United States
Department of
Agriculture

Statistical
Reporting
Service

**Statistical
Research
Division**

**SF&SRB Staff
Report No. 80** 84

Washington, D.C.

March 1984

Review of the Multiple and Area Frame Estimators

John Patrick Nealon

REVIEW OF THE MULTIPLE AND AREA FRAME ESTIMATORS. By John Patrick Nealon; Statistical Research Division; Statistical Reporting Service; U.S. Department of Agriculture; Washington, D.C. 20250; March 1984. SF&SRB Staff Report No. 80.

ABSTRACT

This report discusses in detail the multiple and area frame estimators used by the Statistical Reporting Service (SRS). The advantages and disadvantages of each estimator, particularly with respect to livestock estimation, are presented as well as suggestions for reducing the nonsampling errors associated with each estimator. The livestock estimates from each estimator since 1978 are compared with SRS's official statistics in order to ascertain how much SRS has relied on each estimator. Finally, recommendations are made concerning the livestock estimation program.

CONTENTS

SUMMARY

A unified discussion of the open, closed, weighted and multiple frame estimators is presented in this report. An overview is given not only of the origin of the four estimators in the estimation programs of the Statistical Reporting Service (SRS), but also of the research conducted with these estimators since the mid-1950's. A thorough presentation of the advantages and disadvantages associated with each estimator, particularly with respect to livestock estimation, is given. Also, suggestions for reducing the nonsampling errors inherent in the estimators are outlined.

The estimates from the four estimators were compared with SRS's official statistics to determine how much SRS has relied on each estimator historically when estimating livestock inventories. The analysis showed that SRS has relied on the multiple frame livestock estimator much more than the area frame livestock estimators. Of the three area frame estimators, the closed estimator has been relied on slightly more than the open and weighted estimators for cattle, and the weighted estimator has been used the most for hogs.

Review of the Multiple and Area Frame Estimators

John Patrick Nealon

INTRODUCTION

Almost all statistics published by the Statistical Reporting Service (SRS), U.S. Department of Agriculture, prior to the 1960's were based on nonprobability surveys. Over the past 20 years, SRS has developed and implemented a series of probability surveys to provide estimates for crop acreages, grain stocks, livestock numbers, and other agricultural items such as farm production expenditures and farm labor. These probability surveys are based on area and multiple frame survey designs and the basic estimators generated from these designs are referred to as direct expansion estimators. There are four direct expansion estimators available from area and multiple frame survey designs, namely, the open, closed, weighted and multiple frame estimators. The open and closed estimators are also commonly referred to as the farm and tract estimators, respectively.

One of these estimators-- the closed estimator -- is available to statisticians to arrive at SRS's official statistics for crop acreages at the state, regional and national levels. Also, one direct expansion estimator-- the multiple frame estimator -- is used by statisticians to produce the official statistics for on-farm grain stocks, farm production expenditures and farm labor. On the other hand, statisticians have at their disposal two to four direct expansion estimates from area and multiple frame surveys to generate the official statistics for livestock inventories.

Statisticians combine the various livestock estimates into an official statistic using a subjective process. For example, statisticians in some states may rely almost exclusively on the multiple frame estimate to arrive at the statistic they recommend to the Crop Reporting Board in SRS. On the other hand, statisticians in other states may base their recommendation on some compromise of the available estimates. In 1983, a report entitled "Framework for the Future", which was prepared by the Long-Range Planning Group, recommended that SRS develop an objective procedure for combining the various probability-based estimates into a single, combined estimate for setting each official statistic (1). A review of the multiple and area frame estimators in the livestock program is therefore needed to shed light on which estimators should be given more weight in an objective weighting scheme.

The Long-Range Planning Group also recommended restructuring SRS's current series of surveys into an Integrated Survey Program (ISP). If SRS shifts to an ISP, the number of estimators generated for livestock items in each state might be reduced or changed. Therefore, a unified review of the multiple and area frame estimators in the livestock program is also needed to determine which estimators should be included in the ISP for livestock.

Over the years, research has been conducted to evaluate the multiple and area frame estimators used by SRS with special attention given to livestock estimation. Many of the advantages and disadvantages of each estimator have been mentioned in various research reports. However, no documentation exists that discusses how well the multiple and area frame livestock estimates have compared historically with the official statistics. The Survey Research Section decided to conduct a formal review of the multiple and area frame estimators in the livestock program with the following three objectives in mind:

(1) Present a unified discussion of the advantages and disadvantages for each of the four estimators, focusing primarily on livestock estimation. In addition, recommend procedures, when possible, for reducing the effects of the disadvantages.

(2) Determine how much SRS has relied on each of the four livestock estimators. This will be measured by how close the livestock estimates have been historically to the official statistics and how correlated the estimates from each of the four estimators have been with the official statistics in the time-trend charts used by SRS.

(3) Recommend a course for the future regarding which estimator(s) should be used in the livestock program.

ORIGIN OF THE ESTIMATORS

SRS's involvement in area frame sampling methods dates back to 1954 when research was conducted to evaluate estimates of crop acreages using 703 segments in 10 states (40). By 1965, area frame sampling methods had been expanded to the 48 conterminous states. In that year, the first national area frame survey was conducted, providing crop acreage estimates at the state, regional and national levels. In 1967, the estimation program was expanded to a series of surveys to provide area frame estimates for crops, livestock and other agricultural items (40). Two estimators, the open and closed estimators, were used in the operational program. Both of these estimators were used for livestock estimation while only the closed estimator was used for acreage estimation. The open estimator was relied on exclusively to obtain estimates for items only associated with the farm as a whole such as farm size, income and labor.

In the early 1960's, the theory of multiple frame sampling and estimation was developed at Iowa State University through a cooperative agreement with SRS (46). Multiple frame methods had some distinct advantages for SRS's applications, particularly for livestock estimation, since many survey items such as hog and cattle inventories were not efficiently estimated by the area frame. Multiple frame methods were first evaluated by SRS during 1965 in Wyoming (8) and were introduced into the livestock estimation program in four states beginning in 1969 (39). By 1974, multiple frame livestock estimation had expanded to include 14 states in the hog program and 28 states in the cattle program, representing 85 to 90 percent of the U.S. hog and cattle inventories. In addition, multiple frame methods were implemented nationally by 1975 to provide estimates for farm production expenditures and farm labor. Finally, multiple frame methods were also developed in the 1970's to provide acreage estimates in some states for specialty crops such as potatoes and rice.

A third area frame estimator-- the weighted estimator --was introduced and evaluated in 1962 (38). This estimator was later used in the multiple frame livestock program to provide estimates for the nonoverlap domain. Weighted estimates based solely on the area frame were integrated into the hog and cattle estimation program during 1977 and included only the 14 hog multiple frame states. Starting in 1979, the weighted estimator was also used to estimate the number of farms in these 14 states.

Moving to the present--1984--the closed estimator is still the only estimator used in the national crop acreage estimation program. Also, the multiple frame estimator is the only estimator in the estimation programs for farm production expenditures, farm labor and on-farm grain stocks. To estimate the number of farms, both the open and weighted estimators are used in 10 states while only the open estimator is used in the remaining 38 states. The livestock estimation program is a much more complex program. All 48 states collect data for the open and closed estimators for hogs and cattle, 23 states also have a multiple frame hog estimator, 34 states have a multiple frame cattle estimator, and 10 of the multiple frame states also provide weighted hog and cattle estimates from the area frame.

OVERVIEW OF PREVIOUS RESEARCH

During the early years of area frame methods in SRS, research activities in the area of estimation focused on comparing the estimates obtained from the open and closed estimators for crop acreages and livestock inventories. As early as 1956, SRS decided to abandon the use of the open estimator for crop acreage estimation since the closed estimator was considered superior (38).

Several research projects were conducted from the mid-1950's to the mid-1960's to compare the open and closed estimators for estimating livestock inventories (35,36,37,45). These studies concluded that livestock inventories were more accurately reported using the closed segment approach. The open livestock estimates were believed to be consistently underestimating livestock inventories while the closed estimates, particularly for cattle, were considered to be estimating livestock inventories correctly. In 1966, the Standards and Research Division recommended that SRS use the closed estimator to estimate livestock inventories (45).

In the latter half of the 1960's, research efforts were directed toward multiple frame sampling methods with special emphasis on livestock estimation. Research was primarily concerned with appraising the multiple frame livestock estimates, evaluating available list frame sources, testing procedures for building and updating list frames, and investigating alternative schemes for domain determination (25,26,27,28,29, 34,42,43,53).

During the next decade, SRS undertook the difficult task of developing a general-purpose list frame of all known farmers in each state for use in the multiple frame and nonprobability mail surveys. As the Research Division expanded in the 1970's, research into multiple frame livestock methods accelerated (3). Research efforts concentrated on identifying nonsampling errors associated with multiple frame procedures and recommending procedures to minimize these errors. A wide range of multiple frame survey procedures was evaluated. These included the design of the questionnaires (14,24,30,31,41,47), the adherence to survey concepts (6,22), problems associated with domain determination (22,48,49), optimization of the size for the list frame (7,51), effects of nonresponse (9,10,15,16,17,18), biases in the estimation of the weighted nonoverlap domain (21), rotation group bias (13,19), and the effects of data collection procedures (2,20).

Finally, as SRS entered the 1980's, research with multiple and area frame estimation has centered upon investigating alternative weighted estimators (4,11,32). This work has been pursued since research has repeatedly shown that the weight in the operational weighted estimator is not obtained accurately (6,12,21,33,52).

ESTIMATORS

Each of the four estimators--open, closed, weighted and multiple frame--will be discussed in this section. The discussion for each estimator will include a definition of the estimator, the formula used to arrive at the direct expansion estimate, the advantages and disadvantages of each estimator, particularly with respect to livestock estimation, and suggestions for reducing the adverse effects of some of the disadvantages.

The open, closed and weighted estimates generated for livestock items are loosely referred to as area frame estimates by SRS when in reality they are multiple frame estimates. The area frame sample in all states is supplemented by a list of large hog or cattle operators, known as extreme operators. The area frame data from each extreme operator in the sample is set to zero when computing the area frame estimate, and then the list frame estimate for the extreme operators is added to this area frame estimate to arrive at what SRS refers to as an area frame livestock estimate. This distinct characteristic of the area frame livestock estimates should be kept in mind since it differs from the traditional area frame estimates used for crop acreages, farm size and farm numbers, which are based solely on the area frame. In this report, the precision levels stated for the area frame livestock estimates will always refer to the area frame estimates that have input from the extreme operators. These estimates will be more precise than the traditional area frame livestock estimates. On the other hand, the definitions, formulas, advantages and disadvantages presented in this section for each area frame estimator will pertain to the traditional area frame estimators, which have no list frame component.

Open

The open estimator requires that data associated with the total acres operated be obtained from each agricultural operation which satisfies the "headquarters" rule. For SRS surveys, the headquarters rule is satisfied if the operator of the agricultural operation resides in the sampled segment. Therefore, data are collected only from resident agricultural operators (RAO's). To arrive at an estimate for the state, region or nation using the open segment approach, the data pertaining to the total acres operated for each RAO is first added to the segment level and multiplied by the inverse of the probability of selection for the segment, which is called the expansion factor. Then, the expanded segment totals from the RAO's are added together to arrive at the estimate for the desired inference level such as the state.

The sample estimate of a total from the open estimator for SRS's area frame design can be expressed in the following notation for a given state:

$$\hat{Y}_o = \sum_{i=1}^{s} \sum_{j=1}^{p_i} \sum_{k=1}^{r_{ij}} y_{ijk}$$

where

$$y'_{ijk} = \begin{cases} e_{ijk} \sum_{\ell=1}^{g_{ijk}} y_{ijk\ell} & \text{if } g_{ijk} > 0, \\ 0 & \text{if } g_{ijk} = 0, \end{cases}$$

g_{ijk} = the number of RAO's in segment k, paper stratum j and land-use stratum i,

$y_{ijk\ell}$ = the value of the survey item on the total acres operated for the ℓ^{th} RAO in the k^{th} segment, j^{th} paper stratum and i^{th} land-use stratum,

e_{ijk} = the expansion factor for the k^{th} segment in the j^{th} paper stratum in the i^{th} land-use stratum,

r_{ij} = the number of sample replicates or segments in the j^{th} paper stratum in the i^{th} land-use stratum,

p_i = the number of paper strata in the i^{th} land-use stratum,

s = the number of land-use strata in the state.

There are three advantages associated with the open estimator. These advantages are:

(A1) Data Collection Savings: Only RAO's need to be interviewed when using the open estimator. On the other hand, the closed and weighted estimators require that information be obtained from all agricultural operators with any land in the sampled segments regardless of where they reside. The exact cost savings obtained by using the open approach instead of the closed or weighted approach are not known. For SRS's area frame design, the savings are subjectively estimated to be between 5 and 20 percent depending on the state.

(A2) Flexibility: The reporting unit for the open estimator is the "total land operated". Since all agricultural items can be associated with the farm operation as a whole such as farm income and cattle deaths, the open estimator can be used to estimate for any agricultural item. For some survey items such as gross value of sales, the operator is only able to report accurately if the reporting unit is the total land operated.

(A3) Familiar Reporting Unit: Operators are generally accustomed to thinking in terms of the total land they operate and to providing information about their entire operation.

