ESTIMATORS FOR SAMPLES SELECTED FROM TWO OVERLAPPING FRAMES

Wayne A. Fuller, Iowa State University
Leon F. Burmeister, University of Iowa

1. INTRODUCTION

The use of two frames or lists to select a sample of a population was, perhaps, first used by the United States Bureau of the Census in the "Sample Survey of Retail Stores" conducted in 1949 and described by Bershad in Hansen, Hurwitz and Madow [1, p. 516]. Hartley [7] considered the design and estimation for such surveys. Lund [10], Cochran [3] and Williams [12] have also studied the problem and offered alternative estimation procedures.

We assume two frames, A and B, containing N_A and N_B elements respectively are available. We adopt the notation of Hartley [7] and denote by N_{ab} the number of elements included in both frame A and frame B. N_a is the number of elements occurring only on frame A and N_b is the number of elements occurring only on frame B. Thus

$$N_A = N_a + N_{ab},$$

$$N_B = N_b + N_{ab}$$

and the total number of elements in the population, N, is given by

$$N = N_a + N_b + N_{ab} = N_a + N_B = N_b + N_A.$$

We call the elements contained only on frame A domain a, the elements only on frame B domain b and those elements on both frames A and B domain ab, or the overlap domain. We assume that simple random samples of sizes n_A and n_B are drawn from frames A and B respectively. The number of elements sampled from frame A and contained in domain a is denoted by n_a. The number of elements sampled from frame A and contained in domain ab is denoted by n'_{ab}. The number of sampled elements in domains ab and b drawn from frame B are denoted by n''_{ab} and n_b respectively. Thus

$$n_A = n_a + n'_{ab}, \text{ and}$$

$$n_B = n''_{ab} + n_b. \tag{1.1}$$

We consider the estimation of the number of elements in the overlap domain, N_{ab}, and the estimation of the total of a characteristic, Y. We develop estimators which to order (n) are superior to those appearing in the literature.

2. ESTIMATION OF N_{ab}

2.1 Estimation of N_{ab} when Duplicated Items are Ignored

We first consider the problem of estimating the number of elements in the overlap domain, N_{ab}. Hartley [7] proposed the estimator

$$\hat{N}_{ab,H} = \frac{pn'_{ab}N_A}{n_A} + \frac{qn''_{ab}N_B}{n_B}, \tag{2.1}$$

where p and q are constants with the property

$$p + q = 1.$$

This estimator makes no use of elements in the sample that were selected both from frame A and from frame B. The variance of Hartley's estimator is given by

$$\text{Var}(\hat{N}_{ab,H}) = p^2f_A^{-2}\,\text{Var}(n'_{ab}) + q^2f_B^{-2}\,\text{Var}(n''_{ab}). \tag{2.2}$$

Since n'_{ab} and n''_{ab} are hypergeometric random variables, it follows that the p which minimizes the variance is given by

$$p'_{oH} = \frac{n_AN_bg_B}{n_AN_bg_B + n_BN_ag_A}, \tag{2.3}$$

where

$$g_A = \frac{N_A - n_A}{N_A - 1}$$

and

$$g_B = \frac{N_B - n_B}{N_B - 1}.$$

For this value of p the variance of the estimated overlap domain is given by

$$\text{Var}_o(\hat{N}_{ab,H}) = \frac{N_{ab}N_aN_bg_Ag_B}{n_AN_bg_B + n_BN_ag_A}. \tag{2.4}$$

To obtain an estimator of N_{ab} that is a function of the sample data only we set

$$\hat{N}_b = N_B - \hat{N}_{ab},$$

$$\hat{N}_a = N_A - \hat{N}_{ab}$$

and substitute expression (2.3) for p into (2.1). Expression (2.1) then reduces to a quadratic in $\hat{N}_{ab}$ and we define a new estimator by this quadratic,

$$[n_Ag_B + n_Bg_A]\hat{N}^2_{ab,s} - [n_AN_Bg_B + n_BN_Ag_A + n'_{ab}N_Ag_B + n''_{ab}N_Bg_A]\hat{N}_{ab,s} + [n'_{ab}g_B + n''_{ab}g_A]N_AN_B = 0. \tag{2.5}$$

It can be shown that the roots of (2.5) are always real and that the largest root is always greater than or equal to the minimum of N_A and N_B. The smallest root is always contained in the interval zero to minimum (N_A, N_B), inclusive. Hence we take the left (smallest) root of (2.5) as the estimator of N_{ab}. Note that the Hartley estimator (2.1) with fixed p, unlike the estimator defined by (2.5), does not always fall in the range of feasible values for N_{ab}.

