THE EFFECT OF MULTIPLE WEIGHTING STEPS ON VARIANCE ESTIMATION

Richard Valliant
Westat, 1650 Research Boulevard, Rockville, Maryland 20850
Joint Program for Survey Methodology, University of Maryland

Key Words: Ineligibility; Jackknife; Leverage; Nonresponse; Replication variance estimate

1. Introduction

Multiple steps in weighting are common in survey estimation. Each step usually introduces a source of variability in an estimator that may be important to reflect when estimating variances. A typical sequence of weighting steps in a probability sample is this:

1. Compute base weights.
2. Adjust weights to account for units with unknown eligibility.
3. Adjust weights for nonresponse.
4. Use auxiliary data.

The variance of an estimator is affected by the population structure of the variables being estimated, the complexity of the design used to collect data, and the form of the estimator itself, including the weighting steps above. Intuition may lead us to believe that a variance estimator that somehow incorporates all of these complications is better than one that does not. However, literature that directly addresses this question is limited. The collection on survey nonresponse by Groves, Dillman, Eltinge, and Little (2002), for example, does not include any articles on the effect on variance estimates of multiple steps in weighting.

The two major competitors in finite population variance estimation are replication and linearization. For replication variance estimators there is evidence in particular cases that it is necessary to repeat each step of estimation separately for each replicate subsample in order to produce a consistent or approximately unbiased variance estimate. Empirical results, however, are not uniform. Lemeshow (1979) is an early paper illustrating by simulation that this is necessary for the BRR method.

Valliant (1993) showed theoretically and empirically that poststratification factors must be recomputed for every replicate in order for the BRR or jackknife estimators to be consistent in two-stage sampling. Yung and Rao (1996) obtained similar results for the jackknife in stratified, multistage sampling. Yung and Rao (2000) studied the poststratified estimator when weighting class nonresponse adjustments were made. They proved that the jackknife is consistent if the nonresponse adjustment factors and the poststratification factors are recomputed for each replicate subsample.

There are a number of articles that cover some, but not all, of the four steps when applying Taylor series variance estimators. Lundström and Särndal (1999) study the use of a linearization estimator for the general regression (GREG) estimator when there is nonresponse. Rao (1996) derived a modified linearization variance estimator that accounted for mean imputation.

Shortcut implementations of linearization estimators, that ignore some steps in weighting, are fairly common in practice for at least two reasons. First, linearizing complex estimators is difficult and commercial software packages limit how faithfully a user can reflect the complexities of a design and an estimator. Few studies report direct comparisons of shortcut linearization estimators and more elaborate replication estimators. This paper attempts fill that gap by empirically comparing some alternative variance estimators systematically for various combinations of eligibility rate, response rate, type of estimator, and sample size.

2. Notation and an Estimator of a Total

For the illustrations in this paper we consider only stratified and unstratified, single-stage sampling but include all four of the weighting steps listed in section 1. Suppose that the strata are numbered $h=1,\ldots,H$, the frame size in stratum h is N_h, the number of initial sample units is n_h, and the set of initial sample units is s_h. Denote the base weight for sample unit hi as w_{hi}. Define the following sets of sample cases:

s_{ER} = set of eligible sample respondents;
s_{ENR} = set of eligible sample nonrespondents;
s_{IN} = set of sample units known to be ineligible; and
s_{UNK} = set of sample units whose eligibility status is unknown.

The full sample s is the union of these four sets. The set of units whose eligibility status is known is $s_{KN}=s_{ER}\cup s_{ENR}\cup s_{IN}$.

Suppose that the sample is also divided into classes, $c=1,\ldots,C$, that are used for the unknown eligibility adjustment. Another set of classes, $d=1,\ldots,D$, is used for the nonresponse adjustment. Both of these sets of classes may cut across strata. In practice, the eligibility adjustment and nonresponse adjustment classes may often be the same. Let s_c denote the set of sample units in class c and $s_{c,KN}=s_c\cap s_{KN}$ the set with known eligibility in class

c. Then, the unknown eligibility adjustment for sample units in class c is

$$a_{1c} = \frac{\sum_{(hi)\in s_c} w_{hi}}{\sum_{(hi)\in s_{c,KN}} w_{hi}} \quad (1)$$

and the eligibility-adjusted weight for a unit with known eligibility in class c is

$$w_{1hi} = w_{hi} a_{1c} \qquad (hi) \in s_{c,KN} .$$

The summations over $(hi) \in A$ for some set A means to sum over all strata and the units within each stratum that are members of the set. After this step, the units with unknown eligibility are eliminated.

