# ON THE EFFICIENCIES OF UNBIASED REGRESSION AND REGRESSION—TYPE ESTIMATORS IN FINITE POPULATIONS

ARIJIT CHAUDHURI
Indian Statistical Institute, Calcutta
and
ARUN KUMAR ADHIKARY<sup>1</sup>
Operations Research Group, Baroda

#### SUMMARY

Postulating linear regression models efficiencies of exactly 'design-unbiased' 'regression' and 'regression-type' estimators for finite population means are studied vis-a-vis the 'design-biased' regression estimator based on SRSWOR Scheme.

Keywords: Finite population; Unbiased regression estimator, Super-population models.

## Introduction

Singh and Srivastava [10] suggested two sampling schemes ensuring design-unbiasedness of a 'regression' and a 'regression-type' estimator for a finite population mean. Here the purpose is to investigate the efficiency of each relative to that of the design-biased 'regression' estimator based on SRSWOR Scheme. A usual 'linear regression model' in a super-population set-up is postulated and their relative performances found out.

<sup>1.</sup> When the paper was written, the second author was at Indian Statistical Institute, Calcutta.

# 2. Formulation of the Problem and the Results

Let Y and X be real-variates with values  $Y_i$ ,  $X_i$  (i = 1, ..., N) for a finite population of N units with means  $\overline{Y}$ ,  $\overline{X}$ . For any sample (generically, each supposed to be of a fixed size n) the regression estimator for  $\overline{Y}$  is

$$t = \bar{y} + b \; (\bar{X} - \bar{x})$$

where  $\bar{y}$ ,  $\bar{x}$  are sample means and  $b = sxy/s_x^2$  where

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$
,  $sxy = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}) y_i$ 

where  $\Sigma'$  is sum for a sample,  $y_i$ ,  $x_i$ 's are sampled values.

Writing 
$$x_i' = x_i - \overline{x}$$
,  $X_i' = X_i - \overline{X}$ ,  $y_i' = y_i - \overline{Y}$ ,  $Y_i' = Y_i - \overline{Y}$ ,  $\overline{x}' = \overline{x} - \overline{X}$ ,  $\overline{y}' = \overline{y} - \overline{Y}$  and

$$S_a^2 = \frac{1}{N-1} \sum_{i} X_i^{\prime 2}$$

where  $\Sigma$  is sum over the population, Singh and Srivastava [10] gave a scheme for which selection probability for s is

$$p(s) = \frac{1}{\binom{N}{n}} \frac{s_x^2}{S_x^2}$$

For this  $E_r(t) = \vec{Y}$  but  $E_c(t) \neq \vec{Y}$  where  $E_p$ ,  $E_c$  are expectation-operators with respect to designs p and equal probability sampling (SRSWOR), respectively.

For their another scheme the selection probability is

$$q(s) = \frac{1}{\binom{N-1}{n-1}} \frac{\sum' \chi_i'^2}{\sum \chi_i'^2}$$

Their proposed regression-type estimator is

$$t_1 = \frac{n (N-1)}{N (n-1)} [\bar{y} - \bar{x}' \Sigma' y_t x_i' / \Sigma' x_i'^{2}]$$

for which  $E_q$   $(t_1) = \overline{Y}$  i.e.  $t_1$  is design (q)—unbiased for  $\overline{Y}$ .

Writing  $V_p$ ,  $V_c$ ,  $V_q$  respectively as variance-operators with respect to designs p, SRSWOR and q it may be checked that

$$V_{\mathfrak{p}}(t) = \frac{1}{S_x^2} E_{\mathfrak{p}}[\bar{y}' s_x^2 - 2 \bar{y}' \bar{x}' sxy + \bar{x}'^2 s_{xy}^2 / s_x^2]$$

$$V_{q}(t_{1}) = \frac{n}{N} \cdot \frac{(N-1)^{2}}{(n-1)^{2}} \frac{1}{\sum X_{i}^{2}} E_{e} \left[ \{ \bar{y} - \bar{x}' \sum_{i}' y_{i} x_{i}' | \sum_{i}' x_{i}^{2} \}^{2} \sum_{i}' x_{i}^{2} \right] - \bar{Y}^{2}.$$