The large sample properties of $\hat{N}_{ab,s}$ are given in the following theorem.

Theorem 1: The sequence of estimators, $\hat{N}_{ab,s}$, defined by (2.5) satisfies

$$\operatorname{Var}(\hat{N}_{ab,s}) = \frac{N_{ab}N_aN_bg_Bg_A}{n_AN_bg_B + n_BN_ag_A} + O(1), \qquad (2.6)$$

and

$$E(\hat{N}_{ab,s}) = N_{ab} + O(\tfrac{1}{n}).$$

Thus $\hat{N}_{ab,s}$ has the same limiting variance as the linear estimator (2.1) employing optimal weights. Further the bias of $\hat{N}_{ab,s}$ is of low order compared to the variance of $\hat{N}_{ab,s}$. Based on these results we recommend $\hat{N}_{ab,s}$ as an estimator for N_{ab} when it is judged impractical to identify elements entering the sample from both frames.

For replacement sampling the estimator $\hat{N}_{ab,s}$ reduces to the maximum likelihood estimator presented by Williams [12].

2.2 Estimation of N_{ab} when duplicated elements are identified

Assume that the n'_{ab} sample elements from frame A are compared with the n''_{ab} elements from frame B and it is determined that n_d of these elements are common. The probability that an element in domain ab will be selected in the sample from frame A and from frame B is $(n_A/N_A)(n_B/N_B)$. Hence $n_d f_A^{-1} f_B^{-1}$ is an unbiased estimator of N_{ab}. We consider the estimator of N_{ab} constructed as the linear combination of three unbiased estimators:

$$\hat{N}_{ab,1} = pf_A^{-1}n'_{ab} + rf_B^{-1}n''_{ab} + (1 - p - r)f_A^{-1}f_B^{-1}n_d. \qquad (2.9)$$

The variance of $\hat{N}_{ab,1}$ is given by

$$\operatorname{Var}\{\hat{N}_{ab,1}\} = p^2C_A + r^2C_B + (1 - p - r)^2 [C_A + C_B + C_{AB}] + 2p(1 - r - p)C_A + 2r(1 - r - p)C_B, \qquad (2.10)$$

where

$$C_A = f_A^{-2}\operatorname{Var}\{n'_{ab}\} = g_An_A^{-1}N_aN_{ab},$$

$$C_B = f_B^{-2}\operatorname{Var}\{n''_{ab}\} = g_Bn_B^{-1}N_bN_{ab}$$

and

$$C_{AB} = N_{ab}f_A^{-1}f_B^{-1}g_Ag_B(1 - \frac{1}{N_A} - \frac{1}{N_B} + \frac{N_{ab}}{N_AN_B})$$

$$= N_{ab}f_A^{-1}f_B^{-1}g_Ag_B + O(1).$$

Minimizing the variance with respect to p and r we obtain

$$p_o = \frac{C_AC_B + C_{AB}C_B}{C_AC_B + C_AC_{AB} + C_{AB}C_B}$$

and

$$r_o = \frac{C_AC_B + C_{AB}C_A}{C_AC_B + C_AC_{AB} + C_{AB}C_B} \ . \qquad (2.11)$$

Substituting the optimum values of p and r into equation (2.10), we obtain

$$\operatorname{Var}_o\{\hat{N}_{ab,1}\} = \frac{N_{ab}N_aN_bg_Ag_B}{f_Af_BN_aN_b + n_BN_ag_A + n_AN_bg_B} \ . \qquad (2.12)$$

Comparing the variance expression (2.12) with that of $\hat{N}_{ab,H}$, (2.4), we see that identifying the duplicated items reduces the variance by a fraction depending upon the sampling rates and the proportion of the population in the overlap domain. That is,

$$\operatorname{Var}\{\hat{N}_{ab,1}\} = [f_B(1 - \alpha)g_A + f_A(1 - \beta)g_B] \operatorname{Var}\{\hat{N}_{ab,H}\}[f_Af_B(1 - \alpha)(1 - \beta) + f_B(1 - \alpha)g_A + f_A(1 - \beta)g_B]^{-1} \ . \qquad (2.13)$$

