Paper 1627-2014

A Rake-Trim SAS® Macro and Its Uses at Westat

Lou Rizzo, Westat

ABSTRACT

Raking (iterative proportional fitting) is a procedure that takes sampling weights from complex sample surveys and adjusts them so that they add to known control totals. This process reduces variance and adjusts for undercoverage. But raking in multiple dimensions can lead to extreme weights, which increase variance. Trimming is another sample weighting procedure that reduces extreme weights to cutoffs, thereby improving variance properties while potentially introducing bias. The RAKE-TRIM macro combines raking and trimming in an iterative algorithm to achieve these two goals simultaneously. The raking reduces the bias potential from trimming, and the trimming reduces the variance inflation from raking. When convergence occurs, the final weights aggregate to the control totals, as well as respect the trimming limits. SAS® macros are well suited for this kind of envelope program: the larger macro consists of the integration of component macros that were developed for other applications. A parameter specification sheet enables users to provide all of the parameters needed to define the algorithm for their particular situation, and, if necessary, to alter the parameters to facilitate convergence. Diagnostics are included when convergence fails. Microsoft Excel tables are imported to provide the cell structure and are exported to provide statistics for the algorithm's results. This RAKE-TRIM macro was first developed in 2010 for the 2009 National Household Transportation Survey and has been used in other studies as well. The paper describes the algorithm and discusses our experiences with it.

INTRODUCTION

At Westat, we have long had a SAS macro we developed called FSRAKE that carried out iterative proportional fitting ("raking") of sampling weights to fixed control totals. A companion macro REPRAKE carried out a parallel raking of replicate weights for valid variance estimation using raked weights. In applying FSRAKE and REPRAKE to the survey weights from the 2009 National Household Travel Survey[1] (NHTS), sponsored by the Federal Highway Administration, we found a need for a capacity for trimming the weights in addition and in conjunction with the raking. That led to the development of the 'RAKE-TRIM' macro.

This paper puts this macro into the context of the history of raking and trimming, describes other software which carries out similar purposes, and describes the macro itself and its features. Below is provided a short history of raking and trimming, a description of the macro and its characteristics, an outline of the performance of the macro under simplifying scenarios, a discussion of aspects of the macro which facilitate convergence, and a final discussion and summary.

SHORT HISTORY OF RAKING AND TRIMMING FOR WEIGHT CALIBRATION

Raking, or iterative proportional fitting (IPF), was first presented in papers by Deming and Stephan (1940) and Stephan (1942). This method has been used extensively in the fitting of models to contingency tables (cite for example Bishop et al. (1975)). Its application to surveys is in the calibration of survey weights to external control totals, which represent population totals with no sampling error (or at least sampling error much lower than the survey in question). If this calibration is done to one set of control totals (one 'dimension'), then this is called poststratification. If the calibration is done to more than one set of control totals (multiple 'dimensions'), then raking is generally the accepted methodology. Deville and Särndal (1992) and Deville, Särndal, and Sautory (1993) studied in the context of survey weight calibration raking and a number of alternatives to raking based on slightly different algorithms. The starting point is a set of design weights d_k which are the reciprocals of the design probabilities of selection. Weighted sums of survey characteristics y_k using these weights are sample-design unbiased (the well-known Horvitz-Thompson estimator). A set of control totals $\mathbf{t}_x = \sum_U \mathbf{x}_k$ for a control vector $\mathbf{x}$ are available based on an external source (where the summation is taken over the finite population). A set of calibrated weights w_k can be

[1] U.S. Department of Transportation, Federal Highway Administration, 2009 National Household Travel Survey. Available at http://nhts.ornl.gov

created such that the weighted sum of the sample $\mathbf{x}_k$ values using the w_k are exactly equal to the control totals $\mathbf{t}_x$ ($\sum_s w_k \mathbf{x}_k = \sum_U \mathbf{x}_k$), while being as close as possible to the design weights d_k. Closeness is defined in terms of a distance measure. Deville and Särndal (1992) proved under certain conditions on the population, the design weights, and the distance measures, that weighted estimates for any characteristics y_k using the calibrated weights will be asymptotically design-unbiased, and in particular any characteristics y_k which are correlated to the control vector $\mathbf{x}_k$ in the population will have reduced variance as compared to the weighted estimator based on the design weights d_k. In the extreme case the weighted sum of the characteristic $\mathbf{x}_k$using the calibrated weights will have variance zero, as this by definition is equal to the fixed control totals $\mathbf{t}_x$.

