Multilevel Regression and Poststratification (MRP) for Small Area Estimation: An Application to Estimate Health Insurance Coverage Using Geocoded American Community Survey Data

Xingyou Zhang[ǂ*], Samuel Szelepka[ǂ], Blandine Bawawana[ǂ], Alfred O. Gottschalck[ǂ]

Abstract

Knowledge of the geographic distributions of population socioeconomic and health outcomes is critical for public social and health policy deliberation, formulation, delivery and program planning and evaluation. More granular or local population socioeconomic and health data are often needed but usually do not exist. A variety of small area estimation techniques have been developed to address this significant data gap. Sociodemographic and health surveys have become routinely geocoded in federal statistical agencies, which means that we could have both individual characteristics of survey respondents from the survey itself but also their geographic context that might have great influence on their individual social, economic and health behaviors. Thus, we are developing and validating an innovative multilevel regression and poststratification (MRP) approach that applies multilevel regression models to geocoded surveys; takes account for both individual characteristics and area level factors at multiple geographic levels; predicts individual-level social, economic and health outcomes in a multilevel modeling framework; and estimates the geographic distributions of population socioeconomic and health outcomes. We applied this innovative multilevel approach for small area estimation using geocoded American Community Survey (ACS) data. We demonstrate that MRP provides a flexible statistical linkage and modeling platform that makes full use of geocoded ACS data and available geodemographic data to generate small area estimates of percentages of the population without health insurance coverage. We will also compare our model-based health insurance coverages with those based on the current SAHIE model and direct ACS survey estimates.

Key Words: Multilevel Regression and Poststratification (MRP), Small area estimation, Unit-level Logistic Mixed Model, ACS, SAHIE, Parametric Bootstrapping

1. Introduction

Model-based small area estimation techniques have been widely used to produce small area statistics to meet the local data needs for public policy or program planning, development and evaluation. The U.S. Census Bureau has two well established model-based small area data products: Model-based Small Area Income & Poverty Estimates (SAIPE) for school districts, counties, and states (https://www.census.gov/did/www/saipe/) and Model-based Small Area Health Insurance Estimates (SAHIE) for counties and states (https://www.census.gov/did/www/sahie/).

SAIPE takes an empirical Bayesian approach and SAHIE employs a fully-Bayesian approach for model fitting and estimation. From a modeling perspective, both SAIPE and SAHIE are employing an area-level (also called Fay-Herriot (FH) Model) small area model framework. Specifically, current SAHIE implements an area-level model with errors-in-

Disclaimer: The views expressed on statistical issues in this paper are those of the authors and not necessarily those of the U.S. Census Bureau or Economic Research Service, U.S. Department of Agriculture.

[ǂ]U.S. Census Bureau, Washington, DC 20233.

*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.

JSM 2017 - Survey Research Methods Section

variables via a full-Bayesian approach to produce county-level health insurance coverage for detailed demographic domains. Like other area-level models for small area estimation, sampling variances of direct survey estimates have been assumed to be known but the small sample sizes make the typical estimators of sample variance estimation very problematic. In practice, this area-level modeling approach has limited flexibility in producing reliable estimates for potential smaller domains, such as sub-county (e.g. census tract) or sub-demographic groups within a county (e.g. race-specific estimates). In order to create race-specific small area estimates (SAEs), a separate state-level model has to be constructed and fitted for SAHIE. This modeling strategy would result in a potential inconsistency between direct state-level model-based estimates and those state-level estimates aggregated from county-level model-based estimates. When new demographic domain SAEs of SAHIE outcomes are requested, we have to repeat the entire modeling process and deal with the following inconsistency between SAEs from different small area model specifications. Furthermore, we could not make full use of all the survey data in hand for those domains with very small sample sizes, because direct estimates for these domains at the area level of interest make no sense at all.

On the other hand, the unit-level modeling approach aims to predict the expected responses (or values) of the target population using observed survey data at the level of survey respondent unit. For SAHIE, we could predict each individual's health insurance status based on who they are (i.e., by sex, age, race/ethnicity, education, and poverty status) and where they live (e.g. neighborhood (census tract), county, state). With a fitted unit-level model, we could obtain the SAEs for any demographic domain at any geographic level of interest (Guadarrama, Molina, and Rao 2016). This is only limited by the availability of detailed geo-demographic population counts. In the United States, population counts by sex, age, race/ethnicity are available at census block level from the Decennial census and at the county-level from annual Census Vintage population estimates. Since census block is the basic (smallest) geographic unit for census geography in the United States, a census block measure could be conveniently aggregated to any geographic level of interest for any demographic domains by sex, age and race/ethnicity.