We now consider the maximum likelihood estimator of N_{ab}. The probability of obtaining a sample with the given number of elements selected from the overlap domain is given by

$$L\{n'_{ab},n''_{ab},n_d;N_{ab}\} = L(n'_{ab};N_{ab})L(n''_{ab},n_d|n'_{ab};N_{ab})$$

$$= \frac{\binom{N_a}{n_a}\binom{N_{ab}}{n'_{ab}}}{\binom{N_A}{n_A}} \ \frac{\binom{N_b}{n_b}\binom{N_{ab} - n'_{ab}}{n''_{ab} - n_d}\binom{n'_{ab}}{n_d}}{\binom{N_B}{n_B}} \ .$$

Setting the ratio of $L(n'_{ab},n''_{ab},n_d;N_{ab})$ to $L(n'_{ab},n''_{ab},n_d;N_{ab} - 1)$ equal to one, we find that the value of N_{ab} maximizing the likelihood is given as the solution to the quadratic equation[1]

$$(n_a + n_{ab} + n_b)\hat{N}^2_{ab,m} - [n_aN_B + n_{ab}(N_A + N_B) + n_bN_A - n_an_b]\hat{N}_{ab,m} + n_{ab}N_AN_B = 0, \qquad (2.14)$$

where

$$n_{ab} = n'_{ab} + n''_{ab} - n_d.$$

It is possible to show that $\hat{N}_{ab,m}$ is always given by the left root of (2.14).

Theorem 2: The sequence of maximum likelihood estimators defined by (2.14) satisfies

$$\mathrm{Var}\{\hat{N}_{ab,m}\} = \frac{N_{ab}N_aN_bg_Ag_B}{f_Af_BN_aN_b + n_BN_ag_A + n_AN_bg_B} + 0(1)$$

and

$$E\{\hat{N}_{ab,m} - N_{ab}\} = (f_Af_BN_{ab}^2N_aN_bg_Ag_B) \Big/$$

$$\{(f_Af_BN_aN_b + n_BN_ag_A + n_AN_bg_B)$$

$$[-2N_{ab}(n_A - n_B + f_Af_BN_{ab}) + (n_A + f_BN_{ab}$$

$$- f_Af_BN_{ab})N_B + (n_B + f_AN_{ab} - f_Af_BN_{ab})N_A$$

$$- f_Af_BN_aN_b]\} + 0(\tfrac{1}{n}).$$

3. ESTIMATORS OF THE POPULATION TOTAL

3.1 Duplicated items not identified

We now consider methods of estimating the population total, Y. It is well known that, for example

$$E\{\bar{y}_a | n_a, n'_{ab}, n''_{ab}, n_b\} = \bar{Y}_a.$$

Therefore given an estimator of N_{ab} we consider the estimator

$$\hat{Y}_s = (N_A - \hat{N}_{ab,s})\bar{y}_a + \hat{N}_{ab,s}\bar{y}_{ab,s}$$

$$+ (N_B - \hat{N}_{ab,s})\bar{y}_b, \qquad (3.1)$$

where

$$\bar{y}_{ab,s} = w\bar{y}'_{ab} + (1 - w)\bar{y}''_{ab},$$

$$\bar{y}'_{ab} = \frac{1}{n'_{ab}} \sum_{i=1}^{n'_{ab}} y_i$$ = mean of elements in domain ab selected from frame A,

$$\bar{y}''_{ab} = \frac{1}{n''_{ab}} \sum_{j=1}^{n''_{ab}} y_j$$ = mean of elements in domain ab selected from frame B

and

$$w = \frac{n'_{ab}(1 - f_B)}{n'_{ab}(1 - f_B) + n''_{ab}(1 - f_A)} .$$

Clearly w has been chosen to minimize the variance of $\bar{y}_{ab,s}$. If the finite correction term can be ignored the estimator of $\bar{Y}_{ab}$ reduces to the mean of all $n'_{ab} + n''_{ab}$ elements selected from domain ab.

Note that, given $\hat{N}_{ab,s}$, the estimator $\hat{Y}_s$ is linear in the observations y_i. Since the weights are not a function of the characteristic they apply equally well for all y-characteristics.

