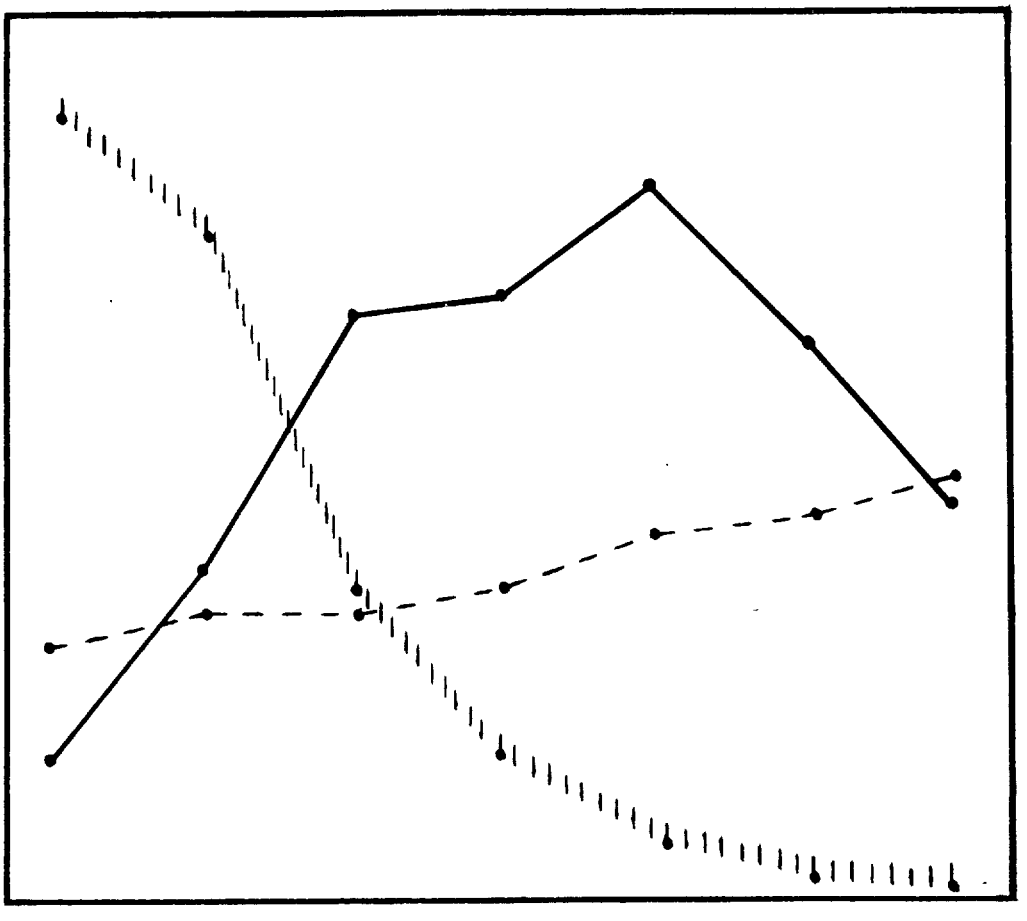
#10

FOR OFFICIAL USE ONLY

ANALYSIS OF 1973 NEBRASKA JUNE ENUMERATIVE SURVEY AND MULTIPLE FRAME SURVEY LIVESTOCK ESTIMATES

REDUCED LIST SAMPLE CONCEPT



Sample Survey Research Branch
Research Division
Statistical Reporting Service
U.S. Department of Agriculture
Washington, D.C.

March 1974

ANALYSIS OF 1973 NEBRASKA JUNE ENUMERATIVE SURVEY AND MULTIPLE FRAME SURVEY LIVESTOCK ESTIMATES

REDUCED LIST SAMPLE CONCEPT

By

Raymond R. Bosecker

Frederic A. Vogel

Sample Survey Research Branch
Research Division
Statistical Reporting Service
U.S. Department of Agriculture
Washington, D.C.

March 1974

CONTENTS

Page

FOREWORD..... i

SUMMARY AND CONCLUSIONS..... 1

INTRODUCTION..... 2

Illustration 1: Reduced List Alternative..... 3

METHODS AND PROCEDURES..... 3

Illustration 2: List-JES Cross-reference..... 5

REDUCED LIST ANALYSIS..... 6

Figure 1: Summary of List Sizes and Sampling Errors, Nebr. 1973 Multiple Frame Hogs and Pigs Estimates..... 7

Figure 2: Summary of List Sizes and Sampling Errors, Nebr. 1973 Multiple Frame Cattle and Calves Estimates..... 8

COST ANALYSES..... 9

NONSAMPLING ERRORS..... 12

RECOMMENDATIONS..... 13

PROPOSALS FOR FURTHER STUDY..... 14

Appendix:

Table 1.--Summary of estimates as list becomes smaller - Nebraska 1973 JES and Multiple Frame hog and pig estimates..... 17

Table 2.--Summary of estimates as list becomes smaller - Nebraska 1973 JES and Multiple Frame cattle and calf estimates..... 18

Table 3.--List and area frame estimates by livestock strata-Nebraska 1973 JES list and area frame estimates..... 19

Table 4.--List size as corresponding hog and cattle strata are added to the list for multiple frame sampling..... 20

Table 5.--Estimated annual survey costs for hog and cattle multiple frame surveys by list and area frame combinations-Nebr., 1973. 21

Table 6.--Nonsampling errors detected in JES..... 22

Table 7.--Nebraska MF survey out-of-pocket costs associated with each group of list strata..... 23

CONTENTS (Contd.)

Page

Table 8.--Nebraska MF survey out-of-pocket costs resulting from personally interviewing all additional NOL tracts..... 24

Table 9.--Nebraska MF survey out-of-pocket costs resulting from mail, telephone and interview of additional NOL tracts in September and March..... 25

Abbreviations:

(1) JES - June Enumerative Survey - June area frame probability survey

(2) OL and NOL - Overlap and Nonoverlap domains - positive or negative name match respectively between area frame and list frame

(3) MF - Multiple Frame - combination of list and area frames

(4) EO - Extreme Operator - livestock operation in highest size group strata

(5) CV - Coefficient of Variation - sampling error as percent of direct expansion

FOREWORD

Estimating livestock inventories by multiple frame methodology was initiated on an operational basis in the mid 60's. Currently the 29 States in the program account for 90 percent of the livestock industry. The multiple frame estimator is a major indicator of livestock numbers and prompts considerable expenditures of funds. During the period that multiple frame estimation was developing as the basic operational tool, the concept that the list frame should be as complete as possible for the item of interest also gradually evolved. As new States were added to the program, SRS policy became one of attempting to achieve a list frame containing at least 90 percent of the item being estimated. As a result some States have more names on their lists than there are farms and in many instances are still not able to account for 90 percent of the livestock.

Perhaps more interest than ever before has been demonstrated in the official SRS estimates due to recent events in the livestock industry culminating in the late summer of 1973 with price freezes and apparent shortages of fat cattle and hogs in the marketplace.

In this setting it is only natural that research of multiple frame methodology has gained top priority. Before initiating large scale research projects, it seemed advisable to analyze existing data in an attempt to guide further research. This study has thus undertaken to examine the optimum mix of list and area sampling frames and to test the assumption that the list frame should be 90 percent complete.

The Nebraska SSO graciously provided their data and assistance in this analysis. Without their effort, this study would not have been possible. Nonsampling errors found in this analysis should not be viewed as a reflection on the quality with which the Nebraska SSO conducts the multiple frame survey but rather a result of the operational procedures used by SRS. It is important that we realize total error of an estimate is composed of sampling error and nonsampling error. Procedures that reduce sampling error by one or two percent may not be justified if we unknowingly increase nonsampling error.

