for census-designated places (CDPs) (e.g., Glide, OR). This analysis excludes instances where MHI estimates are unavailable for a place/CDP.

Table 7.3 Urban-rural population classification for selected states and USA, 2010

	Total population	Urban population	Urban population, urbanized areas (UA)	Urban population, urban clusters (UC)	Rural population	Population share, rural and urban cluster (UC)
New Hampshire	1,316,470	793,872	623,168	170,704	522,598	52.7%
North Carolina	9,535,483	6,301,756	5,232,799	1,068,957	3,233,727	45.1%
Oregon	3,831,074	3,104,382	2,393,393	710,989	726,692	37.5%
Wisconsin	5,686,986	3,989,638	3,173,382	816,256	1,697,348	44.2%
USA	308,745,538	249,253,271	219,922,123	29,331,148	59,492,267	28.8%

Source: US Census Bureau (2010a)

7.5 Statistical Uncertainty in Rural America

The level of statistical uncertainty for MHI estimates among communities with less than 20,000 people is shown in Table 7.4. Expressed through mean and median descriptive statistics, the data illustrate three important points. First, MHI estimates became more reliable for rural and small-town communities during the two periods, 2006–2010 and 2013–2017 (i.e., median CV declined from 13.6 to 12.4). This is principally due to the change in ACS sampling strategy, which effectively "borrowed strength" from more populated urbanized areas to improve statistical reliability for less populated areas. Second, the increased reliability in MHI estimates for rural and small-town communities varied by state. Across rural and small towns in New Hampshire, for example, the reliability of MHI estimates improved the most (i.e., median CV from 18.0 to 13.7), while reliability remained largely constant for Wisconsin communities (i.e., median CV from 9.7 to 9.6). Third, although MHI estimates have improved over the past decade, the median level of statistical uncertainty for the most current MHI estimates remains "somewhat reliable" for rural and small-town communities in New Hampshire, North Carolina, and Oregon, as well as the USA at large.

The statistical uncertainty of ACS estimates varies considerably between rural and small-town communities and more densely populated urban areas. To illustrate this difference, I compared MHI estimates and the accompanying MOE and CV statistics for three different sized communities: (1) small towns (around 1000 people), (2) urban clusters (around 20,000 people), and (3) urbanized areas (around 100,000 people). As Table 7.5 shows, MHI estimates for urban clusters and urbanized areas are statistically reliable across the states analyzed. Here, CV values range from 1.6 to 7.5. MHI estimates for small towns, on the other hand, are considerably less reliable. Consider, for example, the CV values for two communities in Oregon:

Table 7.4 Coefficient of variation (CV) for median household income (MHI), places with population less than 20,000

		2006–2010	2013–2017
New Hampshire	Mean	20.8	18.1
	Median	18.0	13.7
North Carolina	Mean	22.3	16.7
	Median	14.9	13.6
Oregon	Mean	22.0	16.9
	Median	12.8	12.3
Wisconsin	Mean	13.9	12.5
	Median	9.7	9.6
USA	Mean	22.3	15.8
	Median	13.6	12.4

Source: US Census Bureau (2017 and 2010c), American Community Survey (ACS), 2006–2010 and 2013–2017 (5-year combined estimates). Calculations by author

Table 7.5 Median household income (MHI) estimates with corresponding MOE and CV, 2013–2017

	Estimate	MOE	CV
New Hampshire			
Bethlehem (943)	$53,542	±$10,301	11.7
Portsmouth (21,644)	$72,384	±$4716	4.0
Manchester (110,601)	$56,467	±$1605	1.7
North Carolina			
Aulander (962)	$26,731	±$8033	18.3
Havelock (20,404)	$49,604	±$3012	3.7
High Point (109,849)	$44,642	±$1337	1.8
Oregon			
Drain (931)	$39,583	±$9631	14.8
Roseburg (22,013)	$42,507	±$5258	7.5
Eugene (163,135)	$47,489	±$1357	1.7
Wisconsin			
Amherst (1090)	$45,658	±$7604	10.1
Germantown (19,956)	$79,553	±$6970	5.3
Green Bay (104,796)	$45,473	±$1218	1.6