However, the flexible unit-level small area model approach has its own challenges in practice. First, the unit-level model assumes that the survey design is ignorable after conditioning on the covariates included in the model (Hobza and Morales 2016, Gelman 2007). Thus, survey weights are often excluded in the model fitting process, which may produce estimates with substantial bias; other complex survey design components, such as clustering or stratification, are rarely accounted for in the model fitting process. Second, for a generalized linear mixed model (GLMM), such as a logistic mixed model for binary data, the mean squared error estimation often involves complex numerical optimization of the maximum likelihood function without analytical closed forms and is computationally prohibitive in terms of time and resources in practice. The computation of mean squared errors (MSEs) of SAEs is not as intuitive as those based on the linear or linear mixed models. The empirical best linear unbiased prediction (EBLUP) method proposed by Prasad and Rao is not appropriate for common non-continuous outcomes in practice (Prasad and Rao 1990). For unit-level linear mixed models, Molina and Rao applied an intuitive parametric bootstrapping approach for estimating MSEs (Molina and Ra 2010) and Molina et al. employed a full-Bayesian approach that obtains the MSEs for poverty indicators (Molina, Nandram, and Rao 2014) . For unit-level non-linear mixed models, parametric bootstrapping was introduced to obtain the MSEs for SAEs (Gonzalez-Manteiga et al. 2007, Hobza and Morales 2016, Molina, Saei, and Lombardia 2007); their parametric bootstrapping makes a prediction for each individual in the target population and then takes random samples from the population with predicted outcomes and refits the

JSM 2017 - Survey Research Methods Section

bootstrapped samples. The entire bootstrapping computation process is very time consuming. For a very large population like our case, the entire population in the United States, the parametric bootstrapping of this kind is so time-consuming that it becomes infeasible in practice.

This study aims to deal with the two practical challenges for unit-level non-linear models for small area estimation: incorporating survey weights and survey design effects and obtaining the appropriate MSEs via appropriate parametric bootstrapping. This paper is organized as follows. Section 2 briefly summarizes basic ideas of unit-level Multilevel Regression and Poststratification (MRP) for small area estimation, and introduces the logistic mixed model for binary data in American Community Survey (ACS). Three specific models are constructed to account for survey design: a model without individual ACS respondent weights, a model with individual ACS respondent weights, and a model with individual ACS respondent weights while incorporating survey design effects. We also provided more details for the modified parametric bootstrapping for estimating MSEs of final SAEs and the basic strategy to reduce computational time for parametric bootstrapping. Section 3 presents the model-based small area estimation results based on 2015 ACS for the proportions of population without health insurance coverage for population under 65 years old in the United States. The MRP SAEs are compared with direct ACS estimates and SAHIE estimates at the national, state and county levels. Section 4 gives a brief discussion and concludes with future work to further improve the developed methodology.

2. Multilevel Regression and Poststratification (MRP) for Small Area Estimation

2.1 Basic Ideas of MRP

The unit level MRP approach takes the basic unit-level small area model assumptions: The main assumption for the unit-level model is that estimates for a model fitted to the sample should be close to the estimates that would be obtained if that model were fitted to the entire population (Little 2012). In other words, the model constructed for survey data is also applicable to the target population for the survey. Overall, we use a multilevel regression model to quantify the relationships between the modeling outcome and its related individual characteristics and contexts that hold for all the units in the survey and apply this quantitative relationship for all the units in the population of interest. Specifically, the MRP takes two basic steps:

1) Construct and fit a survey sampling unit-level multilevel model to a survey
   - $Y(C) = Y(C_o + C_u)$
   - $Y|C_o = X\beta + G + R$

For a population outcome of interest $Y(C)$, it has two exclusive components: the observed in the sampled population (survey) $Y(C_o)$ and the unknown in the non-sampled population $Y(C_u)$. A generalized linear mixed model can be fitted to the outcome of interest $Y|C_o = X\beta + G + R$ for the sample population in a survey, where X is the known characteristics matrix for each individual in the survey sample; β is the estimable regression coefficients vector associated with X; G is the G side (also called structured) random effects matrix; and R is the residual random effects matrix.

2) Make predictions for each individual in the target population from the fitted model based on the survey ($\hat{Y}(C|C_o) = X\hat{\beta} + \hat{G}$).

JSM 2017 - Survey Research Methods Section

Usually, the finite population estimator $\hat{Y}(C)$ combines both the observed $Y(C_o)$ from the survey (sampled population) and the out-of-sample predictions $\hat{Y}(C_u|C_o)$ for the unknown population ($\hat{Y}(C) = Y(C_o) + \hat{Y}(C_u|C_o)$). When the sampling fraction for a domain is not small, then the finite population estimator is robust under model misspecification. More often, the sampling fraction for a domain is small for most surveys, and the finite population estimator could be approximated by the full model-based prediction for the entire finite population ($\hat{Y}(C) \approx \hat{Y}(C|C_o)$). We could aggregate individual-level predictions of the target population to produce final SAEs of interest.

