COMBINING PRELIMINARY ESTIMATORS OF TOTALS FOR LIVESTOCK SURVEYS BY CONVEX PROGRAMMING

Lynn Kuo, University of Connecticut

1. Introduction. The June Enumerative Survey conducted by the Statistical Reporting Service (SRS) at USDA is a multi-purpose probability survey for reporting inventory on crops, livestock, and other agricultural items. Samples are selected from an area frame. For some items, additional June survey is conducted from list frame. The area frame is stratified by land use, whereas the list frame is stratified by size of the farm. Information on the items of interest is obtained by personal interview near the end of May for the area sample. That from the list sample is obtained by mail, telephone, or personal visit around the June 1 reference date.

Different estimators are often produced for the same characteristics. For example, four estimators, tract, farm, weighted, and multiple frame screening estimators are produced for livestock items for each of the 10 major states. These 10 states usually account for more than 80 percent of the U.S. hogs and cattle inventory. Three of the estimators are derived from the same primary sampling units. Due to different methods of associating the farm products with the segments (primary sampling units) from the area frame, three different estimators are produced. The tract estimator counts only the farm inventory within the segment. The farm estimator would include the farm inventory beyond the segment, so long as those farm products belong to the same operator residing in the segment. The weighted estimator uses a formula depending on the acreages to prorate the farm inventory to tract level. A fourth estimator called multiple frame screening estimator is predominantly computed from data of the list sample. Moreover, a small portion of the weighted estimator from the area frame is also added to the list estimator to compensate for the incompleteness of the list frame.

One of the problems faced by the statistician at SRS is finding a method to combine these estimators into one. A composite estimation model is proposed here. This composite estimator is motivated by minimizing the mean squared errors of a family of weighted averages of the four preliminary estimators.

Other approaches such as empirical Bayes and linear Bayes were also explored by the author to solve this problem. Composite estimation has been pursued. The strictly frequentist and nonparametric features of the composite estimation are also shared by that of the classical survey sampling. These two features give composite estimation the greatest potential for implementation by SRS.

As can be seen from the numerical results in Section 6, not only variances but also nonnegligible biases affect the accuracy of the preliminary estimators. Consequently, analysis of biases has to be incorporated. When all nonsampling errors are considered, it is assumed that the tract estimator is unbiased, and all other estimators are biased. This assumption is also supported by Nealon (1984), where discussion on the biases of the weighted and multiple frame screening estimator can be found. The tract estimator by design is unbiased. It is also least susceptible to nonsampling errors. An unbiased estimator of the bias squared term developed in Section 4 is used for the biased preliminary estimators. The composite estimation developed in this paper provides adjustment for component weights depending on biases.

Mosteller (1948) discusses the desirability of pooling the data. He describes several ways of pooling data from two samples to estimate the mean of one of the populations. He illustrates it by using data from the normal distribution, but his ideas are applicable in a broader context. A stout believer in unbiasedness would only use the tract estimator. However, most statisticians are willing to accept some bias to reduce the mean squared error. This is done by pooling all the available data.

Theoretical work on composite estimation for independent observations from the normal distribution is given by Graybill and Deal (1959). To combine two independent unbiased preliminary estimators for the common mean, they show the composite estimator has uniformly smaller variance than any of the preliminary estimators so long as each sample size is greater than 10. Further improvement and other related references are given by Brown and Cohen (1974). Although the situation at SRS is much more complicated, these theoretical works shed light on the advantage of intelligently combining estimators.

Composite estimation has been used by numerous statisticians in applications. Schaible (1978 and 1979) uses it to estimate small area statistics for the Health Interview Survey. Brock, French, and Peyton (1980) provide an empirical evaluation of mean squared errors of composite estimators, and suggestions for component estimators for small area estimation. Cohen and Sommers (1984) provide empirical evaluation of composite estimation of cost weights for the Consumer Price Index. There is also extensive literature on composite estimation for the Current Population Survey for panel studies and rotation designs. See Wolter (1979) for the theory, applications, and other references.

