Graham Kalton, University of Michigan

Although a considerable amount of research has been carried out in recent years on procedures for compensating for missing survey data, there still remains a good deal to be done in developing improved procedures and in clarifying our understanding of the relative advantages and disadvantages of alternative procedures. Rod Little's paper is a useful contribution in this area.

I will take up three of the issues that Rod addresses, but before doing so I would like to comment briefly on his choice of reference distribution. He obtains results conditional on $\mathbf{n} = (n_1, n_2, \ldots, n_C)$ and $\mathbf{n}_R = (n_{1R}, n_{2R}, \ldots, n_{CR})$ as well as $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ and $\mathbf{r} = (r_1, r_2, \ldots, r_N)$. While these results provide insights into biases and variances of estimators conditional on $\mathbf{n}$ and $\mathbf{n}_R$, it seems to me that the unconditional results are also needed as summaries; indeed Rod does end up giving some average results, both theoretically and in the simulations, across values of $\mathbf{n}$ and $\mathbf{n}_R$. The results I give below are not conditional on $\mathbf{n}$ and $\mathbf{n}_R$. Some of Rod's conditional biases translate into components of variance in the unconditional approach.

I will now turn to the following issues from Rod's paper: the formation of adjustment cells; the relationship between weighting adjustments and imputation; and the effect of compensation procedures on subclasses of the sample that cut across the adjustment cells.

THE FORMATION OF ADJUSTMENT CELLS

Rod discusses two alternative strategies for constructing adjustment cells: they may be formed by stratifying on predictors of the y variable or by stratifying on the response propensity. The justification for these two alternatives can be seen by considering the unconditional biases of the respondent mean ($\bar{y}_R$), the adjusted mean ($\bar{y}_A$) and the poststratified mean ($\bar{y}_S$) (see Thomsen, 1973; Kalton, 1983):

$$\text{Bias}(\bar{y}_R) = [\Sigma P_c(\bar{Y}_{cR} - \bar{Y}_R)(B_c - \bar{B})/\bar{B}] + \Sigma P_c(1-B_c)(\bar{Y}_{cR} - \bar{Y}_{cM})$$

$$= A + B$$

$$\text{Bias}(\bar{y}_A) = \text{Bias}(\bar{y}_S) = B.$$

Here the notation is as in Rod's Table 3, with the addition of the symbol $\bar{Y}_{cM}$ for the mean of the nonrespondents in adjustment cell c, and $B = B_A = B_S$ is what Rod terms the large sample bias (LSB). The equations show that the effect of the weighting or poststratification adjustments is to change the bias of the sample mean from (A + B) to B.

If A and B are of different signs, either form of adjustment may increase the absolute bias. If A and B have the same sign, the adjustment reduces the absolute bias if $A \neq 0$. The term A is a covariance-type term. Two conditions are required for it to be non-zero: (1) the response rates B_c must differ between cells and (2) the respondent means must differ between cells. The formation of cells by stratifying on predictors of the y variable or on the response propensity ensures that one of these conditions holds.

Note that if cells are formed to have varying values of $\bar{Y}_{cR}$, but the response rates do not vary, no bias reduction is obtained by weighting or poststratification: the poststratification adjustment will, however, lead to a slight increase in the precision of the estimator. If the response rates vary between strata, but the means do not, there will again be no bias reduction; however, in this case the estimators $\bar{y}_A$ and $\bar{y}_S$ will be less precise than $\bar{y}_R$ (assuming constant variance within cells).

One consideration in forming adjustment cells is that the variable y is conditionally independent of response within cells (so that $\bar{Y}_{cR} = \bar{Y}_{cM}$). But in practice this is hard to assess. Therefore attention is mainly given to forming cells with different response rates and/or different means. Whether the emphasis is placed on response propensity or predictor variables for forming cells depends on the situation. In the case of unit nonresponse, there are usually only limited data available for forming cells, these data are generally only weakly related to the survey variables, and the adjustment is wanted for all the survey variables simultaneously; consequently the emphasis is mainly on forming cells with different response rates. In the case of item nonresponse, however, there are often several closely related variables available from which the cells can be formed; consequently the emphasis is on forming cells in terms of the predictor variables. In the extreme case, the cells may be formed so that the respondents' y-values are the same within each cell: a nonrespondent from a given cell is then assigned the respondent y-value from that cell. This might appear to be an error-free imputation, but it must be remembered that it depends on the conditional independence assumption: respondents and nonrespondents need not have the same values within cells.

In passing, it is worth observing that the equivalence of effect on bias of weighting adjustments and poststratification noted above should not be misinterpreted. The two techniques have different data requirements which affect the choice of adjustment cells: for weighting adjustments information is needed to assign both respondents and nonrespondents to the cells, whereas for poststratification only the respondents need to be assigned to cells, but

external information on the population distribution across the cells is required. An important difference between the two forms of adjustment is that poststratification handles noncoverage as well as nonresponse, whereas weighting adjustments handle only nonresponse.