Theorem 3: The sequence of estimators defined by (3.1) satisfies

$$\mathrm{Var}\{\hat{Y}_S\} = N_a(f_A^{-1} - 1)S_a^2 + [(1 - f_B)f_A +$$

$$+ (1 - f_A)f_B]^{-1}(1 - f_A)(1 - f_B)N_{ab}S_{ab}^2$$

$$+ N_b(f_B^{-1} - 1)S_b^2$$

$$+ (\bar{Y}_{ab} - \bar{Y}_a - \bar{Y}_b)^2 \frac{N_{ab}N_aN_bg_Ag_B}{n_AN_bg_B + n_BN_ag_A} + 0(1) \qquad (3.2)$$

and

$$E\{\hat{Y}_S\} = Y + 0(\tfrac{1}{n}) ,$$

where

$$S_a^2 = \frac{1}{N_a - 1} \sum_{i=1}^{N_a} (y_{ai} - \bar{Y}_a)^2,$$

$$S_{ab}^2 = \frac{1}{N_{ab} - 1} \sum_{i=1}^{N_{ab}} (y_{ab,i} - \bar{Y}_b)^2$$

and

$$S_b^2 = \frac{1}{N_b - 1} \sum_{i=1}^{N_b} (y_{bi} - \bar{Y}_b)^2.$$

Hartley [7] originally suggested the estimator

$$\hat{Y}_H = \frac{N_A}{n_A}(n_a\bar{y}_a + pn'_{ab}\bar{y}'_{ab})$$

$$+ \frac{N_B}{n_B}(qn''_{ab}\bar{y}''_{ab} + n_b\bar{y}_b),$$

and Lund [10] suggested the modification

$$\hat{Y}_L = f_A^{-1}n_a\bar{y}_a + (pf_A^{-1}n'_{ab} + qf_B^{-1}n''_{ab})\bar{y}_{ab,L}$$

$$+ f_B^{-1}n_b\bar{y}_b, \qquad (3.3)$$

where

$$\bar{y}_{ab,L} = \frac{n'_{ab}\bar{y}'_{ab} + n''_{ab}\bar{y}''_{ab}}{n'_{ab} + n''_{ab}} .$$

The variance of Lund's estimator

$$\mathrm{Var}(\hat{Y}_L) = N_a(f_A^{-1} - 1)S_a^2 + (f_A + f_B)^{-2}[f_A(1 - f_A)$$

$$+ f_B(1 - f_B)]N_{ab}S_{ab}^2 + N_b(f_B^{-1} - 1)S_b^2$$

$$+ g_A n_A^{-1} N_{ab} N_a [\bar{Y}_a - p\bar{Y}_{ab}]^2$$

$$+ g_B n_B^{-1} N_{ab} N_b [\bar{Y}_b - q\bar{Y}_{ab}]^2$$

is never greater than the variance of Hartley's estimator.

Lund gave the optimum value of p as

$$p_{oL} = \frac{f_A^{-1}(1-\alpha)\bar{Y}_a + f_B^{-1}(1-\beta)(\bar{Y}_{ab} - \bar{Y}_b)}{[f_A^{-1}(1-\alpha) + f_B^{-1}(1-\beta)]\bar{Y}_{ab}}$$

and suggested the estimator

$$\hat{p}_{oL} = \frac{n_B f_A^{-1} n_a \bar{y}_a + n_A f_B^{-1} n_b (\bar{y}_{ab} - \bar{y}_b)}{[n_B f_A^{-1} n_a + n_A f_B^{-1} n_b]\bar{y}_{ab}} .$$

Substitution of $\hat{p}_{oL}$ into $\hat{Y}_L$ (3.3) gives an estimator of the same form as $\hat{Y}_s$ of equation (3.1).

The bias in Lund's estimator is 0(1) while that of $\hat{Y}_s$ is 0(1/n). Further the estimator of $\bar{Y}_{ab}$ employed in $\hat{Y}_s$ is more efficient than that in $\hat{Y}_L$ for nonreplacement sampling. Also the linear form of (3.1) is clearly a computational advantage in large scale surveys.

3.2 Duplicate Items Identified

If the items entering the sample from both frame A and frame B are identified, the estimator $\hat{N}_{ab,m}$ is superior to $\hat{N}_{ab,s}$. Also the mean of the distinct units is superior to $\bar{y}_{ab,s}$ as an estimator of the mean of the overlap domain. Therefore the estimator

$$\hat{Y}_m = (N_A - \hat{N}_{ab,m})\bar{y}_a + \hat{N}_{ab,m}\bar{y}_{ab} + (N_B - \hat{N}_{ab,m})\bar{y}_b,$$

where

$$\bar{y}_{ab} = \frac{1}{n_{ab}} \sum_{i=1}^{n_{ab}} y_i = \text{mean of distinct units}$$