The four preliminary estimators presently in use at SRS are described in Section 2. A review of composite estimation and its specialization to SRS applications are given in Section 3. Estimation of the second moment term needed in composite estimation is discussed in Section 4. Variance and mean squared error evaluations of the composite estimators are discussed in Section 5. Numerical results for total hogs and pigs inventory from the 1984 June Enumerative Survey are given in Section 6. Finally, the conclusion is given in Section 7.

2. Description of Presently Used Estimators. As mentioned earlier, both area and list frames are used by SRS to select samples for probability surveys.

The area frame for each state used by SRS is stratified by land use; for example, more than 75 percent cultivated, 50-74 percent cultivated, 15-49 percent cultivated, agriculture mixed with urban, and non-agricultural land. Each stratum is further subdivided into more homogeneous geographic substrata called paper strata. Segments (parcels of land) treated as the primary sampling units are selected as a simple stratified sample from each paper stratum. A detailed description on how the segments are constructed from aerial photographs with identifiable boundaries, how segment sizes and the number of segments are determined, and how the segments are selected via count units can be found in Houseman (1975) and Geuder (1984). For rotational purposes, the first segment selected in each paper stratum is designated as replicate 1, the second as replicate 2, etc. Approximately 20 percent of the segments are replaced annually on a rotational basis.

The list frame consisting of names of farmers is stratified by the size of farms contained in the control information. For example, for hogs and pigs inventory, typical strata are no hogs, 1-99 hogs, 100-199 hogs, 200-399 hogs, 400-999 hogs, 1000-2499 hogs, more than 2500 hogs. Systematic sampling from each stratum is usually used to select the list sample. See Section 5 of June Supervising and Editing Manual (1984).

For each area sample, there are three different methods of evaluating the farm inventory. A tract is a piece of land within the boundary of the segment under one management. A tract may be the entire farm if all of it is in the segment, or a portion of the farm, if the farm's boundary extends to outside of the segment. The area tract estimator is expanded by inventory on all the tracts of the selected segments. The area farm estimator is expanded by inventory on the corresponding farms provided the operator who resides in the segment. The area weighted estimator is computed from each farm inventory weighted by the ratio of tract acreage to farm acreage, regardless of the residency of the operator. There are no such complications for the list sample. The list sample uses the entire inventory of the farm.

Three different domains are needed to explain the four estimators presently in use. Domain D1, the nonoverlap domain, refers to the farms not in the list frame. (This domain is automatically in the area frame, since the area frame is complete.) Domain D2 refers to the farms in both frames but is not classified as "extreme operators." Domain D3 refers to the extreme operators in both frames. (Extreme operators are farmers with very large livestock inventories. The exact definition for the list sample in the Domain D3 will be given later.)

The operational tract, farm, and weighted estimators denoted by $\hat{Y}_1$, $\hat{Y}_2$, and $\hat{Y}_3$ can be expressed as follows:

$$\hat{Y}_i = \hat{Y}_{D1\cup D2,Ai} + \hat{Y}_{D3,L} \ , \qquad (2.1)$$

where i = 1, 2, or 3.

The estimator $\hat{Y}_{D1\cup D2,Ai}$ is computed by

$$\hat{Y}_{D1\cup D2,Ai} = \sum_{h\varepsilon H} e_h \sum_{k=1}^{n_h} y_{i,hk} \ , \qquad (2.2)$$

where H = the collection of paper strata,

e_h = the inverse of the probability of selection of each segment in the hth paper stratum,

n_h = the number of segments sampled in the hth paper stratum,

$$y_{1,hk} = \sum_{m=1}^{g_{hk}} t_{hkm}\delta_{hkm} \ ,$$

$$y_{2,hk} = \sum_{m=1}^{g_{hk}} f_{hkm}d_{hkm}\delta_{hkm} \ ,$$

$$y_{3,hk} = \sum_{m=1}^{g_{hk}} f_{hkm}(a_{hkm}/b_{hkm})\delta_{hkm} \ , \text{ with}$$

t_{hkm} = the value of the characteristic for the mth tract in kth segment of hth stratum

f_{hkm} = the value of the characteristic for the mth farm overlap with the kth segment of the hth stratum,

a_{hkm} = acreage of the hkmth tract,

b_{hkm} = acreage of the hkmth farm,

g_{hk} = total number of tracts in the hkth segment,

$$d_{hkm} = \begin{cases} 1 \text{ if the operator of hkm}^{th} \text{ farm resides} \\ \quad \text{in the hk}^{th} \text{ segment} \\ 0 \text{ otherwise ,} \end{cases}$$

$$\delta_{hkm} = \begin{cases} 1 \text{ if hkm}^{th} \text{ farm is in } D_1 \cup D_2 \\ 0 \text{ otherwise .} \end{cases}$$

