INTERSURVEY CONSISTENCY IN SCHOOL SURVEYS

Albert Holt, Synectics for Management Decisions; Steven Kaufman, NCES; Fritz Scheuren, George Washington University; and Wray Smith, Synectics for Management Decisions

Fritz Scheuren, 1402 Ruffner Rd., Alex. Va. 22302 U.S.A. or Email at scheuren@aol.com

KEY WORDS: Generalized least squares; Raking ratio estimation; Weight attenuation and weight bounding

1. Introduction and Background

For the first time, in 1993-1994, the private school components of the Schools and Staffing Survey(SASS) and the Private School Survey(PSS) are being fielded in the same school year. Even though these two NCES surveys measure some of the same variables, the results between the surveys will not agree.

As the PSS is used for the SASS sampling frame, the PSS results are likely to be the more accurate. Under these circumstances, it makes sense to explore whether the introduction of PSS totals into SASS might lead to improvements. Traditional post-stratification methods exist to employ auxiliary information at the estimation stage in surveys. These, however, cannot be applied to SASS without modification.

In particular, PSS and SASS both measure numbers of schools, numbers of teachers, and numbers of students. Conventional simple or raking ratio adjustment procedures could be used to adjust sample weights so that the SASS estimates agreed with PSS for each of the three totals separately. Such approaches do not work, though, if the weights are to be adjusted so that all three SASS estimates agree simultaneously.

Alternatives are possible, though, that permit simultaneous estimation. For example, the Generalized Least Squares(GLS) techniques advocated by Deville and Särndal(1992) can be used, as in Imbens and Hellerstein (1993). While the asymptotic properties of GLS and GLS-like estimators are attractive, their finite sampling properties are not necessarily desirable. Possible operational concerns with GLS procedures include: (1)Some of the resulting weights may be less than one or even may be negative.(2)The procedure may be difficult to carry out, especially when excessively small weights arise.(3)The effect on estimates not directly adjusted is unknown and could be harmful.

Modified GLS.--To discuss the basic algorithm employed in Generalized Least Squares, it is necessary to define some notation; in particular --

w_i is the original SASS weight for the ith SASS observation, i=1,...,n.

t_i is the SASS total of teachers for ith SASS observation, i=1,...,n.

s_i is the SASS total of the students for the ith SASS observation, i=1,...,n.

N is the total estimated number of schools, as given by PSS.

T is the total estimated number of teachers, as given by PSS.

S is the estimated total number of students, as given by PSS.

In reweighting SASS three constraints are imposed on the new weights u_i,

$$\sum u_i = N$$

$$\sum u_i t_i = T$$

$$\sum u_i s_i = S$$

For our application the new weights u_i, subject to these constraints, are to be chosen to minimize a loss function which can be written as the sum of squares

$$\sum (u_i - w_i)^2$$

This is perhaps the simplest and most straightforward loss function that might be chosen. Motivating it here is outside our present scope, except to say that the sensitivity of the final results to the loss function chosen seems not to be too great(but this is an application issue and will be among the areas for future study, as set forth at the end of this paper). As the literature on GLS methods also makes clear(Deville, Särndal, and Sautory, 1993), the loss function chosen determines the form of the estimators eventually developed and those obtained using squared error loss are particularly convenient in a SASS setting.

Now the usual Lagrange multiplier formulation of this problem yields, after some algebra, that the new weights are of the form

$$u_i = w_i + \lambda_1 + \lambda_2 t_i + \lambda_3 s_i ,$$

where the λ's are obtained from the matrix expression

$$\underline{d} = M\lambda$$

with the vector $\underline{d}$ consisting of three elements, each a difference between the corresponding PSS and SASS totals for schools(first component), teachers(second component), and students(third component); in particular

$$N - \sum w_i$$

$$T - \sum w_i t_i$$

$$S - \sum w_i s_i$$

The matrix M is given by

$$\begin{matrix} n & \sum t_i & \sum s_i \\ \sum t_i & \sum t_i & \sum t_i s_i \\ \sum s_i & \sum t_i s_i & \sum s_i \end{matrix}$$

and $\underline{\lambda}$ is the vector of unknown GLS adjustment factors obtained from

$$\underline{\lambda} = M^{-1}\,\underline{d}.$$

The M matrix is based solely on the unweighted sample relationships among schools, teachers and students. This is not an essential feature of our approach; and, indeed a weighted version of the M matrix has been tried, as discussed later.

