Improved Sampling Weight Calibration by Generalized Raking with Optimal Unbiased Modification

A.C. Singh, N Ganesh, and Y. Lin
NORC at the University of Chicago, Chicago, IL 60603
singh-avi@norc.org; nada-ganesh@norc.org; yongheng-lin@norc.org

Abstract

Traditional methods for sampling weight adjustment involve weighting class adjustment for nonresponse bias reduction, followed by post-stratification (raking-ratio or regression) adjustment for coverage bias reduction, and then trimming (or winsorization) of extreme weights for variance reduction followed by final post-stratification to meet desired control totals and for further variance reduction. Using calibration methods (or generalized raking), the nonresponse weight adjustment can be considerably simplified and improved by relying on external control totals for nonresponse predictors (or auxiliary variables) instead of frame level information for the full sample, resulting typically in a rich set of auxiliary variables. Moreover, this allows in general for a simplification of the process by eliminating the post-stratification step after the nonresponse bias adjustment as long as the set of auxiliary variables used is deemed adequate for post-stratification. However, in calibration methods, there is no built-in mechanism for ensuring variance reduction although generally it does lead to variance reduction. Besides, the nature of the commonly used method of weight trimming before final post-stratification is ad hoc and likely to introduce bias although it is expected to reduce variance. We propose modeling to smooth extreme weights instead of trimming, and introduce new (super) stratum-specific scale parameters in the calibration (or generalized raking) model to capture possibly varying design characteristics by strata or super-strata. The new calibration model with extra parameters maintains approximate unbiasedness of calibration estimators for which the new parameters are estimated outside the calibration equations by minimizing the generalized variance of key study variables or alternatively the unequal weighting effect for simplicity. Using a hypothetical calibration problem based on the 2011 National Immunization Survey (NIS) public use file data, an illustrative example of the proposed method is presented.

Key Words: Weighting Class; Raking Ratio; Weight Trimming; Model-based Weight Smoothing; Stratum-Specific Scale Parameters.

1. Introduction

Many surveys use a weighting class approach when adjusting for bias due to sample attrition at different stages until the subset of complete respondents is reached. For example, in the case of random digit dial telephone surveys, sample attrition may be due to non-resolution of released telephone numbers, nonresponse to screener interview, ineligibility after screening, and nonresponse to the main interview; and for list frames, sample attrition may be due to incorrect location, noncontact, nonresponse and ineligibility to screener, and finally refusal. The weighting class method uses a cross-classification of variables which often results in insignificant interaction terms being included in the model, which, in turn, tends to increase both the variability of the adjusted weights and variance for key survey estimates. Moreover, the weighting class method requires frame level covariate information (i.e., for the variables used to create the non-response adjustment cells) for both respondents and non-respondents—here the term nonresponse is used in a generic sense to refer to sample attrition due to various reasons. In practice, however, there is often a paucity of frame level information about useful nonresponse predictors or covariates for nonrespondents, which affects the validity of bias adjustment at each stage of sample attrition.

An alternative to the weighting class method for nonresponse adjustment was proposed by Folsom and Singh (2000; see also the WTADJUST procedure of SUDAAN v10, RTI International, 2010) which generalizes the weight calibration approach for post-stratification (Deville and Särndal, 1992) to handle the nonresponse bias problem such that the adjustment factor is greater than 1. The calibration controls for auxiliary variables needed for model fitting are allowed to be random because they can be obtained from a larger sample or fixed if they are obtained from census or administrative sources. The main advantage of using a calibration method for fitting nonresponse models is that one can work with only information about model covariates from the sample of respondents as long as calibration control totals for the target population are available. Moreover, it allows for retaining only those covariates in the model that are significant (typically main effects and lower order interactions), and thus reduces the variability of the adjusted weights by fitting a parsimonious model. In addition, this method easily allows for continuous variables in addition to the usual categorical variables to be included in the model.