There are numerous disadvantages associated with the open estimator that collectively outweigh the advantages. These disadvantages are:

(D1) Lack of Precision: The open estimator is the most imprecise of the four estimators. Estimates based on the open segment approach often lack the precision needed to set

reliable state statistics. For example, Table 1 shows the distributions of the coefficient of variation (C.V.) for three important hog items based on the 1983 June Enumerative Survey (JES). Only the 23 multiple frame hog states, which contain about 95 percent of the nation's hogs and pigs, are included in this table. The C.V.'s for the three hog items were almost always in double-digits. With C.V.'s at these high levels, estimates often fluctuate too much from year to year, making time trends difficult to recognize when setting official statistics. However, the estimates for many hog items are precise (C.V. less than 5 percent) when the inference level is national or when the 23 multiple frame states are combined. For example, the national C.V.'s from the 1983 JES for total hogs and pigs, farrowing intentions for the next quarter, and the previous quarter's farrowings were 3.8, 4.2 and 4.3 percent, respectively.

Table 1: The distributions of the coefficient of variation from the open estimator for three hog items in the 23 multiple frame hog states; 1983 JES.

Coefficient of Variation (%)	Total Hogs and Pigs	Farrowing Intentions (June-August)	Previous Farrowings (March-May)
		Number of States	
< 5.0	0	0	0
5.0-9.9	1	1	1
10.0-19.9	15	12	15
20.0-29.9	5	7	3
≥ 30.0	2	3	4
Total	23	23	23

Cattle estimates are more precise than hog estimates, but still not precise enough in many of the major cattle producing states. Table 2 gives the distributions of the C.V. from the 1983 JES for three important cattle items using the 28 original multiple frame states, which account for almost 90 percent of the U.S. cattle. Many of the C.V.'s were greater than 10 percent at the state level for these items. For the 1982 December Enumerative Survey (DES), even more of the states had C.V.'s in the double-digits for these three cattle items. Cattle estimates, however, are precise at the national level or for the 28 multiple frame states combined. For example, the national C.V.'s from the 1983 JES for total cattle and calves, calves born since 1983, and cows and heifers remaining to calve in 1983 were 2.3, 2.9 and 3.1 percent, respectively.

Table 2: The distributions of the coefficient of variation from the open estimator for three cattle items in the 28 multiple frame cattle states; 1983 JES.

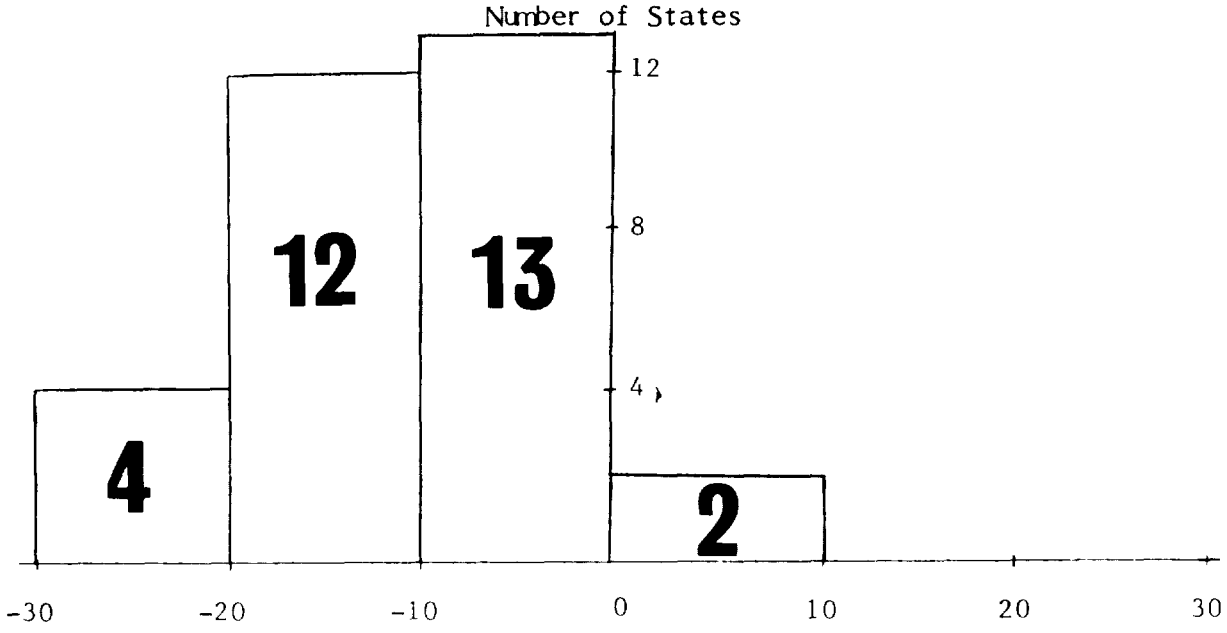
Coefficient of Variation (%)	Total Cattle and Calves	Calves Born Since 1983	Cows and Heifers Remaining to Calve in 1983
		Number of States	
<5.0	0	0	0
5.0-9.9	14	10	6
10.0-19.9	14	16	18
20.0-29.9	0	2	4
≥30.0	0	0	0
Total	28	28	28

(D2) Association Errors: As mentioned earlier, the open estimator depends on data from all RAO's. Therefore, this estimator requires that the correct operator be identified for each operation and that the operator be properly classified as living inside or outside the segment. Enumerators who participated in the 1982 panel studies (52) mentioned that they sometimes identify the wrong person as the operator and discussed the difficulty of determining the proper operator for complex management structures such as livestock-only partnerships and corporations. Nonsampling errors occur when the association between the farm operation and the location of the operator's headquarters is not properly linked. The effect of association errors on the estimates is not known.

(D3) Undercoverage: Experience with the open segment approach has revealed a strong tendency toward undercoverage (23). Operators of small farming operations do not always view themselves as farm operators and, therefore, may be classified as non-farm. Enumerators in the 1982 panel studies discussed the difficulty of detecting small operations (52). Also, multiple land-operating arrangements within the same household can go undetected. Another source of undercoverage exists in residential, industrial and commercial areas because these areas are very difficult to enumerate thoroughly. The effect of the undercoverage varies from one survey item to another. For example, when a small farm is not detected, the estimate of the number of farms is affected more than the livestock estimates. The undercoverage problem associated with the open estimator can be illustrated by examining the estimates for farm numbers. The national open estimate for the number of farms from the 1983 JES was 9.7 percent below the official statistic, which was more than eight standard errors below the official statistic. Figure 1 illustrates the undercoverage for the 31

states with at least 25,000 farms. These states contain over 90 percent of the U.S. farms. This figure shows the frequency distribution of the relative difference between the 1983 JES open estimate and the official statistic for the number of farms. Notice that the number of farms was undercounted in 29 of the 31 states.

Figure 1: The distribution of the relative difference between the open estimate and the official statistic for farm numbers in the 31 states with at least 25,000 farms; 1983 JES.



Relative Difference = 100* (Open - Official)/Official

(D4) Reporting Errors: Data for the entire farm are susceptible to reporting errors if the respondent is not knowledgable about the entire operation or does not understand the survey concepts, e.g. reporting regardless of ownership. For example, a small scale reinterview study in Nebraska (6) after the 1974 JES found that 34.0 percent of the JES responses for total acres operated, 28.0 percent for total hogs and pigs, and 36.7 percent for total cattle and calves were incorrect. Response errors associated with the reporting of hogs and cattle on the entire farm were also found in a three-state study in 1976 (21). Many of the reporting errors associated with livestock have resulted because of the failure to report all livestock on land operated regardless of ownership. Studies have shown (6,45) that the respondents' inclination is to report only their own livestock regardless of where the livestock are located.

(D5) Subjective Imputations: When an enumerator encounters a nonrespondent, visual inspection or observation is restricted since the reporting unit is the total acres operated. Therefore, the enumerator and office personnel often have no current information about the entire operation. If data are imputed manually, as is the case with SRS's area frame surveys, the imputation process is generally very subjective and difficult.

(D6) Sensitive Nature of the Reporting Unit: The operator is asked to provide data pertaining to the total land operated not only for the open estimator, but also for the weighted and multiple frame estimators that will be discussed later. The author believes that operators are more reluctant or more likely to refuse to supply information for their entire operation than for only a part of their operation.

Continual efforts should be made to minimize the adverse effects of these disadvantages. Association errors--identifying the wrong RAO's--result from data collection mistakes. As recommended by the 1982 panel studies (52), the JES questionnaire should be redesigned to identify the operator as early as possible in the interview. This action might reduce the number of association errors. Training schools and the enumerator's manual should emphasize why the correct identification of RAO's is essential to the validity of the open estimator. The 1982 panel studies also revealed that enumerators are confused on how to identify the operator for complex management structures such as corporations. These problems need to be addressed in greater detail at the training schools and in the enumerator's manual.

Reporting errors often result because the respondent does not understand the survey concepts. Important concepts such as why farmers need to report all livestock regardless of ownership should be emphasized to the enumerators. In addition, enumerators should be instructed to familiarize the respondents with these concepts.

Segment canvassing procedures should be stressed at the training schools to address the undercoverage problem. Enumerators discussed during the panel studies the difficulties of detecting small operations. Input should be sought from the enumerators on how to deal with this issue. Also, in nonrural areas where undercounts are believed to be more prevalent, current screening procedures should be reviewed. Finally, some enumerators in the panel studies commented that the questions used to detect multiple operations within the same household are difficult for respondents to understand. Input from the enumerators should be sought to determine if these questions can be improved.

Missing entire farm data due to nonrespondents will always exist and will affect the estimator if the imputed information

differs from reality. The operational imputation procedure lacks uniformity not only across states but also within states. Therefore, objective imputation methods should be investigated. Research in this area began during the 1983 JES in six states.

Finally, the lack of precision at the state level for the open estimator can only be addressed by large increases in the number of segments sampled. As long as budget restraints limit the size of the sample, the open estimator will continue to provide imprecise state estimates for most survey items.

Closed

The closed estimator requires that data be obtained from each agricultural operation with land in a sampled segment. The land within the segment associated with an operation is referred to as a tract, and the operator of a tract is called a tract operator. The reporting unit for the closed segment approach is the tract acres operated. That is, the information reported by the tract operator pertains only to what is in the tract, regardless of ownership. The closed estimate of a total for a survey item is derived by: (1) adding the tract data to the segment level, (2) multiplying the segment total by the expansion factor for the segment, and (3) adding the expanded segment totals to obtain the estimate for the desired inference level.

The formula for computing the state estimate of a total for a survey item based on the closed estimator is:

$$\hat{Y}_c = \sum_{i=1}^{s} \sum_{j=1}^{p_i} \sum_{k=1}^{r_{ij}} y'_{ijk}$$

where

$$y'_{ijk} = \begin{cases} e_{ijk} \sum_{\ell=1}^{f_{ijk}} t_{ijk\ell} & \text{if } f_{ijk} > 0, \\ 0 & \text{if } f_{ijk} = 0, \end{cases}$$

f_{ijk} = the number of tract operations in the k^{th} segment, j^{th} paper stratum and i^{th} land-use stratum,

$t_{ijk\ell}$ = the value of the survey item on the total tract acres operated for the ℓ^{th} tract operation in the k^{th} segment, j^{th} paper stratum and i^{th} land-use stratum,

e_{ijk}, s, p_i and r_{ij} are defined the same as with the open estimator.

Difficulties with the open estimator led to the use of the closed estimator whenever it was possible for the reporting unit to be the tract acres operated. There are numerous advantages associated with the closed segment approach. These are:

(A1) Simplicity: The closed approach is generally less troublesome for enumerators to apply since the enumerators do not have to account for the entire operation and do not have to be concerned with issues such as determining whether the operator satisfies the headquarters rule.

(A2) Not Affected By Association Errors: The closed segment approach avoids the issue of establishing the location of the operator's headquarters. Therefore, association errors, which affect the open estimates, have no effect on the closed estimates.

(A3) Visual Verification: By observing the land associated with the tract, enumerators can detect gross response errors. This advantage is unique to the closed segment approach.

(A4) Coverage Errors Minimized: The closed segment approach, where applicable, has proven to be far superior to the open approach with regard to coverage errors (23). The closed approach avoids most of the coverage problems caused by the ambiguity of what is a farm, the unclear linkage between operators and farms, and the accounting of all operations.

(A5) Observable Data: The enumerator is in a position to obtain some tract information for nonrespondents by observing the tract. Therefore, office personnel often are provided some information to assist them with the data imputations for nonrespondents. The quality of observed data depends on the survey item. For example, accurate crop acreages can be obtained by observation while livestock inventories can be more difficult to observe and count.

(A6) Less Sensitive: The operator is only asked to report for part of the operation in most instances using the closed approach. The author believes that respondents are less reluctant to provide the data since the information does not pertain to the entire operation. Also, if the operator is not available, surrogate respondents such as hired hands are sometimes willing to provide the tract data, but are either unwilling or unable to provide data for the entire operation. Therefore, response rates are probably higher with the closed segment approach.

