Published in final edited form as:

Appl Geogr. 2014 January ; 46: 147–157. doi:10.1016/j.apgeog.2013.11.002.

Patterns and causes of uncertainty in the American Community Survey

Seth E. Spielman^{a,b}, David Folch^b, and Nicholas Nagle^c

^aGeography Department, University of Colorado, 110 Guggenheim Hall, Box 260 UCB, Boulder, CO 80309, USA

^bInstitute of Behavioral Science, University of Colorado, 110 Guggenheim Hall, Box 260 UCB, Boulder, CO 80309, USA

^cDepartment of Geography, University of Tennessee, USA

Abstract

In 2010 the American Community Survey (ACS) replaced the long form of the United States decennial census. The ACS is now the principal source of high-resolution geographic information about the U.S. population. The margins of error on ACS census tract-level data are on average 75 percent larger than those of the corresponding 2000 long-form estimate. The practical implications of this increase is that data are sometimes so imprecise that they are difficult to use. This paper explains why the ACS tract and block group estimates have large margins of error. Statistical concepts are explained in plain English. ACS margins of error are attributed to specific methodological decisions made by the Census Bureau. These decisions are best seen as compromises that attempt to balance financial constraints against concerns about data quality, timeliness, and geographic precision. In addition, demographic and geographic patterns in ACS data quality are identified. These patterns are associated with demographic composition of census tracts. Understanding the fundamental causes of uncertainty in the survey suggests a number of geographic strategies for improving the usability and quality ACS.

Keywords

American Community Survey; Census Tract; Small Area Estimation; Margin of Error; Uncertainty

Introduction

In 2010 the American Community Survey (ACS) replaced the long form of the United States decennial census as the principal source of high-resolution geographic information about the U.S. population. The ACS fundamentally changed the way data about American communities are collected and produced. The long form of the decennial census was a large-sample, low-frequency national survey; the ACS is a high-frequency survey, constantly measuring the American population using small monthly samples. One of the primary challenges for users of the ACS is that the margins of error are on average 75 percent larger

than those of the corresponding 2000 long-form estimate (Alexander 2002; Starsinic 2005). The practical implications of this increase are that users often face data like those in Table 1, which shows the ACS median income estimates for African American households for a contiguous group of census tracts in Denver, Colorado. Income estimates range from around \$21,000 to \$60,000 (American Factfinder website accessed 7/15/2013). Without taking account of the margin of error, it would seem that Tract 41.06 had the highest income, however, when one accounts for the margin of error, the situation is much less clear – Tract 41.06 may be either the wealthiest or the poorest tract in the group.

Some degree of uncertainty is inherent in surveys like the ACS, however the amount of uncertainty in the ACS has far exceeded the United States Census Bureau's (USCB hereinafter) expectations. Initial expectations were that the amount of uncertainty (margin of error) in the ACS would be 33 percent greater than the decennial census long form (Navarro 2012). This loss in precision was justified by the increase in timeliness of ACS estimates which are released annually compared to the once a decade long form. This tradeoff prompted Macdonald (2006) to call the ACS a “warm” (current) but “fuzzy” (imprecise) source of data. Unfortunately, those early expectations were too optimistic, the actual uncertainty in the ACS is much more than 33 percent greater than the census long form. While there are clear advantages to working with fresh data, the ACS margins of error are so large that for many variables at the census tract and block group scales the estimates fail to meet even the loosest standards of data quality.

The ACS is the primary national source of geographically and demographically detailed information about the American population. It is an essential resource for cartographers, geographers, or anyone interested in understanding neighborhood scale social and economic patterns. Savvy use of the ACS requires understanding the nature of the uncertainty in the ACS and its causes- the purpose of this paper is to provide users with a readable account of the causes of uncertainty in the ACS and to suggest some potential solutions to the problem(s). This paper assumes readers have no prior background in survey statistics, however a basic familiarity with survey methods is necessary to understand uncertainty in the ACS, so this paper explains basic survey-statistical concepts in plain English (section 1.1). Section 2, “The construction of the ACS,” provides a high-level overview of the methods underlying the ACS and discusses how these methods contribute to uncertainty in ACS data. Section 3 discusses how factors beyond the USCB's control contribute to uncertainty in ACS data. Finally, some potential solutions to the ACS's problems will be offered in Section 4. It is important to note that the causes of uncertainty in the ACS are not entirely clear, a full enumeration of the causes of uncertainty in the ACS would run many hundreds of pages. While not complete picture of the causes and nature and nature of uncertainty in the ACS this article provides a detailed framework and vocabulary for understanding uncertainty and suggests geographic strategies for dealing with it.

1.1 Uncertainty in Surveys

Like the decennial long form before it, the ACS is a sample survey. Unlike complete enumerations¹, sample surveys do not perfectly measure the characteristics of the population —two samples from the same population will yield different estimates. This sample-to-

sample variability creates some uncertainty about a population's true characteristics, therefore survey-based estimates are usually accompanied by a margin of error. While it was not commonly acknowledged, even the decennial census long form data came with instructions for estimating margins of error. In the ACS, the margin of error for a given variable expresses a range of values around the estimate within which the true value is expected to lie. The margin of error reflects the variability that could be expected if the survey were repeated with a different random sample of the same population. This variability is referred to as *sampling error* and is measured as *standard error* (SE). Calculating standard errors for a complex survey like the ACS is not a trivial task, the USCB uses a simulation procedure called Successive Differences Replication to produce variance estimates (Wolter, 1984; Fay & Train, 1995; Judkins, 1990). The margins of error reported by the USCB with the ACS estimates are simply 1.645 times the standard errors.