Next, denote the set of cases in class d as s_d, those that are known to be eligible in class d as $s_{d,E} = s_d \cap (s_{ER} \cup s_{ENR})$ and the set of eligible respondents in class d to be $s_{d,ER} = s_d \cap s_{ER}$. The nonresponse adjustment for units in class d is

$$a_{2d} = \frac{\sum_{(hi)\in s_{d,E}} w_{1hi}}{\sum_{(hi)\in s_{d,ER}} w_{1hi}} . \quad (2)$$

The nonresponse adjusted weight is then

$$w_{2hi} = \begin{cases} w_{1hi} a_{2d} & (hi) \in s_{d,E} \\ w_{1hi} & (hi) \in s_{IN} \end{cases} \quad (3)$$

i.e., the weights for eligible respondents are adjusted while the weights for known ineligibles remain the same as they were after the unknown eligibility adjustment. The nonrespondents, s_{ENR}, are eliminated. After this step, the units with nonzero weight, which are used in estimation, are s_{ER} and s_{IN}. The known ineligibles are retained on the grounds that their presence in the sample is a reflection of other nonsample ineligibles in the frame.

To illustrate the use of auxiliary data, we take the case of the poststratified estimator, denoted by $\hat{T}_{PS}$, in single-stage sampling. If we let $k = 1, \ldots, K$ index the poststrata and $s_{PS,k}$ be the set of population units in poststratum k, then define the g-weight for a unit as

$$g_{hi} = \begin{cases} N_k / \hat{N}_k & (hi) \in s_{PS,k} \cap (s_{ER} \cup s_{IN}) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where N_k is the population count in poststratum k (which may include some ineligibles) and $\hat{N}_k = \sum_{(hi)\in s_{PS,k} \cap (s_{ER} \cup s_{IN})} w_{2hi}$, i.e., the estimate of the poststratum count based on eligible responding sample units and the sample units that are known to be ineligible.

Computing g_{hi} using $s_{ER} \cup s_{IN}$ presumes that the population control counts, N_k, include some units that are actually ineligible but cannot be separated out. This can occur if the population counts are made from a frame that is somewhat out-of-date. If the population controls include only eligibles, then the g-weight would be computed based only on the eligible respondents, s_{ER}. After the g-weight adjustment, the weight for sample unit i is

$$w_{3hi} = w_{2hi} g_{hi} \qquad (hi) \in s_{ER} \cup s_{IN} .$$

The final weights after the three stages of adjustment would, thus, be defined by

$$w_{hi}^{*} = \begin{cases} a_{1c} a_{2d} g_{hi} w_{hi} & (hi) \in s_{c,KN} \cap s_{d,ER} \\ a_{1c} g_{hi} w_{hi} & (hi) \in s_{c,KN} \cap s_{IN} \\ 0 & \text{otherwise} \end{cases} . \quad (5)$$

For estimation the eligibles are treated as a domain, and the ineligibles are assigned data values of zero both for point estimation and variance estimation. After these sequential adjustments, even an estimated total of the form $\hat{T} = \sum_{(hi)\in s_{ER} \cup s_{IN}} w_{hi}^{*} Y_{hi}$ is nonlinear in the design-based sense because the weights involve various sample-dependent ratio adjustments.

3. Variance Estimators

We will study several variance estimators that, in varying degrees, account for the complexity of the design and the estimator of the total.

