Combining Probability and Non-Probability Samples Using Small Area Estimation

N. Ganesh[1], Vicki Pineau[1], Adrijo Chakraborty[1], J. Michael Dennis[1]
[1]NORC at the University of Chicago, 4350 East-West Highway, Suite 800, Bethesda, MD 20814

Abstract
Given the high cost associated with probability samples, there is increasing demand for combining larger non-probability samples with probability samples to increase sample size for low incidence studies and/or key analytic subgroups. Given bias and coverage error inherent in non-probability samples, use of traditional weighted survey estimators for data from such surveys may not be statistically valid. In this paper, we discuss the use of small area models and estimation methods to combine a probability sample with a non-probability sample assuming the (smaller) probability sample yields unbiased estimates. We consider two distinct small area models: (a) Fay-Herriot model with the probability sample point estimate as the dependent variable and the non-probability sample point estimate as a covariate in the model, and (b) Bivariate Fay-Herriot model that jointly models the probability sample point estimate and the non-probability sample point estimate, and accounts for the bias associated with the non-probability sample.

Key Words: AmeriSpeak Panel, composite estimator, EBLUP, non-probability sample, Small Area Estimation, web survey

1. Introduction

Given the increasing cost associated with fielding a probability-based sample, some studies use a combination of probability and non-probability samples to meet the study requirements. Furthermore, some studies target low incidence populations or require large oversamples of specific subpopulations that make it costly to only field a probability-based sample. A major concern with fielding a non-probability sample is how to account for the bias associated with survey estimates produced using a non-probability sample. In this paper, we discuss using small area models to derive model-based estimates that combine both the probability sample estimate and the non-probability sample estimate to produce unbiased estimates for the target population of interest.

There are several approaches to combining a probability sample with a non-probability sample. Some approaches use explicit statistical models to derive model-based estimates while other methods use statistical models to derive survey weights (using calibration or propensity methods) for the combined sample. Elliott (2009) proposed a method to derive pseudo-weights for the non-probability sample when there are shared covariates between the non-probability and probability samples, and when those covariates are predictive of the probability of selection or substantive variable of interest. This approach provides a weighting solution for combining the two sample sources.

Wang et. al. (2015) used a multilevel regression model with post-stratification (MRP) to predict the outcome of the 2012 Presidential election; the only data source (Xbox user data) in this example was a non-probability sample. Their approach involved first fitting a logistic regression model to predict the proportion of the vote for both (Obama and Romney) major party candidates, and then modeling the proportion of vote for Obama given that the respondent supports a major party candidate. They used the MRP model to generate predicted estimates for the proportion of Obama's vote share for ~176,000 cross-classified cells, and then aggregated those cell level estimates to estimate the proportion of Obama's vote share for each state and the entire nation.

Fahimi et. al. (2015) recommended including calibration variables that differentiate the selection and response mechanism associated with the probability and non-probability samples as a way to adjust for the bias associated with the non-probability sample. In addition to raking the probability and non-probability samples to standard socio-demographic variables (such as age, gender, education, race/Hispanic ethnicity, and geography), Fahimi et. al. (2015) suggested calibrating the non-probability sample using the following variables:

1. Number of online surveys taken in a month
2. Hours spent on the Internet in a week for personal needs
3. Interest in trying new products before other people do;
4. Time spent watching television in a day;
5. Using coupons when shopping; and
6. Number of relocations in the past 5 years.

Benchmarks for the above variables would be obtained from the associated probability sample.

Our approach to combining the probability and non-probability samples is similar to Wang et. al. We use small area estimation models to: (a) model the probability sample estimate as a dependent variable with the non-probability sample estimates as covariates in the model, and (b) jointly model (with a bivariate model) the probability and non-probability sample estimates as dependent variables, and account for the bias associated with the non-probability sample estimates. In Section 2, we provide details on our data application. In Section 3, we discuss the two small area models for combining probability and non-probability samples. In Section 4, we discuss results and compare the two models against a standard weighting approach similar to Fahimi et. al. Finally, in Section 5, we provide some concluding remarks.