Sampling error has two main causes. The first is the sample size - the larger the sample the smaller the standard error, intuitively more data about a population leads to less uncertainty about its true characteristics. The second main cause of sampling error is heterogeneity in the population being measured (Rao 2003). Consider two jars of U.S. coins, one contains U.S. pennies and the other contains a variety of coins from all over the world. If one randomly selected 5 coins from each jar, and used the average of these 5 to estimate the average value of the coins in each jar, then there would be more uncertainty about the average value in the jar that contained a diverse mixture of coins. If one took repeated random samples of 5 coins from each jar the result would always be the same for the jar of pennies but it would vary substantially in the diverse jar, this variation would create uncertainty about the true average value². While the ACS is much more complicated than pulling coins from a jar, this analogy helps to understand the standard error of ACS estimates. Census Tracts and block groups are like jars of coins. If a tract is like the jar of pennies, then the estimates will be more precise, whereas if a tract is like the jar of diverse coins, then the estimate will be less precise.

While the simple example is illustrative of important concepts it overlooks the central challenge in conducting surveys; many people who will be included in a sample will choose not to respond to the survey. While a group's odds of being included in the ACS sample are proportional to its population size, different groups of people have different probabilities of responding. Only 65% of the people contacted by the ACS actually complete the survey (in 2011, 2.13 million responses were collected from 3.27 million samples). Some groups are more likely to respond than others, this means that a response collected from a hard to count group is worth more than a response from an easy to count group. These differential response rates are controlled by *weighting* each response. In the ACS each completed survey is assigned a single weight through a complex procedure involving dozens of steps. The

¹A complete enumeration of large and diverse population may not be possible. The Decennial census while a complete enumeration in principle, has always under/over counted particular populations (Freedman 1991).

²The USCB generally is not actually estimating the "average" value, they are estimating the "total" value of coins in the jar. Repeatedly grabbing five coins and computing the average will over many samples get you a very precise estimate of the average value, but it will give you no information on the total value. To get the total value, you need a good estimate of the average AND a good estimate of the total number of coins in the jar. The loss of contemporaneous population controls caused by decoupling the ACS from the Decennial enumeration means that the census does not have information about the number of coins in the jar. This is discussed in section 2.

important point, as far as this paper is concerned, is that these weights are *estimated* and uncertainty about the appropriate weight to give each response is an important source of uncertainty in the published data.

The concepts of sampling error and weighting are central to understanding uncertainty in the ACS but they are not the entire story. Some of the factors affecting uncertainty in the ACS are the result of decisions and tradeoffs made by the USCB, whereas other factors affecting uncertainty are the result of circumstances beyond the control of the USCB. Section 2 provides a high-level overview of the methods used to construct the ACS and discusses how these methodological choices affect uncertainty and Section 3 discusses how factors beyond the control of the USCB further shape the quality of the survey.

2. The construction of the ACS

As the primary source for high-resolution, socio-spatial information throughout the U.S., the ACS has a profound impact on research and practice across the social sciences. It is extremely difficult to produce timely, detailed data for small geographic areas in a country as large and diverse as the United States. The data-quality problems in the ACS are the direct result of decades of innovation in national surveys. Four innovations in particular are combined in the ACS; the use of sampling, the provision of small-area estimates, the release of annual estimates, and the use of weighting to adjust the importance of individual responses. These innovations make sense individually but become problematic in combination. Understanding the motivation for each of these innovations and how they interact is essential to understanding the problems with ACS tract-level estimates. This section discusses how specific methodological decisions made by the USCB contribute to uncertainty in the ACS.

2.1 Sampling

Before 1940, there was no “long form” or “short form” of the U.S. decennial census; each housing unit (HU) received the same questionnaire. By 1940 the census forms had become long and complicated, through the gradual accrual of demographic and economic questions. In response, the questionnaire was split in 1940 into a short set of questions asked of 100 percent of the population and an additional “long form” administered to a subset of the population. Originally, this long form was administered to a 5 percent random sample, but in later years it was sent to one HU in six (Andersen et al. 2011). Before 1940 any error in the data could be attributed either to missing or double counting an HU, to incorrect transcription of a respondent’s answer, or to intentional/unintentional errors by the respondent. After 1940, however, the adoption of statistical sampling introduced new sources of uncertainty for those questions on the long form. The bifurcation into short and long forms continued through the 2000 census, but in 2010 the decennial census returned to a single, short questionnaire. The detailed demographic, economic, and housing data previously collected on the long form are now provided by the ACS.

There were some important advantages to administering the long and short forms of the census simultaneously. In particular, the short form population counts could serve as controls for long-form based estimates. The decoupling of the sample from the complete

enumeration is an important source of uncertainty in the ACS, accounting for 15 to 25 percent of the difference in margin of error between the ACS and the decennial long form (Navarro 2012). Population controls are essential to the ACS sample weighting process. Prior to the ACS population controls for the long form sample weights were available from the short form, now population controls are only available for relatively large geographic areas such as municipalities and counties.

However, there were also disadvantages to administering the long and short forms simultaneously. Because the long form data were available only once every decade and often took two years to tabulate, in rapidly changing areas the data were out of date before they were publically released (Citro and Kalton 2007). The decennial census is constitutionally mandated to count every person in the U.S., this complete enumeration philosophy bled over into the sample-based long form. Enormous expense was incurred to ensure that every house sampled for the long form is surveyed. In the ACS this philosophy changed, only a fraction of non-responders are tracked down by the USCB. The USCB estimates that the decision not to follow up with all sample households contributes 25 to 28 percent of the increase in ACS margins of error relative to the long form (Navarro 2012).

