A Simulation Study of Cell Collapsing in Poststratification

Jay J. Kim, National Center for Health Statistics, Centers for Disease Control
Linda Tompkins, National Center for Health Statistics, Centers for Disease Control
Jianzhu Li, Joint Program for Survey Methodology, University of Maryland
Richard Valliant, Survey Research Center, University of Michigan

ABSTRACT

A standard procedure in poststratification is to collapse or combine cells when the sample sizes fall below some minimum or the weight adjustments are above some maximum. Collapsing may decrease the variance of an estimate but may simultaneously increase its bias. We study the effects on bias and variance of this type of dynamic cell collapsing through simulation using a population based on the 2003 National Health Interview Survey.

Key Words: bias, combining cells, coverage ratio, poststratification, variance, weight trimming

1. Introduction

Raking and poststratification are two common techniques used in survey weighting. These weighting methods can serve to (1) reduce variances or (2) adjust for deficient coverage by the sample of some groups in the target population. In household surveys in the U.S. the second purpose is especially important because some demographic groups, like young Black males, are covered less well than others (e.g., see Kostanich and Dippo 2000, ch. 16). Adjusting for undercoverage can lead to differential weights, which may correct for bias but will also increase standard errors. Practitioners often avoid making extreme weight adjustments, in effect trading-off some bias reduction in order to keep variances under control.

One method of controlling the size of weight adjustments is to collapse the initial raking or poststratification cells together if the adjustment in a cell is too large. Little (1993) and Lazzeroni and Little (1998) cover methods of collapsing categories of ordinal poststratifiers. Other strategies for how to collapse strata have been suggested by Kalton and Maligalig (1991), and Tremblay (1986). Kim, Thompson, Woltman, and Vajs (1982) give some practical applications. In this paper, we study the effects on bias and variance of combining cells, assuming that more finely defined cells would be preferable if the sample sizes and sizes of weight adjustments were acceptable. This paper is a sequel to Kim (2004) which provides more detailed background.

The *inverse coverage ratio* is the usual statistic used to determine whether cells should be collapsed and is defined as the ratio of the control count to the initially weighted sample count for the row/column. If the ratio is either too large or too small, a row or column may be collapsed with another. The Current Population Survey (CPS) conducted by U.S. Bureau of the Census uses 2 and 0.6, respectively, for the thresholds for collapsing (see Kostanich and Dippo 2000, p. 10-7). A row or column may also be collapsed when its raw sample count is below some minimum. Collapsing serves to restrict the range of weight adjustments used due to poststratification.

This paper demonstrates some of the weaknesses of the current cell collapsing procedures and proposes alternatives. Section 2 discusses the effects on weights of collapsing cells in poststratification and raking. Section 3 analyzes the bias that can be introduced by collapsing. Some alternative ways of restricting the size of weights are introduced in section 4. Empirical properties of the standard and alternative methods are investigated through simulation in section 5. We conclude in section 6 with some recommendations for determining whether cells can be safely collapsed or not.

2. Effect of Weight Redistribution on Coverage Corrections

Weight redistribution due to collapsing of cells can have a substantial effect in poststratification as sketched in this section.

2.1 Poststratification

To illustrate the effects of cell collapsing on coverage corrections, we first consider poststratification. Suppose that there are $i = 1,\dots,I$ poststrata. Let N_i be the control count for poststratum i and $w_i^{(0)}$ be the total initially weighted sample count for cell i. The quantity $w_i^{(0)}$ is the sample estimate of the number of units in the cell in the population. After poststratification using all I cells individually, the weighted count for cell i is

$$w_i^{(1)} = w_i^{(0)} \frac{N_i}{w_i^{(0)}} \equiv w_i^{(0)} f_i \,.$$

We will call $f_i = N_i \big/ w_i^{(0)}$ the *initial adjustment (or ratio) factor* (IAF) for cell i. If cells 1 and 2 are collapsed and $N_2 = cN_1$, then the adjustment factor for the combined cell can be written as