Source: US Census Bureau (2017), American Community Survey (ACS), 2013–2017 (five-year combined estimates)

Note: MOE values that accompany the population estimates are not reported in Table 7.5. Calculations by author

1.7 for Eugene, OR, and 14.8 for Drain, OR. What are the on-the-ground implications of this statistical difference? In Drain, community officials, planners, and business leaders have to navigate a much larger degree of statistical uncertainty in their MHI estimate. When the MOE is considered, the MHI estimate for Drain ranges from $29,952 ($39,583 − $9631) to $49,214 ($39,583 + $9631). These data underscore the statistical precariousness that many small towns face.

The challenges of statistically uncertain MHI estimates are apparent. Equally important is knowing: (1) to what extent have MHI estimates become more reliable for rural and small towns? and (2) how many rural and small-town communities continue to struggle with unreliable MHI estimates?

First, due to the change in the ACS sampling strategy between the two ACS periods, statistical reliability for rural communities and small towns has improved. Table 7.6 illustrates that MHI estimates have generally become more reliable and the number of communities with unreliable MHI estimates has dropped considerably. In the 2006–2010 period, for example, 11,415 rural and small-town communities had

Table 7.6 CV reliability for median household income (MHI), places <20,000

	2006–2010		2013–2017		Change between periods	
	Number	Percent	Number	Percent	Numeric	Percent
New Hampshire						
Reliable (0–12%)	33	37.5%	38	45.2%	5	15.2%
Somewhat reliable (12–40%)	46	52.3%	37	44.1%	(9)	−19.6%
Unreliable (40%+)	9	10.2%	9	10.7%	0	0.0%
North Carolina						
Reliable (0–12%)	252	36.4%	273	41.1%	21	8.3%
Somewhat reliable (12–40%)	362	52.3%	361	54.3%	(1)	−0.3%
Unreliable (40%+)	78	11.3%	30	4.6%	(48)	−61.5%
Oregon						
Reliable (0–12%)	158	46.9%	153	47.2%	(5)	−3.2%
Somewhat reliable (12–40%)	137	40.6%	143	44.2%	6	4.4%
Unreliable (40%+)	42	12.5%	28	8.6%	(14)	−33.3%
Wisconsin						
Reliable (0–12%)	460	62.7%	464	63.4%	4	0.9%
Somewhat reliable (12–40%)	241	32.8%	247	33.7%	6	2.5%
Unreliable (40%+)	33	4.5%	21	2.9%	(12)	−36.4%
USA						
Reliable (0–12%)	11,415	43.0%	12,068	48.2%	653	5.7%
Somewhat reliable (12–40%)	12,191	46.0%	11,715	46.7%	(476)	−3.9%
Unreliable (40%+)	2923	11.0%	1282	5.1%	(1641)	−56.1%

Source: US Census Bureau (2017 and 2010c), American Community Survey (ACS), 2006–2010 and 2013–2017 (five-year combined estimates). Calculations by author

reliable MHI estimates, compared to 12,068 in the 2013–2017 period. This represents a 5.7% increase in reliable MHI estimates between the two periods. Also, where almost 3,000 rural and small towns had unreliable MHI estimates during the 2006–2010 period, the number dropped by more than half (1282) in the most recent 2013–2017 ACS (Table 7.6). However, the increase in statistical reliability was uneven across the states analyzed. For rural communities and small towns in Oregon and Wisconsin, there was essentially no change in the number of statistically reliable MHI estimates. The largest increases in reliable MHI estimates were among rural and small towns within North Carolina and New Hampshire, increasing by 8.3% and 15.2%, respectively.