In theory the ACS sample of 3.54 million households is drawn from a list of all American households. In practice, however, no such universal list of households exists. The ACS is a sample from a continuously updated list of housing units and group quarters called the Master Address File (MAF). The MAF is more than a mere list, it includes detailed, field verified, geographic information about the location of housing unit addresses. The MAF is the most comprehensive list of U.S. addresses in existence, but it is susceptible to both undercoverage and overcoverage. The USCB has estimated a total undercoverage of 6.4 percent, of which 4.7 percent are omissions, that is, addresses not present in the MAF, and 1.7 percent are erroneous exclusions, that is, addresses in the MAF but not used by the ACS because of technical errors (Bates and Hartman 2012). Information on the coverage rates in the MAF by socio-economic group is not available. Overcoverage, or a single HU occurring more than once in the database, is estimated to be 10.2 percent (Bates and Hartman 2012). It is important to note that the ACS is a sample of housing units not people, each housing unit completes a single survey; this is why figure 1 shows the number of housing unit surveys returned, and this is also why the MAF is so important for the ACS. One benefit of the housing unit approach is that the ACS can directly measure vacancy, a housing unit is vacant if it does not contain a household. While many local, state, and federal data sources contribute to the MAF, the primary source is the U.S. Postal Service. The MAF is also continuously updated from the field reports of USCB enumerators. Maintaining the MAF is expensive. A 2009 report by the General Accounting Office raised concerns about the cost and quality of the MAF, noting that its \$444 million price tag (as of 2009) represented a 25 percent cost overrun (General Accounting Office 2009).

2.2 Small Area Estimates

The geographic resolution of ACS estimates also contributes to uncertainty. Tract-level census tabulations were first proposed by Walter Laidlaw, a Presbyterian minister, in 1906. The *New York Times* in 1923, clearly excited by the potential of Laidlaw's tract system,

noted that it allowed one to “know *precisely* and for the first time what is meant by New York.” While survey methodology has evolved substantially over the past century, the practice of reporting estimates for small geographic units has remained fundamentally unchanged. The definitions of some geographic units have changed, and new types of geographic units have been created, but the basic idea of tabulating results using relatively static geographic zones remains constant. The principle is that tract (and smaller) geographic units are designed using criteria that are largely exogenous to the detailed demographic data collected by the census. The ACS and the decennial census use the same system of blocks, block groups, and census tracts for reporting estimates. Hereafter we refer to census tracts and block groups collectively as “small areas.”

Prior to the advent of sampling, the complete count census data could, in principle, be tabulated using any sort of geographic zone. The advent of sampling made small area census data less precise. Since there are a finite number of samples in any geographic area, as tabulation zones become smaller sample sizes decline, making the ACS estimates more uncertain. The rise uncertainty is greater for small populations; for instance the effects of reducing a sample size from 200 to 100 is much greater than the effect of reducing a sample size from 20,000 to 10,000. The USCB counteracts this decline in sample size by pooling small area samples over multiple years, thus diluting the temporal resolution of the estimates. Larger areas do not require multiyear pooling. For large municipalities, counties and states the ACS is a distinct improvement over the decennial census because it provides high quality annual data, however this may not be true at the tract/block group scale where even the 5 year pooled data have much larger standard errors and smaller sample sizes than the 2000 long form data (figure 1).

The USCB attempts to manage the quality of the ACS small area data through trade-offs between geographic resolution and temporal resolution. The decision to maintain the system of census tracts and block groups used by the decennial census is one of the primary reasons for the data quality problems in the ACS. At current funding levels the sample size is simply insufficient for one to “know precisely” the characteristics of America’s census tracts. Substantial increases in the geographic density of samples would be necessary for high quality small area estimates. For example, in the 2007-2011 ACS, the average census tract had 135 completed surveys over the 5 year period (median 124 surveys, Figure 1), while the 2000 decennial long form had an average of 280 housing units (median 249). The USCB estimates that 25 percent of the increase in the margin of error is due to decrease in sample size (Navarro 2012). For some purposes a sample of 135 housing units per tract seems adequate, however one must remember that this sample is spread across all of the various groups for which the ACS provides estimates (age groups, racial groups, ethnic groups, income groups, housing type, education, etc.).

If the philosophy behind the geographic zones used to report ACS data were to change it might be possible to improve the usability of the ACS. For example, if an additional meso-level tabulation unit was created -- larger than a tract but smaller than a county³ -- equivalent to clusters of 3-5 tracts, it might have sufficient sample size to release precise 5 year

³This strategy is only feasible in counties that contain more than 1 census tract. Some rural counties in the US consist of a single tract.

estimates and/or annual estimates. Currently, small area tabulation units boundaries are largely exogenous to the collected data, tract boundaries are not “data driven.” Data-driven geographic units could be designed around statistical criteria such as the minimization of variance. Through such units the loss of information through aggregation could be minimized. However, such units would potentially be unstable over time and might create disclosure risks.

2.3 Annual Estimates

Small-area tabulations and the survey-based long form had become cornerstones of the decennial census by the late twentieth century. However, data users were increasingly concerned about the timeliness of the once-a-decade data (Alexander 2002). In 1985 Congress authorized, but never funded, a mid-decade census to address this problem. Throughout the 1980s and early 1990s interest in a “continuous measurement” model evolved within the USCB, leading to the proposal of the Intercensal Long Form (Alexander 1993). Continuous measurement was inspired by Leslie Kish, a statistician who developed the theory and methods for rolling surveys (Kish 1990).