Illustrative Example.--To fix ideas, consider the following "toy" example that may help illustrate the method being employed. In particular, suppose a SASS subgroup has ten observations; written below as column vectors where the components

$$\begin{matrix} x \\ y \\ z \end{matrix}$$

correspond to SASS schools, teachers, and students respectively:

$$\begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 6 & 2 & 7 & 3 & 8 & 4 & 9 & 5 & 10 \end{matrix}$$

Aggregating the three SASS components yields

$$\begin{matrix} 10 \\ 55 \\ 55 \end{matrix}$$

Suppose further that the PSS totals for this subgroup are

$$\begin{matrix} 10 \\ 50 \\ 50 \end{matrix}$$

Notice, the SASS school total has already been set equal to that in the PSS. This has been done so that the example starts where a standard SASS estimation procedure might end.

For the "modified GLS" the elements of the matrix M and the vector $\underline{d}$ need to be obtained. It is immediate that $\underline{d}$ is

$$\begin{matrix} 0 \\ -5 \\ -5 \end{matrix}$$

For the matrix M, after some calculation, the values are

$$\begin{matrix} 10 & 55 & 55 \\ 55 & 385 & 355 \\ 55 & 355 & 385 \end{matrix}$$

For the inverse of M^{-1}, the values turn out to be

$$\begin{matrix} .5481 & -.0407 & -.0407 \\ -.0407 & .0204 & -.0130 \\ -.0407 & -.0130 & .0204 \end{matrix}$$

Thus, solving

$$\underline{\lambda} = M^{-1}\underline{d}$$

the vector is $\underline{\lambda}' = (.4074, -.0370, -.0370)$ and the modified GLS weights are of the form

$$u_i = w_i + .4074 - .0370t_i - .0370s_i.$$

Additional General Considerations.--So far the GLS algorithms have been discussed as if the issues are simply computational. In point of fact, the real challenges arising in any SASS implementation require statistical judgments. Among these are:

- Deciding on the level of SASS at which the constraints are to be imposed. For example, from a subject-matter perspective, it seems appropriate to do GLS estimation separately within the nine private school types. For some of the larger typologies, maybe even finer groupings might be attempted (say, school level or urbanicity). At what point will the potential benefits of a GLS adjustment outweigh the harm?

- Avoiding weights that are negative or too small (i.e., given that each SASS observation always represents at least itself, a natural requirement to impose is that $u_i \geq 1$ for all i). This concern is particularly troublesome because of the seemingly ad hoc flavor of what may be needed to get acceptable weights.

While the guidance of earlier GLS practice elsewhere is available(e.g., Bankier, 1992; Fuller et al, 1994)), neither of these challenges can be resolved for SASS, except "in the doing." Among the factors to consider are obvious ones such as --

- How difficult(expensive) is the method to implement, including to explain?

- How statistically sensitive are the constrained estimates to seemingly small but arbitrary decisions in the way the method is applied?

2. An Initial SASS Application

The basic approach taken in this Section is to analyze a small but real data set, so as to develop an understanding of the operating characteristics of the modified GLS approach being looked at here for potential use in the 1993-1994 NCES school surveys. To this end, consider, as a test, data on Catholic schools taken from the 1991-1992 PSS and the 1990-1991 SASS. These schools for SASS and PSS are divided into three subgroups: parochial, diocesan, and private. The weighted data on the last of these groups, Private Catholic Schools, are displayed below.