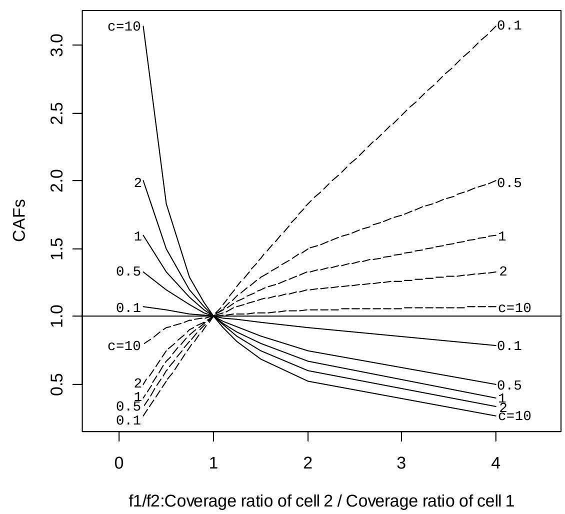
$$\frac{N_1 + N_2}{w_1^{(0)} + w_2^{(0)}} = \frac{f_1 f_2 (1 + c)}{c\, f_1 + f_2}. \qquad (1)$$

This adjustment would be applied to the weights for all units in the collapsed cell. In particular, units in cell 1 would receive the IAF of f_1 times a collapsing adjustment factor (CAF) of $\frac{f_2(1+c)}{c\, f_1 + f_2}$. Similarly, units in cell 2 receive the IAF of f_2 times a CAF of $\frac{f_1(1+c)}{c\, f_1 + f_2}$. The ratio of the CAF for cell 2 to that of cell 1 is simply f_1/f_2, which is also the ratio of the coverage rate for cell 2 to that for cell 1.

Figure 1 graphs the CAF for units in cell 1 for different values of *c* and f_1/f_2, the ratio of coverage rates for cells 1 and 2. If the coverage rates are equal, $f_1/f_2 = 1$, then collapsing does no harm as far as adjusting for undercoverage within cell goes. Assuming that f_1 is the appropriate coverage adjustment for cell *i*, CAF > 1 implies that weights for units in that cell are increased too much to correct for a coverage error. A CAF > 1 for cell 1 corresponds to $f_1/f_2 < 1$, i.e., to a case where cell 2 is covered less well than cell 1. In general, when cell 1 is over-corrected due to collapsing, cell 2 will be under-corrected and vice versa. The range of CAF's in the figure is substantially larger for cell 1 over the range of coverage ratios when cell 2 is much larger than cell 1, e.g., *c* = 10, than when the reverse is true, e.g., *c* = 0.1.

Example 1. As a simple illustration of the numerical effect of collapsing, suppose that $N_1 = 200$, $w_1^{(0)} = 50$, (i.e., f_1 = 4), $N_2 = 150$, $w_2^{(0)} = 150$, and ($f_2 = 1$). In this case, $f_1/f_2 = 4$ and c = 0.75. Assume that an equal probability sample has been selected with $n_1 = 5$ and $n_2 = 15$ so that the initial weights are 10 for units in both cells 1 and 2. If the two cells are collapsed, the adjustment in (1) is $f_1 CAF_1 = f_2 CAF_2$ = 1.75, and the implied CAF's are CAF_1 = 0.438 for cell 1 and CAF_2 = 1.75 for cell 2. After collapsing the weight for each unit in cells 1 and 2 is 17.5 (10*1.75). The estimated population counts in cells 1 and 2 are 88 and 263 rather than the population counts of 200 and 150. Thus, after collapsing the estimated counts are about 56% too low in cell 1 and 75% too high in cell 2. In both cases, the impact is large.

Figure 1. Collapsing adjustment factors (CAF's) for cells 1 and 2 plotted versus the ratio of coverage rates for cell 2 versus cell 1, f_1/f_2. Solid lines are for cell 1 (CAF1); dashed lines are for cell 2 (CAF2).



3. Bias of Poststratified Estimators with Collapsing

In computing the bias and variance of an estimator in the presence of nonresponse and undercoverage, choices have to be made about what the mechanism is for computing expectations. In this section *i* denotes a poststratum and *k* is a unit within poststratum *i*. We assume that π_{ik} can be either the selection probability of unit (*ik*) or the propensity of providing data for the survey. In the latter case, π_{ik} is the selection probability divided by a nonresponse (NR) weight adjustment. The inverse of the NR weight adjustment is interpreted as the propensity of responding given that a unit is selected for the sample.

We assume that calculations of expectations can be done using the non-interview adjusted inclusion probability in the same way that a selection probability can be used. This, in effect, is assuming that the propensity of responding is correctly captured by whatever nonresponse adjustment is used. On the other hand, the undercoverage adjustments will be treated as fixed for the calculations here. This is, of course, unrealistic since the undercoverage adjustment is an estimate and will vary from one sample to another, but treating the adjustments as constants allows some simple points to be made.