(A7) Fewer Reporting Errors: In the earliest years of area frame research in SRS, it was concluded that livestock inventories and especially crop acreages for the tract were susceptible to fewer reporting errors than the responses for the farm. Research during 1974 in Nebraska (6) showed that

there were more reporting errors for farm acres, farm cattle and farm hogs than their corresponding tract values. A reinterview study in Indiana, North Carolina and Oklahoma during 1976 (21) showed that there were more reporting errors associated with farm hogs than tract hogs. The results from these studies may not be applicable, however, to all states. For example, in western states with range areas, it is often difficult to associate livestock with the tract. In these states, reporting errors for the tract might exceed those for the farm.

(A8) Duplication Reduced: Duplication of data occurs when a person provides data that should be reported by another operation. Duplication is likely more pronounced in complex management structures such as partnerships and landlord-tenant arrangements. The closed segment approach is believed to be affected the least by duplication since the reporting unit is restricted to what is physically located on the tract land rather than the entire operation.

(A9) More Precise Than The Open Estimator: For characteristics that are highly correlated with farmland, the closed estimator provides large gains in precision over the open estimator. Also, the closed segment approach imposes some control over the maximum value of each survey item since the reporting unit is the tract land operated. Therefore, fewer large data values occur with the closed approach than the open approach. Both of these factors--correlation and reduction of large values--usually result in improved precison for the closed estimator. For example, the closed estimator provides much more precise estimates than the open estimator for acreage items such as corn and soybean acreages. The difference in the levels of precision between the open and closed estimators is not as great for livestock items since livestock are limited to a lesser degree by the amount of land in the segment and are therefore less correlated with farmland and more susceptible to large values. The closed estimator provides greater increases in precision over the open estimator for cattle than for hogs since many cattle operations are relatively stable from tract to tract while hog operations tend to be more concentrated.

As mentioned earlier, SRS uses only the closed segment approach to provide national crop acreage estimates. The precision levels from the open and closed estimators can be compared, however, for corn and soybean acreages using data from a five-state research study conducted during the 1982 JES. Table 3 shows the C.V.'s from each state and the five states combined for the open and closed estimators. The C.V.'s for corn and soybean acreages were much lower using the closed segment approach. At the five state level, the C.V.'s for corn and soybeans were 6.4 and 6.7 percent, respectively, for the open estimator compared with only 2.4 and 2.5 percent for the closed estimator.

Table 3: The coefficient of variation from the open and closed estimators for corn and soybean acreages; 1982 JES.

State	Coefficient of Variation (%)			
	Corn		Soybeans	
	Open	Closed	Open	Closed
Georgia	20.2	9.2	20.7	7.6
Indiana	11.0	3.2	10.6	4.5
Missouri	14.4	8.4	14.4	4.7
North Carolina	17.9	7.6	21.3	6.9
Ohio	11.6	4.3	11.8	4.9
Five States	6.4	2.4	6.7	2.5

The levels of precision that are attained for major crops by the closed method in the operational program are illustrated in Table 4 for soybean and corn acreages from the 1983 JES. The C.V.'s are only presented for those states which account for over 90 percent of the acreage for each crop. The C.V.'s were less than 10 percent in all states for soybeans and in most states for corn. The C.V. at the national level for both crops was 1.3 percent, which is very precise. For minor crops that are naturally less correlated with farmland, the C.V.'s were much higher. For example, the national C.V.'s from the 1983 JES for rye and peanut acreages were 7.0 and 5.0 percent, respectively.

Table 4: The distributions of the coefficient of variation from the closed estimator for corn and soybean acreages for the major producing states; 1983 JES.

Coefficient of Variation (%)	Corn for Grain	Soybeans for Beans
	Number of States	
< 5.0	7	6
5.0-6.5	3	4
6.6-8.1	2	2
8.2-9.7	2	5
≥ 10.0	2	0
Total	16	17

The open and closed estimators can be compared for livestock inventories using data from the operational program. The precision of the estimates for total cattle and calves was better for the closed estimator in 44 of the 48 states from the 1983 JES. At the national level, the C.V. for total cattle and calves was 2.3 percent for the open estimator compared with 1.2 percent for the closed estimator. On the other hand, the closed estimates for total hogs and pigs were more precise than the open estimates in only 29 of the 48 states and the national C.V. was only reduced from 3.8 to 3.4 percent. Table 5 gives the distributions of the C.V. for the closed estimator based on the 1983 JES for total cattle and calves and total hogs and pigs for the same states used in Tables 1 and 2. The C.V.'s for cattle were much lower using the closed segment approach. The precision levels for total hogs and pigs were not improved much by the closed estimator so were still very imprecise at the state level. For livestock items other than total inventory, the C.V.'s were naturally higher.

Table 5: The distributions of the coefficient of variation from the closed estimator for total cattle and calves and total hogs and pigs for selected states; 1983 JES.

Coefficient of Variation (%)	Total Cattle and Calves	Total Hogs and Pigs
	Number of States	
< 5.0	3	0
5.0-9.9	22	2
10.0-19.9	3	14
20.0-29.9	0	7
≥ 30.0	0	0
Total	28	23

The disadvantages associated with the closed segment approach are:

(D1) Restricted: Data cannot be obtained accurately on the tract basis for many survey items. Therefore, it is impractical to use the closed segment approach for items such as farm production expenditures and cattle deaths. For this reason, SRS has only used the closed segment approach for acreages and livestock inventories.

(D2) Generally Imprecise: Even though the closed estimator has greater precision than the open estimator, the closed estimates are often not precise at the state level. For example, none of the C.V.'s for any cattle or hog items from

the 1982 DES were less than 5 percent in the multiple frame states. Closed acreage estimates for major crops, however, are often precise in the major producing states.

(D3) Photo Shortcomings: Response errors can result because of problems associated with the aerial photographs. Obsolete photos can make segments somewhat unrecognizable to a respondent if the land has changed significantly over the years. Bad segment boundaries, e.g. a boundary cutting through a corn field or range area, can also cause reporting mistakes.

(D4) Data Collection Costs Higher Than The Open Approach: The closed segment approach requires that information be obtained for each operation with land in the segment while the open approach only requires data for operations satisfying the headquarters rule. Therefore, roughly twice as many interviews are required by the closed method. The closed approach does provide some time savings since tract data in many instances is easier to obtain than farm data and since farm operators who have no agricultural activity in the tract do not have to be interviewed. As mentioned earlier, data collection costs increase roughly between 5 and 20 percent depending on the state when using the closed rather than the open segment approach. For crop acreage and cattle estimates, the closed approach is more efficient (precision achieved for a fixed cost) than the open approach in most cases. However, neither approach is generally more efficient for estimating hog numbers.

(D5) Unfamiliar Reporting Unit: Operators must acclimate themselves to the tract concept, which in many cases is unfamiliar to them. If the enumerators do not take the time to ensure that each operator understands the tract boundaries, reporting errors may occur. For example, the respondent must understand the tract boundaries in order to report accurately the number of cattle inside the tract or with access to the tract.

Remedial measures are possible for the third disadvantage -- photo shortcomings. Obsolete aerial photographs will always be a problem if more recent photography is not available. When new photography is available for a state, this photography should be reviewed to determine if enough change has occurred in the land composition for any segments to warrant using the new photography.

Weighted

The weighted segment approach requires the following two pieces of information for each tract operation in a sampled segment: (1) data pertaining to the total land operated for each survey item, and (2) the total tract and farm acreages operated, which are used to prorate the entire farm data for each survey item to a tract basis. For example, if a farm

operator has 150 hogs located on his entire operation of 50 acres and operates 10 acres in the tract, then the weighted value for the operation is: 150*10/50 = 30 hogs. The weighted estimate is computed by: (1) adding the weighted values from the tract operations to the segment level, (2) multiplying the weighted segment data by the expansion factor for the segment, and (3) adding the expanded weighted segment data to the desired inference level such as the state.

The sample estimate of a total using the weighted estimator can be expressed by the following formula for a given state:

$$\hat{Y}_w = \sum_{i=1}^{s} \sum_{j=1}^{p_i} \sum_{k=1}^{r_{ij}} y'_{ijk}$$

where

$$y'_{ijk} = \begin{cases} e_{ijk} \sum_{\ell=1}^{f_{ijk}} a_{ijk\ell}\, z_{ijk\ell} = e_{ijk} \sum_{\ell=1}^{f_{ijk}} w_{ijk\ell} & \text{if } f_{ijk} > 0, \\ 0 \quad \text{if } f_{ijk} = 0, \end{cases}$$

$a_{ijk\ell}$ = the weight for the ℓ^{th} agricultural operation with land in the k^{th} segment, j^{th} paper stratum and i^{th} land-use stratum. This weight is the ratio of the total acres operated in the tract to the total acres operated in the entire operation. The sum of the weights over all segments for a given operation must be equal to one,

$z_{ijk\ell}$ = the value of the survey item on the total acres operated for the ℓ^{th} operation with land in the k^{th} segment, j^{th} paper stratum and i^{th} land-use stratum,

$w_{ijk\ell}$ = the weighted value of the survey item for the ℓ^{th} operation with land in the k^{th} segment, j^{th} paper stratum and i^{th} land-use stratum,

e_{ijk}, r_{ij}, p_i, s and f_{ijk} are as previously defined.

The weighted estimator possesses the following advantages:

(A1) Improved Precision Over The Open And Closed Estimators: The weighted segment approach imposes some control on the maximum value of survey items by prorating the entire farm data to a tract basis. This proration reduces the impact of large farm responses that increase the variability of the open estimates and reduces the influence of large tract values that adversely affect the precision of the closed estimates. Table 6 shows the C.V.'s for the three area

frame estimators from the 1983 JES for total cattle and calves and for total hogs and pigs in the ten states that use these three estimators. In every instance the C.V. was smallest for the weighted estimator.

Table 6: The coefficient of variation from the three area frame estimators for total cattle and calves and total hogs and pigs in the 10 states; 1983 JES.

State	Coefficient of Variation (%)					
	Total Cattle and Calves			Total Hogs and Pigs		
	Open	Closed	Weighted	Open	Closed	Weighted
Georgia	12.0	9.4	7.2	12.1	12.0	9.8
Illinois	11.5	9.8	6.5	15.3	15.1	9.0
Indiana	10.3	10.1	7.0	11.2	11.9	7.9
Iowa	6.8	6.1	4.1	8.7	7.9	5.6
Kansas	11.3	5.9	4.3	15.2	16.0	12.3
Minnesota	8.4	7.3	5.4	10.3	10.6	7.7
Missouri	10.5	5.6	4.1	16.8	15.2	8.4
Nebraska	8.4	5.6	4.2	11.3	10.6	8.7
North Carolina	12.5	8.9	6.5	12.6	7.4	5.2
Ohio	10.5	8.4	6.4	16.5	15.2	10.5
Ten States	3.5	2.3	1.7	4.5	4.2	2.8

For major crop acreage items, the weighted estimator generally provides only a slight increase in precision over the closed estimator. For example, the C.V.'s for corn and soybean acreages at the five-state level from a research project conducted during the 1982 JES were 2.4 and 2.5 percent for the closed estimator and 2.2 and 2.3 percent for the weighted estimator. If nonsampling errors are also considered, the closed estimator is preferred over the weighted estimator for estimating acreages such as corn and soybean acres.

(A2) Not Affected By Association Errors: The weighted segment approach, like the closed approach, avoids the problem of establishing the farm headquarters.

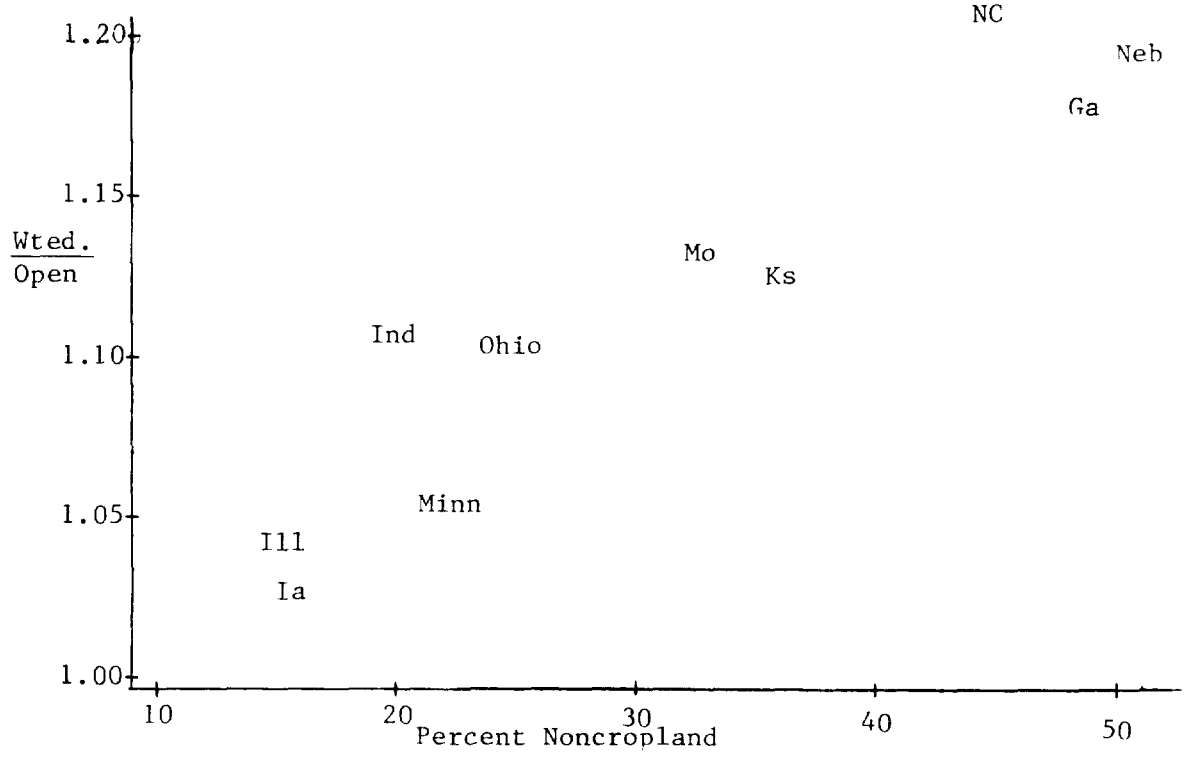
(A3) Flexibility: The reporting unit for the weighted estimator is the total acres operated. Therefore, like the open estimator, the weighted estimator can be used to estimate for any agricultural item.