The results of this analysis should provide basic facts upon which to develop additional research projects as well as provide guidance in operational policies if so desired. It should not be used to show that general purpose lists are unnecessary but that the entire list need not be sampled and used to determine the overlap domain for a specific item of interest.

Chief, SSRB *NORMAN D. BELLER*

SUMMARY AND CONCLUSIONS

This study was undertaken to examine the optimum list size to use in livestock multiple frame (MF) surveys. It is also part of an overall effort to identify reasons for level differences in livestock estimates from the list and area frames. Initial findings from this research on one State, Nebraska, are as follows:

A. Smaller size group strata may be excluded from the list sample and estimated by the area frame as part of the nonoverlap (NOL) domain by allowing a relatively small increase in the sampling error.

B. Deletion of the lower size group strata significantly reduces the size of the list for MF sampling. Reduction in the size of the list frame may allow more time for duplication removal, identification and handling of joint arrangements, the identification of overlap tracts, and the detection of nonsampling errors.

C. Deletion of list strata by substituting the area frame changes the level of the MF estimate. This, in effect, artificially reduces the difference between the estimates by allowing the area frame to contribute more to the MF estimate.

D. Sampling from the universe list of nearly all farms in the State makes nonoverlap a rare item to be estimated by the area frame which is inefficient at estimating rare items. (Also, States checking only nonrotated segments for NOL contribute to reducing NOL representation.)

E. More should be done during the computer edit and summary phase of the survey to alleviate the data manipulation now required of the statistician.

F. Two types of nonsampling errors were encountered:

   (1) Extreme operator data summarized in the area frame.

   (2) Misclassification of NOL tracts because the list for the MF sample was not separate and unique from the universe list.

   Correction of these errors moved the list frame and area frame estimates closer together.

G. Mail and telephone data collection procedures for March and September NOL tracts should be seriously considered. This, combined with sampling fewer list strata, can result in significant cost savings.

H. A code box on JES questionnaires to cross-reference overlap tracts with the list and designate the corresponding livestock strata and overlap status would help in the following ways:

   (1) Provide consistency in overlap designation

   (2) Aid in data handling by:

      a. Designating EO's for automatic exclusion

b. Designating nonoverlap tracts (including hog or cattle strata below a specified cutoff) for automatic nonoverlap estimation.

(3) Permit further research of the reduced list concept.

INTRODUCTION

Enumerative and multiple frame surveys provide the only indications for estimating livestock inventories in the major producing States. These surveys are based on probability sampling designs that provide unbiased estimators for the universe or population. In theory, the only difference that should exist between the area frame and the multiple frame estimators is caused by sampling variation in the two independent samples of the overlap portion of the universe.

In practice, the levels from these estimators are further apart than can be attributed to sampling variation alone. This difference is therefore due to problems in population definition, sample selection, or survey procedures, which lead to a departure from sampling theory. An indepth analysis of the JES and MF survey data for one State was therefore undertaken to provide guidance in future research and perhaps operational decisions pertaining to the most efficient and effective use of list and area frames.

The primary objective of this initial study was to provide a basis for determining the optimum size of list frame to use in conjunction with the enumerative survey. A secondary objective was to identify and evaluate some sources of nonsampling errors that are contributing to the level differences.

The primary objective relative to the best list size for multiple frame surveys is a test of the current assumption that the States must maintain as complete a list as possible for sampling the MF survey. Additional assumptions currently being made with regard to MF and enumerative surveys will be tested in future studies.

Many States are now working with a list of 50,000 to 100,000 names and sample sizes of around 2,000 names. A large proportion of these list names are classified in strata with few livestock. If the list can be reduced to a more manageable level, more attention could be given to nonsampling errors. Resources saved from working with a smaller list could then be used to reduce total error. The changes in survey costs which would occur with a smaller list and a larger NOL sample from the area frame are presented in the report.

Under current instructions for multiple frame and area frame surveys, the NOL domain is to be as small as possible (i.e. the overlap (OL) domain is maximized.) The current estimator results in using the area frame to estimate a rare item where it is very inefficient. The Livestock Branch has noted wide fluctuations in the NOL portion of the MF surveys. A balance between the list estimator for the more specialized operations where it is most efficient and the area estimator for the general farms should be established so they complement each other and provide stability.

The following illustration shows the reduced list sample concept which is under investigation in this report.

Illustration 1: Reduced List Alternative

List strata	Hogs	Cattle	Current surveys MF	Current surveys ES	Possible alternative surveys MF	Possible alternative surveys ES
79	4,000+	25,000+	EO	EO	EO	EO
78	500-3,999	2,500-24,999	EO		EO	
6	300-499	200-2,499	6	OL	6	OL
5	200-299	100-199	5		5	
4	125-199	50-99	4		4	
3	1-124	25-49	3		3	
2	No hogs	0-24	2			
1	No lvstk.	No lvstk.	1		NOL	NOL
88	Nonoverlap	Nonoverlap	NOL	NOL		

The NOL domain would account for a larger proportion of all operators but the list would still account for the greatest share of the livestock. The extreme operator (EO) portion is identical for both list and area frames and is really part of the list. This points out the fact that the enumerative surveys are also multiple frame in that the EO list is combined with tract and farm area expansions for the enumerative estimates.

The second objective, the evaluation of nonsampling errors, is being undertaken in an effort to alert the States to potential problem areas and suggest ways of catching these errors before they are summarized.

Nebraska fully participated in the search for nonsampling errors for this study.

METHODS AND PROCEDURES

Nebraska was chosen for special JES and MF analysis for several reasons:

(1) The staff indicated a real interest in the study and gave their full assistance and cooperation;

(2) Nebraska had a new interpenetrating area frame sample for 1973 JES;

(3) A complete current NOL check was done on all 1,683 June survey tracts;

(4) A listing of all JES tracts had been prepared and cross-referenced with the MF universe list for overlap tracts.

The following data were made available to the Research Division from Nebraska:

(1) Universe list on magnetic tape and printout

(2) Listing of 1973 JES area frame tracts

(3) Listing of EO names and addresses

(4) 1973 JES questionnaires

(5) Pink nonoverlap (A) questionnaires for both OL and NOL tract operators

(6) State Farm Census data for calendar year 1972

(7) Data cards for Dec. 1972 MF Survey

(8) Edit tapes for Cattle and Hog MF surveys.

In addition, 1973 JES data for Nebraska was obtained from the regional data tapes in Washington, D. C. The Statistical Analysis System (SAS) program package was used to examine the data.

A SAS program was prepared to summarize the tract data to an expanded segment total and compute sums and variances for each land use stratum. Using SAS it was then possible to duplicate the work of the JES summary system and also display and examine the data in many different ways.

The first step in assimilating the different data sets into a single package was to use the list ID number and segment tract cross-reference to match the JES data with the list control data. A copy of the resulting printout with list information on the left and JES data on the right is shown in Illustration 2. After the appropriate livestock strata were associated with the JES hog and cattle data for overlap tracts and the original JES NOL tracts were coded, the farm and tract NOL expansions could be recomputed through any level of livestock strata.

The NOL contribution to the MF estimate in this study was computed for the tract and entire farm in the same way as for the overall area frame estimates. This differs from the current MF procedure in two ways.

