Macrothink Institute™

Journal of Sociological Research
ISSN 1948-5468
2014, Vol. 5, No.1

Number of Allocations as a Measure of "Within-Person Item-Incompleteness" (WPII): 2005 to 2010 Data from the American Community Survey

Carlos Siordia[1*]
[1] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

[*]Address correspondence to Carlos Siordia, PhD, 130 North Bellefield Ave, Pittsburgh, PA 15213.
Email: cas271@pitt.edu

Doi:10.5296/jsr.v5i1.6112 URL: http://dx.doi.org/10.5296/jsr.v5i1.6112

Abstract
The production of complete datasets from survey research requires that missing, illogical, or ambiguous values be "edited." The American Community Survey (ACS) in the United States (US) creates "flag" variables to indicate when a response has been "allocated" (i.e., edited in such a way that it differs from the original value). Identifying different data editing levels through easy-to-understand statistics may provide useful products for policy makers. This report sums the number of flags by person in order to measure item-incompleteness (i.e., number of questions requiring editing). By using ACS six single-year Public Use Microdata Use (PUMS) files from 2005 to 2010, the investigation shows how data quality (measured by within-person item-incompleteness "WPII") is most concentrated in male racial-ethnic minority groups. Quantifying the variability introduced by item-missingness in the production of population estimates from sample data is difficult. Identifying high concentrations of item missingness may be possible if detailed flag variables identifying edited data are made prevalent in secondary data sources.

Key words: Imputation; Allocation; Data Quality; Acs; Pums; Flags;

1. Introduction

The United States (US) has collected social, economic, and housing data for 163 years (i.e., since 1850). From 1940 to 2000, the "long form" questionnaire had been used every decennial to collect detailed data on the US population. Having detailed demographic data on the US population every 10 years was determined too large of a time gap to be considered informative for policy related purposes. In 1976, the US government directed federal agencies to determine if more timely data could be create (Kish, 1981). The desire to obtain "mid-decade" estimates (Torrieri, 2007) on the characteristics of the US population from samples gained momentum when it was discovered that the 1990 US decennial Census had an unexpectedly low unit (person) response rate. Leslie Kish, a University of Michigan professor at the Institute for Social Research, proposed the use of "cumulated rolling samples" as a solution for producing mid-decade estimates (Kish, 1981).

The argument was that accumulating rolling samples would, after several years, provide timelier and more useful estimates than decennial censuses (Siordia, 2014a). The view and proposed solution to producing mid-decade population estimates from samples did face opposition. For example, it was argued that the quality of the estimates from rolling samples could be heavily influenced by the US Census Bureau's ability to maintain a professional staff capable of meeting the demands exerted by the sophisticated methods proposed in the rolling-sample approach. In the end, the American Community Survey (ACS), sometimes referred to as the inter-decennial census, was eventually founded on the idea that "continuous measurement" was possible via the use of a rolling sample (Torrieri, 2007).

Planning for ACS's operational feasibility began in the 1990s and was first field tested in 1996 in four US administrative geographies (i.e., counties) and was later tested in 2000 over 1,000 counties. After trials in the 1990s and early 2000s, the ACS entered into "full implementation" in January of 2005. Since then, the ACS has collected detailed demographic and economic information on the US population from month-to-month. A 'timely and detailed profile' of the US population via the year-to-year rolling sample is now made possible by administering the long form survey and creating yearly population profiles of American communities.

Constitutional and federal legal mandates determine which question will be asked in the ACS—i.e., no data on the US population is collected via the ACS unless it meets a legal need. The US Congress, Office of Management and Budget, and the Department of Commerce fund the US Census Bureau (headquartered in Suitland, MD, USA) to administer the ACS. Every year, more than three million households enter the ACS sample and help develop population estimates for the more than three hundred million US residents. The ACS is a massive and challenging enterprise that merits both respect and scrutiny—as its data products have the ability to influence macro level economic and social events which in turn have the potential to impact the lives of individuals (i.e., micro events).