The estimator $\hat{Y}_{D3,L}$ is computed from the list samples in the extreme operator (EO) strata:

$$\hat{Y}_{D3,L} = \Sigma_{\ell\varepsilon EO}(N_\ell/n_\ell) \sum_{k=1}^{n_\ell} y_{\ell k} \qquad (2.3)$$

where $y_{\ell k}$ = the value of the kth farm in the ℓth stratum,

N_ℓ = the population size of the ℓth stratum,

n_ℓ = the sample size of the ℓth stratum,

EO = collection of list strata with extreme operators.

The definition of EO strata from the list population depends on the state. For example, the EO strata for Indiana hogs consists of 3 strata defined by the size of the farms: 1000-1999 hogs, 2000-4999 hogs, and more than 5000 hogs. The biggest stratum is sampled with probability 1. The rest of the EO strata are sampled at the rate of approximately one-quarter and one-half, respectively, for each of the strata.

The above three estimators are area-oriented. The fourth estimator is list-oriented. A version of it can be written as

$$\hat{Y} = \hat{Y}_{D1,A3} + p'\hat{Y}_{D2,A3} + (1-p')\hat{Y}_{D2,L} + p\hat{Y}_{D3,A3} + (1-p)\hat{Y}_{D3,L}$$

where $\hat{Y}_{Di,A3}$ denotes the weighted area estimator for domain Di, and $\hat{Y}_{Di,L}$ denotes the list estimator for domain Di. The constants p and p' are set to zero in the present procedures. Therefore, the fourth estimator, called multiple frame screening estimator, is given by

$$\hat{Y}_4 = \hat{Y}_{D1,A3} + \hat{Y}_{D2,L} + \hat{Y}_{D3,L} \qquad (2.4)$$

The component $\hat{Y}_{D1,A3}$ is defined as

$$\hat{Y}_{D1,A3} = \sum_{h\varepsilon H} e_h \sum_{k=1}^{n_h} \sum_{m=1}^{g_{hk}} f_{hkm}(a_{hkm}/b_{hkm})\delta'_{hkm} , \qquad (2.5)$$

where

$$\delta'_{hkm} = \begin{cases} 1 & \text{if } hkm^{th} \text{ farm } \varepsilon \text{ D1} \\ 0 & \text{otherwise} \end{cases}$$

and all the other terms are defined as before.

The component $\hat{Y}_{D2,L}$ is defined as $\hat{Y}_{D3,L}$ in equation (2.3) except the summation is over $\ell \ \varepsilon \ EO^c$. The set EO^c denotes the collection of the list strata which are not the EO strata.

The estimators $\hat{Y}_i$, i = 1, 2, or 3, are basically derived from the area frame. However, the list estimator replaces the area estimator for the farmers classified as extreme operators. This perhaps could be interpreted as a robust procedure taken by SRS to reduce the influence of the big farms in the area sample. Further study of robust estimation in surveys is needed.

The variances and covariances of the four preliminary estimators denotes by $\hat{v}_{ij}$ are estimated by SRS and given in Kuo (1986a or b).

3. Composite Estimation. In this section, composite estimation is explained and is specialized to the SRS situation. A heuristic argument for composite estimator for the simplest case is given below.

