ESTIMATING THE VARIANCE OF THE SLOPE OF A LINEAR REGRESSION IN A STRATIFIED RANDOM SAMPLE WITH THE BALANCED HALF-SAMPLE TECHNIQUE

Stanley Lemeshow, University of Massachusetts/Amherst

Abstract

Estimation of the variance of the slope of the linear regression under a variety of computer generated situations with the Balanced Half-Sample procedure is considered. Three estimates for the population slope β, each of which is optimal for different situations, are presented. The method of applying the Balanced Half-Sample technique with each of these estimates is investigated and then evaluated with a Monte Carlo experiment.

The results of the investigation show that variance estimates of the slope are highly biased and very unstable unless sizeable numbers of observations are selected from each stratum. The choice of the best estimator of β from the three presented depends on the particular situation under consideration.

1. Introduction

The balanced half-sample (BHS) technique has been used for some time to estimate the variance of the combined ratio estimate in such large-scale sample surveys as the Health Examination Survey (HES) and the Health Interview Survey (HIS) of the National Center for Health Statistics (NCHS). Other large-scale surveys have used variance estimation techniques such as a Taylor Series expansion or the linearization method for the same purpose. Proponents of the BHS technique have claimed that an estimate of the variance of any non-linear estimate of interest could be obtained without having to derive new expressions for the approximations to the variances as would be the case with the linearization method. The properties of the BHS technique have been documented by McCarthy (1966, 1969) and Lemeshow and Epp (1977) and its properties for the ratio estimate have been presented by Lemeshow and Levy (1977).

This paper considers the slope of the linear regression as a particular non-linear estimate. The BHS technique is used to estimate its variance in a variety of computer generated situations. The ability of this method to effectively estimate the variance of the slope is carefully considered and evaluated through the use of Monte-Carlo experiments. This is done in the context of a stratified random sample.

Specifically, consider a population subdivided into L strata of equal weight. A random sample of size n is drawn from each stratum and observations denoted (x_{ij},y_{ij}), $i=1,\ldots,L$, $j=1,\ldots,n$ are made. The pertinent population parameters are denoted by μ_x, μ_y, σ_x^2, σ_y^2, σ_{xy}, $\rho=\sigma_{xy}/\sigma_x\sigma_y$ and $\beta=\sigma_{xy}/\sigma_x^2$. The strata parameters and observations are illustrated as below:

Strata	Pop. Size	Strata Parameters	Sample Observations
1	N_1	$\mu_x^{(1)},\mu_y^{(1)},\sigma_{xx}^{(1)},\sigma_{yy}^{(1)},\sigma_{xy}^{(1)},\beta^{(1)},\rho^{(1)}$	$(x_{11},y_{11}),\ldots,(x_{1n},y_{1n})$
2	N_2	$\mu_x^{(2)},\mu_y^{(2)},\sigma_{xx}^{(2)},\sigma_{yy}^{(2)},\sigma_{xy}^{(2)},\beta^{(2)}\ \rho^{(2)}$	$(x_{21},y_{21}),\ldots,(x_{2n},y_{2n})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
L	N_L	$\mu_x^{(L)},\mu_y^{(L)},\sigma_{xx}^{(L)},\sigma_{yy}^{(L)},\sigma_{xy}^{(L)},\beta^{(L)},\rho^{(L)}$	$(x_{L1},y_{L1}),\ldots,(x_{Ln},y_{Ln})$
	N		

2. Estimating β

Let the population slope be defined as

$$\beta = \sigma_{xy}/\sigma_x^2 . \tag{2.1}$$

The following three estimates are considered for this parameter:

(i)
$$\hat{\beta}_1 = \frac{\sum_{i=1}^{L}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})(y_{ij}-\bar{y}_{..})}{\sum_{i=1}^{L}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})^2} \tag{2.2}$$

where the deviations are taken about the overall means.

(ii)
$$\hat{\beta}_2 = \frac{\sum_{i=1}^{L}\left[\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})(y_{ij}-\bar{y}_{i.})\right]}{\sum_{i=1}^{L}\left[\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2\right]} \tag{2.3}$$

where the deviations are taken about the within-stratum means.

(iii)
$$\hat{\beta}_3 = \frac{1}{L}\sum_{i=1}^{L}\left\{\frac{\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})(y_{ij}-\bar{y}_{i.})}{\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2}\right\} \tag{2.4}$$

which is the average of the strata slopes.