2. Data Application

NORC conducted a Food Allergy Survey on behalf of Northwestern University using NORC's AmeriSpeak® Panel and SSI's non-probability web panel. The main focus of the research was to measure the adult and child prevalence of self-reported and doctor-diagnosed food allergies, both current and outgrown, allergy reactions, experiences in allergy treatments, events coinciding with development or outgrowing a food allergy, and perceived risks associated with food allergies. For the data application that we considered for this paper, we only analyzed data for adults 18+ years. There were 7,218 adult survey completes from the AmeriSpeak Panel and 33,331 adult survey completes from the SSI non-probability web panel.

Funded and operated by NORC at the University of Chicago, AmeriSpeak® is a probability-based panel sample designed to be representative of the U.S. household population. Randomly selected U.S. households are sampled with a known, non-zero probability of selection from the NORC National Frame, and then contacted by U.S. mail, telephone interviewers, overnight express mailers, and field interviewers (face-to-face). AmeriSpeak panelists participate in NORC studies or studies conducted by NORC on behalf of NORC's clients.

The sample frame for the AmeriSpeak is the NORC National Frame, an area probability sample frame constructed by NORC providing sample coverage of 97 percent of U.S. households. The NORC National Frame itself contains almost 3 million households, including over 80,000 rural households added through in-person listing of households that were not recorded on the USPS Delivery Sequence File (see Pedlow and Zhao, 2016).

Once the sample is selected from the National Frame, AmeriSpeak Panel sample recruitment is a two-stage process: initial recruitment using less expensive methods and then non-response follow-up using personal interviewers. For the initial recruitment, sample addresses are invited to join AmeriSpeak by visiting the panel website AmeriSpeak.org or by telephone (in-bound/outbound). As of July 2017, the AmeriSpeak Panel weighted AAPOR 3 response rate was 33.5% (Montgomery, Dennis, and Ganesh, 2017). For further details on AmeriSpeak, please see Dennis (2017) and http://amerispeak.norc.org/about-amerispeak/panel-design/.

For our analysis of the Food Allergy study data, we used the following substantive variables:

- Ever had a food allergy
- Peanut allergy
- Milk allergy
- Either biological parent has a food allergy
- Either biological parent has an environmental allergy

3. Small Area Models

In this section, the two modeling approaches are discussed for the proportion of adults who "ever had a food allergy". Similar models were fitted for the other substantive variables of interest (see Section 2 for the five substantive variables that we analyzed). The first model referred to as the **Fay-Herriot model** (Fay and Herriot, 1979) involves modeling the domain-level point estimate from the probability sample (AmeriSpeak) for proportion of adults who "ever had a food allergy". The domains are a cross-classification of socio-demographic variables. For example, as domains for this data application, we used a cross-classification of:

- Age (18-34 years, 35-49 years, 50-64 years, 65+ years),
- Education (Some college or less, college graduate or higher),
- Race/Hispanic ethnicity (Hispanic, non-Hispanic Black, non-Hispanic All Other), and
- Gender (male, female)

Thus, we created 48 domains, and generated the point estimates from the probability sample for each of the 48 domains. The choice of domains was motivated by "sufficient" sample size for the probability sample adult prevalence rate in each domain but also to capture the variation in the adult prevalence rates across domains. Ideally, domains would be selected such that there is minimal variation in the prevalence rates within a domain and large between domain variation in the prevalence rates.

When using the Fay-Herriot model, we modeled as the dependent variable the domain-level point estimate from the AmeriSpeak sample for "ever had a food allergy" with the following variables as potential explanatory variables:

- Fixed effects for race, age, gender, and education categories.
- Non-probability sample point estimates at the domain level for all five measures of interest (see Section 2).

The point estimates obtained from the probability and non-probability samples were derived using final survey weights that were raked to external population benchmarks from the Current Population Survey. Final survey weights were raked to age, gender, education, race/Hispanic ethnicity, and Census Division. In addition, the non-probability sample weights were calibrated to benchmarks obtained from the probability sample for three additional raking variables corresponding to "early adopter of technology". These early adopter of technology questions were thought to differentiate the probability and non-probability sample respondents (these additional variables are motivated by Fahimi et. al., 2015).