In particular, Deville, Särndal and Sautory (1993) studied several alternatives to raking based on slightly different algorithms, which were implemented in a specialized software package called CALMAR. One of these was the 'logit (L,U) method', which used the same basic algorithm as raking, but put lower and upper bounds on the raking adjustments. This was designed to avoid extreme weights coming out of the raking process. The methodology does nothing about weights that are extreme or close to extreme before beginning the calibration, as it only bounds *the adjustments*, thus there is no complete control over the final weight distribution.

In general, these alternatives in Deville and Särndal (1992) and Deville, Särndal, and Sautory (1993) were found to make very small differences in the estimators ultimately as compared to the well-established raking methodology. Later research began to focus on stronger differences which had a greater effect on the weights, but did not have the asymptotic unbiasedness property of the earlier estimators. Särndal (2007) provides a very accessible summary of this work. One alternative of interest is the 'instrument vector method'. Under this methodology, weights w_k are developed which satisfy the calibration equations $\sum_s w_k \mathbf{x}_k = \sum_U \mathbf{x}_k$. In this case however the w_k are defined as $w_k = d_k(1 + \boldsymbol{\lambda}'\mathbf{z}_k)$, where $\mathbf{z}_k$ is an instrumental vector not necessarily equal to $\mathbf{x}_k$, and $\boldsymbol{\lambda}$ is solved to satisfy the calibration equations $\sum_s w_k \mathbf{x}_k = \sum_U \mathbf{x}_k$. Generally $\mathbf{z}_k$ is some function of $\mathbf{x}_k$. As Särndal describes, even apparently extreme definitions of $\mathbf{z}_k$, such as ones in which $\mathbf{z}_k$ is nonzero for only a small number of values, give surprisingly good results, as long as the calibration equations are satisfied. But asymptotic design-unbiasedness is lost.

The most comprehensive early paper for summarizing methods for trimming in surveys was Potter (1990). The general approach is to trim sample-design weights beyond a specified cutpoint to be equal to that cutpoint, followed by a re-proportioning of the weights below the cutpoint so that the weight sum does not change. Note that this may become an iterative process, as some re-proportioned weights may now exceed the cutpoint. This will reduce variance, but opens up the possibility of bias. Trimming of any magnitude will lead to weights that cannot be justified as being sample-design unbiased. Potter (1990) recommends the explicit computation of mean-squared-error estimates for important survey characteristics using several proposed cutpoints in order to inform a rational decision about this variance-bias tradeoff.

Some form of trimming is used in many large-scale surveys that have a wide weight distribution. That has become the general practice in the presence of extreme weights. For example, Chowdhury et al. (2007) indicates that the National Immunization Survey trims weights that are larger than the median plus six times the interquartile range.

RAKING AND TRIMMING: ALWAYS TOGETHER?

Raking and trimming both have inherent weaknesses which can be mitigated by the other. Raking can lead to extreme weights if the marginal weighting constraints are difficult to satisfy with the distribution of original weights as given. Trimming is a natural mechanism for alleviating this problem in raking. In addition, trimming causes bias if weights are trimmed back for observations with certain characteristic values. Raking can alleviate this by re-balancing the trimmed weights to fixed control totals. Though the pairing of raking and trimming seems natural, it doesn't seem that the two have been paired as a matter of course. The references to trimming such as Potter (1990) recommend rebalancing the weights after trimming so that the summation of the final weights is the same as the summation of the original weights, but this rebalancing is only done at the universe level.