The nonresponse-adjusted Horvitz-Thompson estimator is $\hat{X} = \sum_i \sum_{k \in s_i} \frac{x_{ik}}{\pi_{ik}}$ where s_i is the set of sample units in cell *i* and x_{ik} is the value observed for unit (*ik*). $\hat{X}$ is clearly biased if there is either undercoverage or overcoverage. Suppose that the poststrata are collapsed into $g = 1,\ldots,G$ cells with A_g denoting the set of poststrata assigned to group *g*. If f_g is the coverage adjustment applied to units in collapsed cell *g*, then the adjusted estimator of a total is

$$\hat{X}^* = \sum_g \sum_{i \in A_g} \sum_{k \in s_i} \frac{x_{ik} f_g}{\pi_{ik}} .$$

Treating π_{ik} as described above, the design-expectation of $\hat{X}^*$ is $E_\pi\left(\hat{X}^*\right) = \sum_g f_g X_g^c$ where X_g^c is the total for units in the portion of the population covered by the frame. Further, suppose that N_g is the number of units in the target population in cell *g*, X_g and $\bar{X}_g$ are the total and the mean per unit in that group, and that N_g^c and $\bar{X}_g^c$ are the corresponding figures for the covered population. The design-bias of $\hat{X}^*$ is

$$E_\pi\left(\hat{X}^* - X\right) = \sum_g \left(f_g X_g^c - X_g\right).$$

If f_g equals $X_g \big/ X_g^c = N_g \bar{X}_g \big/ \left(N_g^c \bar{X}_g^c\right)$, then the collapsed stratum estimator will be unbiased. Since the coverage adjustment f_g is computed as $N_g \big/ \hat{N}_g^c$ with $\hat{N}_g^c = \sum_{i \in A_g} \sum_{k \in s_i} 1/\pi_{ik}$, the adjustment used in standard practice will yield unbiased estimates if $\bar{X}_g^c = \bar{X}_g$, i.e., if the mean for the covered population in group *g* is the same as the mean for the target population. This is a strong assumption, and thinking of cases where it is dubious is not hard. For example, telephone surveys in the U.S. exclude non-telephone and cell-phone-only households. Internet surveys exclude persons without Internet access. In both instances, the responding sample may be poststratified to the entire population, but the excluded part of the population may be different in many ways from the covered part.

We can also analyze the bias of $\hat{X}$ with respect to an underlying model. Suppose that the population is reasonably described by a model in which units do have a common mean within the original poststrata, i.e., $E_M\left(x_{ik}\right) = \mu_i$ for units in both the covered and non-covered parts of the population. Then, the model-bias of $\hat{X}^*$ is

$$E_M\left(\hat{X}^* - X\right) = \sum_g \sum_{i \in A_g} \mu_i \left(f_g \hat{N}_i^c - N_i\right) \qquad (2)$$

where $\hat{N}_i^c$ is the estimated number of units in cell *i* that are in the covered portion of the population and N_i is the number of units in the full population. If $N_i \big/ \hat{N}_i^c$ is equal to f_g for each poststratum in the collapsed group *g*, then $\hat{X}^*$ will be model-unbiased. However, the more the cell coverage ratios, $N_i \big/ \hat{N}_i^c$, differ from the overall ratio, $f_g = N_g \big/ \hat{N}_g^c$, the more biased $\hat{X}^*$ will be as long as the cell means, μ_i, differ among the cells collapsed into a group. If $\mu_i \equiv \mu_g$ for each cell in group *g*, then the collapsed cell estimator will also be unbiased because $\sum_{i \in A_g}\left(f_g \hat{N}_i^c - N_i\right) = f_g \hat{N}_g^c - N_g = 0$, similar to the design-based result requiring $\bar{X}_g^c = \bar{X}_g$ for unbiasedness.

In summary, the collapsed stratum estimator will be

(1) Design-unbiased if the finite population means for the covered population and the target population are equal in each collapsed group that is formed, and
(2) Model-unbiased if either the cell coverage ratio or the cell mean is the same in each individual cell in a collapsed group.