Postulate a model enabling us to write

$$Y_i = \alpha + \beta X_i + e_i$$

where  $\alpha$ ,  $\beta$  are unknown regression-parameters and  $e_i$  is a random variate with conditional (given  $X_1$  values) expectation and variance (with respect to an implied super-population) as 0 and  $\delta X_i^g$  with  $\delta > 0$  and  $0 \le g \le 2$ . Writing  $E_M$  as the expectation-operator under this model (assuming  $X_i$ 's as non-stochastic) and incorporating this model into the system work out the following:

$$E_M V_p(t) = \frac{\delta}{S_x^2} E_o \left[ \left( \frac{1}{n^2} \sum_{i} x_i^g - \frac{2}{Nn} \sum_{i} x_i^g + \frac{1}{N^2} \sum_{i} X_i^g \right) s_a^2 \right.$$

$$\left. - 2 \left( \frac{1}{n} - \frac{1}{N} \right) \bar{x}' \cdot \frac{1}{n-1} \sum_{i} x_i^g (x_i - \bar{x}) \right.$$

$$\left. + \frac{1}{(n-1)^2} \cdot \frac{\bar{x}'^2}{s_x^2} \sum_{i} x_i^g (x_i - \bar{x})^3 \right]$$

$$= \frac{n}{n-1} \left( \frac{1}{n} - \frac{1}{N} \right) \delta \text{ (in case } g = 0) = I \text{ (say)}$$

$$\approx \left( \frac{1}{n} - \frac{1}{N} \right) \delta \text{ for large } n$$

$$E_{M}V_{q}(t_{1}) = \frac{n}{N} \frac{(N-1)^{2}}{(n-1)^{2}} \frac{1}{\sum X_{i}^{'2}} E_{o} \left[ (\alpha + \beta \bar{x})^{2} \sum_{i}' x_{i}^{'2} + \frac{\delta}{n^{2}} \sum_{i}' x_{i}^{o} \sum_{i}' x_{i}^{o} \sum_{i}' x_{i}^{'2} - 2 (\alpha + \beta \bar{x}) \bar{x}' \sum_{i}' x_{i}' (\alpha + \beta x_{i}) \right]$$

$$- 2 \frac{\delta}{n} \bar{x}' \sum_{i}' x_{i}' x_{i}^{o} + \frac{\bar{x}'^{2} \{ \sum_{i}' x_{i}' (\alpha + \beta x_{i}) \}^{2} + \bar{x}'^{2} \delta \sum_{i}' x_{i}'^{2} x_{i}^{o}}{\sum_{i}' x_{i}'^{2}} \right]$$

$$= (\alpha + \beta \bar{x})^{2} \left[ \frac{n}{n-1} \frac{N-1}{N(n-1)} \frac{1}{s_{x}^{2}} E_{o} \left\{ \frac{\sum_{i}' X_{i}'^{2} - n \bar{x}'^{2})^{2}}{\sum_{i}' x_{i}'^{2}} \right\} - 1 \right]$$

$$+ \frac{n}{n-1} \left( \frac{1}{n} - \frac{1}{N} \right) \delta$$

$$= (\alpha + \beta \bar{x})^{2} (A - 1) + \frac{n}{n-1} \left( \frac{1}{n} - \frac{1}{N} \right) \delta = II \text{ (say)}$$

For t based on SRSWOR, an asymptotic expression for  $V_o(t)$  is well-known (vide Cochran [4] as  $[(N-n)/Nn] S_y^2 (1-\rho^2)$ , where

$$S_y^2 = \frac{1}{N-1} \sum Y'^2$$
,  $\rho = \frac{S_{xy}}{S_x S_y}$ , where  $S_{xy} = \frac{1}{N-1} \sum Y'_i X'_i$  and  $E_M V_o(t) = \left(\frac{1}{n} - \frac{1}{N}\right) \delta \left[\frac{1}{N} \sum X'^0_i - \frac{1}{S_x^2} \frac{1}{(N-1)^2} \sum X'^0_i X'^2_i\right]$ ,