(A4) Familiar Reporting Unit: Operators generally are familiar with thinking in terms of their entire operation and accustomed to reporting for their farm as a whole.

The weighted estimator used by SRS has several disadvantages that can negate its advantages. These disadvantages are:

(D1) Biased Weight: A reinterview study during 1976 in Indiana, Oklahoma and North Carolina showed that total land operated, which is the denominator of the weight, was underreported by 3, 6 and 11 percent, respectively, in these states (21). The major reason for these undercounts was the operators' failure to include land not actively in use such as woodland and idleland. The effect of underreporting the total acreage is to bias the weight and, therefore, the weighted estimates, upward. Subsequent research (11,12) supported the assertion that the weighted estimates were biased upward for acreage and livestock items. Also, research in 1982 showed that the bias was more pronounced in states where a large percentage of the farmland is devoted to noncropland (11). Other evidence that noncropland causes the weighted estimates to be biased is shown in Figure 2 for the 10 states that compute weighted estimates. This figure shows that as the percentage of noncropland in a state increases, the difference between the weighted and open estimates for the number of farms also increases. Of course, the open estimator may have its own biases so Figure 2 does not necessarily prove the weighted estimator is biased. However, coupled with the findings from the various research projects (11,12,21), Figure 2 gives a strong indication that the upward bias in the weights is related to the omission of noncropland when farmers report their total land.

Figure 2: The ratio of the weighted estimates to the open estimates for number of farms plotted against the percent of farmland devoted to noncropland.



(D2) Reporting Errors: The weighted segment approach is affected by the reporting errors cited earlier for the open approach such as errors in reporting cattle on the total land operated. In addition, the weighted approach is susceptible to errors associated with the weight. For example, if the total acreage in the tract is reported erroneously, then the weight is affected. A reinterview study in 1981 showed that the responses for entire farm acres--the denominator of the weight-- differed between the JES and reinterview for 69 percent of the operations in the three states evaluated (33). Other studies (6,21) have also shown that the entire farm acreage is not obtained accurately. Enumerators in the 1982 panel studies (52) claimed that entire farm acreage was one of the most difficult items to obtain accurately on the JES questionnaire.

(D3) Coverage Errors: The weighted estimator, like the open estimator, is affected by coverage problems such as undercounting operations in commercial, residential and industrial areas, not detecting small farms and overlooking multiple land operations in the same household. However, coverage errors are believed to have less effect on the weighted estimates than the open estimates.

(D4) More Expensive Than The Open And Closed Approaches: The weighted segment approach not only requires that entire farm data be obtained for each survey item from each operation with land in a sample segment, but also requires that the total tract and farm acreages be collected in order to compute the weight for each operation. Therefore, data collection costs are highest for the weighted segment approach. As stated earlier, the costs associated with the weighted segment approach are roughly 5 to 20 percent higher than the open approach depending on the state. Compared to the closed segment approach, the data collection costs from the weighted approach are only marginally higher. For a given cost, the weighted estimator is more efficient than the open estimator for livestock and acreage items and more efficient than the closed estimator for livestock items. Research with major crop acreages has indicated that the efficiency is about the same for the weighted and closed estimators for acreages such as corn and soybeans.

(D5) Subjective Imputations: When a nonrespondent is encountered, visual observation of the survey items is restricted since the reporting unit is the total land operated. This problem was cited earlier as a disadvantage for the open estimator. It is a greater disadvantage for the weighted estimator since a value also has to be imputed for the weight.

(D6) Not Understood: Another disadvantage with the weighted segment approach is this approach is often not understood. In 1974, a survey concepts test administered to survey

statisticians in SRS showed that only 69 percent knew that the weight in the weighted estimator was the ratio of tract acres to farm acres (44). Enumerators are even more uninformed about the weighted segment approach. No mention is made in the enumerator's manual that total land is used to prorate farm data to a tract basis and that the accuracy of the response for total land affects the livestock and number of farms estimates. Therefore, enumerators cannot comprehend the importance of obtaining total land as accurately as possible. The training schools and enumerator's manual should include an explanation of the weighted approach so that enumerators will understand its importance and be motivated to obtain as accurate a response as possible.

The most serious problem with the weighted estimator is that the underreporting of total acres causes the weighted estimates to be biased upward. Obviously, the questions currently used to obtain the entire farm acreage are not working well. Several enumerators in the 1982 panel studies (52) said they reword the questions and ask them in a different sequence. Input should be sought from enumerators on alternative approaches to obtaining "total land operated". Hopefully, the questionnaire can be improved so that underreporting is minimized. If the bias in the weighted estimates persists and this bias varies from year to year so that the bias can not be properly adjusted for in the time-trend charts, SRS should consider using an alternative weight.

Multiple Frame

Multiple frame sampling in SRS uses two sampling frames -- an area frame and a list frame. SRS employs multiple frame methods to estimate livestock numbers, on-farm grain stocks, farm production expenditures, farm labor, and acreages for a few specialty crops. The screening estimator rather than the full multiple frame estimator is used by SRS (5). The multiple frame screening estimates are derived in the following manner. First, each operator identified in the area frame sample is name-matched against the list frame. Next, the area frame operators who were not found on the list frame are classified as nonoverlap. An estimate is then obtained for the list frame and is added to the area frame nonoverlap estimate to arrive at the multiple frame screening estimate. The open, closed or weighted segment approach can be used to provide the area frame nonoverlap estimate.

The sample estimate of a state total for a survey item based on the screening estimator can be expressed in the following notation for SRS's area and list frame designs:

$$\hat{Y}_{mf} = \hat{Y}_{\ell ist} + \hat{Y}_{nol}$$

where

$$\hat{Y}_{\ell ist} = \sum_{h=1}^{H} e_h \sum_{i=1}^{n_h} y_{hi},$$

y_{hi} = the value of the survey item for the i^{th} list frame respondent in list frame stratum h,

e_h = the expansion factor for list frame stratum h,

n_h = the number of respondents in list frame stratum h,

H = the number of list frame strata,

$$\hat{Y}_{nol} = \sum_{i=1}^{s} \sum_{j=1}^{p_i} \sum_{k=1}^{r_{ij}} e_{ijk} \sum_{\ell=1}^{h_{ijk}} y''_{ijk\ell}$$

$$h_{ijk} = \begin{cases} e_{ijk} & \text{if the open segment approach is used for the area frame nonoverlap,} \\ f_{ijk} & \text{if the closed or weighted approach is used for the area frame nonoverlap,} \end{cases}$$

g_{ijk}, f_{ijk}, e_{ijk}, r_{ij}, p_i and s are as previously defined,

$$y''_{ijk\ell} = \begin{cases} y'_{ijk\ell} & \text{if the } \ell^{th} \text{ agricultural operation in segment k, paper stratum j and land-use stratum i is classified as nonoverlap with the list frame,} \\ 0 & \text{otherwise,} \end{cases}$$

$$y'_{ijk\ell} = \begin{cases} y_{ijk\ell} & \text{if the open segment approach is used,} \\ t_{ijk\ell} & \text{if the closed segment approach is used,} \\ w_{ijk\ell} & \text{if the weighted segment approach is used,} \end{cases}$$

$y_{ijk\ell}$, $t_{ijk\ell}$ and $w_{ijk\ell}$ are as previously defined.

SRS uses the open and weighted segment approaches to estimate for the nonoverlap domain. The open nonoverlap estimator is utilized exclusively for surveys dealing with farm production expenditures and farm labor while only the weighted nonoverlap estimator is used for on-farm grain stocks. The cattle multiple frame program, which involves 34 states, uses the weighted nonoverlap estimator in 28 states and the open nonoverlap estimator in 6 states. The hog multiple frame program uses the weighted nonoverlap estimator in 13 states and the open nonoverlap estimator in 10 states.

The advantages and disadvantages mentioned earlier for the

open and weighted estimators also affect the multiple frame estimates. The effect of the area frame component depends upon how much the area frame nonoverlap estimate contributes to the multiple frame estimate. For example, the weighted nonoverlap estimates in New Mexico and Tennessee for total cattle and calves from the 1982 DFS were 7.6 and 41.8 percent of the multiple frame estimates, respectively. Therefore, the various advantages and disadvantages associated with the weighted estimator have negligible impact on the multiple frame estimates in New Mexico, but play an important role in Tennessee.

Multiple frame estimation possesses some distinct advantages that have led SRS to employ this approach instead of or in addition to area frame methods alone. These advantages are:

(A1) Most Efficient: For a given cost, the multiple frame approach yields more precise estimates than any of the area frame estimators for many survey items such as livestock inventories because of two factors. First, the list frame is capable of providing more precision than the area frame, especially if some measure of size is available for a variable related to the survey items that can be used to stratify the list frame efficiently. The area frame is often not stratified specifically for the items of interest, and therefore is not designed to estimate with a high degree of precision items such as livestock numbers and minor crop acreages unless an unreasonably large sample is selected. The second factor contributing to the efficiency of the multiple frame approach is that list frame data can be obtained at less expense than area frame data. The area frame requires that each operator be located and personally interviewed while data from the list frame sample can be obtained using a combination of mail, telephone and personal interviews. However, if the questionnaire for the list frame is complex, nonsampling error considerations may restrict or eliminate the use of mail and telephone interviewing.

The precision of the multiple frame livestock estimator is illustrated in Table 7. This table compares the C.V.'s from the multiple frame estimator with the most precise area frame estimator--the weighted estimator--for total cattle and calves and total hogs and pigs for states that compute both estimators. The precision levels for hogs and pigs are based on the 1983 JES while the levels for cattle pertain to the 1982 DFS since multiple frame cattle estimates are no longer generated in conjunction with the JES. In every instance, the multiple frame estimate was at least as precise as the weighted estimate.

The multiple frame livestock estimator is generally much more precise than the open and closed livestock estimators. This statement is supported by Table 8 which compares the distributions of the C.V. from the open, closed and multiple

frame livestock estimators based on the 1982 DES. For the estimate of total cattle and calves, 21 of the 28 states had a C.V. less than 5 percent using the multiple frame estimator while the open and closed estimates were never that precise. For total hogs and pigs, only one state had a C.V. less than 10 percent based on the open and closed estimators compared with 12 states for the multiple frame estimator.

Table 7: The coefficient of variation from the weighted and multiple frame estimators for total cattle and calves and total hogs and pigs in the 10 states.

State	Coefficient of Variation (%)			
	Total Cattle and Calves 1/		Total Hogs and Pigs 2/	
	Weighted	Multiple Frame	Weighted	Multiple Frame
Georgia	7.5	4.7	9.8	7.9
Illinois	9.2	4.7	9.0	3.8
Indiana	7.7	4.3	7.9	4.5
Iowa	5.6	3.1	5.6	3.4
Kansas	4.8	3.2	12.3	5.7
Minnesota	6.3	3.8	7.7	6.4
Missouri	6.3	4.1	8.4	6.8
Nebraska	8.8	8.1	8.7	4.5
North Carolina 3/	--	--	5.2	5.2
Ohio	6.9	4.4	10.5	7.8
Ten States	2.6	2.0	2.8	1.7

1/ 1982 DES 2/ 1983 JES
3/ Multiple frame cattle estimates are not generated in North Carolina.

Table 8: The distributions of the coefficient of variation for the open, closed and multiple frame estimators for total cattle and calves and total hogs and pigs; 1982 DES.

Coefficient of Variation (%)	Total Cattle and Calves			Total Hogs and Pigs 1/		
	Open	Closed	Multiple Frame	Open	Closed	Multiple Frame
	Number of States					
< 5.0	0	0	21	0	0	5
5.0-9.9	6	17	7	1	1	7
10.0-19.9	22	10	0	14	15	9
≥ 20.0	0	1	0	6	5	0
Total	28	28	28	21	21	21

1/ Only 21 of the 23 multiple frame hog states are included since two of the states were not multiple frame states in 1982.

Although the multiple frame livestock estimator is more efficient than the area frame livestock estimators, it is not known how the efficiency of a multiple frame acreage approach would compare with the operational closed acreage approach for major crop acreages. Even if the multiple frame acreage approach is more efficient, nonsampling error considerations could make the closed segment approach more desirable for major crop acreages.