Governments around the world make use of large scale survey data on their citizenry to inform social policy and distribute economic resources. One of the main task of the federal US government is to distribute tax dollars across states (e.g., to build infrastructure). The ACS is influential in how billions of dollars are distributed. For example, in 2008, the ACS influenced the distribution of $562.2 billion in grants and $520.7 billion in direct payments (Reamer, 2010). Population estimates from the ACS sample are used to identify which entities are eligible for funds, how much should be allocated to them, and even play a role in determining the interest rates of federal loans. For example, the ACS estimates informed the formulas that allocated $261.1 billion to the Department of Health and Human Services for Medical Assistance Program (e.g., medical care for children in low income families). The ACS has the potential to significantly influence macro- and micro-social and economic events in the US. The accuracy of population estimates from ACS data is consequential. The core assumption on which this analysis is built on is the view that data quality in the sample has a direct effect on the level of precision in population estimates. Until now, no public report has attempted to quantify if and how "within-person" data quality (via a measure of questionnaire incompleteness) varies as a function of demographic characteristics (e.g., sex).

Unit-nonresponse (i.e., person as unit) and item-nonresponse (i.e., question as unit) have probably been present in survey research since its birth. Discussions on how missing values should be treated date back many decades (e.g., Little and Rubin, 1989) and discussions on how to "edit" data have been around for many years (see de Waal, 2013; and Van de Pol and Bethlehem, 1997). Abstracting "reality" into a numerical world via questionnaires is complex (Siordia, 2014b). Discussions on how modes of data collection influence the production of data date back decades (e.g., Deming, 1944; and Hochstim, 1967) and continue to receive specialized attention (e.g., West, Kreuter, and Jaenichen, 2013). In the US, ACS data is widely used by the government, private, and research sectors. Thankfully, a growing number of highly trained statisticians, demographers, economist, geographers, and many others are joining the ranks of the admirable ACS team charged with administering the survey and insuring its quality.

The production of population estimates from survey samples can be aided by complete data (i.e., no missing responses). Complete data is useful when attempting to calculate the confidence intervals around an estimate—to ascertain its level of precision. Creating complete datasets demands that missing, illogical, or ambiguous values be “edited” (assigned a useable value). Quantifying the variability introduced by item-missingness in the production of population estimates from sample data is difficult. Identifying high concentrations of item missingness may be possible if detailed ‘flags’ (variables identifying edited responses) are made available in data sources (Siordia & Young, 2013). Because the large scale ACS provides high quality data and seeks to be transparent in all its regards with the public, data creators at the US Census Bureau provide “flag” variables to indicate when a response has been “allocated” (i.e., edited in such a way that it differs from the original value). This is crucial to the ascertainment of within-person data quality, as others have explained that “item allocation rates are final measures of completeness that quantify how frequently allocation was the source of data in the production” of a particular estimate (McGovern & Griffin, 2003).

The amount of allocations is important for various reasons (Siordia & Le, 2013). For example, quantifying the precision of a population estimate from a sample only accounts for random phenomena in the probability of selecting a unit (i.e., person) and not the variability (i.e., bias) introduced by item (question) non-random processes (e.g., non-response). Thus, producing a simple-to-understand measures of potential contamination (i.e., bias) in the estimate from item-allocations is crucial to understanding how the validity of confidence intervals may vary as a function of population size and demographic characteristics.

The specific aim of this study is to use ACS flag variables to sum the number of allocations (referred by some as imputations) within an individual in order to determine if and how levels of data edits vary as a function of basic demographic factors. Please note the paper is not directly concern with “item nonresponse” as data editing algorithms have the ability to change recorded responses (Siordia, 2014b). In this study, “allocations” include changes to both responses and nonresponses—where high within-person level of allocation may signal a “high level of questionnaire incompleteness”. A high level of questionnaire incompleteness within-person has the potential to influence the production of low quality data. By summing number of allocations within-person, a racial-ethnic grouping scheme stratified by sex, can be used to understand if allocations are more prevalent in certain groups, i.e., if the variation in level of questionnaire incompleteness has detectable patterns. Basic statistical properties on the measures of interest are discussed in order to engage non-technicians who may be key stake-holders on the topic at hand.

The core argument of the project is that ‘allocation flags’ are highly instrumental as they help understand and measure within-person item-incompleteness (WPII—i.e., number of questions not requiring editing). In closing, data creators will be encouraged to create detailed data editing flags in order for secondary data users to understand the quality and reliability of data. Please note that at no point in this paper is the argument that ACS data is highly fallible being made. The author holds ACS data in the most highest of regards and considers it a gold standard in the field. The esoteric and microscopic technicalities being discussed here are intended to inform those who which to contribute to the production of high quality national data. The current endeavor is motivated by the belief that high quality national data has the potential to advance the administrative duties of a government in the production of a stable and democratic state.

