THE EFFECT OF FISHER'S z-TRANSFORMATION ON CONFIDENCE INTERVALS FOR THE CORRELATION COEFFICIENT

Mary H. Mulry and Kirk M. Wolter
U.S. Bureau of the Census

1. Introduction

This paper describes a Monte Carlo study of three estimators of variance for the estimated correlation coefficient in finite population sampling. The variance estimators are the random group, jackknife and first order Taylor series estimators. The comparison of the estimators is based on two criteria: 1) the properties of the variance estimator itself, including its bias, variance and mean square error (MSE), and 2) the properties of confidence intervals formed using the variance estimator. Our primary objective is to investigate the effect of Fisher's z-transformation on the properties of the variance estimators and on confidence intervals formed with them.

Previous authors who have presented empirical results about variance estimators in finite population sampling include Frankel (1971), Mellor (1973), Bean (1975), and Campbell and Meyer (1978). Theoretical properties of the variance estimators have been discussed by Rao and Krewski (1978, 1979). In all these studies no one estimator has emerged superior overall. The choice of a variance estimator seems to depend on the parameter to be estimated, its estimator, the sampling design, and the population at hand. Nevertheless, it has been established that variance estimators such as those studied in this paper have reasonably small bias and can be used for making inferential statements. There has been little previous study of the effects of transformations on variance estimation in finite population sampling.

The random group, jackknife and Taylor series variance estimators are defined in Section 2. The data used in the Monte Carlo study are discussed in Section 3, and Section 4 contains the results of our investigation.

2. The Estimators

Throughout this paper we assume that a simple random sample of size n is selected without replacement from a finite population of size N. We assume that a bivariate characteristic is attached to each unit in the population, where (X_i, Y_i) denotes the value of the i-th unit.

The finite population correlation coefficient is

$$\rho = \frac{\sum_i^N (X_i-\bar{X})(Y_i-\bar{Y})}{\{(\sum_i^N (X_i-\bar{X})^2\}^{\frac{1}{2}}\{\sum_i^N (Y_i-\bar{Y})^2\}^{\frac{1}{2}}} .$$

The usual estimator of ρ, say $\hat{\rho}$, and the random group, jackknife, and Taylor series estimators of $\mathrm{Var}\{\hat{\rho}\}$ are given by

$$\hat{\rho} = \frac{\sum_i^n (x_i-\bar{x})(y_i-\bar{y})}{\{\sum_i^n (x_i-\bar{x})^2\}^{\frac{1}{2}}\{\sum_i^n (y_i-\bar{y})\}^2} ,$$

$$v_{rg}(\hat{\rho}) = \frac{1}{k(k-1)} \sum_\alpha^k (\hat{\rho}_\alpha-\hat{\rho})^2 ,$$

$$v_j(\hat{\rho}) = \frac{1}{k(k-1)} \sum_\alpha^k (\hat{\rho}_\alpha-\hat{\rho})^2 ,$$

and

$$v_{ts}(\hat{\rho}) = \frac{1}{n(n-1)} \sum_i^k \hat{r}_i^2 ,$$

respectively.

For the random group estimator, the sample is divided at random into k groups of size m (we assume n = mk), and $\hat{\rho}_\alpha$ is the estimator of ρ obtained from the α-th group.

For the jackknife estimator, the sample is also divided at random into k groups, and the pseudovalue $\hat{\rho}_\alpha$ is defined by

$$\hat{\rho}_\alpha = k\hat{\rho} - (k-1)\hat{\rho}_{(\alpha)},$$

where $\hat{\rho}_{(\alpha)}$ is the estimator of ρ obtained from the sample after deleting the α-th group.

For the Taylor series estimator, we express $\hat{\rho}$ as follows:

$$\hat{\rho}(\bar{u},\bar{v},\bar{w},\bar{x},\bar{y}) = \frac{\bar{w} - \bar{x}\bar{y}}{(\bar{u}-\bar{x}^2)^{\frac{1}{2}}(\bar{v}-\bar{y}^2)^{\frac{1}{2}}} ,$$

where $U_i = X_i^2$, $V_i = Y_i^2$, and $W_i = X_iY_i$. Then,

$$\hat{r} = \hat{d}_1u_i + \hat{d}_2v_i + \hat{d}_3w_i + \hat{d}_4x_i + \hat{d}_5y_i,$$

where $(\hat{d}_1, \hat{d}_2, \hat{d}_3, \hat{d}_4, \hat{d}_5)$ is the vector of partial derivatives of $\hat{\rho}$ with respect to its five arguments evaluated at the point $(\bar{u}, \bar{v}, \bar{w}, \bar{x}, \bar{y})$.