One is a variation of the jackknife that divides the units within a stratum into random groups and deletes one group at a time. If the initial sample is divided into G_h random groups within each stratum, then the delete-one-group jackknife is defined as

$$v_J = \sum_h \frac{G_h - 1}{G_h} \sum_{g=1}^{G_h} \left[\hat{T}_{(hg)} - \hat{T} \right]^2 , \quad (6)$$

where $\hat{T}_{(hg)}$ is the estimated total based on deleting group (hg). All weighting steps—base weight calculation, adjustment for unknown eligibility, nonresponse adjustment, and use of auxiliary data using the retained units—are repeated in the calculation of $\hat{T}_{(hg)}$. This is done without regard to the disposition of the initial unit as a respondent, a nonrespondent, an unknown, or an ineligible. This procedure is available in WesVar® (Westat 2000) and SUDAAN® (Shah, et al 1996).

The total number of groups is $G = \sum_h G_h$. If $G_h = n_h$ and the groups are disjoint, then (6) is just the standard delete-one jackknife. Many variants of the grouped jackknife may be used in practice (e.g., see Rust and Rao 1996). For example, the random groups may include sample units from different strata. In the simulations we will consider only the case of disjoint random groups formed within each stratum with

$G_h = \bar{G}$, i.e., an equal number of groups per stratum. The union of the groups in a stratum is the initial stratum sample.

Several versions of linearization and related variance estimators might be used when the basic design is single-stage stratified sampling. Use of some of the choices would simply be a mistake but could be selected by a naïve user of some software packages. Other choices might be reasonable if some steps in weighting make small contributions to the variance. The simplest variance estimator is one that would be appropriate for the Horvitz-Thompson estimator in a stratified sample selected with varying probabilities and with replacement (Särndal, Swensson, and Wretman 1992, expression 2.9.9). If we interpret the weight w_{hi}^* as the inverse of an adjusted selection probability and add an *ad hoc* finite population correction factor, this variance estimator is

$$v_{naive1} = \sum_h \left(1 - f_{h,ER\cup IN}\right) \frac{n_{h,ER\cup IN}}{n_{h,ER\cup IN} - 1} \times \sum_{i \in s_{h,ER\cup IN}} \left(w_{hi}^* Y_{hi} - \frac{1}{n_{h,ER\cup IN}} \times \sum_{i \in s_{h,ER\cup IN}} w_{hi}^* Y_{hi} \right)^2 \qquad (7)$$

where

$s_{h,ER\cup IN}$ = set of sample eligible respondents and known ineligibles in stratum h,

$n_{h,ER\cup IN}$ = sample size in stratum h of eligible respondents and known ineligibles, and

$f_{h,ER\cup IN} = n_{h,ER\cup IN} / N_h$.

This estimator is available in SUDAAN using the option `DESIGN = STWOR` in a procedure statement, in STATA™ (Stata Corporation 2001) using the procedure `svytotal`, and in SAS™ `PROC SURVEYMEANS` (SAS Institute 2001). Units that are in s_{IN} (known ineligibles) have their Y values set to zero so that the eligibles are appropriately treated as a domain. We label this variance estimator "naïve" because it treats the resulting sample of eligible respondents and ineligibles as a with-replacement sample and ignores the adjustments for unknown eligibility and nonresponse along with the poststratification step (or other use of auxiliary data).

Another variant would be to exclude the s_{IN} cases and treat s_{ER} as a stratified without-replacement (*stwor*) sample. This estimator would be

$$v_{naive2} = \sum_h \left(1 - f_{h,ER}\right) \frac{n_{h,ER}}{n_{h,ER} - 1} \times \sum_{i \in s_{h,ER}} \left(w_{hi}^* Y_{hi} - \frac{1}{n_{h,ER}} \sum_{i \in s_{h,ER}} w_{hi}^* Y_{hi} \right)^2 \qquad (8)$$

where $s_{h,ER}$, $n_{h,ER}$, and $f_{h,ER}$ are defined in terms of the set of eligible respondents in stratum h. This estimator will typically be smaller than v_{naive1} since the ineligibles do not enter the calculation as zeroes. It is possible to compute this estimator in SAS, STATA, and SUDAAN by restricting the dataset to the eligible respondents only.