First, the MF NOL estimate uses the weighted approach where the number of head on the entire farm is ratioed to the tract by the proportion of tract acres to total acres. This estimate ignores the differences between the tract and entire farm estimates.

Secondly, the weighted NOL expansions are currently treated as a simple random sample of tracts (Code 88) for the MF survey. In this study NOL was stratified by land use strata as are all items in the area frame. This affects only the variance computations, not the expanded totals. However, it does maintain comparability between the multiple frame and area frame standard errors in this study.

Costs incurred in the larger sample from the nonoverlap domain were compared with the costs of enumerating names selected from the list frame.

Illustration 2: List-JES Cross-reference

STATISTICAL ANALYSIS SYSTEM

	List									JES											
OBS	UNIVID	SEG	TRACT	DIST	COUNTY	TOTLAND	CATTLE	HOGS	NAME	I840	I846	I900	I2101	I2117	I2505	I3005	I3505	I3710	CTLSTR	HOGSTR	PART
1	40689	3001	1	30	27	260	16	152		160.0	0	160.0	0	0	37.0	223	17	37	0	2	4	1.0
2	26188	3001	2	30	27	280	25	194		320.1	0	0.0	0	0	47.0	0	0	0	0	3	4	1.0
3	900001	3001	3							160.0	0	0.0	0	0	0.0	0	0	0	0	0	0	1.0
4	47984	3002	1	30	3	1720	77	52		150.0	0	1260.0	0	0	245.0	0	0	245	15	4	3	1.0
5	900005	3002	2							160.0	0	0.0	0	0	0.0	0	0	0	0	0	0	1.0
6	47983	3002	3	30	3	480	47	70		156.0	0	0.0	0	0	0.0	0	0	0	0	3	3	1.0
7	24354	3002	4	30	3	320	20	106		312.0	0	572.0	156	50	26.0	156	50	26	3	2	3	1.0
8	900004	3002	5							1.0	0	76.0	0	0	0.0	0	0	0	0	0	0	1.0
9	69706	3003	1	30	21	400	35	152		80.0	0	400.0	64	0	24.0	64	0	91	0	3	4	0.5
10	7120	3003	2	30	21	400	28	203		80.0	0	0.0	0	0	0.0	0	0	0	0	3	5	1.0
11	35301	3003	3	30	21	480	29	0		79.9	0	0.0	0	0	55.0	0	0	0	0	3	2	1.0
12	61329	3003	4	30	21	575	121	0		134.0	0	0.0	0	0	0.0	0	0	0	0	5	2	1.0
13	30914	3003	5	30	21	280	37	0		25.0	0	0.0	0	0	0.0	0	0	0	0	3	2	1.0
14	900008	3003	6							18.0	0	18.0	0	0	6.0	0	0	6	0	0	0	1.0
15	28373	3003	7	30	21	280	0	0		60.0	0	0.0	0	0	0.0	0	0	0	0	1	1	1.0
16	23357	3003	8	30	21	560	53	328		79.0	0	0.0	0	0	0.0	0	0	0	0	4	6	1.0
17	35545	3003	9	30	21	80	5	111		80.0	0	0.0	0	0	0.0	0	0	0	0	2	3	1.0
18	61263	3004	1	30	119	400	29	321		237.0	0	400.0	192	21	93.0	192	21	93	3	3	6	1.0
19	14288	3004	2	30	119	320	2	260		80.0	0	0.0	0	0	0.0	0	0	0	0	2	5	1.0
20	14282	3004	3	30	119	360	9	244		120.0	0	600.0	138	0	45.0	138	0	64	12	2	5	0.5
21	5138	3004	4	30	119	260	4	113		40.0	0	0.0	0	0	0.0	0	0	0	0	2	3	1.0
22	900012	3004	5							2.0	2	0.0	0	0	0.0	0	0	0	0	0	0	1.0
23	25985	3004	6	30	119	480	21	297		159.0	0	0.0	68	0	51.0	0	0	0	0	2	5	1.0
24	18604	3005	1	10	181	1360	7	0		568.0	0	0.0	0	0	81.5	0	0	0	0	2	2	1.0
25	43870	3005	2	10	161	2300	13	78		80.0	0	0.0	0	0	0.0	0	0	0	0	2	3	1.0
26	17219	3006	1	10	105	5000	79	0		494.9	0	4880.0	0	0	0.0	0	0	171	0	4	2	1.0
27	900013	3006	2							12.0	12	0.0	0	0	0.0	0	0	0	0	0	0	1.0
28	17554	3006	3	10	105	6040	113	0		160.0	0	0.0	0	0	0.0	0	0	0	0	5	2	1.0

REDUCED LIST ANALYSIS

The results of allowing the Nebraska area frame to estimate for each succeeding livestock list stratum from one through six are demonstrated in Table 1 for hogs and pigs and Table 2 for cattle and calves. (All tables are in the appendix.) In Table 1, under the original multiple frame concept, the tract direct expansion estimate of hogs and pigs was 3,071,200 head with a coefficient of variation (CV) of 4.2 percent. As the NOL domain became larger, i.e. estimating for each succeeding hog stratum, the multiple frame estimate moved upward. This reflects the higher level of hogs estimated by the area frame as compared with the list frame. The CV increases at a rather slow rate as the portion of the total universe estimated by the NOL domain increased with elimination of the lower list strata.

When the NOL domain was extended from the original NOL to include hog strata one to three, the CV was approximately 6 percent while the list frame sample size for the hog MF survey was reduced from 1,748 to 773. The number of nonoverlap tracts increased from the original 350 to 1,104 NOL tracts. If the universe list consisted only of strata four and larger hog operators the total size of the list would be about 11,000 names. The multiple frame expansions and sampling error coincide with the area frame estimates when the NOL domain is extended through livestock strata six. The CV increased from the original 4.2 percent for the combined list and nonoverlap estimate to about 10 percent for the JES area frame tract estimate. Table 1 results are presented graphically in Figure 1.

Table 2 for cattle and calves shows the original multiple frame estimate resulted in a direct expansion of 7,608,800 head with a CV of 3.5 percent using a tract NOL expansion. The direct expansion then increased rapidly as the NOL domain estimated for cattle and calf strata one through four to a high of 8,376,000 head, well above the total JES area frame estimate. It then dropped for the last two strata to the June Enumerative Survey area frame tract estimate of 7,929,300 head with a CV of 5.4 percent.

Comparable data is also shown for the entire farm estimate. The MF CV using the entire farm data began at 5.1 percent and increased to 10.5 percent for the area frame. When nonoverlap estimates for cattle strata one through three, the list sample decreased by 687 names and the total cattle list size became 10,467 names. Direct expansions, sample errors and list sizes from Table 2 are shown graphically in Figure 2.

If the area frame and list frame estimated approximately the same level for each stratum, the direct expansion lines in Figures 1 and 2 would be more nearly horizontal. For cattle especially, it was interesting to note the wide fluctuation in the direct expansion level for different combinations of the list and area frames while the CV remained fairly consistent. This emphasizes that something more than sampling error is responsible for differences between the list frame and area frame estimates.