The second model referred to as the **Bivariate Fay-Herriot model** (Rao, 2003) involves jointly modeling the domain-level point estimates from the probability sample (AmeriSpeak) and non-probability sample for the proportion of adults who "ever had a food allergy". The domains that we used were the same 48 domains as previously described. For the Bivariate Fay-Herriot model, as explanatory variables, we only used fixed effects for the probability and non-probability samples for race, age, gender, and education categories (i.e., we did not include any other explanatory variables from other national surveys).

3.1 Fay-Herriot Model

Typically, when modeling proportions, the point estimates are transformed using an arcsine transformation (see Jiang et al., 2001). The arcsine transformation preserves the bounds of 0 and 1 for a proportion. Thus, the modeled estimates for "ever had a food allergy" are guaranteed to be between 0 and 1. If, instead, the untransformed point estimates are modeled, the estimation methodology described below may yield estimates outside the bounds of 0 and 1. The transformed point estimate for "ever had a food allergy" is given by:

$$y_d^P = 2\sin^{-1}\sqrt{z_d^P}, \qquad (1)$$

where z_d^P is the point estimate from the probability sample for the proportion of adults who "ever had a food allergy", and $d=1,...48$ indexes the domains (the superscript of 'P' denotes the probability sample).

The arcsine transformed point estimates for all domains were modeled using the Fay-Herriot model:

$$y_d^P = \alpha_d + x_d'\gamma + v_d + e_d^P \qquad (2)$$

where α_d is a domain-level fixed effect and is parametrized as

$$\alpha_d = race_d + age_d + sex_d + educ_d \qquad (3)$$

where $race_d, age_d, sex_d, educ_d$ are fixed effects for the race, age, sex, and education categories associated with domain d, and x_d is a vector of covariates for domain d. The set of possible covariates were the non-probability sample point estimates for all five measures of interest (see Section 2). Among the set of possible fixed effects and covariates, the "best" set of fixed effects and covariates were selected based statistical significance (p-value less than 0.1). For simplicity, we did not include other covariates at the domain-level from the Census and large national surveys such as the American Community Survey, National Health Interview Survey, Current Population Survey, etc. Including covariates from such surveys should improve the model fit and explanatory power of the model. Furthermore, the non-probability samples estimates that were included in the model as explanatory variables are measured with error. One future modification we plan to consider is to incorporate measurement error associated with the covariates (see Ybarra and Lohr, 2008).

In the above model, the v_d's are random effects that capture the domain-specific effect not captured by the regression component $\alpha_d + x_d'\gamma$; and e_d^P is the sampling error associated with y_d^P, the transformed probability sample point estimate. Standard distributional assumptions of normality were assumed for the domain-specific random effects, i.e., $v_d \sim N(0, \sigma_v^2)$, where σ_v^2 is an unknown variance parameter. Furthermore, the v_d's and the e_d^P's are pairwise mutually independent, and $e_d^P \sim N(0, \psi_d^P)$. As mentioned previously, since e_d^P is the sampling error, ψ_d is simply the sampling variance associated with the transformed point estimate, and is estimated by

$$\psi_d^P = \frac{1 + CV_d^2}{n_d^P}, \qquad (4)$$

where n_d^P is the sample size associated with the probability sample for domain d, and $1 + CV_d^2$ is the weighting effect and is used as an approximation of the design effect for domain d, and CV_d is the co-efficient of variation of the final survey weights for the probability sample for domain d. The above estimate for the sampling variance follows from a Taylor-Series approximation for the variance of y_d^P. We also note that the sampling variance ψ_d^P is assumed to be known without error even though it is estimated as described above. This assumption of known sampling variance is typically made in small area models (see Rao, 2003).

The model given by (2) can also be expressed as

$$y_d^P = \vartheta_d + e_d^P \qquad (5)$$

where ϑ_d is the true (but unknown) value for the arcsine transformed proportion of adults who "ever had a food allergy" in domain d, and

$$\vartheta_d = \alpha_d + x_d'\gamma + v_d \qquad (6)$$

The true arcsine transformed proportion excludes the sampling error (e_d^P) associated with the transformed point estimate. Typically, ϑ_d is the parameter of interest in a given small area model. However, in our context, ϑ_d is the true arcsine transformed proportion of adults who "ever had a food allergy" in domain d, but we are interested in the untransformed proportion, and thus, after deriving the model-based estimate for ϑ_d, that estimate would have to be transformed back to obtain an estimate for the proportion of adults who "ever had a food allergy".