The variance estimator that is usually referred to as the linearization estimator is

$$v_L = \sum_h \left(1 - f_{h,ER\cup IN}\right) \frac{n_{h,ER\cup IN}}{n_{h,ER\cup IN} - 1} \times \sum_{i \in s_{h,ER\cup IN}} \left(w_{2hi} r_{hi} - \frac{1}{n_{h,ER\cup IN}} \times \sum_{i \in s_{h,ER\cup IN}} w_{2hi} r_{hi} \right)^2 \qquad (9)$$

where $r_{hi} = Y_{hi} - \hat{Y}_{hi}$ with $\hat{Y}_{hi} = \hat{\mu}_k$ for $(hi) \in s_{PS,k} \cap (s_{ER} \cup s_{IN})$, $\hat{\mu}_k = \sum_{(hi) \in s_{ER\cup IN}} w_{2hi} Y_{hi} / \hat{N}_k$, and w_{2hi} is the weight for unit hi after adjustment for unknown eligibility and nonresponse.

The jackknife linearization estimator is very similar to (9), but uses the w_{hi}^* weights:

$$v_{JL} = \sum_h \left(1 - f_{h,ER\cup IN}\right) \frac{n_{h,ER\cup IN}}{n_{h,ER\cup IN} - 1} \times \sum_{i \in s_{h,ER\cup IN}} \left(w_{hi}^* r_{hi} - \frac{1}{n_{h,ER\cup IN}} \times \sum_{i \in s_{h,ER\cup IN}} w_{hi}^* r_{hi} \right)^2 \qquad (10)$$

Yung and Rao (2000) defined v_{JL} without the finite population correction factor (*fpc*); insertion of the *fpc* in v_{JL} may be useful when the sampling fraction is large. The jackknife linearization estimator does not make separate adjustments for unknown eligibility and nonresponse for each replicate.

The special case of v_{JL} that is appropriate for poststratification is available in `SUDAAN`, using the `POSTWGT` and `POSTVAR` options of some procedures. For poststratification, the difference between the linearization estimator v_L and the jackknife linearization estimator is that v_{JL} includes a factor $\left(N_k / \hat{N}_k\right)^2$ for each unit that is in poststratum k. This inclusion imparts better conditional properties to v_{JL} in samples where $N_k / \hat{N}_k$ is not near 1. Accounting for poststratification does not appear to be possible in STATA v.7 or SAS v.8 unless the user writes his/her own code.

Another, related approximation to the jackknife was derived by Valliant (2002). This estimator adjusts each weighted residual using a leverage, $\Delta_{hi} = w_{2hi}\mathbf{x}'_{hi}\mathbf{A}^{-1}\mathbf{x}_{hi}/v_{hi}$, associated with sample unit (*hi*):

$$v_J^* = \sum_h \left(1-f_{h,ER\cup IN}\right)\times \sum_{s_{h,ER\cup IN}}\left(\frac{w_{hi}^* r_{hi}}{1-\Delta_{hi}} - \frac{1}{n_{h,ER\cup IN}}\times \sum_{s_{h,ER\cup IN}}\frac{w_{hi}^* r_{hi}}{1-\Delta_{hi}}\right)^2 \qquad (11)$$

Since $\Delta_{hi} < 1$, v_J^* will be larger than v_{JL}, which will typically lead to higher confidence interval coverage rates. As the number of eligible respondents and known ineligibles increases, $\Delta_{hi} \to 0$ so that the difference between v_J^* and v_{JL} will diminish.

4. Empirical Evaluation

To compare the different variance estimators, we conducted a simulation study using a poststratified population similar to ones found in human populations in which groups of units have different means.

4.1 Poststratified Population

A stratified population, with specifications shown in table 1, was generated in which poststratification was appropriate. The population has five design strata and five classes that are used as poststrata. The poststrata cut across the strata. The variable *Y* used in estimation is a 0-1 Bernoulli variable with means, P_k, ranging from 0.1 to 0.5 across the poststrata. The proportion of the population in each poststratum ranges from 0.30 to 0.12. For each unit in the population, a poststratum indicator was generated independently of design stratum membership. Each design stratum has $N_h = 1000$ units and the mean of *Y* in the realized population ranges from 0.243 to 0.259 across the design strata. In other words, there is little difference among the design strata in the means of the estimation variable while there is considerable difference in the poststrata.

