A SIMULATION MODEL FOR THE COMPARISON OF SAMPLING STRATEGIES USED IN ESTIMATING TOTAL RESIDENTIAL MARKET VALUE FOR A GEOGRAPHIC AREA

James M. Kraushaar

Department of Accounting and Quantitative Studies
California State University, Fresno

INTRODUCTION AND STATEMENT OF PROBLEM

The total property market value for a taxing district is used to: distribute federal, state, and county taxes; apportion joint school district tax levies; limit taxing and borrowing powers. This total market value cannot be calculated by simple summing the individual property values as listed on the local assessment roll because assed values may or may not reflect current market values, depending on many factors related to the assessment process. Consequently, total district assessed value cannot be a basis for tax allocation.

Appraising the market value for each individual property in the district would be too costly and time consuming because of the large number of properties in a district. Instead, a random sample of properties is valued by expert appraisers and when augmented with recent sales information, can be used to estimate the total market value of the district. Since conscientious property appraisal is expensive, only a small sample of individual properties is selected for appraisal. This means that the estimate of total market value is subject to large sampling error.

Stratified random sampling and ratio estimators are often used to reduce the large sampling error. However, measuring the precision (i.e., bias and variance) of these ratio estimators is difficult when samples are small because the second product moment equation may be unreliable.(1) Without a measure of precision, the effectiveness of various sampling design factors (e.g., number of strata, stratum boundary rule, and type of ratio estimator) can not be evaluated.

METHODOLOGY

A simulation model composed of three separate modules was developed to measure the precision of the total market value estimators for a population of properties.

MODEL OF RESIDENTIAL ASSESSED AND MARKET VALUES

The assessed value, A_k, of the k-th property in the population is randomly generated from a triangular distribution:

$$A_k = TRI(a,b,c)$$

where TRI refers to the triangular distribution and a, b, and c to the lower limit, mode, and upper limit parameters respectively.

The mean ratio of assessed to market value, $\bar{r}_a$, for a given level of assessed value is assumed to be linearly related to assessed value:

$$\bar{r}_a = B_o + B_1 A_k$$

where B_o and B_1 are parameters that describe the intercept and slope of the linear relation. The individual ratios of assessed to market value, r_k, are assumed to be normally distributed around the conditional mean, $\bar{r}_a$ with a constant standard deviation and are randomly generated from this distribution. Therefore, the ratio for the k-th property in the population is:

$$r_k = \bar{r}_a + E_k$$

where E_k is a random disturbance term generated from a normal distribution with mean zero and standard deviation, σ, as:

$$E_k = N(o,\sigma)$$

For each property in the population, the market value is determined by its randomly generated ratio of assessed to market value and its randomly generated assessed value:

$$M_k = \left[A_k / r_k \right]$$

By repeating the above procedure once for each property, a population of assessed and market values can be simulated.(2)

Two populations of residential property values were generated using this residential model and parameters which were based upon sample data supplied by the New York State Board of Equalization and Assessment and the Westchester County Tax Commission. They were not generated to represent the residential properties of any specific town or village. Instead, they were generated to represent several general characteristics common to many towns and villages.

One population with B_1 = .000005 and B_0 = .3 is representative of districts whose ratios of assessed to market value tend to increase with an increase in an assessed value.(3) The other population has an intercept of .3 and -.000005 change in ratio per dollar increase in assessed value and represents districts where individual ratios of assessed to market value decrease as assessed value increases.(4)

The other parameters (e.g., distribution shape and error term) of the residential model were held constant at levels representative of real world data. A triangular distribution with a lower limit of $2,000, mode of $9,000 and an upper limit of $50,000 provided a reasonable hypothetical residential population of assessed values. A Normal distribution with a mean of zero and standard deviation of .035 was used for the ratio error terms. This implies that the relative variation as measured by the coefficient of variation will decrease as assessed value increases since the standard deviation is constant.(5) For simplicity, a constant standard deviation of ratios was assumed in this study.

A MODEL FOR STRATIFYING, SAMPLING, AND ESTIMATING

Module two of the model can be subdivided into three subsections. The first subsection stratifies, according to assessed value, the hypothetical population of residential properties generated in module one. Two rules have been used for stratification; equal number of properties per stratum and equal assessed value per stratum. The second subsection selects a random sample of properties from the stratified population using the equal allocation method. With this random sample of property values, the third subsection calculates estimates of total residential market value using the separate ratio, combined ratio, and unbiased mean expansion estimators.

