Assessing the Impact of Simplified Design Assumptions When Analyzing Data from Public-Use Complex Surveys[1]

Van Parsons[2], Jennifer Parker[2]
[2]National Center for Health Statistics, Centers for Disease Control and Prevention, 3311 Toledo Rd, Hyattsville, MD 20782

Abstract
Large national population-based surveys are often based on multi-stage cluster sampling. When using design-based estimation methods on such surveys, knowledge of the complete hierarchical sampling structure is required to correctly assess the sampling distributions of estimators. However, for public-use data releases, the survey design information for analyses is substantially simplified. For example, the NHIS is described as a stratified set of independently sampled first-stage clusters with the final adjusted survey weights to be used as fixed sampling weights. While such simplified design structures can be justified under hypothetical sampling conditions, those sampling conditions rarely hold in practice. To study the impact of simplified survey analytic structures, a finite pseudo population of 340,000 households has been created using nine years of NHIS. This population captures many of the clustering features present in the U.S. population. For this paper, sampling methods and estimation methods consistent with those of the NHIS are studied on this pseudo population. Evaluations are presented under true and simplified assumptions.

Key Words: BRR, linearization, post-stratification, degrees-of-freedom

1. Introduction

Large scale government surveys like the National Health Interview Survey (NHIS), Botman et al. (2000), are based on complex sampling designs involving several levels of stratification and multiple levels of probability sampling. Furthermore, sampling weights are adjusted to account for non-response and to agree with external population control totals. For confidentiality reasons, much of the implemented design information is not released to the public. Typically, coarsened design structures, e.g., final survey weights, strata and first-stage clusters or alternatively, final survey weights and replicate weights, are provided for design-based analyses on public-use micro data. In the case of the NHIS, public data users are directed to treat the design as having two clusters sampled with replacement from each stratum, and to treat the final adjusted weights as pure sampling weights. (Note, surveys that provide only replicate weights frequently assume a with-replacement structure for replicate creation).

While final survey weights are targeted at providing unbiased and improved precision over base weight estimators, the estimation of precision, e.g., variances, standard errors, coefficients of variation and confidence intervals, requires knowledge of both the

[1] *The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control*

weighting and clustering components of the design. At the theoretical level, the true measures of design-based precision are mathematically intractable due to the complexity of sampling and weighting, and consequently simplified variance estimation structures that are accepted as reasonably accurate and amenable to computation with existing survey software packages, e.g., *SUDAAN*®, *SAS*® , *R*, must be recommended for assessing precision. A commonly used simplifying design assumption is that a design using a hierarchical without replacement sampling sequence at all stages can at some stage of sampling be treated as sampling with replacement, and the higher levels of clustering can be ignored in the functional form of the variance estimator. To justify such an approximation, it is assumed that if $\hat{v}$ is the estimator of a true variance, v, then $E(\hat{v}) \geq v$, but with only a minor positive bias. Cochran (1977), section 10.4, discusses this approximation for a special case of simple random sampling.

For a real population and associated complex design the evaluation quantities, $E(\hat{v})$ and v, needed to assess the quality of a variance estimator are mathematically intractable, but for artificial populations the $\hat{v}$ and v can be assessed by simulations. For this paper we have created an artificial population with structural clustering somewhat consistent with that of the NHIS population, and we define a 5-level multi-stage sampling rule that is consistent with the actual NHIS multi-stage sampling for both the all person per household sample and the one-adult per household sample. Sampling weights are adjusted by post-stratification factors. For variance estimation the multistage sampling is simplified as a "with-replacement sample of 2-clusters per stratum" design as is done in the actual NHIS. The variance estimators for a sample proportion, $\hat{p}$, and a regression estimator, say $\hat{\beta}$, are assessed, as these are the types of statistics that are commonly analyzed on the NHIS data systems. While this current study is limited in scope, the pseudo universe and sampling may be considered as taking a sample of about 1000 households from a generic state from the NHIS. As this work is the result of a 2012 JSM poster session, the layout of this paper will be somewhat succinct with emphasis on tables and short commentary.