The current version of SUDAAN (SUDAAN 11.0) provides raking algorithms (WTADJUST and WTADJX) for survey weights with a 'pre-trim' option. This pre-trim option allows the initial weights to be bounded by lower and upper bounds, and adjusts the initial weights by resetting weights outside of these limits to these bounds. These trimmed weights become the input for the raking algorithms. In addition, the raking algorithm implements the logit (L,U) method as defined in Deville et al. (1993), which allows a bounding of the raking adjustments. SUDAAN sees the raking procedure as a necessary complement to the trimming procedure, and in fact does not offer a stand-alone trimming option (trimming must be used in conjunction with a following poststratification step).

The limitation of the SUDAAN procedure is that the weights coming out of this pre-trim—raking algorithm may in fact still be extreme. Further iterations between trimming and raking are necessary to guarantee a full bounding. RAKE-TRIM does this.

THE RAKE-TRIM ALGORITHM

The RAKE-TRIM algorithm is an 'envelope' program which puts a stand-alone raking SAS Macro called FSRAKE into an iterative SAS macro, with trimming code included preceding the main iteration, and within the main iteration. The general sequence is then a pre-trim step followed by a raking step followed by a post-trim step followed by raking/post-trim pairs, until convergence of the process is achieved. The raking/post-trim pair is called a 'cycle', to distinguish it from the iterations within each raking step.

THE RAKE-TRIM ALGORITHM: TRIMMING STEPS

The trimming steps allow for both an upper and a lower bound on the weights. In many trimming applications, only an upper bound is set, with trimming of extremely large weights. Extremely small weights can also be a source of loss of efficiency however: even though the sample units are in the sample, their weights are so small as to make their marginal contribution to the weighted estimate very small. This makes the sampling variance larger as much as having extremely large weights. A lower bound for which weights smaller than this lower bound are increased up to equal this bound will reduce the negative effects of this on sampling variance.

Lower and upper bounds are defined in terms of multiples of the median of the weights within a defined trimming cell. The user specifies this, but for the upper bound it is generally in survey applications something like 3.5, 4.0, 4.5. The lower bound, if specified, might be the reciprocal of the upper bound factor, but this is not necessary. Another option (recommended) is to specify a maximum percentage of trimmed observations as well as a cutoff. If the percentage of trimmed observations (observations with weights larger than an upper bound or smaller than a lower bound) is greater than this maximum, the cutoff is moved (upwards for an upper bound, and downwards for a lower bound) so that only the maximum percentage is trimmed. This kind of maximum percentage bounds in the potential bias from trimming, as it makes sure that only a specified percentage of sample units have altered weights. Maximum percentages might be 1%, or 5%, and are set by the user.

The trimming cells should be mutually exclusive and exhaustive. In many applications, appropriate trimming cells will be sampling strata, especially if the sampling rates (and weights) differ between strata. If one has for example a stratum with a much lower sampling rate (and much higher weights), then trimming may result in the entire stratum's weights being trimmed back. This will lead to the entire stratum being systematically underrepresented. The stratification structure and its relative sampling rates and weights should be preserved by selecting trimming cells that are nested within sampling strata.

The replicate weights are trimmed using the same trimming ratios as were computed for the main sample weights. This follows standard practice for trimming in the presence of replicate weights. No alternatives to this have been developed as of yet, to the best knowledge of the author.

THE RAKE-TRIM ALGORITHM: RAKING STEPS

The raking portion of the program uses iterative proportional fitting to adjust the current weights to the control totals. The control totals are input to the algorithm using Microsoft© EXCEL workbooks, utilizing the easy interface of SAS with Microsoft© EXCEL. There is a separate EXCEL sheet for each dimension of the raking, and the EXCEL sheets contain the cells and the control totals associated with the cells. These EXCEL sheets also have columns for collapsed cells for each dimension. This may be necessary to facilitate convergence of the algorithm. The program automatically defines SAS fields and aggregates control totals for the collapsed cells, to ease the burden of this necessary work on users.