In this paper, we propose to enhance the above calibration approach to nonresponse adjustment for handling all types of sample attrition before reaching a complete case through a single step of weight adjustment. The basic idea is that although models for the attrition bias adjustment may be different at different stages of attrition due to differences in the underlying phenomena, the rich set of covariates available from the final complete cases contains good predictors of sample attrition for all earlier stages. Therefore, fitting a single model to adjust for different types of sample attrition is expected to provide a reasonable bias adjustment without introducing too many parameters and thus avoiding possibly instability in calibrated weights due to estimation of many parameters. Moreover, since the controls used in the above calibration approach typically coincide with the ones used for coverage bias adjustment via post-stratification, the proposed one step sample attrition bias adjustment in fact also adjusts for the compound bias due to nonresponse and non-coverage (over or under).

The second enhancement to the calibration approach we consider is to introduce new parameters in the model to capture the impact of possibly different sampling designs in different strata (or super-strata—these are simply groups of strata with similar designs). The ultimate goal of any estimation exercise is, of course, to reduce bias and variance to the extent possible. However, commonly used calibration methods of weight adjustment (see Kott, 2006, and Särndal, 2007 for reviews) attempt to take care primarily of the bias problem but do not have any built-in mechanism to reduce variance although, in practice, bias reduction is generally accompanied by some variance reduction also. It is desirable to introduce new parameters in the bias adjustment model so that their values can be chosen to minimize the variance without affecting unbiasedness. In other words, the goal is to minimize the variance of an estimator in the class of unbiased estimators obtained by varying the values of the new parameters which, in turn, also impact estimates of the original model parameters. A useful way to interpret the role of new parameters is that they allow weights of strata with higher relative effective sample size be adjusted less than weights for strata with lower effective sample size, and thus build in some variance optimality in the calibration method. This feature is accomplished by introducing strata-specific scale parameters (or super-strata-specific to reduce the number of new parameters) along the lines of Singh and Wu (2003) in dealing with differential designs in the context of dual frame calibration estimation.

The third enhancement we propose deals with improving the customary but somewhat ad hoc method of weight trimming or winsorization of extreme weights whenever they occur after the sample attrition bias adjustment. Although the calibration estimator after bias adjustment is approximately unbiased jointly under the design and the bias model, it is known that weight trimming is expected to introduce bias but is implemented with the hope that it might be more than offset by the resulting variance reduction. There does not seem to exist a reasonable solution to the bias problem due to weight trimming although it is commonly performed in all surveys. Along the lines of Beaumont (2008), we propose a model for weights conditional on the sample such that extreme weights in the estimator can be replaced by model-based smoothed weights, and the resulting estimator remains approximately unbiased under design and the two models—one for sample attrition and the other for extreme weights. It is observed that a calibration method can also be used to fit the weight smoothing model where the calibration controls are now obtained internally from the weighted estimates of covariates from the sample itself. However, after weight smoothing, the respondent sample will no longer be balanced as in the case of weight trimming and so another step of calibration is performed for the purpose of post-stratification to the desired control totals. It is remarked that the goal of variance reduction in the calibration step for post-stratification can be maintained since after the model-based extreme weight smoothing, the estimator still remains approximately unbiased. Incidentally, a Hajek-ratio adjustment to match the population count before post-stratification is also desirable for improved variance reduction. Furthermore, to minimize variance further, the new scale parameters mentioned earlier as part of the second enhancement are suitably chosen using a grid search so that a common set of values for these parameters for all steps of calibration minimizes the variance of the final estimator for a given study variable. We remark that in practice for simplicity the unequal weighting effect can be used as a surrogate for the generic variance quantity for the minimization purpose.

The new weight calibration method proposed in this paper is termed generalized raking with optimal unbiased modification (or GROUM) signifying its optimality in the class of approximately unbiased calibration estimators. Although the main features of GROUM described above are in the context of a single frame, they also carry over to multiple frame surveys where frames can be deemed as super-strata albeit overlapping. Moreover, frame-specific modeling for weight smoothing will generally be needed for multiple frames, and a multiple frame version of the unequal weighting effect for variance minimization.