2.0 Methods

2.1 Pseudo-Population

A pseudo-population with geographical clustering and intra-class correlations similar to that of the U.S. population is created by combining both design and interview data from the 1997-2005 NHIS. With only a few exceptions, all of the original NHIS Primary Sampling Units (PSUs) and Secondary Sampling Unit (SSU) clusters within the PSUs define the geographical clusters for the pseudo-population. The NHIS interviewed households and sampled persons are the population elements. For stratification, the original NHIS treats the largest metropolitan area PSUs as self-representing (SR) and combines the other PSUs into non-self-representing (NSR) strata defined by state, metropolitan status and select demographic characteristics. For the pseudo-population, the three largest NHIS metropolitan PSUs, New York, Los Angeles, and Chicago, have sample person counts much larger the other PSUs, and these three are defined as SR strata. The remaining PSUs are partitioned into NSR strata defined by geographical region, metropolitan status and minority population proportions. This stratification results in a population consisting of 14 Strata, of which 3 are SR strata and 11 are NSR strata. In totality there are 297 PSUs, 6,680 SSU clusters within the PSUs, 340,348 households within the SSUs, with each SSU having a minimum of 20 households, and a total population count of 879,342 Persons within households. For this study, the variables, age, race/ethnicity (Hispanic, non-Hispanic black, non-Hispanic other), gender, household poverty index, health insurance coverage status, and fair/poor health status are associated with each person. Person variables are taken from the public-use micro databases. NHIS households with unit non-response are omitted from the pseudo-population. As the actual NHIS oversampled blacks and Hispanics, this pseudo-population will have higher proportions of those targeted groups than the true U.S. population.

2.2 Multi-Stage Sampling and Weighting

A multi-stage sampling procedure with 5 sampling levels, which is consistent with NHIS sampling, is used to collect a simulated complex pseudo-sample from the population.
The sequence of sampling is as follows:

1. Sample two PSUs with probability proportional to population size from each NSR Stratum using Brewer's method (Cochran1977, Section 9.A.8).
2. Within each PSU, systematically sample SSUs to satisfy a self-weighing selection, Probability (select SSU j | within PSU i) · Probability (PSU i) = 1/185
3. Within each SSU, sample about 10 households.
4. Within the SSU household sample, oversample for black households 2:1 over non-black households, i.e., screen-in all black households and retain half of the non-black households in the sample.
5. Sample all persons per household, and then select one adult aged 18 years or older using simple random sampling within the household.

After screening in step 4 above, about 1000 households are in the sample. The complex pseudo-sample consists of about 2800 sampled persons for the full household sample and about 1000 sampled adults for the adult sample.

The person and adult base weights are the products of the inverses of the respective probabilities of selection, and as in the original NHIS, a final post-stratification weighing adjustment is applied to the base weights. For the full person sample, 45 age-race/ethnicity-gender population controls are used in post-stratification, and for the adult sample 28 such controls are used.

The multistage sampling and weighting adjustments should capture many of the features of the actual NHIS. Tillé's *R*-package "*sampling*" is used for probability sampling, and 1000 such complex samples are generated to study the sampling properties of design-based estimators.

2.2 Simplified Design Variances

As is done in the original NHIS, the generated sample is assumed to be approximated by a "Two-PSU sample per stratum with replacement sampling" design. The NSR strata have a natural structure for this. For the SR strata the systematically selected SSUs are paired using the technique discussed in Wolter (2007) expression (8.2.8). Thus, data analyses are based on having 14 strata and 2 PSUs within each stratum. When using a Taylor linearization approach for variance estimation, the post-stratified weights are usually treated as base weights when implemented with commonly used survey analysis software packages. In this evaluation a Taylor linearization that accounts for post-stratified weights is not considered. Replication procedures have the capability of applying all weighting adjustments to each replicate to possibly capture another source of variability in design-based estimation. For this study we use a Balanced Repeated Replicate (BRR) procedure along with a post-stratification weighting adjustment for each replicate. BRR Fay-factors of c = 0.3 and c = 0.5 are also considered. Here, if *0* and *2* are the traditional BRR weighting factors for the two PSUs within a stratum, then *c* and *(2-c)* are the corresponding Fay-weighting factors. Lumley's *R* package "*survey*" is used to create replicate weights.