The original idea of MRP was introduced by Little from a Bayesian modeling perspective (Little 1993). Gelman and Little further developed this idea and demonstrated its great statistical power and flexibility in model-based predictions for the outcomes of interest for detailed geo-demographic groups (Gelman and Little 1997). Gelman and his colleagues have applied MRP for a variety of social and political outcomes, such as voter turnout and vote choice (Ghitza and Gelman 2013, Park, Gelman, and Bafumi 2004) and public opinion on the death penalty (Shirley and Gelman 2015). The MRP or similar approach has been adopted for various applications for population health outcomes (Congdon 2009, Malec et al. 1997, Twigg, Moon, and Jones 2000, Wang et al. 2015, Yu et al. 2007, Zhang et al. 2014) and socioeconomic outcomes, such as poverty indicators (Elbers, Lanjouw, and Lanjouw 2003, Haslett and Jones 2005, Molina, Nandram, and Rao 2014, Molina and Ra 2010) and unemployment (Molina, Saei, and Lombardia 2007).

2.2 ACS Data

ACS is the largest demographic survey in the United States and has become the main database to provide the annual updates of population and housing characteristics, including the health insurance coverage status of interest in this study. The ACS independent housing unit address samples are randomly selected for each of the 3,142 counties and county equivalents in the U.S., including the District of Columbia. In 2015 ACS, 5,404,658 sampled individuals with a valid health insurance status. Our target population is the under 65 year-old population; thus, 4,399,937 individuals under 65 years old with a valid health insurance status were included in this study. 2015 ACS was sampled from all 50 states and all 3,142 counties. The state-level sample size has a mean of 86,273 and ranges from 8,467 to 488,453 with a median of 62,566, and the county-level sample size has a mean of 1,400 and range from 4 to 135,288 with a median of 449.

Table 1. 2015 ACS sample size distribution for the population under 65 years old

Geography	N	Min	Q1	Median	Q3	Max	Mean
State	51	8,467	23,385	62,566	111,869	488,453	86,273
County	3,142	4	218	449	1,119	135,288	1,400

2.3 Unit-level logistic mixed model for health insurance status

In ACS, the binary variable of interest for health insurance coverage status in the ACS sample takes the values of one (an individual without health insurance coverage) and zero (an individual with health insurance coverage). Thus, logistic mixed models are constructed to generate county-level sub-demographic domain estimates of interest: percentages of uninsured population for children 0-17 years old, adults 18-64 years old as

JSM 2017 - Survey Research Methods Section

well as population under 65 years old. Let y_{dij} denote an individual (j) response to health insurance coverage status from an age-sex-race/ethnicity specific population group (d) in county (i), i=1, …, 3,142, and j=1, .., n_{di}, the sampled population with county (i) from population group (d). Thus, conditionally on p_{di}, the probability of an individual without health insurance coverage from an age-sex-race/ethnicity specific demographic group (d) in county (i), y_{dij}'s are independent Bernoulli random variables with $P\left(y_{dij} = 1 \middle| p_{di}\right) = p_{di}$.

$$logit(p_{di}) = x\beta + u = \beta_0 + age + sex + race + u_i \quad (1)$$

where β_0 is the regression intercept; age is the age group variable with 16 categories, sex is the sex variable with two categories, $race$ is the race/ethnicity group variable with six categories; county specific random effects are normally distributed with zero means and variances $\sigma^2 : u_i \sim iidN(0, \sigma^2)$. The detailed 16 age groups (years) include 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, and 60-64; the six race/ethnicity groups include non-Hispanic white alone, black alone, American Indian and Alaska Native alone, Asian and Pacific Islander alone, other races and Hispanic. Thus, there are possible 192 (16 x 2 x 6) demographic groups within a county.

The likelihood for our logistic mixed model (equation 1) is given by

$$L(\beta, \sigma^2 | y) = \prod_{i=1}^{3142} \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} \frac{exp\{y_{ij}(x\beta + u)\}}{1 + exp\{x\beta + u\}} \times \frac{e^{-\frac{u_i^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} du_i \quad (2)$$

where n_i is the number of sampled individuals by ACS in county (i). We take a frequentistic approach to fit the above logistic mixed model using SAS proc GLIMMIX. The above likelihood function (equation 2) cannot be evaluated in closed form. Numerical methods must be used to estimate the model parameters that maximize the likelihood function above. SAS proc GLIMMIX applies a pseudo-likelihood approach to fit the above logistic mixed model. The likelihood formulation of equation 2 ignores sample selection probabilities and sample dependencies other than within-county correlations, such as unequal sampling probabilities, within household clustering effects, and other dependencies introduced by all other weight adjustment procedures. We will address both selection probability and dependency via incorporating survey weights and design effects into model fitting (see details in Section 2.4).