Parameters in the simulation were the proportion by design-stratum whose status was known, the proportion eligible, and the proportion responding as shown in table 2. For each unit in the population Bernoulli random variables were generated with probabilities given in table 2 to determine whether a unit had a known status, was eligible, and was a respondent. This procedure was repeated for every sample that was selected.

Table 1. Specifications for the poststratified population

	Design stratum or poststratum				
	1	2	3	4	5
N_h	1000	1000	1000	1000	1000
P_h	0.247	0.259	0.256	0.248	0.243
N_k/N	0.30	0.24	0.18	0.16	0.12
P_k	0.1	0.2	0.3	0.4	0.5

Table 2. Simulation parameters for the proportions known, eligible, and responding in the poststratified population

	Design stratum				
	1	2	3	4	5
Known status	0.70	0.75	0.80	0.85	0.90
Eligible	0.70	0.75	0.80	0.85	0.90
Responding	0.60	0.65	0.70	0.75	0.80

Stratified simple random samples (*stsrs*) of size $n = 100$, 250, and 500 were selected without replacement. An equal number of sample units was allocated to each of the five design strata. Four versions of the grouped jackknife were computed: *G* = 10, 25, 50, 100. In each case, the initial sample within each stratum was divided into $G/5$ random groups. For the jackknife, 1000 *stsrs*'s were selected for each sample size. For the other variance estimators, 4000 samples were selected.

The adjustment for unknown eligibility, defined by (1), and the nonresponse adjustment, defined by (2), were made within each design-stratum. The eligible responding units and the known ineligible units were then poststratified as shown in (4).

4.2 Simulation Results—Ignorable Nonresponse

Figure 1 summarizes results for the poststratified population for ignorable nonresponse. The figure gives columns for the relative bias (relbias) of a variance estimator with respect to the mean square error, coverage of 95 percent confidence intervals (CI's), mean half-width of the confidence intervals, and the standard error of the half-widths. (The estimators labeled v.Pi and v.ssw shown in the figure were part of a larger simulation study and are not discussed here.)

The 95 percent confidence interval using a variance estimate *v* was calculated as $\hat{T}_s \pm t_{DF}\sqrt{v}$ where t_{DF} is a multiplier from the *t*-distribution with *DF* degrees of freedom, and *DF* is the degrees of freedom associated with *v*. The half-width of an interval is $t_{DF}\sqrt{v}$. For the jackknife variance estimates, we used $DF = \sum_h (G_h - 1) = G - H$. For the other variance estimators, we used $DF = \sum_h \left(n_{h,ER\cup IN} - 1\right) = n_{ER\cup IN} - H$.

Table 3 lists the mean and range of the number of eligible respondents plus known ineligibles across the samples. There is a substantial reduction from the initial sample size because of cases that had unknown status or were nonrespondents. Thus, a major source of variation is the number of sample units used in evaluating both the estimate of the total and the variance of that estimate.

Table 3. Sample sizes used in estimation in poststratified population

n	Number of eligible respondents + known ineligibles, $n_{ER \cup IN}$	
	Mean	Range
100	61.5	(42, 77)
250	153.6	(126, 181)
500	307.2	(270, 345)

The variance estimators are sorted in figure 1 by the relbias obtained for samples of size n = 100. For n = 100 v_L, v_{JL}, v_π, and v_{SSW} all have negative biases with the linearization variance estimator being the worst at –15.7 percent. The relbiases for the other estimators range from 8.4 percent for $v_J\left(G=10\right)$ to 20.4 percent for $v_J\left(G=100\right)$. The full delete-one jackknife, thus, has the largest relbias when n = 100. The biases diminish for n = 250 and 500 although the pattern of negative biases for v_L and v_{JL} and positive biases for the other variance estimators persists.