and

$$n_{ab} = n'_{ab} + n''_{ab} - n_d,$$

clearly has smaller variance than $\hat{Y}_s$. We have

$$\text{Var}(\hat{Y}_m) = N_a f_A^{-1}(1 - f_A)S_a^2 + (1 - f_A - f_B + f_A f_B)$$

$$(f_A + f_B - f_A f_B)^{-1} N_{ab} S_{ab}^2 + N_b f_B^{-1}(1 - f_B)S_b^2$$

$$+ \frac{(\bar{Y}_{ab} - \bar{Y}_a - \bar{Y}_b)^2 N_{ab} N_a N_b (1 - f_A)(1 - f_B)}{[f_A f_B N_a N_b + n_B N_a (1 - f_A) + n_A N_b (1 - f_B)]}$$

$$+ 0(1).$$

4. ESTIMATION FOR GENERAL SAMPLING PROCEDURES

We now relax the assumption of simple random sampling in each frame. We assume that sampling is such that unbiased estimators of totals whose error is $0_p(n^{1/2})$ and estimators of the variances whose error is $0_p(n^{1/2})$ of these estimators are available. Let

- $\hat{Y}_a$ be an unbiased estimator of the total for domain a constructed from the sample of frame A,
- $\hat{Y}_b$ be an unbiased estimator of the total for domain b constructed from the sample of frame B,
- $\hat{Y}'_{ab}$ be an unbiased estimator of the total of domain ab constructed from the sample of frame A, and
- $\hat{Y}''_{ab}$ be an unbiased estimator of the total of domain ab constructed from the sample of frame B.

We assume that the observational unit is the same for the sample selected from frame A or for that selected from frame B[2] and let

- $\hat{N}'_{ab}$ be an unbiased estimator of the number of observational units in domain ab constructed from the sample of frame A,
- $\hat{N}''_{ab}$ be an unbiased estimator of the number of observational units in domain ab constructed from the sample of frame B,
- $\hat{N}_a$ be an unbiased estimator of the number of observational units in domain a constructed from the sample of frame A, and
- $\hat{N}_b$ be an unbiased estimator of the number of observational units in domain b constructed from the sample of frame B.

In the previous sections an estimator of N_{ab} was constructed from the two unbiased estimators $\hat{N}'_{ab}$ and $\hat{N}''_{ab}$. Similarly an estimator of $\bar{Y}_{ab}$ was constructed from the two unbiased estimators $\bar{y}'_{ab}$ and $\bar{y}''_{ab}$. Since the expected value of $\bar{y}'_{ab}$ and $\bar{y}''_{ab}$ given $\hat{N}'_{ab}$, $\hat{N}'_{ab} \neq 0$, and $\hat{N}''_{ab}$, $\hat{N}''_{ab} \neq 0$, is $\bar{Y}_{ab}$, it was possible to treat the estimation of N_{ab} and $\bar{Y}_{ab}$ separately.

If the sampling is other than simple random, $\hat{N}'_{ab}$ and $\hat{N}''_{ab}$ may be correlated with $\bar{y}'_{ab} = \hat{Y}'_{ab}/\hat{N}'_{ab}$ and $\bar{y}''_{ab} = \hat{Y}''_{ab}/\hat{N}''_{ab}$. Therefore, to be efficient, it is necessary to use all auxiliary information in the estimation of, for example, Y. This suggests the estimator

$$\tilde{Y}_r = \hat{Y}_a + \hat{Y}_B + \beta_1(\hat{N}'_{ab} - \hat{N}''_{ab}) + \beta_2(\hat{Y}'_{ab} - \hat{Y}''_{ab}), \quad (4.1)$$

where

$$\hat{Y}_B = \hat{Y}_b + \hat{Y}''_{ab}.$$

The optimal $\hat{\beta}' = (\hat{\beta}_1, \hat{\beta}_2)$ is given by

$$\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} = \begin{bmatrix}\hat{Var}(\hat{N}'_{ab} - \hat{N}''_{ab}) & \hat{Cov}(\hat{N}'_{ab} - \hat{N}''_{ab}, \hat{Y}'_{ab} - \hat{Y}''_{ab})\\ \text{sym.} & \hat{Var}(\hat{Y}'_{ab} - \hat{Y}''_{ab})\end{bmatrix}^{-1}$$

$$\begin{bmatrix}-\hat{Cov}(\hat{Y}_a, \hat{N}'_{ab}) + \hat{Cov}(\hat{Y}_B, \hat{N}''_{ab})\\ -\hat{Cov}(\hat{Y}_a, \hat{Y}'_{ab}) + \hat{Cov}(\hat{Y}_B, \hat{Y}''_{ab})\end{bmatrix} . \qquad (4.2)$$