Kish’s basic idea was that a population could be divided into a series of non-overlapping annual or monthly groups called subframes. Each subframe would then be enumerated or sampled on a rolling basis. If each subframe were carefully constructed so as to be representative of the larger population, then the annual estimates would also be representative, and eventually, the entire population would be sampled. The strength of this rolling framework is its efficient use of surveys. The decennial census long form had to sample at a rate appropriate to make reasonable estimates for small geographic areas such as census tracts, which contain on average 4,000 people. Therefore, citywide data released for a municipality of, say, 1 million people would be based on considerably more samples than necessary. Spreading the samples over time lets larger areas receive reasonable estimates annually, while smaller areas wait for more surveys to be collected. The rolling sample therefore increases the frequency of data on larger areas. The primary cost comes in the temporal blurring of data for smaller areas.

Rolling sampling is straightforward in the abstract. For example, suppose that there are $K=5$ annual subframes, that the population in a tract is known ($N=1000$), that the sampling rate is $r=1/6$, and that the response rate is 100 percent; then one would sample $n=N/(K*I/r)$ people per year. Over a 5 year period $1/6$ of the population would be sampled and each returned survey would represent $w=(N/n)/K$ people, where w is the weight used to scale survey responses up to a population estimate. In this simple case, the weight assigned to each survey would be the same. For any individual attribute y , the tract level estimate would be $y_t = \sum w_i y_i$ (equation 1), a weighted summation of all i surveys collected in tract t . If the weights are further adjusted by ancillary population controls X , then the variance of the estimate is $\sum w_i^2 \text{VAR}[y_i|X]$ (equation 2; Fuller 2011, assuming independence.). If the rolling sample consisting of long-form-type questions were administered simultaneously with a short form census, then all the parameters in our simple example (N , K , X) would be known. However, in the ACS good population controls are not available for small areas (N and X are unknown) because, unlike the long form, the survey is not contemporaneous with the

complete enumeration decennial census. Thus weights (w) for each response must be estimated and this is an important source of uncertainty in the ACS.

2.4 Weighting

In the ACS each completed survey is assigned a weight (w) that quantifies the number of persons in the total population that are represented by a sampled household/individual. For example, a survey completed by an Asian male earning \$45,000 per year and assigned a weight of 50 would in the final tract-level estimates represent 50 Asian men and \$2.25 million in aggregate income. The USCB's careful attention to survey weights ensures that the final tract-level estimates are as accurate as possible, i.e., unbiased. Variations in sampling rate, the lack of good estimates of tract-level population controls, and variations in response rate all necessitate a complex method to estimate w .

The construction of ACS weights is described in the ACS technical manual (which runs hundreds of pages, USCB 2009a). The complexity of the ACS weighting process is motivated by an effort to reduce bias, each step can be seen as an attempt to control some form of bias. For example, the temporal adjustments applied to weights are an attempt to control for bias that might arise due to seasonal fluctuations in population, "mode bias" adjustments are made to control for differences in surveys completed by mail, phone, internet, and in-person interview. Individually these steps make sense but they are so numerous and technically complex that in the aggregate they make the ACS estimation process nearly impenetrable for even the most sophisticated data users.

The cost of extensive tweaking of weights is more than just lack of transparency and complexity. Reducing bias by adjusting weights carries a cost. Any procedure that increases the variability in the survey weights also increases the uncertainty in tract-level estimates (Kish 2002). Embedded in this process is a trade-off between estimate accuracy (bias) and precision (variance/margin of error), refining the survey weights reduces bias in the ACS but it also leads to variance in the sample weights. Increasing the variation in the sample weights increases the margin of error of the final estimates (Alexander 2002; Kish 2002) (see equation 2). The weighting procedures applied by the USCB in the estimation of the ACS can then be seen as an implicit policy statement that unbiased (accurate) estimates are more important than precise (low-variance) estimates. While weighting procedures are not solely responsible for the low precision in ACS small-area estimates, the current bias-versus-variance calculation may need to be reevaluated. There are other census programs that publish small area estimates with lower variance but with higher bias, one such example is the Small Area Income and Poverty Estimates (SAIPE) program. It is not clear from the published material about the ACS how much each weighting step reduces bias and increases variance.

3. Exogenous sources of uncertainty in the ACS

The ACS in its current form represents a nuanced and sensitive set of compromises. With a limited budget the USCB is trying to provide timely demographically detailed data for over 200,000 block groups and over 70,000 census tracts. The USCB is fully aware of the implications of the specific methodological decisions about the survey. These

methodological decisions are neither right nor wrong, they are best seen as compromises. For example, reducing the geographic or temporal resolution of the ACS data in order to improve the precision of the estimates might upset a large portion of users – the current methodology represents a compromise between resolution and precision. However, there are also sources of uncertainty in the ACS that are exogenous to the survey itself. In this section we describe two of these, public support for the survey and the internal composition of census tabulation units.

3.1 Public support for the survey

Public support for the ACS can be understood as both direct financial support and the willingness of the public to respond to the survey. In 2011 the ACS' sampling budget increased and the number of households sampled each year rose from 2.9M to 3.5M -- a relatively small increase if one considers the number of small areas for which estimates are produced. Further increases in the ACS's budget would, under current methodology, improve the usability of ACS estimates by reducing the margins of error for small areas.