If either or both of these is violated by the collapsing algorithm, then biased estimates of totals will result. Note that if a (full population) proportion or mean is estimated by $\hat{p}^* = \hat{X}^* \big/ \hat{N}^*$ where $\hat{N}^* = \sum_g f_g \hat{N}_g^c$, the same analysis as for $\hat{X}^*$ applies because $\hat{N}^* = N$, a constant.

4. Alternative Estimators and Remedies

We examine two alternative methods of weight computation when collapsing of poststrata is used. The alternatives are designed to be compromises between (a) use of all poststrata and the potential for large weight adjustments and (b) collapsing of strata yielding less variable weights but potentially biased estimates. The two alternatives presented in this section use cell collapsing but retain a larger share of the weight adjustment for individual cells than does the standard collapsing method. We refer to these as *weight restriction* (WR) methods.

The first alternative is denoted PS.WR1 and consists of the following algorithm. Denote the maximum allowable weight adjustment by $f_{\max}$.

(1) Compute the IAF's for the full set of poststrata. Any cell with $IAF > f_{\max}$ is designated as "sparse" and will be collapsed with another cell.
(2) Determine the "neighbors" of the cells that should be collapsed, i.e., the cell(s) they will be collapsed with.
(3) Censor any IAF greater than $f_{\max}$ to $f_{\max}$ and adjust each weight in the corresponding initial cell to $\tilde{w}_{ik} = w_{ik} f_{\max}$. For units in cells with $IAF \le f_{\max}$, set $\tilde{w}_{ik} = w_{ik}$.
(4) Cycle through the sparse cells one cell at a time. Collapse each sparse cell with the neighbor having the smallest value of the IAF. Compute the CAF for a collapsed group *g* as $\tilde{f}_g = N_g \big/ \tilde{w}_g^{(0)}$ where $\tilde{w}_g^{(0)} = \sum_{i \in A_g} \sum_{k \in s_i} \tilde{w}_{ik}$.
(5) The final adjusted weight is then $\tilde{w}_{ik} \tilde{f}_g$ for unit (*ik*) in group *g*.

This method will reduce the largest values of the final weight adjustment below the without-collapsing CAF's, though there may be one or more groups that have CAF's greater than the $f_{\max}$ cutoff.

In the two cell example in section 2.1, it can be shown that, if cell 1 has $IAF > f_{\max}$ and cell 2 does not, then the collapsing adjustment factors for cells 1 and 2 are

$$CAF_1 = \frac{f_{\max} f_2 (1+c)}{cf_1 + f_{\max} f_2} \text{ and } CAF_2 = \frac{f_1 (1+c)}{cf_1 + f_{\max} f_2} . \quad (3)$$

The final adjustments to weights in cells 1 and 2 are then $f_1 CAF_1$ and $f_2 CAF_2$. Example 2 applies PS.WR1 to the simple case presented in Example 1 of section 2.1.

Example 2. Setting $f_{\max} = 2$ and substituting the values from Example 1 into (3) yields CAF_1 = 0.7 and CAF_2 = 1.4. These compare to the initial factors of 0.438 and 1.75 and in the earlier example. Thus, PS.WR1 shrinks the collapsing adjustment factors toward 1, i.e., the final adjustments are somewhat closer to the initial individual cell adjustments before collapsing. Assuming that the IAF's are the best coverage corrections, PS.WR1 retains more of the individual cell coverage corrections than does standard collapsing. In this case, the final weight adjustments are $f_1 CAF_1$ = 2.8 and $f_2 CAF_2$ = 1.4 compared to a common weight adjustment of 1.75 with standard collapsing in Example 1. As a result, PS.WR1 does create a disparity in weights but not as great as would result if the two poststrata were kept separate.

The second alternative is denoted PS.WR2 and is intended to exercise more control over the size of the final weight adjustment than does PS.WR1. In Example 2 the final adjustment was 2.8 for cell 1 rather than $f_{\max} = 2$. PS.WR2 limits the final adjustment to 2 or some other maximum set in advance. The general idea is to first determine which cells should be collapsed together, as was done for PS.WR1. Then weights in the sparse cells are multiplied by $f_{\max}$. The weights in the non-sparse cell in a collapsed group are then adjusted by a constant factor to bring the estimated population count in the group to the control count. The detailed algorithm for computing weights for PS.WR2 is the following:

