Unit-Level Logistic Mixed Effects Models for Small Area Estimation of Poverty Estimates[1]

Bavwidinsi B. Bawawana,[2] Samuel Szelepka,[2] and Xingyou Zhang[3]

Abstract The American Community Survey (ACS) publishes annual poverty estimates for large counties. In response to the demand for poverty estimates for U.S. counties and school districts, the Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program produces poverty estimates for all U.S. counties and school districts from area-level Fay-Herriot models. This paper focuses on estimating county-level poverty rates for the year 2014, using the unit-level logistic mixed effects model. We fit an unweighted multilevel logistic regression (MLR) model with demographic predictors and state- and county-level random effects. To account for the design of the ACS survey, we consider models with weights scaled based not only on the person sample size (PSS), but also on the housing unit sample size (HUSS). Given the individual demographic characteristics, we estimate the predicted probability that an individual is in poverty. An aggregation within county is made to generate the corresponding county-level poverty rates using the U.S. Census Bureau post-censal population estimates. In comparing ACS direct estimates, SAIPE estimates, and the estimates from the MLR models, the distributions of all three estimates are similar for large counties. For small size counties, the distributions of the estimates produced via MLR models tend to have smaller ranges compared with the distribution of ACS direct estimates. MLR models with weights scaled based on the person sample size almost always yield estimates with smaller mean absolute differences with the ACS.

Key words: multilevel logistic model, small area estimation, American Community Survey, Small Areas Income and Poverty Estimates, Fay-Herriot Model, Proc Glimmix

1 Introduction

To respond to the growing demand for annual, up-to-date data on social, economic, and housing characteristics, the U.S. Census Bureau developed a nationwide large scale-survey, the ACS (https://www.census.gov/programs-surveys/acs/). This survey is carried out yearly, and timely estimates are produced for different geographic units. Among the many data produced from the ACS data, the U.S. Census Bureau generates poverty estimates for the nation, all U.S. states, and counties. However, owing to the fact that some county populations are small, the ACS program only releases yearly poverty estimates for large counties with populations of at least 65,000. Combining surveys from multiple years, ACS releases poverty estimates for all counties from 5-year data. In 2014, the ACS released 1-year poverty estimates for 817 large counties out of 3,142 counties. To fill a gap created by the lack of annual poverty estimates for counties with small populations, the SAIPE (https://www.census.gov/did/www/saipe/) program produces model-based estimates of poverty and income for all counties. The SAIPE program generates its county estimates by fitting a Fay-Herriot model (Fay and Herriot, 1979). To address the insufficiency in sample size for small counties, the Fay-Herriot model borrows strength from

[1]**Disclaimer**: This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical, methodological, technical, or operational issues are those of the authors and do not necessarily represent the official position of the U.S. Census Bureau or Economic Research Service, U.S. Department of Agriculture. [2]U.S. Census Bureau, Washington, DC 20233. [3]Economic Research Service, U.S. Department of Agriculture, Washington, DC, 20024. Xingyou Zhang was employed with U.S. Census Bureau when this work was done.

other counties through the linking model (Rao, 2003) with covariates from administrative records and the decennial census long-form survey. All covariates and the response variable (ACS poverty counts) are log-transformed. The SAIPE model combines the ACS direct estimates with the synthetic estimates, the final estimates are the weighted combination of the ACS direct estimates and the regression estimates with area level covariates and random effects. Nevertheless, for any need arising for new geographic and/or demographic domain, the Fay-Herriot model calls for a new model fit (Guadarrama et al., 2016). SAIPE publishes poverty estimates for age groups 0-17, 5-17 (not in foster care), and all ages for every state and county in the nation. An additional set of estimates for age group 0-4 is published for states. To produce these estimates, SAIPE fits a total of eight separate models. For counties, a model is fit for each of the three published age groups. For state estimation, separate models are fit for the disjoint age groups 0-4, 5-17, 18-64, and 65+, as well as a fifth model for related children aged 5-17. The state estimates for age groups 0-17 and all ages are estimated as aggregates.

Throughout the years, many researchers (Ghosh and Meeden, 1986), (Battese et al., 1988), (Prasad and Rao, 1990), (Malec et al., 1999), (Malec, 2005), and (Hobza and Morales, 2016) have explored unit-level models with continuous or binary response variables to generate small area estimates, using empirical and/or hierarchical Bayesian methods. The motivation for this paper is to use ACS person-level data and U.S. Census Bureau's Population Estimates Program (PEP) data to produce reliable area-level poverty estimates from a unit-level logistic mixed effects model, with county and state random effects and fixed demographic covariates. We use the GLIMMIX procedure of SAS® software to fit models. Compared to area-level models, unit level models have the advantage of not requiring the refit of new models to produce estimates for new domains. They have the flexibility of the building block estimators, where a unique unit-level model can respond to the demand of any geographic and domain level estimates.