THE RAKE-TRIM ALGORITHM: THE FULL ITERATION

The RAKE-TRIM algorithm has three basic versions, governed by a parameter called FLAG: FLAG=1, FLAG=2, and FLAG=3. The FLAG=1 run does the pre-trimming step (if pre-trimming is needed), followed by a first raking step. This first FLAG=1 program is designed to see if the single raking run converges. If it does not, then collapsing of cells in some of the dimensions may be necessary. The output of the raking program allows the user to see which cells might be problematic. There are options for producing extensive intermediate check files to check all iterations of the raking program to find problematic cells. Several runs of this first FLAG=1 program may be necessary to define a cell structure which allows for successful convergence.

The FLAG=2 run carries out the pre-trimming step, the first raking step, and the first raking step for replicate weights. Further collapsing of cells may be necessary to make sure that each replicate weight converges as well as the

primary weights. This allows for valid variance estimation. The program is now ready for the full RAKE-TRIM run: FLAG=3.

The FLAG=3 run carries out the full set of cycles: pre-trim followed by pairs of raking—post-trim cycles. Convergence is determined by a TRIM_DELTA parameter. After doing the most recent raking run, the sum of weights within each raking cell will add to the control totals up to a raking tolerance factor (which can be set very small). The next post-trim is done, and the trimming factors are compared to a trimming tolerance TRIM_DELTA. If some trimming factors are outside of the interval [1-TRIM_DELTA, 1+TRIM_DELTA], then the post-trim is carried through, and the next cycle is initiated (starting with the next raking iteration) on the post-trimmed weights. If all trimming factors are within the interval [1-TRIM_DELTA, 1+TRIM_DELTA], then the algorithm is viewed as convergent, and the weights at the end of last raking run are viewed as final. The post-trim is not carried through (so that the weight sums can match the control totals). TRIM_DELTA should be fairly large (0.01 for example). A too tight TRIM_DELTA (such as 0.0000001) makes convergence impossible without a huge number of iterations. This means in practice that some of the weights will be slightly beyond the trimming cutoffs by this margin. If it is absolutely necessary to achieve certain trimming bounds exactly, then tighter trimming bounds than desired should be set, and the algorithm carried through with these tighter bounds.

The program allows for differing trimming parameters for the initial 'pre-trim' and the followup 'post-trim' steps. This reflects the differing purposes of these two trimming steps. The pre-trim step is designed to do a trimming of the initial weights preceding any raking. This will cover any extreme base weights, or extreme weights coming out of weighting procedures such as nonresponse adjustments preceding the raking step. The post-trim steps on the other hand are designed to handle any extreme weights which are coming out of the raking process itself. This is a somewhat different purpose, and differing parameters for pre-trim and post-trim may be useful.

It should be noted that for the post-trim the maximum percentage of trimmed values is assigned for each iteration of post-trim alone. Some units will be trimmed in some post-trim steps and not trimmed in others. Thus the total percentage of units which had post-trimming at some point in the iterative process will generally be somewhat larger than the maximum percentage set for the post-trim process. It is hard to anticipate how many will be ultimately trimmed at least once, thus if an overall maximum percentage needs to be imposed for post-trimming, it may be necessary to do several runs with varying parameter values to achieve the goal desired. In addition, sample units may or may not be both pre-trimmed and post-trimmed. The program produces a final file which contains the 'omnibus adjustment': the ratio of the final weight for the sample unit as compared to the initial weight. This is the product of all of the adjustment factors throughout the iterative algorithm.

SCENARIOS

Following Brick et al. (2003) we can write models for the unadjusted and raked weighted percentages for a two-dimensional population:

$$ln(e_{ij}) = \alpha_i + \beta_j + (\alpha\beta)_{ij} \quad ln(u_{ij}) = \alpha'_i + \beta'_j + (\alpha\beta)'_{ij}$$

Any convergent RAKE-TRIM process will give the same raked row and column marginal multiplicative factors α'_i and β'_j. The primary purpose of raking is to create a new set of percentages u_{ij} that are close to the e_{ij}, but with new row and column marginal multiplicative factors α'_i and β'_j replacing the old row and column factors α_i and β_j. These new factors are in complete agreement with the control totals. The differences between RAKE-TRIM and simple raking is entirely in the $(\alpha\beta)'_{ij}$ factors.