(1) Compute the IAF's for the full set of poststrata. Any cell with $IAF > f_{\max}$ is designated as sparse. Denote the set of sparse cells as C_{sp}.
(2) Determine the neighbors of the cells that should be collapsed.
(3) Cycle through the sparse cells one cell at a time ($i \in C_{sp}$). Collapse each sparse cell with the non-sparse neighbor having the smallest value of the IAF. Suppose that, at step *k*-1, two or more cells were collapsed to form a group, *g*. If, at step *k*, the non-sparse neighbor of sparse cell i' with the minimum IAF is in *g*, then join cell i' to group *g* to form a larger group.
(4) The collapsing process terminates when all sparse cells are collapsed into groups. Most groups will have one non-sparse cell; in some cases a group can have only sparse cells. Denote the set of sparse cells in group *g* as $A_{g,sp}$ and the non-sparse cell in group *g* as $A_{g,\overline{sp}}$.
(5) In a group containing at least one non-sparse cell, compute the control total in group *g* as $N_g = \sum_{i \in A_g} N_i$ and the adjusted weight for all units (*ik*) in $A_{g,sp}$ as $\tilde{w}_{ik} = w_{ik} f_{\max}$. Compute the adjusted weight for all units (*ik*) in $A_{g,\overline{sp}}$ as $\tilde{w}_{ik} = w_{ik} \left(N_g - \hat{\tilde{N}}_{g,sp} \right) \Big/ \hat{N}_{g,\overline{sp}}$ where $\hat{N}_{g,\overline{sp}} = \sum_{i \in A_{g,\overline{sp}}} \sum_{k \in s_i} w_{ik}$ and $\hat{\tilde{N}}_{g,sp} = \sum_{i \in A_{g,sp}} \sum_{k \in s_i} \tilde{w}_{ik}$. In a group with only sparse cells, compute the weight using regular poststratification, i.e., $\tilde{w}_{ik} = w_{ik} N_g \big/ \hat{N}_g$.

The final adjusted weight is then $\tilde{w}_{ik}$ for unit (*ik*) in group *g*.

Example 3. Continuing with the same data as in Examples 1 and 2, suppose the desired weight adjustment for units in cell 1 is $f_{\max} = 2$. The weight adjustment for units in cell 2 is 1.667. These compare to a common weight adjustment of 1.75 with standard collapsing in Example 1 and adjustments of 2.8 and 1.4 for PS.WR1. Thus, the weight disparity created by PS.WR2 is not as great as would result if the two poststrata were kept separate or if PS.WR1 were used.

5. An Empirical Investigation

To test some of the ideas presented earlier, we conducted a simulation study of the bias properties of alternative methods of poststratification. We also examined the performance of some of the variance estimators that are often used in practice.

5.1 Study Population

The population used in the simulation was extracted from the 2003 National Health Interview Survey (NHIS) person public-use file. A subset of the NHIS was created with 21,664 persons. These were divided into 25 strata with each having six primary sampling units (PSU's). The strata and PSU's are based on those in the NHIS public use file, but sets of three strata were collapsed together to create new strata for the study population. We used four binary variables (0-1 characteristics) for the simulation, each of which is based on a person's self-report:

Health insurance coverage (NOTCOV)—whether a person was covered by any type of health insurance;

Physical, mental, or emotional limitation (LA1AR)—whether a person was limited in any of these ways;

Medical care delayed (PDMED12M)—whether a person delayed medical care or not because of cost in last 12 months

Overnight hospital stay (PHOSPYR)—whether a person stayed overnight in a hospital in last 12 months

Table 1 shows the percentages of persons with these four characteristics in cells formed by age and sex. These 16 (age × sex) cells are the initial set of poststrata used in estimation. The percentages can vary substantially among the cells, depending on the characteristic. For, example, 18-24 year olds are much more likely to have no health insurance; children under age 5 and the elderly age 65 and over are much more likely to have had a hospital stay. Collapsing some of these cells together for estimation has the potential to introduce bias, as noted earlier.

5.2 Sample Design

Two sample PSU's were selected in each stratum with probability proportional to size (PPS) with the size being the count of persons in each PSU. Sampling of PSU's was done with-replacement to simplify variance estimation. In each sample PSU, 20 persons were selected by simple random sampling without replacement for a total of 1,000 persons in each sample. For each combination of parameters discussed below, 2,000 samples were selected.