The unknown parameters $(race_d, age_d, sex_d, educ_d, \gamma, \sigma_v^2)$ in model (2) were estimated by the maximum likelihood estimator. An estimate $\tilde{\vartheta}_d$ of ϑ_d was derived using a best linear unbiased prediction (BLUP; Rao, 2003) approach. Since the BLUP $\tilde{\vartheta}_d$ depends on the unknown variance parameter σ_v^2, an empirical BLUP (referred to as an EBLUP), $\hat{\vartheta}_d$, is obtained by substituting the maximum likelihood estimate $\hat{\sigma}_v^2$ for σ_v^2.

The EBLUP $\hat{\vartheta}_d$ corresponds to the arcsine-transformed proportion of adults who "ever had a food allergy". Thus, the estimate for the proportion of adults who "ever had a food allergy" in each domain was obtained by transforming back from arcsine to a proportion. That is,

$$\hat{p}_d = \sin^2\left(\frac{\hat{\vartheta}_d}{2}\right). \tag{7}$$

Finally, to derive a national-level estimate, we aggregated the domain-level estimates. That is, the national-level estimate for the proportion of adults who "ever had a food allergy" is given by

$$\hat{p} = \sum_{d=1}^{48} \frac{N_d}{N} \hat{p}_d \tag{8}$$

where N_d is the number of adults 18+ years in domain d, as estimated using the Current Population Survey, and N is the total number of adults 18+ years.

An estimate for the variance of $\hat{\vartheta}_d$ for each domain was estimated using Rao (2003), and then the variance of $\hat{p}_d$ was derived by accounting for the arcsine-transformation. The variance of $\hat{p}$ was derived assuming the $\hat{p}_d$'s were independent across domains. We assumed this for simplicity, however, this assumption can easily modified to incorporate the covariances among the $\hat{p}_d$'s.

3.2 Bivariate Fay-Herriot Model

Similar to the Fay-Herriot model, we transformed the point estimates. However, for the bivariate model, we transformed both the probability and non-probability sample point estimates. The transformed point estimates for "ever had a food allergy" are given by:

$$y_d^P = 2\sin^{-1}\sqrt{z_d^P}, \tag{9}$$

$$y_d^{NP} = 2\sin^{-1}\sqrt{z_d^{NP}}, \tag{10}$$

where z_d^P, z_d^{NP} are the point estimates from the probability sample and the non-probability sample for the proportion of adults who "ever had a food allergy", and as before, $d=1,...48$ indexes the domains (the superscript of 'NP' denotes the non-probability sample).

The arcsine transformed point estimates for all domains were modeled using the Bivariate Fay-Herriot model:

$$y_d^P = \alpha_d + v_d + e_d^P \tag{11}$$
$$y_d^{NP} = \alpha_d + \beta_d + v_d + e_d^{NP} \tag{12}$$

where α_d, β_d are domain-level fixed effects and are parametrized as

$$\alpha_d = race_d + age_d + sex_d + educ_d \tag{13}$$

$$\beta_d = race_d^b + age_d^b + sex_d^b + educ_d^b \quad (14)$$

where $race_d, age_d, sex_d, educ_d$ are fixed effects for the race, age, sex, and education categories associated with domain d, and $race_d^b, age_d^b, sex_d^b, educ_d^b$ are fixed effects parameters for the bias associated with the non-probability sample for the race, age, sex, and education categories for domain d. Among the set of possible fixed effects, the "best" set of fixed effects were selected based statistical significance (p-value less than 0.1). As mentioned previously, for simplicity, we did not include covariates from other national-level surveys.

The v_d's are domain level random effects; and e_d^P, e_d^{NP} are the sampling errors associated with y_d^P, y_d^{NP}. Similar to the Fay-Herriot model, $v_d \sim N(0, \sigma_v^2)$, where σ_v^2 is an unknown variance parameter. Furthermore, the v_d's and the e_d^P's and the e_d^{NP}'s are pairwise mutually independent, and $e_d^P \sim N(0, \psi_d^P), e_d^{NP} \sim N(0, \psi_d^{NP})$. The sampling variances ψ_d^P, ψ_d^{NP} were estimated using a similar methodology as (4).