Let us assume there are two independent and unbiased estimators $\hat{Y}_1$ and $\hat{Y}_2$ for the same parameter Y with known variances σ_1^2 and σ_2^2, respectively. Let us propose

$$\hat{Y}_c = c\hat{Y}_1 + (1-c)\hat{Y}_2 ,$$

where c is a constant with values between 0 and 1. Then

$$E\hat{Y}_c = Y$$

$$V(\hat{Y}_c) = c^2\sigma_1^2 + (1-c)^2\sigma_2^2 . \qquad (3.1)$$

To minimize $V(\hat{Y}_c)$, we should choose c to be

$$c_0 = \sigma_2^2/(\sigma_1^2 + \sigma_2^2) .$$

The minimal variance can be obtained from (3.1):

$$V(\hat{Y}_{c_0}) = \sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2) . \qquad (3.2)$$

Note that the expression of (3.2) is always smaller than σ_1^2 and σ_2^2. See Schaible (1982 and 1979) and Royall (1979, pp. 85-86) for more discussion on composite estimation.

In general, the variances of $\hat{Y}_1$ and $\hat{Y}_2$ are unknown. However, they can be estimated from the data. The estimated variances are denoted by $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$. Therefore, the composite estimator is given by

$$\hat{Y}_c = \hat{\sigma}_2^2(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)^{-1}\ \hat{Y}_1 + \hat{\sigma}_1^2(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)^{-1}\ \hat{Y}_2 .$$

Since the weight for the composite estimator is now a function of the data, equation (3.1) can no longer be used to evaluate the variance of the composite estimator. Nevertheless, the variance of the composite estimator can be estimated by sample reuse methods such as jackknife, bootstrap, random group, and balanced repeated replication.

To generalize the above idea to the situation at SRS, let us propose a family of linear combinations of the four preliminary estimators:

$$\hat{Y}_{\underset{\sim}{w}} = \sum_{i=1}^{4} w_i\hat{Y}_i \qquad (3.3)$$

where $0 \le w_i \le 1$ for all i and $\Sigma w_i = 1$.

We search for the one which minimizes the mean squared errors (MSE's) of the estimators in the linear family. Note that

$$\begin{aligned} f(\underset{\sim}{w}) &= \text{MSE of } \hat{Y}_{\underset{\sim}{w}} = E(\hat{Y}_{\underset{\sim}{w}} - \hat{Y})^2 \\ &= \sum_{i=1}^{4} w_i^2 E(\hat{Y}_i - Y)^2 + \sum\sum_{i\neq j} w_i w_j E(\hat{Y}_i - Y)(\hat{Y}_j - Y) \qquad (3.4) \end{aligned}$$

where Y denotes the population total.

Since all the second moment terms are unknown, they have to be estimated from the data. The estimation of the second moment terms will be treated in the next section. Let $\hat{m}_{ij}$ denote the estimated term $\hat{E}(\hat{Y}_i - Y)(\hat{Y}_j - Y)$. The composite estimator, denoted by $\hat{Y}_{\underset{\sim}{w}_0}$ is derived from minimizing

$$\hat{f}(\underset{\sim}{w}) = \sum_{i=1}^{4} \sum_{j=1}^{4} w_i w_j \hat{m}_{ij} \qquad (3.5)$$

subject to linear constraints $0 \le w_i \le 1$, for i = 1 to 4, and $\Sigma w_i = 1$.

A further refinement, motivated by the limited translation idea in Efron and Morris (1971, 1972) and Fay and Herriot (1979), is used to derive the final composition estimation. It depends on a "safety factor" K, a positive number specified in advance.

$$\hat{Y}_f = \begin{cases} \hat{Y}_{\underset{\sim}{w}_0} & , \text{ if } |\hat{Y}_{\underset{\sim}{w}_0} - \hat{Y}_1| < K\cdot SD(\hat{Y}_1) \\ \hat{Y}_1 - K\cdot SD(\hat{Y}_1), & \text{if } \hat{Y}_1 - \hat{Y}_{\underset{\sim}{w}_0} > K\cdot SD(\hat{Y}_1) \\ \hat{Y}_1 + K\cdot SD(\hat{Y}_1), & \text{if } \hat{Y}_1 - \hat{Y}_{\underset{\sim}{w}_0} < -K\cdot SD(\hat{Y}_1), \end{cases} \qquad (3.6)$$

where the estimated standard error $SD(\hat{Y}_1)$ is the square root of $\hat{v}_{11}$ given in Kuo (1986a or b).

This refinement, limiting the amount that composite estimator can deviate from the unbiased estimator, is employed to guard against instability. One can still achieve substantial gain from the composite estimation.