In case g = 0, we have

$$E_M V_o(t) = \frac{N-2}{N-1} \left( \frac{1}{n} - \frac{1}{N} \right) \delta = III \text{ (say)}.$$

For a theoretical comparison of strategies, consider throughout the case g=0 (the homoscedastic case). Singh and Srivastava [10] also illustrated a special (normal) case of this. Though values of g over the interval [1, 2] are more realistic we find it difficult to derive elegant results with such values of g. However in the Appendix we present a Monte Carlo study illustrating the performance of these strategies with g in [0, 2]

By Cauchy-Schwarz's inequality we get

$$\begin{split} E_{\bullet}\left(\Sigma' \, x_{i}^{\prime 2}\right) \, E_{o} \left\{ \frac{\Sigma' \, x_{i}^{\prime 2} - n \, \bar{x}^{\prime 2})^{2}}{\Sigma' \, x_{i}^{\prime 2}} \right\} &\geqslant E_{o}^{2} \left(\Sigma' \, x_{i}^{\prime 2} - n \, \bar{x}^{\prime 2}\right) \\ \text{i.e. } E_{o} \left\{ \frac{\Sigma' \, x_{i}^{\prime 2} - n \, \bar{x}^{\prime 2})^{2}}{\Sigma' \, x_{i}^{\prime 2}} \right\} &\geqslant \frac{E_{o}^{2} \left(E' \, x_{i}^{\prime 2} - n \, \bar{x}^{\prime 2}\right)}{E_{o} \left(\Sigma' \, x_{i}^{\prime 2}\right)} \\ \text{Now } E_{o} \left(\Sigma' \, x_{i}^{\prime 2} - n \, \bar{x}^{\prime 2}\right) &= E_{o} \, \Sigma' \, (x_{i} - \bar{x})^{2} = (n - 1) \, S_{x}^{2} \\ \text{and } E_{o} \left(\Sigma' \, x_{i}^{\prime 2}\right) &= E_{o} \, \left[\Sigma' \, (x_{i} - \bar{x})^{2} + n \, (\bar{x} - \bar{X})^{2}\right] \\ &= (n - 1) \, S_{x}^{2} + n \left(\frac{1}{n} - \frac{1}{N}\right) \, S_{x}^{2} \\ \text{so that } E_{o} \left\{ \frac{\Sigma' \, x_{i}^{\prime 2} - n \bar{x}^{\prime 2})^{2}}{\Sigma' \, x_{i}^{\prime 2}} \right\} \geqslant \frac{n - 1}{n} \, \frac{N \, (n - 1)}{N - 1} \, S_{x}^{2} \end{split}$$

implying that  $A \geqslant 1$  and hence

Result 1 (a) 
$$II > I > III$$
;  
(b)  $II > I \simeq III$ .

The result (b) is more discernible because the asymptotic expression for  $V_{\bullet}(t)$  is valid only for large samples and it underestimates the variance of t for small samples. In fact, it is more important to note that I = III for large n (and hence large N). To  $0(1/n^2)$  also, if g = 0,  $E_M V_c(t) = n - 2/n - 3(1/n - 1/N) \delta = I$ . So, it is not quite helpful to learn from Singh and Srivastava [10] that  $V_c/V_p = 1.70$  for n = 4 when 11 finite populations of size N = 20 are generated each from bivariate normal distributions, their apparent contradictions with ours.