FIGURE 1

SUMMARY OF LIST SIZES AND SAMPLING ERRORS, NEBR. 1973

MULTIPLE FRAME HOGS AND PIGS ESTIMATES

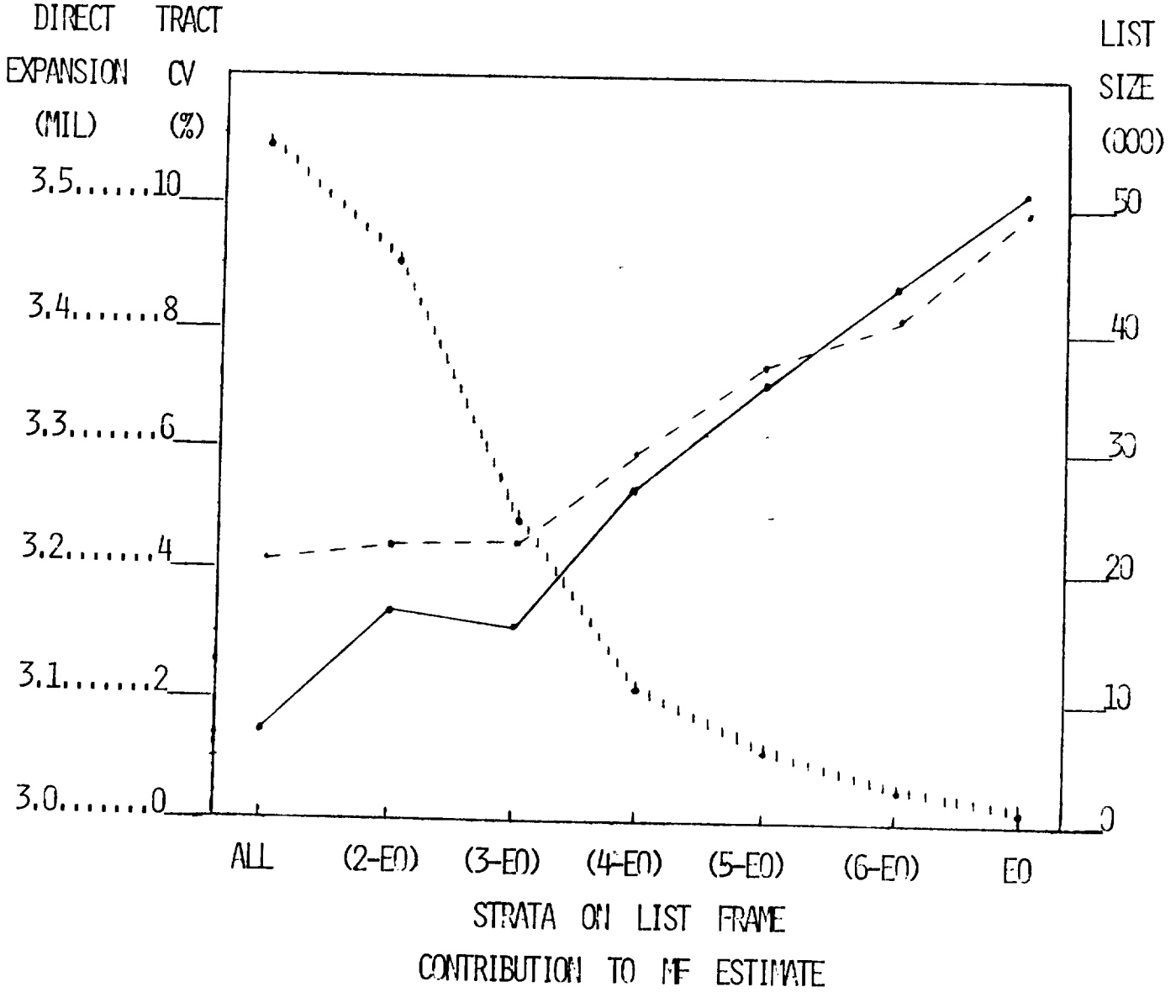
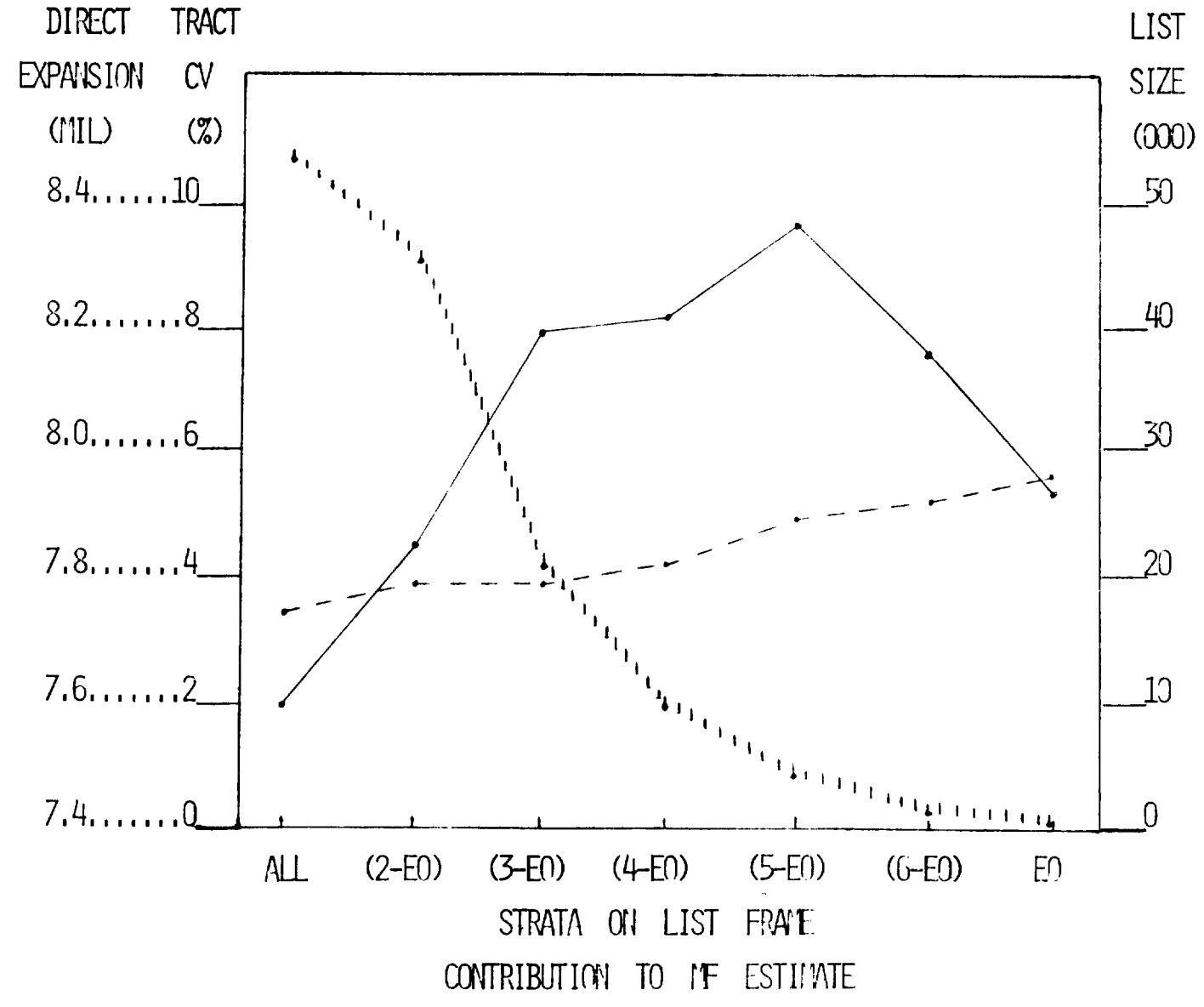


FIGURE 2

SUMMARY OF LIST SIZES AND SAMPLING ERRORS, NEBR. 1973

MULTIPLE FRAME CATTLE AND CALVES ESTIMATES



DIRECT EXPANSION _____

CV % - - - - - - - -

LIST SIZE . . . . . .