The response rates for the ACS are very low, considering participation is mandated by federal law. There is a large discrepancy between the published response rate of around 97%⁴ and the effective response rate of 65%. This discrepancy is rooted in non-response follow-up. If people do not respond to mail, telephone, or internet solicitations then they are placed into a pool for potential non-response follow-up by a USCB field worker. These in-person contacts are expensive and only a portion of the non-responders are contacted by the USCB. The low response rate can be interpreted as lack of public support. Increases in the response rate would improve ACS data substantially, for little or no additional cost. Recently, the ACS added an internet response option, early evidence suggests this may facilitate data collection and improve the response rates (Ruiter et al. 2012).

3.2 Spatial and Demographic Patterns in Data Quality

One critical aspect of using the ACS is that these trade-offs between spatial and temporal resolution and uncertainty are not uniform; some places and some populations have more precise data. For example, Figure 2 shows the distribution of the coefficient of variation (CV) of tract-level estimates of median income for African-American households. We have chosen this variable for the simple reason that it is an important indicator in urban and population geography in the U.S. Like the margin of error, the CV is a measure of uncertainty; it is simply the ratio of the standard error to the estimated value. The CV is a useful statistic because it expresses uncertainty as a relative percent of the estimate: a SE of \$10,000 would be a CV of 0.5 (50%) for a \$20,000 estimate and 0.1 (or 10%) for a \$100,000 income estimate. Each bar in Figure 2 represents five percent of census tracts after they have been sorted by African American median household income, such that the left-most bar represents the poorest five percent of tracts and the right most bar represents the wealthiest five percent of tracts. Within each bar, we can see the relative distribution of quality for these census tracts, where quality is measured by the CV. The ACS user's guide presents the CV as a way to assess an estimate's "fitness for use" (U.S. Census Bureau 2009b, A-3), but

⁴Response rates online at http://www.census.gov/acs/www/methodology/response_rates_data/ accessed 10/21/2013.

does not provide formal criteria for what constitutes “fit” data. The National Research Council (NRC) suggests that a CV of 10 to 12 percent (or less) is a “reasonable standard of precision” (Citro 2007, p. 64). Environmental Systems Research Institute (ESRI) (2011) states that a CV less than 12 percent means “high reliability,” a CV in the 12 to 40 percent range means “moderate reliability,” and a CV over 40 percent means “low reliability.” Figure 2 shows that the “reliability” of these ACS data is strongly dependent upon the economic conditions of the tract; very poor or very wealthy tracts yield lower quality (higher CV) estimates. For African American median household income, more than 75 percent of all census tracts in the United States fail to meet the NRC “reasonable” standard of precision. These estimates are especially bad for the poorest 15 percent of census tracts, where more than half fail to meet ESRI’s more liberal standard of “moderate reliability.”

The relationship between data quality and income in Figure 2 is troubling and difficult to attribute to a single factor. While the USCB does not explicitly over- or under-sample areas based on their socio-economic characteristics, they do allocate more ACS samples to areas with low response in order to equalize coverage rates. In the 2007-2011 ACS tract data there is a very weak, substantively negligible, correlation between sample size and the median household income of the tract ($r=0.02$, $p\text{-value} < 0.001$). This suggests that systematic variation in sample size does not explain the gradient in Figure 2, that is poorer and richer neighborhoods do not have appreciably different sample sizes. Overall coverage rates for the African American population are low in the ACS (89% in 2011, United States Census Bureau 2013a), this type of omission may bias the ACS but it does not directly increase uncertainty in small area estimates. Another potential explanation lies in the fact that some questions are not completed by the survey respondent, causing the USCB to “impute” the value based on a set of decision rules. The imputation rates for income variables are high, approaching 20%, however this rate is not available disaggregated by race or income (United States Census Bureau 2013b).

The pattern of association between income and data quality holds for the entire population, not just African-American households. Figure 3 shows CV of median household income for all census tracts in the 150 largest U.S. cities (defined using the USCB’s Metropolitan Statistical Area (MSA) boundaries). Income is scaled such that for each tract we calculate its income percentile relative to all tracts in the MSA. The leftmost bar in figure 3 represents the first percentile in each MSA, that is, the poorest 1 percent of tracts in each city, the dollar amount defining the first percentile category is city specific. This approach was taken because there is significant regional variation in income and without this correction the relationship between income and data quality is difficult to see. This type of correction was not necessary for figure 2. Figure 3 shows that poorer tracts, in the 150 largest cities, have lower quality estimates than wealthier tracts. The bars in figure 3 show the interquartile range for each percentile, the median is denoted with a white line.

This social pattern in data quality translates into geographic patterns as well. The CV for median household income, considering all tracts in the United States, has a Moran’s I of 0.22 ($p\text{-value} < 0.001$) meaning that tracts with a high (or low) CV tend to be surrounded by tracts with similar CVs. In particular, tracts in the center of cities have a higher CV for income than tracts in the suburbs (figure 4). Figure 4 shows the CV of median household

income estimates for all tracts in the 150 largest metropolitan areas in the US. All tracts in each city are ordered based upon their distance from the “center” of the city, where center is determined using the city’s coordinates as listed in the USGS Geographic Names Information System. The use of a single center for a large polycentric city may be problematic, but the figure is illustrative of patterns nonetheless. Like figure 3 the first bar represented the first percentile, in this case the first percentile (leftmost bar) contains the tracts closest to the city center. The rightmost bar contains the 1 percent of tracts in each MSA that are furthest from the center. Relative distances were used on the X axis to control for the significant variation in MSA extent. Figure 4 shows that data quality for median household income estimates varies systematically within urban areas.