(A2) Flexibility: The reporting unit is the total land operated for all multiple frame surveys that SRS conducts. Therefore, the multiple frame approach, like the open and weighted approaches, can be used to estimate for any agricultural item.

(A3) Familiar Reporting Unit: Another characteristic shared by the open, weighted and multiple frame approaches is that the reporting unit is familiar to farmers since they are generally accustomed to thinking in terms of their farm as a whole.

(A4) Imputation Simplified: As mentioned earlier, the manual imputation process for area frame nonrespondents that is used by SRS for the open and weighted approaches is often subjective and difficult. On the other hand, statisticians do not manually impute data for nonrespondents from the list frame. Instead, the expansion factor for each stratum with any nonrespondents is adjusted by the computer program. Therefore, the imputation procedure is easier for the office personnel.

Efforts to improve the precision of the estimates often lead to a more complex survey design and estimation procedure such as the multiple frame approach. The added complexity can create situations that may increase nonsampling errors. A multiple frame approach usually has more potential for nonsampling errors than a single frame approach because of errors peculiar to each frame as well as errors that result from combining the frames (3). The disadvantages of the multiple frame approach are:

(D1) Domain Determinations: The process of classifying each operation from the area frame sample into either the overlap or nonoverlap domain is generally considered the most difficult task to execute correctly. Factors such as incomplete name and address information, incorrect spellings of names, the use of nicknames, nonperson names and names generated primarily for legal purposes, and the existence of complex management structures lead to problems with matching names correctly between the area and list frames. Also, confusion concerning what constitutes a partnership may result in joint land operations being treated as individual operations and vice versa, which can affect the domain determinations. Studies by SRS have provided ample evidence that errors

occur with domain determinations that can bias the multiple frame estimates (22,48,49).

(D2) Possible Loss of Frame Independence: The area frame sample is used to estimate the incompleteness of the list frame. Therefore, the area and list frames must be kept independent (50). If area frame information is used to update the list frame, bias will be injected into the multiple frame estimates. Research studies have indicated that the frames possibly were not being maintained independently (3). For example, a study during 1975 showed that the size of the nonoverlap domain decreased the longer the operations from the area frame were in the sample, thereby implying either that area frame information was used to update the list frame or that the domain determinations were not done accurately when new segments were rotated into the sample (13).

(D3) Biased Downward: The imputation procedure employed by SRS for list frame nonrespondents assumes that nonrespondents in a given stratum are like the respondents in the same stratum. A study in 1977 discovered that multiple frame livestock estimates are generally biased downward because nonrespondents from the list frame tend to have more livestock than respondents (18). Subsequent research identified a 2 to 6 percent downward bias in the list frame livestock estimates due to the current imputation procedure (9,10). This bias occurred because the percent of nonrespondents having livestock (63 percent) was much larger than the percent of respondents having livestock (28 percent).

(D4) Reporting Errors: As mentioned earlier, research has shown that farmers frequently make errors when reporting for their entire operation. Therefore, multiple frame estimates, like the area frame estimates, are susceptible to reporting errors. Reporting errors may be more pronounced for multiple frame surveys due to the various data collection methods used. For example, survey concepts are more difficult to convey by mail. A study in Wyoming during 1973 found that questionnaires returned by mail required more editing than those by telephone and personal interviews (47). Also, SRS does not train telephone interviewers as thoroughly as personal interviewers, which may result in more data errors.

(D5) List Frame Deficiencies: The coverage provided by the list frame deteriorates rapidly with time as does the accuracy of the control data used for stratification. Therefore, if the list frame is not continually updated, the multiple frame approach loses efficiency. The list frame often contains misleading information about joint operations which causes errors in estimation (6). Duplication of responses is also believed to be a more serious problem with the list frame than with the area frame. Finally, the association of

the sampling unit with the reporting unit is prone to more errors for the list frame than the area frame (3).

The harmful effects of nonsampling errors on the multiple frame estimates can be reduced. Nonsampling errors associated with domain determinations can possibly be decreased by reducing the size of the list frame. Research has shown that the size of the list frame can be drastically decreased with little change in the precision of the multiple frame estimates (7,51). Reducing the size of the list frame simplifies the name-matching process. SRS has recently implemented a reduced list frame concept in a few states. Efforts should continue in the area of optimizing the size of the list frame.

A procedure was devised in 1978 to reduce the downward bias in the multiple frame estimates caused by nonrespondents (10). Although SRS computes livestock estimates based on this procedure, these estimates have not been integrated into the operational program. These estimates should be used when setting the official statistics.

A considerable amount of the data from the list frame sample is obtained by telephone interviews. For example, about 60 percent of the livestock interviews are completed by telephone. The Long-Range Planning Group recommended that specialized training be developed for telephone interviewers to improve their skills (1). This training should decrease the number of reporting errors. Also, the use of computer-assisted telephone interviews (CATI) has great potential for minimizing response errors through on-line data comparisons with current and historic survey data. Finally, an alternative design for the mail questionnaire should be explored in an attempt to provide the operator a clearer presentation of survey concepts such as "land operated" and "regardless of ownership" which are difficult to convey by mail.

One "very" important issue to be considered that affects the quality of the estimates from each of the four direct expansion estimators is the complexity of the current survey procedures. For example, data on livestock inventories are obtained for all four direct expansion estimators in some states. Enumerators and respondents are expected to understand fully various survey concepts and be capable of supplying accurate data for each concept. The more complex the survey procedures, the greater the potential for nonsampling errors. Survey procedures would be simplified greatly if only one estimator was needed for a survey item, thereby resulting in fewer potential nonsampling errors. Therefore, SRS should attempt to reduce the number of estimators used in order to simplify the survey procedures for the enumerators and respondents.

USE OF THE ESTIMATORS

Estimates from the four direct expansion estimators will be compared with the official statistics (revised through 1983) for two important livestock survey items--total cattle and calves and total hogs and pigs. Other pertinent survey items such as crop acreages, on-farm grain stocks and farm production expenditures were not evaluated since these items rely on only one direct expansion estimator. The purpose of the comparisons between the livestock estimates and official statistics is to ascertain how much each estimator is relied on or used by SRS. The "use" of the four livestock estimators was measured by:

(1) how close the estimates have been to the official statistics, and

(2) how correlated the estimates have been with the official statistics.

The second criterion--correlation--will show which estimator provides the most consistent relationship with the official statistics in the time-trend charts used by SRS. Finally, the estimates will also be reviewed to see if any of the estimators have been consistently above or below the official statistics when estimating livestock inventories.

As mentioned by the Long-Range Planning Group (1), a predefined, documented and objective set of procedures are not used to arrive at the official statistics. Therefore, there are limitations when attempting to quantify analytically how much SRS relies on each estimator. For example, an estimate may be close to the official statistic for a given year by coincidence rather than by choice. Thus, the inferences in this report concerning which estimators are used the most are open to some criticism. It is important to point out at the outset that the analysis in this section IS NOT EVALUATING HOW GOOD EACH ESTIMATOR IS since the official statistics are subject to error. This point should be kept in mind when interpreting the results presented for the various estimators.

Cattle estimates are generated from multiple and/or area frame surveys semi-annually. In January, official statistics are set based on these surveys for each state, for the 28 original multiple frame states combined, for the "other 20" states combined and for the nation (48 conterminous states). In July, official statistics are only made for the nation. The direct expansion estimates will be compared with the official statistics for each of these inference levels. State comparisons between the estimates and official statistics will be restricted to the 34 multiple frame cattle states, which contain about 95 percent of the U.S. cattle.

Hog surveys are conducted quarterly (March, June, September and December). In March and September, official statistics are set only for the 10 major hog producing states (known as the

quarterly hog states) and for these 10 states combined using the multiple frame estimator. Both area and multiple frame estimates are computed in December and June. Official statistics are made for each state, for the 10 quarterly hog states combined, for the "other 40" states combined and for the nation in December. In June, official statistics are set for each of the 10 quarterly hog states, the 10 states combined, the other 40 states combined and the nation. The analysis of the hog estimators will be based on the December and June surveys since both the area and multiple frame estimators are generated at these times. Analysis for individual states in December will be restricted to the 23 multiple frame hog states, which contain about 95 percent of the U.S. hogs.

Cattle

Comparisons between the official statistics and the various direct expansion estimates for the 34 individual states were made using January data from 1978 to 1983. The states examined first were the 19 multiple frame states that compute multiple frame (weighted nonoverlap), closed and open estimates.

Table 9 shows the average absolute relative difference between the official statistics and each of the direct expansion estimates for each state. This difference measures how close the various estimates have been to the official statistics. In every state, the multiple frame estimate was consistently much closer to the official statistic than the closed and open estimates. The multiple frame estimates were always within 4 percent of the official statistics while the closed estimates were within 4 percent in only one state and the open estimates were never within 4 percent. Clearly, the multiple frame estimator is relied on very heavily by SRS for these states. Overall, the closed estimates were closer to the official statistics than were the open estimates.

The relationship between the estimates and the official statistics since 1978 has been the strongest for the multiple frame estimator in most states. The correlation coefficient averaged .9 across the 19 states for the multiple frame estimator but only about .5 for the closed and open estimators. Table 10 shows the distributions of the correlation coefficient for the three estimators. None of the coefficients were below .6 for the multiple frame estimator while more than half of the states had correlation coefficients less than .6 for the closed and open estimators. Therefore, the multiple frame estimates were not only consistently the closest to the official statistics, but were also best correlated with the official statistics, thereby providing the best time-trend charts.

Finally, none of the estimators consistently provided estimates above or below the official statistics for the 19 states. However, the multiple frame and closed estimates tended to be above the official statistics, and the open estimates were generally below the official statistics.

Table 9: The average absolute relative difference between each estimator and the official statistics in January for total cattle and calves in the 19 states; 1978-1983.

State	Average Absolute Relative Difference (%) 1/		
	Multiple Frame	Closed	Open
Alabama	1.5	16.5	8.9
Arkansas	2.4	3.5	11.1
California	0.5	6.3	9.6
Colorado	2.7	5.7	13.7
Florida	3.3	12.8	9.2
Idaho	3.1	13.5	8.3
Kentucky	2.3	6.0	11.3
Missisippi	1.0	8.4	12.3
Montana	0.6	9.2	21.5
New Mexico	1.5	9.8	17.3
New York	3.8	11.1	6.5
North Dakota	1.9	4.9	6.2
Oklahoma	2.3	4.5	6.6
Pennsylvania	2.5	8.2	5.5
South Dakota	1.9	4.7	11.3
Tennessee	3.0	16.1	10.1
Texas	0.4	4.2	14.6
Wisconsin	3.6	7.6	9.3
Wyoming	2.8	9.9	18.4
Average	2.2	8.5	11.1

1/ Absolute Relative Difference = $100 * \frac{\text{Absolute Value (Estimate-Official Statistic)}}{\text{Official Statistic}}$

Table 10: The distributions of the correlation coefficient in January for the three estimators in the 19 states; 1978-1983.

Correlation Coefficient	Multiple Frame	Closed	Open
		Number of States	
> .9	13	2	2
.6-.9	6	7	7
< .6	0	10	10
Total	19	19	19

The next group of states reviewed generates all four direct expansion estimates. The average absolute relative difference between the estimates from each estimator and the official statistics is given in Table 11 for each state. The multiple frame estimates (weighted nonoverlap) were closest to the official statistics in 5 of the 9 states. However, in two of the states where the multiple frame estimates were not the closest--Illinois and Iowa--the correlation coefficients between the multiple frame estimates and official statistics were very high (.88 in Illinois and .97 in Iowa), implying that the multiple frame estimator is also heavily relied upon by SRS for these two states. In fact, the relationship between the estimates and the official statistics was the most stable overall for the multiple frame estimator as illustrated in Table 12. The average of the correlation coefficients across the nine states was over .8 for the multiple frame estimator compared with about .6 for the area frame estimators. Finally, the closed and open estimates were neither consistently above nor below the official statistics in the 9 states. On the other hand, the multiple frame estimates were generally below and the weighted estimates usually above the official statistics.

Table 11: The average absolute relative difference between each estimator and the official statistics in January for total cattle and calves in the 9 states; 1978-1983.

State	Average Absolute Relative Difference (%)			
	Multiple Frame	Closed	Open	Weighted
Georgia	4.1	10.7	5.6	23.7 1/
Illinois	6.8	4.0	7.4	4.8
Indiana	3.1	7.4	8.6	9.6
Iowa	6.5	5.7	4.5	5.2
Kansas	3.8	6.6	11.2	5.2
Minnesota	2.1	1.8	5.4	5.8
Missouri	1.5	6.9	7.3	3.6
Nebraska	1.1	5.3	4.1	8.9
Ohio	4.4	5.6	8.5	3.0
Average	3.7	6.0	7.0	7.8

1/ Data for only 3 years available in Georgia for the weighted estimator. The high average absolute relative difference resulted because the absolute relative difference is 1982 was 54.1 percent.

Table 12: The distributions of the correlation coefficient in January for the four estimators in the 9 states; 1978-1983.