The model given by (11)-(12) can also be expressed as
$$y_d^P = \vartheta_d + e_d^P \quad (15)$$

$$y_d^{NP} = \vartheta_d + \beta_d + e_d^{NP} \quad (16)$$

where ϑ_d is the true (but unknown) value for the arcsine transformed proportion of adults who "ever had a food allergy" in domain d, and β_d is a domain-level fixed effect that captures the domain-level bias associated with the non-probability sample, and

$$\vartheta_d = \alpha_d + v_d \quad (17)$$

Similar to the Fay-Herriot model, the unknown parameters in model (11)-(12) were estimated by the maximum likelihood estimator. An empirical BLUP, $\hat{\vartheta}_d$, was used to estimate ϑ_d in (17). The estimate for the proportion of adults who "ever had a food allergy" in each domain and at the national-level was derived similar to (7)-(8). An estimate for the variance was also derived using an identical methodology as used for the Fay-Herriot model.

4. Results

We found significant reductions in standard error for all five substantive variables of interest for both small area models relative to estimation using only the AmeriSpeak probability sample. Table 1 provides the national-level adult prevalence rates and the associated 95% confidence intervals for the five substantive variables of interest that we analyzed. In Table 1,

- "AmeriSpeak Sample" refers to the national-level point estimate and associated confidence interval from only the AmeriSpeak sample. The final survey weights for the AmeriSpeak sample were raked to age, gender, education, race/Hispanic ethnicity, and Census Division.

- "Non-Probability Sample" refers to the national-level point estimate and associated confidence interval from only the non-probability sample; we generated a pseudo confidence interval using a Taylor Series method and assuming that the non-probability sample is like a "probability sample". The final

survey weights for the non-probability sample were raked to age, gender, education, race/Hispanic ethnicity, and Census Division; the non-probability sample was also calibrated to the probability sample point estimate for "early adopter of technology".

- "Combined Sample" refers to the national-level point estimate and associated confidence interval from the combined probability (AmeriSpeak) and non-probability samples; we generated a pseudo confidence interval using a Taylor Series method and assuming that the combined sample is like a "probability sample". The combined sample weights were derived by computing an "optimal" composition factor that minimized the mean squared error when combining the weights associated with the probability and non-probability samples; mean squared error was minimized over a key set of survey variables.

- "Fay-Herriot" refers to the national-level model-based estimate and associated confidence interval under the Fay-Herriot model.

- "Bivariate Fay-Herriot" refers to the national-level model-based estimate and associated confidence interval under the Bivariate Fay-Herriot model.

For all five substantive variables, the model-based estimate obtained under the Fay-Herriot model and the Bivariate Fay-Herriot model were similar to the AmeriSpeak sample point estimate. For most variables, the point estimate obtained using the non-probability sample was very different compared to the AmeriSpeak sample point estimate. Given how we derived the combined sample weights, the point estimate obtained from the combined sample was always in-between the AmeriSpeak sample point estimate and the non-probability sample point estimate. The Fay-Herriot model point estimates had similar half-width confidence intervals as the combined sample point estimates, and significantly shorter half-width confidence intervals when compared to those from the AmeriSpeak sample. The Bivariate Fay-Herriot model point estimates had uniformly shorter half-width confidence intervals compared to the half-width confidence intervals associated with the Fay-Herriot model point estimates and the non-probability sample point estimates.

We verified the normality assumption for the residuals for the Fay-Herriot and Bivariate Fay-Herriot models; a Q-Q plot of the standardized residuals and a plot of the residuals against the predicted residuals indicated the normality assumption was reasonable for both models and all five substantive variables of interest. Also, a normality test rejected the hypothesis that the data were not normally distributed.