One interesting hypothesis for the patterns in figures 2-4 is that samples of the same size yield estimates that vary in quality because of systematic variations in the demographic composition of census tracts. Keeping sample size constant, a sample taken from a diverse population will have more sampling error, and hence uncertainty, than one from a homogenous population (see section 1.1).

Figure 5 shows that there is an association between the median household income in a tract and the amount of diversity in household incomes; Figure 5 shows both the gini coefficient on household income and the median household income for each tract in the US from the 2007-2011 ACS. The gini coefficient is a measure of the equality of a distribution, if income were evenly distributed in the population such that all households had the same annual income the gini coefficient would equal 0 and the variance (or heterogeneity) in income would be zero. One the other hand if a single household earned all of the income in a tract the gini coefficient would equal 1 and the variance in income would be high. One of the key determinants of uncertainty in survey estimates is the amount of heterogeneity (variance) in the population (see section 1.1 for a discussion).

The pattern in figure 5 parallels the pattern in figure 2, low income and high income tracts have a higher gini coefficient and a higher variance in income. Middle income tracts, which have a lower gini coefficient have a substantially more even income distribution and a lower income variance. Sample size and estimation procedures are roughly constant across the range of median household incomes. It is possible that the pattern in figures 2-3 exists because low and high income neighborhoods have more variance in income than middle income neighborhoods—holding sample size constant more variance in the population means more uncertainty in the estimate. Additionally, this may hold for figure 4, tracts in the center of cities may have more income diversity than those in the suburbs. Income diversity gradients have long been a part of American urban life and there is evidence of increasing neighborhood-level diversity in the US (Spielman and Logan 2013; Farrell and Lee 2011). Variation in the composition of neighborhoods, especially in the level of heterogeneity, can have an affect on ACS data quality.

The variations in neighborhood composition is one potential cause of the systematic variations in data quality. Given the complex design of the ACS it is difficult to estimate the net effect of neighborhood-level diversity but it remains an import potential source of uncertainty in the ACS small area estimates. This income-diversity gradient is one of many

possible demographic gradients that may affect ACS data quality, these demographic correlates of ACS data quality are underexplored in the literature and are an important area for future research.

4. Geographic strategies for improving the ACS

The ACS is not an incremental change to census methodology, but a fundamental shift in the way demographic data are collected, organized, and disseminated. The ACS design is informed by decades of theoretical and empirical findings inside and outside the USCB. The final ACS products are the result of many decisions, each of which involves balancing both statistical and financial costs and benefits. Many of the problems the ACS faces are closely related to long running threads of geographic research, in this section we outline a few geographic strategies and areas of research that may improve the ACS.

The ACS sample size is simply insufficient to provide high-frequency data at high spatial resolution with low uncertainty levels. The best available solution to this problem is to manage uncertainty in the data through aggregation. Aggregation works because estimates supported by larger sample sizes tend to have smaller margins of error. There are a variety of ways to boost sample sizes through aggregation. Aggregation by attribute, e.g., merging male and female counts on an educational attainment attribute, can reduce uncertainty because the combined estimate is supported by larger number of samples. The cost of this approach is demographic granularity. Geographic aggregation is grouping census geographies into new larger composite geographic units. The USCB provides simple formulae for re-calculating the standard errors of geographically aggregated variables (U.S. Census Bureau 2009b). Geographic aggregation carries certain risks. Any aggregation of census small areas will lead to a reduction in standard errors, however it also creates modifiable areal unit problems. That is, it is possible to substantially alter the spatial pattern in the input data by, for example, combining a high income tract and low income tract into a new moderate income “region.” While this would reduce uncertainty, it would also dilute important socio-spatial patterns in the data.

Aggregation can be done using “intelligent” computational methods that preserve spatial patterns in the input data. Prior to the 2001 UK census the AZP algorithm (Openshaw and Rao 1995) was used by Martin to design a new set of data-driven geographic units for dissemination (Martin 1998, 2001). A similar effort might be beneficial in the U.S., especially given the ability to link census and ACS returns to the highly disaggregated and geo-located MAF. It may be possible to design geographies using the MAF via a computational procedure that maximized the ACS’s limited sampling budget by ensuring that geo-located responses were combined into small area estimates using an optimal geography. Where “optimal” might be defined as geographic zones that minimize uncertainty in the final estimates for a given spatial resolution.

A twist on the geographic aggregation strategy is the joining of noncontiguous tracts to form substantive geo-demographic clusters (Harris, et al. 2005). Geodemographic systems are multivariate typologies of geo-located individuals or administrative units. Such multivariate groupings would de-emphasize single variable estimates and their margins of error, in favor

of an emphasis on local community types as revealed through combinations of variables, for the past 20 years this strategy has been used effectively in the U.K. where an official public geodemographic classification is available and widely used (Vickers and Rees 2007; Spielman and Singleton 2013).

In addition to these strategies of altering the dissemination of existing data, it is also possible that geospatial research could contribute to estimation methodologies themselves. Current estimation methodologies make limited use of geospatial data. It may also be possible to reduce the margins of error in ACS estimates by developing (or applying) new estimation methodologies. Innovations in methodology are a relatively inexpensive way to improve the survey (see Ghosh and Rao 1994 and Little 2012). In particular there are several *geographic* methodologies could improve the quality of the ACS: The development of accurate small area population controls and the use of ancillary data to improve the estimation of sample weights.