The following theorem is presented without proof:

Theorem 1: If n observations are randomly selected from each of L strata and β is defined as in (2.1), then as $n\to\infty$,

(i) $\hat{\beta}_1 \xrightarrow{p} \beta$ irrespective of the distribution of x or y

(ii) $\hat{\beta}_2 \xrightarrow{p} \beta$ if $\mu_x^{(i)}=\mu_x^*$ for all i, where μ_x^* is any arbitrary constant

(iii) $\hat{\beta}_3 \xrightarrow{p} \beta$ if $\mu_x^{(i)}=\mu_x^*$ and $\sigma_{xx}^{(i)}=\sigma_{xx}^*$ for all i, where μ_x^* and σ_{xx}^* are any abribrary constants.

A proof of the theorem is given by Lemeshow (1976).

The choice of the appropriate method of estimating β is not always clear because the parameters of the independent variable in each stratum are often unknown. In certain cases the choice is clear. For instance, if x and y are bivariate normal, and if the distribution of x is the same in each stratum, then $\hat{\beta}_3$ is the maximum likelihood estimate of β and as such is known to be the minimum variance unbiased estimate. If we only have $\mu_x^{(i)}=\mu_x^*$ in each stratum, then both $\hat{\beta}_1$ and $\hat{\beta}_2$ are consistent. Consistency is always assured using $\hat{\beta}_1$, but clearly, use of this estimate may provide an unnecessary loss of precision.

3. Estimating the Variance of $\hat{\beta}$ with the BHS Technique

In the half-sample method, assume the n observations from each stratum are divided into two groups of r=n/2 observations each. Let x_{ijw}=wth observation in the jth group of stratum i, i=1,...,L, j=1,2,w=1,...,r. The balanced half-sample method can be used to estimate the variance of $\hat{\beta}$ when $\hat{\beta}$ is computed using any of the three estimates (2.2), (2.3), or (2.4).

3.1 Method 1 (deviations computed about overall means):

Let $\hat{\beta}_{(p)}$ be the pth half-sample estimate of β corresponding to the estimate defined in (2.2). That is,

$$\hat{\beta}_{(p)} = \frac{\sum_{i=1}^{L}\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{w=1}^{r}(x_{ijw}-\bar{x}_{...}^{(p)})(y_{ijw}-\bar{y}_{...}^{(p)})}{\sum_{i=1}^{L}\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{w=1}^{r}(x_{ijw}-\bar{x}_{...}^{(p)})^2}$$

where

$$\bar{x}_{...}^{(p)} = \frac{1}{Lr}\sum_{i=1}^{L}\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{w=1}^{r}x_{ijw} ,$$

$$\bar{y}_{...}^{(p)} = \frac{1}{Lr}\sum_{i=1}^{L}\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{w=1}^{r}y_{ijw} ,$$

and

$$\delta_{ij}^{(p)} = \begin{cases} 1 \text{ if the } j^{th} \text{ group of } i^{th} \text{ stratum is in the } p^{th} \text{ half-sample} \\ 0 \text{ if the } j^{th} \text{ group of } i^{th} \text{ stratum is not in the } p^{th} \text{ half-sample} \end{cases} \qquad (3.1)$$

Then, letting M = total number of half-samples computed,

$$\hat{V}_{B1}(\hat{\beta}_1) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_{(i)}-\bar{\beta})^2 , \quad \bar{\beta} = \frac{1}{M}\sum_{i=1}^{M}\hat{\beta}_{(i)}$$

and

$$\hat{V}_{B2}(\hat{\beta}_1) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_{(i)}-\hat{\beta}_1)^2 ,$$ $\hat{\beta}_1$ defined in (2.2).

3.2 Method 2 (deviations computed about within-stratum means):

Let $\hat{\beta}(p)$ be the pth half-sample estimate of β corresponding to the estimate defined in (2.3).

$$\hat{\beta}_{(p)} = \frac{\sum_{i=1}^{L}\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{w=1}^{r}(x_{ijw}-\bar{x}_{ij.})(y_{ijw}-\bar{y}_{ij.})}{\sum_{i=1}^{L}\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{w=1}^{r}(x_{ijw}-\bar{x}_{ij.})^2}$$

where

$$\bar{x}_{ij.} = \frac{1}{r}\sum_{w=1}^{r}x_{ijw}$$

$$\bar{y}_{ij.} = \frac{1}{r}\sum_{w=1}^{r}y_{ijw}$$

and $\delta_{ij}^{(p)}$ is defined as in (3.1).

