# DISCUSSION

Keith Rust, Westat, Inc.
1650 Research Blvd., Rockville, MD 20850

As a discussant for a contributed paper session, I am fortunate that the papers presented in this session have very much a common theme, moreso than is often the case for invited paper sessions on a particular topic. The general theme most closely tying these presentations together is the issue of precision of sampling error estimation for complex sample surveys. Each of these papers is thought-provoking, and each sheds a somewhat different perspective on this common theme. In my discussion, I will concentrate on the questions which occurred to me as I reviewed the papers.

### Kott Paper

Philip Kott uses a population modelling approach to evaluate the properties of linearization variance estimators for a ratio estimator. This approach of using a population model to evaluate a variance estimator (but not as the basis for the derivation of the estimator) has been used successfully in the past particularly as a means of comparing alternative approaches (for example, Brewer, 1963).

The approach Kott takes is to use model expectation, rather than the randomization expectation, to evaluate the properties of the variance estimator with respect to a given sample. While this gives more insight than merely considering the randomization expectation, the difficulty with the approach is that the results depend upon the choice of model. For the finite population surveyed in a given case, it can be difficult to evaluate the appropriateness of the model used.

The model used is

$$y_{kj} = \mu\, m_{kj} + e_{kj}$$

where

$$E(e_{kj}^2) = c m_{kj}.$$

It would be interesting to consider the effect of using a more general model. For instance, one could consider

$$y_{kj} = \alpha_k + \mu_k\, m_{kj} + e_{kj}$$

where

$$E(e_{kj}^2) = c_k\, m_{kj}^{\gamma k}$$

but this is probably too unwieldily and overparameterized. A useful compromise might be to use

$$y_{kj} = \mu\, m_{kj} + e_{kj}$$

with

$$E(e_{kj}^2) = c m_{kj}^{\gamma}$$

Putting $\gamma = 1$ gives the model used by Kott. Kott gives

$$RMV(V_D) = 2 \frac{\Sigma m_k^2}{\left(\Sigma m_k^2\right)}$$

I speculate that for the more general model

$$RMV(V_D) = 2 \frac{\Sigma m_k^{2\gamma}}{\left(\Sigma m_k^{\gamma}\right)^2}$$

This expression could be used to estimate the appropriate number of degrees of freedom, as proposed by Kott, after establishing empirically an appropriate value for $\gamma$.

Kott proposed the use of both the relative model bias (RMB) and the relative model variance (RMV) to adjust the usual linearization variance estimator and associated confidence intervals.

He proposes using $V_D^* = V_D/(1 + RMB\ (V_D))$ to adjust the variance estimator to account for the model bias. Since the value of RMB is dependent upon the choice of the model, and the extent to which this adjustment improves the variance estimator (or otherwise) is dependent upon the appropriateness of the model, I would hesitate to use this adjustment in cases other than where there was good empirical or theoretical validation of the model fit.

The use of RMV is proposed as a means of deriving an approximation of the number of degrees of freedom of the variance estimator (d) through the expression

$$d = 2/RMV(V_D)$$

In view of the absence of the other good approximations to the value of d, I think that this approach will prove very useful. Also I am hopeful that the estimation of d will be somewhat robust to model misspecification. I would encourage Dr. Kott to pursue an investigation of this approach.

### Marker Paper

David Marker's presentation emphasizes the importance of considering the balance between cost and precision of variance estimation, in the formation of sample replicates for use in replicated variance estimation. Needed precision can be lost, perhaps for little real gain, by arbitrarily reducing the number of replicates utilized, and Mr. Marker has shown an artificial but effective example of this. Particular caution is needed when separate estimates are required for subgroups which are highly correlated with the survey strata (e.g., region estimates) as very imprecise variance estimates can easily result if less than full replication is used.

For some discussion of approaches which can be used to reducing the extent of replication without losing needed precision, see Lee (1972, 1973) for BRR, and Rust (1986, 1984). With reasonable knowledge of relative contributions of the survey strata to total sampling variance, adequate precision for creating confidence intervals can generally be obtained using only a few dozen replicates. Caution is required, however, and Marker's warning against ill-considered reduction in the numbers of replicates to be used should be taken seriously by all practitioners. I note that in Dr. Fay's paper this careful consideration has been given to the level of replication required.

One final note in relation to this issue is to emphasize the point that when replication with a paired design is carried out to the fullest extent for linear estimators, it is immaterial which PSU in each pair is denoted at '1' and which as '2'. This designation can be purposive without introducing bias. When reduced replication is used, this assignment must be random, or bias will result from the fact that cross-product terms between strata no longer have zero expectation.

### Kish Paper

The task of grappling with the impact of design considerations on survey inference has been made noticeably easier over the years as a result of Leslie Kish's efforts in popularizing the concept of the design effect. This present paper gives many points of helpful summarization as to the uses of design effects and also points out some of their limitations. I would suggest that survey practitioners at all levels of experience will benefit from a reading of Professor Kish's summarizations.