Item	PSS	SASS
Schools	901	894
Teachers	22340	22340
Students	354040	365367

The modified GLS application might be started by first scaling up the school total from SASS to that for PSS or simply leaving the total as is (the course taken here). In any event, after suitable calculations, familiar from Section 1, the GLS weights are obtained from the expression

$$u_i = w_i + .0415 + .0767t_i - .0046s_i .$$

One of the λ is negative; hence the u_i could be too small or even negative for a particular combination of original weight, teacher and student total. However, this did not occur.

The Private Catholic typology has the smallest sample size(at 112) and was chosen for that reason. Now three constraints are being imposed and sample size "rules of thumb" suggest that the average sample size per constraint be on the order of 25 or more. Here the average is 112/3 = 37, so reasonably good results might be expected at least on this score, provided SASS and PSS are consistent(i.e., that SASS can be treated as a representative sample of the larger PSS). Since the surveys are for different years this last condition is not guaranteed(see Section 3). Figures 1 and 2 below suggest, though, that SASS and PSS are roughly consistent, at least in this case. The SASS scatterplot lies well within that for PSS and is oriented along the same axis. Indeed, the average student/teacher ratios from the two surveys(both at about 16-to-1) are almost identical

3. A Second SASS Application

In this Section, a second GLS application is taken from the 1990-91 SASS and 1991-92 PSS. Here Nonsectarian Special Emphasis Schools are examined. That group was chosen because the weighted SASS and PSS counts are quite far apart(see below). If a problem with the GLS approach were to show up, it might well be in this group.

Item	PSS	SASS
Schools	1810	1700
Teachers	13724	18717
Students	202178	212433

First GLS Attempt.--The Nonsectarian Special Emphasis Typology has a somewhat larger sample size(at 205) than for Private Catholic Schools. Hence, standard concerns about overconstraining small numbers of cases do not bind here; indeed, it would even be possible to attempt to introduce still more PSS data into the SASS estimation --a point we will come back to later.

The modified GLS was solvable, leading to weights of the form

$$u_i = w_i - .0254 + .0101t_i - .0008s_i.$$

If sample size were our only consideration, the GLS weights should work well; however, they do not. As a matter of fact, nearly one third of these weights were less than one and many (22 in all) were negative. The SASS data are just not consistent with those from PSS. For example, the student teacher ratio in PSS is about 15 to 1; for SASS, on the other hand, it is closer to 11 to 1.

In the PSS and particularly in SASS, outliers exist which are well outside the point clouds of either source(see figures 3 and 4). One of these, circled in the SASS data is quite damaging since it has a weight of about 14 and a teacher count of 208 combined with a student count of 78--probably a data error of some sort.

Subsequent Attempts.--Removing the outlier yields the totals below.

Item	PSS	SASS
Schools	1809	1686
Teachers	13516	15836
Students	202100	211353

It would be great if we could now say that negative GLS weights or weights less than one had, with this single change, been eliminated. This did not turn out to be true; nonetheless, the results were encouraging. The number of "small" or negative weights was cut way down(from over eighty to under two dozen -- still quite sizable, however).

An examination of the SASS cases that had GLS weights that were too small revealed two patterns that might be mentioned: (1)Most of the cases were ones where the original SASS weight was close to one to begin with. (2)Some of the cases with negative weights had student/teacher ratios, that put them near the edge of the SASS and PSS point clouds -- making them possible candidates for outlier treatment too.

A series of alternatives were tried, including the use of different GLS loss functions(See Scheuren, 1994). Eventually, we settled on an alternative that fit a GLS estimator to the smaller two-thirds of the schools. The larger schools were simply too inconsistent to be fit with a GLS estimator; instead, an imputation approach was considered that might have future promise in the sample regions where the 1993-1994 SASS cases have weights of nearly one to begin with. More is said about this in the concluding section.