As is stated in Brick et al. (2003), the raking process does not determine the $(\alpha\beta)'_{ij}$ factors, and tends to leave them close to the original $(\alpha\beta)_{ij}$ factors. This will facilitate approximate design-unbiasedness, as the $(\alpha\beta)_{ij}$ come from the unadjusted sample-design weights. The RAKE-TRIM algorithm by trimming some weights will tend to give a different set of $(\alpha\beta)'_{ij}$ factors which will tend to be more divergent from the $(\alpha\beta)_{ij}$ factors.

Several scenarios will be analyzed. The first set of scenarios (in Section 4.1) describe scenarios where there are extreme weights going into the algorithm (a need for pre-trimming). The second set of scenarios (in Section 4.2) studies the raking/post-trim cycles, assuming there are either no extreme initial weights, or these weights are successfully dealt with in the pre-trimming step.

SCENARIOS WITH PRE-TRIMMING NECESSARY

The first scenario is the 'haphazard large and small weights' scenario. Under this scenario, the extreme weights on either end are generated by sampling anomalies whose presence can be seen as entirely random with respect to the raking cells, and to most survey characteristics. In this case, trimming will tend an average to leave the $(\alpha\beta)_{ij}$ factors

alone, as the presence of extreme weights is uniform across the ij-cells. Iterating between raking and trimming will not affect the final distributions to the ij-cells, but the extreme weights within each ij-cell will be trimmed back. Very limited biases at most would be endured, and precision improved. A winning scenario for rake-trim certainly. Also, since the presence of extreme weights is scattered over the ij -cells, convergence will occur very quickly (within one or two iterations of the algorithm).

The ‘haphazard large and small weights’ scenario is the best from the viewpoint of the results of the algorithm on sampling properties and in terms of convergence. On the other extreme of the spectrum is the ‘marginal extreme weights’ scenario. Again we assume that the sampling strata are one dimension in the raking process, but the trimming cells are not chosen to be identical to the sampling strata. This scenario is pathological in the sense that the algorithm may not be able to converge at all, and may in fact bounce back and forth between raking and trimming adjustments indefinitely.

In the simplest form of this scenario, we have one or more sampling strata with a much lower sampling rate or a much higher sampling rate, so that most of the base weights (reciprocals of the probability of selection) are in fact beyond the trimming bound. For example, the trimming bound may be 10 times the median weight in a trimming cell, and one stratum may have ***all*** its base weights greater than 20 times the median base weight. If the rake-trim algorithm is used in this case, the trimming adjustment will trim all the weights back uniformly, but then the raking step will increase the weights immediately back to what they were before (as the sampling stratum is a margin whose summation needs to be respected). The two steps will contradict each other, and no convergence will result.

In terms of the multiplicative factors analysis for the two-dimensional case, the trimming makes it impossible to achieve one or more factors α_i' or β_j'. The algorithm has to fail. In this case, one of the conditions imposed on the system has to be relaxed. If the sampling strata are to be respected (necessary for approximate sample design unbiasedness), then the larger or smaller weights for sampling strata with differing sampling rates have to be respected. Trimming cannot be forced on these weights as a group. This can be done in practice by making sure the trimming cells are the same as the sampling strata.

These two scenarios mark opposite extremes. In one case, the presence of an extreme weight is entirely uncorrelated with the raking cells. In the other case, the presence of extreme weights is highly correlated with one or more of the marginal cells: one of the marginal cells has all extreme weights. The third scenario is the ‘marginal balance—cross-cell imbalance’ scenario. This is the most difficult scenario. Convergence will occur, but it may require many iterations. In this case, the extreme weights are well-distributed across the marginal cells, but tend to pile up within particular cross cells. In this case, the rake-trim algorithm will tend to re-distribute weight between the cross-cells as it matches the marginal cell totals following each raking step. For example, suppose cell i^*j^* has a high prevalence of very large weights. A raking run without any trimming will achieve

$$ln\left(u_{i^*j^*}\right) = \alpha_{i^*}' + \beta_{j^*}' + (\alpha\beta)_{i^*j^*}'$$