Second, despite the improvements in statistical reliability, many rural communities and small towns still suffer from statistically unreliable MHI estimates. More than half (12,997 or 51.8%) of rural and small-town communities across the USA have either somewhat reliable or unreliable MHI estimates in the 2013–2017 ACS (Table 7.6). A greater share of rural and small-town communities in New Hampshire (54.8%), North Carolina (58.9%), and Oregon (52.8%) have somewhat reliable or unreliable MHI estimates. The number of communities across rural America struggling with this issue is significant. Without changes to program eligibility criteria (e.g., considering MOE), changes in sampling strategy, and/or increasing the ACS sample size—pending available government funding—it is essential that states consider policies to assist rural communities and small towns with statistically unreliable MHI estimates.

7.6 Support for Helping Rural Communities with Unreliable ACS Estimates

In Oregon, the state's economic development agency, Business Oregon (through the Infrastructure Finance Authority [IFA]),[11] forges strategic partnerships and makes funding available to rural communities interested in conducting an income census/survey. Oregon IFA, like other state economic development agencies, covers some costs of conducting income censuses/surveys.[12] A February 2015 (p. 7) article in the Oregon League of Cities (LOC) magazine[13] recommends the following approach to unreliable ACS estimates in rural communities in Oregon:

> . . .communities that believe the ACS contains incorrect data and will be adversely impacted should contact their IFA regional coordinator, who can provide guidance and instruction on available options, which may include conducting a local survey. As noted previously, the survey has to be conducted according to the strict requirements of the specific federal agency.

[11]The chief aim of the Oregon IFA is to help Oregon communities apply, receive, and manage federal and state loan/grant funds for water, sewer, roads, and other infrastructure development.

[12]In the recent past, Oregon IFA covered up to $7500 of costs for a census enumeration for cities with a population less than 500 and 50% of costs up to $5000 for a survey with cities with a population of 500 or more. Currently (through June 2019), Oregon IFA covers up to $1000 in costs for conducting a survey.

[13]This link also contains information for conducting an income census/survey (for communities in the Great Lakes RCAP region): http://greatlakesrcap.org/uploads/PDF/Winter2014RCAPConnectionFINAL.pdf

7.7 Summary and Recommendations

State and federal infrastructure grant/loan programs are essential lifelines for improving water, sewerage, and transportation infrastructure for communities across rural America. Eligibility for these programs is often determined according to income estimates from the ACS. The problem, especially for rural communities, is that the statistical uncertainty of ACS data is higher for rural communities, and MOE values are often disregarded for determining program eligibility. This chapter contextualizes the breadth of this problem among small town and rural communities. In the 2013–2017 period, more than half (12,997 or 51.8%) of rural and small-town communities across the USA have either somewhat reliable or unreliable income estimates. Equally important, the reliability of income estimates varies by state. In New Hampshire, North Carolina, and Oregon, for example, a greater share, 54.8%, 58.9%, and 52.8%, respectively, of rural and small-town communities have somewhat reliable or unreliable income estimates.

Communities can hire a consultant to conduct an income census or survey to formally challenge the USCB's ACS estimate. Many rural communities, however, elect not to pursue this approach. Often, they are either unaware that some grant/loan programs reimburse costs for conducting an income census/survey or are dissuaded by the necessary time and resources.