A MODEL FOR ESTIMATING THE SAMPLING ERROR OF THE RATIO ESTIMATOR

The usual second product moment approximation to the variance of the ratio estimator generally provides a relatively precise estimate of true sampling error if the sample size is reasonably large and the population has a bivariate normal distribution. However if the sample size is small as it is when estimating total residential property value for a town or village, the second product moment approximation to the variance may not be an accurate estimate of true sampling error for at least two reasons. First, the bias of the ratio estimator may not be negligible relative to the standard error.(6) Therefore, the mean square error rather than the variance is the appropriate measure of precision. Second, the usual second product moment approximation to the mean square error may not be close to the true precision since product moments beyond second order may not be negligible.(7)

More precise estimates of sampling error may be provided by including product moments beyond second order. However, these equations become increasingly complex as more terms in the Taylor series are included.(8) Also, it is difficult to know when enough terms in the series have been added so that the precision of the estimate of sampling error is acceptable.

MONTE CARLO SAMPLING METHOD

The procedure used in this study to estimate the true sampling error of the ratio estimator is dissimiliar to the higher order product moments method. This method employs a simulated sampling distribution for the ratio estimator by empirical sampling. Small random samples are repeatedly selected from a population of residential properties stratified according to assessed value. For each sample, an estimate of total residential market value is calculated. The variability in these estimates provides a simple estimate of true sampling error without resorting to the complex estimators which have product moments beyond second order.

The mean square error for an estimator of total residential market value is estimated by randomly selecting n_r samples of size n from all possible samples and using the formula:

$$\hat{MSE}_{\hat{T}g} = \sum_{j=1}^{n_r} (T_{jg} - T)^2 / n_r$$

where T_{jg} is the estimate of total residential market value from the j-th sample and for the g-th estimator type (g = 1, 2 or 3; unbiased mean expansion, combined ratio, separate ratio). T is the known total market value for the simulated residential population:

$$T = \sum_{k-1}^{N} m_k$$

where m_k is the market value for the k-th residential property and N is the total number of simulated residential properties.

The estimated variance, $\hat{V}(\hat{T}_g)$, for an estimator of total residential market value based on n_r estimates of total value is:

$$\hat{V}(\hat{T}_g) = \sum_{j=1}^{n_r} (\hat{T}_{jg} - \hat{\bar{T}}_g)^2 / (n_r - 1)$$

where the mean of the n_r estimates of total value for the g-th estimator type is:

$$\hat{\bar{T}}_g = \sum_{j=1}^{n_r} \hat{T}_{jg} / n_r$$

The bias exhibited by the two ratio estimators of total residential value can be estimated by:

$$\hat{BIAS}_{\hat{T}_g} = \hat{\bar{T}}_g - T$$

where g = 1, 2 or 3 and as defined above, $\hat{\bar{T}}_g$ is the mean of the n_r estimates of total residential market value and T is the known total residential value.

The technique used in this study to estimate the sampling error of ratio estimators in the small sample case has been widely used by other researchers,(9) but it has not previously been applied to sampling methods that involve both stratification and ratio estimation. Extending the technique seems worthwhile because stratification and ratio estimation are often used together when sampling is expensive and sample size is subsequently small.

The simulation model of all three modules was programmed in Fortran IV. Program debugging was performed on an IBM 370/165 with a WATFIV compiler. The actual experimental runs were made on the same IBM 370/165 with the IBM Fortran G Level compiler.

The basis for the random number generators used in this study is a uniform random generator described by Hutchison.(10) Lewis, Goodman, and Miller (11) give a modification of the generator for IBM's 360 and 370 machines where there are 31 bits available for computation in the 32 bit general registers, one bit being the sign bit. The sequence of random numbers, U_k, is generated by the equation:

$$U_{k+1} = AU_k \pmod{p}$$

where p is the prime number $(2^{31} - 1)$ and A (16807) is a positive root of p. Various seeds, U_0, (481694049, 864268549, 13034519) were used with this formula and the resulting random number sequences were extensively tested by the International Mathematical and Statistical Libraries, Inc.(12) The subroutine used in this study to generate uniform random numbers between 0 and 1 is called MRANDU and it is an extension of the Lewis, Goodman, and Miller IBM 360/370 generator often called RANDU.(13)

The finite population of assessed property values is randomly generated in this study from a triangular distribution by the inverse transformation method. The normally distributed error term used in the generation of market values from assessed values is the adjusted sum of 12 uniform random numbers.