WEIGHTING AND IMPUTATION

Rod draws attention to the close relationship between weighting and imputation. He notes, in particular, that the sample mean based on a cell weighting adjustment is equivalent to a sample mean based on the respondents' values and the missing observations assigned the values of their respondent cell means. This equivalence is often cited, but I think there is a danger that it can be misinterpreted. The equivalence applies only for the sample mean. The two procedures produce different estimators for other parameters. In particular, researchers are usually also interested in the distribution of y. The weighting adjustment retains the respondents' y-distribution within cells, and providing the conditional independence of y and r holds, the overall weighted sample distribution estimates the population distribution and the weighted sample variance estimates the population variance. However, imputing the cell mean for missing observations in that cell distorts the y-distribution, creating spikes at the cell means and attenuating the variance. Cell mean imputation is consequently normally avoided in practice, and instead some form of stochastic imputation that retains the variance of the respondents' values within cells - a hot-deck imputation - is generally preferred.

One type of hot deck imputation is to take some form of equal probability sample from the respondents within a cell and assign their values to the nonrespondents. When this procedure is used, the sample mean has the same bias for estimating the population mean as the one using imputed cell mean values. The sampling of respondents leads to an increase in the variance of the mean, however. It would not be difficult to extend Rod's research to include an evaluation of some form of hot-deck imputation, and in view of the practical importance of this form of imputation it would seem useful to do so.

EFFECTS OF COMPENSATION ON SUBCLASSES

I very much welcome the section of Rod's paper that deals with the effects of compensation on subclass estimates when the subclass cuts across the adjustment cells. As he observes, most of the literature on nonresponse adjustments has focussed on estimates of population means and totals. There has been too little attention given to the effects of adjustments on other statistics; in particular, the dangers of imputation for correlation and regression analyses do not seem to be fully appreciated (see Santos, 1981; Kalton and Kaspryzk, 1982).

As Rod shows, weighting adjustments by cell will not make the correct compensation for a subclass that cuts across the cells unless the response rate for the subclass is the same as - or at least proportional to - that for the cell as a whole. Rod gives an example to illustrate this point. Another simple example is as follows. Suppose there are two adjustment cells, and subclasses of men and women as in Table 1.

Table 1

Sex	Cell 1		Cell 2	
	Respondents	Total sample	Respondents	Total sample
Men	10	30	20	20
Women	10	10	20	30
Total	20	40	40	50

Weighting by b_c^{-1} gives a weight of 2 to respondents in cell 1 and a weight of 1.25 to respondents in cell 2. Thus the weighted distribution is given in Table 2.

Table 2

Sex	Weighted totals	
	Cell 1	Cell 2
Men	20	25
Women	20	25
Total	40	50

While the overall weighted distribution over the cells corresponds to the desired total sample distribution (40:50), that does not hold for the subclass taken separately: men are underrepresented and women overrepresented in cell 1, whereas the reverse holds in cell 2. This situation arises because of the difference in the response rates for men and women. In general, this problem will be avoided only if adjustment cells are formed so that response rates are the same in each cell for all subclasses used in the analysis.

Rod also points out that although weighting adjustments and imputation yield the same estimates $\bar{y}_A$, they produce different results in subclasses. With imputation of the cell mean, the male nonrespondents in cell 1 would be assigned the mean of the 20 male and female respondents in that cell, and the female nonrespondents in cell 2 would be assigned the mean of the 40 male and female respondents in that cell. If the two sexes have different means, this imputation will cause a distortion in the subclass means for men and women. In general the difference between subclass means is made smaller by the imputation. As a simple illustration, suppose that in each cell among male respondents the proportion answering "Yes" is 0.8 and among female respondents it is 0.2. Thus half of the respondents in each cell answered "Yes", so that the 20 male nonrespondents in cell 1 and the 20 female nonrespondents in cell 2 would all be assigned values of 0.5 (or with a controlled hot-deck, half would be assigned "Yes" and half "No" answers). The overall proportion of men answering "Yes" is

tnus estimated as (8+10+16)/50 = 0.68, whereas the overall proportion of women answering "Yes" is estimated as (2+4+5)/40 = 0.275. This attenuation of the difference between the subclass proportions has extremely important consequences for survey analysis, which typically is much concerned with subclass estimates and their comparisons. It is a special case of the general problem of attenuation of covariances caused by imputation. A good deal more research is needed in this area.

Rod raises a number of other interesting points on which I would like to comment. However, I feel that I should stop at this point to give others the opportunity to take up the discussion. I would like to thank Rod for a stimulating paper.

REFERENCES

Kalton, G. (1983). Compensating for Missing Survey Data. Institute for Social Research, University of Michigan.

Kalton, G. and Kasprzyk, D. (1982) Imputing for missing survey responses. Proceedings of the Section on Survey Research Methods, American Statistical Association, 1982, 22-31.

Santos, R. (1981). Effects of imputation on regression coefficients. Proceedings of the Section on Survey Research Methods, American Statistical Association, 1981, 140-145.

Thomsen, I. (1973). A note on the efficiency of weighting subclass means to reduce the effects of nonresponse when analyzing survey data. Statistisk Tidskrift, 4, 278-283.