The ACS samples with unequal selection probabilities related to the size of the smallest sampling entity; areas with smaller populations are sampled at higher rates than areas with larger populations (for details about the ACS design and weights, refer to chapters 4 and 11 in https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html). The ACS data provide individual weights that account for the unequal selection probabilities. The weights are adjusted for non-response, and then post-stratified to ensure that the totals coincide with their corresponding demographic domain within the ACS weighting area. In the presence of complex survey data, fitting population models may lead to bias of point estimators. Many studies (Lavalle and Beaumont, 2015), (Pfeffermann, 1993), (Pfeffermann et al., 1998), (Pfeffermann and Sverchkov, 2007), (Verret et al., 2015), and (West et al., 2015) have enjoined the use of the weights in modeling complex survey data to shield against point estimate bias. To account for the ACS survey design, we produce poverty estimates by fitting weighted models. The weights are scaled to prevent the overstatement of the sample size within the area of interest. The scaled weights are transformed to sum to the ACS sample size within the scaling area rather than the population size. The estimates, produced by fitting the MLR unweighted model along with a set of models with scaled weights, are compared with ACS direct estimates (published and unpublished) and SAIPE area-level model-based estimates, using mean absolute differences (MADs) as a performance metric.

This paper is structured such that Section 2 introduces the data, Section 3 describes the model and different weights' scaling schemes, and Section 4 compares the ACS, SAIPE and MLR estimates. We conclude this paper with a brief discussion.

2 Data

We use two datasets, the 2014 1-year American Community Survey and the Vintage 2015 bridged-race postcensal population estimates. The ACS is a continuous large-scale survey, where sampled participants respond throughout the year. Actually, ACS program publishes reliable 1-year estimates for areas with at least 65,000 people every year.[4] The sample is drawn from people living in group-quarters and housing units. Beginning in 2011, the annual ACS sample selected consists of 3.54 million addresses, from five exclusive rotating sub-frame addresses across the United States. The 2014 ACS person-level data have a housing unit sample size of 2,403,614 for a total of 5,454,957 people. Of the 5,454,957 observations, 257,687 have missing poverty status. All states and counties are sampled. At the housing unit level, all members in a selected housing unit are included in the survey. Thus, there is total dependence of poverty status among related members of a housing unit, because poverty status is assigned to the family as a whole. Each family's poverty status is determined by comparing the household pre-tax income to the federal poverty line associated with its size. Provided that the family pre-tax income is less than the corresponding federal poverty line, all related members in the family are considered to be under poverty. Otherwise, the entire related members of the household are above poverty. The federal poverty line, also known as the poverty threshold, is the minimum level of necessary resources to meet a family or housing unit's basic needs. It is adjusted annually for inflation from the base-year 1982 using the average 12 months of consumer price index.

The Vintage 2015 bridged-race postcensal population estimates data have two data files, ages 0 to 85 years and over population estimates and ages 85 to 100 years over population estimates. For this study, we use ages 0 to 85 years and over population estimates. The file contains annual population estimates of the residents of the United States between April 1^{st} of the last decennial census (2010), to July 1^{st} of the indicated year (2015) by county, single-year of age (0, 1, 2, ..., 85 years and over), bridged-race category (White, Black or African American, American Indian or Alaska Native, Asian or Pacific Islander), Hispanic origin (not Hispanic or Latino, Hispanic or Latino), and sex. The PEP produces the bridged-race postcensal population estimates at the county level in collaboration with the National Center for Health Statistics (NCHS), but does not release them due to reliability concerns. Estimates for single-year of age are released to the public by the NCHS (https://www.cdc.gov/nchs/nvss/bridged_race.htm). Each year PEP produces and releases the unbridged-race postcensal population estimates at the county level by five-year age-group, race, Hispanic origin, and sex. However, the single-year of age bridged-race data used to conduct this study have been provided by the PEP program. The data have one additional category for race (Two or More Races). For simplicity, the population estimates in our study are considered as known constants.

The 2014 1-year ACS and the 2014 single-year of age 0 to 85 and over population estimates data both include an age variable, which provides a specific single age for each individual. We created a categorical variable "age-group" with 21 categories (0-4, 5-9, 10-14, 15-17, 18, 19, 20, 21-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+). The motivation behind the creation of these age-groups is to coincide with, by aggregation, analogous age-groups within the age-groups included during SAIPE production cycles (all, 0-4, 0-17, 5-17, 18-64, 65+). We also group race/ethnicity into non-Hispanic sub-groups (Non-Hispanic White, Non-Hispanic Black, Non-Hispanic

[4] ACS released the 2015 1-year supplemental estimates for areas with population of 20,000 to 64,999. (for details visit https://www.census.gov/acs/www/data/data-tables-and-tools/supplemental-tables/).