3.0 Evaluations

3.1 Statistics

Sample proportions are perhaps the most commonly computed statistics from NHIS data. Here, the following survey estimates of proportions and variance will be evaluated for both the full person sample and the one-adult per household sample.

1. Proportion of persons (adults) having fair or poor health
2. Proportion of persons (adults) less than 65 years old who have no health insurance.
3. Proportion of persons (adults) less than 65 years old who have no health insurance, but restricted to living in 6 selected geographical strata.

Regressions are also frequently performed on complex survey data. For regressions the population parameters are defined $\beta = (X'X)^{-1} X'Y$, where the Y vector is a complete population variable and the X matrix consists of columns of complete population covariates. The β parameters are estimated from the sample by the corresponding expression $(x'wx)^{-1}x'wy$, where the x and y are now the sampled components of X and Y and w are the survey weights.

The estimated proportion of persons (adults) having fair or poor health, $\hat{p}$, will be considered to be a linear function of sex, race/ethnicity, age and poverty as in the following model specification:

$E(\hat{p})$ ~ Intercept + sex (M, F) + race (Other, Black, Hispanic) + age + poverty.

This model is considered for both the full person sample and the one-adult per household sample.

3.2 Evaluation Quantities

The sampled proportion, $\hat{p}$, and its variance estimator, $\widehat{Var}(\hat{p})$, will be evaluated with respect to the sampling distribution measures of *Bias* and *Variance* along with functions of such measures:

Bias($\hat{p}$) = (E($\hat{p}$) – p), with p equal to the true population proportion,

Relative-Bias($\widehat{Var}(\hat{p})$) = $E(\widehat{Var}(\hat{p}))$ / *Var*($\hat{p}$), with *Var*($\hat{p}$) equal to the sampling variance of $\hat{p}$, true coefficient of variation of $\hat{p}$, $CV(\hat{p}) = SE(\hat{p}) / E(\hat{p})$, where $SE(\hat{p}) = \sqrt{Var(\hat{p})}$, expected estimated $CV(\hat{p}) = E\ (\widehat{SE}(\hat{p}) / \hat{p}\ )$, where $\widehat{SE}(\hat{p}) = \sqrt{(\widehat{Var}(\hat{p})}$, and the

Satterthwaite *degrees-of-freedom* of the variance estimator ,

$DF = 2 \cdot [E\left(\widehat{Var}(\hat{p})\right)]^2 / Var(\widehat{Var}(\hat{p}))$. This is equivalent to $2/CV^2(\widehat{Var}(\hat{p}))$.

For standard regression packages, typical outputs include $\hat{\beta}$, $\widehat{SE}(\hat{\beta})$ and $T = \hat{\beta} / \widehat{SE}(\hat{\beta})$. To evaluate these components, the same functional forms used for $\hat{p}$ will also be used for each single regression parameter, β, except, rather than considering the $CV(\hat{\beta})$, the *T*-statistic for testing a "zero" β value will be considered. These will be reciprocals of the CV's as stated above.

The true *T*-statistic will be defined as $E(\hat{\beta})\ /\ SE(\hat{\beta})$, where $SE(\hat{\beta}) = \sqrt{Var(\hat{\beta})}$, and the expected estimated *T*-statistic will be defined as $E(\hat{\beta}/\widehat{SE}(\hat{\beta}))$, where $\widehat{SE}(\hat{\beta}) = \sqrt{(\widehat{Var}(\hat{\beta})}$.