A consistent estimator of the variance of the estimator is

$$\hat{Var}(\tilde{Y}_r) = \hat{Var}(\hat{Y}_a) + \hat{Var}(\hat{Y}_B) +$$

$$+ \hat{\beta}_1[\hat{Cov}(\hat{Y}_a, \hat{N}'_{ab}) - \hat{Cov}(\hat{Y}_B, \hat{N}''_{ab})]$$

$$+ \hat{\beta}_2[\hat{Cov}(\hat{Y}_a, \hat{Y}'_{ab}) - \hat{Cov}(\hat{Y}_B, \hat{Y}''_{ab})]. \qquad (4.3)$$

The estimator (4.1) is clearly recognizable as a multiple regression estimator. Hartley proposed the estimator

$$\hat{Y}_H = \hat{Y}_a + \hat{Y}''_{ab} + \hat{Y}_b + p(\hat{Y}'_{ab} - \hat{Y}''_{ab}). \qquad (4.4)$$

Hartley's estimator is inefficient relative to $\tilde{Y}_r$ if the partial correlation between $\hat{Y}_a + \hat{Y}_B$ and $\hat{N}'_{ab} - \hat{Y}''_{ab}$ is not zero.

It is also clear that if other y-characteristics are observed in the survey it may be possible to further decrease the variance of the estimator (4.1) by including other unbiased estimators of zero in the regression estimator.

The estimator (4.1) is not linear in y since $\hat{\beta}_2(\hat{Y}'_{ab} - \hat{Y}''_{ab})$ is not a linear function of y. However it is 'nearly linear' in the sense that given the differences $\hat{N}'_{ab} - \hat{N}''_{ab}$ and $\hat{Y}'_{ab} - \hat{Y}''_{ab}$ the estimator is linear in y. Thus in a large scale survey one might be willing to choose a few y values to serve as auxilliary variables for all computations. This would simplify the computations and guarantee numerical consistency between tables.

Acknowledgment

This research was partially supported by the Bureau of the Census through the Joint Statistical Agreement J.S.A. 72-4.

FOOTNOTES

1. In order to simplify our presentation we ignore the fact that the estimate of N_{ab} should be an integer.

2. This assumption can be modified so that conclusions are applicable for cases wherein the observational units are different.

REFERENCES

[1] Bershad, Max A., "A Sample Survey of Retail Stores," *Sample Survey Methods and Theory*, Vol. 1, New York: John Wiley and Sons, Inc, 1953, 516-558.

[2] Bryant, Edward C. and King, Donald W., "Estimation from Populations Identified by Overlapping Sample Frames," Unpublished paper presented at the American Statistical Association meeting, Palo Alto, California, 1960.

[3] Cochran, Robert S., "Theory and Application of Multiple Frame Surveys," Ph.D. dissertation, Iowa State University Library, Ames, 1965.

[4] Cochran, Robert S., "The Estimation of Domain Sizes When Sampling Frames are Interlocking," Proceedings of the Social Science Section of the American Statistical Association meeting, Washington, D. C., 1967.

[5] Cochran, William G., *Sampling Techniques*, 2nd ed., New York: John Wiley and Sons, Inc., 1963.

[6] Fuller, W. A., "Properties of Some Estimators for the Errors in Variables Model. Submitted for Publication, 1972.

[7] Hartley, H. O., "Multiple Frame Surveys," Proceedings of the Social Science Section of the American Statistical Association meeting, Minneapolis, Minnesota, 1962.

[8] Horvitz, D. G. and Thompson, D. J., "A Generalization of Sampling Without Replacement from a Finite Universe," *Journal of the American Statistical Association*, 47, 1952, 663-685.

[9] King, Donald W., "Variance Estimation in Populations Identified by Multiple Sampling Frames," M.S. Thesis, University of Wyoming Library, Laramie, 1960.

[10] Lund, Richard E., "Estimators in Multiple Frame Surveys," Proceedings of the Social Science Section of the American Statistical Association meeting, Pittsburg, Pennsylvania, 1968.

[11] Steinberg, J., "A Multiple Frame Survey for Rare Population Elements," Proceedings of the Social Science Section of the American Statistical Association meeting, Philadelphia, Pennsylvania, 1965.

[12] Williams, Robert E., "Estimation of Overlapping Strata Boundaries," M.S. Thesis, University of Wyoming Library, Laramie, 1957.