Asian or Pacific Islander, Non-Hispanic Indian or Alaska Native, Non-Hispanic Two or More Races), and Hispanic, for a total of six racial/ethnic groups. Gender has two categories, male and female. A total of 252 ($21 \times 6 \times 2$) demographic groups are created as the cross product of the three demographic indicators.

3 The Model and Scaling of the Weights Schemes

A logistic mixed effects model with state and county random effects and demographic fixed effects is fitted to the ACS data. The poverty status Y_{scdi} of sample unit i, in demographic domain d, county c and state s is equal to 1, noted $Y_{scdi} = 1$, if the housing unit income falls below the federal poverty line, and $Y_{scdi} = 0$, otherwise. Thus, under this particular model

$$\begin{array}{lcl} y_{scdi} & \sim & Bernouilli(p_{scd}) \\ logit(p_{scd}) & = & \log\left(\frac{p_{scd}}{1-p_{scd}}\right) = x'_{scdi}\beta + u_s + v_c, \end{array}$$

$s = 1, \cdots, 51$ states,

$c = 1, \cdots, C$ $(3,141$ total, excluding Kalawao,[5] Hawaii),

$d = 1, \cdots, 252$ demographic domains,

$i = 1, \cdots, n_{scd}$ individuals,

where x is a vector of individual demographic predictors (age, gender, and ethnicity). The parameters u_s and v_c are state and county random-effects, $u_s \overset{iid}{\sim} N(0, \sigma_u^2),\ v_c \overset{iid}{\sim} N(0, \sigma_v^2)$. The predicted probability for domain d in county c to be under poverty is

$$\hat{p}_{scd} = \frac{e^{x'_{scd}\hat{\beta} + \hat{u}_s + \hat{v}_c}}{1 + e^{x'_{scd}\hat{\beta} + \hat{u}_s + \hat{v}_c}}.$$

The ACS is a complex survey with a weighting scheme that accounts for not only the differential selection probabilities, but also for the adjustment due to housing unit nonresponse. The final weights are post-stratified to ensure that the weighted estimates of the housing units and persons by age, sex, race and ethnicity conform to estimates from the PEP for the year of interest within the weighting area (United States Census Bureau, 2014). Many studies (Lavalle and Beaumont, 2015), (Pfeffermann, 1993), (Rabe-Hesketh and Skrondal, 2006), and (Verret et al., 2015) have indicated that when a survey design is not simple, the estimators obtained under the assumption of a SRS can be biased. To acknowledge the sample design, it is generally recommended that all design covariates or the survey weights be incorporated into the model. To safeguard against the informativeness of the sample and any misspecification arising while modeling the population model, (Pfeffermann, 1993) has recommended that one introduce the weights in the model specification. (Pfeffermann et al., 1998) have argued that in the case of complex survey data, estimates from a multilevel model may be biased if unequal selection probabilities at any stage of the sampling are not controlled for in the covariates. They have proposed not only to use the survey weights, but also to scale the weights as remedy. (Verret et al., 2015) have mentioned in their simulation study that ignoring the survey design of informative sampling may result in sizable incidence on point estimate bias and mean squared error of the empirical best linear unbiased predictor (EBLUP). (Chambless and Boyle, 1985) cited

[5] Counts from the 2010 Decennial Census indicate that the population of Kalawao, HI comprised 90 residents. Due to its former status as a leper's colony, state law prohibits individuals under sixteen years of age to reside in Kalawao, HI.

by (Pfeffermann, 1993) have demonstrated in an empirical study that when modeling a binary response variable with complex survey design, the intercept estimate obtained under the population model has a large bias. The bias of the intercept would severely impact the quality of predicted probabilities. Thus, they have suggested the use of survey weights to account for the survey design while fitting the model. (Pfeffermann and Sverchkov, 2007) have noted that failure to take into account the sampling scheme produces biased predictors if the response variable and the sampling weights are correlated given the covariates and the random effect. To account for the ACS design, we fit weighted models. We use the ACS individual weights in estimating the weighted models. The models are fitted with SAS PROC GLIMMIX, which uses pseudo maximum likelihood. (Zhang et al., 2014) have pointed out that the WEIGHT statement in SAS GLIMMIX procedure, would prompt SAS to handle the weights variable as frequency weights. Without any scaling of the weights, the standard errors associated with the model parameters would be underestimated, due to the weights summing up to the population size instead of the sample size. Hence, we scale the ACS weights to sum to the sample size rather than the population size. The scaling is done in two ways; first we scale the weights based on the person sample size (PSS) of 5,584,987 people. Second, we scale the weights based on the housing unit sample size (HUSS) of 2,403,614 units. Each ACS weight is multiplied by a scaling factor $\delta_{scdi} = \frac{n_{area}}{\sum\limits_{i \in Area} ACSwgt_{scdi}}$, where n_{area} is the person or housing unit sample size in the geographic boundary of interest and $ACSwgt_{scdi}$ is the ACS weight for person i. The scaling factor (δ) is referred to as the scaling Method 2, in (Pfeffermann et al., 1998), and served as a means of reducing bias created when the area sample sizes (n_{area}) are not large and the sample design is informative. Thus, we scale the weights within ACS weighting areas and strata. Strata are based on the sizes of the small sampling entities. The ACS has 16 strata, each with a different selection probability. The ACS weighting area is either a county or a group of less populous counties. In addition to the two schemes, we also scale the weights within the nation, states, counties, and housing units. The final weights are given by $Weight_{scdi} = \frac{n_{area}}{\sum\limits_{i \in Area} ACSwgt_{scdi}} \times ACSwgt_{scdi}$.