Then,

$$\hat{V}_{B1}(\hat{\beta}_2) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_{(i)}-\bar{\beta})^2 , \quad \bar{\beta} = \frac{1}{M}\sum_{i=1}^{M}\hat{\beta}_{(i)}$$

and

$$\hat{V}_{B2}(\hat{\beta}_2) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_{(i)}-\hat{\beta}_2)^2 ,$$ $\hat{\beta}_2$ defined in (2.3).

3.3 Method 3 (average of the strata slopes):

Let $\hat{\beta}_{(p)}$ be the pth half-sample estimate of β corresponding to the estimate defined in (2.4).

$$\hat{\beta}_{(p)} = \frac{1}{L}\sum_{i=1}^{L}\left\{\frac{\sum_{j=1}^{2}\delta_{ij}^{(p)}\sum_{k=1}^{r}(x_{ijw}-\bar{x}_{ij.})(y_{ijw}-\bar{y}_{ij.})}{\sum_{i=1}^{2}\delta_{ij}^{(p)}\sum_{k=1}^{r}(x_{ijw}-\bar{x}_{ij.})^2}\right\},$$

where

$\bar{x}_{ij.}$, $\bar{y}_{ij.}$, $\delta_{ij}^{(p)}$ are defined as before.

Then,

$$\hat{V}_{B1}(\hat{\beta}_3) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_{(i)}-\bar{\beta})^2 , \quad \bar{\beta} = \frac{1}{M}\sum_{i=1}^{M}\hat{\beta}_{(i)}$$

and

$$\hat{V}_{B2}(\hat{\beta}_3) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_{(i)}-\hat{\beta}_3)^2, \quad \hat{\beta}_3 \text{ defined in}$$

(2.4).

4. The Sampling Experiment

The sampling experiment consists of randomly selecting n observations for each of L strata of infinite size whose parameters are precisely specified. On the j-th draw from the i-th stratum the random pair (x_{ij}, y_{ij}) is observed where

$$\begin{pmatrix} x_{ij} \\ y_{ij} \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_x^{(i)} \\ \mu_y^{(i)} \end{pmatrix}, \begin{pmatrix} \sigma_{xx}^{(i)} & \sigma_{xy}^{(i)} \\ \sigma_{xy}^{(i)} & \sigma_{yy}^{(i)} \end{pmatrix}\right].$$

Although in this experiment, the parameters are known, estimates of them are obtained. The process was then repeated K times and the distribution of the estimates of β were studied.

The strata correlations, $\rho^{(i)}$ are set equal to .9 for all strata and values are specified for $\mu_x^{(i)}$, $\sigma_{xx}^{(i)}$, $\beta^{(i)}=\sigma_{xy}^{(i)}/\sigma_{xx}^{(i)}$ and $\alpha^{(i)}=\mu_y^{(i)}-\beta(i)\mu_x^{(i)}$. By fixing these parameters, the values of $\mu_y^{(i)}$, $\sigma_{yy}^{(i)}$ and $\sigma_{xy}^{(i)}$ are determined.

A variety of "situations," covering a range of parameters, were considered. These can be summarized as follows:

Situation (i) L=3, n=20, K=1200:

	$(\mu_x^{(1)},\mu_x^{(2)},\mu_x^{(3)})$	$(\sigma_{xx}^{(1)},\sigma_{xx}^{(2)},\sigma_{xx}^{(3)})$	$(\beta^{(1)},\beta^{(2)},\beta^{(3)})$	$(\alpha^{(1)},\alpha^{(2)},\alpha^{(3)})$
Situation (i-1)	(5,5,5)	(1,1,1)	(1,1,1)	(0,0,0)
Situation (i-2)	(5,5,5)	(1,1,1)	(1,2,3)	(0,0,0)
Situation (i-3)	(5,5,5)	(1,1,1)	(1,1,1)	(0,1,2)
Situation (i-4)	(5,5,5)	(1,1,1)	(1,2,3)	(0,1,2)
Situation (i-5)	(5,5,5)	(1,2,3)	(1,1,1)	(0,0,0)
Situation (i-6)	(5,5,5)	(1,2,3)	(1,2,3)	(0,0,0)
Situation (i-7)	(5,5,5)	(1,2,3)	(1,1,1)	(0,1,2)
Situation (i-8)	(5,5,5)	(1,2,3)	(1,2,3)	(0,1,2)
Situation (i-9)	(5,10,15)	(3,6,9)	(1,1,1)	(0,0,0)
Situation (i-10)	(5,10,15)	(3,6,9)	(1,2,3)	(0,0,0)
Situation (i-11)	(5,10,15)	(3,6,9)	(1,1,1)	(0,1,2)
Situation (i-12)	(5,10,15)	(3,6,9)	(1,2,3)	(0,1,2)