3.3 Targets of Evaluation Quantities

The Bias of the estimators $\hat{p}$ or $\hat{\beta}$ will be directly compared to the true population value which can be directly computed from the population. The target value for Bias is 0. The target for Relative-Bias ($\widehat{Var}(\hat{p})$) is 1.0; values less than unity indicate that the variance estimator is an underestimate on average, and values larger than unity indicate the variance estimator is an overestimate on average. An underestimate raises more concerns as the situation falsely appears to indicate better precision than is warranted, e.g., the *CV*'s are underestimating the true values. In the case of testing a regression parameter to be zero, the Type-I error will be larger than the nominal value. The Satterthwaite degrees-of-freedom (*DF*) reduces to "(number of PSU's – the number of strata)" if the PSU sampling within strata is indeed with replacement and the magnitude of the sampling variance is constant over the different strata. For the generated sample, the target DF = 14 when all PSUs contain the targeted characteristics of estimation.

The nominal coverage of a confidence interval is compared to the true coverage. If $\hat{p} \pm t\,\widehat{SE}(\hat{p})$ are the two tails of estimated coverage, then the position of the true value of p in relation to the tails can be calculated, and thus the coverage can be determined. For the generated sample, the nominal t cutoff will correspond to a *T*-distribution with d degrees-of-freedom and the Probability ($T_d \geq t$) = .05, i.e., the nominal confidence level will be 0.90.

4.0 Findings

The set of Tables 1-3 and Tables 4-7 provide the evaluation quantities for the selected estimates of proportions and estimates of regression parameters, respectively, for the sample of persons and the sample of adults (18+ years). To study the impact of post-stratification, the original base weights without any post-stratifications are also considered as sole-weighing factors. These analyses are designated as BW in column (2) of the tables. The rows for linearization and traditional BRR with post-stratified weights (PS) are highlighted as those are the commonly released structures in the public-use data files. Since the type of variables and statistics considered are limited in scope, broad-based generic conclusions should not be made. Below are some general observations.

1. An examination of column (3) indicates no or minimal bias (except for the age β in Table 6.1) for $\hat{p}$ and $\hat{\beta}$ for all weighting and variance estimation methods.

2. For proportions $\hat{p}$, the PS-weighted estimates have smaller CV's, (column (6)), than the BW-weighted estimates. This finding is consistent with the motivation of using PS-weighting to reduce variance, without modifying bias in $\hat{p}$.

3. For the cases studied, the traditional BRR is more conservative than the BRR with Fay factor 0.3 or 0.5, however, no general conclusions about BRR versus linearization can be made.

4. An examination of columns (4) shows that cases exist for both non-negligible positive and negative biases of the variance estimator. Any assumption that the coarsening of design information for variance estimation results in conservative standard errors should be viewed with caution. This current study is too limited to discuss conditions for directional variance bias.

5. The Satterthwaite degrees-of-freedom of column (5) is less than the "rule-of-thumb" degrees-of-freedom defined by the (number of PSUs – the number of strata). For inference, data users often use a t-distribution or the normal z-distribution cutoff for a confidence interval or *T*-test. From the simulations, the *T*-statistic frequently has a true skewed distribution. The *p*-values or confidence level cutoffs based on the nominal *z*- or *t*-distributions, may be inaccurate. In particular, one-sided tests may have very large biases.

6. The confidence interval coverage of column (8) has components of relative bias of standard errors, degrees-of-freedom and sampling distribution tail weight. It can be seen that sampling distributions for $\hat{p}$ and $\hat{\beta}$ have tails that are non-symmetric. Extreme cases of asymmetry occur in Table 6.1 for the tails of the β-age parameter estimate when using BRR replicates. For proportions, the estimated no insurance variable coverage based on nominal cutoffs (Tables 2.1, 2.2, 3.1, 3.2) appears to underestimate tail probabilities.

4.1 Limitations

1. The pseudo-population and sampling results are "smoother" than the actual NHIS components. For example, the true population has ineligible households, empty SSUs and more size variations among geographical units. The true NHIS sample is subject to non-response and additional weighting adjustments.
2. For proportions, performing a logistic regression would be preferable to performing a standard regression as is done in this study.
3. The stated "true" results of the tables are subject to simulation error. A simulation with 1000 repetitions has a simulation standard error of about 0.95% for a 10% population characteristic.