Theoretical considerations of the proposed GROUM calibration method are provided in Section 2 followed by a stepwise description of GROUM in Section 3. An illustrative hypothetical example based on the 2011 NIS PUF is presented in Section 4 followed by concluding remarks in Section 5.

2. Theoretical Considerations underlying the Proposed Method

As mentioned in the introduction, in any survey, there are successive stages of sample attrition. For example, sample attrition in list frame surveys involves incorrect location, noncontact, nonresponse to screener, ineligibility, and refusal while for RDD telephone surveys involves non-resolution, screener noncontact, screener refusal, ineligibility, interview noncontact, and refusal. Often weighting class adjustment methods are used for attrition bias. It is simple and appealing as it tries to model each stage but may entail serious compromises in weighting class formation due to lack of predictors required to be known for the attritted part of the sample.

We propose a calibration approach to sample attrition bias and wish to adjust initial weights of the respondent subsample directly for all stages of attrition in one step. For this purpose, we need adjustment factors greater than 1 and external control totals (fixed or random) for all covariates in the model. To this end, we can model the joint probability of attrition occurrence or case retention at successive stages via a logit or a restricted logit model as in Folsom and Singh (2000). This will only be an approximation because attrition phenomena at different stages could be driven by different factors. However, since the respondent subsample (s_r) will be used for the calibration method for attrition bias adjustment and covariate information about nonrespondents in the calibration approach is not needed, a rich set of covariates or predictors can be available for the bias adjustment model which are expected to be good predictors for all stages of sample attrition. Therefore a single adjustment step for all stages of sample attrition may be quite reasonable.

2.1 Model-based Initial Weight Smoothing before Nonresponse Bias Adjustment

It is observed that since the nonresponse bias adjustment factor should be greater than 1, we need the total estimate (t_x) from the respondent subsample to be less than the control total (T_x) for all covariates x under the calibration approach to model fitting. This condition may not be satisfied due to the likely presence of some large initial weights. So smoothing of large initial weights before any nonresponse adjustment might be needed. However, for this purpose, the usual method of weight trimming or winsorization may not be desirable because it may introduce further bias which we are trying to reduce. We propose an alternative to smoothing weights by modeling. In particular, within each super-stratum h, we model initial weights w_{kh} conditional on the sample (somewhat analogous to the method in Beaumont, 2008) as

$$E(w_{kh}|s_{rh}) = w^*_{kh} a_{kh}(\lambda_h) \text{ if } if\ k\ \in s_{rh} \qquad (1)$$

where $a_{kh}(\boldsymbol{\lambda}_h) = \exp(x'_{kh}\boldsymbol{\lambda}_h)$, and w^*_{kh} denote known constants representing an initial smoothing of weights before their adjustment by the factor $a_{kh}(\boldsymbol{\lambda}_h)$. The constants w^*_{kh} could be specified by the customary weight winsorization where all weights beyond $med\ \pm 3\, IQR$ are shrunk to the nearest boundary (left or right) and where IQR denotes Interquartile Range, or by using a new type of winsorization (termed non-extreme and extreme weight winsorization) in the sense that all the initial weights are first partitioned into weight intervals with boundaries defined by $med\ \pm a \times IQR$ where the factor a can take values in increments of .5; i.e., 0.5, 1, 1.5, 2, and 2.5. Now each weight is truncated to either the left or the right boundary of the interval it belongs such that it is shrunk toward the median. A third option is to let w^*_{kh} be common for all cases in the stratum and equal to the average weight.

We fit the above weight smoothing model (1) using a calibration approach for which random controls t_{xh} are obtained internally from the sample s_r of respondents itself; in other words, the sample totals are maintained which is reasonable since the model is defined conditional on the sample s_{rh} of respondents in the super-stratum h. It may be of interest to note that the calibration equations for fitting the log linear model (1) is equivalent to maximum likelihood equations assuming a Poisson distribution for initial weights treated as count data. After weight smoothing, large initial weights are replaced by smoothed weights but other weights are not smoothed. After weight smoothing, it is expected that the condition $t_x < T_x$ will be satisfied for all x's used in the nonresponse bias adjustment as required. Incidentally, the set of x's chosen for calibration for nonresponse bias or coverage bias adjustments can be identical to the set of covariates selected for the weight smoothing model.