Preliminary tests of the computer program were used to examine the simulator costs, fluctuation in experimental response, and validity of the simulation program code. Calculational checks against several computer runs indicated that the computer program exactly followed the model represented earlier. Costs and experimental response variation determined that there should be approximately 800 random samples for each response.

EXPERIMENTAL DESIGN

An experiment conducted to ascertain the effect on mean square error of four factors: number of strata, stratum boundary rule, estimator type and distribution of assessed to market value ratios. A full factorial design was used to examine these four factors at the levels listed below:

<u>Factor A</u>: Number of Strata (six levels)

A-1. No stratification. (one strata)
A-2. Two strata.
A-3. Three strata.
A-4. Four strata.
A-5. Six strata.
A-6. Twelve strata.

<u>Factor B</u>: Stratum Boundary Rule. (two levels)

B-1. Equal total assessed value per strata. (EA)

B-2. Equal number of properties (parcels) per strata. (EP)

<u>Factor C</u>: Estimator. (three levels)

C-1. Simple mean expansion. (SME)
C-2. Combined ratio. (CR)
C-3. Separate ratio. (SR)

<u>Factor D</u>: Distribution of assessed to market value ratios. (two levels)

D-1. Model I: $R_i = .3 + .000005A_i + E_i$.

D-2. Model II: $R_i = .8 - .000005A_i + E_i$.

The factorial design model indicating the main effect of each factor and the interaction effect between factor levels, is expressed as follows:

$$Y_{abcdt} = \mu + A_a + B_b + C_c + D_d + (AB)_{ab} + (AC)_{ac} + (AD)_{ad} + (BC)_{bc} + (BD)_{bd} + (CD)_{cd} + (ABC)_{abc} + (ABD)_{abd} + (ACD)_{acd} + (BCD)_{bcd} + (ABCD)_{abcd} + E_{abcdt}$$

where Y_{abcdt} is the observed model response for each replicated simulation trial, t. Each trial is composed of 800 random samples generated at one set of factor levels.

μ is the overall mean.

A_a, B_b, C_c, D_d are the main effects of factors A, B, C, and D respectively with subscripts indicating the number of levels for each factor.

The first order interaction effects are $(AC)_{ac}$, $(AD)_{ad}$, $(BC)_{bc}$, $(BD)_{bd}$, $(CD)_{cd}$. For example, $(AB)_{ab}$ is the interaction effect between factors A and B. This measures the lack of independence in the effects of the two factors. If the interaction effect is significantly different from zero, the combination of A and B has some effect over and above the individual effects of A and B separately. With significant interactions, statistical conclusions regarding the main effects of A and B become confounded. Therefore, any conclusions regarding main effects A and B become hazardous, since the averages of specific treatment means may not be meaningful measures. (14)

$(ABC)_{abc}$, $(ABD)_{abd}$, $(ACD)_{acd}$, $(BCD)_{bcd}$ are the second order interactions. They have interpretations similar to $(AB)_{ab}$ but for combinations of three factors.

$(ABCD)_{abcd}$ is the third order interaction. It has an interpretation similar to $(AB)_{ab}$ but for combinations of four factor levels.

E_{abcdt} represents the experimental error for the a-th, b-th, c-th and d-th level of the four factors and the t-th simulation run.

A complete factorial experiment was carried out with two replications in order to estimate main effects and all the interaction effects. This experiment required 48 runs with each combination of factor levels being measured twice. The three types of estimators are calculated for each run rather than in separate runs. The simulator program was written this way because the trade off between time saved and increased space required to store the estimators seems to favor this approach.

RESULTS AND CONCLUSIONS

The results of the experiment indicated non-normal residuals, non-homogeneous residual variances, and widely varying mean square errors. The ANOVA analysis of the response variable may not be appropriate. The size of the residuals indicate the basic problem. The small difference between responses for each replication may be causing all the higher level interactions to be significant.

Each replication results from a different group of random samples taken from the population, but there is little variability in the response between replications. This may be caused by several factors. First, all three estimates of total value are calculated from the same 800 random samples. Therefore, these responses are not independent of each other. Second, a unique series of random numbers is used for each random variable in the model. Therefore, when a replication of a given set of experimental conditions is obtained, the only change in the simulator is the 800 random samples drawn. The sequence of random numbers remains unchanged for all other random variables in the model. This maximizes the positive correlation between simulation runs and reduces experimental error. Third, the estimated variance obtained from the repeated random estimates of total residential market value is very precise because the sample size is large. Each estimated variance is calculated from 800 estimates of total value.