5.0 Conclusions

This study considered statistical estimation and inference when coarsened design structures are used for variance estimation. While limited in scope, the results of this study may be considered applicable to design-based estimation for a generic state within the NHIS having about 1000 sampled households. The findings suggest that the weighting strategies lead to approximately unbiased estimates of proportions and regression parameters, but using the coarsened design information does not allow any broad based generic conclusions about standard error or inference accuracy. Under such design informational conditions, analysts should be cautious in making inference. The use of additional levels of sampling may be needed for more accurate variance estimation and this will be a future direction of study. Confidentiality concerns may restrict such a finer design release on public-use micro data.

References

Botman, S., Moore, T., Moriarity, C., and Parsons, V. (2000). Design and Estimation for the National Health Interview Survey, 1995-2004, *Vital and Health Statistics*, 2(130).

Cochran, W. G (1997). *Sampling Techniques*, John Wiley and Sons, New York, 3rd edition.

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL http://www.R-project.org.

Wolter, K.M. (2007). Introduction to Variance Estimation, Second Edition, Springer.

Table 1.1 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
All persons in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) CV	(7) CV	(8)		
Var Est Method	Wgt	$E(\widehat{p})$	$\frac{E(\widehat{Var})}{Var(\hat{p})}$	DF	$\frac{SE(\hat{p})}{E(\widehat{p})}$	$E(\widehat{CV})$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.10	1.06	10.4	8.8	8.9	0.05	0.04	0.91
Linear	PS	0.09	1.00	9.5	8.3	8.1	0.07	0.04	0.89
BRR(0)	BW	0.10	1.06	10.5	8.8	8.9	0.05	0.03	0.91
BRR(0)	PS	0.09	1.09	9.8	8.3	8.4	0.06	0.03	0.91
BRR(.3)	PS	0.09	1.04	9.7	8.3	8.3	0.06	0.04	0.90
BRR(.5)	PS	0.09	1.02	9.6	8.3	8.2	0.07	0.04	0.90
Target		0.09	1.00	14			0.05	0.05	0.90

Table 1.2 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
One sample adult (age 18+) in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) CV	(7) CV	(8)		
Var Est Method	Wgt	$E(\widehat{p})$	$\frac{E(\widehat{Var})}{Var(\hat{p})}$	DF	$\frac{SE(\hat{p})}{E(\widehat{p})}$	$E(\widehat{CV})$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.12	1.03	10.8	10.3	10.3	0.06	0.05	0.90
Linear	PS	0.12	0.97	10.6	9.9	9.6	0.06	0.04	0.90
BRR(0)	BW	0.12	1.03	10.8	10.3	10.3	0.06	0.05	0.90
BRR(0)	PS	0.12	1.10	11.0	9.9	10.2	0.05	0.04	0.91
BRR(.3)	PS	0.12	1.03	10.8	9.9	9.9	0.06	0.04	0.90
BRR(.5)	PS	0.12	1.00	10.7	9.9	9.7	0.06	0.04	0.90
Target		0.12	1.00	14			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) CV = True SE of estimator / True mean of estimator
(7) Sampling mean of estimated CV
(8) True coverage for *p* based on confidence intervals using lower and upper t(df=14) tail cutoffs

Note: Highlighted rows represent commonly released design structures on public-use data files.

Table 2.1 Evaluations for $\hat{p}$ = proportion of persons age < 65 with no health insurance, All persons in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) CV	(7) CV	(8)		
Var Est Method	Wgt	$E(\widehat{p})$	$\frac{E(\widehat{Var})}{Var(\hat{p})}$	DF	$\frac{SE(\hat{p})}{E(\widehat{p})}$	$E(\widehat{CV})$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.26	0.91	11.5	7.1	6.7	0.08	0.05	0.87
Linear	PS	0.26	0.89	13.2	6.6	6.1	0.09	0.05	0.86
BRR(0)	BW	0.26	0.91	11.4	7.1	6.7	0.08	0.05	0.87
BRR(0)	PS	0.26	0.97	13.0	6.6	6.4	0.08	0.04	0.88
BRR(.3)	PS	0.26	0.92	13.2	6.6	6.2	0.09	0.04	0.87
BRR(.5)	PS	0.26	0.91	13.2	6.6	6.2	0.09	0.04	0.87
Target		0.26		14			0.05	0.05	0.90