Figure 1 gives a plot of the domain-level ratio of the standard error of the AmeriSpeak sample point estimate for "ever had a food allergy" and the standard error for "ever had a food allergy" under the Fay-Herriot model for adults 18+ years. Ratios greater than one indicate that the standard error under the Fay-Herriot model is smaller than the standard error associated with the AmeriSpeak sample point estimate. Domains are ordered based on the AmeriSpeak domain sample size. Domains with small sample sizes see significant reduction in standard error under the Fay-Herriot model. Domains with much larger sample sizes, have similar standard errors under both the design-based approach (AmeriSpeak sample) and the Fay-Herriot model. Typically, when using small area models, domains with smaller sample sizes see a much larger reduction in standard error compared to domains with larger sample sizes. The median ratio of the standard errors across all domains was 2.1; for 34 out of the 48 domains, the ratio of the standard errors was greater

than 1.5. Similar results were observed for the other four substantive variables of interest. The Bivariate Fay-Herriot model resulted in even smaller standard errors when compared to the standard errors obtained under the Fay-Herriot model (not shown here).

Figure 2 gives a plot of the domain-level differences between the AmeriSpeak sample point estimate for "ever had a food allergy" and the model-based estimate for "ever had a food allergy" under the Fay-Herriot model for adults 18+ years. Once again, domains are ordered based on the AmeriSpeak domain sample size. As the domain sample sizes increase, the difference between the AmeriSpeak sample point estimate and the model-based estimate tends toward zero. This observation is typical of small area models. The mean and median difference across all domains was approximately 0. Similar results were observed for the other four substantive variables of interest.

Table 1: Comparison of national-level adult prevalence rates and associated 95% confidence intervals.

| Variable	AmeriSpeak Sample	Non-Probability Sample	Combined Sample	Fay-Herriot	Bivariate Fay-Herriot
Ever had an food allergy	21.6 ± 1.5	28.1 ± 0.6	23.9 ± 1.0	21.1 ± 1.0	21.3 ± 0.6
Peanut allergy	1.7 ± 0.4	5.1 ± 0.3	2.9 ± 0.3	1.4 ± 0.2	1.4 ± 0.1
Milk allergy	4.4 ± 0.5	6.5 ± 0.3	5.2 ± 0.4	4 ± 0.4	4.1 ± 0.1
Biological parent has a food allergy	11.9 ± 1.1	14.7 ± 0.5	12.9 ± 0.7	11.5 ± 0.8	11.5 ± 0.5
Biological parent has an environmental allergy	28.6 ± 1.9	29.6 ± 0.7	28.9 ± 1.2	28.2 ± 1.2	27.9 ± 0.6
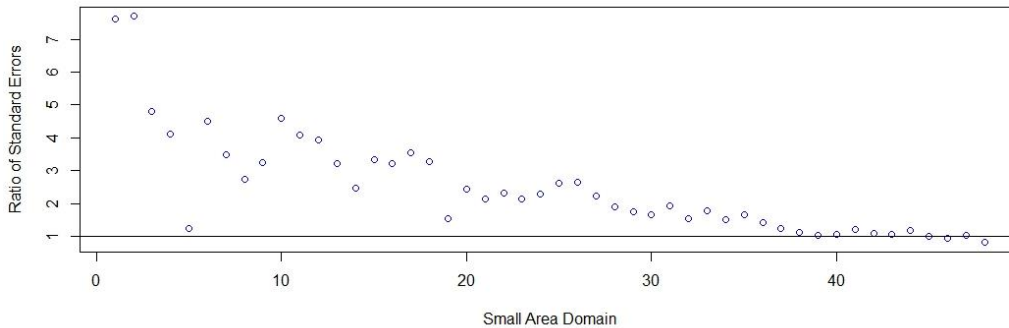
Figure 1: Plot of the domain-level ratio of the standard error of the AmeriSpeak sample point estimate for "ever had a food allergy" and the standard error for "ever had a food allergy" under the Fay-Herriot model for adults 18+ years. Domains are ordered based on the AmeriSpeak domain sample size.
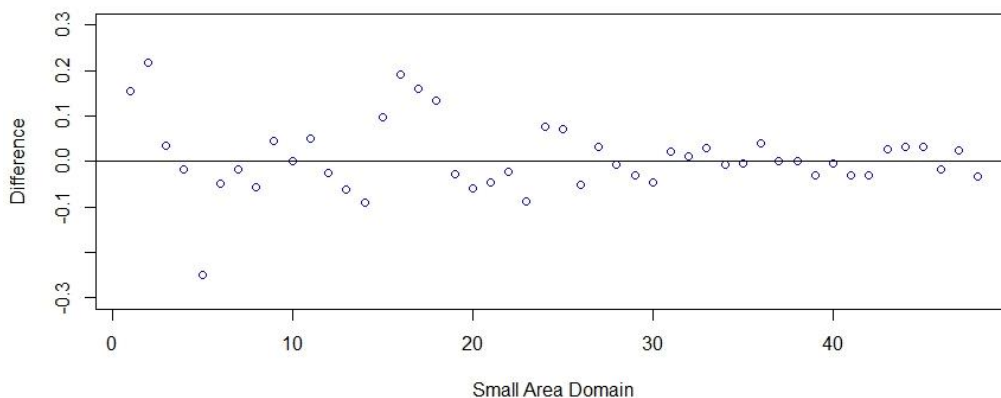
Figure 2: Plot of the domain-level difference between the AmeriSpeak sample point estimate for "ever had a food allergy" and model-based estimate for "ever had a food allergy" under the Fay-Herriot model for adults 18+ years. Domains are ordered based on the AmeriSpeak domain sample size.