2. Methods

2.1. Data

The analysis uses six single-year Public Use Microdata Use (PUMS) ACS files from the years 2005 to 2010. ACS PUMS files are made available to anyone with an internet connection[1]. Downloaded zip files contain errata notes and directions for codebooks and other technical documentation. All data management and analysis was conducted using Statistical Analysis Software (SAS) 9.3®. Independent estimates are produced for each year. Please note that ACS PUMS files may be updated when errata are noted. The PUMS files being used here were all downloaded in December of 2013 from the US Census website as "UNIX" zip files and have the following "last updated" dates: 2005 updated on November 9, 2010 ; 2006 updated on December 17, 2009; 2007 updated on September 23, 2008; 2008 file updated on December 9, 2010; 2009 updated on November 2, 2010; and 2010 updated on October 12, 2011.

2.2. Sample

The analytic sample from each of the six years from 2005 to 2010 only includes individuals at or over the age of 21 (legal age of consent in the US) who reside within the US mainland (contiguous states). This is done in order to reduce any potential effects non-mainland data collection efforts may have on the variability of the estimates. It may be that geographically disconnected areas experience different administrative teams, survey participant response profiles, and perceptions of the US federal government than the mainland. For example, some residents in Puerto Rico (a commonwealth of the US) may view any efforts by the US federal government with some suspicion (see Torruella, 1985). Because the following "racial" and "ethnic" groups represent almost 9 in every 10 US mainland residents (Siordia, 2014a), only their "racial-ethnic" groups are included in the analysis: Non-Latino-White (NLW—the majority group in the US); Non-Latino-Black (NLB—currently the largest minority group in the US); and Mexican-Latinos/as (MEX—the soon to be largest minority group in the US).

The US federal government allows individuals to be identified using these non-biological categories of race and ethnicity in order to determine who is a "minority" and "majority"—decisions which influence the formation of voting districts and allocation of civil right resources for the advancement of socioeconomic equality. Here are the sizes of the six analytic samples from ACS 2005 to 2010 PUMS files: 2005 has 1,872,520 actual people who are said to represent 180,909,153 people in the US mainland when the population weight is applied (variable PWGTP); 2006 has 1,930,232 actual people ("unweighted count") and which represent 187,579,032 individuals ("weighted count"); in 2007 the unweighted count is 1,946,115 and the weighted count is 189,004,341; in 2008 the unweighted count is 1,958,305 and the weighted count is 190,966,711; in 2009 the unweighted count is 1,977,431 and the weighted count is 192,632,594; and in 2010 the unweighted count is 1,993,020 and the

[1] As of December 2013, the 2005 file can be downloaded from: http://www2.census.gov/acs/downloads/pums/2005/; the 2006 files from http://www2.census.gov/acs/downloads/pums/2006/; the 2007 files from http://www2.census.gov/acs2007_1yr/pums/; the 2008 files from http://www2.census.gov/acs2008_1yr/pums/; the 2009 files from http://www2.census.gov/acs2009_1yr/pums/; and the 2010 files http://www2.census.gov/acs2010_1yr/pums/.

Journal of Sociological Research
ISSN 1948-5468
2014, Vol. 5, No.1

weighted count is 193,266,002. In short, the analysis makes use of response profiles for 11,677,623 actual people spread over 6 years to determine if the level of questionnaire completeness varies by race-ethnicity stratified by sex. The size of the overall analytic sample of ~11.7 million people should give some credence to the results. As mentioned earlier, no comparable data in terms of size, transparency, and availability can be found in the US.

2.3. Measuring Incompleteness via "Flag Variables"

Each of the six single-year PUMS ACS files under investigation contains hundreds (> 400) of variables, from these, there are only 42 flag variables present in all the years under analysis—from 2005 to 2010. The 'allocation flag variables' are displayed in detail in Appendix A. Allocation flags indicate when a response was edited in such a way as to have been altered from its original value. Because within-person data quality is measured in this analysis by within-person item-incompleteness (WPII), the 42 flags are summed within the individual to compute WPII for each the ~11.7 million persons in the analysis. People with a WPII score of 0 have no allocations over the 42 items—i.e., WPII=0 indicates the lowest level of questionnaire *completeness* being measured. In contrast, people with a WPII score of 42 would contain the maximum number of allocations—i.e., the highest level of questionnaire *incompleteness* being measured. Note that allocations may not refer to 'reassigned' values (through explicit or implicit logic in data processing algorithms) but only capture data edits—i.e., where value reassignment is produced from the use of hot-deck (geographically and demographically aware imputations) or cold-deck (geographically unaware and limited demographic awareness) procedures. More detailed documentation in ACS product is required to disentangle these assumptions.