Table 2.2 Evaluations for $\hat{p}$ = proportion of persons age < 65 with no health insurance, One sample adult (age 18+) in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) CV	(7) CV	(8)		
Var Est Method	Wgt	$E(\widehat{p})$	$\frac{E(\widehat{Var})}{Var(\hat{p})}$	DF	$\frac{SE(\hat{p})}{E(\widehat{p})}$	$E(\widehat{CV})$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.28	0.93	12.3	7.4	7.0	0.07	0.06	0.88
Linear	PS	0.28	0.88	13.1	6.9	6.4	0.07	0.06	0.86
BRR(0)	BW	0.28	0.93	12.3	7.4	7.0	0.07	0.06	0.88
BRR(0)	PS	0.28	1.02	12.9	6.9	6.8	0.06	0.05	0.89
BRR(.3)	PS	0.28	0.94	13.1	6.9	6.6	0.07	0.06	0.88
BRR(.5)	PS	0.28	0.91	13.1	6.9	6.5	0.07	0.06	0.87
Target		0.28		14			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) CV = True SE of estimator / True mean of estimator
(7) Sampling mean of estimated CV
(8) True coverage for *p* based on confidence intervals using lower and upper t(df=14) tail cutoffs
Note: Highlighted rows represent commonly released design structures on public-use data files.

Table 3.1 Evaluations for $\hat{p}$ = proportion of persons age < 65 with no health insurance
All persons in sampled household restricted to a 6 stratum domain, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) CV	(7) CV	(8)		
Var Est Method	Wgt	$E(\widehat{p})$	$\frac{E(\widehat{Var})}{Var(\hat{p})}$	DF	$\frac{SE(\hat{p})}{E(\widehat{p})}$	$E(\widehat{CV})$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.29	0.96	4.7	11.1	10.3	0.08	0.06	0.86
Linear	PS	0.29	0.89	5.5	10.7	9.7	0.09	0.06	0.85
BRR(0)	BW	0.29	0.97	4.7	11.1	10.4	0.08	0.05	0.86
BRR(0)	PS	0.29	0.99	5.5	10.7	10.3	0.09	0.05	0.86
BRR(.3)	PS	0.29	0.94	5.5	10.7	10.0	0.09	0.06	0.86
BRR(.5)	PS	0.29	0.92	5.5	10.7	9.9	0.09	0.06	0.86
Target		0.29		6			0.05	0.05	0.90

Table 3.2 Evaluations for $\hat{p}$ = proportion of persons age < 65 with no health insurance,
One sample adult (age 18+) in sampled household restricted to a 6 stratum domain

(1)	(2)	(3)	(4)	(5)	(6) CV	(7) CV	(8)		
Var Est Method	Wgt	$E(\widehat{p})$	$\frac{E(\widehat{Var})}{Var(\hat{p})}$	DF	$\frac{SE(\hat{p})}{E(\widehat{p})}$	$E(\widehat{CV})$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.31	1.02	4.7	11.2	10.9	0.07	0.05	0.88
Linear	PS	0.31	0.92	5.7	11.0	10.2	0.07	0.06	0.87
BRR(0)	BW	0.31	1.03	4.7	11.2	10.9	0.07	0.05	0.88
BRR(0)	PS	0.31	1.07	5.8	11.0	11.0	0.06	0.04	0.90
BRR(.3)	PS	0.31	0.99	5.8	11.0	10.6	0.06	0.06	0.88
BRR(.5)	PS	0.31	0.95	5.7	11.0	10.4	0.06	0.06	0.88
Target		0.31		6			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) CV = True SE of estimator / True mean of estimator
(7) Sampling mean of estimated CV
(8) True coverage for *p* based on confidence intervals using lower and upper t(df=14) tail cutoffs
Note: Highlighted rows represent commonly released design structures on public-use data files.