The predicted probability ($\hat{p}_{di}$) of an individual without health insurance coverage from demographic group (d) in county (i) is:

$$\hat{p}_{di} = \frac{e^{\hat{\eta}}}{1 + e^{\hat{\eta}}} = \frac{e^{X\hat{\beta} + \hat{u}_i}}{1 + e^{X\hat{\beta} + \hat{u}_i}} \quad (3)$$

where the estimated linear predictor $\hat{\eta} = X\hat{\beta} + \hat{u}_i$, X is the covariate matrix, including intercept ones, age, sex and race/ethnicity.

In order to generate the final small area estimates, we need to sum the population weighted probabilities for all or some of demographic groups (d) within a county. The predicted proportion of individuals without health insurance coverage in county (i) under an empirical Bayesian predictor is

$$\hat{P}_i = N_i^{-1}\left(\sum_{j=1}^{n_{di}} y_{dij} + \sum_{i=1}^{d} (N_{di} - n_{di}) * \hat{p}_{di}\right) \quad (4)$$

where n_{di} is the number of sampled individuals in ACS from demographic group (d) in county (i) , d is the number of demographic groups in county (i), and $N_i = \sum_{i=1}^{d} N_{di}$. We obtain the population count N_{di} demographic group (d) in county (i) from the U.S. Census

JSM 2017 - Survey Research Methods Section

Bureau's postcensal population estimates data. In practice, we often assume that the population counts by age, sex and race/ethnicity at the county level are known without error, and this a strong assumption.

Since the number of sampled individuals are much smaller than the population size of the corresponding demographic group, we have an empirical Bayesian estimator (equation 3) reduced to a multilevel regression estimator (equation 5):

$$\hat{P}_i = N_i^{-1}(\sum_{j=1}^{n_{di}} y_{dij} + \sum_{i=1}^{d}(N_{di} - n_{di}) * \hat{p}_{di}) \approx N_i^{-1}(\sum_{i=1}^{d} N_{dis} * \hat{p}_{di}) \quad (5)$$

The variance of $\hat{P}_{cs}$ is

$$Var(\hat{P}_i) = N_i^{-2}\left(Var\left(\sum_{i=1}^{d} N_{di} * \hat{p}_{di}\right)\right) \quad (6)$$

$Var(\hat{P}_i)$ has a non-linear non-closed form expression, thus we apply a parametric bootstrapping to approximately estimate it. The predicted MSEs is the square root of its estimated variance ($Var(\hat{P}_i)$). Unit-level logistic mixed model could be conveniently specified and fitted using SAS proc GLIMMIX, a SAS statistical procedure for generalized linear mixed models. By default, the GLIMMIX procedure estimates the parameters of logistic mixed models by applying pseudo-likelihood techniques (Breslow and Clayton 1993, Wolfinger and Oconnell 1993).

2.4 Model fitting with survey weights

The best way to incorporate survey weights in unit-level small area models is an open research question. Unit-level small area models often ignore the survey weights in modeling fitting, since they assume survey design is ignorable after conditioning the covariates in the unit-level models. Actually, this is a general controversial issue: should we implement regression modeling for survey data with or without survey weights? (Gelman 2007). In practice, adding weight should cause additional model fitting complexity for generalized linear models. The SAS proc GLIMMIX procedure has a weight statement to account for individual weights into its pseudo-likelihood optimization. Since the proc GLIMMIX procedure WEIGHT statement treats weight variables as frequency weights, direct use of survey weights in the WEIGHT statement will substantially underestimate standard errors associated with model parameters. Thus, we should rescale the original ACS survey weights to actual ACS sample size. We compared different rescaling schemes for small area estimation using ACS and compared their performance (see details in Blandine et al. report 2017). We selected a within-state rescaling weighing scheme, and it defined the new weight for a ACS respondent as follows (equation 7):

$$W\_new(i) = \frac{W_{acs(i)}}{\sum W_{acs(i)}} S\_acs \quad (7)$$

where $W_{acs(i)}$ is the individual ACS survey weight, $\sum W_{acs(i)}$ is the sum of all individual ACS survey weights within a state, and S_acs is the ACS sample size within a state. In model fitting with the proc GLIMMIX prodedure, we could add the WEIGHT statement with the rescaled weights $W\_new(i)$ and account for the ACS weighting impact into logistic mixed model fitting.