For t based on SRSWOR, we also work out an exact expression for  $E_M V_c(t)$  as follows:

$$E_{M} V_{c}(t) = \delta E_{c} \left[ \left( \frac{1}{n^{2}} \sum_{i} x_{i}^{q} - \frac{2}{Nn} \sum_{i} x_{i}^{q} + \frac{1}{N^{2}} \sum_{i} x_{i}^{q} \right) - 2 \overline{x}' \left( \frac{1}{n} - \frac{1}{N} \right) \frac{\sum_{i} x_{i}^{q} (x_{i} - \overline{x})}{\overline{\Sigma}' (x_{i} - \overline{x})^{2}} \overline{x}'^{2} \frac{\sum_{i} x_{i}^{q} (x_{i} - \overline{x})^{2}}{\{\Sigma' (x_{i} - \overline{x})^{2}\}^{2}} \right]$$

$$= \left( \frac{1}{n} - \frac{1}{N} \right) \delta + \delta E_{c} \left[ \frac{\overline{x}'^{2}}{\Sigma' (x_{i} - \overline{x})^{2}} \right] \text{ (in case } g = 0)$$

If following Avadhani and Srivastava [2] we assume n and N so large that for each s we may ignore the error in neglecting  $\bar{x}'$ , then we have approximately

$$E_M V_c(t) \simeq \left(\frac{1}{n} - \frac{1}{N}\right) \delta = III_A(\text{say})$$

With this assumption

$$E_M V_q(t_1) \simeq \frac{N-n}{N(n-1)} \left(\alpha + \beta \overline{X}\right)^2 + \frac{n}{n-1} \left(\frac{1}{n} - \frac{1}{N}\right) \delta = I V_A \text{ (say)}$$

and hence

Y

Result 2 (a)  $IV_A > I \simeq III \simeq III_A$  for large n

(b) 
$$E_M V_g(t_1) > E_M V_c(t) > E_M V_r(t)$$
 for small or moderate n

(c) 
$$E_m V_q(t_1) > E_m V_c(t) \approx E_M V_p(t)$$
 for large  $n$  if  $g = 0$ .

If further it is assumed that  $X_i$ 's are random variates distributed independently identically normally with a mean m and variance  $\sigma^2$  (say), then denoting the model-expectation (for variation of  $e_i$  given  $x_i$  and then over  $X_i$ ) by  $E_M$  and assuming (on the strength of the law of large numbers)  $\overline{X}$  (for finite population with N large) to equal m, we have

$$E'_{M}V_{c}(t) \simeq \delta \left[ \left( \frac{1}{n} - \frac{1}{N} \right) + \frac{1}{n(n-3)} \right] = III_{B} \text{ (say)},$$

$$E'_{M}V_{c}(t) \simeq \frac{2(n+1)}{(n-1)^{2}(n-2)} (\alpha + \beta_{m})^{2} + \frac{n}{n-1} \left( \frac{1}{n} - \frac{1}{N} \right) \delta$$

$$= IV_{B} \text{ (say)}$$

on neglecting the error in writing 1 for N-1/N and  $\sum X_i^2 \simeq N \sigma^2$  (using the fact that  $\sum X_i'^2/\sigma^2 = \sum (X_i - m)^2/\sigma^2$  is a chi-square variate with N degree of freedom). So

Result 3  $III_B > I$ ; but

$$IV_B - III_B = \frac{2(n+1)}{(n-1)^2(n-2)} (\alpha + \beta_m)^2 - \frac{\delta}{n-1} \left[ \frac{2}{n(n-3)} + \frac{1}{N} \right]$$

Which may be either positive or negative.

Incidentally, we note that  $E_M E_o$   $(t - \overline{Y}) = 0$ , although  $E_o$   $(t - \overline{Y}) \neq 0$  i.e. t is 'model-design-unbiased' though it is 'design-biased'. We also note that  $E_M E_T$   $(t - \overline{Y}) = 0$  and  $E_M$   $(t - \overline{Y}) = 0$  i.e. t is 'model-design (p)-unbiased' as well as 'model-unbiased'. But  $E_M E_q$   $(t_1 - \overline{Y}) = 0$  although  $E_M$   $(t_1 - \overline{Y}) \neq 0$  i.e.  $t_1$  is 'model-design(q)-unbiased' though it is not 'model-unbiased'.

So it is concluded that two strategies due to Singh and Srivastava [10] may offer fare worse than the classical strategy of regression estimator with SRSWOR Scheme; however the latter also does not always beat the former two.