2.4. Person Inflation Ratio

For each of the six years under analysis, both weighed and unweighted sample counts are presented. Details on how the population weight variable is created are available elsewhere (US Census Bureau, 2011). The weighted population counts from PUMS data will not be identical to those produce by internal files at the US Census Bureau since PUMS population weights account for the fact that public microdata is only a sub-sample of the complete ACS file (Keathley, Navarro, & Asiala, 2010). In order to show readers the average number of people represented by each actual respondent, the Person Inflation Ratio (PIR) is presented (see Siordia, 2013c), which uses the following formula: (weighted count ÷ weighted total population). As PIR numbers increase, the average number of people represented by each person increase—an increase in PIR may signal an increase on the potential for person associated measurement errors to propagate and contaminate the production of a population estimate and its confidence intervals. In other words, when both a high PIR and WPII are present, the potential for the propagation of non-random effects (bias) in the estimates may increase.

2.5. Describing Concentrations of Incompleteness

The weighted and unweighted count of individuals in each of the three racial-ethnic groups (NLW, NLB, MEX) are stratified by sex and presented for each of the six years under analysis (2005-2010). The WPII variable mean and standard deviation for each race-ethnic group stratified by sex for each year is provided in the tables. The project investigates the level of questionnaire incompleteness and "qualitatively" infers potential consequences for the precision of estimates in a way that invites non-technicians to understand the issues at hand. More technically, popular frequentist statistical approaches are not used to infer the statistical significance of changes over time or between race-ethnic groups. Stablishing the statistical significance of differences or change is avoided as its use may obfuscate the value of data for informing policy—where statistical inference is not always a priority or a part of the common

Journal of Sociological Research
ISSN 1948-5468
2014, Vol. 5, No.1

language.

The mean is the arithmetic average across all the observations by year, race-ethnic group, and sex is presented along with other easy-to-understand statistical measures. The mean is presented as it is one of the most widely used measures of central tendency. The standard deviation (SD) is another popular measure and represents the square root of the variance—a measure of the spread of observations over WPII. Larger SDs signal observations are more evenly spread out over the possible scores of WPII (from 0 to 42).

Because the mean is sensitive to both extremely large and small values, the race-ethnic and sex skewness and kurtosis of the WPII distribution is presented for each year. Skewness measures the degree and direction of asymmetry—i.e., how normally distributed WPII is in the group. Because a normal distribution has a skewness of zero, all the positively skew distributions of WPII below indicate the mean is greater than the median. Kurtosis measures the 'heaviness' of tails in a distribution, where a normal distribution has a kurtosis of zero. All the positive kurtosis scores in the tables indicate tails in the distribution of WPII are "lighter" than in a normal distribution—i.e., distribution of WPII is leptokurtic.

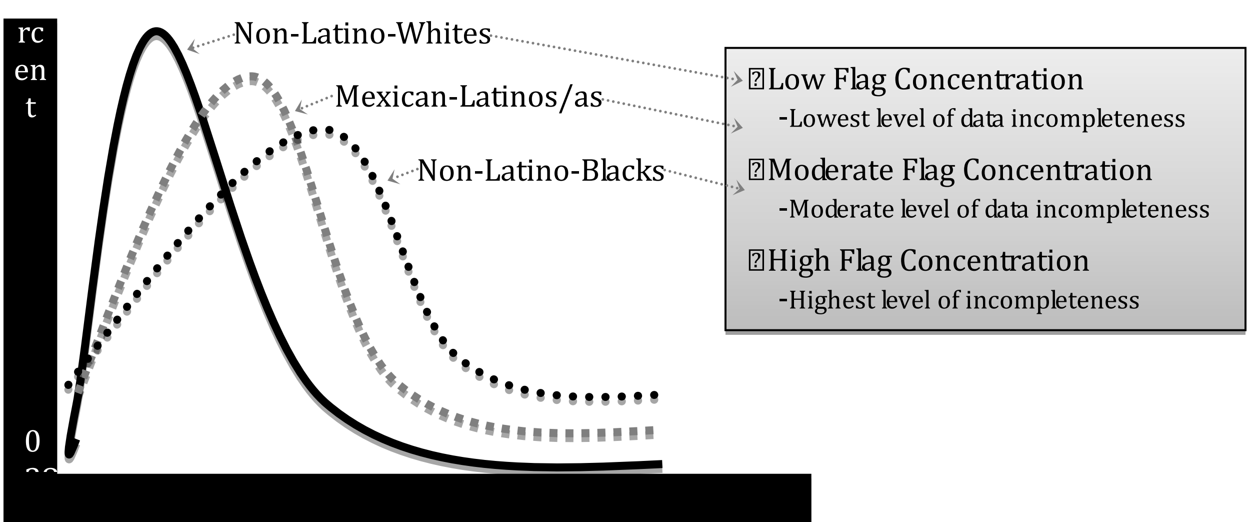
The potential range of WPII is 0 to 42. Because all the sub-populations across the years have at least one person with 0 allocations, only the 'maximum' of the range is presented in the tables. A broad sense on the volatility in the precision of a population estimate from a particular race-ethnic-sex-year group may be gathered by evaluating a sample's PIRs and their WPII's mean, SD, skewness, kurtosis.