What specific policies can states implement to assist rural communities with unreliable estimates and secure infrastructure funding that advances economically healthy, resilient, and sustainable communities? First, a coordinating agency, presumably the state's economic development agency, should reach out to key partners engaged in rural economic funding development issues (e.g., water, sewerage, transportation, and housing infrastructure) at state, regional, and local levels. Consideration should also be given to partnering with national organizations like the Rural Community Assistance Partnership (RCAP). A 501(c)(3) nonprofit organization, RCAP has contract agreements with the US Departments of Health and Human Services (HHS), USDA Rural Development, and the Environmental Protection Agency (EPA). Another potential national partner is the USCB's State Data Center (SDC) program, which is working collaboratively with the Bureau to assess the viability of Rural Statistical Area (RSA) geographies. This approach minimizes any duplication of efforts by identifying potential partners—both technical and nontechnical—and improving overall efficiency. Second, universities should be considered as census/survey partners. Universities, especially land-grant institutions with rural extension programs or urban-serving institutions working statewide, often have locally specific knowledge from community-engaged scholarship. Having universities involved also provides experiential learning opportunities for students. These learning opportunities may be directly tied to conducting censuses/surveys and may also result from additional community research needs (e.g., housing, transportation, health, and workforce training).

Third, states should adopt guidelines for prioritizing how to best allocate state funds to communities requesting an income census/survey. This is a critically important step because statistical uncertainty of income estimates varies according

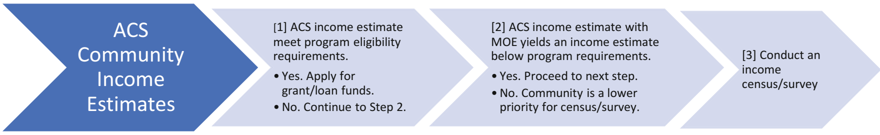


Fig. 7.2 Recommended steps for prioritizing census/survey need

to community size. I recommend states to prioritize community need for an income census/survey according to the approach outlined in Fig. 7.2.

The steps outlined in Fig. 7.2 are as follows:

1. If a community meets program eligibility requirements, then proceed by applying for grant/loan infrastructure funds. If the community does not meet program eligibility requirements, proceed to Step 2.
2. Determine if the ACS income estimate (along with the corresponding MOE) yields an income estimate below the program income threshold.

   A. If a community's ACS income estimate, along with the corresponding MOE, yields an income value below the program income requirement, it means that the "actual" value could qualify the community for grant/loan funding. Proceed to Step 3.
   B. If a community's ACS income estimate, along with the corresponding MOE, *does not* yield an income value below the program income requirement, it means that the "actual" value, more than likely, would *not* qualify the community for grant/loan funding. These communities should be lower priorities for receiving funding for conducting a survey/census.

3. Conduct an income survey/census for communities identified in 2(A).

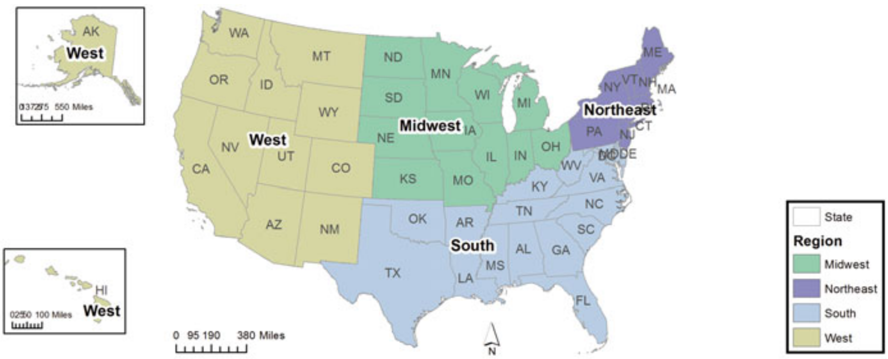
To explain how this proposal would work, consider a grant/loan program with an income estimate threshold of $35,000. The towns of Admiral and Birdtown each have income estimates of $36,500, but the MOE for Admiral and Birdtown is ±$1200 and ±$3500, respectively. Under this scenario, neither community meets the eligibility requirements for the grant/loan program because MOE is ignored (Step 1).