ANOVA analysis of the response data is limited because of the small and systematically variable replications error. Therefore, descriptive rather than inferential analysis will be used to interpret the results of this study. Several other similar studies have made this same assumption.(15)

One of the main objectives of this study was to compare small sample experimental results with large sample theory. These comparisons indicate no major differences between large sample theory and small sample results for the populations examined in this study.

Cochran has showed that the variance of the unbiased stratified mean estimator decreases inversely as the square of the number of strata, dividing a rectangular distribution, increases. Even for skewed distributions with optimum boundaries and Neyman allocation this relationship continued to hold approximately. Cochran also suggests that when some other variable, for example assessed value, is used to construct the strata, a point is reached where only a trivial proportional reduction in the variance of the mean expansion estimator if the number of strata are further increased. He suggests that six strata is a practical maximum unless the correlation between the variable being estimated and the variable used for stratification exceeds .95.(16)

The results of this study indicate that the relationship is inversely proportional to the number of strata for the stratified small sample ratio estimator with few strata. However, as the number of strata increases there is a tendency for the rate of decrese to become slightly less than the reciprocal of the number of strata. The results of this study suggest that for the conditions examined, the rate of decrease in estimator variance is much less than $1/L^2$. The results also indicate that Cochran's rule of thumb also roughly holds for stratified small sample ratio estimators. In populations like model I and II where even though the correlation coefficients are very high and the relationships between market and assessed value are almost linear, there is little to be gained with more than a few strata. In fact, generally two strata provide the main reduction in variance. Table 1 indicates that all but one treatment combination had its highest percentage decrease with only two strata. Ten of the 12 treatment combinations had over 50% of the total decrease in variance occur with only two strata. All 12 combinations had at least 80% of the total reduction in variance occur with six strata. It seems Cochran's rule of thumb for number of strata may apply to stratified small sample ratio estimation as well as stratified unbiased mean estimation.

Large sample theory indicates the separate ratio will have a lower variance than the combined ratio for populations where the ratio estimator should be used. However, this difference will be small unless the stratum ratios of market to assessed value are very different. The small sample results of this study support the large sample theory. For both populations I and II the separate ratio had a slightly lower variance than the combined ratio. There were moderate differences in stratum ratios of market to assessed value, but obviously for the populations examined in this study, they must be larger for the separate ratio to have any substantial advantage over the combined ratio when sample sizes are small.

Large sample theory also suggests that the separate ratio estimator may be subject to larger bias than the combined ratio especially with numerous strata. The results of this study with small sizes concur with the large sample theory since the separate ratio did have slightly larger biases than the combined ratio. However, there was no trend for the bias of the separate ratio to increase as the number of strata increased. This is probably because the two population models used in this study led to practically unbiased ratio estimators. Small sample conclusions or recommendations concerning populations that lead to large biases cannot be made from this study.

SAMPLING STRATEGIES

Another major objective of this study was to determine the best combination, as measured by precision per unit cost, of sample design factor levels for residential populations similar to populations I and II. The experimental results indicate two general strategies could be used to estimate total residential market value with a standard error of about 3 to 4% of the true total value (see Table 2). A third strategy could obtain more precision than either of the first two strategies. Cost considerations determine which of these three strategies would be most appropriate. A summary of the sampling technique recommendations suggested by the results of this study is presented in Table 3.

The mean expansion estimator could be used with stratification for the first strategy. This would require that assessed values be recorded for all properties as is currently done. However, the assessed values would not be needed for the calculation of the estimate of total value. This may result in a cost savings over the ratio estimator which requires both assessed and market values for calculation of the estimate of total value.

For population model I six strata are enough to get about the same precision as the ratio estimator without stratification. However, with model II it takes 12 strata to achieve the same precision for the ratio estimator without stratification.

Generally, if this first strategy is selected based on cost considerations, the equal number of properties boundary rule should be used for populations similar to population model I. However, for model II type populations the equal assessed value per stratum rule should be used.

Second, the ratio estimator could be used to estimate total residential market value without using stratification. This technique would require only the proper association of market and assessed values for the properties in the sample. There would be no stratification costs involved. This technique seems to be especially effective for population II. Also, without stratification the cumulative bias problem which sometimes affects the separate ratio estimator when stratified sampling is used would be of no concern. Since the bias has also been shown to be negligible when compared to the standard error for population models I and II without stratification, the ratio estimator can be used without fear of bias. This second strategy obviously requires no boundary rule.