Situation (ii) L=3, n=2, repeated for all 12 sets of parameters as in Situation (i)

Situation (iii) L=3, n=4, " " " " " " " " "

Situation (iv) L=3, n=8, " " " " " " " " "

Situation (v) L=3, n=12, " " " " " " " " "

Situation (vi) L=3, n=16, " " " " " " " " "

Situation (vii) L=3, n=100, " " " " " " " " "

Situations (viii-1)-(viii-12) correspond, for L=4 strata, to the situations described in Situations (i-1)-(i-12).

Situations (ix-1)-(ix-12) correspond, for L=15 strata, to the situations described in Situations (i-1)-(i-12).

The method used for generating the random pair (x_{ij}, y_{ij}) is not described in detail here. All normal deviates were independently generated by the method of Marsaglia (1973).

The validity of the sampling experiments were checked in a variety of ways. These are described by Lemeshow (1976). There was close agreement between theoretical and simulated results providing reassuring evidence that the simulations reported here operated correctly.

5. Results

To assess the relative advantages of the different estimates of β defined in (2.2), (2.3) and (2.4), it is necessary to compute the population value of β in each of the sampling situations under consideration. This value, in terms of the parameters fixed in each stratum, is as follows:

$$\beta = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\sum_{i=1}^{L}\sigma_{xy}^{(i)} + \sum_{i=1}^{L}(\mu_x^{(i)}-\mu_x)(\mu_y^{(i)}-\mu_y)}{\sum_{i=1}^{L}\sigma_{xx}^{(i)} + \sum_{i=1}^{L}(\mu_x^{(i)}-\mu_x)^2}. \quad (5.1)$$

The discussion here is restricted to the situations with L=3. Situations (i-1)-(i-4) correspond to having $\mu_x^{(i)}=\mu_x^*$ and $\sigma_{xx}^{(i)}=\sigma_{xx}^*$, $i=1,\ldots,L$. From Theorem 1 we expect all three estimates of β to be consistent and, since sampling is from bivariate normal populations, $\hat{\beta}_3$ should be the minimum variance unbiased estimate. Situations (i-5)-(i-8) have $\mu_x^{(i)}=\mu_x^*$, $i=1,\ldots,L$. From Theorem 1, only $\hat{\beta}_1$ and $\hat{\beta}_2$ will be consistent estimates of the population β. Situations (i-9)-(i-12) correspond to all strata having different means and variances. From Theorem 1, only $\hat{\beta}_1$ should be consistent for β.

Values of $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ were calculated for each of the twelve situations (i-1)-(i-12). The means, $\hat{E}(\hat{\beta}_i)$, and variances, $\hat{V}(\hat{\beta}_i)$, $i=1,2,3$, from the 1200 repetitions were computed. Table 1 presents these results along with the value of β computed using (5.1).

It is clear that these results agree with the conclusions of Theorem 1. In Situations (i-1)-(i-4), $\hat{E}(\hat{\beta}_i)\approx\beta$, $i=1,2,3$. That is, each of the $\hat{\beta}_i$ appears to be an unbiased estimate of β. In addition, except in the rather uninteresting Situation (i-1) in which all strata are identical, $\hat{\beta}_1$ has larger variances than $\hat{\beta}_2$ or $\hat{\beta}_3$. In Situations (i-2) and (i-4) where each stratum has a different slope, $\hat{\beta}_3$ has the minimum variance of the unbiased estimates. In Situations (i-5)-(i-8), $\hat{\beta}_1$ and $\hat{\beta}_2$ are consistent by Theorem 1 and, in the sampling experiment, $\hat{E}(\hat{\beta}_i)\approx\beta$, $i=1,2$. However, in Situations (i-6) and (i-8) in which the strata slopes are not all equal, $\hat{E}(\hat{\beta}_3)\neq\beta$. Note that $\hat{V}(\hat{\beta}_2)$ never exceeds $\hat{V}(\hat{\beta}_1)$. In Situations (i-9)-(i-12), $\hat{\beta}_1$ appears unbiased and has smaller variance than the others. $\hat{\beta}_2$ and $\hat{\beta}_3$ are biased whenever the strata have different linear regressions.