3. Results

The general patterns of WPII distributions over the ~11.7 million people shown in Table 1 and 2 are visually represented in Figure 1. The conceptual representation of how flag concentrations vary between racial-ethnic groups indicates that in general NLWs (a racial-ethnic majority group in the US) have the highest level of data completeness from 2005 to 2010, followed by MEXs (an ethnic minority group in the US) and then NLBs (a racial-ethnic minority group in the US). If the racial-ethnic hypothetical lines were drawn by sex, we would see more flags (i.e., item editing) in males than in females. A series of complex comparisons are made possible by the detailed tables. Brief examples of interpretations are provided here.

Figure 1

Conceptual representation of "flag concentration" and "level of completeness"



Journal of Sociological Research
ISSN 1948-5468
2014, Vol. 5, No.1

From Table 1 and in year 2005, we see NLW females only have a PIR of 90 compared to NLBs (121) and MEX (136) females. This means that on average, NLBs and MEXs females represented more of their counterparts in the population than NLWs. After observing that

Table 1

Counts of females by race-ethnicity and Within-Person Item-incompleteness descriptive

2005	NLW[1]	NLB[2]	MEX[3]	2008	NLW	NLB	MEX
W[4]	74,082,162	12,420,634	7,318,796	W	76,835,301	13,189,769	8,310,267
UW[5]	825,696	102,437	60,809	UW	842,733	111,618	71,635
PIR[6]	90	121	136	PIR	91	118	116
Mean	0.90	1.78	1.21	Mean	1.41	2.52	1.79
SD[7]	2.96	4.51	3.82	SD	3.89	5.52	4.85
Skewness	5.95	3.83	5.26	Skewness	4.52	3.10	4.20
Kurtosis	41.92	15.77	30.99	Kurtosis	23.20	9.70	18.96
Maximum	37	37	38	Maximum	37	36	37

2006	NLW	NLB	MEX	2009	NLW	NLB	MEX
W	76,128,551	12,913,304	7,670,690	W	77,229,032	13,399,326	8,543,293
UW	836,752	108,900	64,820	UW	846,380	113,894	74,942
PIR	91	119	118	PIR	91	118	114
Mean	1.01	1.91	1.26	Mean	1.11	1.98	1.52
SD	3.16	4.73	3.89	SD	3.39	4.80	4.44
Skewness	5.61	3.73	5.24	Skewness	5.35	3.70	4.27
Kurtosis	37.43	14.96	30.91	Kurtosis	33.28	14.65	24.35
Maximum	37	38	37	Maximum	36	36	37

2007	NLW	NLB	MEX	2010	NLW	NLB	MEX
W	76,445,723	13,019,710	7,891,225	W	76,698,926	13,742,814	9,226,619
UW	842,300	109,327	67,009	UW	844,307	118,141	78,834
PIR	91	119	118	PIR	91	116	117
Mean	1.02	1.93	1.33	Mean	1.36	2.31	1.83
SD	3.18	4.77	4.09	SD	3.94	5.35	5.12
Skewness	5.58	3.69	5.10	Skewness	4.72	3.35	4.12
Kurtosis	36.86	14.55	29.06	Kurtosis	24.77	11.53	17.71
Maximum	38	37	38	Maximum	37	37	37