Survey design effects for survey data should not be ignorable in a regression model. For ACS, within household correlations are common and strong for many socioeconomic outcomes of interest, such as poverty and health insurance overage status. We take

JSM 2017 - Survey Research Methods Section

advantage of the property of frequency weights in proc GLIMMIX procedure WEIGHT statement, and we could rescale the original ACS weights into an effective sample size within a state to address the ACS survey design effects. State effective sample size is the state actual ACS sample divided by the state-level design effects. Thus, the new weight that accounts for both survey weight and survey design effects is defined as follows (equation 8):

$$W^*_{\_new(i)} = \frac{W_{acs(i)}}{\sum W_{acs(i)}} \left(\frac{S_{acs}}{DE_{acs}}\right) \qquad (8)$$

Here DE_{acs} is state-level design effect for the outcome of interest. Designs varies by outcome as well as geography. For 2015 ACS, its state-level design effects for health insurance coverage ranges from 1.96 to 6.01 and has a mean of 3.28 with a median of 3.22; its state-level design effects for poverty status ranges from 3.65 to 6.59 and has a mean of 4.78 with a median of 4.73. The larger design effect for poverty status is expected, since poverty status is defined for the entire household unit, while persons within a household may have different health insurance status based on whether they are an adult or child.

With consideration of survey weights and design effects, we fit three unit-level logistic mixed models for small area estimation:

- Logistic mixed model without ACS survey weights (UNW)
- Logistic mixed model with ACS survey weights ($W_{\_new(i)}$) that are rescaled to within state ACS sample size (WGT)
- Logistic mixed model with ACS survey weights ($W^*_{\_new(i)}$) that are rescaled to within state ACS effective sample size (WTD)

2.5 Mean square error (MSE) estimation via parametric bootstrapping

For small area estimation, a Bayesian approach would be very intuitive to produce final valid predicted MSEs (Molina, Nandram, and Rao 2014). But the Bayesian approach for model fitting via MCMC could be computationally very intensive and make it almost infeasible in practice for large datasets like ACS. In addition, it is also very difficult to fit unit-level logistic models with individual survey respondent weights in a Bayesian setting. Thus, we fit our logistic mixed models under a fequentistic approach using SAS proc GLMMIX procedure. This approach is pretty efficient and could conveniently incorporate survey weights and survey design effects as we described in section 2.4. It takes about two hours to fit the logistic mixed model using the entire ACS dataset on our current hardware. But this efficiency and convenience comes with a price that proc GLIMMIX could not directly produce the predicted MSEs for the final SAEs of interest. Actually, it is always challenging to obtain valid MSE estimates for small area estimators under a frequentistic modeling framework, especially for the logistic mixed model estimator because of its non-linearity nature.

With the breakthroughs in computer simulation in the 1990s and cheaper and easier accessibility to supercomputing power in 2000s, computationally intensive resampling methods, including bootstrapping, and Bayesian computation, it has become more and more popular for quantifying the statistical properties of non-linear statistics in practice. Parametric Bootstrapping has been used in small area estimation to obtain the predicted MSEs (Gonzalez-Manteiga et al. 2007, Hall and Maiti 2006, Hobza and Morales 2016, Pfeffermann and Correa 2012, Pfeffermann and Tiller 2005, Molina and Ra 2010). None of these parametric bootstrapping algorithms have considered accounting for individual survey weights in model fitting as well as survey design effects. These common parametric

JSM 2017 - Survey Research Methods Section

Bootstrapping algorithms in small area estimation involve generating the virtual original sample and refitting the small area models to estimate the statistical uncertainties in final SAEs. Since the virtual sample comes from the predicted target population, it could have a very different demographic composition from the original sample for a small area. The parametric bootstrapping has to ignore the original survey weights and ignore survey weighting in model fitting. If survey weighting has to be considered in model fitting, new survey weights must be recomputed for the virtual bootstrap sample, but this could become extremely complicated. This, actually, is one main reason that bootstrapping methods has limited use in complex survey data.

In order to keep flexibility of incorporating the ACS survey weights and survey design effects, we revised the current parametric bootstrapping algorithm as follows:

Step 1. Fit the logistic mixed model to the original ACS sample and calculate the predicted linear predictor ($\hat{\eta}_{ij} = x'_{ij}\hat{\beta} + z'_i\hat{u}$) and predicted standard error ($\hat{\sigma}_{ij}$) using SAS proc GLIMMIX procedure.

Step 2. Take a random sample for ($\hat{\eta}^*_{ij}$) from $N(\hat{\eta}_{ij}, \hat{\sigma}^2_{ij})$ and calculate predicted probability $\hat{p}^*_{ij} = \frac{e^{\hat{\eta}^*_{ij}}}{1+e^{\hat{\eta}^*_{ij}}}$.

Step 3. Generate the bootstrap binary outcome $\hat{y}^*_{ij}|\hat{p}^*_{ij} = (1,0)$ for each original ACS respondents.

Step 4. Refit the logistic mixed model with ($\hat{y}^*_{ij}$) and calculate predicted probability ($\hat{p}^{b*}_{ij}$).