In Table 3, the tract and farm area frame expansions for each livestock stratum are compared with the list estimate. The coefficients of variation in strata one and two for hogs and cattle are not materially different for the two frames. However, the levels of the estimates in the zero livestock strata differ by nearly 100 thousand hogs and 230-400 thousand cattle. The list and area frame estimates will most assuredly move closer together without much loss in the CV of the multiple frame estimator when the lower strata are dropped from the list frame. However, in so doing, the assumption will have been made that the larger estimate from the area frame for the lower livestock strata is superior, i.e. more correct, than the smaller estimate given for the same strata by the list frame.

The area frame was above the list frame expansion for most strata. The notable exceptions are strata 5 and 6 for cattle where the area frame expansions were well below the list frame. These results for one State suggest that the expansions obtained by the procedures used in the two frames may be dependent on the size of the livestock operation. Direct expansions of hogs and cattle in the zero livestock stratum using the list sample were about one-third that of the area frame expansions. On the other hand, cattle operations indexed at over 100 head expanded to nearly 50 percent more than in the area sample.

Sample sizes from the list and area frames are also provided for comparison. As expected, there are a large number of tracts in the smaller size group strata and few in the larger size groups. The relationship between farms and tracts in the area frame was about 3.25 tracts per farm. The expanded number of overlap farms from the area frame was approximately 53,200 compared to the list size of about 54,200.

Tables 1 and 2 demonstrate that replacing the lower livestock strata on the list with the nonoverlap area frame estimator increased the CV moderately while it substantially reduced the list size. Table 3 shows the list estimator is clearly more efficient than the area frame in the larger size livestock strata.

The bulk of the names which must be maintained and updated are in the lower livestock strata. Table 4 shows the size of the list for each level of cattle and hog cutoff limits. For example, if the cattle lower limit was an index of 50 and the hog index was 125, the Nebraska list would have been comprised of 19,200 names rather than 54,000 names. Other combinations are also possible. If cattle had a lower limit index of 100 and hogs a lower limit index of 125, computations showed the list size would be approximately 14,500 names. The possibilities of duplication within the list or mistakes in checking nonoverlap may be substantially diminished by using a list of 14,500 names rather than a list of 54,000 names.

COST ANALYSES

Analyses presented in previous sections showed that some of the small size group strata could be eliminated from the list frame without seriously affecting sampling error. The decrease in sample size from the list frame will be offset by a larger number of area frame tracts that become NOL.

One purpose of decreasing the size of the list frame for sampling purposes is the elimination of some nonsampling errors resulting from the survey workload. Just as important, however, is the effects of such an action on the survey costs. For example, under current survey procedures sample units from the list surveyed by relatively inexpensive procedures are replaced by sample units from the area frame surveyed by a more expensive procedure - personal enumeration. Therefore, the problem is to weigh the merits of a reduction in list size in terms of survey costs.

The most difficult part of the cost analysis was determining the cost to apply to each questionnaire returned by mail, telephone, or personal interview. Cost structure and response rates by enumeration technique vary by State. In addition, MF cost data becomes intertwined with costs of conducting other surveys. Only data collection costs were considered. Savings which could result from maintaining and updating a smaller list were not considered in this analysis.

The estimated MF survey costs associated with sampling fewer list strata are presented in Table 5. Total MF survey cost is comprised of the cost for the list sample and the cost of collecting data for area frame NOL tracts. As fewer list strata are sampled the list cost decreases and the cost for the NOL domain increases. The average costs per completed questionnaire were estimated to be $.50 for each mail return, $1.00 for each completed telephone interview, and $15.00 for a personal interview. The number of list frame questionnaires and area frame tracts which contribute to the MF costs shown in Table 5 are presented in the Appendix Tables 7, 8 and 9 so survey costs may be recomputed under different rates per completed questionnaire.

Alternative survey costs for various reductions in the list sample and two different data collection procedures for NOL tracts are presented in Table 5. Under current survey procedures all NOL domain data is obtained by personal interview. Costs associated with this procedure are shown in column 2. The MF survey costs resulting from replacing the list sample with NOL tracts enumerated at $15.00 per interview are increased as demonstrated in column 4. To illustrate, if list strata 1 and 2 were not sampled but all NOL tracts in these strata were personally interviewed in December (cattle or hogs) and March-September (hogs only) then the MF surveys for the year would cost $8,356 more than at present.

An alternative would be to use the mail and telephone enumeration procedures to obtain the data for nonoverlap tracts. This procedure is feasible for March and September because the name, address, and telephone number of the tract operator along with tract and farm acreage have been established by personal interview in June and can be verified in the December Enumerative Survey. Only a change in the entire farm acreage reported in the March or September surveys would require a personal interview to determine if the tract acreage was affected for proration purposes.

Column 3 shows the costs associated with the nonoverlap domain under the assumption that NOL data for weighted estimates are obtained by mail, telephone and interview with response rates comparable to the March and September 1973 list frame returns in those strata. With this alternative procedure, the additional nonoverlap tracts in livestock strata 1-4 would cost $10,893 over a year's time. Since it costs $20,049 less to collect the data for the list sample when the lower four livestock strata are deleted from the list, this results in an overall reduction in total cost of $9,156 during the year.

The cost differences between columns 2 and 3 show that considerable savings are possible if total enumeration of NOL tracts by interview can be avoided. Subsamples of 151 and 174 original NOL tracts were surveyed in March and September 1973 respectively. As shown for the entire list, assuming a response rate comparable to strata 1-2, $4,155 in savings results from a mail, telephone and interview contact of these 325 original NOL tracts.

The results of this analysis do not mean MF surveys should be curtailed to save money but that funds freed by a change in procedure without sacrificing sampling error could be reinvested in additional area samples to benefit the entire SRS estimating program and provide MF cattle and hog estimates with a lower total error.

The cost analysis in Table 5 is based on several additional assumptions:

(1) No list survey costs were associated with the sample questionnaires estimated, inaccessible, or known zero but refusals were included in the count of list names surveyed.

(2) No cost was given to the additional entire farm supplement "pink" questionnaires which would be required in June for more nonresident NOL tracts.

(3) No charge to the MF survey was made for NOL tracts enumerated as part of the June survey. However, the NOL tracts occurring in the December Enumerative Survey were included in the cost analysis.

(4) An estimated 40 percent of any additional NOL agricultural tracts are already in the December sample currently as overlap tracts. It is unlikely that all of the remaining tracts would have to be included in the December sample but was assumed for this cost study. The December contribution to higher cost under this assumption is at least partially offset by improved area frame estimates for all items.

(5) Nonoverlap tracts are currently subsampled in the March and September hog and pig surveys. An approximate sampling rate of 50 percent of the additional NOL tracts for each survey was assumed. This is consistent with current samples and allows all those reporting hogs on the farm in June to be surveyed in September and March plus a sizable sample of those reporting no hogs.

(6) Hog and cattle strata remaining as the list becomes smaller would be sampled at current rates and maintain a comparable CV.

(7) The JES and DES currently provide estimates of the NOL domain for both the June-December hog surveys and the July-January cattle MF surveys. The cost analysis assumes this practice would continue if the nonoverlap domain were allowed to increase in size.

Additional savings, regardless of method of enumeration, would result by further subsampling the increased number of nonoverlap tracts in the March, September and December surveys, especially those reporting zero livestock in June. Costs presented in Table 5 are considered to be near maximum costs under a reduced list sampling plan.