Correlation Coefficient	Multiple Frame	Closed	Open	Weighted
		Number of States		
>.9	4	0	1	0
.6-.9	4	4	3	6
<.6	1	5	5	3
Total	9	9	9	9

The final group of multiple frame cattle states--Arizona, Louisiana, Michigan, Oregon, Virginia and Washington--currently relies on the multiple frame (open nonoverlap), closed and open estimators in January. These three estimators were integrated into the estimation program in Louisiana, Michigan and Virginia during the 1982 DFS and in the other three states during the 1983 DFS. Therefore, historic comparisons of the three estimators with the official statistics are not available for these states. It is worth noting that the C.V.'s from the closed and open estimators for total cattle and calves from the 1982 DFS in Louisiana, Michigan and Virginia were in double-digits. Therefore, the closed and open estimates in these states will probably fluctuate quite a bit from year to year, thereby making time trends difficult to recognize. In these same states, the area frame nonoverlap estimates accounted for about half of the multiple frame estimates, which was the highest percentage of all multiple frame states. This caused the C.V.'s from the multiple frame estimator to be higher than desirable in these states (9.2 percent in Virginia, 9.6 percent in Michigan and 9.8 percent in Louisiana). Therefore, the multiple frame estimator may not be relied on as much in these states in the future as it has been in the original 28 multiple frame states.

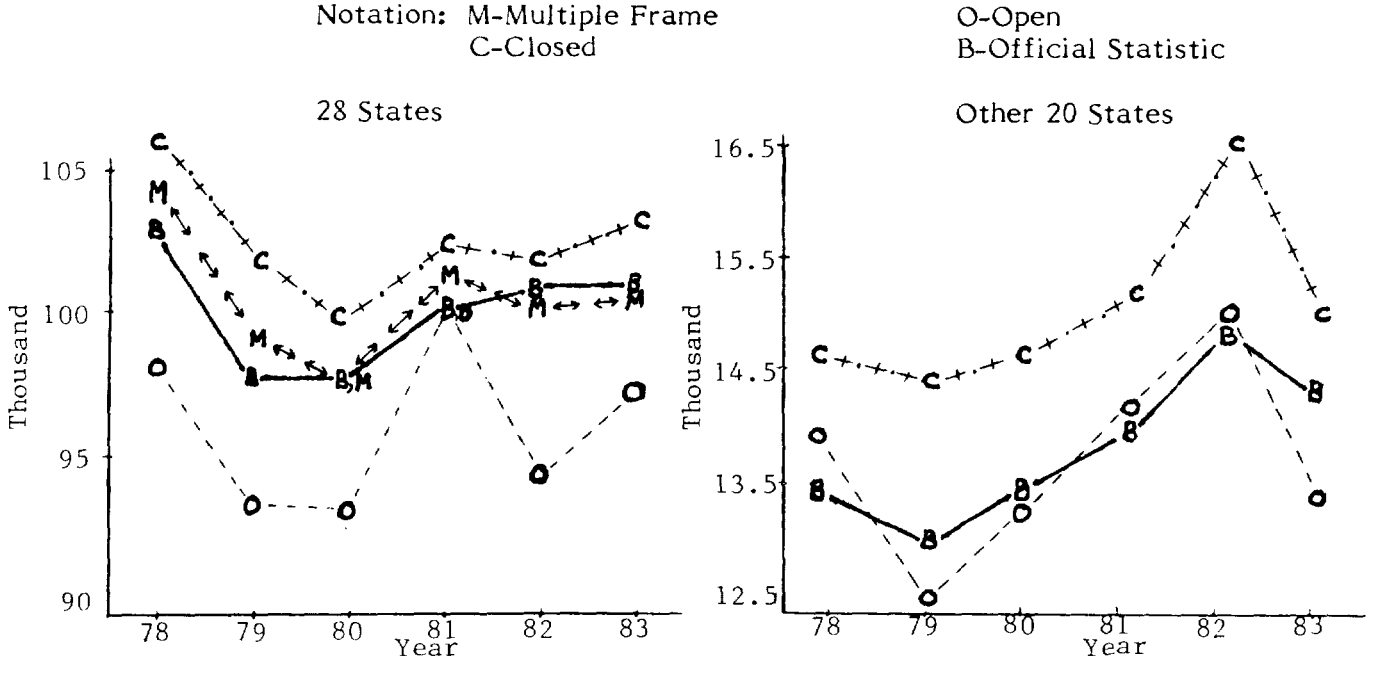
Next, the January estimates and official statistics were compared for the 28 original multiple frame states combined and for the other 20 states combined. The 28 states generate multiple frame, open and closed estimates. The other 20 states only use the open and closed estimates. Figure 3 shows graphically the estimates and official statistics since 1978.

For the 28 multiple frame states combined, which contain about 90 percent of the U.S. cattle, Figure 3 shows that the multiple frame estimates have historically been the closest to the official statistics. The closed estimates have always been higher than the official statistics while the open estimates have been far below the official statistics (except 1981). The multiple frame and closed estimates have been well correlated with the official statistics since 1978. The open estimates have followed a similar curve except for 1981. Therefore, although the levels from the estimators differ, they tend to move in a similar manner from year to year. Under the

assumption that the official statistics are correct, Figure 3 suggests that consistent factors, namely, nonsampling errors, separate the lines. The cumulative effect of the nonsampling errors results in an average relative difference between the estimates and official statistics of -4.1 percent for the open, 2.3 percent for the closed and 0.3 percent for the multiple frame estimator at the 28 state level. The negative relative difference for the open estimator seems to confirm the undercoverage problem cited earlier. The reason why the closed estimates have been higher than the official statistics is not apparent. Possibly, there is a tendency for farm operators to report more cattle than are physically located in or have access to the tract. Finally, the closed estimates may be consistently higher than the multiple frame and open estimates because of the different methods of imputing for nonrespondents. As mentioned earlier, nonresponse imputations are more accurate for the closed segment approach than the multiple frame, open and weighted approaches.

Figure 3 shows that the closed estimates were always higher than the open estimates for the other 20 states combined, which contain about 10 percent of the U.S. cattle. The closed estimates were always higher than the official statistics while the open estimates have been close to and neither consistently above nor below the official statistics. Both estimators have been well correlated with the official statistics in the time-trend charts. Contrary to the pattern for the 28 states combined, SRS has followed the open estimator more than the closed estimator for the other 20 states combined.

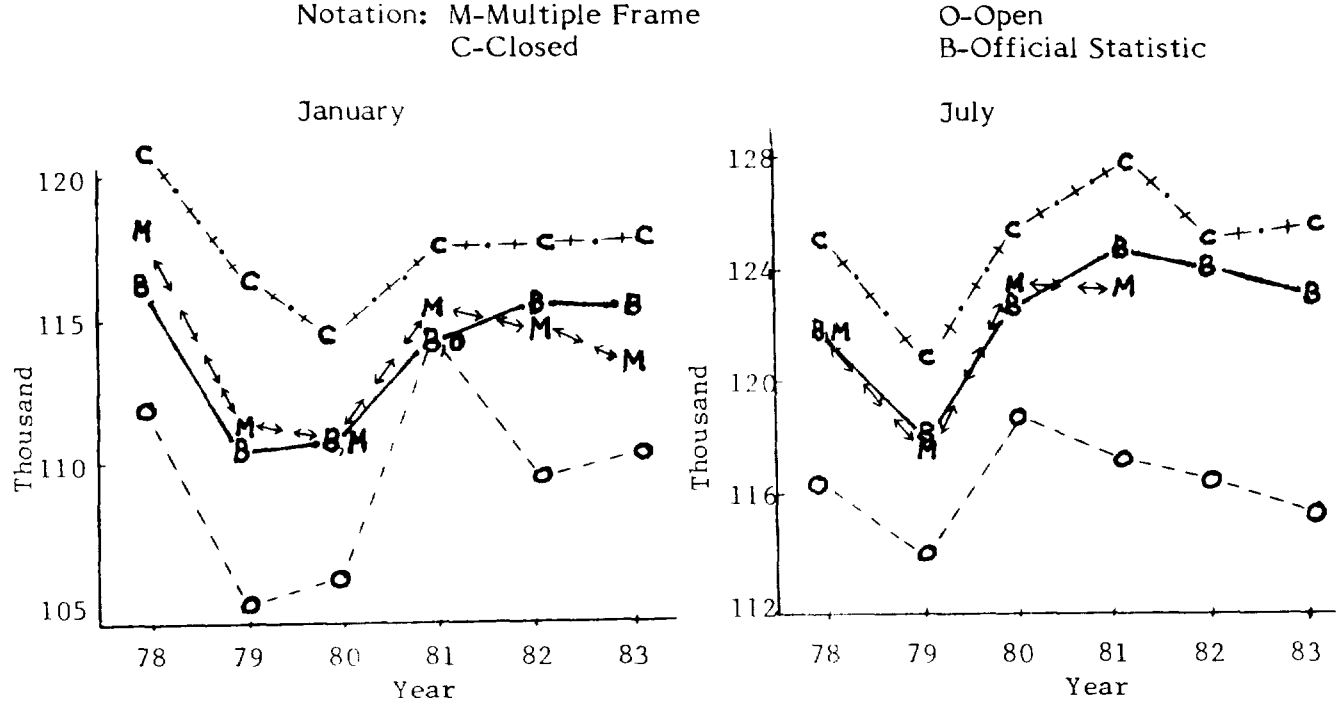
Figure 3: The direct expansion estimates and official statistics for total cattle and calves in January for the 28 multiple frame states and the other 20 states; 1978-1983.



Finally, the national estimates were compared with the official statistics in January and July. Two points are worth mentioning about the July cattle estimation program. First, the multiple frame approach was terminated in 1982 due to budget cuts. Therefore, comparisons between the multiple frame estimates and official statistics in July are restricted to 1978-1981. Second, weighted cattle estimates are presently generated in 10 states even though no official statistics are made at the state level in July. Therefore, the weighted cattle estimates serve no official purpose in July.

Figure 4 summarizes the comparisons between each of the three estimators and the official statistics at the national level (the 48 conterminous states). The multiple frame estimates are not "pure" multiple frame estimates since the multiple frame approach is not used in every state. The multiple frame estimates shown in Figure 4 are based on the multiple frame estimates from the 28 original multiple frame states and the open estimates from the remaining 20 states. Figure 4 shows that the closed estimates have always been higher than the official statistics while the open estimates have been below the official statistics. Again, the multiple frame estimates have been consistently the closest to the official statistics. In July, the closed estimates have been much closer to the official statistics than the open estimates. Finally, all three estimators are well correlated with the official statistics in the national charts, indicating again that although the estimates are different, the estimates of change are relatively stable from one year to the next.

Figure 4: The direct expansion estimates and official statistics for total cattle and calves for the nation; 1978-1983.



In summary, SRS relies on the multiple frame cattle estimator more than the area frame cattle estimators. At each inference level--state, 28 states combined and national-- the multiple frame estimates were the closest to the official statistics. Also, at the state level, the relationship between the estimates and official statistics in most states was the best for the multiple frame estimates. When the states are combined, a strong relationship exists between the official statistics and each estimator. Finally, none of the area frame estimators performed consistently better than the other area frame estimators. Overall, however, SRS has relied on the closed estimator slightly more than the open estimator.

Hogs

The direct expansion estimates and official statistics for total hogs and pigs were compared using data in December from 1978 to 1982 and in June from 1978 to 1983. Comparisons between each estimator and the official statistics were made for each inference level cited earlier, namely, each of the 23 major hog producing states, the 10 quarterly hog states combined, the "other 40" states combined and the nation.

The first group of states examined was the 10 quarterly hog states, which contain over three-fourths of the U.S. hogs. The average absolute relative difference, which measures how close the estimates have been to the official statistics, is given in Table 13 for each state. This table shows that the multiple frame estimates (weighted nonoverlap) were closest to the official estimates in all states except Illinois and Iowa where the weighted estimates were closest. The percentage of noncropland associated with farmland in Illinois and Iowa is the smallest of all the quarterly states. The bias in the weighted estimates in these two states is believed to be negligible since the bias-causing component--noncropland--is small [11]. Therefore, it is not surprising that the weighted estimates have been close to the official statistics in Illinois and Iowa. Of the area frame estimators, the weighted estimates were closer to the official statistics in most states than either the open or closed estimates.

The correlation between the estimates and official statistics in the 10 states was the strongest for the multiple frame estimator as shown in Table 14. The average correlation coefficient for the multiple frame estimator was .95 for the 10 states. Also, the correlation from the time-trend charts was the highest for the multiple frame estimator in Illinois and Iowa, the only two states where the multiple frame estimates were not the closest to the official statistics. Therefore, based on the two criteria--closeness and correlation--SRS relies on the multiple frame hog estimator much more than the area frame estimators for the 10 quarterly hog states. The weighted estimator is used the most of the three area frame estimators. Finally, none of the estimators consistently provided estimates above or below the official statistics in the

10 states. However, the multiple frame hog estimates have generally been lower than the official statistics.

Table 13: The average absolute relative difference between each estimator and the official statistics for total hogs and pigs in the 10 quarterly hog states; 1978-June 1983.