A key source of error in the ACS is lack of cotemporaneous population controls for small areas like census tracts. The lack of such controls accounts 15-25% of the increase in variance between the decennial long form and the ACS (Navarro 2012). The Master Address File provides a geographically referenced list of housing units, which could, in theory, be used to estimate tract-level populations. The USCB has determined that the MAF is not suitable for estimating small area populations, though the reasons for this are not entirely clear and do not seem to be in the public record. Perhaps the MAF update cycle is too slow, or the 10.2 percent MAF over count has been deemed unacceptable. Swanson and McKibben (2010) have suggested that enhancing the Master Address File with auxiliary information would make it a suitable source of small area population controls. Since Wright (1936), geographers have been developing methods to improve small area population estimates with ancillary spatial information. Continued innovation in methods for spatially linking ancillary data to surveys is one of the ways the discipline of geography can contribute to the improvement of the ACS through the development of accurate small area population controls (e.g. Nagle et al. *in press*).

Part of the reason it is difficult to estimate survey weights is that there is little information available about survey respondents prior to their selection into the sample. The USCB uses a method called Generalized Regression (GREG) to reduce uncertainty in small-area estimates by reducing the variance of the sample weights. The ACS GREG procedure incorporates person-level administrative data from federal agencies on age, race, and gender. Estimates for large counties and cities do not use GREG. In order to maintain a single methodology for the entire country, the GREG procedure only uses variables with national coverage. The limited set of variables currently used could potentially be expanded to include real estate sales or other transactional databases. For example, home prices exhibit strong spatial autocorrelation and therefore it might be possible to use a geostatistical framework to estimate home prices- an important aspect of a household's economic characteristics. Through the currently used GREG procedure such a database would have the potential to improve the ACS weights. Research into both of these strategies - of defining new geographic units through aggregation, and improving spatial data-fusion techniques to

improve the estimation ACS weights— are areas in which geographers can play a unique role.

5.0 Conclusion

The ACS is the best available resource for geographers seeking to study small area social and demographic variation. However, ACS small area estimates are plagued by attribute uncertainty. For some variables the uncertainty (margin of error) is so large that the data are difficult to use. Uncertainty in the ACS can be decomposed into two broad categories— uncertainty that is rooted in methodological decisions made by the USCB (section 2) and uncertainty that is beyond the Bureau’s control (section 3).

There is no reason to think that the design of the ACS is fundamentally flawed. The heavy emphasis placed on bias reduction in the ACS design means that, on average, the estimates are likely reasonable approximations of the populations they represent. Nonetheless, the margins of error are troubling from a geographic perspective because of the inverse relationship between geographic and attribute resolution. Little (2012) argues that a fundamental philosophical shift is necessary within both federal statistical agencies and among data users, “we should see the traditional survey as one of an array of data sources, including administrative records, and other information gleaned from cyberspace. Tying this information together to yield cost-effective and reliable estimates...” However, Little also notes that for the Census “combining information from a variety of data sources is attractive in principle, but difficult in practice” (Little 2012, p.309). By understanding the causes of uncertainty in the ACS and some potential solutions, geographers can use the survey more effectively and potentially devise methods to improve the estimates.

References

- Alexander CH. Still rolling: Leslie Kish’s” rolling samples” and the American Community Survey. *Survey Methodology*. 2002; 28(1):35–42.
- Alexander CH. A continuous measurement alternative for the US Census. *Proceedings of the American Statistical Association Survey Research Methods Section*. 1993:486–491.
- Anderson, MJ.; Citro, C.; Salvo, JJ. *Encyclopedia of the US Census: From the Constitution to the American Community Survey*. CQ Press; 2011.
- Bates, L.; Hartman, J. Research on Master Address File Quality – Implications for the American Community Survey. United States Department of Commerce; 2012 Feb 3. American Community Survey Research and Evaluation Report Memorandum Series #ACS12-RER-03.
- Citro, CF.; Kalton, G., editors. *Using the American Community Survey: benefits and challenges*. National Academy Press; 2007.
- Environmental Systems Research Institute (ESRI). [2/1/2013] The American Community Survey: An ESRI White Paper. 2011. <http://www.esri.com/library/whitepapers/pdfs/the-american-community-survey.pdf>
- Farrell CR, Lee BA. Racial diversity and change in metropolitan neighborhoods, *Social Science Research*. 2011; 40(4):1108–1123.
- Fay RE, Train GF. Aspects of survey and model-based postcensal estimation of income and poverty characteristics for states and counties. *Proceedings of the Government Statistics Section, American Statistical Association*. 1995:154–159.
- Fuller, WA. *Sampling statistics*. Wiley; 2011.
- General Accounting Office. 2010 CENSUS: Efforts to Build an Accurate Address List Are Making Progress, but Face Software and Other Challenges. 2009 Report # GAO-10-140T.