Cell i^*j^* will have its very large weights likely unchanged. The rake-trim algorithm will give the same marginal factors, but a different interaction factor:

$$ln\left(u_{i^*j^*}\right) = \alpha_{i^*}' + \beta_{j^*}' + (\alpha\beta)_{i^*j^*}''$$

This new interaction factor $(\alpha\beta)_{i^*j^*}''$ will be smaller than $(\alpha\beta)_{i^*j^*}'$: some of the weight in cell i^*j^* will be re-distributed to other cells, while the marginal sums will be unchanged. Whether this will generate potential bias depends on the nature of the survey characteristic y being estimated using the weights. Suppose for example $y^{(1)}$ with regard to the two dimensions satisfies a main-effects model and $y^{(2)}$ on the other hand has significant interactions:

$$\bar{y}_{ij}^{(1)} - \bar{y}_{i.}^{(1)} + \bar{y}_{.j}^{(1)} = 0 \qquad \bar{y}_{ij}^{(2)} - \bar{y}_{i.}^{(2)} + \bar{y}_{.j}^{(2)} \neq 0$$

The estimate for $y^{(1)}$ will be improved by the rake-trim algorithm, as the extreme weights will be trimmed back (improving precision), and there is no bias as the y-characteristic satisfies a main-effects model. For $y^{(2)}$ however the improvement in precision from the trimmed weights will be offset by the bias induced by changing the weights for cell i^*j^*, as this cell has a mean value differing from the simple sum of the marginal means.

SCENARIOS FOR THE RAKING/POST-TRIM CYCLES

Suppose now that either there are no extreme weights going into the algorithm, or pre-trimming was used and the initial extreme weights are successfully taken care of. Raking itself can in many situations ***lead*** to extreme weights, as the raking algorithm may need to re-distribute weights between cells in sometimes a fairly extreme way in order to match the raking cell totals. The raking algorithm matches the control cells, and also perfectly preserves the weight distribution within every cross-cell defined by intersections of the marginal cells (the cells with control totals). The ratios of weights between items within cross-cells will be unchanged by the algorithm: if two units in a common cross-

cell have a weight ratio of 2 going into the raking step, they will continue to have a weight ratio of 2 after the raking step is completed)[2]. Also, cross-cells with zero weight before raking (no sample units) will still have zero weight after raking. But the distributions across cross-cells will certainly be altered, and extreme weights may be generated in doing this. For example, a particular cross-cell may have its weight sum increased in the process of matching to the marginal control cell totals in each raking step. The largest weights in that cross-cell may end up being too large then.

The post-trimming after each raking step can trim back weights that become too large or too small after the raking process is completed. Following this post-trimming, a new raking step should always be executed to re-calibrate the weights to the marginal control cell totals. The final set of weights will meet both the raking constraints and the trimming constraints if the algorithm converges. The weights will tend to be re-distributed both across cross-cells, and within cross-cells.

If the trimming cells are identical to one dimension of the raking margins, then the final weight coming out of the rake-trim algorithm will be a function of the cross-cell $i_1, \dots i_K$ and monotone function of the weight ratio within the cross-cell. For example, suppose the initial weights within the cross-cell $i_1, \dots i_K$ are given by $IW_{i_1,\dots i_K}(1), \dots\dots, IW_{i_1,\dots i_K}(n)$. Each raking step in the overall rake-trim iteration has the effect of shifting these weights by a constant multiplicative factor (which differs across cross-cells), while preserving the ratios between the weights completely. This follows from the nature of the iterations of the raking algorithm, which multiply all of the weights within the cross-cell by a constant ratio determined for each margin one by one. The overall multiplicative factor is the product of these constant ratios determined by each marginal adjustment.