2.2 One Step Sample Attrition Bias Adjustment

We next consider the proposed one step sample attrition bias adjustment via calibration by generalized raking in the form of inverse logit:

$$a_{kh}(\boldsymbol{\lambda}) = 1 + \exp(-x'_{kh}\boldsymbol{\lambda}) \ \ if\ k \in U_h\ , \qquad (2)$$

where U_h denotes the stratum h universe, and some of the model parameters $\boldsymbol{\lambda}$ could be stratum-specific if calibration controls were available at the stratum level; although typically calibration controls are not available at the stratum level but they may be available at the super-stratum level. It follows from (2) that by construction, the adjustment factors will be greater than 1 as desired.

The above nonresponse bias adjustment also adjusts for under-or over-coverage bias because the covariates used in coverage bias models are also deemed good predictors of nonresponse and vice-versa, and can be common with the covariates for the nonresponse bias model under the calibration approach. Thus, if there are no extreme weights after the attrition bias adjustment, the calibration process can stop.

2.3 Extreme Weight Smoothing after Bias Adjustment and Hajek-Ratio Adjustment

If there are extreme weights after the bias adjustment as is usually the case in practice, then we need to perform a separate weight smoothing step like the previous one such that we maintain approximate unbiasedness under the design and models employed for nonresponse bias and extreme weights before nonresponse bias adjustment. This weight smoothing step is expected to reduce variance without introducing bias. Further variance reduction can be achieved by calibration for post-stratification because after weight smoothing, desired calibration controls will no longer be satisfied as extreme weights identified after nonresponse adjustment are replaced by smoothed weights. However, before any post-stratification adjustment, it is desirable to perform the Hajek ratio adjustment so that estimated population count matches with the known population count. This adjustment can be super-stratum specific and tends to improve the variance reduction feature of post-stratification because in the case of regression calibration, it induces desirable centering in the working variance-covariance structure which leads to optimal regression in the special case of simple random sampling; see Singh (1996).

2.4 Post-stratification for Variance Reduction

For the post-stratification step, usual raking-ratio (equivalent to log linear modeling) is performed where the adjustment factor can be less than 1. That is

$$a_{kh}(\boldsymbol{\lambda}) = \exp(x'_{kh}\boldsymbol{\lambda}) \; if \; k \in U_h, \tag{3}$$

where some $\boldsymbol{\lambda}$s can be stratum-specific if corresponding controls are included in calibration. After calibration for post-stratification, if there are not too many extreme weights and their impact is not deemed to be too great as measured by the unequal weighting effect (UWE), then the calibration process can be stopped. If not, then it is customary to iterate the steps of weight smoothing, Hajek-ratio adjustment, and post-stratification a few times.

2.5 Design-based Variance Optimality in the Class of Unbiased Calibration Estimates

So far we introduced enhancements to the customary calibration method by model-based weight smoothing and one step sample attrition bias adjustment. However, there is no design-based variance optimality feature in usual calibration methods which can capture possibly different designs in different super-strata. For example, if a stratum has much larger effective sample size than the other, then it is desirable for the calibration method not to adjust much the initial weights of that stratum. However, any such down-weighting of the adjustment factors for specific strata should maintain approximate unbiasedness of the overall estimator. Following the Singh-Wu technique for dual frame calibration, we propose to introduce pre-specified scaling factors η_h (representing relative effective sample size, $0 < \eta_h < 1$) in the weight adjustment model so that the resulting estimates remain approximately unbiased. These new parameters are specified via a grid search external to calibration equations by minimizing the generalized variance over a key set of study variables (i.e, the sum of their variances) or the final UWE for simplicity. Specifically, the model (2) is revised for each super-stratum *h* as follows:

$$a_{kh}(\lambda) = 1 + \exp\left(\eta_h^{-1} x'_{kh}\lambda\right) if \; k \in U_h \tag{4}$$