[1] Non-Latino-White; [2] Non-Latino-Black; [3] Mexican-Latino; [4] Weighted number of people (using single population weight); [5] Unweighted counts, i.e., actual number of individuals in sample; [6] Average number of people represented by each actual respondent, i.e., "Person Inflation Ratio" = (weighted count ÷ weighted total population); [7] Standard deviation

NLBs and MEXs females both have more right-shifted WPII distributions (higher WPII mean, SD, and lower skewness and kurtosis), it may be possible to conclude that any within-person bias introduced via editing protocols has the potential to propagate to influence the production of point estimate but not its confidence intervals—a key point since the precision of the estimate can only be captured with the latter.

By looking at PIR values from 2005 to 2010 we can see how the average number of people for each survey participant has been decreasing—i.e., the inflation of potential within-person biases has diminished over time (PIR has decreased). Unfortunately, the results shown in Table 1 indicate that the average WPII (i.e., level of within-person item-incompleteness) increased during the 2005 to 2010 period—suggesting the need for data editing has increased since 2005. For example, in 2005 and for NLWs, the average WPII was 0.90 and was 1.36 by 2010. Amongst females and years under investigation, the highest level of item-incompleteness is found in 2008 for NLBs (average WPII=2.52).

From Table 2 we see males have higher levels of PIRs and WPIIs—this suggests the potential for propagation of within-person bias into estimates may be greater from males than females. From year 2008, we also see NLW males had a PIR of 94—a lower value when compared to NLBs (124) and MEX (131) males. WPII is most concentrated in NLB males while PIR is highest in MEX males throughout the 2005-2010 period. Note how positive skewness seems to be decreasing (WPII distribution shifting to the right indicating more data editing) from 2005 to 2010 with a peculiar bump in the trend in the year 2008—the same is present in females. No clear explanation for the '2008 exception' is evident from the data. Perhaps and administrative event during the 2008 survey year occurred that reduced the real amount of data editing or synthetically altered the amount of documented data editing.

Table 2

Counts of males by race-ethnicity and Within-Person Item-incompleteness descriptive statistics

2005	NLW[1]	NLB[2]	MEX[3]
W[4]	69,016,148	9,813,825	8,257,588
UW[5]	748,168	74,183	61,227
PIR[6]	92	132	135
Mean	0.98	2.02	1.61
SD[7]	3.13	4.87	4.49
Skewness	5.78	3.56	4.56
Kurtosis	38.68	13.18	22.48
Maximum	35	36	37

2008	NLW	NLB	MEX
W	71,936,457	11,202,349	9,492,568
UW	769,141	90,496	72,682
PIR	94	124	131
Mean	1.44	2.89	2.24
SD	4.01	5.90	5.47
Skewness	4.49	2.73	3.69
Kurtosis	22.53	7.16	14.03
Maximum	35	36	36

2006	NLW	NLB	MEX
W	71,102,665	10,861,008	8,902,814
UW	763,852	87,526	68,382
PIR	93	124	130
Mean	1.10	2.29	1.77
SD	3.35	5.20	4.74
Skewness	5.37	3.19	4.32
Kurtosis	33.39	10.36	20.01
Maximum	37	36	36

2009	NLW	NLB	MEX
W	72,385,045	11,420,943	9,654,955
UW	773,147	93,014	76,054
PIR	94	123	127
Mean	1.20	2.30	1.91
SD	3.59	5.21	5.00
Skewness	5.11	3.25	4.17
Kurtosis	29.62	10.88	18.40
Maximum	35	37	37

Journal of Sociological Research
ISSN 1948-5468
2014, Vol. 5, No.1

2007	NLW	NLB	MEX
W	71,484,396	11,009,646	9,153,641
UW	769,280	88,118	70,081
PIR	93	125	131
Mean	1.12	2.29	1.86
SD	3.41	5.18	4.87
Skewness	5.30	3.17	4.17
Kurtosis	32.35	10.20	18.59
Maximum	36	37	37

2010	NLW	NLB	MEX
W	72,172,711	11,670,221	9,754,711
UW	774,553	97,172	80,013
PIR	93	120	122
Mean	1.46	2.53	2.25
SD	7.14	5.59	5.67
Skewness	4.53	3.09	3.69
Kurtosis	22.35	9.48	13.75
Maximum	36	36	37