Table 4.1 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: Sex(F) – Sex(M)

All persons in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var}(\hat{\beta}))}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\,\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.00	0.97	11.2	0.3	0.3	0.06	0.04	0.90
Linear	PS	0.00	0.94	11.3	0.3	0.3	0.06	0.04	0.90
BRR(0)	BW	0.00	1.05	11.2	0.3	0.3	0.05	0.04	0.91
BRR(0)	PS	0.00	1.12	11.2	0.3	0.3	0.04	0.04	0.93
BRR(.3)	PS	0.00	1.05	11.1	0.3	0.3	0.04	0.04	0.92
BRR(.5)	PS	0.00	1.03	11.1	0.3	0.3	0.04	0.04	0.92
Target		0.00		14			0.05	0.05	0.90

Table 4.2 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: Sex(F) – Sex(M)

One sample adult (age 18+) in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var}(\hat{\beta}))}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\,\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.00	0.97	11.5	0.2	0.2	0.04	0.05	0.91
Linear	PS	0.00	0.92	11.8	0.2	0.2	0.04	0.06	0.90
BRR(0)	BW	0.00	0.98	11.5	0.0	0.0	0.05	0.04	0.91
BRR(0)	PS	0.00	1.06	11.9	0.0	0.0	0.05	0.02	0.93
BRR(.3)	PS	0.00	0.97	11.8	0.0	0.0	0.05	0.03	0.92
BRR(.5)	PS	0.00	0.93	11.7	0.0	0.0	0.06	0.03	0.91
Target		0.00		14			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) T for testing $E(\widehat{\beta}) = 0$, True mean of estimator / True SE of estimator
(7) Sampling mean of estimated T statistic
(8) True coverage for β based on confidence intervals using lower and upper t(df=14) tail cutoffs
Note: Highlighted rows represent commonly released design structures on public-use data files.

Table 5.1 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: race(B) – race(W)
All persons in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var(\hat{\beta})})}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.03	0.98	7.8	1.8	1.9	0.05	0.04	0.91
Linear	PS	0.03	0.96	7.9	1.8	2.0	0.06	0.04	0.90
BRR(0)	BW	0.03	0.74	10.2	1.7	2.1	0.08	0.05	0.87
BRR(0)	PS	0.03	0.79	10.5	1.7	2.0	0.08	0.05	0.87
BRR(.3)	PS	0.03	0.75	10.4	1.7	2.1	0.08	0.06	0.86
BRR(.5)	PS	0.03	0.73	10.4	1.7	2.1	0.08	0.06	0.86
Target		0.03		14			0.05	0.05	0.90

Table 5.2 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: race(B) – race(W)
One sample adult (age 18+) in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var(\hat{\beta})})}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.05	0.98	7.6	1.6	1.8	0.06	0.04	0.90
Linear	PS	0.05	0.93	8.4	1.6	1.8	0.06	0.05	0.89
BRR(0)	BW	0.05	1.01	8.8	1.5	1.6	0.05	0.04	0.91
BRR(0)	PS	0.05	1.08	9.4	1.5	1.5	0.04	0.04	0.92
BRR(.3)	PS	0.05	0.97	9.6	1.5	1.6	0.05	0.05	0.90
BRR(.5)	PS	0.05	0.93	9.5	1.5	1.6	0.05	0.06	0.89
Target		0.05		14			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) T for testing $E(\widehat{\beta})$ = 0, True mean of estimator / True SE of estimator
(7) Sampling mean of estimated T statistic
(8) True coverage for β based on confidence intervals using lower and upper t(df=14) tail cutoffs
Note: Highlighted rows represent commonly released design structures on public-use data files.