Alternative random group or jackknife estimators may be obtained by taking squared deviations from $\hat{\bar{\rho}} = k^{-1} \sum_\alpha^k \hat{\rho}_\alpha$. An alternative Taylor series estimator may be obtained by grouping the $\hat{r}_i$ and then applying the random group or jackknife estimators to the group means. None of these alternatives are discussed specifically in this paper.

In Section 4 we discuss confidence intervals for correlation coefficient. The intervals are of the form

$$\left(\hat{\rho}-c\sqrt{v(\hat{\rho})},\ \hat{\rho}+c\sqrt{v(\hat{\rho})}\right),$$

where c is the tabular value from either the normal or Student's distributions and $v(\hat{\rho})$ is one of the three estimators of $\mathrm{Var}\{\hat{\rho}\}$.

In small samples from bivariate normal populations, Fisher's z-transformation i.e.,

$$z = \phi(\rho) = \frac{1}{2}\log(1+\rho)/(1-\rho),$$

should improve the quality of confidence intervals for ρ, where quality is measured by the discrepancy between actual and nominal coverage rates. We shall investigate whether such improvement occurs in confidence intervals constructed for real survey data.

To obtain a confidence interval for ρ, using the transformation, we construct an interval for $\phi(\rho)$ and then transform back to the original scale. The general form of the interval for $\phi(\rho)$ is

$$\phi(\hat{\rho}) - c\sqrt{v(\phi(\hat{\rho}))},\quad \phi(\hat{\rho}) + c\sqrt{v(\phi(\hat{\rho}))}$$

where $v(\phi(\hat{\rho}))$ is v_{rg} or v_j applied to $\phi(\hat{\rho}_\alpha)$ or v_{ts} applied to $\phi(r_i)$.

3. Data

The data in this study were collected in the 1972-73 Consumer Expenditure Survey, sponsored by the Bureau of Labor Statistics and conducted by the Bureau of the Census. The correlation between monthly grocery store purchases and the sum of selected annual income categories was chosen for investigation. The data refer to 1972 annual income and average monthly grocery purchases during the first quarter of 1973. An experimental file of 4,532 consumer units who responded to all the grocery purchases and income categories during the first quarter of 1973 was created and treated as the finite population of interest. The income categories which were selected are: wages and salary, own business, own farm, interest, social security and railroad pension, regular contributions, federal civil service retirement, state civil service retirement, veterans benefits, armed forces pay, armed forces subsistance allowance.

The population mean of the income variable for the 4,532 consumer units is \$14,006.60 and the standard deviation is \$12,075.42. The mean and standard deviation of monthly grocery store purchases are \$146.30 and \$84.85 respectively. The correlation between annual income and monthly grocery store purchases is ρ = .3584[1]. Figure 1 is a scatter plot of the data.

4. Empirical Results

To investigate the properties of the estimators, 500 samples (srs wor) of size n = 60 were selected from the population described in Section 3. Also, 500 samples of size n = 120 and 1,000 samples of size n = 480 were drawn. These sample sizes correspond roughly to the sampling fractions .013, .026, and .106, respectively.

For each sample size, the following were computed:

a. the mean and variance of $\hat{\rho}$
b. the mean and variance of $v_{rg}(\hat{\rho})$
c. the mean and variance of $v_j(\hat{\rho})$
d. the mean and variance of $v_{ts}(\hat{\rho})$
e. proportion of confidence intervals formed using $v_{rg}(\hat{\rho})$ that contain the true ρ
f. proportion of confidence intervals formed using $v_j(\hat{\rho})$ that contain the true ρ
g. proportion of confidence intervals formed using $v_{ts}(\hat{\rho})$ that contain the true ρ
h. coverage rates in e, f, and g for confidence intervals constructed using Fisher's z-transformation.

For computations involving the random group or jackknife estimator and sample size n = 60 or 120, the group size was m = 5 and k = n/5 = 12 or 24. For n = 480, the group size was m = 15 and k = n/15 = 32.