[1] Non-Latino-White; [2] Non-Latino-Black; [3] Mexican-Latino; [4] Weighted number of people (using single population weight); [5] Unweighted counts, i.e., actual number of individuals in sample; [6] Average number of people represented by each actual respondent, i.e., "Person Inflation Ratio" = (weighted count ÷ weighted total population); [7] Standard deviation

4. Conclusions

Data editing is required in the production of complete datasets as item-missingness maybe prevalent in survey research. The empirical analysis demonstrates that within-person item-incompleteness (WPII)—a proxy measure of within-person data quality—varies in detectable ways: racial-ethnic majority females (i.e., NLWs females) have the smallest level of WPII concentrations from 2005 to 2010 indicating the possibility that any within-person biases with the potential to propagate into population estimates are at the lowest levels within this sub-population. The data also indicate data editing in the ACS may have increased from 2005 to 2010. Because within-person item-incompleteness patterns have the potential to produce non-random effects (biases) difficult to quantify in the production of population estimates and its confidence intervals, more research is needed. If data creators decide to manipulate data to eradicate missing, illogical, or ambiguous values, then within-person item editing should be tracked with highly detailed flag variables—making it possible for secondary data users to identify when abstraction from reality through survey questions may have passed through an additional filter with the potential to introduce systemic error.

The project is limited in that the detailed tables could have been shown by mode of data collection (mail, phone, and in-person). Future work should investigate potential differences in WPII by mode. From other research, it is possible to expect WPII will be higher in mail mode than in phone or in-person modes. The use of "data quality" is not clearly defined early in this report—an omission primarily driven by the need to stay on point. Note how it is assumed, by the treatment of the term, that an individual who provides all the responses for a questionnaire has high within-person data quality. This may not be the case as the validity, reliability, and applicability of each question may fluctuate within and between survey respondents. In other words, the presence of non-edited data does not insure "high data quality." For example, most of the data in the ACS is derived from a proxy (i.e., a person who resides in the same household)(Siordia, 2014c). Future work should explore this topic and specifically focus on the fact that ACS uses a "reference person" to complete the survey for all the individuals in the household. It may be that WPII varies by "self-report" (reference person reports of self) and "proxy-report" (reference person reports on others in the household).

Notwithstanding these limitations, the project is novel in that it focuses on how within-person item-incompleteness can help understand issues related to data quality and how these patterns may have implications for the precision of population estimates from sample data. The current project is made possible because of the high caliber of ACS data. If a reader is tempted to think differently of the ACS, he or she would be encouraged to consider the fact that the vast majority of datasets only make use of a few hundred or thousand people who are selected into the sample using less sophisticated sampling methods than those employed by the ACS. The protocols in the production of data in these relatively small data sets is obscured at best and frequently unavailable for the scrutiny of the public. Delving into the crevices of esoteric survey phenomena is only possible with high quality and transparent data like the ACS. If the production of national level data has the potential to advance the proper fulfillment of administrative duties by governments, and the latter has the ability to help advance the production of a stable and democratic state, then specialist and stake-holders should continue to explore how item-missingness can be reduced and captured in the production of population estimates from samples.

5. References

de Waal, T. (2013). Selective Editin: A Quest for Efficiency and Data Quality. Journal of Official Statistics, 29, 473-488.

Deming, W. E. (1944). On errors in surveys. American Sociological Review 9:359-369.

Hochstim, J. R. (1967). A critical comparison of three strategies of collecting data from households. Journal of the American Statistical Association, 62:976-989.

Keathley, D. H., Navarro, A., & Asiala, M. E. (2010). An Analysis of Alternate Variance Estimation Methods for the American Community Survey Group Quarters Sample. Joint Statistical Meetings, Vancouver. Washington, DC: US Census Bureau

Kish L. Space/time variations and rolling samples. Journal of Official Statistics. 1998;14:31-46.

Kish, L. (1981). Using cumulated rolling samples to integrate Census and survey operations of the Census Bureau: an analysis, review, and response. Washington, DC: U.S. government Printing Office.

Little, R. J., & Rubin, D. B. (1989). The analysis of social science data with missing values. Sociological Methods & Research, 18(2-3), 292-326.

McGovern, P. D. (2004). A quality assessment of data collected in the American Community Survey (ACS) from households with low English proficiency. Survey Methodology, 01.