Assuming again that each trimming cell is also a raking cell, the trimming step following the completion of the raking step has the effect of changing the distribution of the weights within the cross-cell, while retaining the ordering. Note that this property depends on the cross-cell nesting within a trimming cell. The final weights $FW_{i_1,\dots i_K}(1), \dots\dots, FW_{i_1,\dots i_K}(n)$ at the end of the rake-trim algorithm (if it converges) will be equal to the initial weights times a 'universal' multiplicative factor determined by the raking steps, times another 'shrinking' factor which will possibly bring in the extremes (increasing the minimum and the smaller weights, decreasing the maximum and the larger weights) while certainly preserving the ordering of the weights (i.e., if $IW_{i_1,\dots i_K}(j_1) \leq IW_{i_1,\dots i_K}(j_2)$, then $FW_{i_1,\dots i_K}(j_1) \leq FW_{i_1,\dots i_K}(j_2)$).

ASPECTS OF THE MACRO TO EVALUATE AND CONTROL THE PROCESS

The rake-trim algorithm doesn't always converge readily in practice, and thus a machinery is in place to facilitate alterations to allow for convergence. In our applications of this algorithm, we have found that a lack of convergence can be caused by a small number of marginal cells. After a sufficiently large number of iterations, the algorithm can 'settle down' to a steady state where most marginal cells show no change between pre-raking weight totals and post-raking weight totals, but a few marginal cells do show a constant perturbation: the pre-raking totals are off again after the trimming step, and this re-adjustment is similar to the last one. If the magnitude of these adjustments is getting smaller and smaller, then it may just be a question of more cycles necessary (increase the total number of cycles allowed). If on the other hand, no convergence of this kind seems to be obvious, then a fatal lack of convergence may be occurring. This can be apparent by studying the output of the raking steps in each cycle.

In the latter case, a good solution can be to try to collapse the marginal cells in which the problem is occurring. This may require collapsing marginal cells in several dimensions. In many practical situations, this may be quite sufficient to allow the process to converge. The macro allows for easy collapsing of cells by defining two fields on the main data file with the indicators for the marginal cells--an 'initial cell' field and a 'final cell' field—and two fields on the control totals file with two sets of control totals—an 'initial control total file' field and a 'final control total file' field. An EXCEL file drives the program by containing these fields. Table 1 below has an example of what this EXCEL file looks like for a simple example with two dimensions: age and grade.

In this example, the program is run for the initial cells, and we find a lack of convergence which concentrates among the 5 and 6 year old cells, and the Grade 1 and Grade 2 cells (these cells show considerable perturbations in all cycles). The decision is to collapse 5 and 6 year olds, and collapse Grade 1 and Grade 2. The user implements this decision by altering the 'Final Cell' column in the driver EXCEL file. The only necessity is to make the name the same in the cells which one desires to collapse, and the program will take it from there. The program automatically puts a new indicator field (AGE_F) on the main data file which has the indicators for the new cell definitions, and automatically generates a new set of control totals and puts them in the control total file with the specified name (AGE_CT_F for dimension 1, GRAD_CT_F for dimension 2). The new control totals are summations of the old control

[2] Within each cross-cell, each iteration of the raking steps multiplies the weights in the cross-cell by a factor equal to the marginal control total for that dimension and marginal cell divided by the current weight sum for the same marginal cell. This factor is constant within the cross-cell (which nests within every marginal cell containing it) so every weight in a given cross-cell will always be multiplied by a common factor throughout the raking step.

totals over the collapsed marginal cells. In practice, it is important to have this process of updating indicator fields and control total fields automatically, as doing this manually is very tedious and can be error-prone. With this EXCEL driver file, the user is free to experiment with a number of final cell definitions without a lot of tedious effort. He/she only needs to make changes in that one column in the driver EXCEL file.