As Professor Kish points out, care is needed in the use of design effects when the survey data are not approximately self-weighting. Here, both the weighting and the design features of stratification, clustering, and so on, have an impact. Care is needed in defining and interpreting the "simple random sampling variance" -- the denominator of the design effect -- in this case. I believe that there is tendency to compute the ratio of the unbiased sampling variance estimate to the variance estimate which is obtained if unweighted data are used in a statistical software package. This is not the design effect, which compares sampling variance with that which **would have been obtained had a simple random sample been used**. Rather, this is an analysis effect, which compares the true sampling variance to that which **is obtained if the survey design and survey weights are ignored in the analyses.**

To illustrate , consider an estimate of a proportion, $\hat{p}$ of p, from a weighted sample. The design effect is estimated by

$$\text{deff}\,(\hat{p}) \quad = \quad \frac{\text{var}(\hat{p})}{\hat{p}\,(1-\hat{p})/n}$$

The "analysis" effect is

$$\text{aneff}\,(\hat{p}) \quad = \quad \frac{\text{var}(\hat{p})}{p^*\,(1-p^*)/n}$$

where p* is the unweighted sample proportion. With differential weighting $\hat{p}$ and p* can differ substantially, so that the design effect and analysis effect are not one and the same. Thus, I would add as a postscript to Professor Kish's paper a word of caution to distinguish between these two concepts, and to be sure to use the appropriate one in a given case. Those who practise survey design are interested in design effects, but those who analyze survey data will often be more interested in the impact of the sample design and weighting on the analysis of the data at hand, rather than in considering the relative precisions of alternative designs.

### Fay Paper

Dr. Fay's paper deals with issues related to practical problems of variance estimation in a particular application, the Survey of Income and Program Participation. These issues are the use the Durbin-Sanford method to select two PSUs per stratum and the implications for variance estimation, the random reduction of the sample following initial selection, modification of the Yates-Grundy variance estimator to reflect the multistage design the development of replicated variance estimation procedures to reflect these steps, and the choice of Fay's modified method of Balanced Repeated Replication.

The paper describes well the final procedures that were used, but I would be interested to learn more as to the considerations which led to these choice of methods. For example, how was the decision reached to use the Durbin-Sanford method to select PSUs, as opposed to using independent selections, or choosing one PSU per stratum?

The use of Durbin's method permits unbiased estimation of the first stage sampling component of variance. However, this is at the expense of more complicated variance estimation for the second and later stage variance components, and the selection of two PSUs per stratum in this way may perhaps decrease the efficiency of the sample design by not maximizing the use which could be made of information available at the design stage. A discussion of the various trade-offs would be enlightening.

The general approach given by Dr. Fay to approaching the problem of finding an appropriate replication approach for a given complex design appears to be a very useful one. The approach is to formulate the design appropriate explicit variance estimator for a linear estimator as

$$\text{Var}^*\,(\underset{\sim}{1}'\,\underset{\sim}{x}_w) \quad = \quad \underset{\sim}{x}_w'\,C(s)\,\underset{\sim}{x}_w$$

where C(s) is a matrix of quantities dependent upon the sample design but not the sample data. Then appropriate replicated variance estimators can be expressed in terms of eigenvectors and eigenvalues of C(s). This approach gives a method of obtaining an appropriate replication variance estimator for particularly complex designs as illustrated in the paper. I believe that this formulation should also prove useful in the future in comparing alternative "valid" replication approaches (e.g., BRR and jackknife) for a given design.

The modification which Dr. Fay proposes for BRR of weighting cases in a given replicate by factors of 0.5 and 1.5 in place of using 0 and 2 in the case of full BRR suggests that perhaps a "good" BRR procedure is one which is similar to the jackknife. However, the jackknife is known to the inconsistent for variance estimates for quantiles (Brillinger, 1964), whereas BRR is not inconsistent (Rao and Wu, 1987). Fay's approach seems likely to retain the consistency of standard BRR and, thus, be at an advantage over the jackknife. My colleague and Dr. Fay's, David Judkins, is empirically investigating this particular question at present.

Finally, I will point out for the benefit of those who would like to analyze data using Dr. Fay's modified BRR approach that Westat's SAS procedure for variance estimation, PROC WESVAR, has recently undergone a number of enhancements, one of which gives it the capability to implement Fay's modified BRR, given an appropriate set of replicate weights.

This is achieved by providing as input the factor $b_r$ for use in the variance estimator

$$\mathrm{Var}^*(\underset{\sim}{1}'\,\underset{\sim}{x}_w) = \sum_r b_r\,(\underset{\sim}{w}_r\,\underset{\sim}{x} - \underset{\sim}{W}_x)^2,$$

described in the paper.

## Conclusion

Finally, I wish to thank all of the speakers for their thought-provoking presentations and for their efforts in making this what I believe to have been a very useful, interesting and successful session.

## References


Brewer, K.R.W. (1963). Ratio estimation and finite populations: some results deducible from the assumption of an underlying stochastic process. *Australian Journal of Statistics,* 5,93-105.

Brillinger, D.R. (1964). The asymptotic behavior of Tukey's general methods of setting approximate confidence intervals (the jackknife) when applied to maximum likelihood estimates. *Review of the International Statistical Institute*, 32, 202-206.

Lee, K-H (1972). Partially balanced designs for half sample replication method of variance estimation. *Journal of the American Statistical Association,* 67, 324-334.

Lee, K-H. (1973). Using partially balanced designs for the half sample replication method of variance estimation. *Journal of the American Statistical Association*, 68, 612-614.

Rao, J.N.K., and Wu, C.F.J. (1987). Methods for standard errors and confidence intervals from sample survey data: some recent work. *Bulletin of the International Statistical Institute,* 3,5-21.

Rust, K.F. (1984). *Techniques for Estimating Variances for Sample Surveys,* PhD. thesis, University of Michigan

Rust, K. (1986). Efficient replicated variance estimation, *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 81-87.