A program using Lagrange multipliers and the PROC MATRIX procedure in SAS (Statistical Analysis System) has been written by the author to solve equation (3.5), a convex programming problem with constraints. See the Appendix for a detailed explanation.

4. Estimation of Second Moments. Development of the estimation of the second moment terms which incorporates bias analysis is discussed in this section.

As is seen from equation (3.5), there are four MSE's and six mixed central moments to be estimated. To estimate these terms, it is assumed:

$$E\hat{Y}_i = Y + b_i(Y),\ i = 2,\ 3,\ \text{or}\ 4,$$

where $b_i(Y)$ denotes the bias of the i^{th} estimator.

To estimate the second moment terms, we use the following identity.

$$m_{ij} = E(\hat{Y}_i-\hat{Y}_1)(\hat{Y}_j-\hat{Y}_1) + Cov(\hat{Y}_1,\hat{Y}_i) + Cov(\hat{Y}_1,\hat{Y}_j) - V(\hat{Y}_1) \qquad (4.1)$$

for all i and j.

It is verified in Kuo (1986a) by inserting $\hat{Y}_1$ in the first step and then Y in the second step in the expansion of $E(\hat{Y}_i-Y)(\hat{Y}_j-Y)$.

Unbiased estimates of the mixed central moment terms, and refinements over the unbiased estimates of the MSE terms can be obtained as follows.

$$\hat{m}^2_{11} = \hat{v}_{11}\ , \qquad (4.2)$$

$$\hat{m}^2_{ii} = \max\{(\hat{Y}_i-\hat{Y}_1)^2 + 2\hat{v}_{1i} - \hat{v}_{11}, v_{ii}\},\quad i = 2,\ 3,\ \text{or}\ 4, \qquad (4.3)$$

$$\hat{m}_{1j} = \hat{v}_{1j},\ \text{for}\ j \neq 1, \qquad (4.4)$$

$$\hat{m}_{ij} = (\hat{Y}_i-\hat{Y}_1)(\hat{Y}_j-\hat{Y}_1) + \hat{v}_{1i} + \hat{v}_{1j} - \hat{v}_{11},\quad \text{for}\ i,j \neq 1, \qquad (4.5)$$

where relevant $\hat{v}_{ij}$'s are given in Section 2 of Kuo (1986a or b).

The maximum function in $\hat{m}^2_{ii}$ is employed to ensure that the estimators for the bias squared terms are nonnegative.

Equation (4.3) enables us to obtain an unbiased estimate of the bias squared term $b^2_i(Y)$, for $i \neq 1$. A refinement over this unbiased estimate is given by

$$\hat{b}^2_i = \max\{(\hat{Y}_i-\hat{Y}_1)^2 + 2\hat{v}_{1i} - \hat{v}_{11} - \hat{v}_{ii}, 0\}. \qquad (4.6)$$

5. Variance and Mean Squared Error Evaluation of the Composite Estimator.

5.1 Description of the Jackknife Method in General. The Heuristic argument for using composite estimation has been given. The variance and mean squared error estimates for the composite estimator are needed to justify the gain in using composite estimation. The jackknife method is used to estimate variance and mean squared error. This method is adopted because of its simplicity of explanation and ease of programming. See Efron (1982), Wolter (1985) for expositions on sample reuse methods.

Assume the data are divided into g independent groups. Let $\hat{Y}_{(k)}$ be an estimator derived from the data with k^{th} group deleted. The k^{th} pseudo-value of $\hat{Y}$ is defined to be $Y^*_{(k)} = g\hat{Y} - (g-1)\hat{Y}_{(k)}$, where $\hat{Y}$ is the estimator based on the full sample.