where $0 < \eta_h < 1$, $\sum_{h=1}^{H} \eta_h = 1$. For each set of pre-specified values of η_h, the λ −parameters are estimated by calibration equations and their estimates as expected depend on η_h. Resulting estimates remain approximately unbiased because the scaling factors η_h only affect the λ −estimates which remain consistent for large samples. It follows that a class of unbiased estimates can be obtained after the sample attrition bias adjustment for different choices of η_h. For each given set of η_h, we perform weight smoothing, Hajek-ratio adjustment and post-stratification to obtain a final set of weights. That is, the same set of η_h −values used for model (2) are also used for model (3) to obtain the final weights, and then a grid search on η_h −constants over possible values of .1, .2,…, .9 such that $\sum_{h=1}^{H} \eta_h = 1$, can be used to find optimal η_h's. This completes the description of the proposed method of GROUM calibration where the generalized raking model is modified with η_h −parameters to obtain a class of approximately unbiased estimators from which the optimal estimator is selected. This is somewhat reminiscent of optimal regression in survey sampling in that it minimizes variance in the class of linear unbiased estimators. In fact, the idea of scaling factors as a surrogate for effective sample sizes for optimality occurred by considering the calibration form of optimal regression in simple random samples; see Singh and Wu (2003).

3. GROUM—the Proposed Method of Weight Calibration

A stepwise description of GROUM is summarized below.

Step I (Model-based Smoothing of Initial Weights): Start with the base weights of the final respondent subsample. Before the proposed sample attrition bias adjustment or nonresponse adjustment in general, check if the sample estimate for all the potential calibration variables are less than the calibration control totals. If not, then use a log linear model-based weight smoothing by which high extreme weights are replaced by the smoothed weights. If for any x-variable, the required condition of $t_x < T_x$ is still not satisfied, then that x-variable can either be dropped from the nonresponse adjustment model or the weights can be further smoothed by lowering the threshold for identifying extreme weights.

Step II (Compound Bias Adjustment for Nonresponse and Coverage): Use a logit linear model with super-stratum specific scale adjustment factors to adjust weights after smoothing under Step I. The adjustment factor will be greater than 1 by construction. It might be preferable to use bounds (L and U) as in the generalized exponential model of Folsom and Singh (2000) which generalizes Deville-Sarndal's restricted logit model to nonresponse adjustment. Here L is set at 1, but we need to set U and another constant C such that $L < C < U$. To get reasonable values for C and U, the unrestricted logit linear can be fit first using the calibration approach and then C can be set at the median adjustment factor and a suitable conservative (i.e., somewhat tight) choice of U based on observed third quartile of adjustment factors.

Step III (Model-based Smoothing of Bias-adjusted Weights): After Step II, some weights are expected to be extreme. Use a log linear model of expected values of adjusted weights (similar to Step I) as a function of covariates used in calibration for bias adjustment. May need separate models for different super-strata. Replace extreme weights by smoothed weights.

Step IV (Hajek Ratio Adjustment): The smoothed weights after Step III will not satisfy calibration controls. Before calibration for post-stratification, perform Hajek-Ratio adjustment to smoothed weights so that estimated count from the sample matches with the population count. This can be done foreach super-stratum separately if corresponding population counts are known.

Step V (Post-stratification for Variance reduction): For the pre-specified set of values of scaling factors, perform post-stratification using log linear modeling. The adjustment factor is greater than 0 but could be less or greater than 1. It might be preferable to use bounds (L, U) for a restricted logit model as in Deville-Sarndal. Here the center C is set at 1 and initial choices of L and U can be obtained by using first and third quartiles of the adjustment factor under the (unrestricted) log linear model which is fit initially.

Step VI (Grid Search for optimal Scaling factors): For each pre-specified set of scaling factors, compute the generalized variance of calibrated estimates over a set of key study variables to find the optimal choice of scaling factors that minimizes it. In practice, for simplicity, UWE can be used as an alternative to the generalized variance. For stratified designs, the stratified version of UWE will be preferable using super-strata and is given by

$$STUWE = \frac{\sum_{h=1}^{H}(N_h/N)^2 UWE_h/n_h}{\sum_{h=1}^{H}(N_h/N)^2/n_h} \ , \qquad (5)$$

where H denotes the number of super-strata. Now use the optimal choice of scaling factors to obtain the final calibrated weights.