For all confidence intervals, the value of the constant c was taken as the tabular value of Student's t with k-1 degress of freedom.

The Monte Carlo properties of the three estimators of variance and of confidence intervals constructed with them are presented in Table 1. Clearly the quality of confidence intervals (as measured by the discrepancy between actual and nominal coverage rates) constructed with the Taylor series estimator is less than the quality of confidence intervals that use random group or jackknife variance estimators. In terms of the properties of the variance estimators, however, the Taylor series estimator compares favorably with the two replication type estimators. The Taylor series estimator tends to be downward biased, whereas the jackknife estimator tends to be biased upwards. The variance and MSE of the Taylor series and random group estimators are about the same, whereas the corresponding quantities for the jackknife are larger. There is little to choose between random group and jackknife in respect to the quality of confidence intervals. Most of the confidence intervals err on the side of being larger than the true ρ.

The Monte Carlo variance of $\hat{\rho}$ is .0166, .0102, and .0037 for n = 60, 120, and 480, respectively. We note that the relative biases of the variance estimators are around 20 percent, and that they do not decrease with increasing sample size. These results are generally consistent with Frankel (1971).

Properties of the variance estimators based on transformed data, i.e., $v_{rg}(\phi(\hat{\rho}))$, $v_j(\phi(\hat{\rho}))$, and $v_{ts}(\phi(\hat{\rho})$, were also computed but are not presented here. Comparisons between the three variance estimators remain largely as described for the untransformed data.

Table 2 presents the coverage rates for confidence intervals based on transformed data. For the random group and jackknife estimators, there is clear improvement in the quality of these confidence intervals vis-a-vis the intervals based on untransformed data. The only exception is the random group estimator with n = 60. For the Taylor series estimator, the transformation is no help and seems to decrease the quality of the confidence intervals.

We illustrate the distributions of the estimators in Figures 2 and 3. Both figures refer to the distribution of the random group estimator for n = 120. Figure 2 displays $v_{rg}(\hat{\rho})$ and Figure 3, $v_{rg}(\phi(\hat{\rho}))$. Apparently the transformation causes little change in the shape of the distribution for these data. Results are similar for the other estimators and sample sizes.

We illustrate the distribution of

$$\hat{t} = \frac{\hat{\rho} - \rho}{\sqrt{v(\hat{\rho})}}$$

and of

$$\hat{t}_\phi = \frac{\phi(\hat{\rho}) - \phi(\rho)}{\sqrt{v(\phi(\hat{\rho}))}}$$

in Figures 4 and 5, respectively. Both figures again refer to the random group estimator $v_{rg}(.)$ for n = 120. Although the mean of $\hat{t}_\phi$ is closer to zero. Also, the variance of $\hat{t}_\phi$ is less than the variance of $\hat{t}$. These two characteristics explain the improvement in confidence intervals when the transformed data are used. There is little change in the coefficient of skewness or the kurtosis.

An anomaly in Tables 1 and 2 is that the quality of the confidence intervals decreases with increasing sampling fraction. The anomaly was also noted, but never explained, in Frankel's (1971) work. An explanation is that 1) the correlation between $\hat{\rho}$ and $v(\hat{\rho})$ is negative and decreases with increasing sampling fraction, and 2) $\hat{\rho}$ is biased upwards (the Monte Carlo expectation of $\hat{\rho}$ is .4122, .4033, and .3884 for n = 60, 120, 480, respectively). Thus, $\hat{\rho}$ tends to be too large, and when $\hat{\rho}$ is too large $v(\hat{\rho})$ tends to be too small. The confidence interval tends to miss the true ρ, particularly on the high side, and this situation worsens as the sampling fraction increases. This behavior also occurs when the z-transformation is used.

5. Recommendations

Based on the empirical results presented in Section 4, the random group or jackknife estimator can be recommended for estimating the design variance of the estimated finite population correlation coefficient ρ. Both of these estimators seem preferable to the Taylor series estimator. Further, for making inferential statements, Fisher's z-transformation seems to improve the quality of studentized statistics and of confidence intervals for ρ. It should be recognized, however, that the Monte Carlo study discussed here was limited in several important respects, and that the results may not generalize to other variables, other finite populations, or other sampling designs.