As mentioned in Section 2, all members within a sampled housing unit will be surveyed in the ACS. Since related members of the housing unit are totally dependent in terms of poverty status, the classical assumption of independence that holds between observations is violated. In order to remedy the dependence among members of the household, we create, in addition to 12 scaling of weights schemes, a new weights variable. Under the new weighting scheme, members of a household have equal weights HU_Weight_{scdi}, which sum to one within the household. Thus, $HU\_Weight_{scdi} = \frac{1}{n_h}$, where n_h is the housing unit size (HUS). The household sizes range from one to twenty members. Different scaling schemes are unweighted, weights based on HUS, and 12 other schemes listed below.

Weights scaled based on PSS within:

- Nation
- State
- County
- ACS Weighting Area (ACS_WA)
- Stratum
- Housing Unit

Weights scaled based on HUSS within:

- Nation
- State
- County
- ACS Weighting Area (ACS_WA)
- Stratum
- Housing Unit

We fit MLR models with weights from the list above and calculate the predicted probability being in poverty ($\hat{p}_{scd}$) for each demographic domain within county. Final estimated poverty rates are obtained by aggregating appropriate expected counts over all demographic groups within the area of interest and dividing by the corresponding population. The estimated poverty rate for a given county is:

$$\tilde{p}_{sc} = \frac{\sum_d (N_{scd} \times \hat{p}_{scd})}{\sum_d N_{scd}},$$ where d indexes age/race/ethnicity/sex demographic group, and N_{scd} is the population size for demographic group d in county c and state s.

$\hat{p}_{scd}$ is the predicted probability than an individual in state s, county c, and demographic group d is in poverty. Final estimates of poverty rates for any aggregate combination of the 21 age, six race/ethnic, and two sex groups, at any geographic level, are similarly obtained.

4 Poverty Rate Estimates

Using the weights' scaling schemes described in Section 3, we fit fourteen models, including the unweighted model. The estimated national poverty rates are shown in Table 3. For state and county estimates, only mean absolute differences (MADs) are reported. The estimated state and county poverty rates from ACS, SAIPE and four select MLR models are illustrated with box-plots. The select MLR models are chosen as the ones yielding the smallest four MADs between the ACS direct estimates.

4.1 National Poverty Estimates

Table 1: National Poverty Estimates in Percentage and Absolute Difference

	ALL AGES		AGE 0-17		AGE 5-17	
Weights scaled within	Pov	AD	Pov	AD	Pov	AD
ACS Est. Related Children	15.51	0.00	21.34	0.00	20.37	0.00
SAIPE	15.15	0.36	21.58	0.24	20.22	0.15
Unweighted	15.15	0.36	20.09	1.25	19.34	1.03
Nation based on PSS	15.59	0.08	21.67	0.33	20.82	0.45
Nation based on HUSS	15.58	0.07	21.67	0.33	20.82	0.45
State based on PSS	15.59	0.08	21.59	0.25	20.71	0.34
State based on HUSS	15.59	0.08	21.60	0.26	20.71	0.34
County based on PSS	15.64	0.13	21.64	0.30	20.72	0.35
County based on HUSS	15.70	0.19	21.70	0.36	20.77	0.40
ACS_WA based on PSS	15.88	0.37	21.69	0.35	20.78	0.41
ACS_WA based on HUSS	16.53	1.02	21.84	0.50	20.92	0.55
Stratum based on PSS	15.69	0.18	21.62	0.28	20.71	0.34
Stratum based on HUSS	15.87	0.36	21.69	0.35	20.76	0.39
Household based on PSS	15.08	0.43	20.10	1.24	19.31	1.06
Household based on HUSS	16.90	1.39	20.07	1.27	19.29	1.08
Household based on HUS	16.97	1.46	20.05	1.29	19.31	1.06

AD: Absolute difference, Pov: Poverty rate

National estimates from MLR models are consistent with the ACS national estimate. The national poverty estimates from models with weights scaled within the nation and state are closer to the ACS national estimate for PSS and HUSS. The national poverty estimate for the model weighted within household based on HUS has the largest absolute difference (AD) with the ACS poverty estimate. The AD between these two estimates is 1.46 percentage points (pp), while for the other scaling schemes, the largest absolute difference is 0.37pp (other than three with the largest ADs) and the smallest absolute difference is 0.07pp. The largest three absolute differences (AD) between MLR and ACS are from models with weights scaled based on HUSS and HUS (Household HUS, Household HUSS and ACS_WA HUSS). We also note that the ADs between ACS and MLR for the four selected models (0.07pp, 0.08pp, 0.08pp, and 0.08pp) are smaller than the ACS-SAIPE AD of 0.36pp.