It may be recalled that a similar situation prevails in respect of ratio estimators based on SRSWOR and Lahiri-Midzuno-Sen (LMS, say, in brief, [5], [6], [9]) scheme—the latter is design-unbiased, the former is design-biased but in terms of efficiency neither is uniformly better than the other (as may be checked from Arnab [1], Avadhani and Srivastava [2] and Chaudhuri and Adhikary [3] among others).

#### 3. Variance Estimation

Incidentally, it may be noted that the unbiased estimators for  $V_p(t)$  and  $V_o(t_1)$  proposed by Singh and Srivastava [10] following Raj [7] should often turn out negative because it is known that Raj's variance-estimator (for ratio estimator based on LMS-Scheme) is negative for samples with high selection-probabilities. Following Rao [8] and noting that  $V_p(t)$  and  $V_q(t_1)$  equal 0, when  $Y_i = C X_i$  and  $Y_i = C X_i'$ ,  $i = 1, \ldots, N$  respectively (for some  $C \neq 0$ ), these can be written as

$$V_{p}(t) = -\frac{1}{N^{2}} \sum_{i < j=1}^{N} d_{ij} X_{i} X_{j} \left( \frac{Y_{i}}{X_{i}} - \frac{Y_{j}}{X_{j}} \right)^{2}$$
and
$$V_{q}(t_{1}) = -\frac{1}{N^{2}} \sum_{i < j=1}^{N} d'_{ij} X'_{j} X'_{i} \left( \frac{Y_{i}}{X'_{i}} - \frac{Y_{j}}{X'_{j}} \right)^{2},$$
where
$$d_{ij} = \left( \sum_{s \ni i, j} d_{is} d_{js} p(s) \right) - 1 = e_{ij} - 1, \text{ say,}$$

$$d'_{ij} = \sum_{s \ni 1, j} \left( d'_{is} d'_{js} q(s) \right) - 1 = e'_{ij} - 1, \text{ say,}$$

where 
$$d_{is} = 0$$
 if  $s \not\ni i$ 

$$= \frac{N}{n} \left[ 1 - n \, \bar{x}' \, \frac{(x_i - \bar{x})}{\Sigma' \, (x_i - \bar{x})^2} \, \right] \text{ if } s \ni i$$
and  $d'_{is} = 0$  if  $s \not\ni i$ 

$$= \frac{N-1}{n-1} \left[ 1 - n \, \bar{x}' \, \frac{x'_i}{\Sigma' \, x'_i} \, \right] \text{ if } s \ni i$$

then, non-negative quadratic unbiased estimators of  $V_p(t)$  and  $V_q(t_1)$  are necessarily of the form, respectively,

$$\hat{V}_{x}(t) = -\frac{1}{N^{2}} \sum_{i < j} \sum_{i \in s} d_{ij}(s) x_{i} x_{j} \left(\frac{y_{i}}{x_{i}} - \frac{y_{i}}{x_{j}}\right)^{2}$$

$$\hat{V}_{q}(t_{1}) = -\frac{1}{N^{2}} \sum_{i \in s} \sum_{j \in s} d'_{ij}(s) x'_{i} x'_{j} \left(\frac{y_{i}}{x'_{i}} - \frac{y_{i}}{x'_{j}}\right)^{2}$$

where  $d_{is}(s) = 0$  if  $s \ngeq i, j, \sum_{s} d_{ij}(s) p(s) = d_{ij}, i < j$  and  $d'_{ij}(s) = 0$  if  $s \trianglerighteq i, j, \sum_{s} d'_{ij}(s) q(s) = d'_{ij}, i < j$ . We may consider two classes of estimators for each of  $V_{x}(t)$  and  $V_{q}(t_{1})$  of the above-noted form as follows:

$$C_{1}: di_{j}(s) = e_{ij} \left[ \frac{f_{i}}{\pi_{ij}} + \frac{1 - f_{1}}{M_{2}p(s)} \right] - \left[ \frac{f_{2}}{\tau_{ij}} + \frac{1 - f_{2}}{M_{2}p(s)} \right]$$

$$C_{2}: d_{ij}(s) = d_{is} d_{js} - \left[ \frac{f_{3}}{\pi_{ij}} + \frac{1 - f_{3}}{M_{2}p(s)} \right]$$