Table 1. Estimators of Var($\hat{\rho}$) Summary of 500 SRSWOR for n = 60, 120, 1000 SRSWOR for n = 480

Technique	Bias	Variance	MSE	90% Confidence Intervals: Average width	% contain ρ	% ρ ≤ lower bound	% ρ ≥ upper bound	95% Confidence Intervals: Average width	% contain ρ	% ρ ≤ lower bound	% ρ ≥ upper bound
Random Group											
n=60,k=12,m=5	.00139	.000058	.000060	.471	87.4	11.4	1.2	.577	93.8	6.0	0.2
n=120,k=24,m=5	-.00174	.000006	.000009	.313	81.6	16.2	2.2	.377	88.8	10.8	0.4
n-480,k=32,m=15	-.00194	>.000001	.000004	.141	67.2	27.4	5.4	.170	76.6	21.0	2.4
Jackknife											
n=60,k=12,m=5	.00134	.000397	.000407	.464	84.0	15.0	1.0	.569	90.4	9.2	0.4
n=120,k=24,m=5	.00070	.000089	.000089	.335	81.4	16.8	1.8	.405	87.8	11.4	0.8
Taylor Series											
n=60	-.00347	.000051	.000063	.400	77.2	18.4	4.4	.490	83.4	14.2	2.4
n=120	-.00242	.000017	.000023	.294	75.2	21.2	3.6	.355	83.6	14.4	2.0
n=480	-.00114	.000001	.000002	.162	70.4	21.8	7.8	.193	77.6	15.8	6.6

Table 2. Confidence Intervals for z = $\phi(\rho)$ Summary of 500 SRSWOR for n = 60, 120,1000 SRSWOR for n = 480

Technique	90% Confidence Intervals: Average width	% contain z	% z ≤ lower bound	% z ≥ upper bound	95% Confidence Intervals: Average width	% contain z	% z ≤ lower bound	% z ≥ upper bound
Random Group								
n=60,k=12,m=5	.760	94.8	5.0	0.2	.931	98.2	1.6	0.2
n=120,k=24,m=5	.511	93.4	6.2	0.4	.617	96.8	3.2	0.0
n=480,k=32,m=15	.195	75.8	21.4	2.8	.234	83.7	15.2	1.1
Jackknife								
n=60,k=12,m=5	.578	87.8	11.0	1.2	.709	93.4	6.0	0.6
n=120,k=24,m=5	.409	83.6	14.2	2.2	.494	90.0	9.2	0.8
Taylor Series								
n=60	.401	71.0	24.2	4.8	.491	78.2	18.8	3.0
n=120	.295	68.0	27.8	4.2	.356	75.4	21.8	2.8
n=480	.162	62.6	28.8	8.6	.193	71.0	21.5	7.5

Table 3. Correlation Between $\hat{\rho}$ and $v(\hat{\rho})$

Variance Estimator	Sample Size		
	n = 60	n = 120	n = 480
Random Group	-.2815	-.3764	-.4626
Jackknife	-.2120	-.2801	—
Taylor Series	-.1948	-.1928	-.0559