Underestimation by v_L and v_{JL} leads to CI's that cover at less than the nominal rate. When n = 100, the coverage rate with these choices is at most 92 percent. For n = 500 coverage for these estimators is near 95 percent. For the jackknife choices, overestimation of the *mse* does not necessarily lead to overcoverage by the CI's. When n = 100, 15-20 percent relbias produces coverage rates of 93.9 to 94.6 percent for $v_J\left(G=25,\ 50,\ 100\right)$. The least biased of the grouped jackknife choices, $v_J\left(G=10\right)$, has the worst CI coverage at 91.5 percent for n = 100. For the two larger sample sizes, $v_J\left(G=25,\ 50,\ 100\right)$ all have at least 95 percent coverage. The approximate jackknife, v_J^*, performs well, having relbias less than 5 percent for n = 250 and 500 and has coverage rates of 94.0, 94.9, and 95.2 at the three sample sizes.

The average half-widths and standard errors of the half-widths show some differences between the variance estimates. Average lengths are somewhat longer for the jackknife estimates and the related estimate v_J^*, especially for n = 100. Longer intervals are due to the variance estimates and the multipliers from the *t*-distribution being larger for the jackknife estimates. The estimate that stands out for its high variability is $v_J\left(G=10\right)$. The stability of the grouped jackknife increases as the number of groups increases—a phenomenon that is well known among practitioners (see also Wolter 1985, sec.4.2.5).

The estimators, v_{naive1} and v_{naive2}, are theoretically incorrect for the poststratified estimator but are included here since users of some software packages might select them. Both are overestimates since each uses the wrong residual for the poststratified estimator. The relbias of v_{naive1} ranges from about 18 percent at n = 100 to 14.6 percent at n = 250. Note that there is no decrease in relbias when moving from n = 250 to n = 500. v_{naive2} is smaller than v_{naive1} and is actually less biased because it ignores the fact that the eligibles are a domain. The positive bias of v_{naive1} leads to overcoverage by the confidence intervals, although the problem is not severe. At n = 250 and 500, for example, the empirical coverage rate using v_{naive1} is 96.4 percent.

5. Conclusion

Two general types of variance estimators used in survey sampling are ones based on squared residuals, like linearization variance estimators, and replication variance estimators. The simulation results reported here are part of a larger study available from the author.

Theory for the two types of estimators shows that asymptotically there is little difference in large samples with full response. The basic design-based or model-based theory does involve some strong assumptions, e.g., the first-stage is selected with replacement or the first-stage units are independent. More importantly, the possibilities that there are ineligible units and that some units will not respond are often not considered when comparing the variance estimators. For the replication variance estimators, there is literature showing that adjusted data values can be used to create consistent variance estimators. However, these adjustments may vary depending on the form of the basic estimator (total, mean, ratio, etc.) and are not included in the commercial software packages now available.

In the simulations presented here, the linearization estimators and several others based on squared residuals are underestimates. This problem is considerably worse when there is ignorable nonresponse, as opposed to full response, and leads to confidence intervals that cover at less than the nominal rate. The one exception among the squared-residual estimators is a leverage-adjusted estimator that approximates the jackknife and tends to be somewhat of an overestimate.

The grouped jackknife estimator tends to be an overestimate and the degree of overestimation is worse with smaller sample sizes. This overestimation is accompanied by some overcoverage by confidence intervals, although the excess above the nominal level is small.

In summary, when there is ignorable nonresponse, the only estimator in this study that combines reasonably small positive bias with near-nominal confidence interval coverage is the leverage-adjusted estimator v_J^*.

6. References

Groves, R.M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (2002). *Survey Nonresponse*. New York: John Wiley & Sons.

Lemeshow, S. (1979). The Use of Unique Statistical Weights for Estimating Variances with the Balanced Half-Sample Technique. *Journal of Statistical Planning and Inference*, **3**, 315-323.

Lundström, S., and Särndal, C.-E. (1999). Calibration as a Standard Method for Treatment of Nonresponse. *Journal of Official Statistics*, **15**, 305-327.

Rao, J.N.K. (1996). On Variance Estimation with Imputed Survey Data. *Journal of the American Statistical Association*, **91**, 499-506.