The jackknife estimator of the variance of $\hat{Y}$ is given by

$$v_j(\hat{Y}_Q) = g^{-1}(g-1)^{-1} \sum_{k=1}^{g} (Y^*_{(k)}-\bar{Y}^*)^2 \qquad (5.1)$$

where

$$\bar{Y}^* = \sum_{k=1}^{g} Y^*_{(k)}/g\ .$$

If $\hat{Y}$ is an estimator other than $\hat{Y}_1$, then the mean squared error of $\hat{Y}$ can also be estimated by the jackknife method.

$$MSE_J(\hat{Y}) = (\hat{Y} - \hat{Y}_1)^2 + 2g^{-1}(g - 1)^{-1} \cdot \sum_{k=1}^{g} (Y^*_{(k)}-\bar{Y}^*)(Y^*_{1(k)}-\bar{Y}^*) - g^{-1}(g-1)^{-1} \sum_{k=1}^{g} (Y^*_{1(k)}-\bar{Y}^*_1)^2 \qquad (5.2)$$

where $Y^*_{1(k)}$ is the k^{th} pseudo-value of $\hat{Y}_1$.

5.2 Special Construction of the Groups for SRS Data. Due to the different constructions by SRS of the replication codes of the area and list samples, formulations of the groups of data for the jackknife method are slightly different between the area and the list sample. The replication codes in the area sample, which usually run from 1 to 10 or 1 to 5 for each land use stratum, are used. A land use stratum defined at the beginning of Section 2 is a collection of paper strata. The replication codes for each list stratum were generated by the author using random numbers. Several (3 or 4) replications are constructed for each list stratum.

The k^{th} $(1 \leq k \leq d$ where $d \leq g)$ jackknife estimate is computed by deleting the k^{th} replicate of each land use stratum (i.e., deleting each segment from each paper stratum in the same land use stratum). The expansion factor e_h is adjusted by multiplying the number of replicates and dividing by the number of replicates - 1 in each land use stratum. The number d is the total number of replicates for all land use strata.

The k^{th} $(d + 1 \leq k \leq g)$ jackknife estimate is computed by deleting each replicate from each list stratum sequentially. The adjustment on the expansion factor is discussed in Kuo (1986a or b).

No data are deleted from the self-representing stratum (the largest EO stratum).

Variance and mean squared error evaluations for Indiana and Minnesota are given in this paper. The numbers d and g for Indiana are 31 and 55, i.e., the area sample is divided into 31 groups, the nonself-representing list sample is divided into 24 approximately independent groups (6 strata with 4 groups each). The numbers d and g for Minnesota are 30 and 60, where 30 approximately independent groups from the list sample are derived from 10 strata with 3 groups in each stratum.

Empirical evaluations for the two states also reveal that the variance estimates for the composite estimators are more sensitive to outliers from the pseduo-values. Consequently, the Winsorized variance estimates and the mean squared error estimates are used here. Ten percent from each end of the pseudo-values of Indiana's data and 15 percent from that of Minnesota are Winsorized to obtain the variance and covariance estimates.

6. Numerical Results. The data are from the 1984 June Enumerative Survey conducted by SRS. Summary statistics for Indiana and Minnesota are given in Tables 1 and 2. Numerical results for four more states, Iowa, Kansas, Missouri, and Ohio, are given in Kuo (1986a or b).

Seven estimates, denoted by $\hat{Y}_i$, $i = 1,\ldots,7$, are given for the total hogs and pigs inventory. The estimates $\hat{Y}_i$, $i = 1,\ldots,4$, are the tract, farm, weighted and multiple frame screening estimates defined before. The estimate $\hat{Y}_5$ is the composite estimate defined by (3.6) for any $K \geq 1$. The estimate $\hat{Y}_6$ is derived similarly to $\hat{Y}_5$ except by setting $w_2 = w_3 = 0$. The estimate $\hat{Y}_7$ denotes the official CRB statistics published in the Livestock Series: Hogs and Pigs (Crop Reporting Board, 1984). The optimal weights (denoted by $\hat{w}$) for the components of $\hat{Y}_5$ derived from equation (3.5) are given in the tables. The optimal weights for $\hat{Y}_6$ are denoted by $\tilde{w}$.

All the standard errors and root mean squared errors of the four preliminary estimators are obtained by taking square roots of $\hat{v}_{ii}$ and $\hat{m}^2_{ii}$. These estimates are given in the tables denoted by SD_i and $\sqrt{MSQ_i}$. Two estimates of bias of $\hat{Y}_i$, $i = 2, 3, 4$, are given. One is obtained from equation (4.6), i.e., $\hat{b}_i$ or $-\hat{b}_i$. A second estimate is an unbiased estimate of the bias, i.e., $\hat{\hat{b}}_i = \hat{Y}_i - \hat{Y}_1$. The mean squared error matrix of the four estimators with entries $\hat{m}_{ij}$ is also given in each of the tables.