State	Survey	Average Absolute Relative Difference (%)			
		Multiple Frame	Weighted 1/	Closed	Open
Georgia	December	2.0	6.4	7.6	15.2
	June	2.0	6.3	18.6	11.2
Illinois	December	5.4	2.9	6.7	8.8
	June	4.6	4.1	9.5	12.6
Indiana	December	3.6	9.0	16.1	9.2
	June	3.8	9.9	20.7	12.8
Iowa	December	2.6	3.2	5.2	5.1
	June	4.7	2.0	3.8	6.4
Kansas	December	4.9	7.4	15.3	19.8
	June	3.9	8.8	8.2	17.8
Minnesota	December	1.9	9.1	7.2	3.4
	June	2.5	7.6	8.0	7.8
Missouri	December	3.6	3.7	4.5	6.0
	June	3.0	4.8	7.8	6.2
Nebraska	December	4.0	14.3	17.4	14.8
	June	8.3	11.6	12.4	11.0
North Carolina	December	1.1	2.0	14.5	8.3
	June	1.7	3.3	8.1	7.3
Ohio	December	3.0	5.4	11.6	8.3
	June	3.5	3.5	20.6	10.1
Average		3.5	6.3	11.2	10.1

1/ Data for weighted estimates only available in 1981 and 1982 from the DES for Georgia and North Carolina, in 1979, 1981, 1982 and 1983 from the JES for Georgia and from the 1981 to 1983 JES for North Carolina.

Table 14: The distributions of the correlation coefficient for the four estimators in the 10 quarterly hog states; 1978-June 1983.

Correlation Coefficient	Multiple Frame	Closed	Open	Weighted 1/
		Number of States		
>.9	18	3	4	7
.6-.9	2	9	9	6
<.6	0	8	7	3
Total	20	20	20	16

1/ Excludes correlation coefficients based on fewer than 5 observations.

Three estimators--multiple frame (weighted nonoverlap), closed and open--are used in Kentucky, South Dakota and Wisconsin during December. The average absolute relative difference between the hog estimates and official statistics is shown in Table 15 for each estimator. In South Dakota and Wisconsin, the multiple frame estimates were not only very close to (generally below) the official statistics, but also have been almost perfectly correlated with the official statistics (correlation coefficent = .99). None of the estimators performs consistently well in Kentucky. The large contribution to the multiple frame estimates from the area frame nonoverlap domain (40 percent in 1982) may be causing the multiple frame estimates to be unstable.

Table 15: The average absolute relative difference between each estimator and the official statistics in December for total hogs and pigs in the 3 states; 1978-1982.

State	Average Absolute Relative Difference (%)		
	Multiple Frame	Closed	Open
Kentucky	12.5	15.2	10.6
South Dakota	2.4	11.0	6.1
Wisconsin	1.1	9.6	7.0

The last ten multiple frame states currently rely on the multiple frame (open nonoverlap), closed and open estimators in December. Eight of the states began using all three of these estimators in 1982 while the remaining two states started in 1983. Historic information on all three estimators is therefore not available in these states for comparison with the official statistics. Inspection of the estimates from 1982 in the eight states (see Table 16) revealed that the closed and open estimates were very imprecise so these estimators are not expected to be useful. The multiple frame estimates were much more precise, but not as precise as the multiple frame estimates in the states discussed earlier.

Table 16: The coefficient of variation from each estimator for total hogs and pigs in the 8 states; 1982 DES.

State	Coefficient of Variation (%)		
	Multiple Frame	Closed	Open
Alabama	14.8	25.6	25.8
Michigan	13.1	19.1	19.6
Mississippi	14.3	43.8	43.5
Oklahoma	11.9	25.3	22.9
Pennsylvania	18.2	21.5	52.5
South Carolina	11.8	18.7	18.9
Texas	5.1	16.4	12.6
Virginia	16.5	87.0	22.6
Average	13.2	32.2	27.3

The final comparisons of the estimates with the official statistics were made for the following inference levels--the 10 quarterly states combined, the other 40 states combined and the nation. Four points are worth mentioning concerning these comparisons:

(1) Only the multiple frame, closed and open estimates were compared with the official statistics at the 10 state level. SRS does not use the weighted estimates in the time-trend charts since weighted estimates have only been generated in all ten states since 1981. Therefore, the weighted estimator was not included in the analysis.

(2) Only the closed and open estimators are available for comparisons pertaining to the other 40 states combined.

(3) The JES and DES estimates from the 38 states combined are used to set the official statistics for 40 states, and the estimates from the 48 conterminous states combined are used to set the official statistics for 50 states. Therefore, the hog estimates at these inference levels should be slightly below the official statistics.

(4) As is the case with the cattle program, the national multiple frame hog estimates are based on both the multiple frame and open estimators. In June, the national multiple frame estimates include multiple frame estimates from the 10 quarterly hog states and open estimates from the remaining 38 states. In December, multiple frame estimates from 23 states are available for the national multiple frame estimates.

Figure 5 shows how the estimates from each estimator compare to the official statistics in December and June for each inference level. All three estimators are highly correlated with the official statistics, as can be seen by the similar pattern of the lines in the figures.

The multiple frame estimates were consistently below the official statistics for the 10 quarterly states combined. Under the assumption that the official statistics are correct, the average relative difference of -2.9 percent between the multiple frame estimates and the official statistics may be caused by the current nonresponse imputation method, which has been shown to be biased downward. The closed hog estimates have usually been higher than the official statistics in December for the 10 quarterly states, the other 40 states combined and the nation. Recall that the closed cattle estimates were also higher than the official statistics. To the contrary, the closed hog estimates in June tend to be slightly below the official statistics (about 1 percent). Finally, the open estimates have been higher than the official statistics for the 10 states combined and the nation. This pattern is contrary to the behavior of the open estimator for cattle.

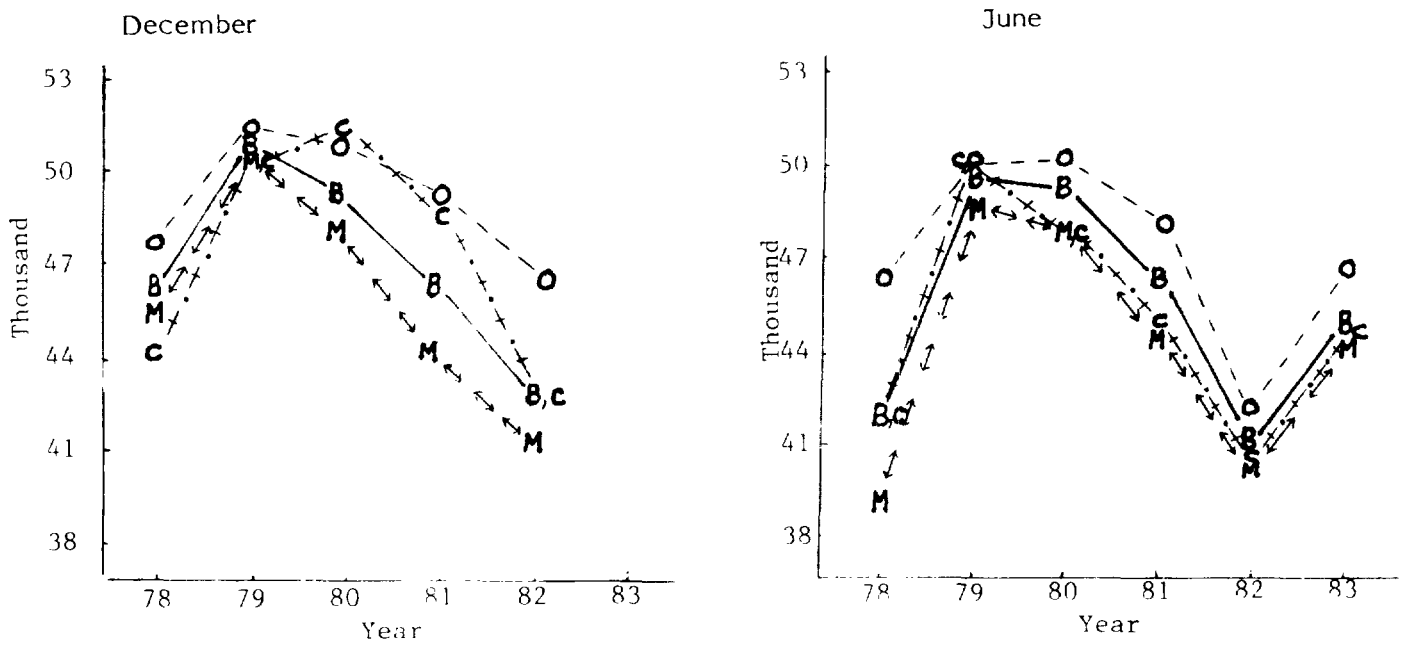
The average absolute relative difference between the estimates and official statistics is given in Table 17 for each inference level. The multiple frame estimates were closest to the official statistics in December while the closed estimates have been the closest in June.

Summarizing the analysis for total hogs and pigs, SRS relies on the multiple frame estimator the most of the four direct expansion estimators. The weighted estimator is used more than the closed and open for the state statistics in the 10 quarterly hog states. Overall, SRS uses the closed estimator more than the open estimator.

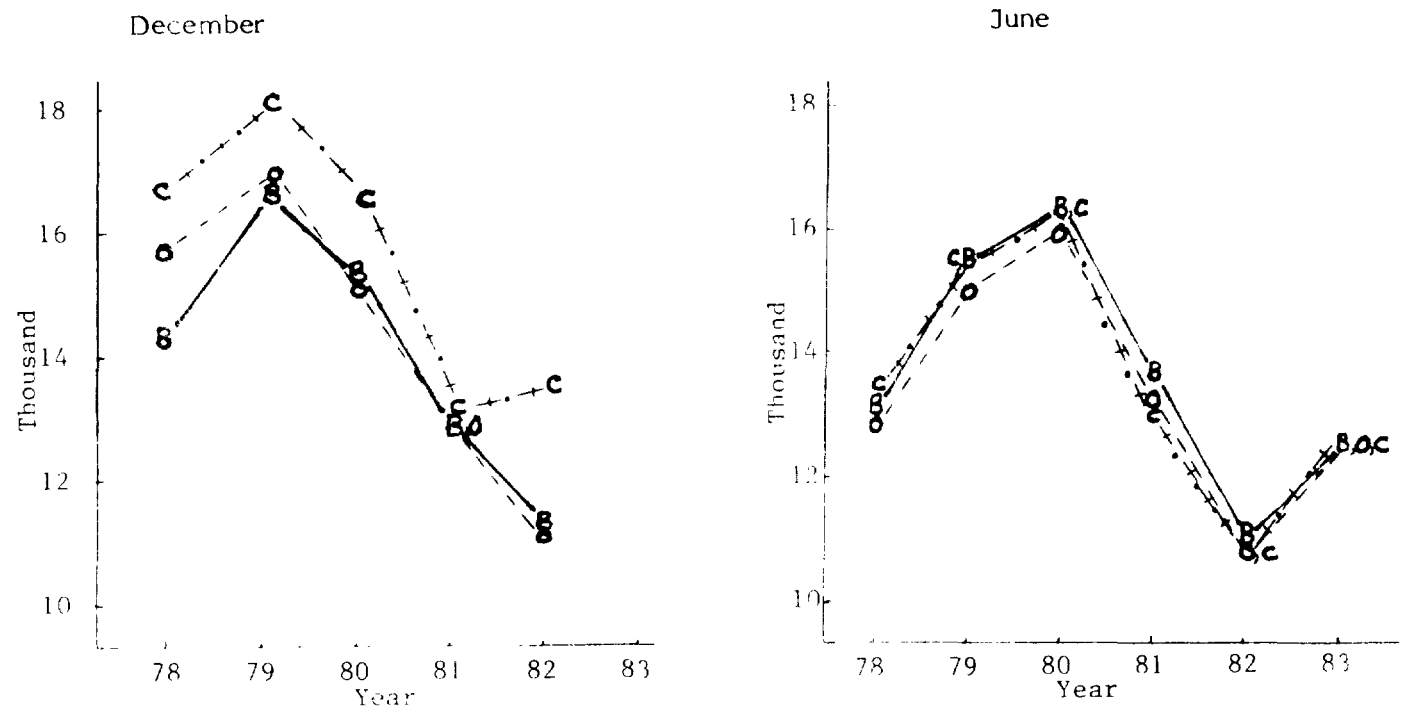
Figure 5: The direct expansion estimates and official statistics for total hogs and pigs for the 10 quarterly states, the other 40 states and the nation; 1978-June 1983.