The cost analysis in Table 5 relied upon estimated average costs incurred by Nebraska. These costs will differ by State. For example, all telephoning in Nebraska is done in the SSO on State and federal leased lines with local telephone enumerators. This results in a substantially different cost function compared with costs incurred by enumerators calling from their homes.

NONSAMPLING ERRORS

The nonsampling errors which were discovered while working with the Nebraska June Survey were all either directly or indirectly associated with the magnitude of the job done in June. In the two-week survey period nearly 1,700 tracts were checked and rechecked against a list of over 54,000 names to determine who was overlap and who was nonoverlap. Basically, two types of oversights were discovered during this analysis with the summary programs which were used. First, extreme operator data remained in the area frame to overstate the tract estimate, and secondly, NOL tracts classified as overlap understated the multiple frame estimate. The results of extreme operator data in the tract questionnaires and misclassification of nonoverlap tracts are shown in Table 6. When these errors were corrected the area frame and multiple frame estimates moved somewhat closer together.

Steps were already being taken by the Nebraska office before this study began to correct the circumstances which led to some of these errors. A unique universe list for multiple frame selection was to be implemented in place of the 1973 universe of 64,000 names of which only 54,000 had a chance for multiple frame selection. Even though those in the universe which had no chance of selection were coded, the code was sometimes missed when checking the area frame against the list frame.

A printout of segment expansions quickly pointed out unusual segments for closer analysis. This led to the discovery of the extreme operator who expanded to 326,000 hogs in the tract. His nickname which he consistently uses locally was recorded on the tract questionnaire and failed to match with his legal name on the extreme operator list.

The tract listing with universe information on the left side matched with June enumerative data on the right side (Illustration 2, page 5) uncovered the other errors shown in Table 6. When a person on the list who was coded as not belonging to the MF sampling universe appeared with an overlap ID it could easily be detected. The new procedure where every name listed in the MF universe has a chance of selection will make this check unnecessary. Likewise, it was easy to spot a report that had control data indicating it was an EO yet showed tract or entire farm data for that specie. Data deletion, proration, and domain classification by computer would remove the errors which now occur from data manipulation and assignment during editing. This would reduce the need for a tract by tract printout from an editing tool to simply a reference source.

RECOMMENDATIONS

On the basis of this study it is recommended that one or two States be placed on a reduced list sample basis. This will permit further investigation of possible gains in resources which might then be applied to reduction of nonsampling errors and analysis of other assumptions being made in the multiple frame surveys. These test States which use a reduced list and thereby increase the NOL domain should attempt to survey the NOL sample in September and March with the same data collection procedures used for the list sample. Detailed cost data should also be obtained for further analysis.

The nonsampling errors discovered during the study in a State which has a very good record in this regard points up several areas where the operational procedures could be improved in all States.

(1) Let machine edit and summary replace a portion of the data alterations now done by the statistician. The summary system employed in this study automatically excluded EO data from being summarized in the area frame and also applied the partial NOL factor to the data where applicable. To do this, a code box needs to be incorporated into the face page of the enumerative survey questionnaire for the MF identification number, the partial NOL factor, and the list livestock control numbers for overlap tracts. A request for the special code box has already been formally presented to the appropriate organizational units.

(2) Allow the States to maintain one MF list which includes the extreme operators who would be identified by their control number being above the cutoff number. A match with the ID number for an EO or a livestock stratum above the EO cutoff would then automatically exclude that specie data from analysis without the statist having to physically cross it off the questionnaire. All index numbers should be comparable. EO's would simply be the upper two strata of the MF list. The MF printout for checking NOL tracts should be unique, contain no other names, and correspond exactly with the universe from which the MF sample was drawn.

(3) Provide States with a program which will summarize tracts to an expanded segment total basis for editing purposes. Nebraska was quite interested in getting such a summary printout when they discovered it was available in D. C. This will uncover busts in the data and provide the States data for analysis purposes prior to making an estimate.

(4) It would be desirable for the States to be able to see the raw or expanded data tract by tract, together with any corresponding MF control data and the MF name versus the JES name on the same printout. This would also be made possible by cross-referencing the MF identification with segment and tract for overlap tracts. States could then be as careful with their classification of overlap as they have in the past with the nonoverlap.

(5) Summarization of area frame overlap data by livestock strata as done in Table 3 would also be a valuable tool for discovering busts or outlier reports in either the area frame or list frames.

Suggested actions on estimating procedures in all States are:

(1) Compute tract and entire farm direct expansions for the NOL domain in June as well as the weighted segment expansion. The tract (closed segment) and farm (open segment) NOL expansions are the NOL contribution to the full area frame estimate while the weighted segment expansion is the NOL contribution to the multiple frame estimate. A comparison of these estimators from exactly the same respondents should be made to determine the impact of each on the respective frames.

If a weighted segment estimate were currently used for the entire area frame then the NOL domain estimated by the area frame would be identical for both frames. The list frame estimate for the EO strata should also be identical for both frames.

(2) Do a current NOL determination on all segments in the June survey rather than on only nonrotated segments. The NOL tract is already a rather rare item so further reduction of the sample for the NOL domain should be avoided.

PROPOSALS FOR FURTHER STUDY

First, research and analysis in other States similar to this Nebraska study is planned and under way. This requires the States to cross-reference the multiple frame list with the June Enumerative Survey overlap tracts. The cattle strata and hog strata for each overlap tract are recorded and any partial nonoverlap identified. The nonoverlap domain will then be resummarized through any given livestock strata. The check for nonsampling errors through misclassification of nonoverlap tracts is done by the States themselves as they are coding the overlap tracts. Results will be issued as a supplement to this report.

Secondly, a project should be initiated to determine if livestock data from the nonoverlap domain can be obtained by mail or telephone. Also, other studies could be made using additional Nebraska information available. This includes:

(1) Matching State Farm Census with Multiple Frame and JES land and livestock data for comparable periods;

(2) Comparing State Farm Census data for Nebraska survey refusals with State Farm Census data for survey respondents;

(3) Post-stratification by livestock strata for the area frame overlap domain;

(4) Resummarizing the JES and MF data under different definitions of overlap and partial overlap;

(5) Comparing data imputation procedures to handle refusals;

(6) Computing weighted area frame expansions for all JES data;

(7) Applying other types of estimators to the current multiple frame data.

Finally, assumptions currently being made which should be tested include:

(1) Sampling unit and reporting unit are equivalent;

(2) List units can uniquely identify the sampling units in the population;

(3) Equivalent livestock data are being obtained by mail, telephone and personal interview;

(4) The same farm operation is described by mail or telephone contact as results from a personal interview with sketches of the land operated;

(5) Partnerships are as clearly defined by mail as they are in an interview;

(6) The overlap domain for the area frame is in fact estimating for the same universe as the list frame;

(7) Samples are being selected with a known probability.

These assumptions may be leading to unknown errors which are forced into the survey by design. There are many different sources of nonsampling errors.