The jackknife method can be used for any estimators from probability surveys. Therefore, for each of the estimators $\hat{Y}_i$, $i = 1,\ldots,6$, we can compute its variance and mean squared error estimates as in Section (5.1) with $\hat{Y}$ replaced by $\hat{Y}_i$. The Winsorized variance estimates and mean squared error estimates are used in Tables 1 and 2. They are denoted by $SDJKR_i$ and $\sqrt{MSQJKR_i}$, $i = 1,\ldots,6$. When $i = 1,\ldots,4$, the quantities $SDJKR_i$ and $\sqrt{MSQJKR_i}$ could be compared to SD_i and $\sqrt{MSQ_i}$, which are computed from the full sample using the stratified design, to determine the goodness of the variance estimates by the jackknife method. All the numbers here except $\hat{w}_i$, $\tilde{w}_i$ and $\hat{m}_{ij}$ are expressed in thousands of heads for easier examinations.

Further research and improvement on the variance estimators by sample reused methods will be needed. However, the numerical results in each of the tables present enough evidence to show that the composite estimator performs very well. Examining the mean squared errors and mixed moments of the preliminary estimators, it can be seen that the composite estimate is very effective in selecting the desirable components, i.e., the components with small mean squared errors or the components which are negatively correlated. Numerical results for four more states given in Kuo (1986a or b) also reveal similar performances for the composite estimator.

7. Conclusion. This paper provides a composite estimation methodology which combines the different preliminary estimators used by SRS into one by minimizing the mean squared errors of the combined estimators. Some numerical results from the 1984 June Enumerative Survey are presented. Further research on related topics is discussed in Kuo (1986a or b).

APPENDIX: Convex Programming to Search for Composite Weights. From equation (3.5) in the text, we need to minimize

$$\hat{f}(\underset{\sim}{w}) = \Sigma\Sigma w_i w_j \hat{m}_{ij} \tag{A.1}$$

subject to $0 \leq w_i \leq 1$ for all i, and $\sum_{i=1}^{4} w_i = 1$. Let the inequality constraint functions be denoted by $g_i(\underset{\sim}{w}) = w_i$ for all i. The Lagrange multiplier technique is applied. A necessary and sufficient condition for the minimum to exist is as follows (see Avriel (1976)): There exists u, w_i and positive constants λ_i, $i = 1,\ldots,4$, such that

$$\begin{cases} \lambda_i g_i(\underset{\sim}{w}) = 0 \text{ for all } i \\ \nabla\hat{f}(w) - \sum_i \lambda_i \nabla g_i(\underset{\sim}{w}) - u\nabla(\Sigma w_i - 1) = 0, \end{cases} \tag{A.2}$$

and the w's satisfy the constraints of (A.1). There are nine equations with nine variables (four w's, four λ's, and one u) in equation (A.2) to be solved simultaneously. The function SOLVE in the PROC MATRIX procedure is used to solve them. Basically, the program searches for the minimal value of (A.1) among all cases of the possible combinations of the preliminary estimators: combination of all the four, three at a time, two at a time, and just the preliminary estimators. A more detailed programming explanation is given in Kuo (1986a or b).

Table 1. Summary Statistics for Indiana

										$\hat{m}_{ij}$			
i	$\hat{Y}_i$ (1,000)	$\hat{w}_i$	$\tilde{w}_i$ (1,000)	SD_i (1,000)	$SDJKR_i$ (1,000)	$\sqrt{MSQ_i}$ (1,000)	$\sqrt{MSQJKR_i}$ (1,000)	$\hat{b}_i$ (1,000)	$\hat{\hat{b}}_i$ (1,000)	1	2	3	4
1	3,367	0	0.829	413	349	413	349	0	0	1.703E11			
2	3,797	0		470	507	588	568	352	430	1.654E11	3.454E11		
3	3,616	1		249	205	249	205	0	249	7.007E10	1.723E11	6.213E10	
4	4,331	0	0.171	178	183	884	909	866	964	1.171E10	4.212E11	1.516E11	7.819E11
5	3,616				253		279						
6	3,532				477		484						
7	4,300												