Figure 1. Grocery store purchases vs Income

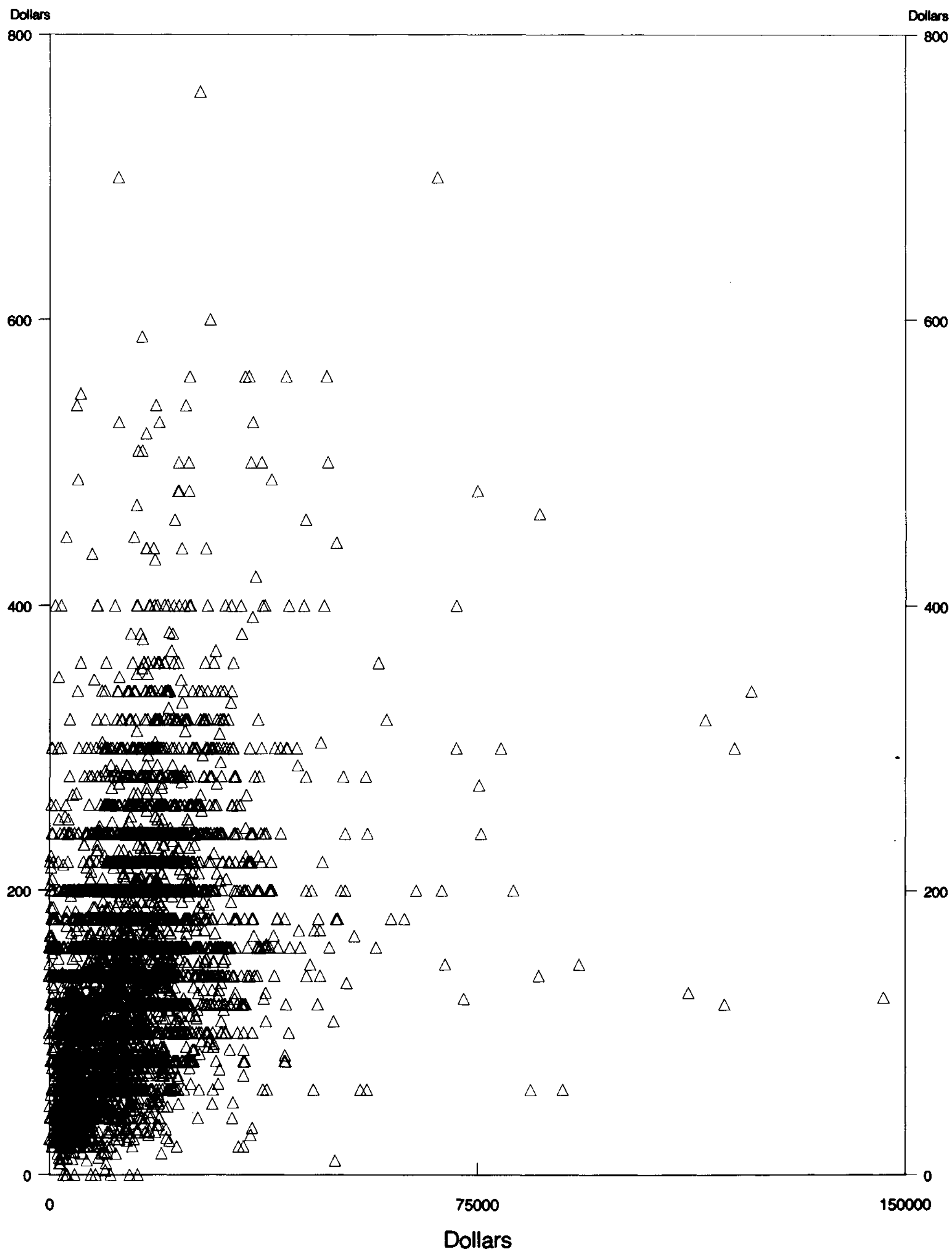


Figure 2. **ESTIMATED VARIANCE OF ESTIMATED RHO**

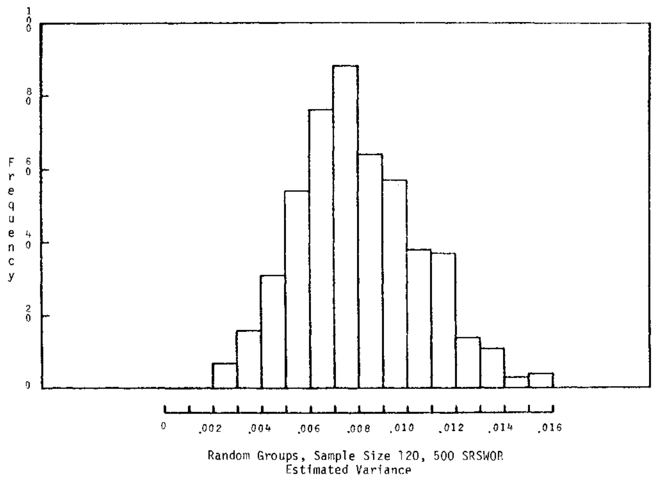


Figure 3. **ESTIMATED VARIANCE OF ESTIMATED Z**

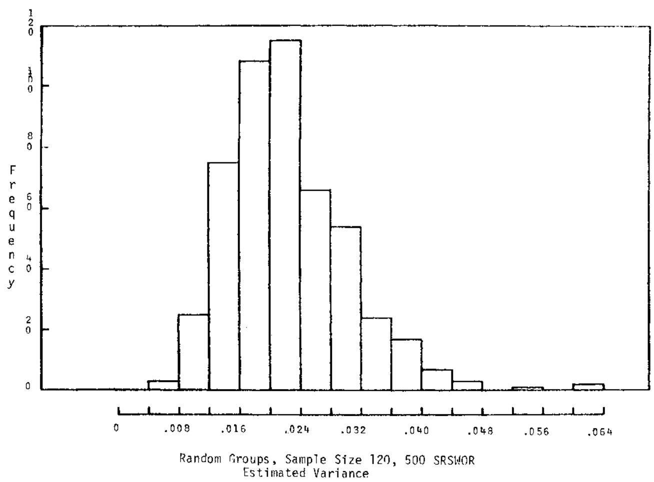


Figure 4. **T STATISTIC FOR RHO**

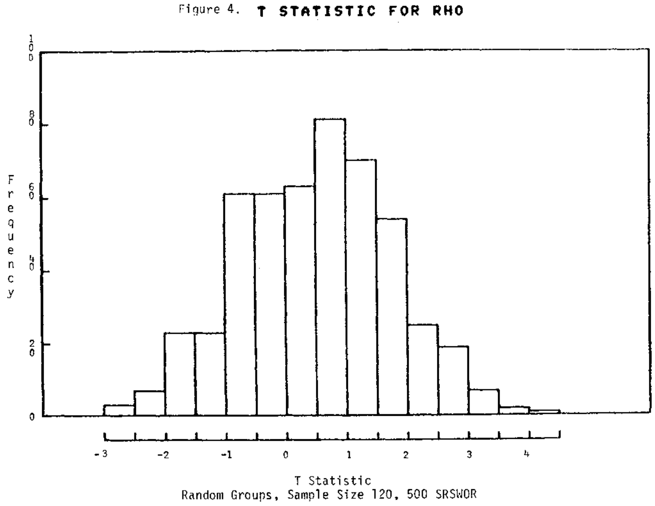
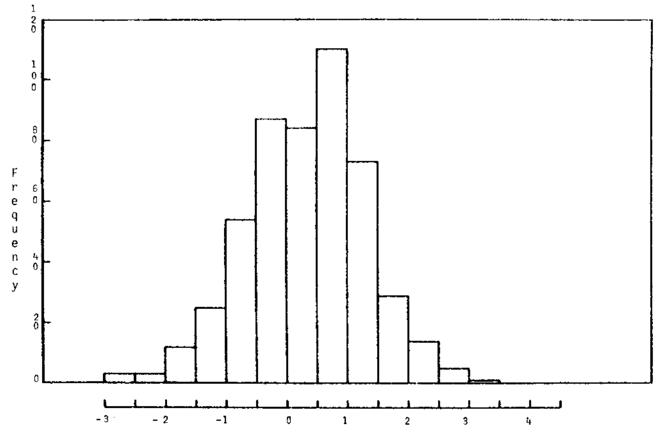


Figure 5. **T STATISTIC FOR Z**



T Statistic
Random Groups, Sample Size 120, 500 SRSWOR

REFERENCES

1. Bean, J.A. (1975), Distribution and Properties of Variance Estimators for Complex Multistage Probability Samples. National Center for Health Statistics, Series 2, Number 65.
2. Campbell, C. and Meyer, M. (1978), Some properties of T confidence intervals for survey data. Proc. Sec. Survey Res. Methods. Amer. Statist. Assoc., Washington, D.C. 437-442.
3. Frankel, M.R. (1971), Inference from Survey Samples. Institute for Social Research, The University of Michigan, Ann Arbor, MI.
4. Mellor, R.W. (1973), "Subsample replication variance estimators." Unplublished Ph.D dissertation. Harvard University.
5. Krewski, D. and Rao, J.N.K. (1978), Inference from stratified samples I: large sample properties of the linearization, jackknife and balanced repeated replication methods. Technical Report, Dept. of Math., Carleton University.
6. Krewski, D. and Rao, J.N.K. (1979), Inference from stratified samples II: small sample properties of the linearization, jackknife and balanced repeated replication methods. Technical Report, Dept. of Math., Carleton University.
7. Lemeshow, S. and Levy, P. (1978), Estimating the variance of ratio estimates in complex sample surveys with two primary units per stratum - a comparison of balanced replication and jackknife techniques. J. Statist. Comp. Simul. A7 1007-1025.

FOOTNOTE

[1]The grocery store purchases include purchases made with food stamps. This probably tends to depress the correlation.