For age-groups 0-17 and 5-17, three of the four selected models are from scaling schemes within state, stratum and county based on the PSS, and one from state based on HUSS. Finally, for age-groups 0-17 and 5-17, the ADs between ACS and SAIPE are lower than those of the four selected MLR models.

4.2 State Poverty Estimates

Table 2: State Mean Absolute Differences ACS, SAIPE and MLRs

Weights scaled within	Mean Absolute Differences					
	ALL AGES		AGE 0-17		AGE 5-17	
	ACS	SAIPE	ACS	SAIPE	ACS	SAIPE
ACS	0.000	0.410	0.000	0.532	0.000	0.652
SAIPE	0.410	0.000	0.532	0.000	0.652	0.000
Unweighted	0.449	0.415	1.353	1.335	1.212	0.987
Nation PSS	0.141	0.529	1.053	0.990	1.072	1.304
Nation HUSS	0.201	0.554	1.121	1.046	1.115	1.346
State PSS	0.136	0.526	1.028	0.961	1.039	1.248
State HUSS	0.179	0.541	1.088	1.007	1.072	1.279
County PSS	0.188	0.577	1.042	0.969	1.051	1.258
County HUSS	0.270	0.654	1.092	1.009	1.091	1.295
ACS_WA PSS	0.401	0.794	1.093	0.984	1.051	1.261
ACS_WA HUSS	1.026	1.419	1.332	1.201	1.206	1.410
Stratum PSS	0.493	0.394	1.343	1.320	1.223	0.989
Stratum HUSS	1.544	1.937	1.717	1.603	1.559	1.394
Household PSS	0.224	0.617	1.031	0.950	1.024	1.228
Household HUSS	0.417	0.811	1.145	1.033	1.082	1.294
Household HUS	1.620	2.013	1.716	1.601	1.547	1.383

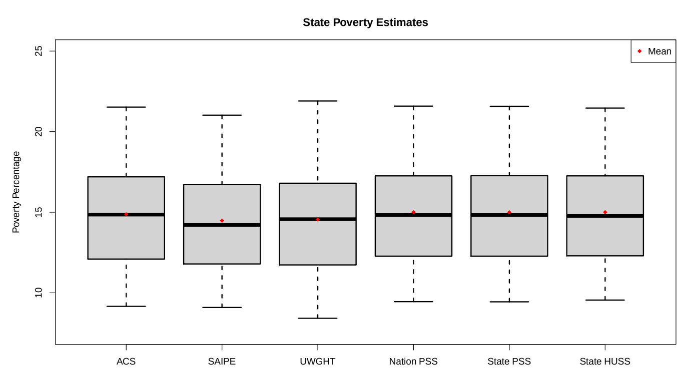

Figure 1: State Pov. Est. ACS, SAIPE, Unweighted, (Nation, State PSS), state HUSS.

Table 2 shows the MADs between state estimates produced from ACS, SAIPE and MLR models by age group. Among estimates for all ages group, the MLR models with the smallest MADs have weights scaled within state, nation, county based on PSS, and one within state based on HUSS. The smallest ACS-MLR MAD is 0.136pp, with the model scaled within state based on PSS. The four selected ACS-MLR's MADs (0.136pp, 0.141pp, 0.179pp and 0.188pp) are smaller compared to the MAD of ACS-SAIPE (0.410pp). The boxplots in Figure 1 show that the distributions of the MLR estimates from the three selected models (nation, state within PSS and state within HUSS) and the unweighted model have similar distributions than those of ACS and SAIPE.

For age-groups 0-17 and 5-17, selected models have weights scaled within state, household, ACS weighting area and county based on PSS. For age-groups 0-17 and 5-17, the MADs of the ACS and SAIPE are smaller than the ones of the ACS and MLR.