Table 2. Summary Statistics for Minnesota

										$\hat{m}_{ij}$			
i	$\hat{Y}_i$ (1,000)	$\hat{w}_i$	$\tilde{w}_i$	SD_i (1,000)	$SDJKR_i$ (1,000)	$\sqrt{MSQ_i}$ (1,000)	$\sqrt{MSQJKR_i}$ (1,000)	$\hat{b}_i$ (1,000)	$\hat{\hat{b}}_i$ (1,000)	1	2	3	4
1	4,899	0	0.652	699	551	699	551	0	0	4.883E11			
2	5,226	0.233		716	765	753	765	234	328	4.741E11	5.673E11		
3	4,645	0.563		395	483	395	483	0	-254	2.151E11	1.179E11	1.559E11	
4	3,753	0.204	0.348	237	279	941	1,054	-911	-1,145	3,116E10	-3.581E11	4.600E10	8.862E11
5	4,598				665		710						
6	4,500				686		777						
7	3,870												

REFERENCES

Avriel, M. (1976). Nonlinear Programming: Analysis. New Jersey: Prentice-Hall.

Brock, D.B., D.K. French & B.W. Peyton (1980). Small Area Estimation: Empirical Evaluation of Several Estimators for Primary Sampling Units. Proceedings, Section of Survey Research Methods. American Statistical Association, 766-771.

Brown, L.D. & A. Cohen (1974). Point and Confidence Estimation of a Common Mean and Recovery of Interblock Information. Annals of Statistics, 2, 963-976.

Cohen, M.P. & J.P. Sommers (1984). Evaluation of Methods of Composite Estimation of Cost Weights for the CPI. Proceedings, Section on Business and Economics, Amer. Statist. Assoc., 466-471.

Crop Reporting Board (1984). Livestock Series, Hogs and Pigs. Statist. Report Service, USDA.

Efron, B. & C. Morris (1971)(1972). Limiting the Risk of Bayes and Empirical Bayes Estimators--Part I: The Bayes Case; Part II: The Empirical Bayes Case. JASA 66, 807-815; 67, 117-130.

Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia: Society for Industrial and Applied Mathematics.

Fay, R.E. III & R.A. Herriot (1979). Estimates of Income in Small Places: An Application of James-Stein Procedures to Census Data. JASA 74, 269-277.

Geuder, J. (1984). Paper Stratification in SRS Area Sampling Frames. Statistical Reporting Service, USDA.

Graybill, F.A. & R.B. Deal (1959). Combining Unbiased Estimators. Biometrics, 15, 543-550.

Houseman, E. (1975). Area Frame Sampling in Agriculture. Statist. Reporting Serv., USDA.

Kuo, L. (1986a and b). Composite Estimation of Totals for Livestock Surveys. a) Statist. Rep. Serv., USDA; b) Tech. Rep. 74, U.Calif., Davis.

Mosteller, F. (1948). On Pooling Data. JASA, 43, 231-242.

Nealon, J. (1984). Review of the Multiple and Area Frame Estimators. Stat. Rep. Serv., USDA.

Royall, R.M. (1979). Prediction Models in Small Area Estimation, in Synthetic Estimators for Small Areas: Statistical Workshop Papers and Discussion. NIDA Research Monograph 24, U.S. Government Printing Office, 63-87.

Schaible, W.J. (1978). Choosing Weights for Composite Estimators for Small Area Statistics. Proceedings, Section on Survey Research Mathods, JASA, 741-746.

Schaible, W. (1979). A Composite Estimator for Small Areas (with discussion), in Synthetic Estimators for Small Areas: Statistical Workshop Papers and Discussion. U.S. Government Printing Office, 36-62.

Statistical Reporting Service (1984). June Supervising and Editing Manual. USDA.

Wolter, K. (1979). Composite Estimation in Finite Populations. JASA, 74, 604-613.

Wolter, K. (1985). Variance Estimation. New York: Springer-Verlag.