A P P E N D I X

NEBRASKA MF SURVEY ANALYSIS

TABLES 1-9

Table 1.--Summary of estimates as list becomes smaller - Nebraska 1973 JES and Multiple Frame hog and pig estimates

Multiple frame	Direct expansions using tract and farm estimates of nonoverlap domain						Universe and sample size	
	Tract (210)			Farm (300)			N	n
	DE	SE	CV	DE	SE	CV		
	(000)	(000)	(%)	(000)	(000)	(%)		
List (Orig.)	2,601.1	82.7	3.2	2,601.1	82.7	3.2	54,193	1,748
NOL (Orig.)	470.1	97.2	20.7	517.2	103.5	20.0	-	350
Total (Orig.)	3,071.2	127.6	4.2	3,118.3	132.5	4.2		
List-Str. 1	2,559.2	81.0	3.2	2,559.2	81.0	3.2	45,795	1,554
NOL+Str. 1	607.0	109.2	18.0	621.2	114.6	18.4	-	493
Total	3,166.2	136.0	4.3	3,180.4	140.3	4.4		
List-Str. 1,2	2,468.0	77.6	3.1	2,468.0	77.6	3.1	24,877	1,286
NOL+Str. 1,2	684.1	112.7	16.5	701.1	117.8	16.8	-	886
Total	3,152.1	136.8	4.3	3,169.1	141.1	4.5		
List-Str. 1-3	1,931.3	69.9	3.6	1,931.3	69.9	3.6	11,130	773
NOL+Str. 1-3	1,336.2	180.6	13.5	1,342.1	185.4	13.8	-	1,104
Total	3,267.5	193.7	5.9	3,273.4	198.1	6.1		
List-Str. 1-4	1,440.5	63.6	4.4	1,440.5	63.6	4.4	6,359	546
NOL+Str. 1-4	1,914.2	234.3	12.2	2,014.9	249.9	12.4	-	1,180
Total	3,354.7	242.8	7.2	3,455.4	257.9	7.5		
List-Str. 1-5	1,023.3	58.1	5.7	1,023.3	58.1	5.7	3,362	360
NOL+Str. 1-5	2,410.6	278.0	11.5	2,627.0	325.3	12.4	-	1,234
Total	3,433.9	284.0	8.3	3,650.3	330.4	9.1		
List-Str. 1-6 (EO)	492.3	48.9	10.5	492.3	48.9	10.5	840	144
NOL+Str. 1-6	3,024.3	343.7	11.4	3,255.5	447.6	13.8	-	1,279
Total	3,516.6	347.2	9.9	3,747.8	450.3	12.0		
JES Area Frame	3,516.6	347.2	9.9	3,747.8	450.3	12.0		
Bd.	3,250			3,250				

Table 2.--Summary of estimates as list becomes smaller - Nebraska 1973 JES and Multiple Frame cattle and calf estimates

Multiple frame	Direct expansions using tract and farm estimates of nonoverlap domain						Universe and sample size	
	Tract (250)			Farm (350)			N	n
	DE	SE	CV	DE	SE	CV		
	(000)	(000)	(%)	(000)	(000)	(%)		
List (Orig.)	5,947.5	163.1	2.7	5,947.5	163.1	2.7	54,180	1,418
NOL (Orig.)	1,661.3	214.1	12.9	1,370.9	339.5	24.8	-	350
Total (Orig.)	7,608.8	269.1	3.5	7,318.4	376.6	5.1		
List-Str. 1	5,851.7	160.3	2.7	5,851.7	160.3	2.7	45,814	1,275
NOL+Str. 1	1,990.7	257.4	12.9	1,869.7	382.4	20.4	-	493
Total	7,842.4	303.2	3.9	7,721.4	414.6	5.4		
List-Str. 1,2	4,926.5	139.6	2.8	4,926.5	139.6	2.8	21,003	964
NOL+Str. 1,2	3,273.5	288.3	8.8	3,016.2	445.7	14.8	-	820
Total	8,200.0	320.3	3.9	7,942.7	467.1	5.9		
List-Str. 1-3	3,999.5	130.2	3.3	3,999.5	130.2	3.3	10,467	731
NOL+Str. 1-3	4,226.3	327.3	7.7	3,813.6	470.6	12.3	-	997
Total	8,225.8	352.2	4.3	7,813.1	488.3	6.2		
List-Str. 1-4	2,948.4	121.1	4.1	2,948.4	121.1	4.1	4,097	503
NOL+Str. 1-4	5,427.6	371.1	6.8	5,208.0	615.8	11.8	-	1,159
Total	8,376.0	390.4	4.7	8,156.4	627.6	7.7		
List-Str. 1-5	1,976.8	112.4	5.7	1,976.8	112.4	5.7	1,357	303
NOL+Str. 1-5	6,187.8	397.8	6.4	5,757.3	698.7	12.1	-	1,234
Total	8,164.6	413.4	5.1	7,734.1	707.7	9.2		
List-Str. 1-6	1,042.0	105.7	10.1	1,042.0	105.7	10.1	220	83
NOL+Str. 1-6	6,887.3	411.8	6.0	6,222.6	764.1	12.3	-	1,293
Total	7,929.3	425.1	5.4	7,264.6	771.4	10.6		
JES Area Frame	7,929.3	425.1	5.4	7,264.6	771.4	10.6		
Bd.	7,300			7,300				

Table 3.--List and area frame estimates by livestock strata-Nebraska 1973 JES list and area frame estimates

Multiple frame strata	List			Tract			Farm			List as % of tract	List as % of farm
	Sample size	DE	CV	No.	DE	CV	No.	DE	CV		
Hogs & Pigs		(000)	(%)		(000)	(%)		(000)	(%)	(%)	(%)
Nonoverlap	349	441.5 1/	19.0	349	470.1	20.7	107	517.2	20.0	93.9	85.4
1 (No livestock)	194	41.9	40.5	147	136.9	38.4	57	104.1	40.2	30.6	40.2
2 (0 hogs + cattle)	268	91.2	25.3	409	77.1	40.8	104	79.8	40.6	118.3	114.3
3 (1-124) 2/	513	536.7	6.3	226	652.1	21.1	75	641.1	21.4	82.3	83.7
4 (125-199)	227	490.8	5.9	77	577.9	22.7	30	672.8	21.6	84.9	72.9
5 (200-299)	186	417.2	6.2	55	496.5	31.2	17	612.1	32.6	84.0	68.2
6 (300 +)	216	533.3	6.0	47	570.6	33.1	13	585.4	36.0	93.5	91.1
Cattle & Calves											
Nonoverlap	357	1,398.6 1/	9.7	357	1,661.3	12.9	107	1,370.9	24.8	84.2	102.0
1 (No livestock)	143	95.8	31.0	147	329.4	27.6	57	498.8	37.0	29.1	19.2
2 (0-24) 3/	311	925.2	8.5	334	1,282.9	11.0	139	1,146.5	13.2	72.1	80.7
3 (25-49)	233	927.0	5.4	182	952.8	14.3	49	797.5	17.5	97.3	116.2
4 (50-99)	228	1,051.1	4.6	170	1,206.5	16.4	42	1,394.4	18.8	87.1	75.4
5 (100-199)	200	971.6	4.6	76	755.0	21.1	9	549.3	36.3	128.7	176.9
6 (200 +)	220	940.8	4.3	63	631.3	24.1	5	465.3	51.2	149.0	202.2

1/ The weighted NOL estimate as computed for the multiple frame survey indication.

2/ Hog index range for this stratum.

3/ Cattle index range for this stratum.