4. Future Plans

At this still early stage it is hard to do more than just conjecture about next steps in terms of the 1993-1994 SASS. Even so, there are some "lessons learned" and a few observations that may be of general interest. This short section makes a beginning summary of these.

First, our test plans call for more of the nine SASS typologies to be GLS-adjusted. It is plausible to speculate that still other methods may occur to us as we tackle these remaining typologies. Preliminary work, though, on some of these other typologies suggests that it is unlikely, for the 1993-1994 SASS, that we will uncover better approaches than those discussed. On the other hand, our sense of how and when to apply these techniques may grow considerably.

Second, we need to display evidence, convincing in the test SASS applications, that a GLS adjustment of the type contemplated will lead to an improvement in the estimates; or, at least, to no(or minimal) harm. On this latter point figures 5 and 6 are encouraging(because these figures show that the GLS weights are only minimally altered from their original values).

Third, methods for variance estimation need exploration . While the general GLS approach is well covered in the literature, an efficient method has to be programmed and tested in the SASS environment. Particular concerns exist, too, about the impact on variance and variance estimation of the various ad hoc adaptations needed to keep the weights reasonable.

Fourth, a general strategy for applying GLS to SASS may emerge from our work; but it appears highly unlikely that GLS procedures for SASS will become automatic any time soon. There is simply not going to be enough of an experience base to make this safe.

Fifth, some improvements in SASS and PSS processing may be a consequence of the study of GLS applications. One of those that has arisen so far is the clear possibility that edit checking could be enhanced if GLS estimation is attempted. A subtler concern is the treatment in SASS of the very largest schools, when these become nonrespondents. Here perhaps an imputation rather than a weighting approach may be preferred -- using , say, the PSS data as a starting point. Among schools above a given size this could have more benefit in reducing SASS mean square error than GLS.

Obviously, still other concerns need to be considered, even if the present modified GLS method were judged desirable; and could be made routine. Among these, of course, are the cost in time and money of its application. So stay tuned.

Selected Bibliography

Bankier, M. (1992), "Two-step Generalized Least Squares Estimation in the 1991 Canadian Census," in *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Burton, R. (1989), Unpublished Memorandum, National Center for Education Statistics.

Casady, R. and Valliant, R.(1993), "Conditional Properties of Post-Stratified Estimators Under Normal Theory," *Survey Methodology*, 19, 183-192.

Deming, W.E. and Stephan, F.F. (1940), "On a Least Squares Adjustment of a Sampled Frequency When the Expected Marginal Tables are Known," *Annals of Mathematical Statistics*, 11, 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.

Fuller, W., Loughin, M., and Baker, H. (1994), "Regression Weighting in the Presence of Nonresponse with Application to the 1987-1988 Nationwide Food Comsumption Survey," *Survey Methodology*, 20, 75-86.

Hansen, M., Hurwitz, W.,and Madow, W.(1953), *Sample Survey Methods and Theory*, New York: Wiley.

Holt, D. and Smith, T.M.F. (1979), "Post-stratification," *Journal of the Royal Statistical Society, Series A*, 142, 33-46.

Imbens, G.W. and Hellerstein, J.K. (1993), "Raking and Regression," *Discussion Paper Number 1658*, Cambridge, MA, Harvard Institute of Economic Research, Harvard University.

Ireland, C.T. and Kullback, S. (1968), "Contingency Tables with Known Marginals," *Biometrika*, 55, 179-188.

Kaufman, S. and Huang, H. (1993), "1990-91 Schools and Staffing Survey: Sample Design and Estimation," *Technical Report NCES 91-127*, Washington, DC: U.S. Department of Education.

Little, R.J.A. (1991), "Post-Stratification: A Modeler's Perspective," in *Proceedings of the Section on Survey Research Methods*, American Statistical Association..