5. Conclusion

We used small area models to combine a probability sample from the AmeriSpeak Panel with a non-probability sample in order to generate unbiased estimates for the target population of interest. The model-based estimates are unbiased only if the point estimates from the probability sample for each domain are also unbiased. Both small area models that we considered produced reasonable reductions in standard error relative to the standard error associated with the AmeriSpeak sample point estimate, especially for domains with smaller sample sizes. At the domain-level, the AmeriSpeak sample standard errors were approximately 2 times larger than the standard errors obtained under the Fay-Herriot model, and approximately 2.9 times larger than the standard errors obtained under the Bivariate Fay-Herriot model. There were also significant reductions in standard error for the national-level estimates under both models. Under the model assumptions, the model-based estimates are unbiased, but this depends on the crucial assumption that the probability sample point estimates are unbiased for each domain.

Our proposed small area method for combining a probability sample with a non-probability sample is still experimental and under development. We also note there are potential improvements that could be incorporated. These include incorporating measurement error for the non-probability sample point estimates in the Fay-Herriot model, including other explanatory variables at the domain-level from large national surveys, and choosing the domains using a more objective criterion.

References

Dennis, J.M. 2017. Technical overview of the AmeriSpeak Panel. Available at
http://amerispeak.norc.org/research

Elliott, M.R. 2009. Combining data from probability and non-probability samples using pseudo-weights. Survey Practice, v. 2 (6).

Fahimi, M., Barlas, F.M., Thomas, R.K., and Buttermore, N. 2015. Scientific surveys based on incomplete frames and high rates of nonresponse. Survey Practice, v. 8 (6).

Fay, R.E., and Herriot, R.A. 1979. Estimates of income for small places: An application of James-Stein procedures to Census data, v. 74 (366), pp. 269-277.

Jiang, J., Lahiri, P., Wan S. and Wu, C. 2001. Jackknifing in the Fay-Herriot model with an example. Proceedings of the Seminar on Funding Opportunity in Survey Research, pp. 98-104.

Montgomery, R., Dennis, J.M., and Ganesh, N. 2016. Response rate calculation methodology for recruitment of a two-phase probability-based panel: The case of AmeriSpeak. Available at http://amerispeak.norc.org/research

Pedlow, S. and Zhao, J. Bias reduction through rural coverage for the AmeriSpeak Panel. Presented at AAPOR 2016, Austin, TX. Available at http://amerispeak.norc.org/research

Rao, J.N.K. 2003. Small area estimation. Hoboken, NJ: John Wiley & Sons.

Wang, W., Rothschild, D., Goel, S., and Gelman A. 2015. Forecasting elections with non-representative polls. International Journal of Forecasting, v. 31 (3), pp. 980-991.

Ybarra, L.M.R. and Lohr, S.L. 2008. Small area estimation when auxiliary information is measured with error. Biometrika, v. 95 (4), pp. 919-931.