Step 5. Calculate the predicted target outcome $\hat{P}^{\mathrm{b}}_{\mathrm{i}} = (\sum_{\mathrm{i=1}}^{\mathrm{k}} \mathrm{n_{ij}}\, \hat{p}^{b*}_{\mathrm{ij}})/\mathrm{N_i}$

Step 6. Repeat step 2 to step 5 many times (b=1, 2, …, B) and the number of bootstrapping B=1,000 in this study. The final point estimate variance: $var(\hat{P}_i) = \frac{1}{B}\sum_{b=1}^{B}(\hat{P}^b_i - \overline{\hat{P}^b_i})^2$ and $\overline{\hat{P}^b_i} = \frac{1}{B}\sum_{b=1}^{B}\hat{P}^b_i$.

Parametric bootstrapping still involves a model refitting which still could be very time-consuming for ACS data. Thus, we split the entire ACS data by Census Division and run the same logistic mixed model for each division's ACS sub-dataset. This modeling strategy reduced the computation time from more than 2,000 hours to less than 100 hours for a 1,000 replicate bootstrapping.

3. Results

This study presents three final outcomes of interest: the percentages of uninsured estimates for populations 1) 0-64 years, 2) 0-17 years old children, and 3) 18-64 years old adults. All the results are based on the parametric bootstrapping with B=1,000 samples. First, we compare the model-based point estimates with reliable direct ACS estimates at national, state and county levels, then we compare their county-level standard error estimates.
Table 2 presents the estimates from different methods at the national level. Logistic mixed models with survey weights produced almost the same point estimates as ACS direct survey estimates. The logistic mixed model without survey weights has produced similar estimates as current SAHIE models.

JSM 2017 - Survey Research Methods Section

Table 2. 2015 ACS sample sizes and ACS model-based uninsured population estimates (%)

	Age Groups		
	0-64 years	0-17 years	18-64 years
ACS Sample Size	4,399,937	1,122,680	3,277,257
ACS	11.26	4.82	13.63
SAHIE	10.92	4.78	13.21
MRP(UNW)	10.58	4.86	12.69
MRP(WGT)	11.23	4.84	13.59
MRP(WTD)	11.25	4.83	13.62

Table 3. Mean absolute differences between model-based and 2015 ACS direct estimates for state-level uninsured estimates (%)

		Age Groups		
Method	N	0-64 years	0-17 years	18-64 years
SAHIE	51	0.32	0.12	0.41
MRP(UNW)	51	1.09	0.98	1.37
MRP(WGT)	51	0.81	0.90	0.99
MRP(WTD)	51	1.21	1.04	1.46

Table 4. Mean absolute differences between model-based and 2015 ACS direct estimates for county-level uninsured estimates (%)

		Age Groups		
Method	N	0-64 years	0-17 years	18-64 years
SAHIE	811	1.17	1.40	1.36
MRP(UNW)	811	1.33	1.90	1.66
MRP(WGT)	811	1.02	1.77	1.36
MRP(WTD)	811	1.68	1.88	2.08

Table 3 shows that logistic mixed model estimators have produced larger mean absolute differences than current SAHIE models at state level. At the county level (Table 4), logistic mixed model with survey weights (MRP(WGT)) and current SAHIE model have similar mean absolute differences for those 811 larger counties (county population >=65,000 in 2015).

Table 5 presents the distributions of predicted MSEs of the county-level uninsured population estimates. Model-based estimates have much smaller standard errors than direct ACS estimates as expected. Logistic mixed model without survey weights (MRP(UNW) and with survey weights (MRP(WGT)) have similar standard errors as current SAHIE model for large age groups (0-64 years and 18-64 years), but they have much smaller predicted standard errors than SAHIE model for small age group (0-17 years old). Logistic mixed model with survey weight and adjusting ACS design effects has produced consistent smaller standard errors than current SAHIE model.

JSM 2017 - Survey Research Methods Section

Table 5. Standard errors for county-level uninsured estimates (%)

Age Group	Methods	N	Min	Q1	Median	Q3	Max	Mean
0-64 years	ACS	3131	0.14	1.37	2.27	3.74	32.82	2.91
	SAHIE	3141	0.14	0.53	0.74	0.93	1.89	0.74
	MRP(UNW)	3142	0.09	0.42	0.68	0.98	5.27	0.79
	MRP(WGT)	3142	0.10	0.42	0.76	1.13	3.86	0.84
	MRP(WTD)	3142	0.09	0.25	0.42	0.65	2.51	0.49
0-17 years	ACS	2713	0.06	1.19	2.33	4.41	34.38	3.48
	SAHIE	3141	0.11	0.62	0.85	1.22	3.35	0.96
	MRP(UNW)	3141	0.05	0.22	0.33	0.50	3.33	0.42
	MRP(WGT)	3141	0.06	0.22	0.37	0.56	2.28	0.43
	MRP(WTD)	3141	0.06	0.14	0.21	0.32	1.47	0.25
18-64 years	ACS	3131	0.17	1.58	2.68	4.38	32.05	3.36
	SAHIE	3141	0.18	0.67	0.94	1.20	2.21	0.95
	MRP(UNW)	3142	0.11	0.49	0.81	1.17	6.48	0.93
	MRP(WGT)	3142	0.13	0.50	0.91	1.34	4.97	1.00
	MRP(WTD)	3142	0.10	0.29	0.50	0.79	2.98	0.58