4.3 All County Poverty Estimates

Table 3: All Counties Mean Absolute Differences ACS, SAIPE and MLRs

Weights scaled within	Mean Absolute Differences					
	ALL AGES		AGE 0-17		AGE 5-17	
	ACS	SAIPE	ACS	SAIPE	ACS	SAIPE
ACS	0.000	2.605	0.000	6.445	0.000	6.652
SAIPE	2.605	0.000	6.445	0.000	6.652	0.000
Unweighted	3.075	1.825	7.107	4.087	7.235	3.540
Nation PSS	2.479	1.745	6.590	3.561	6.601	3.370
Nation HUSS	3.187	2.046	7.199	3.996	7.227	3.695
State PSS	2.378	1.704	6.500	3.526	6.520	3.335
State HUSS	3.046	1.965	7.069	3.908	7.107	3.618
County PSS	2.165	1.681	6.322	3.423	6.331	3.283
County HUSS	2.850	1.891	6.916	3.744	6.935	3.512
ACS_WA PSS	2.213	1.734	6.348	3.470	6.370	3.303
ACS_WA HUSS	2.983	2.065	7.017	3.916	7.058	3.623
Stratum PSS	3.093	1.834	7.108	4.114	7.245	3.577
Stratum HUSS	3.792	2.476	8.032	4.966	8.168	4.366
Household PSS	2.257	1.718	6.385	3.488	6.405	3.321
Household HUSS	2.935	1.945	6.988	3.848	7.028	3.572
Household HUS	3.807	2.505	8.030	4.958	8.159	4.346

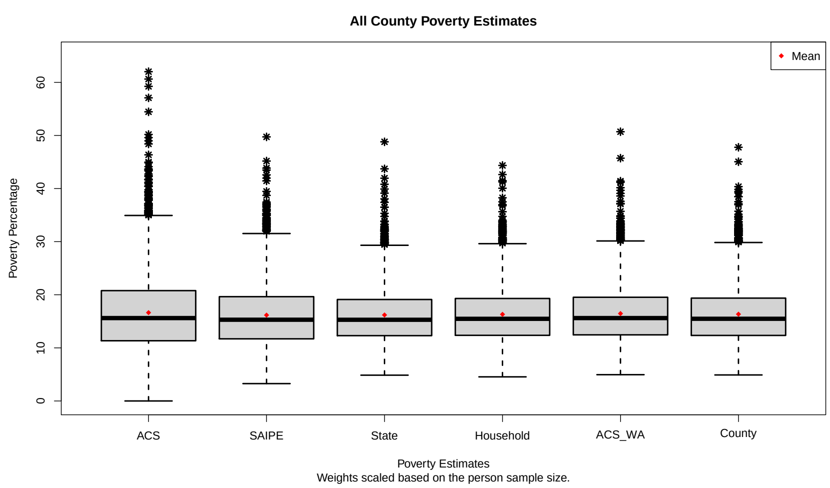

Figure 2: Cty pov. est. ACS, SAIPE, State, HU, ACS_WA, County PSS.

Table 3 shows MADs for county poverty estimates between ACS, SAIPE and MLR approaches by age-group. For all ages group, all four selected MLR models have weights scaled based on PSS within state, county, ACS_WA and household. The ACS-MLR MADs of the four selected models (2.165 pp, 2.213pp, 2.257pp, and 2.378pp) are smaller than the ACS-SAIPE MAD for all ages (2.605 pp). The boxplots in Figure 2 show that the distributions of estimates generated via MLR and SAIPE models have narrow ranges compared with the distribution of the ACS direct estimates.

For age-groups 0-17 and 5-17, the estimates from the four selected models have MADs that are smaller than the MADs of ACS-SAIPE 6.445pp and 6.652pp, respectively. All four of ACS-MLR MADs are produced from models with weights scaled within state, county, ACS weighting area, and household based on the PSS.

4.4 Large County Poverty Estimates

For large published counties, estimates for all ages group from the MLR with the smallest ACS-MLR MADs are from models with weights scaled based on person sample size. The MADs of those estimates are close to 0.5 percentage point and the ACS-SAIPE MAD is 0.86 percentage point. The boxplots in Figure 3 show that all the three methods (ACS, SAIPE and MLR) produce estimates with similar distributions.

Table 4: Large[6] Counties Mean Absolute Differences ACS, SAIPE and MLRs

Weights scaled within	Mean Absolute Differences					
	ALL AGES		AGE 0-17		AGE 5-17	
	ACS	SAIPE	ACS	SAIPE	ACS	SAIPE
ACS	0.000	0.860	0.000	1.949	0.000	2.608
SAIPE	0.860	0.000	1.949	0.000	2.608	0.000
Unweighted	1.310	1.003	3.385	2.413	3.577	2.215
Nation PSS	0.547	0.910	2.969	2.211	2.925	2.468
Nation HUSS	0.982	1.053	3.268	2.427	3.220	2.593
State PSS	0.518	0.907	2.934	2.178	2.907	2.416
State HUSS	0.898	1.025	3.195	2.359	3.164	2.525
County PSS	0.546	0.939	2.955	2.188	2.921	2.431
County HUSS	0.953	1.127	3.252	2.419	3.204	2.592
ACS_WA PSS	0.675	1.103	2.995	2.209	2.962	2.434
ACS_WA HUSS	1.399	1.694	3.385	2.520	3.342	2.654
Stratum PSS	1.324	0.985	3.389	2.423	3.593	2.233
Stratum HUSS	2.261	2.246	4.185	3.118	4.351	2.897
Household PSS	0.560	0.973	2.949	2.189	2.925	2.415
Household HUSS	1.008	1.207	3.252	2.415	3.226	2.557
Household HUS	2.299	2.306	4.182	3.114	4.340	2.889

For age groups 0-17 and 5-17, the four selected are models with weights scaled based on PSS. In addition, the ACS-SAIPE MADs are smaller than the ACS-MLR MADs of the four selected models.