$$C'_{1}: d'_{ij}(s) = e'_{ij} \left[ \frac{f'_{i}}{\pi'_{ij}} + \frac{1 - f'_{i}}{M_{2}q(s)} \right] - \left[ \frac{f'_{2}}{\pi'_{ij}} + \frac{1 - f'_{2}}{M_{2}q(s)} \right]$$

$$C'_{2}: d'_{ij}(s) = d'_{is} d'_{js} - \left[ \frac{f'_{3}}{\pi'_{ij}} + \frac{1 - f'_{3}}{M_{2}q(s)} \right]$$

where

1-

$$M_2 = {N-2 \choose n-2}, \, \pi_{ij} = \sum_{s \ni i, j} p(s), \, \pi'_{ij} = \sum_{s \ni i, j} q(s);$$

 $f_i, f'_i$  (i = 1, 2, 3) are any constants independent of sample chosen and of variate-values chosen arbitrarily before the survey with values of each over the closed interval [0, 1]—simplest and the most practical choice to be recommended is either 0 or 1 (independently of one another) for each. The inclusion-probabilities of first two orders for the sampling schemes

p(s) and q(s) are respectively obtained as

$$\pi_{i} = \sum_{s \ni i} p(s) = \frac{nN - N - n}{N(N - 2)} + \frac{N - n}{N - 2} \frac{X_{i}'^{2}}{\sum X_{i}'^{2}}$$

$$\pi_{ij} = \sum_{s \ni i, j} p(s) = \frac{(n - 2)(Nn - N - 2n)}{N(N - 2)(N - 3)} \frac{(X_{j} - X_{i})^{2}}{N\sum X_{i}'^{2}}$$

$$\cdot \frac{(N - n)(N - n - 1)}{(N - 2)(N - 3)} + \frac{X_{i}'^{2} + X_{j}'^{2}}{\sum X_{j}'^{2}} \frac{(n - 2)(N - n)}{(N - 2)(N - 3)}$$

$$\pi'_{i} = \sum_{s \ni i} q(s) = \frac{n - 1}{N - 1} + \frac{N - n}{N - 1} \frac{X_{i}'^{2}}{\sum X_{i}'^{2}}$$
and 
$$\pi'_{ij} = \sum_{s \ni i, j} q(s) = \frac{n - 1}{N - 1} \frac{n - 2}{N - 2} + \frac{n - 1}{N - 1} \frac{N - n}{N - 2} \frac{X_{i}'^{2} + X_{j}'^{2}}{\sum X_{i}'^{2}}$$

### REFERENCES

- [1] Arnab, R. (1979): On the relative efficiencies of some sampling strategies under a super-population model, J. Ind. Soc. Agri. Statist., 31: 89-96.
- [2] Avadhani, M. S. and Srivastava, A. K. (1972): A comparison of Midzuno-Sen scheme with P.P.S. sampling without replacement and its application to successive sampling, Ann. Inst. Statist. Math., 24: 153-164.
- [3] Chaudhuri, A. and Adhikary, A. K. (1983): On the efficiency of Midzuno-Sen's strategy relative to several ratio-type estimators under a particular regression model. *Biometrika*, 70: 689-693.
- [4] Cochran, W. G. (1963): Sampling Techniques, Wiley and Sons, New York.
- [5] Lahiri, D. B. (1951): A method of sample selection providing unbiased ratio estimates, *Bull. Int. Statist. Inst.*, 33: 133-140.
- [6] Midzuno, H. (1952): On the sampling system with probability proportional to sum of sizes, Ann. Inst. Statist. Math., 3: 99-107.
- [7] Raj, D. (1954): Ratio estimation in sampling with equal and unequal probabilities, J. Ind. Soc. Agri. Statist., 6: 127-138.
- [8] Rao, J. N. K. (1979): On deriving mean square errors and their non-negative unbiased estimators in finite population sampling, J. Ind. Statist. Assoc., 17: 125-136.
- [9] Sen, A. R. (1953): On the estimation of variance in sampling with varying probabilities, J. Ind. Soc. Agri. Statist., 5: 119-127.
- [10] Singh, P. and Srivastava, A. R. (1980): Sampling schemes providing unbiased regression estimators, *Biometrika*, 67: 205-209.