Sixteen initial poststrata were used which were the cross of the eight age groups, shown in Table 2, with gender. In each sample, we computed the estimators of proportions described earlier in sections 2-4—the Horvitz-Thompson ratio estimator, $\hat{p}_{\pi} = \left(\sum_{i\in s} x_i/\pi_i\right) \Big/ \left(\sum_{i\in s} 1/\pi_i\right)$, the poststratified estimator described in sec. 2.1, denoted by $\hat{p}_{PS1}$, that uses all 16 poststrata, the poststratified estimator with collapsing of cells, $\hat{p}_{PS2}$, and the two weight-restricted estimators, $\hat{p}_{PS.WR1}$ and $\hat{p}_{PS.WR2}$. Each of the estimators $\hat{p}_{PS1}$, $\hat{p}_{PS2}$, $\hat{p}_{PS.WR1}$, and $\hat{p}_{PS.WR2}$ has the form $\left(\sum_{i\in s} w_i x_i\right) \Big/ \left(\sum_{i\in s} w_i\right)$ where w_i is the weight for unit *i* computed from the particular method of poststratification. For each of the estimators of a proportion, a linearization variance estimate was calculated. (Note that, for the *PS* estimators, the sum of weights across the full sample is a constant—the population count *N*—and linearization is unnecessary.) The simulation code was written in the R language (R Development Core Team 2005) with extensive use of the R survey package (Lumley 2004).

5.3 Coverage mechanisms

Four sets of coverage mechanisms were employed to filter the population before the PSU's were sampled. The coverage ratios listed in Table 2 varied by poststratum and were different for each of the four characteristics for which proportions were estimated. For example, if the coverage ratio in the poststratum of males younger than 5 years old is 0.9, then 90% of the population in that poststratum were randomly selected to stay in the sampling frame while the rest had a zero probability to be sampled. The coverage ratios, named C1 through C4, in Table 2 were artificially created based on the population means for each age and sex group. Poorer coverage was assigned to groups with larger percentages with a characteristic.

5.4 Collapsing rules

We set up situations where one or both of the conditions for unbiasedness in section 3 can be violated when cells were collapsed in the simulations. Each of the estimators, $\hat{p}_{PS2}$, $\hat{p}_{PS.WR1}$, and $\hat{p}_{PS.WR2}$ involve cell collapses. If the IAF poststratification factor in an initial poststratum, $f_i = N_i / w_i^{(0)}$, exceeds the maximum allowable adjustment, we call this poststratum a "sparse" cell and collapse it with the neighboring cell which has the smallest poststratification factor. In this section, we denote the maximum allowable adjustment by A_U, which corresponds to $f_{\max}$ for PS2, PS.WR1, and PS.WR2. The neighbors of a specific cell are defined as the cells either horizontally or vertically adjacent to it in the age × sex table. For example, the neighbors of the cell 3 are cells 2, 4, and 7, shaded in the following, abbreviated table.

1	5
2	6
3	7
4	8

Two different values of A_U were used in the simulations— A_U =2 and 1.8. Use of A_U =1.8 leads to more collapsing of cells than A_U =2 and exhibits more of the biases noted in section 3.1 caused by combining of cells with different means. Of course, in practice many variations are used to decide which combinations of cells are allowable. We have used just one of the possibilities for illustration in the simulation.

Once all of the sparse cells and their neighbors with the minimum poststratification factor are identified, the collapsing process proceeds sequentially from cell 1 using the methods in sections 3 and 4. In a survey with many potential poststrata defined in advance, these procedures might have to be performed iteratively to obtain weight adjustments that respect the desired bounds. In this simulation, we performed only one round of collapsing.

5.5 Simulation Results

Tables 3-5 summarize results for relative biases of estimated proportions, variances of alternative estimators, and confidence interval coverage using linearization variance estimators. The HT estimates, shown in Table 3, are badly negatively biased since they include no correction for the undercoverage that is highest in cells where the population proportions are highest.

Poststratification with no collapsing of cells (PS1) gives unbiased estimates while the alternatives—PS2, PS.WR1, and PS.WR2—all introduce a bias due to collapsing (see Table 3). The relative biases of PS2, which collapses cells whose PS factor is greater than or equal to A_U, range from -6.1% to -0.6 when A_U =2 and from -8.2 to -0.5 when A_U =1.8. The alternatives, PS.WR1 and PS.WR2, have biases for A_U =2 that are intermediate between PS1 (no collapsing) and PS2 (standard collapsing). PS.WR2 has biases that are greater in absolute value than those of PS.WR1 due to its more extreme weight restrictions. PS.WR1 is generally competitive with PS1 in terms of bias.