Little, R.J.A. and Wu, M. (1991), "Models for Contingency Tables with Known Marginals When Target and Sample Populations Differ," *Journal of the American Statistical Association.*

Oh, H.L. and Scheuren, F. (1978a), "Multivariate Raking Ratio Estimation in the 1973 Exact Match Study," in *Proceedings of the Section on Survey Research Methods*, American Statistical Association..

Oh, H.L. and Scheuren, F. (1978b), "Some Unresolved Application Issues in Raking Ratio Estimation," in *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Olkin,I.(1958), "Multivariate Ratio Estimation for Finite Populations," *Biometrika*, 45,154-165.

Scheuren, F. (1994), "Median Raking and Other Garden Variety Sampling Tools," *Unpublished Draft NCES Working Paper.*

Särndal, C.E., Swensson, B. and Wretman,J. (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.

Figure 1: PSS Teacher Versus Student Total for Private Catholic Schools

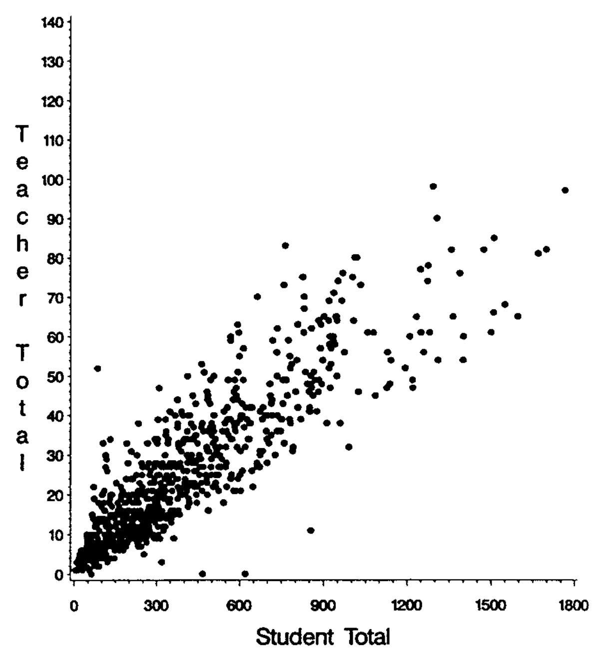

Figure 2: SASS Teacher Versus Student Total for Private Catholic Schools

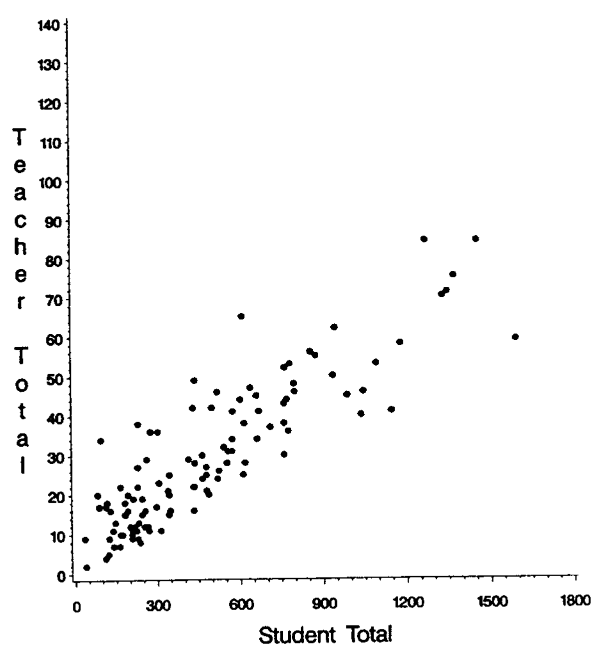

Figure 3: PSS Teacher Versus Student Total for Nonsectarian Special Emphasis Schools