The proposed GROUM method is easily applicable to multiple frames in a manner similar to the method proposed earlier by Singh and Wu (2003) for dual frames. In this case, different frames can be regarded as super-strata so that frame–specific models for weight smoothing can be fit. However, since frames can be overlapping (for example, cell and landline phones), new calibration variables for the overlap domains become part of calibration for which the control total is zero.

4. An Illustrative Example

We construct a hypothetical calibration problem using the 2011 NIS PUF data for the North-East Census Division of children aged 18-35 months in the landline sample. The PUF contains socio-demographic, vaccination information and final calibrated weights for the respondent sample of children in NIS. The calibration in NIS is based on weighting class adjustment for various stages of sample attrition and raking ratio for post-stratification. Table 1 shows the six states along with the number of complete responses for children, and the number of estimated households with landline available from CDC (http://www.cdc.gov/nchs/data/nhsr/nhsr061.pdf). Since PUF did not have information about the base weights, approximate base weights for each state were obtained by setting them proportional to the ratio of the number of households with landline and the number of sample completes. The NE census division was treated as the target population of interest. Although states essentially serve as strata in the RDD sampling design of NIS, we divided the NE division into two super-strata for our purpose of illustrating the GROUM calibration: two large states Connecticut and Massachusetts were grouped in one super-stratum with relatively low sampling rates while the remaining four small states were grouped in the second super-stratum with much higher relative sampling rates.

Table 2 shows the target population size of eligible children as 232404.77—it is not an integer because it was obtained as the calibrated population count from the PUF. The corresponding total sample size of completes is 1680. There are six categorical calibration variables corresponding to Mother's education, Mother's age, Child Age, Child gender, Child Race/ethnicity and Child First Born status each with a number of categories resulting in a total of 13 linearly independent control totals.

The results after GROUM calibration are shown in Tables 3 and 4. There was no need of Step I of initial weight smoothing. The step II of compound bias adjustment for nonresponse and coverage was performed using logit linear model with two scaling factors corresponding to the two super-strata. In step III, only one weight smoothing model was fit for simplicity rather than separate models for each super-stratum. After Hajek-ratio adjustment in Step IV and post-stratification in Step V, a grid search was used to find out the optimal set of η_h's which turns out to be (.6, .4) under all the three types of weight smoothing methods; see Table 3. It is interesting to note that, as expected, the range of calibrated weights decreases as the set of η_h's approach the optimal choice. Among the three smoothing methods, in terms of calibrated weight distribution, Method 2 of model-based smoothing with initial common weights seems to perform best followed by the conventional method of weight trimming. However, point estimates for the six vaccination variables look quite comparable under the three methods of weight smoothing. The purpose of this example was only to illustrate how various methods of weight smoothing with (super-) stratum specific scaling factors behave empirically. Without a suitable simulation experiment, it is difficult to draw definitive conclusions about the relative performance of methods.

5. Concluding Remarks

In this paper, we proposed a single or one step attrition bias adjustment via calibration as an approximation to the product of separate adjustments for successive stages of attrition. Model-based smoothing of high initial weights before attrition bias adjustment helps to satisfy the required condition for total estimates of calibration variables to be less than external control totals in order to ensure that the adjustment factor is greater than 1. Super-Stratum level scaling parameters in the adjustment factor were introduced to build-in optimality of variance reduction by capturing possibly different designs in super-strata. These new parameters are analogous to relative effective sample sizes and can be estimated outside calibration via a grid search. The proposed enhancements to the conventional calibration method lead to a new method termed generalized raking with optimal unbiased modification or GROUM. A simple numerical example based on the 2011 NIS PUF was presented to illustrate the GROUM methodology for a hypothetical calibration problem. Applications of GROUM to the multiple frame calibration problem was also briefly discussed. It is believed that the proposed method of GROUM calibration may have wide applications in the practice of survey sampling.