# **APPENDIX**

# A Monte Carlo Study

In the above theoretical study only the case g=0 is investigated while g is likely to lie between 1 and 2 in many practical situations. So take up a Monte Carlo Study which can be used to confirm the theoretical results as well as to fill up the gap of theoretical study. Consider the following two cases with parametric combinations as given below:

Case I: 
$$N = 5$$
,  $n = 2$ ,  $\alpha = 2$ ,  $\beta = 3$ ,  $\delta = 2$   
Case II:  $N = 6$ ,  $n = 3$ ,  $\alpha = 6$ ,  $\beta = 3$ ,  $\delta = 5$ 

The values of x considered in the two cases are 2, 4, 5, 3, 6, and 6, 9, 12, 18, 21, 24 respectively. Choosing  $e_i$  from N (0,  $\delta x_i^q$ ) the corresponding y-values may be obtained from the relation  $y_i = \alpha + \beta x_i + e_i$ . For example, for g = 1 the corresponding y-values in the two cases respectively turn out as 8.1, 15.7, 20.7, 13.4, 20.9 and 29.2, 36.1, 47.9, 69.0, 69.9, 95.6. But as the expected variances of the estimators are independent of y-values we do not show them for all values of g. For different values of g viz. g = 0.0, 0.5, 1.0, 1.5 and 2.0 we report below in Table 1 the values of  $E_M V_p(t)$ ,  $E_M V_q(t_1)$  and  $E_M V_c(t)$  using exact expressions for them which are denoted by V1, V2 and V3 respectively. For  $E_M V_c(t)$  we have also considered the asymptotic expression for  $V_c(t)$  in which case  $E_M V_c(t)$  is denoted by V4. The figures within the parentheses correspond to Case II.

TABLE 1—EXPECTED VARIANCES OF FOUR SAMPLING STRATEGIES FOR VARIOUS VALUES OF g

	V1	V2	<i>V</i> 3	V4
g = 0.0	1.20 ( 1.25)	64.86 (253.53)	2.82 ( 2.12)	0.45 ( 0.67)
g = 0.5	2.36 ( 9.49)	66.02 (253.98)	5.55 ( 22.81)	0.89 ( 2.52)
g = 1.0	4.80 ( <b>3</b> 6.13)	(8.46 (268.01)	11.29 ( 88.30)	1.80 ( 10.00)
g = 1.5	9.95 (1 <b>54</b> .14)	73.72 (326.97)	23.43 (366.34)	3.73 ( 41.13)
g = 2.0	21.04 (662.52)	85.26 (582.35)	49.61(1583.86)	7.89 (174.50)

Defining the relative efficiencies of strategies as  $Ei = 100 \ (V_2/V_t)$ , i = 1, 2, 3, 4 we present below in Table 2 the values of the relative efficiencies of the strategies for various values of g considered above. Here also the figures within the parentheses correspond to Case II.

TABLE 2—RELATIVE EFFICIENCIES OF SAMPLING STRATEGIES FOR VARIOUS VALUES OF g

	<i>E</i> 1	E2	E3	E4
g = 0.0	5405 (20042)	100 (100)	2300 (11817)	14413 (37392)
g = 0.5	2797 (2676)	100 (100)	1190 (1113)	7418 (10079)
g = 1.0	1426 (742)	100 (100)	606 (303)	3803 (2680)
g = 1.5	741 (212)	100 (100)	315 (89)	19 <b>7</b> 6 (79 <b>5</b> )
g = 2.0	405 (88)	100 (100)	172 (37)	1081 (334)

Note from Table 1 that for g = 0.0 in both the cases

$$V_2 > V_3 > V_1 > V_4$$

which conforms to our theoretical result. The same ordering is also found to hold good for other values of g considered in both the cases excepting the situations g = 1.5 and g = 2.0 in Case II.