A third strategy could be used to obtain more precision than either of the first two strategies. This strategy involves the use of both stratification and ratio estimation. Therefore, it would have stratification costs as well as relatively high estimation costs, but it has somewhat better precision than either strategy one or two.

Two options of this third strategy depend on how the cost due to stratification behaves when the number of strata are increased. Option A is for situations where the cost for two strata is about the same as the cost for 12 strata or at least the cost of stratification goes up only slightly as the number of strata increase. Option B is for situations where the cost of stratification increases at a constant or at least not rapidly declining rate as the number of strata increase.

For populations with characteristics similar to model I the combined ratio with six strata and equal assessed value per stratum should be used. Actually, 12 strata would give a slightly better estimate of total residential market value than six strata. However, since this would result in only one observation per stratum (i.e., sample size is 12 and 12 strata), a reliable estimate of the precision cannot be obtained from the sample since the number of strata is less than 20.(17)

For populations similar to model II the separate ratio with six strata and equal assessed value per stratum should be used. It should also be noted, however, that for population models which indicate possible ratio bias the combined ratio should probably be used for both model I and II since there is little practical difference between the two estimators. In such populations mean square error not variance is the appropriate measure of precision. Since the combined ratio is generally less subject to cumulative stratum bias than the separate ratio, it may give a lower mean square error.

Clearly if there is much increase in stratification cost for two through 12 strata, then the small increase in precision obtained by more strata will not be worthwhile. For populations similar to model I, the combined ratio estimator with two strata and the equal assessed value boundary rule should be used, since after two strata the increase in precision is negligible considering the costs of having more strata.

For populations similar to model II the separate ratio estimator should be used with two strata and equal assessed value per strata. Again the increased precision obtained by increasing the number of strata beyond two is negligible, if there is a large cost increase. It should be noted also that the combined ratio provides virtually the same precision as the separate ratio with equal assessed value and should be used if there is any possibility of bias since it is less susceptible to the bias influence than the separate ratio.

TABLE 1

Decrease in Estimated Standard Error as a Percentage of Total Decrease for Increases in the Number of Strata from 1 to 12

Estimator	Boundary Rule	Change in Number of Strata 1 - 2	2 - 3	3 - 4	4 - 6	6 - 12
		Model I				
SME	EA	43.7	13.6	13.6	7.2	21.9
SME	EP	56.5	19.8	7.1	9.2	7.4
CR	EA	67.5	20.4	-5.1	8,1	9.1
CR	EP	58.9	11.9	18.2	8.2	3.1
SR	EA	54.1	27.9	-2.0	8.6	11.4
SR	EP	70.5	51.9	19.3	3.4	1.6
		Model II				
SME	EA	51.9	15.4	12.2	6.4	14.7
SME	EP	48.0	20.3	7.0	9.2	14.4
CR	EA	64.3	11.1	12.9	7.9	3.7
CR	EP	23.5	30.6	7.6	13.3	24.9
SR	EA	68.1	9.1	13.2	6.7	28.9
SR	EP	35.2	33.0	3.8	11.6	16.3

TABLE 2

Estimated Standard Error as a Percentage of Total Residential Market Value

Estimator	Boundary Rule	Number of Strata: 1	2	3	4	6	12
		Model I					
SME	EA	11.3	7.6	6.5	5.3	4.7	2.9
SME	EP	11.3	6.4	4.7	4.1	3.3	2.7
CR	EA	4.5	3.1	2.6	2.7	2.6	2.4
CR	EP	4.5	3.4	3.1	2.8	2.6	2.6
SR	EA	4.5	3.4	2.8	2.8	2.7	2.4
SR	EP	4.5	3.1	3.0	2.6	2.6	2.5
		Model II					
SME	EA	16.7	9.2	7.0	5.2	4.3	2.3
SME	EP	16.7	9.9	7.1	6.0	4.7	2.7
CR	EA	3.2	2.0	1.9	1.7	1.5	1.4
CR	EP	3.2	2.9	2.4	2.3	2.1	1.7
SR	EA	3.2	2.0	1.8	1.6	1.5	1.4
SR	EP	3.2	2.7	2.2	2.2	2.0	1.7