Acknowledgments

The authors would like to thank Dan Kasprzyk for support and encouragement and Kirk Wolter for useful suggestions about the empirical example.

References

Beaumont, J.-F. (2008). A new approach to weighting and imputation in sample surveys. *Biometrika*, 95, 539-553.

Darroch, J.N. and Ratcliff, D. (1972). Generalized Iterative Scaling for Log-Linear Models. *The Annals of Mathematical Statistics*, 43(5), 1470-1480.

Deville, J.C. and Särndal, C.E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87, 376-382.

Folsom, R.E. and Singh, A.C. (2000). The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and Poststratification. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 598-603.

Kott, P.S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, 32, 133-142.

RTI International (2010). SUDAAN Release 10.0.0., RTP, NC

Sarndal, C.-E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, 33, 99-119.

Singh, A.C. (1996).“Combining Information in Survey Sampling by Modified Regression.” Proceedings of the American Statistical Association, Section on Survey Research Methods, pp. 120-129.

Singh, A.C., and Wu, S. (2003). "An extension of generalized regression estimator to dual frame surveys", Proceedings of the American Statistical Association, Section on Survey Research Methods, 3911-3918

Table 1: NE Census Division as the Target Population, #HH with Landline, and #Sample Completes

State	#HH with Landline	#Sample Completes
Connecticut	869113.5	352
Massachusetts	1483176.5	264
Maine	302277.3	258
New Hampshire	308921.4	217
Rhode Island	253263.1	340
Vermont	151972.9	249

Table 2: Calibration Control Totals based on 2011 NIS PUF for NE Census Division

(Total Count of Children N=232404.77, Number of Sample Completes= 1680)

Aux Variable	Total Count	Aux Variable	Total Count
Mother’s Education	1: 30250.18 2: 57731.18 3: 37682.18 4: 106741.22	Child Gender	1: 119131.43 2: 113273.34
Mother’s Age	1: 3455.46 2: 66412.46 3: 162536.84	Child Race/Ethnic	1: 41541.29 2: 156048.65 3: 14846.56 4: 19968.27
Child Age	1: 68607.93 2: 80732.37 3: 83064.46	First Born Status	1: 145851.58 2: 86553.19

Footnote: Mother’s Ed (< 12 Years, 12 Years, > 12 Years, Non-College Grad, College Grad); Mother’s Age (<= 19 Years, 20 - 29 Years, >= 30 Years); Child Age (19 - 23 Months, 24 - 29 Months, 30 - 35 Months); Child Gender (Male, Female); Child Race/Ethnicity (Hispanic, Non-Hispanic White Only, Non-Hispanic Black Only, Non-Hispanic Other + Multiple Race); First Born Status (No, Yes).

Table 3: Final Calibrated Weights under Different Extreme Weight Smoothing

Cal Wt	Eta=.4			Eta=.5			Eta=.6		
Summary	**R=1**	**3**	**1000**	**R=1**	**3**	**1000**	**R=1**	**3**	**1000**
Weight Smoothing (Method 1: Extreme Weight Winsorization or Trimming)									
Min	6.23	6.67	6.90	14.56	14.34	14.34	28.90	28.66	24.65
Q1	18.83	20.60	21.55	45.57	45.27	45.24	71.42	71.50	72.55
Med	36.36	45.68	52.25	86.63	86.39	86.37	118.61	117.61	109.27
Q3	180.05	181.34	203.23	195.05	200.86	201.35	163.84	165.55	177.78
Max	994.61	706.78	506.45	585.38	492.00	476.66	476.58	406.43	385.40
#EW	112	113	0	6	0	0	30	4	0
UWE	2.79	2.48	2.20	1.81	1.79	1.79	1.41	1.40	1.40
Weight Smoothing (Method 2: Model with Initial Common Weights)									
Min	6.44	7.20	7.48	14.45	13.96	14.12	28.72	28.09	27.82
Q1	22.36	26.47	27.61	46.17	46.32	46.23	71.93	70.88	70.87
Med	48.17	71.94	76.93	90.70	92.08	93.63	118.64	117.06	117.51
Q3	176.66	208.98	235.29	209.04	218.41	222.88	165.21	168.04	168.06
Max	1020.64	665.31	561.06	588.99	526.54	531.23	489.51	427.91	383.97
#EW	108	2	0	2	0	0	28	6	0
UWE	2.64	2.10	1.97	1.76	1.72	1.72	1.40	1.38	1.38
Weight Smoothing (Method 3: Model with Nonextreme and Extreme Weight Winsorization)									
Min	5.96	6.08	6.29	14.71	14.71	14.54	29.30	29.14	28.91
Q1	17.26	17.77	19.33	44.44	44.38	44.90	71.88	71.66	71.19
Med	31.30	32.04	36.78	82.60	84.68	85.74	116.47	117.08	117.06
Q3	172.34	177.38	182.76	176.00	182.09	187.62	163.61	163.45	163.36
Max	1108.90	1053.07	988.53	726.05	627.28	527.95	492.60	466.70	396.01
#EW	117	114	112	71	67	70	66	66	52
UWE	2.99	2.95	2.81	1.92	1.87	1.85	1.44	1.43	1.42