For Admiral, the income estimate and corresponding MOE *does not* meet the program income requirement (the lower-bound estimate is $35,300, based on $36,500 − $1200) (i.e., Step 2[B]). Also, a test for statistical significance indicates that it is *unlikely* that the "true" income value for Admiral is below the $35,000 threshold (comparing $35,000 and $36,500 [±$1200] yields a statistically significant *t*-test value of 2.0563 at 95%). Conversely, Birdtown's income estimate and corresponding MOE *meets* the program income requirement (the lower-bound estimate is $33,000, based on $36,500 − $3500) (i.e., Step 2[A]). And because the estimate for Birdtown contains greater statistical uncertainty (i.e., higher MOE), there is a greater chance–relative to Admiral–that the "true" income value is below the $35,000 threshold (comparing $35,000 and $36,500 (±$3500) yields a non-statistically significant *t*-test value of 0.7050 at 90, 95, and 99%[14]). In the end, Birdtown is an excellent candidate for receiving funds to conduct an income census/survey. Using these statistical tests can help jurisdictions prioritize the likelihood for whether a community has a "true" income value above/below grant/loan income requirements.

Changes to program eligibility criteria—either explicitly considering MOE or increasing the ACS sample size, for example—would be immediate and effective policy remedies to help rural communities and small towns with unreliable ACS estimates. In reality, however, implementing these changes will take time and resources. States, in the meantime, should invest in policies that efficiently and equitably distribute resources for conducting income surveys/censuses. This approach ensures that rural communities secure much-needed capital projects that advances economic development and quality of life for all citizens and, in particular, for the most vulnerable.

Acknowledgments This research was partially funded with support from Oregon State University (OSU) and Portland State University (PSU). Special thanks to the following individuals, who provided valuable insight and constructive feedback on this manuscript: Paul Lask, Beth Emshoff, Nick Chun, Lena Etuk, Mallory Rahe, Charles Rynerson, Jeff Sherman, David Tetrick, Kevin Tracy, and Bruce Weber. Thanks also to Business Oregon. All errors and oversights are solely the responsibility of the author.

Appendix: Census Regions



Source: US Census Bureau

[14]This example uses 90% as the cutoff for statistical significance based on prevailing practice in social science research, but determining the level of statistical significance (i.e., how much uncertainty one is willing to tolerate) is somewhat arbitrary, and ultimately up to each state.

References

Alvarez JA, Salvo J (2014) Navigating reliability of small area data, lessons from New York City. Presented at American Community Survey (ACS) data user group conference, May 29–30, Washington, DC

Asiala, ME (2012) American Community Survey sampling and weighting. Presented at measuring people in place conference, Colorado University, Boulder. 5 October 2012. https://cupc.colorado.edu/workshops/measuring_people_in_place/themes/theme1/asiala.pdf

Carpenter R, Reamer A (2010) Surveying for dollars: the role of the American Community Survey in the geographic distribution of federal funds. Brookings Institution, Washington, DC. https://www.brookings.edu/research/surveying-for-dollars-the-role-of-the-american-community-survey-in-the-geographic-distribution-of-federal-funds/

Caster P (2017) House committee passes Herrera-Beutler bill on flawed income data. http://www.chronline.com/news/house-committee-passes-herrera-beutler-bill-on-flawed-income-data/article_e8c3a330-6d6f-11e7-848b-8f4c80959193.html. Accessed 14 Feb 2019

Citro C, Kalton G (eds) (2007) Using the American Community Survey: benefits and challenges. National Academies Press, Washington, DC. http://www.nap.edu/read/11901/chapter/1#ii

Cromartie J, Swanson L (1996) Defining metropolitan areas and the urban-rural continuum: a comparison of statistical areas based on county and sub-county geography. Staff paper no. AGES9603. Economic Research Service (ERS), U.S. Department of Agriculture, Washington DC

Environmental Systems Research Institute (ESRI) (2014) The American Community Survey. https://www.esri.com/library/whitepapers/pdfs/the-american-community-survey.pdf. Accessed 21 Nov 2018