- Ghosh M, Rao JNK. Small Area Estimation: An Appraisal. *Statistical Science*. 1994; 9(1):55–76.
- Harris, R.; Sleight, P.; Webber, R. *Geodemographics, GIS and neighbourhood targeting*. Wiley; 2005.
- Kish L. Rolling samples and censuses. *Survey Methodology*. 1990; 16(1):63–79.
- Kish L. Combining multipopulation statistics. *Journal of statistical planning and inference*. 2002; 102(1):109–118.
- Little RJ. Calibrated Bayes: an Alternative Inferential Paradigm for Official Statistics. *Journal of Official Statistics*. 2012; 28(3):309–372.
- Martin D. Optimizing census geography: the separation of collection and output geographies. *International Journal of Geographical Information Science*. 1998; 12(7):673–685. [PubMed: 12294535]
- Martin D. Towards the Geographies of the 2001 UK Census of Population. *Transactions of the Institute of British Geographers*. 2001; 25(3):321–332.
- MacDonald H. The American community survey: Warmer (more current), but fuzzier (less precise) than the decennial *census*. *Journal of the American Planning Association*. 2006; 72(4):491–503.
- Nagle, Nicholas N.; Battenfield, Barbara P.; Leyk, Stefan; Spielman, Seth. *Dasymetric Modeling and Uncertainty*. *Annals of the Association of American Geographers*. Forthcoming
- Navarro, F. An Introduction to ACS Statistical Methods and Lessons Learned. *Measuring People in Place Conference*; Boulder, CO. 2012. http://www.colorado.edu/ibs/cupc/workshops/measuring_people_in_place/themes/theme/asiala.pdf
- New York Times. New York's Doomsday Book, or Demonography; 1923 Apr 1.
- Openshaw S, Rao L. Algorithms for reengineering 1991 Census geography. *Environment and Planning A*. 1995; 27(3):425–446. [PubMed: 12346252]
- Rao, JNK. *Small area estimation*. Vol. 327. Wiley-Interscience; 2003.
- Salvo JJ, Lobo AP. Moving from a decennial census to a continuous measurement survey: factors affecting nonresponse at the neighborhood level. *Population Research and Policy Review*. 2006; 25(3):225–241.
- Singleton A, Spielman SE. The Past Present and Future of Geodemographics in the US and UK. *The Professional Geographer*. 2013
- Spielman SE, Logan J. Using High Resolution Population Data to Identify Neighborhoods and Determine their Boundaries. *Annals of the Association of American Geographers*. 2013; 103(1): 67–94. [PubMed: 23279975]
- Starsinic, M. American Community Survey: Improving Reliability for Small Area Estimates. *Proceedings of the 2005 Joint Statistical Meetings on CD-ROM*; 2005. p. 3592-3599.
- Starsinic, M.; Tersine, A. *Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association; 2007. Analysis of variance estimates from American Community Survey multiyear estimates; p. 3011-3017.
- Swanson DA, McKibben JN. New Directions in the Development of Population Estimates in the United States? *Population Research and Policy Review*. 2010; 29(6):797–818. [PubMed: 21076679]
- U.S. Census Bureau. *American Community Survey*. U.S. Government Printing Office; Washington, DC: 2009a. Design and Methodology.
- U.S. Census Bureau. *U S Census Bureau, A Compass for Understanding and Using American Community Survey Data: What Researchers Need to Know*. U. S. Government Printing Office; Washington, DC: 2009b.
- United States Census Bureau. [2/5/2013] *The American Community Survey: Coverage Rates Data*. 2013a. http://www.census.gov/acs/www/methodology/coverage_rates_data/
- United States Census Bureau. [2/5/2013] *The American Community Survey: Item Allocation Rates Data*. 2013b. http://www.census.gov/acs/www/methodology/item_allocation_rates_data/
- United States Census Bureau. [2/5/2013] *The American Community Survey: Item Allocation Rates Data*. 2013c. http://www.census.gov/acs/www/methodology/sample_size_data/
- Webber R. Making the Most of the Census for Strategic Analysis. *The Town Planning Review*. 1978; 49(3):274–284.

Wright JK. A method of mapping densities of population: With Cape Cod as an example. *Geographical Review*. 1936; 26(1):103–110.

Figure 1.
Census tract sample size (housing units) in the 2007-2011 ACS and the 2000 decennial census long form

Figure 2.

This figure uses the 2007–2011 ACS estimates of African-American median Household Income to divide all census tracts into 20 equal size bins (quantiles) such that each bin contains 5% of all census tracts. The leftmost bin contains the poorest 5% of census tracts and the rightmost bin contains the richest 5% of census tracts. The label shows the range of median household incomes for each bin. Within each bin we have calculated the percent of census tracts that belong in each of four data quality categories. The figure shows the inverse relationship between data quality (CV) and African-American median household income.

Figure 3.

This figure shows the relationship between ACS income estimates and ACS income data quality. For each of the 150 largest MSAs tract-level median household income estimates are ordered and placed into 1-percentile bins such that the lowest bin (1st percentile) contains the poorest 1% of tracts and the 99th percentile contains the wealthies tracts in the city. Then each of these city-specific percentile bins are pooled and plotted in the figure. The leftmost bar represents the poorest 1% of all tracts in the 150 largest MSA's, where the percentile is relative to each MSAs income distribution. Bars show the interquartile range for each percentile.

Figure 4.

This figure shows geographic patterns in ACS income data quality. For each of the 150 largest MSAs tracts are ordered and placed into 1-percentile bins on the basis of the distance between the tract centroid and the city center. The lowest bin (1st percentile) contains the tracts closest to the city center and the 99th percentile contains the tracts furthest from the center. Bars show the interquartile range for each percentile.

Figure 5. American Community Survey (2007-2011) Census Tract Economic Diversity by Median Household Income. The vertical axis shows the GINI coefficient (a measure of diversity) constructed using 16 income categories from the 2007-2011 ACS. The figure shows a strong negative association between neighborhood level income diversity and median neighborhood income.

Table 1

ACS Estimates of African-American Median Household Income in a group of contiguous tracts in Denver County, Colorado

Tract Number	African-American Median Household Income	Margin of Error
Census Tract 41.01	\$28,864	\$8,650
Census Tract 41.02	\$21,021	\$4,458
Census Tract 41.03	\$43,021	\$14,612
Census Tract 41.04	\$36,092	\$3,685
Census Tract 41.06	\$60,592	\$68,846