Table 6.1 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: age (recoded)
All persons in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var}(\hat{\beta}))}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.37	0.99	10.8	10.6	11.3	0.04	0.04	0.92
Linear	PS	0.37	0.97	10.8	10.7	11.7	0.05	0.04	0.91
BRR(0)	BW	0.38	0.75	12.2	10.6	13.0	0.03	0.12	0.85
BRR(0)	PS	0.38	0.83	11.9	10.7	12.5	0.02	0.10	0.87
BRR(.3)	PS	0.38	0.78	12.0	10.7	12.9	0.02	0.12	0.86
BRR(.5)	PS	0.38	0.76	12.0	10.7	13.0	0.03	0.12	0.85
Target		0.36		14			0.05	0.05	0.90

Table 6.2 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: age (recoded)
One sample adult (age 18+) in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var}(\hat{\beta}))}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.46	0.95	11.3	6.8	7.5	0.05	0.05	0.90
Linear	PS	0.46	0.92	11.4	6.9	7.7	0.05	0.05	0.90
BRR(0)	BW	0.46	1.00	11.2	6.7	7.2	0.05	0.04	0.91
BRR(0)	PS	0.45	1.09	12.1	6.7	6.8	0.04	0.03	0.92
BRR(.3)	PS	0.46	0.99	12.1	6.7	7.1	0.05	0.04	0.91
BRR(.5)	PS	0.46	0.95	12.0	6.7	7.3	0.06	0.04	0.90
Target		0.46		14			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) T for testing $E(\widehat{\beta})$ = 0, True mean of estimator / True SE of estimator
(7) Sampling mean of estimated T statistic
(8) True coverage for β based on confidence intervals using lower and upper t(df=14) tail cutoffs
Note: Highlighted rows represent commonly released design structures on public-use data files.

Table 7.1 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: poverty (recoded to increasing)
All persons in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var}(\hat{\beta}))}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.03	0.94	7.2	6.6	7.4	0.06	0.03	0.92
Linear	PS	0.03	0.93	7.1	6.7	7.5	0.06	0.03	0.91
BRR(0)	BW	0.03	0.70	11.5	6.7	8.6	0.06	0.09	0.85
BRR(0)	PS	0.03	0.76	11.2	6.8	8.3	0.06	0.08	0.86
BRR(.3)	PS	0.03	0.72	11.2	6.8	8.5	0.06	0.08	0.86
BRR(.5)	PS	0.03	0.70	11.2	6.8	8.6	0.06	0.09	0.85
Target		0.03		14			0.05	0.05	0.90

Table 7.2 Evaluations for $\hat{p}$ = proportion of persons with fair to poor health
$E(\hat{p})$ ~ I + Sex (M,F) + Race(Other, Black, Hispanic) + age + poverty
β: poverty (recoded to increasing)
One sample adult (age 18+) in sampled household, pseudo sample

(1)	(2)	(3)	(4)	(5)	(6) T	(7) T	(8)		
Var Est Method	Wgt	$E(\widehat{\beta})$	$\frac{E(\widehat{Var}(\hat{\beta}))}{Var(\hat{\beta})}$	DF	$\frac{E(\hat{\beta})}{SE(\hat{\beta})}$	$E\frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$	Low Tail	Upper Tail	Conf Level
Linear	BW	0.04	0.94	7.7	5.4	6.0	0.05	0.04	0.92
Linear	PS	0.04	0.89	8.0	5.4	6.2	0.06	0.04	0.91
BRR(0)	BW	0.04	1.00	9.1	5.7	6.0	0.02	0.06	0.92
BRR(0)	PS	0.04	1.08	9.1	5.6	5.7	0.03	0.05	0.92
BRR(.3)	PS	0.04	0.99	9.1	5.6	6.0	0.02	0.06	0.91
BRR(.5)	PS	0.04	0.95	9.1	5.6	6.1	0.03	0.07	0.90
Target		0.04		14			0.05	0.05	0.90

(1) Variance Estimation Method: Linearization or BRR with Fay factor
(2) BW is the base weight before post-stratification, PS is the post-stratified weight
(3) The true sampling mean of the estimator
(4) Relative Bias of variance estimator
(5) Satterthwaite degrees-of-freedom
(6) T for testing $E(\widehat{\beta})$ = 0, True mean of estimator / True SE of estimator
(7) Sampling mean of estimated T statistic
(8) True coverage for β based on confidence intervals using lower and upper t(df=14) tail cutoffs
Note: Highlighted rows represent commonly released design structures on public-use data files.