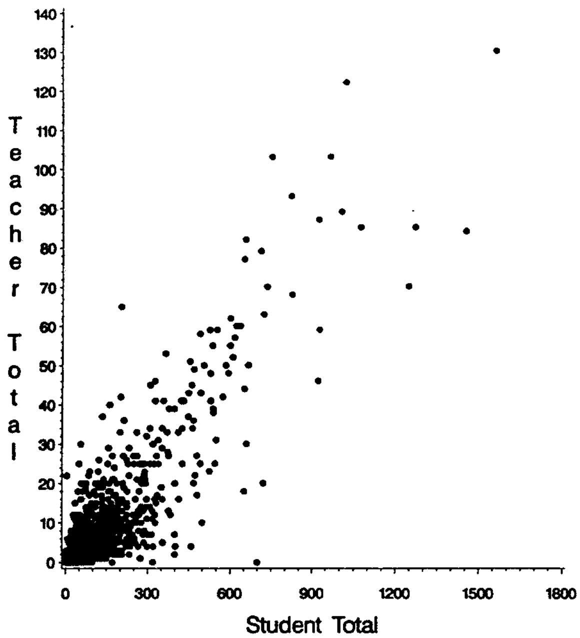

Figure 4: SASS Teacher Versus Student Total for Nonsectarian Special Emphasis Schools

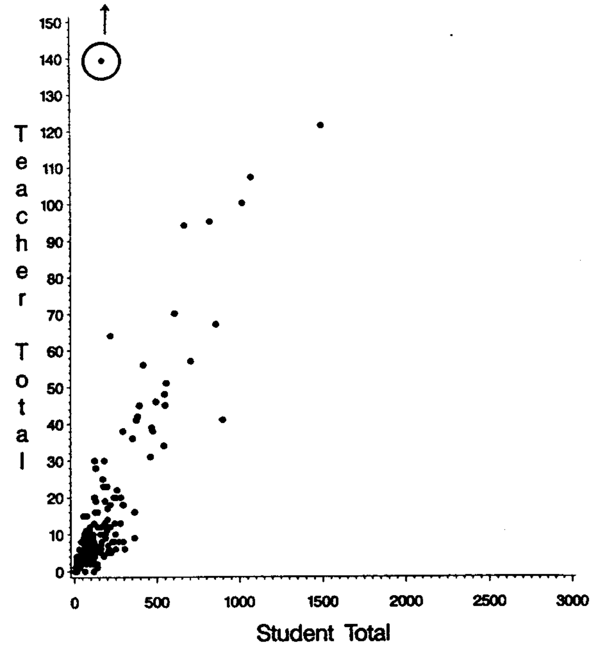

Figure 5: SASS Original Weight vs. GLS Weight for Private Catholic Schools

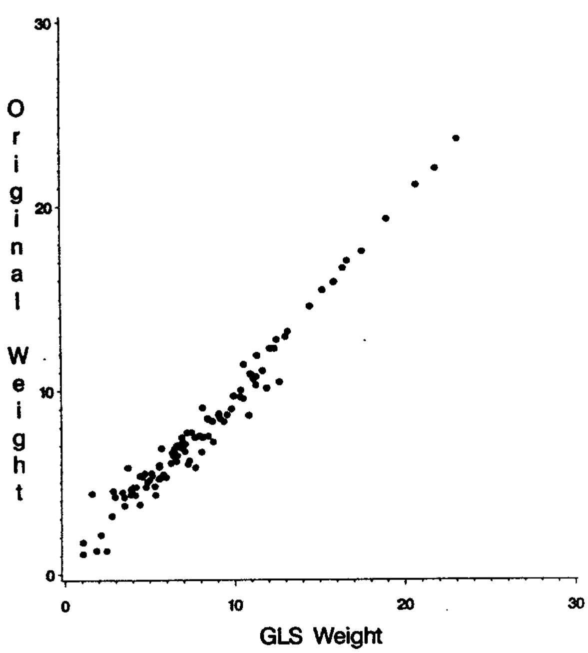

Figure 6: SASS Original Weight vs. GLS Weight for Nonsectarian Special Emphasis Schools (smallest two-thirds)