Reamer AD. Surveying for Dollars: The Role of the American Community Survey in the geographic distribution of federal funds. Metropolitan Policy Program at Brookings, Washington D.C.; 2010.

Siordia C. (2013). Methodological note: allocations with health insurance items in the American Community Survey. Journal of Frailty and Aging, 6:149-153.

Siordia C., and Le V. D. (2013). Precision of Disability Estimates for Southeast Asians in American Community Survey 2008-2010 Microdata. Central Asian Journal of Global Health, 1(2): DOI:10.5195/cajgh.2012.2166-7403.

Siordia C., and Young R. L. (2013). Methodological Note: Allocation of Disability Items in the American Community Survey. Disability and Health Journal, 6(2):149-153.

Siordia C. (2014a). Detecting "real" population changes with American Community Survey data: The implicit assumption of treating between-year differences as "trends." Journal of Sociological Research, 4(2):494-509.

Siordia C. (2014b). An Algorithm for Creating Alternate CES-D Composite Scores: Advancing

Research through Methodological Clarity. Journal of Depression and Anxiety, 3(2):149-159. http://dx.doi.org/10.4172/2167-1044.1000149

Siordia C. (2014c). Proxy-reports in the ascertainment of disability prevalence with American Community Survey Data. Journal Frailty and Aging, 2014, http://www.jfrailtyaging.com/all-issues.html?article=188.

Torrieri, N. K. (2007). America is Changing, and So is the Census. The American Statistician, 61(1).

Torruella, J. R. (1985). The Supreme Court and Puerto Rico: The Doctrine of Separate and Unequal. La Editorial, UPR.

US Census Bureau. (2011). ACS Accuracy of the Data (2011) on http://www.census.gov/acs/www/data_documentation/documentation_main/

Van de Pol, F. and Bethlehem, J. (1997). Data Editing Perspectives. Statistical Journal of the United Nations ECE, 14, 153–171.

West, B., Kreuter, F., and Jaenichen U. (2013). “Interviewer” Effects in Face-to-Face Surveys: A Function of Sampling, Measurement Error, or Nonresponse? Journal of Official Statistics, 29, 277-297.

Appendix A

Allocation flags available in American Community Survey from 2005 to 2010

1. FAGEP = Age allocation flag
2. FANCP = Ancestry allocation flag
3. FCITP = Citizenship allocation flag
4. FCOWP = Class of worker allocation flag
5. FENGP = Ability to speak English allocation flag
6. FESRP = Employment status recode allocation flag
7. FFERP = Children born within the past 12 months allocation flag
8. FGCRP = Responsible for grandchildren allocation flag
9. FHISP = Detailed Hispanic origin allocation flag
10. FINDP= Industry allocation flag
11. FINTP = Interest, dividend, and net rental income allocation flag
12. FJWDP = Time of departure to work allocation flag
13. FJWMNP = Travel time to work allocation flag
14. FJWRIP = Vehicle occupancy allocation flag
15. FJWTRP = Means of transportation to work allocation flag
16. FLANP = Language spoken at home allocation flag
17. FLANXP = Language other than English allocation flag
18. FMARP = Marital status allocation flag
19. FMIGP = Mobility status allocation flag
20. FMIGSP = Migration state allocation flag
21. FMILPP = Military periods of service allocation flag
22. FMILSP = Military service allocation flag
23. FOCCP = Occupation allocation flag
24. FOIP = All other income allocation flag
25. FPAP = Public assistance income allocation flag
26. FPOBP = Place of birth allocation flag
27. FPOWSP = Place of work state allocation flag
28. FRACP = Detailed race allocation flag
29. FRELP = Relationship allocation flag
30. FRETP = Retirement income allocation flag
31. FSCHGP = Grade attending allocation flag
32. FSCHLP = Highest education allocation flag
33. FSCHP = School enrollment allocation flag
34. FSEMP = Self-employment income allocation flag
35. FSEXP = Sex allocation flag
36. FSSIP = Supplementary security income allocation flag
37. FSSP = Social security income allocation flag
38. FWAGP = Wages an salary income allocation flag
39. FWKHP = Usual hours worked per week past 12 months allocation flag
40. FWKLP = Last worked allocation flag
41. FWKWP = Weeks worked past 12 months allocation flag

42. FYOEP = Year of entry allocation flag

Funding

This work was supported by the National Institute of Aging at the National Institutes of Health (grant number T32 AG000181to A. B. Newman).