TABLE 3

Summary of Sampling Recommendations

Strategy	Stratification Costs	Relative Estimation Costs	Recommended Boundary Rule and Number of Strata: Model I	Model II
I. Mean Expansion Estimator With Stratification	Yes	Low	Equal parcels per stratum with six strata	Equal assessed value per stratum with six strata
II. Ratio Estimator Without Stratification	No	High	None	None
III. Ratio Estimator With Stratification				
<u>Option A</u>	Yes [Increasing as the number of strata increase]	High	Equal assessed value per	Equal assessed value per
<u>Option B</u>	Yes [Constant for 2-12 strata]	High	Equal assessed value per stratum with 6 strata and the combined ratio estimator	Equal assessed value per stratum with 6 stata and the combined or separate ratio estimator

REFERENCES

1. Chai, J. and Chottiner, S., "Estimating Mean Square Error of Estimator of the Property Class Ratio of Assessed to Market Values From a Small Sample--An Empirical Study," paper presented at the 1975 AIDS meeting, Las Vegas March, 1975.

2. The idea of generating market value from assessed value and the equalization ratio was used in the report: "New York City: Economic Base and Fiscal Capacity," prepared for the State Study Commission for New York City under the direction of the Metropolitan and Regional Research Center of the Maxwell School and Syracuse University Research Corporation, I-71 - I-75.

3. For example see residential data from Table 4, "List of Residential Appraisal Parcels in 1973," Westchester County Tax Commission, Westchester County, New York.

4. Ross, M. H., "The Property Tax Assessment Review Process: A Case for Regressive Property Taxation?" National Tax Journal, 24, (1971), pp. 37-43; Greytak, D. and Bjornstad, D., "The Assessment of Property: Variation by Location, Property Type and Value," Occasion Paper No. 11, Maxwell School of Citizenship and Public Affairs, Syracuse University, p. 21

5. A recent study indicates that the coefficient of variation was greatest at lower and upper levels of assessed values but higher in the latter. However, they were concerned that interaction between property type, location, and value may have caused this relationship to be more artificial than real. Greytak and Bjornstad, p. 21.

6. However, Chai and Chottiner have indicated for several towns in Westchester County, New York, that the bias is relatively small for samples of twenty-two to thirty-seven residential properties. Chai, J. and Chottiner, S., "Estimating Mean Square Error of the Estimator of the Property Class Ratio of Assessed to Market Values from a Small Sample - An Empirical Study," Proceedings of AIDS National M Meeting, Las Vegas, Nevada, March 20, 1974, p. 1-4.

7. Lauh, E., and W. Williams, "Some Small Sample Results for the Variance of a Ratio," Proceedings of American Statistical Association, Social Statistics Section, 1963, p. 277.

8. Sukhatme, P. V. and Sukhatme, B. V., "Sampling Theory of Surveys with Applications, Iowa State University Press, 1970, p. 141.

9. See for example, Hutchinson, "A Monte Carlo Comparison of Some Ratio Estimators," Biometrika 58, 1971, pp.313-321; or Rao, J. and Beegle, L., "A Monte Carlo Study of Some Ratio Estimators," Sankhya B 29, 1967, pp. 47-57.

10. Hutchinson, D. W., "A New Uniform Pseudo-Random Number Generator," Communications of the ACM 9, 1966, No. 6, pp. 432-3

11. Lewis, P., Goodman, A. and Miller, J., "A Pseudo-Random Number Generator for the System/360," IBM Systems Journal 2, pp. 136-146.

12. International Mathematical and Statistical Libraries IMC., Library 1, "Generation and Testing of Random Numbers; Goodness of Fit," 3rd ed., 1973, pp. 61-67.

13. Mize, J., "Multiple Sequence Random Number Generators," Winter Simulation Conference Proceedings, 1973, pp.67-75.

14. Neter, J. and Wasserman, W., Applied Linear Statistical Models, Richard Irwin, Inc., Homewood IL, 1974, p. 563.

15. See for example, Lauh and Williams, pp. 273-283; Hutchinson, M. C., "A Monte Carlo Comparison of Some Ratio Estimators," Biometrik 58, 1971, 313-321; and Rao, J., and Beegle, L., "A Monte Carlo Study of Some Ratio Estimators," Sankhya B 29, 1967, pp. 47-57.

16. Cochran, W. G., Sampling Techniques, John Wiley and Sons, 1963, pp. 133-135.

17. The technique used to estimate the variance with only one sample observation per stratum is called the method of collapsed strata. See Cochran, Sampling Techniques, pp. 141-142.