4. Discussion and conclusions

The unit-level logistic mixed model small area estimator proposed in this study has great flexibility in incorporating survey weights and design effects. These logistic mixed models has only three individual covariates (age, sex and race/ethnicity) and a county-level random effect. Despite their simple format, these unit-level logistic mixed model estimators take a bottom-top approach in geo-demographic domain aggregation and could produce accurate and consistent point estimates across all demographic domains at all geographic levels. Logistic mixed models with survey weights produce more consistent small area estimates while compared to direct ACS estimates. If necessary, the logistic mixed models could include county-level covariates such as those in current SAHIE area-level models.

We fit these logistic mixed models via parametric bootstrapping that could be conveniently and efficiently implemented in SAS using the GLIMMIX procedure and routine data and random number simulation steps. This is important for routine data production in small area estimation practice. The unit-level logistic mixed model estimators has consistently produced small predicted standard errors than current SAHIE area-level models. The unit-level logistic mixed model with survey weights while adjusting ACS survey design effects has the smallest prediction standard errors. The modified parametric bootstrapping could incorporate survey weights and design effects, and it does not bootstrap model regression coefficients and county-random effects directly and avoids the ignorance of potential correlation between model parameters of fixed and random effects. The parametric bootstrapping approach in this study assumes each random sample for specific demographic group for a county are independently drawn from other random samples for other demographic groups within the same county. Further research is greatly needed to compare different parametric bootstrapping methods for logistic mixed model-based small area estimators.

JSM 2017 - Survey Research Methods Section

Our proposed logistic mixed models with normally distributed county random effects are commonly used in small area estimation for binary outcome data. The county random effects take a flexible and simple normality assumption, and it might become a controversial aspect of the methodology (Diallo and Rao 2014). Some research shows that the non-normality assumption for random effects does not make much difference in real applications (McCulloch and Neuhaus 2011). In a small area estimation context, random effects play a critical role in introducing local variation into the model outcomes of interest; therefore the choice of distribution for random effects may be crucial. Diallo and Rao show that that the normality assumption for unit-level small area model is relatively robust (Diallo and Rao 2014). In the United States, the children's health insurance program (CHIP) has increased health insurance coverage for certain children populations, thus child and adult specific county random effects might be needed to better catch these program impacts.

We take a frequentistic approach to develop and fit the unit-level logistic mixed models and to obtain the model parameters for both fixed and random effects. We adopt a "Bayesian" simulation approach (bootstrapping) for final statistical inference on small area estimates under these fitted models. This approach could make full use of the efficiency of model development and fitting in a freqentistic framework and the flexibility of inference in a Bayesian framework. The combination of freqentistic and Bayesian methods in our model make our MRP approach more practical while keep statistical inference accuracy and flexibility.

Acknowledgements

The authors would like to thank Jerry J. Maples at the U.S. Census Bureau for his helpful comments and suggestions.

References

Breslow, N. E., and D. G. Clayton. 1993. "Approximate Inference in Generalized Linear Mixed Models." *Journal of the American Statistical Association* no. 88 (421):9-25.

Congdon, P. 2009. "A multilevel model for cardiovascular disease prevalence in the US and its application to micro area prevalence estimates." *International Journal of Health Geographics* no. 8.

Diallo, M. S., and J.N.K. Rao. 2014. Small Area Estimation of Complex Parameters Under Unit-level Models with Skew-Normal Errors. Paper read at JSM, at Boston.

Elbers, C., J. O. Lanjouw, and P. Lanjouw. 2003. "Micro-level estimation of poverty and inequality." *Econometrica* no. 71 (1):355-364.

Gelman, A. 2007. "Struggles with survey weighting and regression modeling." *Statistical Science* no. 22 (2):153-164.

Gelman, A., and T. C. Little. 1997. "Poststratification into many categories using hierarchical logistic regression " *Survey Methodology* no. 23 (2):127-135.

Ghitza, Y., and A. Gelman. 2013. "Deep Interactions with MRP: Election Turnout and Voting Patterns Among Small Electoral Subgroups." *American Journal of Political Science* no. 57 (3):762-776.