The conslusions of Theorem 1 and this sampling experiment is that if an estimate of the population β is desired, $\hat{\beta}_1$ is consistent and asymptotically unbiased. The variance of $\hat{\beta}_1$ is generally larger than the variances of the alternative estimates. If it can be assumed that the strata have the same mean for the independent variable, then, in the sampling experiment, $\hat{\beta}_2$

is a better estimate of β than $\hat{\beta}_1$ since it is also unbiased but much less variable. Use of $\hat{\beta}_3$ is not recommended since the necessary assumptions may be too restrictive.

Note that β is estimated by computing deviations about some mean. When this mean is a within-group mean as in $\hat{\beta}_2$ or $\hat{\beta}_3$, at least two observations are needed in each of the 2 groups which were established for use with the balanced half-sample method. Moreover, it is of interest to determine the minimum number of observations per stratum necessary to introduce some degree of stability into the variance estimation calculations.

As described earlier, the sampling experiment was repeated for the L=3 strata situations, with n=4, 8, 12, 16, 20 or 100 observations per stratum. By taking at least n=4 observations per stratum, we are assured of having at least 2 observations in each of the established groups.

Table 2 presented the absolute relative bias and variance of the two estimates of $V(\hat{\beta}_i)$, i=1,2,3, in Situation (i-2), (iii-2), (iv-2), (v-2), (vi-2) and (vii-2). That is, we present the results for those situations in which the means and variances of the independent variable are the same for each of the L=3 strata but, while the intercepts are the same, the slopes differ for the linear regressions in each stratum.

In Table 2 we see that selecting small samples from each stratum can result in estimates of extremely low precision and high variability. For instance, when $\hat{\beta}_1$ was used to estimate β with n=4 observations per stratum, the balanced half-sample estimate of $\hat{V}(\hat{\beta}_1)$ missed this target variance by 93%. The variance of these estimates were quite high. Using an absolute relative bias of .05 as an acceptable level of precision, we observe that, when using $\hat{\beta}_1$ to estimate β, more than 100 observations per stratum were needed. When using $\hat{\beta}_2$ as an estimate of β, at least 20 observations should be used. About 100 observations should be used for the variance estimates when $\hat{\beta}_3$ is used to estimate β. As a general rule, at least n=20 observations from each stratum are needed in order to introduce stability into the variance estimates. When using $\hat{\beta}_3$, variance estimates show greater sensitivity to small n, and a larger n should be selected if possible. All twelve parameterizations were studied in the same way with similar results.

A study of the other situations indicates that contrary to the results of the linear case (Lemeshow and Epp (1977)) and the combined ratio estimates (Lemeshow and Levy (1977)), the two half-sample estimates of the variance of $\hat{\beta}_i$ differed. In the "balanced" situation (i.e., L=3), neither one of variance estimates was consistantly better or worse than the other over all situations considered.

In the "full-matrix" situation (i.e., L=4) it was noted that $\hat{E}[\hat{V}_{B1}(\hat{\beta}_i)]<\hat{E}[\hat{V}_{B2}(\hat{\beta}_i)]$, i=1,2,3. In fact, $\hat{V}_{B1}(\hat{\beta}_i)$ always had a negative bias. However, the variance and mean square error using this balanced half-sample estimate were never greater that the corresponding measurement for $\hat{V}_{B2}(\hat{\beta}_i)$. This corresponds to results for linear and combined ratio estimates presented in the references cited above. When L=15 strata were used, the results described above for L=3 appear to apply.

6. Conclusions

The sampling experiments have demonstrated that each of the variance estimation techniques appear to have the potential of providing usable estimates of the target variance for the slope provided a large enough sample is selected from each stratum. This is very different than previous results for using the balanced half-sample method in the linear case or with the combined ratio estimate. There, even as few as two observations per stratum would result in minimal bias. Here, however, at least 20 observations per stratum are necessary in order to introduce some stability into the variance estimation process for the slope of the linear regression.

Acknowledgements

This research is based, in part, on the author's Ph.D dissertation, which was undertaken at UCLA. Computing assistance was obtained from the Health Sciences Computing Facility (UCLA) under NIH Special Resources Grant RR-3.

References

Lemeshow, S., An Evaluation of the Performance of the Balanced Half-Sample and Jackknife Variance Estimation Techniques. Unpublished doctoral dissertation, University of California, Los Angeles (1976).

Lemeshow, S. and Epp, R., "Properties of the Balanced Half-Sample and Jackknife Variance Estimation Techniques in the Linear Case." To be published in Communications in Statistics A, Vol. 6, Issue 13 (1977).