Heuvelink GBM, Burrough PA (2002) Developments in statistical approaches to spatial uncertainty and its propagation. Int J Geogr Inf Sci 16(2):111–113

Isserman AM (2005) In the national interest: defining rural and urban correctly in research and public policy. Int Reg Sci Rev 28(4):465–499

Jurjevich JR, Griffin AG, Spielman SE, Folch DC, Merrick M, Nagle NN (2018) Navigating statistical uncertainty: how urban and regional planners understand and work with American Community Survey (ACS) data for guiding policy. J Am Plan Assoc 84(2):112–126

McGranahan DA, Hession JC, Hines FK, Jordan MF (1986) Social and economic characteristics of the population in metro and nonmetro counties. 1 970–80 RDR-58. Economic Research Service. U.S. Department of Agriculture, Washington, DC

National Academy of Sciences (2015) Realizing the potential of the American Community Survey: challenges, tradeoffs, and opportunities. National Academies Press, Washington DC

Navarro F (2012) An introduction to ACS statistical methods and lessons learned. Presented at measuring people in place conference. Colorado University, Boulder. 5 October 2012. http://www.colorado.edu/ibs/cupc/workshops/measuring_people_in_place/themes/theme1/navarro.pdf

Nesse K, Rahe M (2015) Conflicts in the use of ACS by federal agencies between statutory requirements and survey methodology. Popul Res Policy Rev 34(4):361–480. https://doi.org/10.1007/s11113-014-9352-7

Oregon League of Cities (LOC) (2015) IFA Offers Guidance on American Community Survey. Available at: http://www.orcities.org/Portals/17/Publications/localfocus/Feb2015web.pdf. Accessed 4 Jan 2019

Ratcliffe M, Burd C, Holder K, Fields A (2016) Defining rural at the U.S. Census Bureau. ACSGEO-1. U.S. Census Bureau, Washington, DC

Spielman SE, Folch DC (2015) Reducing uncertainty in the American Community Survey through data-driven regionalization. PLoS One 10:e0115626. https://doi.org/10.1371/journal.pone.0115626

Spielman SE, Folch DC, Nagle NN (2014) Patterns and causes of uncertainty in the American Community Survey (ACS). Appl Geogr 46:47–157

Torrieri NK (2007) America is changing, and so is the census. Am Stat 61(1):16–21
U.S. Census Bureau (2009) A compass for understanding and using American Community Survey data: what state and local governments need to know. US Government Printing Office, Washington, DC
U.S. Census Bureau (2010a) Urban and rural (P2): Census 2010, summary file 1 (SF1), 100 percent data
U.S. Census Bureau (2010b) Rural-urban classification criteria. Available at: https://www.census.gov/geo/reference/ua/urban-rural-2010.html
U.S. Census Bureau (2010c) American Community Survey (ACS), 2006–2010 combined 5-Year estimates, B19013: median household income (dollars). Generated using American FactFinder. Accessed 14 Sept 2012
U.S. Census Bureau (2017) American Community Survey (ACS), 2013–2017 combined 5-Year estimates, B19013: median household income (dollars). Generated using American FactFinder. Accessed 3 Mar 2019)
U.S. Department of Agriculture (2012) Administrative Notice (AN) 4679. Implementation of 2010 Census Data for Rural Development Programs. http://www.ruralhome.org/storage/documents/rd_obligations/rd_docs/an4679.pdf. Accessed 30 Nov 2019
US Census Bureau (2011) American Community Survey: accuracy of the data. https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/ACS_Accuracy_of_Data_2011.pdf. Accessed 5 Apr 2019
US Census Bureau (2019) History: questionnaires. https://www.census.gov/history/www/through_the_decades/questionnaires/. Accessed 3 Mar 2019
Usowski K, Rodgers L, Lihn M, Kahn P (2008) Using American Community Survey data in HUD's income limits and fair market rents. Cityscape J Policy Dev Res 10(1):199–210