One justification that is conventionally given for collapsing cells is that extreme weights will be reduced and variances of estimates will, in turn, be reduced. Table 4 shows the ratios of the empirical variances of estimated proportions as a proportion of the variance of PS1. The HT estimates do have variances that are 12 to 24% smaller than those of PS1, but, of course, HT is badly biased. There are some precision gains from using PS2. For example, with A_U =1.8 the variance of hospitalized with PS2 is 83.4% of that of PS1, but this is at the expense of a -8.2% bias for PS2. PS.WR1 either reduces variances only slightly or increases variances by 4 to 5% compared to PS1. PS.WR2 reduces variance mainly in cases where it introduces bias.

Table 5 reports the empirical coverages of 95% CI's computed using the estimated proportions and the linearization variance estimator that naturally accompanies each. The HT coverage rates are extremely poor, as expected, ranging from 72.0% to 84.8%. Among the poststratified estimators, PS1 and PS.WR1 cover at near the nominal rate of 95%. In contrast, PS2 coverage is poor for health coverage and hospitalization, especially when A_U =1.8 where the coverages are 87.6% and 85.2%. PS.WR2 has coverage rates nearer to 95% than PS2 and is competitive with PS1 and PS.WR1.

6. Concluding Remarks

Designers of surveys of households or establishments often have a lengthy list of poststrata in mind when they develop weighting systems. However, if the sample size in a poststratum is small or the sample estimate of the population count in a poststratum is much less than an external control count, the poststratum may be collapsed with another adjacent one. The conventional justification for collapsing is that the possibility of creating extreme weights is reduced as are variances of estimates.

However, collapsing has at least two undesirable consequences: (1) deficient frame or sample coverage in some cells is not completely corrected and (2) estimates from the standard approach to collapsing may be quite biased. The latter problem can result in confidence intervals that cover at much less than the nominal rate. Because of points (1) and (2), retention of all individual poststrata may be preferable to collapsing, even though collapsing does, in fact, reduce the variance of estimates.

Practitioners should address two issues before implementing any collapsing procedure:

(1) What are the estimated means of important survey variables within each of the planned poststrata?
(2) What are the estimated population coverage rates, $N/\hat{N}$, in each of the planned poststrata?

Collapsing cells with substantially different means or coverage rates can introduce bias that may not be offset by any reduction in variance due to cell collapsing.

If collapsing of cells is used, the alternative methods of restricting weight adjustments, PS.WR1 and to a lesser extent, PS.WR2, are good options for retaining more of the bias correction afforded by full poststratification with no collapsing.

References

Gonzalez, J.F., Town, M., and Kim, J. (2005). Mean Square Error Analysis of Health Estimates from the Behavioral Risk Factor Surveillance System for Counties along the United States/Mexico Border Region, *Proceedings of the Section on Survey Methods Research*. Alexandria VA: American Statistical Association.

Kalton, G. and Maligalig, D.S. (1991). A Comparison of Methods of Weighting Adjustment for Nonresponse. *Proceedings of the U.S. Bureau of the Census Annual Research Conference*, 409-428.

Kim, J. (2004). Effects of Collapsing Rows/Columns of Weighting Matrix on Weights, *Proceedings of Section on Survey Research Methods*, American Statistical Association.

Kim, J., Thompson, J.H., Woltman, H.F. and Vajs, S.M. (1982). Empirical Results from the 1980 Census Sample Estimation Study. Proceedings of Section on Survey Research Methods, American Statistical Association, pp 170-175.

Kostanich, D., and Dippo, C. (2000). *Current Population Survey: Design and Methodology*. Technical paper 63. Washington DC: Department of Commerce.

Little, R.J.A. (1993). Post-Stratification: a Modeler=s Perspective, *Journal of American Statistical Association*, **88**, 1001-1012.

Lazzeroni, L., and Little, R.J.A. (1998). Random-effects Models for Smoothing Poststratification Weights. *Journal of Official Statistics*, **14**, 61-78.

Lumley, T. (2004). Analysis of complex survey samples, *Journal of Statistical Software*, **9,** 1-19.

R Development Core Team (2005), *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, http://www.R-project.org.

Tremblay, V. (1986). Practical Criteria for Definition of Weighting Classes. *Survey Methodology*, **12**, 85-97.