Table 4: Estimated Vaccination Rates under Different Calibration Methods for the Hypothetical Estimation Problem for the Population of Children in the NE Division

Estimate	Eta=.4			Eta=.5			Eta=.6		
	R=1	3	1000	R=1	3	1000	R=1	3	1000
Weight Smoothing (Method 1: Extreme Weight Winsorization or Trimming)									
4+DT	0.8813	0.8842	0.8882	0.8844	0.8864	0.8865	0.8846	0.8842	0.8845
3+Polio	0.9573	0.9592	0.9608	0.9576	0.9587	0.9588	0.9572	0.9566	0.9566
1+Measles	0.9422	0.9449	0.9472	0.9426	0.9438	0.9439	0.9404	0.9398	0.9406
3+Hib	0.9587	0.9604	0.9619	0.9583	0.9595	0.9596	0.9560	0.9558	0.9562
3+HepB	0.9183	0.9219	0.9247	0.9198	0.9213	0.9214	0.9200	0.9193	0.9191
1+Varicella	0.9139	0.9146	0.9153	0.9075	0.9086	0.9087	0.8950	0.8942	0.8972
Weight Smoothing (Method 2: Model with Initial Common Weights)									
4+DT	0.8755	0.8818	0.8868	0.8853	0.8888	0.8899	0.8842	0.8830	0.8828
3+Polio	0.9556	0.9604	0.9604	0.9588	0.9605	0.9602	0.9570	0.9566	0.9559
1+Measles	0.9420	0.9449	0.9459	0.9444	0.9461	0.9470	0.9407	0.9401	0.9392
3+Hib	0.9572	0.9616	0.9615	0.9597	0.9609	0.9605	0.9558	0.9562	0.9556
3+HepB	0.9169	0.9244	0.9234	0.9223	0.9250	0.9246	0.9194	0.9195	0.9191
1+Varicella	0.9109	0.9104	0.9102	0.9080	0.9091	0.9099	0.8952	0.8943	0.8935
Weight Smoothing (Method 3: Model with Nonextreme and Extreme Weight Winsorization)									
4+DT	0.8805	0.8809	0.8823	0.8810	0.8827	0.8844	0.8848	0.8847	0.8846
3+Polio	0.9569	0.9576	0.9581	0.9556	0.9568	0.9577	0.9576	0.9575	0.9569
1+Measles	0.9405	0.9411	0.9431	0.9400	0.9412	0.9423	0.9406	0.9404	0.9398
3+Hib	0.9584	0.9591	0.9597	0.9562	0.9575	0.9584	0.9561	0.9561	0.9558
3+HepB	0.9161	0.9172	0.9197	0.9159	0.9180	0.9192	0.9209	0.9204	0.9198
1+Varicella	0.9142	0.9142	0.9143	0.9059	0.9068	0.9078	0.8956	0.8951	0.8946