Rust, K., and Rao, J.N.K. (1996). Variance estimation for complex estimators in sample surveys. *Statistics in Medicine*, **5**, 381-397.

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

SAS Institute (2001). *SAS Stat Version 8*, Cary NC: SAS Institute.

Shah, B.V., Barnwell, B., Bieler, G., Boyle, K., Folsom, R., Lavange, L., Wheeless, S., and Williams, R. (1996). *Technical Manual: Statistical Methods and Algorithms Used in SUDAAN*, Research Triangle Park, NC: Research Triangle Institute.

Stata Corporation (2001). *STATA Statistical Software, Version 7*, College Station TX: Stata Corporation.

Valliant, R. (1993). Poststratification and conditional variance estimation. *Journal of the American Statistical Association*, **88**, 89-96.

Valliant, R. (2002). Variance estimation for the general regression estimator. *Survey Methodology*, **28**, 103-114.

Westat (2000). *WesVar 4.0 User's Guide*, Rockville MD: Westat.

Wolter, K. (1985). *Introduction to Variance Estimation*, New York: Springer-Verlag.

Yung, W., and Rao, J.NK. (1996). Jackknife Linearization Variance Estimators under Stratified Multi-stage Sampling. *Survey Methodology*, **22**, 23-31.

Yung, W., and Rao, J.NK. (2000). Jackknife Variance Estimation under Imputation for Estimators Using Poststratification Information. *Journal of the American Statistical Association*, **95**, 903-915.

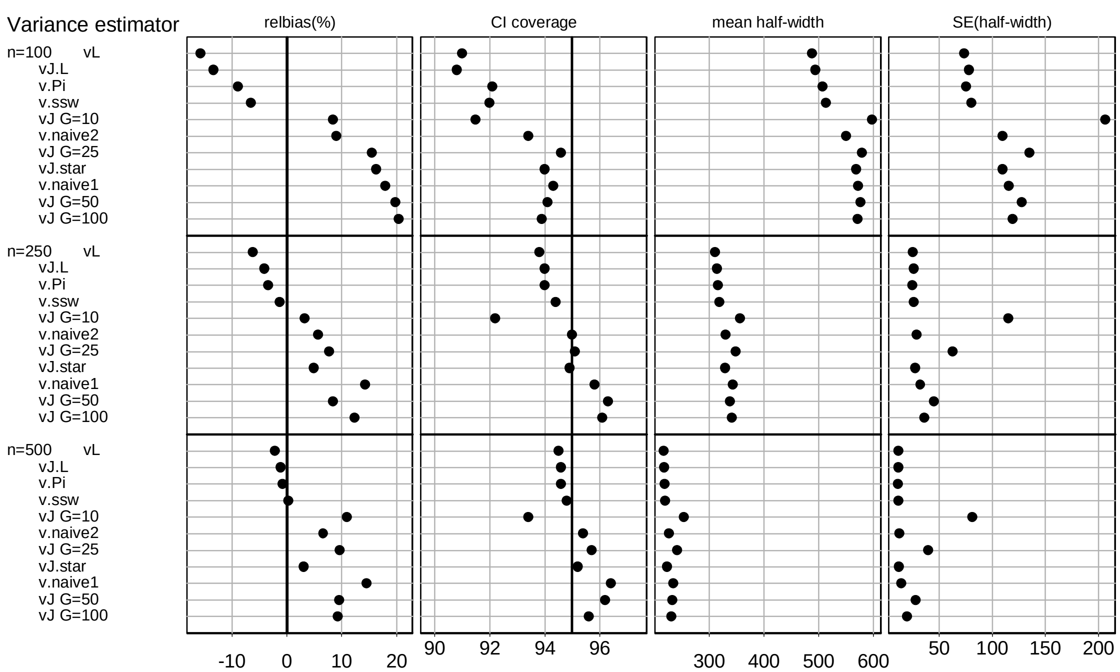

Figure 1. Comparisons of the relbias, mean half-widths of 95 percent confidence intervals, and standard error of half-widths of different estimators of variance for the poststratified estimator of a total