Table 1. Percentages of persons with four health-related characteristics in groups formed by age and sex.

	Not covered by health insurance			Physical, mental, emotional limitations			Delayed medical care in last 12 months			Hospital stay in last 12 months		
Age	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
< 5	10	9	9	4	3	3	3	4	3	17	15	16
5-17	13	14	13	10	6	8	4	4	4	2	1	2
18-24	37	31	34	4	4	4	8	11	9	3	14	8
25-44	28	23	25	7	7	7	9	10	9	3	10	6
45-64	14	14	14	16	19	18	7	11	9	8	10	9
65-69	2	1	2	24	29	27	3	8	6	15	14	14
70-74	1	1	1	34	32	33	2	5	4	18	15	17
75+	1	1	1	41	48	45	2	2	2	22	22	22
Total	18	16	17	12	13	13	6	8	7	7	10	8

Table 2. Coverage ratios used in the simulations.

	C1: Not covered by health insurance		C2: Physical, mental, emotional limitations		C3: Delayed medical care in last 12 months		C4: Hospital stay in last 12 months	
Age	Male	Female	Male	Female	Male	Female	Male	Female
< 5	0.9	0.9	0.9	0.9	0.9	0.9	0.5	0.5
5-17	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
18-24	0.5	0.5	0.8	0.8	0.6	0.5	0.8	0.5
25-44	0.5	0.5	0.8	0.8	0.6	0.5	0.8	0.5
45-64	0.8	0.8	0.7	0.7	0.6	0.5	0.8	0.5
65-69	0.9	0.9	0.6	0.6	0.9	0.5	0.5	0.5
70-74	0.9	0.9	0.5	0.5	0.9	0.7	0.5	0.5
75+	0.9	0.9	0.5	0.5	0.9	0.8	0.5	0.5

Table 3. Percentage relative biases of estimated proportions

Characteristic	Horvitz-Thompson	PS1 (no collapsing)	PS2 (standard collapsing)	PS.WR1 (truncate weights then collapse)	PS.WR2 (fixed maximum weight adjustment)
Adjustment bound = 2					
Health coverage	-11.4	0.3	-4.2	1.1	-1.2
Limitations	-10.4	-0.1	-0.6	0.0	-0.6
Delayed care	-9.6	-0.8	-2.2	-0.3	-1.4
Hospitalized	-13.3	0.0	-6.1	0.8	-2.3
Adjustment bound = 1.8					
Health coverage	-11.5	0.1	-6.4	1.0	-3.3
Limitations	-10.2	0.1	-0.5	0.2	-0.7
Delayed care	-9.4	-0.4	-2.6	0.4	-2.0
Hospitalized	-13.1	-0.1	-8.2	0.5	-4.6

Table 4. Ratio of variances to the variance of the poststratified estimator (PS1) with no collapsing.

Characteristic	Horvitz-Thompson	PS2 (standard collapsing)	PS.WR1 (truncate weights then collapse)	PS.WR2 (fixed maximum weight adjustment)
Adjustment bound = 2				
Health coverage	0.870	1.019	1.031	0.998
Limitations	0.879	0.973	1.003	0.981
Delayed care	0.831	0.966	1.045	0.977
Hospitalized	0.760	0.905	1.036	0.944
Adjustment bound = 1.8				
Health coverage	0.877	0.958	1.038	0.974
Limitations	0.867	0.967	0.994	0.978
Delayed care	0.828	0.941	1.050	0.956
Hospitalized	0.760	0.834	1.027	0.887

Table 5. Coverage of 95% confidence intervals.

Characteristic	Horvitz-Thompson	PS1 (no collapsing)	PS2 (standard collapsing)	PS.WR1 (truncate weights then collapse)	PS.WR2 (fixed maximum weight adjustment)
Adjustment bound = 2					
Health coverage	76.4	94.1	90.2	94.7	93.6
Limitations	76.0	94.0	94.2	94.6	94.2
Delayed care	84.8	94.0	93.0	94.0	93.9
Hospitalized	72.3	95.0	89.0	95.0	94.0
Adjustment bound = 1.8					
Health coverage	76.0	93.9	87.6	94.5	92.8
Limitations	77.7	93.8	94.0	94.0	94.1
Delayed care	84.2	93.3	92.8	93.7	93.3
Hospitalized	72.0	93.7	85.2	93.8	91.7