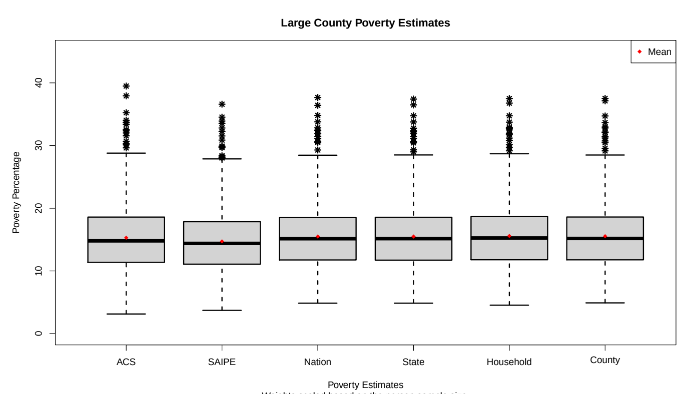

Figure 3: Pov. Est. ACS, SAIPE, Nation, State, HU, County.

[6] Population size greater than or equal to 65,000 in 2014

4.5 Small County Poverty Estimates

Table 5: Small [7]Counties Mean Absolute Differences ACS, SAIPE and MLRs

Weights scaled within	Mean Absolute Differences					
	ALL AGES		AGE 0-17		AGE 5-17	
	ACS	SAIPE	ACS	SAIPE	ACS	SAIPE
ACS	0.000	3.219	0.000	8.030	0.000	8.078
SAIPE	3.219	0.000	8.030	0.000	8.078	0.000
Unweighted	3.696	2.114	8.418	4.676	8.524	4.006
Nation PSS	3.159	2.038	7.866	4.036	7.896	3.687
Nation HUSS	3.963	2.395	8.585	4.547	8.639	4.082
State PSS	3.033	1.984	7.757	4.000	7.793	3.657
State HUSS	3.800	2.295	8.435	4.453	8.496	4.002
County PSS	2.734	1.941	7.508	3.857	7.532	3.582
County HUSS	3.517	2.160	8.208	4.210	8.250	3.835
ACS_WA PSS	2.754	1.956	7.530	3.914	7.571	3.608
ACS_WA HUSS	3.540	2.196	8.297	4.406	8.368	3.964
Stratum PSS	3.715	2.133	8.419	4.708	8.532	4.049
Stratum HUSS	4.330	2.557	9.388	5.615	9.513	4.883
Household PSS	2.853	1.979	7.595	3.944	7.632	3.640
Household HUSS	3.613	2.204	8.305	4.351	8.368	3.929
Household HUS	4.337	2.575	9.387	5.606	9.505	4.858

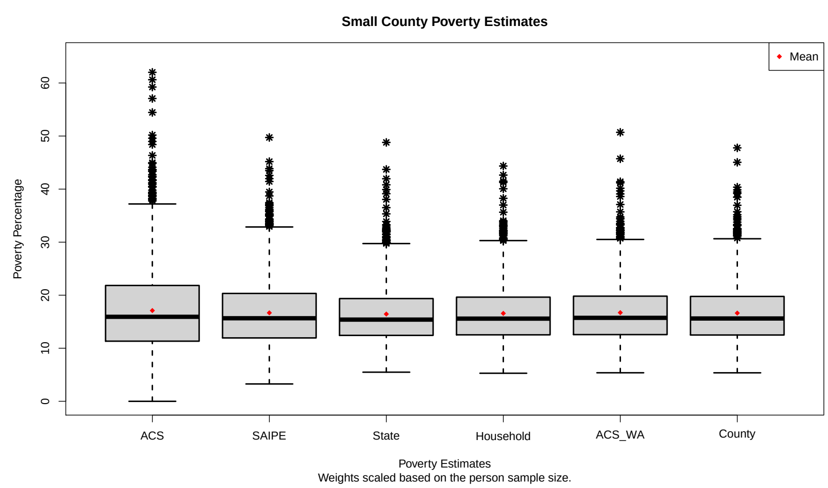

Figure 4: Pov. Est. ACS, SAIPE, MLR: State, HU, ACS_WA, County.

For small unpublished counties, the four selected are models with weights scaled based on PSS. The ACS-SAIPE MADs are greater than those of the four selected MLR models, irrespective of the age group. The boxplots in Figure 4 show the same pattern displayed for all county estimates, whereas the ACS distribution is more spread out with many and larger outliers than the SAIPE and MLR distributions.