Notation: M-Multiple Frame O-Open
C-Closed B-Official Statistic

10 Quarterly States



Other 40 States



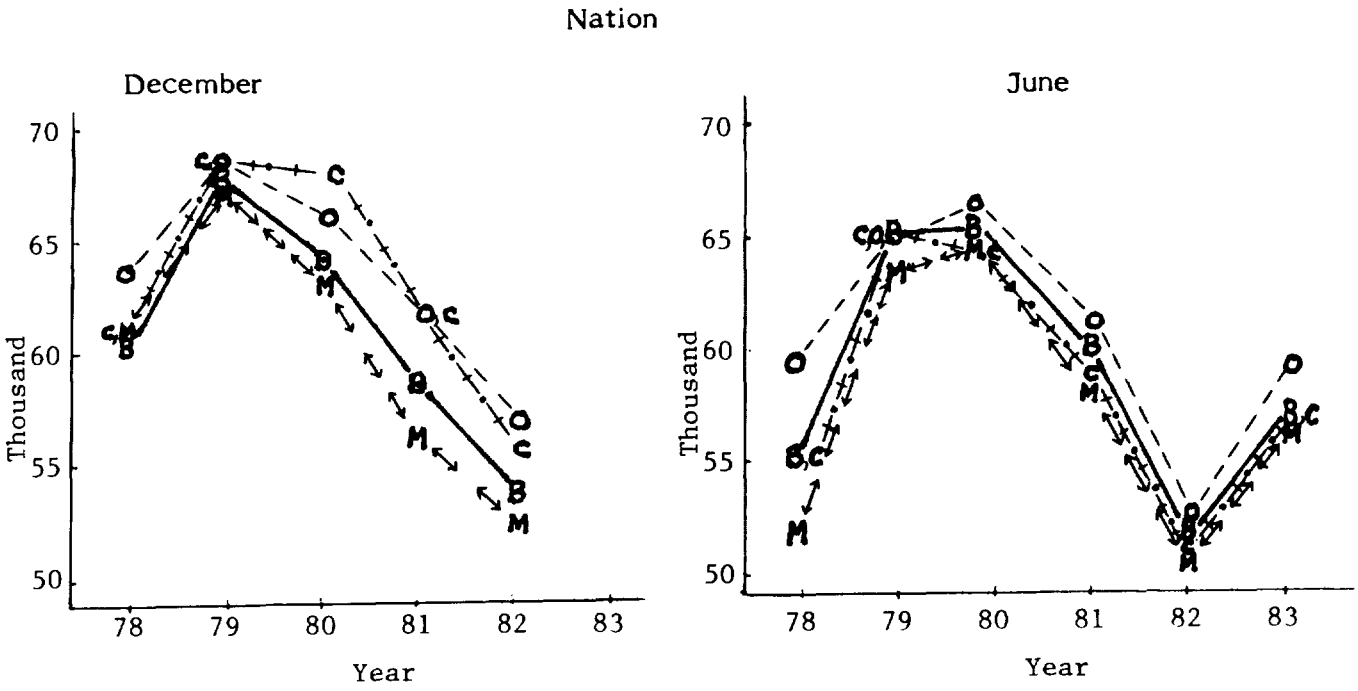


Table 17: The average absolute relative difference between each estimator and the official statistics for total hogs and pigs when the states are combined; 1978-June 1983.

Inference Level	Survey	Average Absolute Relative Difference (%)		
		Multiple Frame	Closed	Open
10 Quarterly States	December	2.2	3.5	4.7
	June	3.5	1.3	3.8
Other 40 States	December	Not Applicable	9.8	2.9
	June		1.3	1.7
National	December	1.5	3.0	3.9
	June	3.2	1.1	2.6

CONCLUSIONS AND RECOMMENDATIONS

The discussions in this report of the advantages and disadvantages of each estimator can be summarized as follows:

(1) The disadvantages of the open estimator outweigh the advantages. In most states, open estimates are too imprecise to provide stable estimates from year to year.

(2) The closed estimator is generally more precise and subject to fewer nonsampling errors than the open estimator. Unfortunately, closed estimates are also unstable at the state level for livestock inventories. The benefits of the closed estimator are most pronounced when estimating crop acreages. For crop acreages such as corn and soybean acres, the closed estimator is superior to the open and weighted estimators.

(3) The major advantage of the weighted estimator over the open and closed estimators is that weighted estimates are more precise. However, this benefit can be negated by the upward bias in the weighted estimates. This bias can be serious in states where a considerable percentage of the farmland is devoted to noncropland.

(4) The multiple frame livestock estimator is the most efficient of all the estimators. However, the multiple frame approach has the potential to increase nonsampling errors, if not executed properly. Also, the nonoverlap component is susceptible to the nonsampling errors of the area frame estimator upon which it is based. For example, a multiple frame crop acreage estimator would be more susceptible to nonsampling errors than the closed estimator.

Numerous suggestions are presented in this report for reducing the nonsampling errors inherent in each estimator. These suggestions should be given serious consideration by SRS.

The comparisons of the livestock estimates from the four estimators with the official statistics showed that:

(1) SRS uses the multiple frame livestock estimates much more than the area frame estimates.

(2) The closed estimator is relied on slightly more than the open and weighted estimators for cattle while the weighted estimator is used the most of the area frame estimators for hogs.

Based on the three factors evaluated in this report for each estimator, namely, the advantages, disadvantages and use, the following changes are recommended for estimating livestock inventories:

(1) For states computing multiple frame livestock estimates, official statistics should be based very heavily or

completely on the multiple frame estimates. Therefore, there is no need to obtain, in addition, two or three area frame estimates in each state to estimate livestock inventories. Area frame estimates for livestock inventories should only be obtained for the nonoverlap domain. Currently, the weighted and/or open segment approach is used to estimate for the nonoverlap domain in a state. It is recommended that all multiple frame states use the same estimator for the nonoverlap domain. It is further recommended that the nonoverlap estimator be the closed or a weighted estimator. The current weighted estimator is not recommended until data collection procedures are developed to eliminate or minimize the bias in this estimator.

(2) For non-multiple frame states, which collectively contain only a very small percentage of the U.S. livestock, both the open and closed estimates are too imprecise and unstable for setting official statistics at the state level. Having two unstable estimates from the same sample does not lead to better state statistics than can be obtained from one estimator. Therefore, each non-multiple frame state should be restricted to a single area frame estimator for livestock inventories. This estimator should be the same area frame estimator used for the nonoverlap estimates in the multiple frame states. This action will simplify the data collection procedures for enumerators and the interview for farm operators, which should result in fewer nonsampling errors associated with the estimates. If, on the other hand, precise livestock estimates are needed in some or all of the non-multiple frame states, then multiple frame methods should be evaluated.

Finally, the following issues concerning the current livestock estimation program should be addressed:

(1) Official cattle statistics are published only at the national level in July. Therefore, the weighted cattle estimates obtained in 10 states during the JES should be dropped since they serve no official purpose.

(2) The time-trend charts for the 10 quarterly hog states combined should include the weighted estimate in the future since the weighted hog estimates were closer to the official statistics than the open and closed estimates in most of these states.

(3) Estimates Division should evaluate the idea of making optimal use of the multiple frame estimates in the time-trend charts when states are aggregated. For example, charts based on the 34 multiple frame cattle states rather than just the 28 states should be explored. Also, charts based on the 23 multiple frame hog states should be evaluated.

REFERENCES

(1) Allen, R., G. Clampet, C. Dunkerley, R. Tortora and F. Vogel. "Framework for the Future". Statistical Reporting Service (SRS), U.S. Department of Agriculture (USDA), March 1983.

(2) Arends, W., R. Addison, R. Young and R. Bosecker. "An Evaluation of Enumeration Techniques and Associated Response Errors and Biases". SRS, USDA, November 1973.

(3) Beller, N. "Error Profile for Multiple Frame Surveys". SRS, USDA, August 1979.

(4) Bosecker, R. "Alternative Weighted Estimator". SRS, USDA, 1980.

(5) Bosecker, R. and B. Ford. "Multiple Frame Estimation with Stratified Overlap Domain". SRS, USDA, July 1976.

(6) Bosecker, R. and W. Kelly. "Summary of Results from Nebraska Survey Concepts Study". SRS, USDA, November 1975.

(7) Bosecker, R. and F. Vogel. "Analysis of 1973 Nebraska June Enumerative Survey and Multiple Frame Survey Livestock Estimates". SRS, USDA, September 1974.

(8) Caudill, C. "Joint Use of Different Sampling Frames". SRS, USDA, 1960's.

(9) Crank, K. "Study of Nonrespondents in Nebraska". SRS, USDA, 1978.

(10) Crank, K. "The Use of Current Partial Information to Adjust for Nonrespondents". SRS, USDA, April 1979.

(11) Dillard, D. and J. Nealon. "Comparison of the Operational and Cropland Weighted Estimators". SRS, USDA, October 1983.

(12) Fecso, R., J. Geuder, B. Hale and S. Pavlasek. "Estimating Dry Bean Acreage in Michigan". SRS, USDA, February 1982.

(13) Ford, B. "Rotation Group Effects in SRS Surveys". SRS, USDA, June 1975.

(14) Ford, B. "The Effect of Land Questions on the Multiple Frame Hog Survey". SRS, USDA, June 1975.

(15) Ford, B. "Missing Data Procedures:A Comparative Study". SRS, USDA, August 1976.

(16) Ford, B. "A General Overview of the Missing Data Problem". SRS, USDA, August 1978.

(17) Ford, B. "Missing Data Procedures:A Comparative Study (Part 2)". SRS, USDA, June 1978.

(18) Gleason, C. and R. Bosecker. "The Effects of Refusals and Inaccessibles on List Frame Estimates". SRS, USDA, August 1978.

(19) Hall, K. "Detecting a Rotational Bias in Multiple Frame Surveys". SRS, USDA, December 1979.

(20) Hall, K. and B. Ford. "The Effects of Data Collection Methods". SRS, USDA, October 1979.

(21) Hill, G. and M. Farrar. "Impact of Nonsampling Errors on Weighted Tract Survey Indications". SRS, USDA, December 1977.

(22) Hill, G. and D. Rockwell. "Associating a Reporting Unit With a List Frame Sampling Unit in Multiple Frame Sampling--Ohio and Wisconsin". SRS, USDA, June 1977.

(23) Houseman, E. "Area Frame Sampling in Agriculture". SRS, USDA, November 1975.

(24) Kelly, W. "Enumerator Evaluation of 1973 December Enumerative and Multiple Frame Hog Surveys". SRS, USDA, March 1974.

(25) Moore, R.P. "Research Report on Four State Multiple Frame Study". SRS, USDA, December 1969.

(26) Moore, R.P. "Research Report on the Amount of Agriculture Associated With Several Incomplete Sampling Frames". SRS, USDA, January 1970.

(27) Moore, R.P. "A Method of Identifying Duplications in Large Lists of Names". SRS, USDA, June 1970.

(28) Moore, R.P. "Research Report on Updating the Tennessee ASCS List". SRS, USDA, July 1970.

(29) Moore, R.P. "The Evaluation of Mail and Telephone Techniques in Snowballing Studies". SRS, USDA, July 1970.

(30) Nealon, J. "The Effects of Omitting Acreage Questions and Modifying the Operation Description Section in Hog Surveys". SRS, USDA, March 1980.

(31) Nealon, J. "The Effects of Omitting Acreage Questions and Modifying the Operation Description Section in Cattle Surveys". SRS, USDA, November 1980.

(32) Nealon, J. "An Evaluation of Alternative Weights for a Weighted Estimator". SRS, USDA, October 1981.

(33) Nealon, J. "Response Errors in the Weighted Estimator". SRS, USDA, July 1982.

(34) "1965 Mississippi Multiple Frame Study". SRS, USDA, January 1966.

(35) "Preliminary Report of Progress". SRS, USDA, July 1966.

(36) "Report on 1956 Research Surveys". SRS, USDA, November 1958.

(37) "Report on 1957 Research Surveys". SRS, USDA, May 1960.

(38) "Report on 1962-1963 Research Projects". SRS, USDA, January 1968.

(39) Rockwell, D. "The Effect on Sampling Errors Resulting From Using a Livestock List With an Area Frame Sample". SRS, USDA, May 1976.

(40) "Scope and Methods of the Statistical Reporting Service". SRS, USDA, September 1983.

(41) Steiner, M. "Study of Hog Balance Sheet and Alternative Ways to Ask Farrowing Questions". SRS, USDA, 1980.

(42) Strand, N., F. Vogel and R.P. Moore. "Frame Construction by the Snowballing Method and Its Evaluation". SRS, USDA, July 1970.

(43) "Summary of Wyoming Multiple Frame Study". SRS, USDA, December 1965.

(44) "Survey and Sampling Concepts Test". SRS, USDA, 1975.

(45) "The June 1965 Study of Livestock Estimates Based on Open and Closed Segment Data". SRS, USDA, March 1966.

(46) Vogel, F. "The Application of Multiple Frame Methodology". SRS, USDA, 1970's.

(47) Vogel, F. "How Questionnaire Design May Affect Survey Data--Wyoming Study". SRS, USDA, March 1974.

(48) Vogel, F. and R. Bosecker. "Multiple Frame Livestock Surveys:A Comparison of Area and List Sampling". SRS, USDA, May 1974.

(49) Vogel, F., R. Bosecker and D. Rockwell. "Multiple Frame Livestock Surveys:An Evaluation of Alternative Methods of Overlap Determination". SRS, USDA, June 1976.

(50) Vogel, F. and D. Rockwell. "Fiddling With Area Frame Information in List Development and Maintenance". SRS, USDA, March 1977.

(51) Vogel, F. and J. Thorson. "Optimizing the Use of List and Area Samples in Livestock Multiple Frame Surveys". SRS, USDA, September 1973.

(52) Weidenhamer, M. "Views on the June Enumerative Survey:A Qualitative Analysis of Discussions With Enumerators and Supervisory Enumerators". SRS, USDA, October 1983.

(53) Young, R. "The Feasibility of Obtaining the Social Security Number From Farm Operators and Its Use in Identifying the Overlap Domain in Multiple Frame Sampling". SRS, USDA, April 1971.