Lemeshow, S. and Levy, P. "Estimating the Variance of Ratio Estimates in Complex Sample Surveys with Two Primary Sampling Units Per Stratum - A Comparison of Balanced Replication and Jackknife Techniques." Submitted for publication.

Marsaglia, G., Ananthanarayanan, K. and Paul, N., How to Use the McGill Random Number Package 'SUPER-DUPER,' four pages typed description, School of Computer Sciences, McGill University, Montreal, Quebec (1973).

McCarthy, P.J., Replication. An Approach to the Analysis of Data from Complex Surveys. Vital and Health Statistics. NCHS, Series 2, No. 14 (1966).

McCarthy, P.J., Pseudoreplication: Half Samples. Review of the International Statistical Institute, Vol. 37, No. 3: 239-264 (1969).

______________, Pseudoreplication: Further Evaluation and Application of the Balanced Half-Sample Technique. Vital and Health Statistics. Series 2, No. 31 (1969).

Table 1: Population β for each of the twelve situations with L=3 strata described in means, $\tilde{E}(\beta_i)$, and variances, $\tilde{V}(\hat{\beta}_i)$, of $\hat{\beta}_i$, i=1,2,3, as estimated from the sampling experiment are presented.

		Means			Variances		
Situation	Pop β	$\tilde{E}(\hat{\beta}_1)$	$\tilde{E}(\hat{\beta}_2)$	$\tilde{E}(\hat{\beta}_3)$	$\tilde{V}(\hat{\beta}_1)$	$\tilde{V}(\hat{\beta}_2)$	$\tilde{V}(\hat{\beta}_3)$
(i- 1)	1.00000	1.00078	1.00034	1.00027	.00410	.00425	.00459
(i- 2)	2.00000	2.00553	2.00030	2.00095	.33297	.04640	.02094
(i- 3)	1.00000	1.00092	.99994	1.00073	.01523	.00417	.00453
(i- 4)	2.00000	2.01199	1.99562	2.00025	.50021	.04376	.02234
(i- 5)	1.00000	1.00016	.99926	1.00062	.00457	.00476	.00424
(i- 6)	2.33333	2.34308	2.32874	1.99972	.18958	.05090	.02015
(i- 7)	1.00000	.99703	.99811	.99680	.01089	.00503	.00477
(i- 8)	2.33333	2.30799	2.31779	1.99410	.27262	.05287	.02220
(i- 9)	1.00000	.99936	1.00159	1.00132	.00110	.00479	.00430
(i-10)	3.55882	3.56746	2.32645	2.00300	.01732	.05320	.02076
(i-11)	1.14706	1.14880	1.00009	1.00123	.00122	.00490	.00447
(i-12)	3.70588	3.72134	2.32342	2.00092	.01911	.04999	.02025

Table 2: Results of sampling experiment in which n=4, 8, 12, 16, 20 or 100 observations were selected from each of L=3 strata. Estimated values based on the sampling experiment are presented for absolute relative bias, and variance using the three methods of estimation. In all cases, $\mu_x^{(i)}=5$, $\sigma_{xx}^{(i)}=1$, $\alpha^{(i)}=0$, $\rho^{(i)}=.9$, $i=1,2,3$. $\beta^{(1)}=1$, $\beta^{(2)}=2$, $\beta^{(3)}=3$.

Method 1

	Absolute Relative Bias (I)		Variance (I)	
n	B1	B2	B1	B2
4	.93	1.11	35.04468	51.53947
8	.34	.38	1.85307	1.99540
12	.24	.27	.54089	.58215
16	.13	.15	.26443	.27707
20	.10	.12	.14486	.14973
100	.07	.08	.00460	.00462

Method 2

	Absolute Relative Bias (I)		Variance (I)	
n	B1	B2	B1	B2
4	1.99	3.15	3.09580	5.48564
8	.20	.42	.02398	.02919
12	.16	.31	.00537	.00626
16	.01	.11	.00308	.00349
20	.01	.06	.00185	.00197
100	.02	.04	.00005	.00005

Method 3

	Absolute Relative Bias (I)		Variance (I)	
n	B1	B2	B1	B2
4	*	*	**	**
8	1.53	2.87	.28248	.84464
12	.49	.92	.00540	.00979
16	.23	.51	.00201	.00323
20	.25	.49	.00083	.00123
100	.03	.06	.00001	.00001

*Absolute relative bias >100

**Variance >100