5 Discussion

When comparing poverty estimates from MLR to the ACS and SAIPE, the MLR models with weights scaled based on the person sample size almost always yield better estimates in terms of mean absolute differences than models with weights scaled based on the housing

[7] Population size less than 65,000 in 2014

unit sample size and the unweighted model. The MLR model with weights within households depend on the housing unit size (HUS) has the worst performance of all. Given large population sizes (nation, states, large counties), ACS, SAIPE and MLR approaches have similar performance, with the distributions of the estimates displaying the same pattern as seen on Figures 1 and 3.

This empirical study seeks to compare the ACS direct survey estimates, SAIPE area-level model-based estimates, and unit-level logistic mixed effects model-based estimates. Given that the population parameters for different geographies of interest are unknown, we cannot definitively affirm from our results which approach is the best for estimating poverty rates. All poverty rates from ACS, SAIPE and MLR are estimates, and the true poverty rates are unknown. Nevertheless, as stated earlier, when the sample sizes are large, the three approaches behave adequately, yielding estimates with similar patterns. For future studies, one direction to take could be to mimic the ACS dataset from a hypothetical population with known poverty rates and generate data from which we would produce the like-ACS, SAIPE and MLR estimates, and compare them with the hypothetical true ones. Another route could be to extend the demographic predictors used in the MLR method, to include county-level auxiliary data used in the SAIPE model (Supplemental Nutrition Assistance Program (SNAP) participation, PEP, Internal Revenue Service (IRS), and the decennial census 2000 long-form survey) to improve the model's strength when the sample sizes are small. Moreover, thirteen models fitted in this study assumed independence within housing unit members, but related members of housing unit are not independent with respect to poverty status. A potential study could be to consider a unit level model that accounts for clustering effects and some correlation structures within related members of the housing unit.

Acknowledgments

We thank Rachel D. Upton and John Posey for taking time to read the manuscript of this paper.

References

Battese, G. E., Harter, R. M., and Fuller, W. A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83(401):28–36.

Chambless, L. E. and Boyle, K. E. (1985). Maximum likelihood methods for complex sample data: logistic regression and discrete proportional hazards models. *Communications in Statistics - Theory and Methods*, 14(1):1377–1392.

Fay, R. E. and Herriot, R. A. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74(366):269 – 277.

Ghosh, M. and Meeden, G. (1986). Empirical Bayes estimation in finite population sampling. *Journal of the American Statistical Association*, 81(396):1058–1062.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. (2009). *Survey Methodology*. Wiley, New Jersey.

Guadarrama, M., Molina, I., and Rao, J. N. K. (2016). A comparison of small area estimation methods for poverty mapping. *Statistics in Transition new series*, 17(1):41–66.

Hobza, T. and Morales, D. (2016). Empiral best prediction under unit-level logit mixed models. *Journal of Official Statistics*, 32(3):661–692.

Lavalle, P. and Beaumont, J.-F. (2015). Why we should put some weight on weights. *Survey Insights*.

Malec, D. (2005). Small area estimation from the american community survey using a hierarchical logistic model of person and housing units. *Journal of Official Statistics*, 21(3):411–432.

Malec, D., Davis, W. W., and Cao, X. (1999). Model based small area estimates of overweight prevalence using sample selection adjustement. *Statistics in Medicine*, 18:3189–3200.

Pfeffermann, D. (1993). The role of sampling weights when modeling survey data. *International Statistical Institute*, 61(2):317–337.

Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H., and Rasbash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Royal Statistical Society*, 60(1):23–40.

Pfeffermann, D. and Sverchkov, M. (2007). Small-area estimation under informative probability sampling of areas and within the selected areas. *Journal of the American Statistical Association*, 102(480):1427–1439.

Prasad, N. G. N. and Rao, J. N. K. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85(409):163–171.

Rabe-Hesketh, S. and Skrondal, A. (2006). Multilevel modeling of complex survey data. *Royal Statistical Society*, 32(3):805–827.

Rao, J. (2003). *Small Area Estimation*. Wiley, New Jersey.

Skinner, C., Holt, D., and Smith, T. (1989). *Analysis of Complex Survey*. Wiley, Essex.

United States Census Bureau (2014). American community survey design and methodology.

Verret, F., Rao, J., and Hidiroglou, M. A. (2015). Model-based small area estimation under informative sampling. *Survey Methodology*, 41(2):333–347.

West, B. T., Beer, L., Gremel, G. W., Weiser, J., Johnson, C. H., Garg, S., and Skarbinski, J. (2015). Weighted multilevel models: A case study. *American Journal of Public Health*, 105(11):2214–2215.

Zhang, X., Holt, J. B., Lu, H., Wheaton, A. G., Ford, E. S., Greenlund, K. J., and Croft, J. B. (2014). Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *American Journal of Epidemiology*, 179:1025–1033.