Table 4.--List size as corresponding hog and cattle strata are added to the list for multiple frame sampling

Cattle & hog strata	Cattle and hog index number combinations	List size (approx. no. names)
EO	Cattle ≥ 2,500 or Hogs ≥ 500	1,050
6 & EO	Cattle ≥ 200 or Hogs ≥ 300	4,700
5 & over	Cattle ≥ 100 or Hogs ≥ 200	9,950
4 & over	Cattle ≥ 50 or Hogs ≥ 125	19,200
3 & over	Cattle ≥ 25 or Hogs ≥ 1	35,450
2 & over	Cattle ≥ 1 or Hogs ≥ 0	45,800
All strata	Cattle ≥ 0 or Hogs ≥ 0	54,200

Table 5.--Estimated annual survey costs for hog and cattle multiple frame surveys by list and area frame combinations - Nebraska, 1973

List sample strata	List frame cost 1/	Area nonoverlap		Total multiple frame survey costs	
		Interview cost	Mail, telephone, interview cost 3/	((1) + (2))	((1) + (3))
	(1)	(2)	(3)	(4)	(5)
	Dol.	Dol.	Dol.	Dol.	Dol.
Entire list	38,025	10,125 4/	5,970 5/	48,150	43,995
List less Str. 1	36,030	13,560	7,535	49,590	43,565
Cost difference 2/	(- 1,995)	(+ 3,435)	(+ 1,565)	(+ 1,440)	(- 430)
List less Str.1-2	31,726	24,780	13,772	56,506	45,498
Cost difference 2/	(- 6,299)	(+14,655)	(+ 7,802)	(+ 8,356)	(+ 1,503)
List less Str.1-3	23,129	29,430	15,967	52,559	39,096
Cost difference 2/	(-14,896)	(+19,305)	(+ 9,997)	(+ 4,409)	(- 4,899)
List less Str.1-4	17,976	31,020	16,863	48,996	34,839
Cost difference 2/	(-20,049)	(+20,895)	(+10,893)	(+ 846)	(- 9,156)

1/ $.50, $1.00 and $15.00 per completed questionnaire by mail, telephone, and interview respectively.

2/ Difference in cost compared to sampling the entire list.

3/ Assuming the same response rate as obtained from list sample and equivalent costs per completed questionnaire.

4/ Cost of personal interview enumeration of 350 NOL tracts in the December survey and the March and September NOL samples totaling 325 tracts.

5/ Cost of mail, telephone and personal interview enumeration of 350 NOL tracts in December and 325 NOL tracts in the March - September surveys with the response rates of combined list strata 1-2.

Table 6.--Nonsampling errors detected in JES

Source	Hogs & pigs				Cattle & calves			
	JES		MF 1/		JES		MF 1/	
	Tract	Farm	Tract NOL	Farm NOL	Tract	Farm	Tract NOL	Farm NOL
	(000)	(000)	(000)	(000)	(000)	(000)	(000)	(000)
Original JES expansions	3,843.2	3,747.8	470.1	517.2	7,929.3	7,264.1	1,661.3	1,370.9
Original MF expansions	-	-	3,071.2	3,118.3	-	-	7,608.8	7,318.4
EO data summarized in area frame 2/	- 326.6	-	-	-	- 33.2	-	-	-
NOL classified as OL 3/	-	-	+ 31.7	+ 31.7	-	-	+ 125.2	+ 356.3
OL with no chance of list selection 4/	-	-	+ 43.1	+ 43.1	-	-	+ 35.0	-
Corrected JES expansions	3,516.6	3,747.8	544.9	592.0	7,896.1	7,264.6	1,821.5	1,727.2
Corrected MF expansions	-	-	3,146.0	3,193.1	-	-	7,769.0	7,674.7

1/ Nonsampling errors described here relate to MF expansions by their effect on the tract and farm estimates of the NOL domain. The original list frame expansions were 2,601,100 hogs and pigs and 5,947,500 cattle and calves.

2/ Hog extreme operator missed because common name was recorded rather than legal name. Discovered by looking at expanded segment totals. Two cattle EO's not discovered in edit. Uncovered when cattle index was printed out beside tract data.

3/ Tracts which were on the universe listing but not designated for MF sampling. Found when tracts were printed with universe code.

4/ One hog operator and one cattle operator who had no chance of selection in MF sample. Found when tract data was matched with list information.

Table 7.--Nebraska MF survey out-of-pocket costs associated with each group of list strata

List strata	Hogs and pigs four surveys				Cattle and calves two surveys				Gross list costs 2/
	Mail	Tele-phone	Inter-view	Costs	Mail	Tele-phone	Inter-view	Costs	
	No.	No.	No.	Dol.	No.	No.	No.	Dol.	Dol.
1	290	369	56	1,354	90	146	30	641	1,995
1-2	654	958	162	3,715	332	498	128	2,584	6,299
1-3	1,323	2,174	500	10,335	466	818	234	4,561	14,896
1-4	1,604	2,816	687	13,923	636	1,098	314	6,126	20,049
All	2,211	4,029	1,424	26,495	970	1,624	628	11,530	38,025

1/ Costs associated with the livestock strata combinations based on 1973 MF hog and cattle survey returns by mail, telephone and interview at $.50 per mail return, $1.00 per telephone response and $15.00 per personal interview. Computations for col. 1 of Table 5 cost analysis.

2/ Costs reflect only the amounts which would not be spent for collecting data in those strata which would become part of the NOL domain. Savings which would result from maintaining a smaller list are not represented.

Table 8.--Nebraska MF survey out-of-pocket costs resulting from personally interviewing all additional NOL tracts 1/

List strata	December survey		Combined March and September hog surveys					Gross add'l. costs 2/
	Add'l. NOL tracts	Interview costs	Add'l. NOL tracts both surveys	Est. response rates: Mail	Telephone	Interview	Total interview costs	
	No.	Dol.	No.	Pct.	Pct.	Pct.	Dol.	Dol.
1	86	1,290	143	-	-	100	2,145	3,435
1-2	441	6,615	536	-	-	100	8,040	14,655
1-3	533	7,995	754	-	-	100	11,310	19,305
1-4	563	8,445	830	-	-	100	12,450	20,895

1/ Costs associated with those area frame tracts which are currently OL but would become NOL upon deletion of the livestock strata from the list sample. A cost of $15.00 per interview is assumed. Computations for col. 2 of Table 5 cost analysis.

2/ Total additional costs assuming the added number of NOL tracts in March and September are personally interviewed. These costs are in addition to $10,125 for personal enumeration of the original 350 NOL tracts in the December survey and the March and September subsamples of 151 and 174 NOL tracts respectively.

Table 9.--Nebraska MF survey out-of-pocket costs resulting from mail, telephone and interview of additional NOL tracts in September and March 1/

List strata	December survey		Combined March and September hog surveys					Gross add'l. costs 2/
	Add'l. NOL tracts	Interview costs	Add'l. NOL tracts both surveys	Est. response rates: Mail	Telephone	Interview	Total M-T-I costs	
	No.	Dol.	No.	Pct.	Pct.	Pct.	Dol.	Dol.
1	86	1,290	143	40	52	8	275	1,565
1-2	441	6,615	536	37	53	10	1,187	7,802
1-3	533	7,995	754	33	54	13	2,002	9,997
1-4	563	8,445	830	30	55	15	2,448	10,893

1/ Costs associated with those area frame tracts which are currently OL but would become NOL upon deletion of livestock strata from the list sample. A cost of $.50 per mail return, $1.00 per telephone response and $15.00 for a personal interview is applied to the additional NOL tracts under the estimated response rates.

2/ Total additional costs assuming the added number of NOL tracts in March and September are contacted by mail, telephone and interview in the percentages shown which are similar to the list strata. These costs are in addition to $5,970 estimated cost for mail, telephone and personal interview enumeration of 350 original NOL tracts in December and 151 NOL tracts in March and 174 NOL tracts in September with the response rates shown for list strata 1-2.