JSM 2017 - Survey Research Methods Section

Gonzalez-Manteiga, W., M. J. Lombardia, I. Molina, D. Morales, and L. Santamaria. 2007. "Estimation of the mean squared error of predictors of small area linear parameters under a logistic mixed model." *Computational Statistics & Data Analysis* no. 51 (5):2720-2733.

Guadarrama, M., I. Molina, and J. N. K. Rao. 2016. "A Comparison of Small Area Estimation Methods for Poverty Mapping." *Statistics in Transition new series* no. 17 (1):41-66.

Hall, P., and T. Maiti. 2006. "On parametric bootstrap methods for small area prediction." *Journal of the Royal Statistical Society Series B-Statistical Methodology* no. 68:221-238.

Haslett, S., and G Jones. 2005. "Small area estimation using surveys and censuses: some practical and statistical issues." *Statistics in Transition new series* no. 7 (3):541-555.

Hobza, T., and D. Morales. 2016. "Empirical Best Prediction Under Unit-Level Logit Mixed Models." *Journal of Official Statistics* no. 32 (3):661-692.

Little, R. J. 2012. "Calibrated Bayes, an Alternative Inferential Paradigm for Official Statistics Rejoinder." *Journal of Official Statistics* no. 28 (3):367-372.

Little, R. J. 1993. "Post-Stratification - a Modelers Perspective." *Journal of the American Statistical Association* no. 88 (423):1001-1012.

Malec, D., J. Sedransk, C. L. Moriarity, and F. B. LeClere. 1997. "Small area inference for binary variables in the National Health Interview Survey." *Journal of the American Statistical Association* no. 92 (439):815-826.

McCulloch, C. E., and J. M. Neuhaus. 2011. "Misspecifying the Shape of a Random Effects Distribution: Why Getting It Wrong May Not Matter." *Statistical Science* no. 26 (3):388-402.

Molina, I., B. Nandram, and J. N. K. Rao. 2014. "Small Area Estimation of General Parameters with Application to Poverty Indicators: A Hierarchical Bayes Approach." *Annals of Applied Statistics* no. 8 (2):852-885.

Molina, I., and J. N. K. Ra. 2010. "Small area estimation of poverty indicators." *Canadian Journal of Statistics-Revue Canadienne De Statistique* no. 38 (3):369-385.

Molina, I., A. Saei, and M. J. Lombardia. 2007. "Small area estimates of labour force participation under a multinomial logit mixed model." *Journal of the Royal Statistical Society Series a-Statistics in Society* no. 170:975-1000.

Park, D. K., A. Gelman, and J. Bafumi. 2004. "Bayesian multilevel estimation with poststratification: State-level estimates from national polls." *Political Analysis* no. 12 (4):375-385.

Pfeffermann, D., and S. Correa. 2012. "Empirical bootstrap bias correction and estimation of prediction mean square error in small area estimation." *Biometrika* no. 99 (2):457-472.

Pfeffermann, D., and R. Tiller. 2005. "Bootstrap approximation to prediction MSE for state-space models with estimated parameters." *Journal of Time Series Analysis* no. 26 (6):893-916.

Prasad, N. G. N., and J. N. K. Rao. 1990. "The Estimation of the Mean Squared Error of Small-Area Estimators." *Journal of the American Statistical Association* no. 85 (409):163-171.

Shirley, K. E., and A. Gelman. 2015. "Hierarchical models for estimating state and demographic trends in US death penalty public opinion." *Journal of the Royal Statistical Society Series a-Statistics in Society* no. 178 (1):1-28.

JSM 2017 - Survey Research Methods Section

Twigg, L., G. Moon, and K. Jones. 2000. "Predicting small-area health-related behaviour: a comparison of smoking and drinking indicators." *Social Science & Medicine* no. 50 (7-8):1109-1120.

Wang, Y., N. A. Ponce, P. Wang, J. D. Opsomer, and H. Yu. 2015. "Generating Health Estimates by Zip Code: A Semiparametric Small Area Estimation Approach Using the California Health Interview Survey." *Am J Public Health* no. 105 (12):2534-40.

Wolfinger, R., and M. Oconnell. 1993. "Generalized Linear Mixed Models - a Pseudo-Likelihood Approach." *Journal of Statistical Computation and Simulation* no. 48 (3-4):233-243.

Yu, H., Y. Y. Meng, C. A. Mendez-Luck, M. Jhawar, and S. P. Wallace. 2007. "Small-area estimation of health insurance coverage for California legislative districts." *Am J Public Health* no. 97 (4):731-7.

Zhang, X., J. B. Holt, H. Lu, A. G. Wheaton, E. S. Ford, K. J. Greenlund, and J. B. Croft. 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." *Am J Epidemiol* no. 179 (8):1025-33.