Dimension	**Initial Cell**	**Final Cell**	**Initial ITERVAR**	**Final ITERVAR**	**Initial EXT**	**Final EXT**	**Initial EXTVAR**	**Final EXTVAR**
1	5_yrs_old	**5&6_yrs_old**	AGE_I	AGE_F	AGE_CT_I	AGE_CT_F	AGE_CT_I	AGE_CT_F
1	6_yrs_old	**5&6_yrs_old**	AGE_I	AGE_F	AGE_CT_I	AGE_CT_F	AGE_CT_I	AGE_CT_F
1	7_yrs_old	7_yrs_old	AGE_I	AGE_F	AGE_CT_I	AGE_CT_F	AGE_CT_I	AGE_CT_F
1	8_yrs_old	8_yrs_old	AGE_I	AGE_F	AGE_CT_I	AGE_CT_F	AGE_CT_I	AGE_CT_F
1	9_yrs_old	9_yrs_old	AGE_I	AGE_F	AGE_CT_I	AGE_CT_F	AGE_CT_I	AGE_CT_F
2	Grade_1	**Grade_1&2**	GRD_I	GRD_F	GRD_CT_I	GRD_CT_F	GRD_CT_I	GRD_CT_F
2	Grade_2	**Grade_1&2**	GRD_I	GRD_F	GRD_CT_I	GRD_CT_F	GRD_CT_I	GRD_CT_F
2	Grade_3	Grade_3	GRD_I	GRD_F	GRD_CT_I	GRD_CT_F	GRD_CT_I	GRD_CT_F
2	Grade_4	Grade_4	GRD_I	GRD_F	GRD_CT_I	GRD_CT_F	GRD_CT_I	GRD_CT_F

Table 1. Excel File for Controlling the Algorithm

There are a number of options for controlling the amount of printouts and EXCEL files that describe the outcomes of the algorithm. If the algorithm is converging successfully, a limited amount of output is necessary. Printouts of the distributions before and after a pre-trimming step, the results of the first raking step and first post-trimming step, and results after the final raking step may be all that is necessary when convergence results. When convergence fails, studying the intermediate cycles may be necessary as well. Printouts and Excel files can be generated for these intermediate cycles by altering the relevant parameters.

DISCUSSION AND SUMMARY

The RAKE-TRIM Macro is a SAS Macro which puts together simple building blocks into an iterative procedure which accomplishes complex goals of survey weight adjustments. The first set of goals (accomplished by raking) is to make certain summations of the weights equal to known control totals for a number of dimensions. This should both improve precision and reduce noncoverage errors. Trimming then will constrain the weight distribution, which should improve precision. (See for example Kalton and Flores-Cervantes (2003) for a discussion of the purposes of raking and trimming). Putting the two processes together also has the tendency to correct the defects of each method alone: trimming leaves weight summations off of their totals, which raking corrects, and raking can create extreme weights, which trimming corrects.

The building blocks are the trimming steps, which are fairly straightforward adjustments to the weight sets, and the raking steps, which in turn are an iteration of simple multiplicative adjustments using ratios of weight sums. The macro system in SAS makes putting these building blocks together a straightforward process. A set of parameters specified by the user in a ‘parameter sheet’ controls the algorithm. In addition, the easy interface with EXCEL makes it straightforward to control the raking cell structure easily. This facilitates achieving convergence of the process.

The new RAKE-TRIM Macro is a powerful tool for improving survey weights in a wide range of practical survey applications.

REFERENCES

- Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice.* Cambridge, MA: MIT Press.
- Brick, J. M., Montaquila, J., and Roth, S. (2003). Identifying problems with raking estimators. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 710-717.
- Chowdhury, S., Khare, M., and Wolter, K. (2007). Weight trimming in the National Immunization Survey. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 2651-2658.
- Deming, W. E., and Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics*, 11(4), 427-444.
- Deville, J.-C., and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.
- Deville, J.-C., Särndal, C.-E., and Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, *88*, 1013-1020.
- Kalton, G., and Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics,* 19(*2*), 81-97.
- Potter, F. (1990). A study of procedures to identify and trim extreme sampling weights. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 225-230.
- Särndal, C.-E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, 33*(*2), 99-119.
- Stephan, F. F. (1942). Iterative method of adjusting frequency tables when expected margins are known. *Annals of Mathematical Statistics,* 13(2), 166-178.

DISCLAIMER

The contents of this paper are the work of the author and do not necessarily represent the opinions, recommendations, or practices of Westat.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Lou Rizzo
Westat
Room RE 454
1600 Research Blvd.
Rockville, MD 20850
Phone: 301-294-4486
Fax: 301-294-2034
Email: RIZZO1@WESTAT.com
Web: Westat